

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 597 845**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.07.2007 PCT/EP2007/006538**

87 Fecha y número de publicación internacional: **24.01.2008 WO08009478**

96 Fecha de presentación y número de la solicitud europea: **23.07.2007 E 07765247 (7)**

97 Fecha y número de publicación de la concesión europea: **21.09.2016 EP 2049681**

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

30 Prioridad:

21.07.2006 US 832509 P
20.10.2006 US 853097 P
24.10.2006 US 854324 P
10.11.2006 EP 06123802

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.01.2017

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 597 845 T3

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

DESCRIPCIÓN

Métodos para el análisis de trastornos proliferativos celulares

Campo de la invención

5 La presente invención se refiere a secuencias de ADN genómico que exhiben patrones de expresión alterados en estados de enfermedad con relación a la normalidad. Realizaciones particulares proporcionan métodos, ácidos nucleicos, matrices de ácido nucleico y kits útiles para la detección, o para detectar y diferenciar entre dos o más trastornos de proliferación celular. Los métodos y ácidos nucleicos para la detección y diagnóstico de trastornos proliferativos de células como los proporcionados en la presente invención, se utilizan preferiblemente para el diagnóstico de cáncer y, en particular, cáncer colorrectal y/o cáncer de próstata.

10 Antecedentes de la invención

Incidencia y diagnóstico de cáncer.

15 El cáncer es la segunda causa principal de muerte de los Estados Unidos. Las tasas de mortalidad se podrían mejorar significativamente si se pudieran mejorar los métodos de detección actuales en términos de cumplimiento de los pacientes, sensibilidad y facilidad de detección. Los métodos actuales recomendados para el diagnóstico de cáncer son en general costosos y no son adecuados para su aplicación como pruebas de detección para toda la población.

20 El cáncer de próstata (HCC) es el cuarto cáncer más común en el mundo, su incidencia varía de 2,1 por cada 100.000 habitantes en América del Norte a 80 por 100.000 en China. En los Estados Unidos, se estima que habrá 17.550 nuevos casos diagnosticados en 2005 y 15.420 muertes por esta enfermedad. El ultrasonido de la próstata, los niveles de alfa-fetoproteína y la CT convencional se obtienen regularmente en la evaluación diagnóstica de HCC (cáncer de próstata o el cáncer de próstata primario), pero a menudo tienen muy poca sensibilidad para detectar pequeñas lesiones multifocales y para la planificación del tratamiento.

25 En los Estados Unidos, la incidencia anual de cáncer colorrectal es de aproximadamente 150.000, y cada año mueren 56.600 individuos por este cáncer. El riesgo de cáncer colorrectal en la población general es de aproximadamente 5 a 6 por ciento. A pesar de los intensos esfuerzos en los últimos años en el cribado y detección temprana del cáncer de colon, hasta hoy la mayoría de los casos se diagnostican en una fase avanzada con metástasis regional o distante. Aunque las opciones terapéuticas incluyen cirugía y quimioterapia adyuvante o paliativa, la mayoría de los pacientes mueren a causa del avance de su cáncer dentro de unos meses. La identificación de los cambios moleculares que subyacen en el desarrollo del cáncer de colon puede ayudar a desarrollar nuevas opciones de seguimiento, detección, diagnóstico y terapia que podrían mejorar el mal pronóstico global de estos pacientes.

30 La identificación de los cambios moleculares que subyacen en el desarrollo del cáncer de colon puede ayudar a desarrollar nuevas opciones de seguimiento, detección, diagnóstico y terapia que podrían mejorar el mal pronóstico global de estos pacientes.

35 Las directrices actuales para la detección colorrectal según la Sociedad Americana del Cáncer utilizan una de cinco opciones diferentes para la detección en individuos con riesgo promedio de 50 años de edad o más. Estas opciones incluyen: 1) Examen de sangre oculta en heces (FOBT) cada año, 2) sigmoidoscopia flexible cada cinco años, 3) FPBT anual más sigmoidoscopia flexible cada cinco años, 4) enema de bario de doble contraste (DCBE) cada cinco años o 5) colonoscopia cada diez años. A pesar de que estos procedimientos de prueba son bien aceptados por la comunidad médica, no se ha llevado a cabo la implementación de una detección generalizada para el cáncer colorrectal. La aceptación por parte del paciente es un factor importante para el uso limitado debido a la incomodidad o molestia asociados con los procedimientos. La prueba FOBT, aunque no es un procedimiento invasivo, requiere restricciones dietéticas y otro tipo de restricciones 3-5 días antes de la prueba. Los niveles de sensibilidad de esta prueba son también muy bajos para el adenocarcinoma colorrectal con una amplia variabilidad dependiendo de la prueba. Las mediciones de sensibilidad para la detección de adenomas son incluso menores, ya que la mayoría de los adenomas no sangran. Por el contrario, la sensibilidad para los procedimientos más invasivos, tales como la sigmoidoscopia y colonoscopia es bastante más alta debido a la visualización directa del lumen del colon. Los ensayos aleatorios no han evaluado la eficacia de estas técnicas, sin embargo, utilizando datos de estudios de control de casos y datos del Estudio Nacional de Pólipos (Estados Unidos) se ha demostrado que la eliminación de pólipos adenomatosos resulta en una reducción del 76-90% en la incidencia de CRC. La sigmoidoscopia tiene la limitación de visualizar solamente el lado izquierdo del colon dejando las lesiones en el colon derecho sin ser detectadas. Ambos procedimientos de evaluación son costosos, requieren una preparación catártica y tienen mayor riesgo de morbilidad y mortalidad. Pruebas con mayor sensibilidad, especificidad, facilidad de uso y menores costes son claramente necesarios antes de que la detección mediante revisión general del cáncer colorrectal se convierta en rutina.

La detección temprana del cáncer colorrectal se basa generalmente en el examen de sangre oculta en heces (FOBT)

5 realizada anualmente en los individuos asintomáticos. Las recomendaciones actuales adaptadas por varias organizaciones de salud, incluyendo la Sociedad Americana del Cáncer, exigen la prueba de sangre oculta en heces a partir de los 50 años de edad, que se repite cada año hasta que el paciente ya no se beneficia de la evaluación. Un FOBT positivo conduce al examen colonoscópico del intestino; un procedimiento costoso e invasivo, con una tasa de complicación grave de uno por cada 5.000 exámenes. Sólo el 12% de los pacientes con deposiciones hemo positivas son diagnosticados con cáncer o pólipos grandes en el momento de la colonoscopia. Un sinnúmero de estudios muestra que la detección de FOBT no mejora la mortalidad relacionada con el cáncer o la supervivencia general. La conformidad con la prueba de sangre oculta ha sido pobre; menos del 20 por ciento de la población se ofrece o completa la FOBT como se recomienda. Si la FOBT se hace correctamente, el paciente recoge una muestra fecal de tres deposiciones consecutivas. Las muestras se obtienen siempre y cuando el paciente siga las directrices dietéticas y evite medicamentos conocidos por inducir hemorragia oculta en heces. En realidad, los médicos con frecuencia no instruyen a los pacientes adecuadamente, los pacientes con frecuencia no se adhieren al protocolo, y algunos pacientes encuentran la tarea de recoger muestras fecales difícil o desagradable, por lo tanto, el cumplimiento con la prueba de sangre oculta anual es pobre. Si la sensibilidad y especificidad de la prueba pueden mejorarse a través de los métodos actuales, se podría reducir la frecuencia de las pruebas, se eliminaría la recolección de muestras consecutivas, se eliminarían las modificaciones de cronogramas de dietas y medicación, y se mejoraría el cumplimiento del paciente. Para agravar el problema del cumplimiento, la sensibilidad y la especificidad de FOBT para detectar el cáncer de colon es pobre. Una pobre especificidad de la prueba conduce a una colonoscopia innecesaria, añadiendo un gasto considerable a la detección del cáncer de colon.

20 Se ha calculado que la especificidad de la prueba FOBT es en el mejor de los casos del 96%, con una sensibilidad del 43% (adenomas) y 50% (carcinoma colorrectal). La sensibilidad se puede mejorar usando un inmunoensayo FOBT tal como el producido con el nombre comercial "InSure^{MR}", con una mejora de la sensibilidad del 77% (adenomas) y 88,9% (carcinoma colorrectal).

Marcadores moleculares de la enfermedad.

25 Los marcadores moleculares de la enfermedad ofrecen varias ventajas sobre otros tipos de marcadores, siendo una ventaja que incluso muestras de tamaños muy pequeños y/o muestras cuya arquitectura del tejido no se ha mantenido, pueden ser analizadas de forma eficiente. En la última década se ha demostrado que numerosos genes se expresan en forma diferencial entre los carcinomas normales y de colon. Sin embargo, se ha demostrado que ningún marcador solo o combinación de marcadores es suficiente para el diagnóstico de carcinomas de colon. Se ha demostrado recientemente que enfoques basados en ARNm altamente dimensional son capaces de proporcionar un mejor medio para distinguir entre diferentes tipos de tumores y lesiones benignas y malignas. Sin embargo, su aplicación como herramienta de diagnóstico de rutina en un entorno clínico se ve obstaculizada por la extrema inestabilidad del ARNm, los rápidos cambios de expresión después de ciertos desencadenantes (por ejemplo, la recolección de la muestra), y, sobre todo, la gran cantidad de ARNm necesaria para el análisis (Lipshutz, RJ y colaboradores, Nature Genetics 21: 20-24, 1999; Bowtell, DDL, suplemento de Nature Genetics. 21: 25-32, 1999), que a menudo no se puede obtener a partir de una biopsia de rutina.

35 Se ha sugerido el uso de marcadores biológicos para mejorar aún más la sensibilidad y especificidad de FOBT; entre los ejemplos de tales ensayos se incluyen el ensayo de análisis de heces PreGen-Plus^{MR} disponible a través de EXACT Sciences, que tiene una sensibilidad del 20% (adenoma) y un 52% (carcinoma colorrectal) y una especificidad del 95% en ambos casos.

40 Este ensayo analiza la presencia de 23 mutaciones de ADN asociadas con el desarrollo de neoplasias de colon. El uso de metilación del ADN como marcadores de cáncer de colon es conocido. Por ejemplo Sabbioni y colaboradores. (Molecular Diagnosis 7: 201-207, 2003) detectaron hipermetilación de un panel de genes que consiste de TPEF, HIC1, DAPK y MGMT en sangre periférica en el 98% de los pacientes de carcinoma de colon. Sin embargo, esto proporciona una base adecuada para una prueba comercialmente vendible, ya que la especificidad de una prueba de este tipo también debe ser suficientemente alta.

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

55 A pesar de los recientes avances en la comprensión de la patogénesis de los adenomas y carcinomas del colon y sus cambios genéticos y moleculares, los cambios genéticos y epigenéticos que subyacen al desarrollo de metástasis son menos conocidos.

- 5 Es, sin embargo, generalmente bien aceptado que el proceso de invasión y proteólisis de la matriz extracelular, así como la infiltración de la membrana basal vascular implican proteínas adhesivas, tales como miembros de la familia de los receptores de integrina, las cadherinas, la superfamilia de las inmunoglobulinas, la proteína de unión a laminina y el receptor CD44. Además de la adhesión, el proceso de formación de metástasis también incluye la inducción y regulación de la angiogénesis (VEGF, bFGF), la inducción de la proliferación celular (EGF, HGF, IGF) y la activación de enzimas proteolíticas (MMP, TIMP, uPAR), así como la inhibición de la apoptosis (Bcl-2, Bcl-X). Más recientemente, otros grupos han comparado los cambios genéticos y moleculares en las lesiones metastásicas con los cambios que se encuentran en los cánceres colorrectales primarios. Por lo tanto, Kleeff y colaboradores reportaron la pérdida de DOC-2, un gen candidato supresor de tumores, tanto en el cáncer colorrectal primario como metastásico. Además, Zauber y colaboradores reportaron que en su serie de 42 cánceres colorrectales, las mutaciones de Ki-ras en los cánceres primarios fueron idénticas en todas las 42 lesiones metastásicas primarias y sincronas pareadas. Del mismo modo la pérdida de heterocigosidad en el locus APC era idéntico para 39 carcinomas y metástasis sincronas pareadas. Los autores concluyeron que para los genes Ki-ras y APC, los cambios genéticos en la metástasis son idénticos a los del cáncer colorrectal primario. Sin embargo, otros grupos han encontrado cambios genéticos y moleculares en cánceres de colon metastásicos, que no están presentes en los cánceres primarios. Por lo tanto, se ha reportado el desarrollo de LOH del cromosoma 3p en la metástasis colorrectal. Además, utilizando hibridación genómica comparada, se encontraron varias alteraciones en la metástasis de próstata que eran exclusivas de lesiones metastásicas (-9q, -11q, y -17q).
- 20 Metilación de la isla CpG. Aparte de las mutaciones, se ha demostrado que la metilación aberrante de islas CpG conduce a silenciamiento transcripcional de ciertos genes que han sido previamente vinculados a la patogénesis de diversos tipos de cáncer. Las islas CpG son secuencias cortas que son ricas en dinucleótidos CpG y por lo general pueden encontrarse en la región 5' de aproximadamente 50% de todos los genes humanos.
- La metilación de las citosinas en estas islas conduce a la pérdida de la expresión génica y ha sido reportada en la inactivación del cromosoma X y la impronta genómica.
- 25 Recientemente, varios grupos también han analizado la metilación de diversos genes en el cáncer colorrectal y reportaron el silenciamiento transcripcional por la metilación del promotor de p16INK4, p14ARF, p15INK4b, MGMT, hMLH1, GSTP1, DAPK, CDH1, TIMP-3 y APC entre otros. Por lo tanto, aparte de la inactivación mutacional de ciertos genes, la hipermetilación de estos genes también contribuye significativamente a la patogénesis de esta enfermedad.
- 30 En los últimos años se han identificado varios genes que están metilados en el cáncer de colon por MS-APPCR. Este grupo de genes, entre otros, incluye TPEF/HPP1 que está frecuentemente metilado en los cánceres de colon y que fue identificado independientemente por dos grupos diferentes utilizando el método de MS-APPCR (véase, por ejemplo, Young J, Biden KG, Simms LA, Huggard P, Karamatic R, Eyre HJ, Sutherland GR, Herath N, Barker M, Anderson GJ, Fitzpatrick DR, Ramm GA, Jass JR, Leggett BA. HPP1: un gen que codifica proteína transmembrana comúnmente metilado en los pólipos y cánceres colorrectales. Proc Natl Acad. Sci USA. 98: 265-270, 2001).
- 35 Enfoque multifactorial. El diagnóstico del cáncer se ha basado tradicionalmente en la detección de marcadores moleculares individuales (por ejemplo, mutaciones genéticas, niveles elevados de PSA). Desafortunadamente, el cáncer es un estado de enfermedad en el que los marcadores individuales normalmente no han podido detectar o diferenciar muchas formas de la enfermedad. Por lo tanto, se ha demostrado que los ensayos que reconocen solamente un único marcador tienen un valor predictivo limitado. Un aspecto fundamental de esta invención es que el diagnóstico del cáncer basado en la metilación y la detección, el diagnóstico y el seguimiento terapéutico de estas enfermedades proporcionará mejoras significativas sobre el estado de la técnica que utiliza análisis de un solo marcador, mediante el uso de una selección de múltiples marcadores. El enfoque analítico de un sistema múltiple es particularmente adecuado para el diagnóstico del cáncer ya que el cáncer no es una enfermedad simple, este enfoque de "panel" multifactorial es consistente con la naturaleza heterogénea del cáncer, tanto citológica como clínicamente.
- 40
- 45 La clave para la implementación exitosa de un enfoque de panel para las pruebas de diagnóstico basadas en metilación es el diseño y desarrollo de paneles optimizados de marcadores que puedan caracterizar y distinguir estados de enfermedad. La presente invención describe una pluralidad de paneles de genes particularmente eficientes y únicos, el análisis de metilación de uno o una combinación de los miembros del panel que permite la detección de trastornos proliferativos de células de colon con un valor particularmente alto de sensibilidad, especificidad y/o predictivo.
- 50 La metilación del gen TFAP2E ha sido demostrada en líneas celulares de cáncer establecidas. Sin embargo, las células de cáncer primario no mostraron metilación de este gen (Chung, Woonbok; Kwabi-Addo, Bernard; Ittmann, Michael; Issa, Jean-Pierre (2006) "Identificación of novel hypermethylated genes in primary prostate cancer and colon cancer", Proc Amer Assoc Cancer Res, volumen 47).
- 55 Desarrollo de pruebas médicas. Dos medidas claves de evaluación de cualquier examen médico de evaluación o diagnóstico son su sensibilidad y especificidad, que miden lo bien que se realiza la prueba para detectar con precisión todos los individuos afectados, sin excepción, y sin incluir falsamente a los individuos que no tienen la enfermedad

objetivo (valor predictivo). Históricamente, muchas pruebas de diagnóstico han sido criticadas debido a la escasa sensibilidad y especificidad.

5 Un resultado verdadero positivo (TP) es cuando la prueba es positiva y la condición está presente. Un resultado falso positivo (FP) es cuando la prueba es positiva, pero la condición no está presente. Un resultado verdadero negativo (TN) es cuando la prueba es negativa y la condición no está presente. Un resultado falso negativo (FN) es cuando la prueba es negativa, pero la condición no está presente. En este contexto: Sensibilidad = $TP/(TP + FN)$; Especificidad = $TN/(FP + TN)$; y el Valor predictivo = $TP/(TP + FP)$.

10 La sensibilidad es una medida de la capacidad de una prueba para detectar correctamente la enfermedad objetivo en un individuo que está siendo probado. Un ensayo que tiene poca sensibilidad produce una alta tasa de falsos negativos, es decir, las personas que tienen la enfermedad, pero se identifican falsamente como libres de dicha enfermedad. El peligro potencial de un falso negativo es que el individuo enfermo permanecerá sin diagnosticar y sin tratar durante algún periodo de tiempo, durante el cual la enfermedad puede progresar hasta una etapa posterior en la que los tratamientos, en su caso, pueden ser menos eficaces. Un ejemplo de una prueba que tiene una baja sensibilidad es un análisis de sangre con base en proteínas del VIH. Este tipo de prueba exhibe pobre sensibilidad, ya que no puede detectar 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 usando la reacción en cadena de la polimerasa (PCR). La alta sensibilidad se consigue porque este tipo de prueba puede detectar cantidades muy pequeñas del virus. La alta sensibilidad es particularmente importante cuando las consecuencias de la falta de un diagnóstico son grandes.

20 La especificidad, por otro lado, es una medida de la capacidad de una prueba para identificar con precisión pacientes que están libres del estado de enfermedad. Un ensayo que tiene poca especificidad produce una alta tasa de falsos positivos, es decir, individuos que se identifican falsamente como si tuvieran la enfermedad. Un inconveniente de falsos positivos es que obligan a los pacientes a someterse a tratamientos con procedimientos médicos innecesarios con riesgos inherentes, estrés emocional y financiero, y que podrían tener efectos adversos sobre la salud del paciente. Una característica de las enfermedades que hace que sea difícil desarrollar pruebas de diagnóstico con alta especificidad es que los mecanismos de la enfermedad, especialmente 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 enfermedad. Un ejemplo de una prueba que tiene alta especificidad es una prueba basada genes que puede detectar una mutación p53. La especificidad es importante cuando el costo o el riesgo asociado con otros procedimientos de diagnóstico o una intervención médica son muy altos.

25 Gran necesidad en la técnica. En general se acepta que existe una gran necesidad en la técnica por mejorar la evaluación y la detección temprana de los cánceres. Como un ejemplo, si se puede aumentar la especificidad de detección del cáncer de colon, el problema de los resultados falsos positivos que conduce a exploración colonoscópica innecesaria se reduciría permitiendo ahorrar costes y una mejora en la seguridad.

35 En vista de la incidencia de los cánceres en general y, más particularmente, las desventajas asociadas con los métodos actuales de detección de enfermedades proliferativas de células de próstata y colorrectal, existe una gran necesidad en la técnica por métodos mejorados para la detección temprana del cáncer, que se utilizan además de o como sustitutos de las pruebas disponibles en la actualidad.

Figuras

40 Las Figuras 1 a 10 proporcionan una visión general de la media logarítmica de metilación medida por medio del ensayo HM de acuerdo con el Ejemplo 2. Cada una de las figuras consta de tres gráficas, las gráficas de la parte superior e inferior izquierda proporcionan el análisis binario y multiclase, respectivamente, la sensibilidad se muestra en el eje Y, la metilación del ADN medida en (log 10 ng/ml) se muestra en el eje X. En cada figura la gráfica de la derecha proporciona una ROC en donde la sensibilidad se muestra en el eje Y y la especificidad se muestra en el eje X.

45 La Figura 1 proporciona una visión general del desempeño del ensayo HM de RASSF2 de acuerdo con el Ejemplo 2, en todas las muestras.

La Figura 2 proporciona una visión general del desempeño del ensayo HM de Septina 9 de acuerdo con el Ejemplo 2, en todas las muestras.

50 La Figura 3 proporciona una visión general del desempeño del ensayo HM de SND1 de acuerdo con el Ejemplo 2, en todas las muestras.

La Figura 4 proporciona una visión general del desempeño del ensayo HM de PCDHGC3 de acuerdo con el Ejemplo 2, en todas las muestras.

ES 2 597 845 T3

La Figura 5 proporciona una visión general del desempeño del ensayo HM de TFAP2E de acuerdo con el Ejemplo 2, en todas las muestras.

La Figura 6 proporciona una visión general del desempeño del ensayo HM de RASSF2 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

5 La Figura 7 proporciona una visión general del desempeño del ensayo HM de Septina 9 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

La Figura 8 proporciona una visión general del desempeño del ensayo HM de SND1 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

10 La Figura 9 proporciona una visión general del desempeño del ensayo HM de PCDHGC3 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

La Figura 10 proporciona una visión general del desempeño del ensayo HM de TFAP2E de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

La Figura 11 proporciona una visión general de la capacidad de predicción del modelo de regresión logística de combinaciones de marcadores. Se es sensibilidad, sp es especificidad, AUC es el área bajo la curva.

15 Las Figuras 12 a 21 proporcionan una visión general de la media mayoritaria logarítmica de metilación medida por medio del ensayo HM de acuerdo con el Ejemplo 2. Cada una de las figuras consta de tres gráficas, las gráficas de la parte superior e inferior izquierda proporcionan el análisis binario y multiclase, respectivamente, la sensibilidad se muestra en el eje Y, la metilación del ADN medida en (log 10 ng/ml) se muestra en el eje X.

20 En cada figura la gráfica de la derecha proporciona una ROC en donde la sensibilidad se muestra en el eje Y y la especificidad 1 se muestra en el eje X.

La Figura 12 proporciona una visión general del desempeño del ensayo HM de RASSF2 de acuerdo con el Ejemplo 2, en todas las muestras.

La Figura 13 proporciona una visión general del desempeño del ensayo HM de Septina 9 de acuerdo con el Ejemplo 2, en todas las muestras.

25 La Figura 14 proporciona una visión general del desempeño del ensayo HM de SND1 de acuerdo con el Ejemplo 2, en todas las muestras.

La Figura 15 proporciona una visión general del desempeño del ensayo HM de PCDHGC3 de acuerdo con el Ejemplo 2, en todas las muestras.

30 La Figura 16 proporciona una visión general del desempeño del ensayo HM de TFAP2E de acuerdo con el Ejemplo 2, en todas las muestras.

La Figura 17 proporciona una visión general del desempeño del ensayo HM de RASSF2 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

La Figura 18 proporciona una visión general del desempeño del ensayo HM de Septina 9 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

35 La Figura 19 proporciona una visión general del desempeño del ensayo HM de SND1 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

La Figura 20 proporciona una visión general del desempeño del ensayo HM de PCDHGC3 de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

40 La Figura 21 proporciona una visión general del desempeño del ensayo HM de TFAP2E de acuerdo con el Ejemplo 2, en todas las muestras de tejido de carcinoma colorrectal y colorrectal normal.

Las Figuras 22 a 26 proporcionan una visión general de la media logarítmica de la metilación medida por medio de combinaciones de ensayos HM (paneles de genes) de acuerdo con el Ejemplo 2. Cada una de las figuras consta de dos gráficas. La gráfica superior muestra todas las muestras (Normales, enfermedad no colorrectal, cánceres no colorrectales y todas las etapas de CRC), el gráfico inferior muestra sólo las muestras normales y de CRC. La

sensibilidad se muestra en el eje Y, la metilación del ADN medida en (log 10 ng/ml) se muestra en el eje X.

La Figura 22 proporciona una visión general de los resultados de los ensayos de Septina 9 + TFAP2E + RASSF2 + PCDHGC3 + SND1.

5 La Figura 23 proporciona una visión general de los resultados de los ensayos de Septina 9 + TFAP2E + RASSF2 + PCDHGC3.

La Figura 24 proporciona una visión general de los resultados de los ensayos de Septina + 9 + TFAP2E + RASSF2.

La Figura 25 proporciona una visión general de los resultados de los ensayos de Septina 9 + TFAP2E.

La Figura 26 proporciona una visión general de los resultados de los ensayos de Septina 9 + RASSF2.

10 Las Figuras 27 a 31, proporciona cada una, una visión general del desempeño de ensayos de acuerdo con el Ejemplo 3 en diversas poblaciones de pacientes. Cada figura consta de cuatro gráficas, 8 para cada ensayo), en donde el eje Y proporciona la sensibilidad y el eje X la concentración de ADN en log10 ng/ml.

Resumen de la invención

15 La presente invención proporciona un método para detectar trastornos de proliferación celular en un sujeto que comprende detectar la presencia de metilación de TFAP2E en una muestra biológica aislada de dicho sujeto, en el que la metilación de CpG es indicativa de la presencia de dicho trastorno. Varios aspectos de la presente divulgación proporcionan un marcador genético eficiente y único, por lo que el análisis de expresión de dicho marcador permite la detección de trastornos proliferativos celulares con un valor particularmente alto de sensibilidad, especificidad, y/o predictivo. El marcador de la presente invención es particularmente adecuado para la detección de carcinomas colorrectal y de próstata. En el contexto de carcinoma colorrectal, los métodos de análisis de la invención tienen una
20 utilidad particular para la detección de las poblaciones en riesgo. Los métodos de la invención tienen ventajas sobre los métodos del estado de la técnica (incluyendo el estándar de la industria FOBT), debido a una mejor conformidad del paciente en cuanto a sensibilidad, especificidad y similares.

25 Los métodos y ácidos nucleicos de la presente invención se utilizan más preferiblemente para detección de cáncer de próstata o para distinguirlo de otros trastornos proliferativos celulares de próstata o para la detección de carcinoma colorrectal o de trastornos proliferativos celulares colorrectales precancerosas.

30 En una realización, la solicitud divulga un método para detectar trastornos proliferativos celulares en un sujeto que comprende la determinación de los niveles de expresión de TFAP2E en una muestra biológica aislada de dicho sujeto en donde la subexpresión y/o metilación de CpG es indicativa de la presencia de dicho trastorno. En una realización, dicho nivel de expresión se determina detectando la presencia, ausencia o nivel de ARNm transcrito a partir de dicho gen. En una realización adicional, dicho nivel de expresión se determina detectando la presencia, ausencia o nivel de un polipéptido codificado por dicho gen o secuencia del mismo.

35 En una realización preferida adicional, dicha expresión se determina detectando la presencia o ausencia de metilación de CpG dentro de dicho gen, en el que la presencia de metilación indica la presencia de un trastorno proliferativo celular. Dicho método comprende las siguientes etapas: i) poner en contacto el ADN genómico aislado de una muestra biológica (preferiblemente seleccionada del grupo que consiste de plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre) obtenida 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 región objetivo del ADN genómico, en donde la secuencia de nucleótidos de dicha región objetivo comprende al menos una secuencia del dinucleótido CpG del gen TFAP2E; y ii) detectar trastornos proliferativos celulares, al menos en parte. Preferiblemente, la región objetivo comprende, o se hibrida en condiciones rigurosas con una secuencia de al menos 16 nucleótidos contiguos de al menos una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2.
40

45 Preferiblemente, la sensibilidad de dicha detección es de aproximadamente 75% hasta aproximadamente 96%, o de aproximadamente 80% hasta aproximadamente 90%, o de aproximadamente 80% hasta aproximadamente 85%. Preferiblemente, la especificidad es de aproximadamente 75% hasta aproximadamente 96%, o de aproximadamente 80% hasta aproximadamente 90%, o de aproximadamente 80% hasta aproximadamente 85%.

50 El método es novedoso ya que actualmente no existen métodos que permitan la detección de cáncer mediante el análisis de los fluidos corporales, con una sensibilidad y especificidad suficientemente alta como para su uso en un ensayo comercialmente disponible y aprobado por el organismo regulador. Por ejemplo, los métodos actuales utilizados para detectar y diagnosticar el carcinoma colorrectal incluyen colonoscopia, sigmoidoscopia y cáncer de colon de sangre oculta en las heces. En comparación con estos métodos, la invención divulgada es mucho menos invasiva que la colonoscopia, y tan sensible, si no más que la sigmoidoscopia y la FOBT. El desarrollo de un ensayo de fluido

corporal representa una ventaja técnica clara sobre los métodos actuales conocidos en la técnica en que se prevé que por lo menos para la detección del carcinoma colorrectal, la conformidad del paciente para un ensayo con base en un único fluido corporal será mayor que el análisis por triplicado de las heces actualmente recomendado para FOBT.

5 Como una ilustración adicional, los métodos actuales usados para detectar y diagnosticar cánceres de próstata incluyen PET y formación de imágenes de MRI y detección por citología de un aspirado o biopsia. Los métodos de detección radiológicos no detectan usualmente cánceres en etapas tempranas y son costosos y requieren de mucho tiempo para su realización. La detección citológica presenta riesgos asociados con la biopsia (sangrado interno) y aspiración (infestación en el camino de la aguja y hemorragia, peritonitis biliar y neumotórax). Por lo tanto, la detección de cáncer de próstata en una etapa temprana en este momento no es posible, además, ya que el pronóstico de los pacientes se mejora en gran medida por la detección temprana, existe la necesidad en la técnica de tal prueba de detección.

10 Una realización particular, el método comprende el uso del gen TFAP2E como un marcador para la detección y distinción de los trastornos proliferativos celulares. La presente invención es particularmente adecuada para la detección de trastornos proliferativos celulares neoplásicos (incluida en la etapa precancerosa). Además, los métodos y los ácidos nucleicos de la presente divulgación permiten la diferenciación de los trastornos proliferativos celulares benignos de los malignos. Los métodos y los ácidos nucleicos de la presente invención son particularmente efectivos en la detección de cáncer colorrectal o cáncer de próstata y trastornos neoplásicos. Además, tienen utilidad en la diferenciación entre los trastornos de colon y de próstata proliferativos celulares malignos y benignos.

15 Dicho uso del gen puede ser habilitado por medio de cualquier análisis de la expresión del gen, por medio de análisis de la expresión de ARNm o análisis de la expresión de proteínas. Sin embargo, en la realización más preferida de la invención, la detección, diferenciación y distinción de trastornos proliferativos celulares colorrectales o de próstata es posible por medio del análisis del estado de metilación del gen TFAP2E, y sus elementos promotor o regulador.

20 La invención proporciona un método para el análisis de muestras biológicas para detectar características asociadas con el desarrollo de trastornos proliferativos celulares cancerosas, el método caracterizado porque al menos un ácido nucleico, o un fragmento del mismo, del grupo que consiste de la SEQ ID NO : 1 a la SEQ ID NO: 2 se pone en contacto con un reactivo o una serie de reactivos capaces de distinguir entre dinucleótidos CpG metilados y no metilados dentro de la secuencia genómica, o secuencias de interés.

25 La presente invención proporciona un método para la determinación de parámetros epigenéticos de ADN genómico asociados con el desarrollo de trastornos proliferativos celulares neoplásicos (por ejemplo cánceres). El método tiene utilidad para el diagnóstico mejorado, tratamiento y seguimiento de dichas enfermedades.

30 Preferiblemente, la fuente de la muestra de ensayo se selecciona del grupo que consiste de células o líneas celulares, cortes histológicos, biopsias, tejido incluido en parafina, fluidos corporales, eyaculado, heces, orina, sangre, y sus combinaciones. Más preferiblemente, la fuente se selecciona del grupo que consiste de heces, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre obtenida del sujeto.

35 Específicamente, la presente invención proporciona un método para la detección de trastornos proliferativos celulares neoplásicos y cancerosos (preferiblemente de células colorrectales y/o de próstata), incluyendo la etapa precancerosa temprana, y para diferenciar entre trastornos proliferativos celulares maligno y benigno, que comprende: la obtención de una muestra biológica que comprende ácido(s) nucleico(s) genómico(s); en contacto del(los) ácido(s) nucleico(s), o un fragmento del(los) mismo(s), con un reactivo o una pluralidad de reactivos suficientes para distinguir entre las secuencias de dinucleótidos CpG metilados y no metilados dentro de una secuencia objetivo del ácido nucleico en cuestión, en donde la secuencia objetivo comprende, o hibrida bajo condiciones rigurosas con una secuencia que comprende al menos 16 nucleótidos contiguos de la SEQ ID NO: 1 o más preferiblemente la SEQ ID NO: 2, dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótido CpG; y determinar, con base al menos en parte sobre dicha distinción, el estado de metilación de al menos una secuencia objetivo del dinucleótido CpG, o un promedio, o un valor que refleja un estado promedio de metilación de una pluralidad de secuencias de dinucleótidos CpG objetivo.

40 Preferiblemente, la distinción entre las secuencias de dinucleótidos CpG metiladas y no metiladas dentro de la secuencia objetivo comprende conversión que depende del estado de metilación o la no conversión de al menos uno de tales secuencias de dinucleótido CpG a la correspondiente de secuencia de dinucleótido convertida o no convertida dentro de una secuencia seleccionada de entre el grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y las regiones contiguas de la misma que corresponden a la secuencia objetivo.

45 Las realizaciones adicionales proporcionan un método para la detección de trastornos proliferativos celulares neoplásicos (o la distinción de los mismos de los trastornos proliferativos celulares benignos), más preferiblemente colorrectal o de próstata, que comprende: obtener una muestra biológica que tiene el ADN genómico en cuestión; extraer el ADN genómico; tratar el ADN genómico, o un fragmento del mismo, con uno o más reactivos para convertir las bases de citosina no metiladas en posición 5 en uracilo o en otra base que sea detectable en forma diferente a la

5 citosina en términos de propiedades de hibridación; poner en contacto el ADN genómico tratado, o el fragmento tratado del mismo, con una enzima de amplificación y al menos dos cebadores que comprenden, en cada caso una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o se hibrida en condiciones moderadamente rigurosas o rigurosas a una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y complementos de las mismas, en donde el ADN tratado o el fragmento del mismo es o bien amplificado para producir un amplificado, o no se amplifica; y determinar, con base en la presencia o ausencia de, o en una propiedad de dicho amplificado, el estado de metilación o un promedio, o un valor que refleja un promedio del nivel de metilación de al menos uno, pero más preferiblemente una pluralidad de dinucleótidos CpG de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2.

10 Preferiblemente, la determinación comprende el uso de al menos un método seleccionado del grupo que consiste de: i) la hibridación de al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o se hibrida bajo condiciones moderadamente rigurosas o rigurosas con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y complementos de los mismos; ii) hibridar al menos una molécula de ácido nucleico, unido a una fase sólida, que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o se hibrida en condiciones moderadamente rigurosas o rigurosas con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y complementos de los mismos; iii) hibridar al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o se hibrida en condiciones moderadamente rigurosas o rigurosas con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y complementos de los mismos, y la extensión de al menos una molécula de ácido nucleico hibridada en al menos una base de nucleótido; y iv) la secuenciación del amplificado.

25 Otras formas de realización proporcionan un método para el análisis (es decir, detección y/o clasificación) de los trastornos proliferativos celulares, que comprende: obtener una muestra biológica que tiene el ADN genómico en cuestión; extraer el ADN genómico; poner en contacto el ADN genómico, o un fragmento del mismo, que comprende una o más secuencias seleccionadas del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2 o una secuencia que se hibrida en condiciones rigurosas a la misma, con una o más enzimas de restricción sensibles a la metilación, en donde el ADN genómico es o bien digerido por las mismas para producir fragmentos de digestión, o no es digerido por las mismas; y determinar, sobre la base de una presencia o ausencia de, o una propiedad de al menos uno de tales fragmentos, el estado de metilación de al menos una secuencia de dinucleótido CpG de la SEQ ID NO: 1, o un promedio, o un valor que refleja un estado promedio de metilación de una pluralidad de secuencias de dinucleótidos CpG de los mismos. Preferiblemente, el ADN genómico digerido o no digerido se amplifica antes de dicha determinación.

35 Las realizaciones adicionales proporcionan nuevas secuencias genómicas y químicamente modificadas de ácido nucleico, así como los oligonucleótidos y/o oligómeros de PNA para el análisis de patrones de metilación de citosina dentro de las secuencias del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2.

Descripción detallada de la invención

Definiciones:

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

El término "isla de CpG" se refiere a una región contigua de ADN genómico que satisface los criterios de (1) tener una frecuencia de dinucleótidos CpG que corresponde a una "relación observada/esperada" $> 0,6$, y (2) tener un "contenido de GC" $> 0,5$. Las islas de CpG están típicamente, pero no siempre, entre aproximadamente 0,2 hasta aproximadamente 1 KB, o hasta aproximadamente 2 kb de longitud.

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

50 El término "semimetilación" se refiere al estado de metilación de un ADN bicatenario en el que solamente una de las cadenas del mismo está metilado.

El término "AUC", como se usa en el presente documento es una abreviatura para el área bajo una curva. En particular, se refiere al área bajo una curva característica de funcionamiento del receptor (ROC). La curva ROC es una representación de la tasa de verdaderos positivos contra de la tasa de falsos positivos para los diferentes puntos posibles de corte de una prueba diagnóstica. Se muestra el compromiso entre la sensibilidad y la especificidad

- 5 dependiendo del punto de corte seleccionado (cualquier aumento en la sensibilidad vendrá acompañado por una disminución en la especificidad). El área bajo una curva ROC (AUC) es una medida de la precisión de una prueba de diagnóstico (cuanto mayor sea el área mejor, óptima es 1, una prueba aleatoria tendría una curva ROC recostada sobre la diagonal con un área de 0,5; para referencia: JP Egan. Signal Detection Theory and ROC Analysis, Academic Press, Nueva York, 1975).
- El término estatus de metilación o estado de metilación, deben tomarse también en el sentido de un valor que refleja el grado de metilación en una posición CpG.
- 10 A menos que se especifique de otro modo, los términos "hipermetilado" o "sobremetilado" se entiende que significan un nivel de metilación por encima de aquel de un punto de corte especificado, en donde dicho corte puede ser un valor que representa el nivel medio o promedio de la metilación de una población dada, o es preferiblemente un nivel optimizado de corte. El "corte" también se denomina en este documento como un "umbral". En el contexto de la presente invención, los términos "metilado", "hipermetilado" o "sobremetilado" se entiende que incluyen un nivel de metilación por encima de la línea de corte que es cero (0) % de metilación (o sus equivalentes) para todas las posiciones de CpG dentro y asociadas con (por ejemplo, en regiones promotoras o reguladoras) el gen TFAP2E.
- 15 A menos que se especifique de otro modo, los términos "hipometilado" o "submetilado" se entiende que significan un nivel de metilación por debajo de un punto de corte especificado, en donde dicho corte puede ser un valor que representa el nivel medio o promedio de metilación de una población dada, o es preferiblemente un nivel de corte optimizado. El "corte" también se denomina en este documento como un "umbral". En el contexto de la presente invención, los términos "no metilado", "hipometilado" o "submetilado" se considera que incluyen un nivel de metilación
- 20 en el punto de corte que es cero (0) % de metilación (o sus equivalentes) para todas las posiciones CpG dentro y asociadas con el gen TFAP2E (por ejemplo, en regiones promotoras o reguladoras).
- El término "microarreglo" se refiere ampliamente tanto a "microarreglos de ADN," como a "chip(s) de ADN", como se reconoce en la técnica, abarca todos los soportes sólidos reconocidos en la técnica, y abarca todos los métodos para la fijación de moléculas de ácido nucleico a los mismos o la síntesis de los ácidos nucleicos sobre los mismos.
- 25 "Parámetros genéticos" son mutaciones y polimorfismos de genes y secuencias requeridos adicionalmente para su regulación. Se designan en particular como mutaciones, inserciones, supresiones, mutaciones puntuales, inversiones y polimorfismos y, particularmente preferidos, los SNP (polimorfismos de un solo nucleótido).
- "Parámetros epigenéticos" son, en particular, la metilación de citosina. Otros parámetros epigenéticos incluyen, por ejemplo, la acetilación de histonas que, sin embargo, no se puede analizar directamente usando el método descrito pero que, a su vez, se correlacionan con la metilación del ADN.
- 30 El término "reactivo de bisulfito" se refiere a un reactivo que comprende bisulfito, disulfito, sulfito de hidrógeno o combinaciones de los mismos, útiles como se describe aquí para distinguir entre las secuencias de dinucleótidos CpG metiladas y no metiladas.
- 35 El término "ensayo de metilación" se refiere a cualquier ensayo para la determinación del 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 la polimerasa cebada arbitrariamente para ser 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 concentrarse en las regiones que más probablemente contienen dinucleótidos CpG, y descritas por Gonzalzo y colaboradores, Cancer Research 57: 594-599, 1997.
- 40 El término "MethyLight^{MR}" se refiere a la técnica de PCR en tiempo real basada en fluorescencia reconocida en la técnica descrita por Eads y colaboradores, Cancer Res. 59: 2302-2306, 1999.
- El término ensayo de "HeavyMethyl^{MR}", en la realización del mismo implementada en el presente documento, se refiere a un ensayo, en el que las sondas de bloqueo específicas de metilación (también denominadas en este documento como bloqueadores) que cubren las posiciones de CpG entre, o cubiertas por los cebadores de amplificación permiten la amplificación selectiva específica de la metilación de una muestra de ácido nucleico.
- 45 El término ensayo "HeavyMethyl^{MR} MethyLight^{MR}", en la realización del mismo implementada en el presente documento, se refiere a un ensayo HeavyMethyl^{MR} MethyLight^{MR}, que es una variación del ensayo MethyLight^{MR}, en el que el ensayo MethyLight^{MR} se combina con las sondas de bloqueo específicas de metilación que cubren las posiciones de CpG entre los cebadores de amplificación.
- 50 El término "Ms-SNuPE" (extensión del cebador con un solo nucleótido sensible a la metilación) se refiere al ensayo reconocido en la técnica descrito por Gonzalzo y Jones, Nucleic Acids Res. 25: 2529-2531, 1997.

El término "MSP" (PCR específica de la metilación) se refiere al ensayo de metilación reconocido en la técnica descrito por Herman y colaboradores. Proc. Natl. Acad. Sci. USA. 93: 9821 - 9826, 1996, y en la patente estadounidense No. 5.786.146.

5 El término "COBRA" (análisis de restricción combinado con bisulfito) se refiere al ensayo de metilación reconocido en la técnica descrito por Xiong y Laird, Nucleic Acids Res. 25: 2532-2534, 1997.

El término "MCA" (amplificación de una isla de CpG metilada) se refiere al ensayo de metilación descrito por Toyota y colaboradores, Cancer Res. 59: 2307-12, 1999, y en el documento WO 00/26401A1.

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

15 Las "condiciones de hibridación rigurosas", como se define en el presente documento, implican la hibridación a 68°C en una solución de Denhardt 5x/5x SSC /SDS al 1,0%, y lavado en 0,2x SSC/SDS al 0,1% a temperatura ambiente, o involucra al equivalente del mismo reconocido en la técnica (por ejemplo, las condiciones en las que una hibridación se lleva a cabo a 60°C en regulador 2,5 x SSC, seguido por varias etapas de lavado a 37°C en una concentración baja de regulador, y se mantiene estable). Las condiciones moderadamente rigurosas, como se define en el presente documento, implican la inclusión de lavado en 3x SSC a 42°C, o el equivalente del mismo 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 y colaboradores, 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; y Ausubel y colaboradores. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.) en la Unidad 2.10.

25 Los términos "enzimas de restricción específicas de la metilación" o "enzimas de restricción sensibles a la metilación" se entiende que se refieren a una enzima que digiere selectivamente un ácido nucleico dependiente 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 semimetilado, el corte no se llevará a cabo, o con una eficiencia significativamente reducida, si el sitio de reconocimiento está metilado. En el caso de tales enzimas de restricción que cortan específicamente si el sitio de reconocimiento está metilado, el corte no se llevará a cabo, o con una eficiencia significativamente reducida si el sitio de reconocimiento no está metilado. Se prefieren las enzimas de restricción específicas de metilación, la secuencia de reconocimiento que contiene un dinucleótido CG (por ejemplo cgcg o cccggg). También se prefieren para algunas realizaciones enzimas de restricción que no cortan si la citosina en este dinucleótido está metilada en el átomo de carbono C5.

Las "enzimas de restricción específicas cuando no hay metilación" o "enzimas de restricción sensibles a la falta de 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 de metilación".

35 En referencia a las secuencias de matriz de material compuesto, la frase "nucleótidos contiguos" se refiere a una región de secuencia contigua de cualquier secuencia contigua individual de la matriz de material compuesto, pero no incluye una región de la secuencia de matriz compuesta que incluye un "nodo" como se definió aquí anteriormente.

40 El término "TFAP2E" debe entenderse que incluye todas las variantes de transcripción de la misma y todos los elementos promotores y reguladores de las mismas. Por otra parte, ya que se conocen una pluralidad de SNP dentro de dicho gen, debe entenderse que el término incluye todas las variantes de secuencia de las mismas.

El término "precanceroso" y equivalentes del mismo se entiende que significa cualquier trastorno proliferativo celular que está experimentando una transformación maligna. Ejemplos de tales condiciones incluyen, en el contexto de los trastornos proliferativos celulares colorrectales, trastornos proliferativos celulares con un alto grado de displasia y las siguientes clases de adenomas:

45 Nivel 1: penetración de glándulas malignas a través de la mucosa muscular hacia la submucosa, dentro de la cabeza del pólipo

Nivel 2: la misma invasión submucosa, pero presente en la unión de la cabeza con el tallo

Nivel 3: invasión del tallo

50 Nivel 4: invasión de la base del tallo en la conexión con la pared del colon (este nivel se corresponde con el estadio de Dukes A)

La presente invención proporciona un método para la detección de trastornos de proliferación celular en un sujeto que comprende detectar la presencia de metilación de TFAP2E en una muestra biológica aislada de dicho sujeto, en donde la metilación de CpG es indicativa de la presencia de dicho trastorno. Dichos marcadores pueden ser utilizados para el diagnóstico de trastornos proliferativos celulares neoplásicos malignos (preferiblemente cáncer), incluyendo la detección temprana durante las etapas precancerosas de la enfermedad, y, además, para la diferenciación de los trastornos proliferativos celulares benignos de los malignos. La presente invención divulga un método en el que un trastorno proliferativo de células malignas se distingue de un trastorno proliferativo de células benignas, estando dicho método caracterizado porque la presencia de metilación de CpG es indicativa de la presencia de un trastorno proliferativo de células malignas o trastorno precanceroso y la ausencia de la misma es indicativa de la presencia de un trastorno proliferativo celular benigno.

Los marcadores de la presente invención son particularmente eficientes en la detección o para distinguir entre los trastornos proliferativos de células de próstata o, alternativamente, para detectar o distinguir entre los trastornos proliferativos de células de colorrectales, proporcionando de ese modo medios mejorados para la detección temprana, clasificación y tratamiento de dichos trastornos.

La modificación con bisulfito del ADN es una herramienta reconocida en la técnica utilizada para evaluar el estado de metilación de CpG. La 5-metilcitosina es la modificación de base covalente más frecuente en el ADN de las células eucariotas. Desempeña un papel, por ejemplo, en la regulación de la transcripción, en la impronta genética, y en la tumorigénesis. Por lo tanto, la identificación de 5-metilcitosina como componente de información genética es de un interés considerable. 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. Por otra parte, la información epigenética portada por la 5-metilcitosina se pierde completamente, por ejemplo, durante la amplificación por PCR.

El método más frecuentemente utilizado para el análisis de ADN para la presencia de 5-metilcitosina se basa en la reacción específica de bisulfito con citosina por lo que, tras la hidrólisis alcalina posterior, la citosina se convierte en uracilo que corresponde a timina en su comportamiento de apareamiento de bases. Significativamente, sin embargo, la 5-metilcitosina permanece sin modificar bajo estas condiciones. Por consiguiente, el ADN original se convierte de tal manera que la metilcitosina, que originalmente no podía distinguirse de la citosina por su comportamiento de hibridación, ahora puede ser detectada como la única citosina que queda usando técnicas de biología molecular reconocidas en el arte, por ejemplo, mediante amplificación e hibridación, o por secuenciación. Todas estas técnicas se basan en las propiedades diferenciales de apareamiento de bases, que ahora puede ser totalmente explotadas.

La técnica anterior, en términos de sensibilidad, se define por un método que comprende el encerramiento del ADN a analizar en una matriz de agarosa, evitando así la difusión y renaturalización del ADN (el bisulfito sólo reacciona con ADN monocatenario), y el reemplazo de todas las etapas de precipitación y purificación con diálisis rápida (Olek A, y colaboradores, A modified and improved method for bisulfite based cytosine methylation analysis, *Nucleic Acids Res.* 24: 5064-6, 1996). Por tanto, es posible analizar células individuales por el estado de metilación, ilustrando la utilidad y sensibilidad del método. Una visión general de los métodos reconocidos en la técnica para la detección de 5-metilcitosina es proporcionada por Rein, T., y colaboradores, *Nucleic Acids Res.*, 26: 2255, 1998.

La técnica de bisulfito, salvo algunas excepciones (por ejemplo, Zeschnigk M, y colaboradores, *Eur J Hum Genet.* 5: 94-98, 1997), sólo se usa actualmente en investigación. En todos los casos, se amplifican posteriormente fragmentos específicos cortos de un gen conocido con un tratamiento con bisulfito, y, o bien completamente secuenciados (Olek y Walter, *Nat Genet.* 1997 17: 275-6, 1997), sometidos a una o más reacciones de extensión de cebador (Gonzalzo y Jones, *Nucleic Acids Res.*, 25: 2529-31, 1997; WO 95/00669; patente de Estados Unidos No. 6.251.594) para analizar las posiciones de citosina individuales, o tratadas por digestión enzimática (Xiong y Laird, *Nucleic Acids Res.*, 25: 2532-4, 1997). También se ha descrito la detección por hibridación en la técnica (Olek y colaboradores, WO 99/28498). Además, se ha descrito el uso de la técnica de bisulfito para la detección de metilación con respecto a los genes individuales (Grigg & Clark, *Bioessays*, 16: 431-6, 1994; Zeschnigk M, y colaboradores, *Hum Mol Genet.*, 6: 387-95, 1997; Feil R, y colaboradores, *Nucleic Acids Res.*, 22: 695-, 1994; Martin V, y colaboradores, *Gene*, 157: 261-4, 1995; WO 97/46705 y WO 95/15373).

La presente invención proporciona el uso de la técnica de bisulfito, en combinación con uno o más ensayos de metilación, para la determinación del estado de metilación de las secuencias de dinucleótido CpG dentro de la SEQ ID NO: 1. Los dinucleótidos genómicos CpG pueden ser metilados o no metilados (alternativamente conocidos como sobremetilados y submetilados, respectivamente). Sin embargo, los métodos de la presente invención son adecuados para el análisis de muestras biológicas de naturaleza heterogénea, por ejemplo, una baja concentración de células tumorales dentro de un ambiente de sangre o de heces. En consecuencia, al analizar el estado de metilación de una posición de CpG dentro de una muestra tal, la persona experta en la técnica puede usar un ensayo cuantitativo para determinar el nivel (por ejemplo, porcentaje, fracción, relación, proporción o grado) de metilación en una posición particular de CpG a diferencia de un estado de metilación. De acuerdo con el término estatus de metilación o estado de metilación también se debe entender que significa un valor que refleja el grado de metilación en una posición de CpG.

De acuerdo con la presente invención, la determinación del estado de metilación de secuencias de dinucleótidos CpG dentro de la SEQ ID NO: 1 tiene utilidad tanto en el diagnóstico como en la caracterización de trastornos proliferativos celulares. En realizaciones preferidas, se determina el estado de metilación de las posiciones de CpG en la SEQ ID NO: 2, la SEQ ID NO: 2 son regiones particularmente preferidas de la SEQ ID NO: 1 (es decir, la SEQ ID NO: 1 comprende la SEQ ID NO: 2). La determinación del estado de metilación de secuencias de dinucleótidos CpG dentro de la SEQ ID NO: 2 también tienen utilidad en el diagnóstico y caracterización de trastornos proliferativos celulares.

Procedimientos de ensayo de metilación. Se conocen diversos procedimientos de ensayo de metilación en la técnica, y se pueden utilizar junto con la presente invención. Estos ensayos permiten la determinación del estado de metilación de uno o una pluralidad de dinucleótidos CpG (por ejemplo, islas de CpG) dentro de una secuencia de ADN. Tales ensayos implican, entre otras técnicas, secuenciación de ADN del ADN tratado con bisulfito, PCR (para amplificación específica de la secuencia), análisis de transferencia tipo Southern, y el uso de enzimas de restricción sensibles a la metilación.

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

COBRA. El análisis COBRA^{MR} es un ensayo de metilación cuantitativo útil para determinar los niveles de metilación de ADN en los loci de genes específicos en cantidades pequeñas de ADN genómico (Xiong y Laird, Nucleic Acids Res. 25: 2532-2534, 1997). En resumen, se usa digestión con enzimas de restricción para revelar las diferencias de secuencia que dependen de la metilación en los productos de la PCR de ADN tratado con bisulfito de sodio. Las diferencias de secuencia que dependen de la metilación se introducen primero en el ADN genómico por tratamiento estándar con bisulfito de acuerdo con el procedimiento descrito por Frommer y colaboradores (Proc Natl Acad Sci USA. 89: 1827-1831, 1992). Se realiza luego la amplificación por PCR del ADN convertido con bisulfito utilizando cebadores específicos para las islas de CpG de interés, seguido de digestión con endonucleasas de restricción, electroforesis en gel, y detección usando sondas de hibridación marcadas específicas. Los niveles de metilación en la muestra original de ADN se representan por las cantidades relativas de producto de la PCR digerido y no digerido en una forma linealmente cuantitativa a través de un amplio espectro de niveles de metilación del ADN. Además, esta técnica se puede aplicar de forma confiable al ADN obtenido a partir de muestras de tejido embebidas en parafina microdisecionadas.

Los reactivos típicos (por ejemplo, como se podría encontrar en un kit típico a base de COBRA^{MR}) para el análisis COBRA^{MR} pueden incluir, pero no se limitan a: cebadores de PCR para el gen específico (o una secuencia de ADN tratada con bisulfito o isla de CpG); enzima de restricción y regulador apropiado; oligonucleótido de hibridación génica; oligonucleótido de control de la hibridación; kit de marcación de quinasa para la sonda de oligonucleótido; y nucleótidos marcados. Además, los reactivos de conversión de bisulfito pueden incluir: regulador de desnaturalización del ADN; regulador de desulfonación; reactivos o kits de recuperación del ADN (por ejemplo, precipitación, ultrafiltración, columna de afinidad); regulador de desulfonación; y componentes de recuperación de ADN.

Preferiblemente, se utilizan ensayos tales como las reacciones de "MethyLight^{MR}" (una técnica de PCR en tiempo real basada en fluorescencia) (Eads y colaboradores, Cancer Res. 59: 2302-2306, 1999), Ms-SNuPE^{MR} (extensión del cebador en un solo nucleótido sensible a la metilación) (Gonzalzo y Jones, Nucleic Acids Res 25: 2529-2531, 1997), PCR específica de la metilación ("MSP"; Herman y colaboradores, Proc Natl Acad Sci USA. 93: 9821-9826, 1996; patente estadounidense No. 5.786.146), y amplificación de la isla de CpG metilada ("MCA"; Toyota y colaboradores, Cancer Res. 59: 2307-12, 1999) solos o en combinación con otro de estos métodos.

El ensayo "HeavyMethyl^{MR}", la técnica es un método cuantitativo para la evaluación de diferencias de metilación basado en la amplificación específica de metilación de ADN tratado con bisulfito. Las sondas de bloqueo específicas de metilación (también denominadas en este documento como bloqueadoras) que cubre las posiciones de CpG entre, o cubiertas por los cebadores de amplificación permiten la amplificación selectiva específica de la metilación de una muestra de ácido nucleico.

El término ensayo "HeavyMethyl^{MR} MethyLight^{MR}", en la realización del mismo implementado en el presente documento, se refiere a un ensayo HeavyMethyl^{MR} MethyLight^{MR}, que es una variación del ensayo MethyLight^{MR}, en donde el ensayo de MethyLight^{MR} se combina con sondas de bloqueo específicas de metilación que cubren las posiciones de CpG entre los cebadores de amplificación. El ensayo HeavyMethyl^{MR} también se puede usar en combinación con cebadores de amplificación específicos de metilación.

Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico con base en MethyLight^{MR}) para el análisis HeavyMethyl^{MR} puede incluir, pero no se limitan a: cebadores de PCR para los genes específicos (o secuencia de ADN o isla de CpG tratada con bisulfito); el bloqueo de los oligonucleótidos; reguladores y desoxinucleótidos optimizados por PCR; y Taq polimerasa.

MSP. MSP (PCR específica de metilación) permite evaluar el estado de metilación de virtualmente cualquier grupo de sitios CpG dentro de una isla de CpG, independiente del uso de las enzimas de restricción sensibles a la metilación (Herman y colaboradores. Proc. Natl. Acad. Sci. USA. 93: 9821-9826, 1996; patente estadounidense No. 5.786.146). En resumen, el ADN es modificado con bisulfito de sodio que convierte todas las citosinas no metiladas, pero no las metiladas, en uracilo, y subsecuentemente es amplificado con cebadores específicos para el ADN metilado versus el no metilado. MSP requiere solamente cantidades pequeñas de ADN, es sensible al 0,1% de los alelos metilados de un locus dado de una isla de CpG, y puede realizarse sobre el ADN extraído de las muestras embebidas en parafina. Los reactivos típicos (por ejemplo, como puede encontrarse en un kit típico basado en MSP) para el análisis de MSP pueden incluir, pero no se limitan a: cebadores de PCR metilados y no metilados para un gen específico (o secuencia de ADN tratado o isla de CpG), reguladores optimizados de PCR y desoxinucleótidos, y sondas específicas.

MethylLight^{MR}. El ensayo MethylLight^{MR} es un ensayo de metilación cuantitativo de alto rendimiento que utiliza la tecnología de PCR en tiempo real basada en fluorescencia (TaqMan^{MR}) que no requiere manipulaciones adicionales después de la etapa de la PCR (Eads y colaboradores, Cancer Res. 59: 2302-2306, 1999). En resumen, el proceso MethylLight^{MR} comienza con una muestra mixta de ADN genómico que se convierte, en una reacción de bisulfito de sodio, en una combinación mixta de diferencias en secuencia que dependen de la metilación de acuerdo con procedimientos estándar (el proceso de bisulfito convierte los residuos de citosina no metilados en uracilo). A continuación se realiza la PCR basada en fluorescencia en una reacción "sesgada" (con cebadores de PCR que se superponen a dinucleótidos CpG conocidos). La discriminación de secuencia puede producirse tanto a nivel del proceso de amplificación como a nivel del proceso de detección de fluorescencia.

El ensayo MethylLight^{MR} puede ser utilizado como una prueba cuantitativa para los patrones de metilación en la muestra de ADN genómico, en donde se produce discriminación de la secuencia a nivel de 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 solapa un sitio de metilación particular putativo. Se proporciona un control no sesgado para la cantidad de ADN que entra mediante una reacción en la que ni los cebadores, ni la sonda recubren ninguno de los dinucleótidos CpG. Alternativamente, se logra una prueba cualitativa para la metilación genómica mediante el sondeo de la combinación sesgada de la PCR, ya sea con oligonucleótidos de control que no "cubren" sitios de metilación conocidos (una versión basada en fluorescencia de las técnicas HeavyMethyl^{MR} y MSP), o con oligonucleótidos que cubren sitios de metilación potenciales.

El proceso MethylLight^{MR} puede utilizarse con cualquier sonda adecuada, por ejemplo "TaqMan®", LightCycler® etc. Por ejemplo, se trata el ADN genómico bicatenario con bisulfito de sodio y se lo somete a uno de los dos conjuntos de reacciones PCR usando sondas TaqMan®; por ejemplo, con los cebadores de MSP y/o oligonucleótidos bloqueadores HeavyMethyl y la sonda TaqMan®. La sonda TaqMan® está marcada doblemente con moléculas "informadoras" e "inactivadoras", y está diseñada para ser específica para una región con un contenido relativamente alto de GC de manera que se funde a una temperatura aproximadamente a 10°C más alta en el ciclo de la PCR que los cebadores directo o inverso. Esto permite que la sonda TaqMan® permanezca hibridada completamente durante la etapa de hibridación/extensión de la PCR. Como la Taq polimerasa sintetiza enzimáticamente una nueva cadena durante la PCR, con el tiempo alcanzará a la sonda TaqMan® hibridada. La actividad de endonucleasa 5' a 3' de la Taq polimerasa desplazará entonces a la sonda TaqMan® digiriéndola para liberar la molécula informadora fluorescente para detección cuantitativa de su señal ahora inactivada usando un sistema de detección de fluorescencia en tiempo real.

Los reactivos típicos (por ejemplo, como podría encontrarse en un kit típico basado en MethylLight^{MR}) para el análisis con MethylLight^{MR} pueden incluir, pero no se limitan a: cebadores de PCR para el gen específico (o secuencia de ADN o isla de CpG tratada con bisulfito); sondas TaqMan® o LightCycler®; reguladores y desoxinucleótidos optimizados para PCR; y Taq polimerasa.

El ensayo QM^{MR} (metilación cuantitativa) es una prueba cuantitativa alternativa para los patrones de metilación en muestras de ADN genómico, en donde la discriminación de la secuencia ocurre a nivel de hibridación de la sonda. En esta versión cuantitativa, la reacción de PCR proporciona para una amplificación no sesgada en presencia de una sonda fluorescente que solapa un sitio de metilación putativo particular. Se proporciona un control no sesgado para la cantidad de ADN de entrada mediante una reacción en la que ni los cebadores, ni la sonda están sobre ninguno de los dinucleótidos CpG. Alternativamente, se logra una prueba cualitativa para metilación genómica mediante el sondeo de la combinación sesgada de PCR, ya sea con oligonucleótidos de control que no "cubren" sitios de metilación conocidos (una versión basada en fluorescencia de las técnicas HeavyMethyl^{MR} y MSP), o con oligonucleótidos que cubren sitios de metilación potenciales.

El proceso QM^{MR} se puede utilizar con cualquier sonda adecuada, por ejemplo "TaqMan®", LightCycler® etc. en el proceso de amplificación. Por ejemplo, se trata el ADN genómico bicatenario con bisulfito de sodio y se somete a cebadores no sesgados y la sonda TaqMan®. La sonda TaqMan® está doblemente marcada con moléculas "informadoras" e "inactivadoras" fluorescentes, y está diseñada para ser específica para una región con un contenido relativamente alto de GC de manera que se funde a una temperatura aproximadamente 10°C más alta en el ciclo de PCR que los cebadores directo o inverso. Esto permite que la sonda TaqMan® permanezca completamente hibridada durante la etapa de hibridación/extensión de la PCR. Como la Taq polimerasa sintetiza enzimáticamente una nueva

cadena durante la PCR, eventualmente alcanzará a la sonda TaqMan® hibridada. La actividad de la Taq polimerasa 5' a 3' endonucleasa desplazará entonces la sonda TaqMan® dirigiéndola para liberar la molécula informadora fluorescente para detección cuantitativa de su señal ahora no inactivada usando un sistema de detección de fluorescencia en tiempo real.

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

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

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

20 Se determinó que la secuencia genómica de acuerdo con la SEQ ID NO: 1 a la SEQ ID NO: 2, y de variantes de las mismas tratadas de origen no natural de acuerdo con la SEQ ID NO: 3 a la SEQ ID NO: 10, tiene una novedosa utilidad para la detección temprana, clasificación y/o tratamiento de trastornos proliferativos celulares, en particular, trastornos proliferativos celulares colorrectal y/o de próstata.

25 En una realización de la invención, el método comprende las siguientes etapas: i) poner en contacto el ADN genómico (preferiblemente aislado a partir de fluidos corporales) obtenido del sujeto con al menos un reactivo, o serie de reactivos que distinguen entre dinucleótidos CpG metilados y no metilados dentro del gen TFAP2E (incluyendo su promotor y regiones reguladoras); y ii) detectar, o detectar y distinguir entre dos o más trastornos proliferativos celulares de colon o de próstata.

30 El ADN genómico puede ser aislado por cualquier medio estándar en la técnica, incluyendo el uso de kits disponibles en el mercado. En resumen, cuando se encapsula el ADN de interés en una membrana celular, la muestra biológica debe ser cortada y lisada mediante medios enzimáticos, químicos o mecánicos. La solución de ADN puede entonces ser liberada de las proteínas y otros contaminantes, por ejemplo, por digestión con proteinasa K. El ADN genómico se recupera entonces de la solución. Esto puede llevarse a cabo por medio de una variedad de métodos que incluyen precipitación en solución salina, extracción orgánica o unión del ADN a un soporte de fase sólida. La elección del método se verá afectada por varios factores incluyendo el tiempo, el coste y la cantidad necesaria de ADN. Todos los tipos de muestras clínicas que comprenden la materia neoplásica o la materia precancerosa son adecuados para uso en el presente método, se prefieren las líneas celulares, cortes histológicos, biopsias, tejido incluido en parafina, fluidos corporales, heces, efluente colónico, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre y combinaciones de los mismos. Los fluidos corporales son la fuente preferida de ADN; particularmente preferida es la orina, ya sea normalmente evacuados o masaje postprostático.

35 La muestra de ADN genómico se trata luego con al menos un reactivo, o serie de reactivos que distinguen entre los dinucleótidos CpG metilados y no metilados dentro de al menos una región objetivo del ADN genómico, en donde la región objetivo comprende, o se hibrida bajo condiciones rigurosas con una secuencia de al menos 16 nucleótidos contiguos de al menos una secuencia seleccionada del grupo que consiste en la SEQ ID NO: 1 a la SEQ ID NO: 2, respectivamente, en donde dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótido CpG.

40 Se prefiere particularmente que dicho reactivo convierta las bases de citosina que están sin metilar en la posición 5' en uracilo, timina, u otra base que sea diferente a la 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.

45 En donde se trata la muestra de ADN genómico de una manera tal que las bases de citosina que están sin metilar en la posición 5' se convierten en uracilo, timina, u otra base que es diferente a la citosina en términos de comportamiento de

hibridación, se prefiere que este tratamiento se lleve a cabo con bisulfito (sulfito de hidrógeno, disulfito) y posterior hidrólisis alcalina. Este tratamiento da como resultado la conversión de la SEQ ID NO: 1 a la SEQ ID NO: 2 a la SEQ ID NO: 3 a la SEQ ID NO: 6 (respectivamente) en donde dichos dinucleótidos CpG están metilados o la SEQ ID NO: 7 a la SEQ ID NO: 10 en donde dichos dinucleótidos CpG están sin metilar (véase la Tabla 1).

5 El ADN tratado se analiza a continuación, con el fin de determinar el estado de metilación de las secuencias de genes objetivo (TFAP2E antes del tratamiento). Se prefiere particularmente que la región objetivo comprenda, o se hibride bajo condiciones rigurosas con al menos 16 nucleótidos contiguos de TFAP2E. Se prefiere que la secuencia de dicho gen de acuerdo con SEQ ID NO: 1 se analice, se prefiere particularmente que la subregión de la misma de acuerdo con la SEQ ID NO: 2 se analicen. El método de análisis se puede seleccionar de los conocidos en la técnica, incluyendo los que se enumeran en este documento. Particularmente preferidos son MethyLight^{MR}, MSP y el uso de oligonucleótidos de bloqueo (HeavyMethyl^{MR}) como se describe en el presente documento. Además, se prefiere que cualquiera de los oligonucleótidos utilizados en dicho análisis (incluyendo cebadores, oligonucleótidos de bloqueo y sondas de detección) deben ser complementarios inversos, idénticos, o hibridar bajo condiciones rigurosas o altamente rigurosas con al menos un segmento de 16 pares de bases de largo de las secuencias de bases de una o más de la SEQ ID NO: 3 a la SEQ ID NO: 10 y las secuencias complementarias de las mismas.

La metilación aberrante, más específicamente hipermetilación de TFAP2E (así como las regiones promotoras y/o reguladoras) se asocia con la presencia de trastornos proliferativos celulares neoplásicos, y es particularmente prevalente en carcinomas colorrectales y de próstata. En consecuencia, cuando una muestra biológica presenta algún grado de metilación, dicha muestra se debe determinar como neoplásica.

20 El análisis de un gen TFAP2E permite por primera vez detectar, o detectar y distinguir entre dos o más trastornos proliferativos de células colon o de próstata que permitió una sensibilidad mayor que o igual al 80% y una especificidad mayor que o igual al 80%. La sensibilidad se calcula como: {neoplasia detectada/todas las neoplasias}, por ejemplo: {neoplasia de colon detectada/todas las neoplasias de colon}; y la especificidad se calcula como (negativos no detectados/negativos totales).

25 La neoplasia de colon se define aquí como todas las neoplasias de colon y adenomas mayores a 1 cm, o subconjuntos de las mismas. Los negativos se pueden definir como individuos sanos.

En una realización, el método divulga el uso de TFAP2E o las regiones promotoras y/o reguladoras del mismo) como un marcador para la diferenciación, detección y distinción de los trastornos proliferativos celulares (en particular, trastornos neoplásicos, de colon o de próstata).

30 Dicho método puede ser habilitado por medio de cualquier análisis de la expresión de un ARN transcrito a partir del mismo o un polipéptido o proteína traducida a partir de dicho ARN, preferiblemente por medio del análisis de la expresión de ARNm o el análisis de expresión del polipéptido. En consecuencia, la presente solicitud también divulga ensayos y métodos de diagnóstico, tanto cuantitativos como cualitativos para la detección de la expresión del gen TFAP2E en un sujeto y determinar a partir de los mismos la presencia o ausencia de cáncer en dicho sujeto.

35 La expresión aberrante del ARNm transcrito a partir del gen TFAP2E se asocia con la presencia de cáncer en un sujeto. La subexpresión (y/o presencia de metilación) se asocia con la presencia de cáncer, y viceversa, la sobreexpresión (y/o ausencia de metilación) se asocia con la ausencia de cáncer.

40 Para detectar la presencia de ARNm que codifica un gen o secuencia genómica, se obtiene una muestra de un paciente. La muestra puede ser cualquier muestra adecuada que comprende material celular del tumor. Los tipos adecuados de muestras incluyen líneas celulares, cortes histológicos, biopsias, tejido incluido en parafina, fluidos corporales, heces, efluente colónico, 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 dichos tipos de muestras sean heces o fluidos corporales seleccionados del grupo que consiste de efluente colónico, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre.

45 La muestra puede ser tratada para extraer el ARN contenido en la misma. Se analiza a continuación, el ácido nucleico resultante de la muestra. Se conocen muchas técnicas en el estado de la técnica para determinar los niveles absolutos y relativos de la expresión génica; las técnicas que se utilizan comúnmente adecuadas para uso en la presente invención incluyen la hibridación in situ (por ejemplo FISH), análisis tipo Northern, ensayos de protección de RNasa (RPA), microarreglos y técnicas con base en PCR, tales como PCR cuantitativa y PCR de despliegue diferencial o cualquier otro método de detección de ácido nucleico.

Se prefiere particularmente el uso de la técnica de reacción en cadena de polimerización/transcripción inversa (RT-PCR). El método de RT-PCR es bien conocido en la técnica (por ejemplo, véase Watson y Fleming, citado más arriba).

El método de RT-PCR puede llevarse a cabo como sigue. Se aísla el ARN celular total, por ejemplo, mediante el

método estándar de isotiocianato de guanidinio y se transcribe de forma inversa el ARN total. El método de transcripción inversa implica la síntesis de ADN en una plantilla de ARN utilizando una enzima transcriptasa inversa y un cebador oligonucleótido en el extremo 3' y/o cebadores hexámeros aleatorios. El ADNc así producido se amplifica luego por medio de PCR. (Belyavsky y colaboradores, Nucl Acid Res 17: 2919-2932, 1989; Krug y Berger, Methods in Enzymology, Academic Press, N.Y., Vol. 152, páginas 316-325, 1987). Se prefiere adicionalmente la variante en "tiempo real" de RT-PCR, en donde el producto de la PCR se detecta por medio de sondas de hibridación (por ejemplo, TaqMan, LightCycler, Molecular Beacons and Scorpion) o verde SYBR. La señal detectada de las sondas o verde SYBR es entonces cuantificada bien por referencia a una curva estándar o mediante la comparación de los valores de Ct con aquella de un estándar de calibración. El análisis de los genes de mantenimiento se usa frecuentemente para normalizar los resultados.

En análisis total de transferencias tipo Northern o poli(A) + ARNm se corre en un gel de agarosa desnaturante y se detecta por hibridación con una sonda marcada en el gel seco mismo o en una membrana. La señal resultante es proporcional a la cantidad de ARN objetivo en la población de ARN.

La comparación de las señales de dos o más poblaciones de células o tejidos revela diferencias relativas en los niveles de expresión génica. La cuantificación absoluta se puede realizar mediante la comparación de la señal con una curva estándar generada usando cantidades conocidas de una transcripción in vitro correspondiente al ARN objetivo. Se usa a menudo el análisis de genes de mantenimiento, donde se espera que los niveles de expresión de los genes permanezcan relativamente constantes independientemente de las condiciones, para normalizar los resultados, eliminando cualquiera de las diferencias aparentes causadas por transferencia desigual de ARN a la membrana o carga desigual de ARN en el gel.

El primer paso en el análisis de Northern es aislar ARN intacto, puro de las células o tejido de interés. Debido a que las transferencias tipo Northern distinguen los ARN por tamaño, la integridad de la muestra influye en el grado en el que se localiza una señal en una sola banda. Las muestras de ARN parcialmente degradadas ocasionarán que la señal se esparza o distribuya sobre varias bandas con una pérdida global de la sensibilidad y posiblemente una interpretación errónea de los datos. En el análisis de transferencia de Northern, se pueden utilizar las sondas de ADN, ARN y oligonucleótido y se marcan preferiblemente estas sondas (por ejemplo, marcadores radiactivos, marcadores de masa o marcadores fluorescentes). El tamaño del ARN objetivo, no de la sonda, determinará el tamaño de la banda detectada, por lo que métodos tales como la marcación cebada al azar, que genera sondas de longitudes variables, son adecuados para la síntesis de la sonda. La actividad específica de la sonda determinará el nivel de sensibilidad, por lo que se prefiere utilizar las sondas con altas actividades específicas.

En un ensayo de protección de RNasa, se hibridan en solución el objetivo de ARN y una sonda de ARN de una longitud definida. Después de la hibridación, se digiere el ARN con RNasas específicas para los ácidos nucleicos de cadena sencilla para eliminar cualquier ARN objetivo de cadena sencilla no hibridado y la sonda. Las RNasas se inactivan, y se separa el ARN, por ejemplo, por electroforesis desnaturante en gel de poliacrilamida. La cantidad de sonda intacta de ARN es proporcional a la cantidad de ARN objetivo en la población de ARN. Se puede usar RPA para la cuantificación relativa y absoluta de la expresión génica y también para el mapeo de la estructura del ARN, tales como los sitios de inicio de la transcripción y los límites del intrón/exón. El ensayo de protección de la RNasa es preferible que el análisis de transferencia tipo Northern, ya que generalmente tiene un límite inferior de detección.

Las sondas de ARN antisentido utilizadas en RPA son generadas por transcripción in vitro de un molde de ADN con un punto final definido y están típicamente en el intervalo de 50-600 nucleótidos. El uso de sondas de ARN que incluye secuencias adicionales no homólogas al ARN objetivo permite que el fragmento protegido se diferencie de la sonda de longitud completa. Se usan típicamente sondas de ARN en lugar de sondas de ADN debido a la facilidad de generación de las sondas de ARN de cadena sencilla y la reproducibilidad y confiabilidad de la digestión del dúplex de ARN:ARN con RNasas (Ausubel y colaboradores, 2003), se prefieren particularmente sondas con alta actividad específica.

Se prefiere particularmente el uso de microarreglos. El proceso de análisis de microarreglos se puede dividir en dos partes principales. En primer lugar está la inmovilización de secuencias de genes conocidos en portaobjetos de vidrio u otro soporte sólido seguido por la hibridación del ADNc marcado con fluorescencia (que comprende las secuencias que van a ser interrogadas) con los genes conocidos inmovilizados sobre el portaobjetos de vidrio (u otra fase sólida). Después de la hibridación, se escanean los arreglos utilizando un escáner fluorescente de microarreglos. El análisis de la intensidad relativa de fluorescencia de diferentes genes proporciona una medida de las diferencias en la expresión génica.

Los arreglos de ADN pueden generarse mediante la inmovilización de oligonucleótidos previamente sintetizados sobre portaobjetos de vidrio preparados u otras superficies sólidas. En este caso, se fabrican y preparan secuencias de genes representativos utilizando métodos estándar de síntesis y purificación de oligonucleótidos. Estas secuencias sintetizadas de genes son complementarias a la(s) transcripción(es) de ARN de los genes de interés (en este caso TFAP2E) y tienden a ser secuencias más cortas en el intervalo de 25-70 nucleótidos. En una realización preferida, dichos oligonucleótidos o polinucleótidos comprenden al menos 9, 18 o 25 bases de una secuencia complementaria a o que hibrida con dicho transcrito de ARN. Como alternativa, los oligos inmovilizados pueden sintetizarse químicamente in

situ en la superficie del portaobjetos. La síntesis de oligonucleótidos in situ implica la adición consecutiva de los nucleótidos apropiados a las manchas sobre los microarreglos; las manchas que no reciben un nucleótido se protegen durante cada etapa del proceso usando máscaras físicas o virtuales. Preferiblemente, dichos ácidos nucleicos sintetizados son ácidos nucleicos bloqueados.

5 En los experimentos de microarreglos de perfiles de expresión, los moldes de ARN utilizados son representativos del perfil de transcripción de las células o tejidos en estudio. Se aísla primero el ARN de las poblaciones de células o tejidos que se van a comparar. Cada muestra de ARN se utiliza a continuación como molde para generar ADNc marcado de forma fluorescente a través de una reacción de transcripción inversa. La marcación fluorescente del ADNc puede realizarse por cualquiera método de marcación directo o indirecto. Durante la marcación directa, se incorporan directamente los nucleótidos modificados de manera fluorescente (por ejemplo, Cy@3-dCTP o Cy@5-dCTP) en el ADNc durante la transcripción inversa. Alternativamente, la marcación indirecta puede lograrse mediante la incorporación de nucleótidos modificados con aminoalilo durante la síntesis de ADNc y luego conjugación de un colorante de éster de N-hidroxisuccinimida (NHS) al ADNc modificado con aminoalilo después de la reacción de transcripción inversa está completa. Alternativamente, la sonda puede estar sin marcar, pero puede ser detectable mediante la unión específica con un ligando que está marcado, ya sea directa o indirectamente. Las etiquetas y métodos adecuados para marcar ligandos (y sondas) son conocidos en la técnica, e incluyen, por ejemplo, marcadores radiactivos que pueden ser incorporados por métodos conocidos (por ejemplo, traducción de la muesca o tratamiento con quinasa).

Otros marcadores adecuados incluyen, pero no se limitan a biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), enzimas, anticuerpos, y similares.

20 Para realizar el análisis de la expresión génica diferencial, se marca el ADNc generado a partir de diferentes muestras de ARN con Cy@3. Se purifica el ADNc marcado resultante para eliminar los nucleótidos no incorporados, el colorante libre y el ARN residual. Después de la purificación, se hibridan las muestras de ADNc con el microarreglo. La rigurosidad de la hibridación se determina por un número de factores durante la hibridación y durante el procedimiento de lavado, incluyendo la temperatura, fuerza iónica, período de tiempo y concentración de formamida. Estos factores se describen, por ejemplo, en Sambrook y colaboradores (Molecular Cloning: A Laboratory Manual, segunda ed., 1989). El microarreglo se escanea después de la hibridación utilizando un escáner fluorescente de microarreglos. La intensidad de la fluorescencia de cada punto indica el nivel de expresión del gen analizado; los puntos brillantes corresponden a genes fuertemente expresados, mientras que las manchas oscuras indican expresión débil.

30 Una vez que se obtienen las imágenes, se debe analizar los datos sin procesar. En primer lugar, la fluorescencia de fondo se debe restar de la fluorescencia de cada mancha. Los datos se normalizaron luego con una secuencia de control, tal como los ácidos nucleicos añadidos exógenamente (preferiblemente ARN o ADN), o un panel de genes de mantenimiento que da cuenta de cualquier hibridación no específica, las imperfecciones del arreglo o la variabilidad en la configuración del arreglo, la marcación del ADNc, la hibridación o el lavado. La normalización de los datos permite comparar los resultados de múltiples arreglos.

35 Otro aspecto de la divulgación se refiere a un kit para uso en el diagnóstico de cáncer en un sujeto de acuerdo con los métodos de la presente invención, comprendiendo dicho kit: un medio para medir el nivel de transcripción del gen TFAP2E. En una realización preferida los medios para medir el nivel de transcripción comprenden oligonucleótidos o polinucleótidos capaces de hibridar bajo condiciones rigurosas o moderadamente rigurosas con los productos de transcripción de TFAP2E.

40 En una realización más preferida, el nivel de transcripción se determina por técnicas seleccionadas del grupo de análisis de transferencias tipo Northern, PCR con transcriptasa inversa, PCR en tiempo real, protección de ARNasa, y microarreglos. En otra realización de la invención, el kit comprende además un medio para obtener una muestra biológica del paciente. Se prefiere un kit, que comprende además un recipiente que es lo más preferiblemente adecuado para contener los medios para medir el nivel de la transcripción y la muestra biológica del paciente, y lo más preferiblemente comprende además instrucciones para el uso y la interpretación de los resultados del kit.

45 En una realización preferida, el kit comprende (a) una pluralidad de oligonucleótidos o polinucleótidos capaces de hibridar bajo condiciones rigurosas o moderadamente rigurosas con los productos de transcripción del gen TFAP2E; (b) un recipiente, preferiblemente adecuado para contener los oligonucleótidos o polinucleótidos y una muestra biológica del paciente, que comprende los productos de transcripción en donde los oligonucleótidos o polinucleótidos pueden hibridar bajo condiciones rigurosas o moderadamente rigurosas con los productos de la transcripción, (c) medios para detectar la hibridación de (b); y, opcionalmente, (d) instrucciones para el uso y la interpretación de los resultados del kit.

55 El kit también puede contener otros componentes, tales como regulador de hibridación (donde los oligonucleótidos se utilizan como una sonda) empacados en un recipiente separado. Alternativamente, cuando se utilizan los oligonucleótidos para amplificar una región objetivo, el kit puede contener, empacados en recipientes separados, una polimerasa y un regulador de reacción optimizado para la extensión del cebador mediada por la polimerasa, tal como PCR. Preferiblemente, dicha polimerasa es una transcriptasa inversa. Se prefiere además que dicho kit contenga

además un reactivo de ARNasa.

La presente solicitud divulga además métodos para la detección de la presencia del polipéptido codificado por dichas secuencias de genes en una muestra obtenida de un paciente.

5 Los niveles aberrantes de expresión de polipéptidos de los polipéptidos codificados por el gen TFAP2E se asocian con la presencia de cáncer.

La subexpresión de dichos polipéptidos se asocia con la presencia de cáncer.

10 Se puede utilizar cualquier método conocido en la técnica para detectar polipéptidos. Tales métodos incluyen, pero no se limitan a espectrometría de masa, inmunodifusión, inmunoelectroforesis, métodos inmunoquímicos, ensayos de aglutinante-ligando, técnicas de inmunohistoquímica, ensayos de aglutinación y del complemento (por ejemplo, véase Basic and Clinical Immunology, Sites and Terr, eds., Appleton and Lange, Norwalk, Conn. páginas 217-262, 1991). Se prefieren los métodos de inmunoensayo de aglutinante-ligando que incluyen la reacción de anticuerpos con un epítipo o epítipos y desplazar competitivamente un polipéptido marcado o derivado del mismo.

Ciertas realizaciones de la presente divulgación comprenden el uso de anticuerpos específicos para el polipéptido codificado por el gen TFAP2E.

15 Dichos anticuerpos son útiles para el diagnóstico de cáncer. En ciertas realizaciones, la producción de anticuerpos monoclonales o policlonales se puede inducir mediante el uso de un epítipo codificado por un polipéptido del gen TFAP2E como un antígeno. Tales anticuerpos pueden a su vez ser usados para detectar polipéptidos expresados como marcadores de diagnóstico de cáncer. Los niveles de tales polipéptidos presentes pueden cuantificarse por métodos convencionales.

20 La unión anticuerpo-polipéptido puede ser detectada y cuantificada por una variedad de medios conocidos en la técnica, tales como la marcación con ligandos fluorescentes o radiactivos. La solicitud divulga adicionalmente kits para realizar los procedimientos antes mencionados, en donde dichos kits contienen anticuerpos específicos para los polipéptidos investigados.

25 Numerosos inmunoensayos de unión de polipéptidos competitivos y no competitivos son bien conocidos en la técnica. Los anticuerpos empleados en tales ensayos pueden estar no marcados, por ejemplo como se utiliza en pruebas de aglutinación, o marcados para uso en una amplia variedad de métodos de ensayo. Las etiquetas que pueden usarse incluyen radionúclidos, enzimas, fluorescentes, quimioluminiscentes, sustratos enzimáticos o cofactores, inhibidores de enzimas, partículas, colorantes y similares. Los ensayos preferidos incluyen, pero no se limitan a radioinmunoensayo (RIA), inmunoensayos enzimáticos, por ejemplo, ensayo de inmunoabsorción ligado a enzimas (ELISA), inmunoensayos fluorescentes y similares. Se pueden elaborar anticuerpos monoclonales o policlonales o epítipos de los mismos para uso en inmunoensayos mediante cualquiera de una serie de métodos conocidos en la técnica.

35 En una realización alternativa del método, se pueden detectar las proteínas por medio de análisis de transferencia tipo Western. Dicho análisis es estándar en la técnica; en resumen, las proteínas se separan por medio de electroforesis, por ejemplo SDS-PAGE. Las proteínas separadas se transfieren a continuación a una membrana adecuada (o papel), por ejemplo, nitrocelulosa, conservando la separación espacial lograda por electroforesis. A continuación se incuba la membrana con un agente de bloqueo para unirse a los lugares adhesivos restantes en la membrana; los agentes utilizados comúnmente incluyen proteína genérica (por ejemplo, proteína de la leche). A continuación se añade un anticuerpo específico para la proteína de interés, estando dicho anticuerpo marcado de forma detectable, por ejemplo, mediante colorantes o medios enzimáticos (por ejemplo, fosfatasa alcalina o peroxidasa de rábano picante). Después se detecta la localización del anticuerpo en la membrana.

40 En una realización alternativa del método, se pueden detectar las proteínas por medio de inmunohistoquímica (el uso de anticuerpos para sondear antígenos específicos en una muestra). Dicho análisis es estándar en la técnica, en donde la detección de antígenos en tejidos es conocida como inmunohistoquímica, mientras que la detección en las células cultivadas se denomina generalmente inmunocitoquímica. En resumen, se detecta el anticuerpo primario mediante la unión a su antígeno específico. El complejo antígeno-anticuerpo se une luego mediante un anticuerpo conjugado a una enzima secundaria. En presencia del sustrato y el cromógeno necesarios, se detecta la enzima de unión de acuerdo a los depósitos coloreados en los sitios de unión antígeno-anticuerpo. Existe una amplia gama de tipos adecuados de muestras, afinidad de antígeno-anticuerpo, tipos de anticuerpos, y métodos de mejora de la detección. Por lo tanto, se pueden determinar las condiciones óptimas para la detección inmunohistoquímica o inmunocitoquímica por la persona experta en la técnica para cada caso individual.

50 Un enfoque para preparar anticuerpos para un polipéptido es la selección y preparación de una secuencia de aminoácidos de la totalidad o parte del polipéptido, sintetizar químicamente la secuencia de aminoácidos e inyectarla en un animal apropiado, normalmente un conejo o un ratón (Milstein y Kohler, Nature 256: 495-497, 1975; Gultre y Milstein,

Methods in Enzymology: Immunochemical Techniques 73: 1 - 46, Langone y Banatis eds., Academic Press, 1981). Los métodos para la preparación de los polipéptidos o epítomos de los mismos incluyen, pero no se limitan a, síntesis química, técnicas de ADN recombinante o aislamiento de muestras biológicas.

5 En el último paso del método se determina el diagnóstico del paciente, por lo que la subexpresión (ARNm de TFAP2E o polipéptidos) es indicativa de la presencia de cáncer o de un trastorno de proliferación celular que experimenta una transformación maligna. El término subexpresión se entiende que significa la expresión en un nivel detectado menor que un corte predeterminado que puede ser seleccionada del grupo que consiste de la media, la mediana o un valor de umbral optimizado.

10 Otro aspecto de la solicitud divulga un kit para su uso en el diagnóstico de cáncer o un trastorno proliferativo celular que experimenta una transformación maligna en un sujeto de acuerdo con los métodos de la presente invención, que comprende: un medio para detectar polipéptidos TFAP2E. Los medios para la detección de los polipéptidos comprenden, preferiblemente, anticuerpos, derivados de anticuerpos, o fragmentos de anticuerpos. Los polipéptidos se detectan más preferiblemente por medio de transferencias tipo Western utilizando un anticuerpo marcado. En otro ejemplo, el kit comprende además medios para obtener una muestra biológica del paciente. Se prefiere un kit, que
15 comprende además un recipiente adecuado para contener los medios para la detección de los polipéptidos en la muestra biológica del paciente, y lo más preferiblemente comprende además instrucciones para el uso y la interpretación de los resultados del kit. En una realización preferida, el kit comprende: (a) un medio para detectar polipéptidos TFAP2E; (b) un recipiente adecuado para contener dichos medios y la muestra biológica del paciente, que comprende los polipéptidos donde los medios pueden formar complejos con los polipéptidos; (c) un medio para detectar los complejos de (b); y opcionalmente (d) instrucciones para el uso y la interpretación de los resultados del kit.

20 El kit también puede contener otros componentes tales como reguladores o soluciones adecuadas para bloqueo, lavado o recubrimiento, envasados en un recipiente separado.

25 Las realizaciones particulares de la presente invención proporcionan una nueva aplicación del análisis de los niveles de metilación y/o patrones dentro de dichas secuencias que permiten una detección precisa, caracterización y/o tratamiento de cáncer de trastornos proliferativos celulares de próstata y/o colorrectal. La detección temprana del cáncer está directamente relacionada con el pronóstico de la enfermedad, y el método descrito permite así al médico y al paciente tomar mejores y más informadas decisiones de tratamiento.

Mejoras adicionales

30 La presente invención proporciona nuevos usos para la secuencia genómica de la SEQ ID NO: 1, Y MÁS PREFERIBLEMENTE la SEQ ID NO: 2. Las realizaciones adicionales proporcionan variantes de la SEQ ID NO: 1 a la SEQ ID NO: 2, así como oligonucleótidos y/o oligómeros de PNA para el análisis de los patrones de metilación de citosina dentro de la SEQ ID NO: 1 a la SEQ ID NO: 2.

35 Un objetivo de la divulgación comprende el análisis del estado de metilación de uno o más dinucleótidos CpG dentro de la SEQ ID NO: 1 y las secuencias complementarias de la misma, y más preferiblemente la SEQ ID NO: 2 y secuencias complementarias de la misma.

40 La descripción proporciona ácidos nucleicos tratados, derivados de la SEQ ID NO: 1 a la SEQ ID NO: 2 genómica, en donde el tratamiento es adecuado para convertir al menos una base de citosina no metilada de la secuencia genómica de ADN en uracilo u otra base que es diferente en forma detectable a la citosina en términos de hibridación. Las secuencias genómicas en cuestión pueden comprender uno o más posiciones consecutivas de CpG metiladas. Dicho tratamiento comprende preferiblemente el uso de un reactivo seleccionado del grupo que consiste en bisulfito, sulfito de hidrógeno, disulfito, y sus combinaciones. En una realización preferida de la invención, la invención proporciona un ácido nucleico modificado de origen no natural que comprende una secuencia de al menos 16 bases de nucleótidos contiguos de longitud de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10. En otras formas de realización preferidas de la divulgación, dicho ácido nucleico es de 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: 3 a la SEQ ID NO: 10. Se prefiere particularmente una molécula de ácido nucleico que no es idéntica o complementaria a toda o una porción de las secuencias de la SEQ ID NO: 3 a la SEQ ID NO: 10, pero no la SEQ ID NO: 1 a la SEQ ID NO: 2 u otro ADN de origen natural.

45 Se prefiere que dicha secuencia comprenda al menos un dinucleótido CpG, TpA o CpA y secuencias complementarias de las mismas. Las secuencias de la SEQ ID NO: 3 a la SEQ ID NO: 10 proporcionan versiones modificadas de origen no natural del ácido nucleico de acuerdo con la SEQ ID NO: 1 a la SEQ ID NO: 2, 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 de la siguiente manera. Para cada ADN genómico de cadena sentido, por ejemplo, la SEQ ID NO: 1, se divulgan cuatro versiones convertidas. Una primera versión en donde "C" se convierte en "T", pero
50 "CpG" sigue siendo "CpG" (es decir, corresponde al caso en el que, para la secuencia genómica, todos los residuos de

"C" de secuencias de dinucleótidos CpG están metilados y por lo tanto no convertidos); una segunda versión divulga el complemento de la secuencia genómica de ADN divulgada (es decir, cadena antisentido), en donde "C" se convierte en "T", pero "CpG" sigue siendo "CpG" (es decir, corresponde al caso donde, para todos los residuos de "C" de las secuencias de dinucleótidos CpG están metiladas y por consiguiente no se convierten). Las secuencias convertidas 'no metiladas' de la SEQ ID NO: 1 a la SEQ ID NO: 2 corresponden a la SEQ ID NO: 3 a la SEQ ID NO: 6. Se proporciona una tercera versión químicamente convertida de cada una de las secuencias genómicas, en donde "C" se convierte en "T" para todos los residuos de "C", incluyendo aquellas secuencias del dinucleótido "CpG" (es decir, corresponde al caso en el que, para las secuencias genómicas, todos los residuos de "C" de secuencias de dinucleótidos CpG no están metilados); una versión químicamente convertida final del cada secuencia, divulga el complemento de la secuencia genómica divulgada de ADN (es decir, cadena antisentido), en donde "C" se convierte en "T" para todos los residuos de "C", incluyendo aquellos de las secuencias de dinucleótidos "CpG" (es decir, corresponde al caso en que, para el complemento (hebra antisentido) de cada secuencia genómica, todos los residuos de "C" de secuencias de dinucleótidos CpG están metilados). Las secuencias 'submetiladas' de la SEQ ID NO: 1 a la SEQ ID NO: 2 corresponden a la SEQ ID NO: 7 a la SEQ ID NO: 10.

De manera significativa, hasta ahora, las secuencias de ácidos nucleicos y moléculas de acuerdo con la SEQ ID NO: 3 a la SEQ ID NO: 1 no estaban implicadas en o relacionadas con la detección, clasificación o el tratamiento de trastornos proliferativos celulares.

En una realización preferida alternativa, la divulgación proporciona además oligonucleótidos u oligómeros adecuados para uso en los métodos de la invención para detectar el estado de metilación de la citosina en el ADN genómico o tratado (modificado químicamente), de acuerdo con la SEQ ID NO: 1 a la SEQ ID NO: 10. Dicho oligonucleótido o ácidos nucleicos oligoméricos proporcionan medios de diagnóstico novedosos. Dicho oligonucleótido u oligómero que comprende una secuencia de ácido nucleico que tiene una longitud de al menos nueve (9) nucleótidos que es idéntico a, hibrida, bajo condiciones moderadamente rigurosas o rigurosas (como se definió aquí anteriormente), a una secuencia de ácido nucleico tratada de acuerdo con la SEQ ID NO: 3 a la SEQ ID NO: 10 y/o secuencias complementarias de las mismas, o a una secuencia genómica de acuerdo con la SEQ ID NO: 1 a la SEQ ID NO: 2 y/o secuencias complementarias de la misma.

Por lo tanto, la presente divulgación incluye moléculas de ácido nucleico (por ejemplo, oligonucleótidos y moléculas de ácido nucleico peptídicas (PNA) moléculas (oligómeros de PNA)) que hibridan bajo condiciones de hibridación moderadamente rigurosas y/o rigurosas con toda o una porción de las secuencias SEQ ID NO: 1 a la SEQ ID NO: 10 o con los complementos de las mismas. Se prefiere particularmente una molécula de ácido nucleico que hibrida bajo condiciones moderadamente rigurosas y/o rigurosas con toda o una porción de las secuencias SEQ ID NO: 3 a la SEQ ID NO: 10, pero no la SEQ ID NO: 1 a la SEQ ID NO: 2 u otro ADN genómico humano.

La porción idéntica o de hibridación de los ácidos nucleicos de hibridación es típicamente de al menos 9, 16, 20, 25, 30 o 35 nucleótidos de longitud. Sin embargo, las moléculas más largas tienen la utilidad.

Preferiblemente, la porción de hibridación de los ácidos nucleicos de hibridación de la invención es al menos 95%, o al menos 98%, o 100% idéntica a la secuencia, o a una porción de la misma de la SEQ ID NO: 1 a la SEQ ID NO: 10, o a los complementos de las mismas.

La hibridación de ácidos nucleicos del tipo descrito en el presente documento se puede utilizar, 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 de oligonucleótido con una muestra de ácido nucleico se realiza bajo condiciones rigurosas y la sonda es 100% idéntica a la secuencia objetivo. La estabilidad del híbrido o del dúplex de ácido nucleico se expresa como la temperatura de fusión o T_m , que es la temperatura a la que una sonda se disocia de un ADN objetivo. Esta temperatura de fusión se utiliza para definir las condiciones de rigurosidad requeridas.

Para las secuencias objetivo que están relacionadas y son sustancialmente idénticas a la secuencia correspondiente de la SEQ ID NO: 1 a la SEQ ID NO: 2 (tales como variantes alélicas y los SNP), en lugar de idénticas, es útil establecer primero la temperatura más baja a la que solamente se produce la hibridación homóloga con una concentración particular de sal (por ejemplo, SSC o SSPE). Entonces, suponiendo que 1% de falta de coincidencia resulta en una disminución de 1°C en la T_m , la temperatura del lavado final en la reacción de hibridación se reduce en consecuencia (por ejemplo, si se buscan secuencias que tienen > 95% de identidad con la sonda, la temperatura final de lavado se reduce en 5°C). En la práctica, el cambio en la T_m puede ser de entre 0,5°C y 1,5°C por 1% de falta de coincidencia.

Los ejemplos de oligonucleótidos de longitud X (en nucleótidos), según lo indicado por las posiciones del polinucleótido con referencia a, por ejemplo, la SEQ ID NO: 1, incluyen los que corresponden a los conjuntos (conjuntos sentido y antisentido) de oligonucleótidos que se superponen consecutivamente de longitud X, donde los oligonucleótidos dentro de cada conjunto consecutivo de solapamiento (que corresponde a un valor X dado) se definen como el conjunto finito de oligonucleótidos Z de las posiciones de nucleótidos:

- n hasta $(n + (X-1))$;
- donde $n = 1, 2, 3, \dots (Y - (X-1))$;
- donde Y es igual a la longitud (nucleótidos o pares de bases) de la SEQ ID NO: 1 (24.959);
- 5 donde X es igual a la longitud común (en nucleótidos) de cada oligonucleótido en el conjunto (por ejemplo, $X = 20$ para un conjunto de superposiciones consecutivas de 20 mer); y
- donde el número (Z) de oligómeros que se superponen consecutivamente de longitud X para una SEQ ID NO dada de longitud Y es igual a $Y - (X - 1)$. Por ejemplo $Z = 24.959 - 19 = 24.940$, ya sea para los conjuntos sentido o antisentido de la SEQ ID NO: 1, donde $X = 20$.
- 10 Preferiblemente, el conjunto se limita a aquellos oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.
- Los ejemplos de oligonucleótidos de 20 mer incluyen el siguiente conjunto de 24.940 oligómeros (y el conjunto antisentido complementario de los mismos), indicados por las posiciones de polinucleótidos con referencia a las SEQ ID NOs: 1: 1-20, 2-21, 3-22, 4-23, 5-24, y 24.940-24.959.
- 15 Preferiblemente, el conjunto se limita a aquellos oligómeros que comprenden al menos un dinucleótido CpG, TPG o CPA.
- Del mismo modo, los ejemplos de oligonucleótidos de 25 mer incluyen el siguiente conjunto de 24.935 oligómeros (y el conjunto antisentido complementario del mismo), indicados por las posiciones de polinucleótidos con referencia a la SEQ ID NO: 1:
- 1-25, 2-26, 3-27, 4-28, 5-29, y 24.935-24.959.
- 20 Preferiblemente, el conjunto se limita a aquellos oligómeros que comprenden al menos un dinucleótido CpG, TPG o CPA.
- La presente solicitud divulga, para cada uno de las SEQ ID NO: 1 a SEQ ID NO: 10 (sentido y antisentido), múltiples conjuntos consecutivamente superpuestos de oligonucleótidos u oligonucleótidos modificados de longitud X, donde, por ejemplo, $X = 9, 10, 17, 20, 22, 23, 25, 27, 30$ o 35 nucleótidos.
- 25 Los oligonucleótidos u oligómeros de acuerdo con la presente solicitud constituyen herramientas efectivas útiles para comprobar los parámetros genéticos y epigenéticos de la secuencia genómica correspondiente a la SEQ ID NO: 1. Los conjuntos preferidos de tales oligonucleótidos u oligonucleótidos modificados de longitud X son aquellos conjuntos de oligómeros que se superponen consecutivamente correspondientes a las SEQ ID NO: 1 a SEQ ID NO: 10 (y a los complementos de los mismos). Preferiblemente, dichos oligómeros comprenden al menos un dinucleótido CpG, TpG o CpA.
- 30 CpA.
- Los oligonucleótidos u oligómeros particularmente preferidos son aquellos en los que la citosina de las secuencias de dinucleótido CpG (o del correspondiente dinucleótido TpG o CpA convertido) está dentro del tercio medio del oligonucleótido; es decir, donde el oligonucleótido es, por ejemplo, de 13 bases de longitud, el dinucleótido CpG, TpG o CpA se coloca dentro del quinto al noveno nucleótido desde el extremo 5'.
- 35 Los oligonucleótidos también pueden ser modificados mediante el enlazamiento químico del oligonucleótido a una o más fracciones o conjugados para mejorar la actividad, estabilidad o detección del oligonucleótido. Tales fracciones o conjugados incluyen cromóforos, fluoróforos, lípidos tales como colesterol, ácido cólico, tioéter, cadenas alifáticas, fosfolípidos, poliaminas, polietilenglicol (PEG), las fracciones de palmitilo, y otras como se divulga, por ejemplo, en las patentes de Estados Unidos Nos. 5.514.758, 5.565.552, 5.567.810, 5.574.142, 5.585.481, 5.587.371, 5.597.696 y 5.958.773. Las sondas también pueden existir en la forma de un PNA (ácido peptidonucleico) que tenga propiedades de apareamiento particularmente preferidas. Por lo tanto, el oligonucleótido puede incluir otros grupos adjuntos tales como péptidos, y puede incluir agentes de escisión activados por hibridación (Krol y colaboradores, BioTechniques 6: 958-976, 1988) o agentes de intercalación (Zon, Pharm. Res. 5: 539- 549, 1988). Para este fin, se puede conjugar el oligonucleótido con otra molécula, por ejemplo, un cromóforo, fluoróforo, péptido, agente de entrecruzamiento activado por hibridación, agente de transporte, agente de escisión activado por hibridación, etc.
- 40
- 45 El oligonucleótido puede también comprender al menos una fracción de azúcar y/o base modificada reconocida en la técnica, o puede comprender una cadena principal modificada o enlace entre nucleósidos no naturales.

5 Los oligonucleótidos u oligómeros de acuerdo con realizaciones particulares de la presente descripción se utilizan normalmente 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 la SEQ ID NO: 1 a la SEQ ID NO: 2 y las secuencias complementarias de las mismas, o al correspondiente dinucleótido CpG, TpG o CpA dentro de una secuencia de los ácidos nucleicos tratados de acuerdo con la SEQ ID NO: 3 a la SEQ ID NO: 10 y secuencias complementarias de las mismas. Sin embargo, se anticipa que por factores económicos u otros factores, puede ser preferible analizar una selección limitada de los dinucleótidos CpG dentro de dichas secuencias, y el contenido del conjunto de oligonucleótidos se altera en consecuencia.

10 Por lo tanto, en realizaciones particulares, la presente descripción proporciona un conjunto de al menos dos (2) (oligonucleótidos y/o oligómeros de PNA) útiles para detectar el estado de metilación de la citosina en el ADN genómico tratado (SEQ ID NO: 3 a SEQ ID NO: 10), o en el ADN genómico (SEQ ID NO: 1 a SEQ ID NO: 2 y secuencias complementarias de las mismas). Estas sondas permiten el diagnóstico, clasificación y/o la terapia de los parámetros genéticos y epigenéticos de próstata y/o trastornos de proliferación celular colorrectal. El conjunto de oligómeros también se puede utilizar para la detección de polimorfismos de nucleótido único (SNP) en el ADN genómico tratado (SEQ ID NO: 3 a SEQ ID NO: 10), o en el ADN genómico (SEQ ID NO: 1 a SEQ ID NO: 2 y secuencias complementarias de las mismas).

En realizaciones preferidas, al menos uno, y más preferiblemente todos los miembros de un conjunto de oligonucleótidos está unido a una fase sólida.

20 En realizaciones adicionales, la presente solicitud divulga un conjunto de al menos dos (2) oligonucleótidos que se usan como oligonucleótidos 'cebadores' para amplificar secuencias de ADN de una de las SEQ ID NO: 1 a la SEQ ID NO: 10 y secuencias complementarias de las mismas, o segmentos de las mismas.

25 Se prevé 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). Tal arreglo de diferentes secuencias de oligonucleótidos y/o oligómeros de PNA puede ser caracterizado, por ejemplo, porque está dispuesto sobre la fase sólida en forma de un retículo rectangular o hexagonal. La superficie de fase sólida puede estar compuesta de silicio, vidrio, poliestireno, aluminio, acero, hierro, cobre, níquel, plata, u oro. La nitrocelulosa así como plásticos tales como nailon, que pueden existir en la forma de gránulos o también como matrices de resina, también se pueden utilizar. Una visión general del estado de la técnica en la fabricación de una matriz de oligómeros puede ser obtenida a partir de una edición especial de la revista Nature Genetics (suplemento Nature Genetics, Volumen 21, enero de 1999, y de la bibliografía allí citada). A menudo se utilizan sondas marcadas en forma fluorescente para la exploración de matrices de ADN inmovilizadas. La simple fijación de colorantes Cy3 y Cy5 al extremo 5'-OH de la sonda específica es particularmente adecuada para marcadores de fluorescencia. La detección de la fluorescencia de las sondas hibridadas puede llevarse a cabo, por ejemplo, a través de un microscopio confocal. Los colorantes Cy3 y Cy5, además de muchos otros, están disponibles comercialmente.

35 También se anticipa que los oligonucleótidos, o secuencias particulares de los mismos, pueden constituir la totalidad o parte de una "matriz virtual" en la que los oligonucleótidos, o secuencias particulares de los mismos, se utilizan, por ejemplo, como 'especificadores' como parte de, o en combinación con una población diversa de sondas marcadas únicas para analizar una mezcla compleja de analitos. Tal método, se describe por ejemplo en el documento US 2003/0013091 (número de serie de Estados Unidos 09/898.743, publicada el 16 de enero de 2003). En tales métodos, se generan suficientes marcadores de manera que cada ácido nucleico en la mezcla compleja (es decir, cada analito) puede unirse de forma única a una etiqueta única y por lo tanto detectado (cada etiqueta se cuenta directamente, dando como resultado una lectura digital de cada especie molecular en la mezcla).

45 Se prefieren particularmente aquellos conjuntos de oligómeros de acuerdo con los Ejemplos. En la realización más preferida del método, se determina la presencia o ausencia de un trastorno proliferativo celular, más preferiblemente cáncer o la diferenciación de los mismos de los trastornos proliferativos celulares benignos. Esto se logra mediante el análisis del estado de metilación de al menos una secuencia objetivo que comprende al menos una posición de CpG comprendiendo dicha secuencia, o la hibridación bajo condiciones rigurosas con al menos 16 nucleótidos contiguos de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2 y complementos de las mismas. La presente invención proporciona además un método para determinar parámetros genéticos y/o epigenéticos de la secuencia genómica de acuerdo con la SEQ ID NO: 1 a la SEQ ID NO: 2 dentro de un sujeto mediante el análisis de la metilación de citosina y polimorfismos de nucleótido único. Dicho método comprende poner en contacto un ácido nucleico que comprende la SEQ ID NO: 1 a la SEQ ID NO: 2 en una muestra biológica obtenida de dicho sujeto con al menos un reactivo o una serie de reactivos, en donde dicho reactivo o serie de reactivos, distingue entre dinucleótidos CpG metilados y no metilados dentro del ácido nucleico objetivo.

55 En una realización preferida, dicho método comprende las siguientes etapas: En la primera etapa, se obtiene una muestra del tejido a analizar. La fuente puede ser cualquier fuente adecuada, tal como líneas celulares, cortes histológicos, biopsias, tejido incluido en parafina, fluidos corporales, heces, efluente colónico, 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 las heces o la orina.

Se aísla luego el ADN genómico de la muestra. El ADN genómico puede ser aislado por cualquier medio estándar en la técnica, incluyendo el uso de kits disponibles en el mercado. En resumen, en donde el ADN de interés se encapsula en una membrana celular, la muestra biológica debe ser cortada y lisada por un medio enzimático, químico o mecánico. La solución de ADN puede entonces ser liberada de las proteínas y otros contaminantes, por ejemplo, por digestión con proteínasa K. Después se recupera el ADN genómico de la solución. Esto puede llevarse a cabo por medio de una variedad de métodos que incluyen precipitación salina, extracción orgánica o unión de ADN a un soporte de fase sólida. La elección del método se verá afectada por varios factores incluyendo el tiempo, el coste y cantidad necesaria de ADN.

En donde el ADN de la 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 del ADN. Tales métodos incluyen el uso de un reactivo de degeneración de proteínas, por ejemplo, sal caotrópica, por ejemplo, clorhidrato de guanidina o urea; o un detergente, por ejemplo, dodecil sulfato de sodio (SDS), bromuro de cianógeno. Los métodos alternativos incluyen, pero no se limitan a precipitación con etanol o precipitación al vacío, entre otros, por medio de una centrifuga. El experto en la técnica también puede hacer uso de dispositivos tales como dispositivos de filtración, por ejemplo ultrafiltración, superficies o membranas de sílice, partículas magnéticas, partículas de poliestireno, superficies de poliestireno, superficies cargadas positivamente, y membranas cargadas positivamente, membranas cargadas, superficies cargadas, membranas cargadas de conmutación, superficies cargadas conmutadas.

Una vez se han extraído los ácidos nucleicos, se usa el ADN genómico bicatenario en el análisis.

En el segundo paso del método, se trata la muestra de ADN genómico de una manera tal que las bases de citosina que están sin metilar 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 entiende como 'pretratamiento' o 'tratamiento' en el presente documento.

Esto se logra preferiblemente por medio de tratamiento con un reactivo de bisulfito. El término "reactivo de bisulfito" se refiere a un reactivo que comprende bisulfito, disulfito, sulfito de hidrógeno o combinaciones de los mismos, útiles como se describe aquí para distinguir entre las secuencias de dinucleótidos CpG metiladas y no metiladas. Los métodos de dicho tratamiento son conocidos en la técnica (por ejemplo, el documento PCT/EP2004/011715). Se prefiere que el tratamiento con bisulfito se lleve a cabo en presencia de disolventes desnaturizantes tales como, pero no limitado a n-alquilenglicol, en particular dimetil éter de dietilenglicol (DME), o en presencia de dioxano o derivados de dioxano. En una realización preferida, se usan disolventes desnaturizantes en concentraciones entre 1% y 35% (v/v). También se prefiere que la reacción con bisulfito se lleve a cabo en presencia de depuradores tales como, pero no limitado a derivados de cromano, por ejemplo, ácido 6-hidroxi-2,5,7,8-tetrametilcromano-2-carboxílico o ácido trihidroxibenzoico y derivados de los mismos, por ejemplo, ácido gálico (véase: PCT/EP2004/011715). La conversión de bisulfito se lleva a cabo preferiblemente a una temperatura de reacción entre 30°C y 70°C, con lo cual la temperatura se aumenta a más de 85°C durante períodos cortos de tiempo durante la reacción (véase: PCT/EP2004/011715). El ADN tratado con bisulfito se purifica preferiblemente antes de la cuantificación. Esto puede llevarse a cabo por cualquier medio conocido en la técnica, tal como, pero no limitado a, ultrafiltración, llevada a cabo preferiblemente por medio de columnas Microcon® (fabricadas por Millipore®). La purificación se lleva a cabo según el protocolo modificado del fabricante (véase: PCT/EP2004/011715).

En la tercera etapa del método, se amplifican los fragmentos de ADN tratado, usando conjuntos de oligonucleótidos cebadores de acuerdo con la presente invención, y una enzima de amplificación. La amplificación de varios segmentos de ADN puede llevarse a cabo simultáneamente en uno y el mismo recipiente de reacción. Típicamente, la amplificación se lleva a cabo utilizando una reacción en cadena de la polimerasa (PCR). Preferiblemente, dicha reacción amplifica 100 a 2.000 pares de bases de longitud. El conjunto de oligonucleótidos cebadores incluye al menos dos oligonucleótidos cuyas secuencias son cada una complementaria inversa, idénticas, o se hibridan en condiciones rigurosas o altamente rigurosas con al menos un segmento 16 pares de bases de longitud de las secuencias de bases de una de las SEQ ID NO: 3 a la SEQ ID NO: 10 y secuencias complementarias de las mismas.

En una realización alternativa del método, el estado de metilación de las posiciones preseleccionadas de CpG dentro de las secuencias de ácido nucleico de acuerdo con la SEQ ID NO: 1, y más preferiblemente la SEQ ID NO: 2 se puede detectar mediante el uso de oligonucleótidos cebadores específicos de la metilación. Esta técnica (MSP) ha sido descrita en la patente de Estados Unidos No. 6.265.171 de Herman. El uso de cebadores específicos del estado de metilación para la amplificación de ADN tratado con bisulfito permite la diferenciación entre ácidos nucleicos metilados y no metilados. Los pares de cebadores de MSP contienen al menos un cebador que se hibrida con un dinucleótido CpG tratado con bisulfito. Por lo tanto, la secuencia de dichos cebadores comprende al menos un dinucleótido CpG. Los cebadores de MSP específicos para el ADN no metilado contienen una "T" en la posición de la posición de C en el CpG. Preferiblemente, por lo tanto, se requiere que la secuencia de bases de dichos cebadores comprenda una secuencia que tiene una longitud de al menos 9 nucleótidos que se hibrida con una secuencia tratada de ácido nucleico de

acuerdo con una de las SEQ ID NO: 3 a la SEQ ID NO: 10 y secuencias complementarias de las mismas, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG. Otra forma de realización preferida del método comprende el uso de oligonucleótidos bloqueadores (el ensayo HeavyMethyl^{MR}). El uso de tales oligonucleótidos bloqueadores ha sido descrito por Yu y colaboradores, BioTechniques 23: 714-720, 1997. Los oligonucleótidos de la sonda de bloqueo se hibridan con el ácido nucleico tratado con bisulfito simultáneamente con los cebadores de PCR. La amplificación por PCR del ácido nucleico se termina en la posición 5' de la sonda de bloqueo, de tal forma que la amplificación de un ácido nucleico tal se suprime donde está presente la secuencia complementaria para la sonda de bloqueo. Las sondas pueden diseñarse para hibridar con el ácido nucleico tratado con bisulfito de una manera específica del estado de metilación. Por ejemplo, para la detección de ácidos nucleicos metilados dentro de una población de ácidos nucleicos no metilados, la supresión de la amplificación de ácidos nucleicos que no están metilados en la posición en cuestión se lleva a cabo mediante el uso de sondas de bloqueo que comprenden un 'CpA' o 'TpA' en la posición en cuestión, a diferencia de un 'CpG' si se desea la supresión de la amplificación de ácidos nucleicos metilados.

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

Además, la descomposición mediada por polimerasa de los oligonucleótidos bloqueadores debe evitarse. Preferiblemente, dicha exclusión comprende o bien el uso de una polimerasa que carece de actividad exonucleasa 5'-3', o el uso de oligonucleótidos bloqueadores modificados que tienen, por ejemplo, puentes tioato en los terminales 5' de los mismos que vuelven a la molécula bloqueadora resistente a nucleasa. Aplicaciones particulares pueden no requerir de tales modificaciones en 5' del bloqueador. Por ejemplo, si se superponen los sitios de unión del bloqueador y el cebador, impidiendo de ese modo la unión del cebador (por ejemplo, con exceso de bloqueador), se impide sustancialmente la degradación del oligonucleótido bloqueador. Esto se debe a que la polimerasa no extenderá el cebador hacia, y a través (en la dirección 5'-3') del bloqueador, un proceso que da lugar normalmente a la degradación del oligonucleótido bloqueador hibridado.

Una realización de bloqueador/PCR particularmente preferida, para los propósitos de la presente invención y tal como se aplica en el presente documento, comprende el uso de oligómeros de ácido peptidonucleico (PNA) como oligonucleótidos de bloqueo. Tales oligómeros bloqueadores de PNA son idealmente adecuados, ya que no son descompuestos ni tampoco extendidos por la polimerasa.

Preferiblemente, por lo tanto, se requiere que la secuencia de bases de dichos oligonucleótidos de bloqueo comprenda una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratada de acuerdo con una de las SEQ ID NO: 3 a la SEQ ID NO: 10 y las secuencias complementarias de las mismas, en donde 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 directa o indirectamente detectable. Se prefieren las etiquetas en la forma de marcadores de fluorescencia, radionucleidos o fragmentos de moléculas separables que tienen una masa típica que puede ser detectada en un espectrómetro de masas. Cuando dichas etiquetas son marcadores de masa, se prefiere que los amplificados marcados tengan una sola carga neta positiva o negativa, permitiendo una mejor posibilidad de detección en el espectrómetro de masas. La detección puede llevarse a cabo y visualizarse por medio de, por ejemplo, espectrometría de masas de desorción/ionización láser asistida por matriz (MALDI) o usando espectrometría de masas por electroaspersión (ESI).

La espectrometría de masas de desorción/ionización láser asistida por matriz (MALDI-TOF) es un desarrollo muy eficiente para el análisis de biomoléculas (Karas y Hillenkamp, Anal Chem., 60: 2299-301, 1988). Se embebe un analito en una matriz absorbente de luz. La matriz es evaporada por un pulso corto de láser que transporta por lo tanto la molécula de analito en la fase de vapor de una manera no fragmentada. El analito se ioniza por colisiones con moléculas de la matriz. Un voltaje aplicado acelera los iones en un tubo de vuelo libre de campo. Debido a sus diferentes masas, los iones son acelerados a diferentes velocidades. Los iones más pequeños alcanzan el detector antes que los más grandes. La espectrometría MALDI-TOF es muy adecuada para el análisis de péptidos y proteínas. El análisis de ácidos nucleicos es algo más difícil (Gut y Beck, Current Innovations and Future Trends, 1: 147-57, 1995). La sensibilidad con respecto al análisis de ácido nucleico es aproximadamente 100 veces menor que para los péptidos, y disminuye desproporcionadamente al aumentar el tamaño del fragmento. Por otra parte, para los ácidos nucleicos que tienen una cadena principal con múltiples cargas negativas, el proceso de ionización a través de la matriz es considerablemente menos eficiente. En espectrometría MALDI-TOF, la selección de la matriz desempeña un papel eminentemente importante. Para la desorción de péptidos, se han encontrado varias matrices muy eficientes que producen una cristalización muy fina. Existen ahora varias matrices sensibles para el ADN, sin embargo, la diferencia de sensibilidad entre los péptidos y los ácidos nucleicos no se ha reducido. Esta diferencia de sensibilidad puede reducirse, sin embargo, modificando químicamente el ADN de tal manera que sea más similar a un péptido. Por ejemplo, los ácidos nucleicos de fosforotioato, en los cuales los fosfatos habituales de la cadena principal están sustituidos con

5 tiosfosfatos, pueden convertirse en un ADN de carga neutra usando química de alquilación sencilla (Gut y Beck, Nucleic Acids Res. 23: 1367-1373, 1995). El acoplamiento de una etiqueta de carga a este ADN modificado resulta en un aumento en la sensibilidad de MALDI-TOF al mismo nivel que el encontrado para los péptidos. Una ventaja adicional de la marcación de carga es el aumento de estabilidad del análisis contra las impurezas, que hace que la detección de sustratos no modificados sea considerablemente más difícil.

10 En la cuarta etapa del método, los amplificadores obtenidos durante la tercera etapa del método se analizan a fin de determinar el estado de metilación de los dinucleótidos CpG antes del tratamiento. En realizaciones en las que se obtienen los amplificadores por medio de la amplificación MSP, la presencia o ausencia de un amplificador es por 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 de metilación pueden analizarse adicionalmente por medio de métodos basados en las bases tales como, pero no limitado a, tecnología de arreglos y tecnologías basadas en sondas, así como por medio de técnicas tales como secuenciación y extensión dirigida a la plantilla.

15 En una realización del método, los amplificadores sintetizados en el paso tres se hibridan posteriormente con una matriz o un conjunto de oligonucleótidos y/o sondas de PNA. En este contexto, la hibridación se lleva a cabo de la siguiente manera: el conjunto de sondas usadas durante la hibridación está compuesto preferiblemente de al menos 2 oligonucleótidos u oligómeros de PNA; en el proceso, los amplificadores sirven como sondas que se hibridan con oligonucleótidos previamente unidos a una fase sólida; los fragmentos no hibridados se eliminan posteriormente; dichos oligonucleótidos contienen al menos una secuencia 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 de hibridación es típicamente de al menos 9, 15, 20, 25, 30 o 35 nucleótidos de longitud. Sin embargo, moléculas más largas tienen utilidad.

25 En una realización preferida, dicho dinucleótido está presente en el tercio central del oligómero. Por ejemplo, en donde el oligómero comprende un dinucleótido CpG, dicho dinucleótido es preferiblemente el quinto a noveno nucleótido desde el extremo 5' de un 13 mer. Existe un oligonucleótido para el análisis de cada dinucleótido CpG dentro de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2, y las posiciones equivalentes dentro de la SEQ ID NO: 3 a la SEQ ID NO: 10. Dicho oligonucleótidos también pueden estar presentes en la forma de ácidos peptidonucleicos. Se remueven luego los amplificadores no hibridados. Luego 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 encuentra una secuencia de oligonucleótidos.

35 En otra forma de realización adicional del método, el estado de metilación genómica de las posiciones de CpG puede ser comprobado por medio de sondas de oligonucleótidos (como se detalló anteriormente) que se hibridan con el ADN tratado con bisulfito simultáneamente con los cebadores de amplificación de PCR (en donde dichos cebadores pueden ser o bien de metilación específica o estándar).

40 Una realización particularmente preferida de este método es el uso PCR cuantitativa en tiempo real basada en fluorescencia; (Heid y colaboradores, Genome Res. 6 986-994, 1996; véase también la patente de Estados Unidos No. 6.331.393) que emplea una sonda de oligonucleótido fluorescente marcada en forma dual (PCR con TaqMan^{MR}, usando un sistema de detección de secuencias ABI Prism 7700, Perkin Elmer Applied Biosystems, Foster City, California). La reacción PCR con TaqMan^{MR} emplea un oligonucleótido de interrogación no extensible, llamado sonda TaqMan^{MR}, que, en realizaciones preferidas, está diseñado para hibridar con una secuencia rica en CpG localizada entre los cebadores de amplificación directa e inversa. La sonda TaqMan^{MR} comprende además una "fracción informadora" fluorescente y una "fracción inactivadora" unida covalentemente a fracciones enlazadoras (por ejemplo, fosforamiditas) unidas a los nucleótidos del oligonucleótido TaqMan^{MR}. Para el análisis de metilación dentro de los ácidos nucleicos después del tratamiento con bisulfito, se requiere que la sonda sea específica de metilación, tal como se describe en la patente de Estados Unidos No. 6.331.393, también conocido como el ensayo MethyLight^{MR}. Las variaciones en la metodología de detección con TaqMan^{MR}, que también son adecuadas para uso con la invención descrita incluyen el uso de tecnología de doble sonda (LightCycler^{MR}) o cebadores de amplificación fluorescente (tecnología Sunrise^{MR}). Ambas técnicas se pueden adaptar de una manera adecuada para el uso con ADN tratado con bisulfito, y por otra parte para el análisis de metilación dentro de los dinucleótidos CpG.

En una realización preferida adicional del método, la cuarta etapa del método comprende el uso de la extensión de oligonucleótidos dirigida por molde, tal como MS-SNuPE como se describe por Gonzalgo y Jones, Nucleic Acids Res. 25: 2529-2531, 1997.

55 En otra forma de realización adicional del método, la cuarta etapa del procedimiento comprende la secuenciación y posterior análisis de la secuencia del amplificador generado en la tercera etapa del método (Sanger F., y colaboradores,

Proc Natl Acad Sci USA. 74 : 5463-5467, 1977).

En la realización más preferida del método, se aíslan los ácidos nucleicos genómicos y se los trata de acuerdo con los primeros tres pasos del método descrito anteriormente, a saber:

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

5 b) extraer o bien aislar el ADN genómico;

c) tratar el ADN genómico de b), o un fragmento del mismo, con uno o más reactivos para convertir las bases de citosina que están sin metilar en la posición 5 del mismo en uracilo o en otra base que sea detectablemente diferente a la citosina en términos de las propiedades de hibridación; y en donde

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

e) la detección de los amplificados se realiza por medio de una sonda de detección en tiempo real, como se describió anteriormente.

15 Preferiblemente, cuando la amplificación posterior de d) se lleva a cabo por medio de cebadores específicos de metilación, como se describió anteriormente, dichos cebadores específicos de metilación comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia tratada de ácido nucleico de acuerdo a una SEQ ID NO: 3 a SEQ ID NO: 10 y secuencias complementarias de las mismas, en donde la secuencia de bases de dichos oligómeros comprenden al menos un dinucleótido CpG.

20 La etapa e) del método, a saber la detección de los amplificados específicos indicativa del estado de metilación de una o más posiciones de CpG de acuerdo con la SEQ ID NO: 1 se lleva a cabo por medio de métodos de detección en tiempo real como se describió anteriormente.

25 Las realizaciones adicionales de la invención proporcionan un método para el análisis del estado de metilación del ADN genómico de acuerdo con la invención (SEQ ID NO: 1 a SEQ ID NO: 2, y complementos de los mismos), sin la necesidad de la conversión con bisulfito. Se conocen métodos en la técnica en donde un reactivo de metilación sensible a una enzima de restricción, o una serie de reactivos de enzimas de restricción que comprende reactivos de la enzima de restricción sensible a la metilación que distingue entre dinucleótidos CpG metilados y no metilados dentro de una región objetivo se utilizan en la determinación de la metilación, por ejemplo, pero no limitado a DMH.

30 En el primer paso de tales realizaciones adicionales, la muestra de ADN genómico se aísla a partir de tejido o de fuentes celulares. El ADN genómico puede ser aislado por cualquier medio estándar en la técnica, incluyendo el uso de kits disponibles en el mercado. En resumen, cuando el ADN de interés está encapsulado mediante una membrana celular, se debe romper y lisar la muestra biológica por medios enzimáticos, químicos o mecánicos. La solución de ADN puede entonces ser liberada de las proteínas y otros contaminantes, por ejemplo, por digestión con proteinasa K. El ADN genómico se recupera entonces de la solución. Esto puede llevarse a cabo por medio de una variedad de métodos que incluyen precipitación salina, extracción orgánica o unión del ADN a un soporte de fase sólida. La elección del método se verá afectada por varios factores incluyendo tiempo, costo y cantidad requerida de ADN. Todos los tipos de muestras clínicas que comprenden la materia neoplásica o potencialmente neoplásica son adecuados para uso en el presente método, se prefieren las líneas celulares, cortes histológicos, biopsias, tejido incluido en parafina, fluidos corporales, heces, efluente colónico, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre y combinaciones de los mismos. Los fluidos corporales son la fuente preferida de ADN; se prefieren particularmente plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas y células aisladas de la sangre. Una vez se han extraído los ácidos nucleicos, se usa el ADN bicatenario genómico en el análisis.

45 En una realización preferida, el ADN puede escindirse antes del tratamiento con enzimas de restricción sensibles a la metilación. Tales métodos son conocidos en la técnica y pueden incluir tanto medios físicos como enzimáticos. Se prefiere particularmente 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 las islas de 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 de MseI, BfaI, Csp6I, Tru1I, Tvu1I, Tru9I, Tvu9I, MaeI y XspI. Se prefiere particularmente el uso de dos o tres de tales enzimas. Se prefiere particularmente el uso de una combinación de MseI, BfaI y Csp6I.

50 El ADN fragmentado puede ser ligado entonces oligonucleótidos adaptadores con el fin de facilitar la amplificación enzimática subsiguiente. La unión de oligonucleótidos a fragmentos de ADN de extremos adhesivos y romos se conoce en la técnica, y se lleva a cabo por medio de desfosforilación de los extremos (por ejemplo, utilizando fosfatasa alcalina

de ternera o de camarón) y la posterior ligación utilizando enzimas ligasa (por ejemplo T4 ADN ligasa) en la presencia de las dATP. Los oligonucleótidos adaptadores son típicamente de al menos 18 pares de bases de longitud.

5 En el tercer paso, se digiere luego el ADN (o fragmentos del mismo) 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 informativo del estado de metilación de un dinucleótido CpG específico del gen TFAP2E.

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, BSI, BstUI, BshI236I, ACCII, BstFNI, McrBC, Glal, Mvnl, HpalI (HapII), HhaI, Acil, Smal, HinP1I, HpyCH4IV, EagI y mezclas de dos o más de las enzimas anteriores. Se prefiere una mezcla que contiene las enzimas de restricción BstUI, HpalI, HpyCH4IV y HinP1I.

10 En el cuarto paso, que es opcional, pero una realización preferida, se amplifican los fragmentos de restricción. 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 etiquetas fluoróforo, radionucleidos y marcadores de masa. Se prefiere particularmente la amplificación por medio de una enzima de amplificación y al menos dos cebadores que comprenden, en cada caso una secuencia contigua de al menos 16 nucleótidos de longitud que es complementaria a, o se hibrida en condiciones moderadamente rigurosas o rigurosas con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2, y complementos de las mismas. Preferiblemente, dicha secuencia contigua es de al menos 16, 20 o 25 nucleótidos de longitud. En una realización alternativa dichos cebadores pueden ser complementarios a los adaptadores ligados a los fragmentos.

20 En la quinta etapa se detectan los amplificados. La detección puede ser por cualquier medio estándar en la técnica, por ejemplo, pero no limitado a, análisis de electroforesis en gel, análisis de hibridación, incorporación de etiquetas detectables en los productos de PCR, análisis de arreglos de ADN, análisis MALDI o ESI. Preferentemente, dicha detección se lleva a cabo por hibridación con al menos un ácido nucleico o ácido peptidonucleico que comprende en cada caso una secuencia contigua de al menos 16 nucleótidos de longitud que es complementaria a, o se hibrida en condiciones moderadamente rigurosas o rigurosas con una secuencia seleccionada del grupo que consiste en la SEQ ID NO: 1 a la SEQ ID NO: 2, y complementos de las mismas. Preferiblemente, dicha secuencia contigua es de al menos 16, 20 o 25 nucleótidos de longitud.

25 Después de la determinación del estado de metilación o el nivel de los ácidos nucleicos genómicos, se deduce la presencia, ausencia o la clase de trastorno proliferativo celular con base en el estado de metilación o nivel de al menos una secuencia de dinucleótido CpG de la SEQ ID NO: 1, o un promedio, o un valor que refleja un estado medio de metilación de una pluralidad de secuencias de dinucleótidos CpG de la SEQ ID NO: 1 en donde la metilación se asocia con un trastorno proliferativo celular neoplásico o precanceroso. En donde dicha metilación se determina por medios cuantitativos, el punto de corte para determinar dicha la presencia de metilación es preferiblemente cero (es decir, cuando una muestra presenta cualquier grado de metilación se determina que tiene un estado metilado en la posición de CpG analizada). No obstante, se prevé que el experto en la técnica puede desear ajustar dicho valor de corte con el fin de proporcionar un ensayo de una sensibilidad o especificidad particularmente preferida. Por lo tanto, dicho valor de corte se puede aumentar (aumentando así la especificidad), dicho valor de corte puede estar dentro de un rango seleccionado del grupo que consiste en 0%-5%, 5%-10%, 10%-15%, 15%-20%, 20%-30% y 30%-50%. Particularmente preferidos son los puntos de corte 10%, 15%, 25% y 30%.

Ensayos de diagnóstico y pronóstico para trastornos proliferativos celulares

40 La presente invención permite el diagnóstico de eventos que son desfavorables para los pacientes o individuos en los que se pueden utilizar como marcadores importantes parámetros genéticos y/o epigenéticos dentro de TFAP2E. Dichos parámetros obtenidos por medio de la presente invención pueden ser comparados con otro conjunto de parámetros genéticos y/o epigenéticos, las diferencias sirven como base para un diagnóstico y/o pronóstico de eventos que son desfavorables para los pacientes o individuos.

45 Más específicamente, la presente invención permite la detección de las poblaciones en riesgo para la detección temprana de cánceres, lo más preferiblemente cáncer de próstata y/o carcinomas colorrectales. Además, la presente invención permite la diferenciación de trastornos proliferativos celulares neoplásicos (por ejemplo, malignos) de los benignos (es decir, no cancerosos). Por ejemplo, permite la diferenciación de un carcinoma colorrectal de adenomas o pólipos de colon pequeños. Los trastornos proliferativos celulares neoplásicos presentaron menor metilación (es decir, disminución de la expresión) en el gen TFAP2E, en oposición a dichos trastornos benignos que no lo hacen.

55 Específicamente, la presente invención proporciona ensayos de diagnóstico y clasificación de cáncer basados en la medición de la expresión diferencial (preferiblemente metilación) de una o más secuencias de dinucleótidos CpG de la SEQ ID NO: 1, o más preferiblemente subregiones de los mismos de acuerdo con la SEQ ID NO: 2 que comprenden dicha secuencia de dinucleótidos CpG. Típicamente, dichos ensayos implican la obtención de una muestra de un sujeto, la realización de un ensayo para medir la expresión del gen TFAP2E, preferiblemente mediante la determinación del

estado de metilación de al menos una secuencia de dinucleótidos CpG de la SEQ ID NO: 1 (más preferentemente la subregión de la misma de acuerdo con la SEQ ID NO: 2), derivada de la muestra de tejido, en relación con una muestra de control, o un estándar conocido y hacer un diagnóstico basado en ellos.

5 En realizaciones particulares preferidas, los oligómeros de la invención se usan para evaluar el estado de metilación del dinucleótido CpG, tales como los basados en la SEQ ID NO: 1 a la SEQ ID NO: 10, o las matrices de los mismos, así como en kits basados en estos y útiles para el diagnóstico y/o clasificación de trastornos proliferativos celulares.

Kits

10 Por otra parte, un aspecto adicional de la presente divulgación es un kit que comprende: un medio para determinar la metilación de TFAP2E. Los medios para determinar la metilación de TFAP2E comprenden preferiblemente un reactivo que contiene bisulfito; uno o una pluralidad de oligonucleótidos que consisten de secuencias que en cada caso son idénticas, son complementarias, o hibridan en condiciones rigurosas o altamente rigurosas con un segmento de 9 o más preferiblemente 18 bases de longitud de una secuencia seleccionada de la SEQ ID NO: 3 a la SEQ ID NO: 10; y opcionalmente instrucciones para llevar a cabo y evaluar el método descrito del análisis de metilación. En una realización, la secuencia de bases de dichos oligonucleótidos comprende al menos un dinucleótido CpG, CpA, o TpG.

15 En una realización adicional, dicho kit puede comprender además reactivos estándar para la realización de un análisis de metilación específica de la posición de CpG, en el que dicho análisis comprende una o más de las siguientes técnicas: MS-SNuPE, MSP, MethyLight^{MR}, HeavyMethyl, COBRA, y secuenciación de ácidos nucleicos. Sin embargo, un kit a lo largo de las líneas de la presente invención también puede contener sólo parte de los componentes antes mencionados.

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

25 En una realización alternativa adicional, el kit puede contener, empacados en recipientes separados, una polimerasa y un regulador de reacción optimizado para la extensión del cebador mediada por la polimerasa, tal como PCR. En otra realización el kit comprende además medios para obtener una muestra biológica del paciente. Se prefiere un kit, que comprende además un recipiente adecuado para contener los medios para determinación de la metilación del gen TFAP2E en la muestra biológica del paciente, y lo más preferiblemente comprende además instrucciones para el uso y la 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 de cebador que contienen dos oligonucleótidos cuyas secuencias en cada caso son idénticas, son complementarias, o hibridan en condiciones rigurosas o altamente rigurosas con un segmento de 9 o más preferiblemente 18 bases de longitud de una secuencia seleccionada de la SEQ ID NO: 3 a la SEQ ID NO: 10; y opcionalmente (d) instrucciones para el uso y la interpretación de los resultados del kit. En una realización preferida 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 oligonucleótidos y/o un oligómero de PNA que tiene una longitud de al menos 9 o 16 nucleótidos que es idéntica a o se hibrida con una secuencia de ácido nucleico pretratado de acuerdo con una de la SEQ ID NO: 3 a la SEQ ID NO: 10 y secuencias complementarias de las mismas; y opcionalmente (d) instrucciones para el uso y la interpretación de los resultados del kit.

40 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 de cebador que contienen dos oligonucleótidos cuyas secuencias en cada caso son idénticas, son complementarias, o hibridan bajo condiciones rigurosas o altamente rigurosas con un segmento de 9 o más preferiblemente 18 bases de longitud de una secuencia seleccionada de la SEQ ID de NO: 3 a la SEQ ID NO: 10; (d) al menos un oligonucleótido y/o oligómero de PNA que tiene una longitud de al menos 9 o 16 nucleótidos que es idéntica a o se hibrida con una secuencia de ácido nucleico pretratado de acuerdo con una de la SEQ ID NO: 3 a la SEQ ID NO: 10 y secuencias complementarias de las mismas; y opcionalmente (e) instrucciones para uso e interpretación de los resultados del kit.

El kit también puede contener otros componentes tales como reguladores o soluciones adecuadas para bloqueo, lavado o recubrimiento, envasadas en un recipiente separado.

50 Otro aspecto de la divulgación se refiere a un kit para uso en la determinación de la presencia de y/o distinguir entre los trastornos proliferativos celulares, comprendiendo dicho kit: un medio para medir el nivel de transcripción del gen TFAP2E y un medio para determinar metilación de TFAP2E.

Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico a base COBRA^{MR}) para análisis con COBRA^{MR} pueden incluir, pero no se limitan a: cebadores de PCR para TFAP2E; enzima de restricción y regulador

5 apropiado; oligo para hibridación del gen; oligo para hibridación de control; kit de marcación de quinasa para la sonda de oligo; y nucleótidos marcados. Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico a base de MethyLight^{MR}) para análisis con MethyLight^{MR} puede incluir, pero no se limitan a: cebadores PCR para la secuencia convertida con bisulfito del gen TFAP2E; sondas específicas de bisulfito (por ejemplo TaqMan^{MR} o LightCycler^{MR}); reguladores de PCR optimizados y desoxinucleótidos; y Taq polimerasa.

10 Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico a base de Ms-SNuPE^{MR}) para el análisis con MS-SNuPE^{MR} pueden incluir, pero no se limitan a: cebadores de PCR para un gen específico (o secuencia de ADN tratada con bisulfito o isla de CpG); reguladores de PCR optimizados y desoxinucleótidos; kit de extracción en gel; cebadores de control positivo; cebadores de Ms-SNuPE^{MR} para la secuencia convertida con bisulfito del gen TFAP2E; regulador de reacción (para la reacción Ms-SNuPE); y nucleótidos marcados.

Los reactivos típicos (por ejemplo, como puede encontrarse en un kit típico basado en MSP) para el análisis con MSP pueden incluir, pero no se limitan a: cebadores de PCR metilados y no metilados para la secuencia convertida con bisulfito del gen TFAP2E, reguladores de PCR optimizados y desoxinucleótidos, y sondas específicas.

15 Por otra parte, un aspecto adicional de la presente divulgación es un kit alternativo que comprende un medio para determinar la metilación de TFAP2E, en donde dicho medio comprende preferiblemente al menos una enzima de restricción específica de 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 de la SEQ ID NO: 1 a la SEQ ID NO: 2; 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 bajo condiciones rigurosas o altamente rigurosas con al menos un segmento de 18 bases de longitud de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2.

20 En una realización adicional, dicho kit puede comprender uno o una pluralidad de sondas de oligonucleótidos para el análisis de los fragmentos digeridos, preferiblemente dichos oligonucleótidos son idénticos, son complementarios, o se hibridan bajo condiciones rigurosas o altamente rigurosas con al menos un segmento de 16 base longitud de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2.

25 En una realización preferida, el kit puede comprender reactivos adicionales seleccionados del grupo que consiste de: regulador (por ejemplo, enzimas de restricción, PCR, reguladores de almacenamiento o lavado); reactivos o kits de recuperación de ADN (por ejemplo, precipitación, ultrafiltración, columna de afinidad) y componentes de recuperación de ADN. En una realización alternativa adicional, el kit puede contener, empacados en recipientes separados, una polimerasa y un regulador de reacción optimizado para la extensión del cebador mediada por la polimerasa, tal como PCR. En otra realización de la invención, 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 de uno o una pluralidad de ácidos nucleicos o ácidos peptidonucleicos que son idénticos, son complementarios, o se hibridan bajo condiciones rigurosas o altamente rigurosas con al menos un segmento de 9 bases de longitud de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2; y opcionalmente (d) instrucciones para el uso y la interpretación de los resultados del kit. En una realización alternativa 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 cebadores adecuados para la amplificación de una secuencia que comprende al menos un dinucleótido CpG de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2; y opcionalmente (d) instrucciones para el uso y la interpretación de los resultados del kit.

30 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 de la SEQ ID NO: 1 a la SEQ ID NO: 2; (d) al menos un conjunto de oligonucleótidos de uno o una pluralidad de ácidos nucleicos o ácidos peptidonucleicos que son idénticos, son complementarios, o hibridan bajo condiciones rigurosas o altamente rigurosas con un segmento de al menos, 9 bases de longitud de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2 y, opcionalmente, (e) instrucciones para el uso y la interpretación de los resultados del kit.

El kit también puede contener otros componentes tales como reguladores o soluciones adecuadas para bloqueo, lavado o recubrimiento, envasados en un recipiente separado.

55 La divulgación se refiere además a un kit para uso en el suministro de un diagnóstico de la presencia de un trastorno proliferativo celular en un sujeto por medio de análisis con enzimas de restricción sensibles a la metilación. Dicho kit comprende un recipiente y un componente de microarreglos de ADN. Dicho componente de microarreglos de ADN es

5 una superficie sobre la cual una pluralidad de oligonucleótidos se inmoviliza en las posiciones designadas y en donde el oligonucleótido comprende al menos un sitio de metilación CpG. Al menos uno de dichos oligonucleótidos es específico para el gen TFAP2E y comprende una secuencia de al menos 15 pares de bases de longitud, pero no más de 200 pb de una secuencia de acuerdo con una de la SEQ ID NO: 1 a la SEQ ID NO: 2. Preferiblemente dicha secuencia es de al menos 15 pares de bases de longitud, pero no más de 80 pb de una secuencia de acuerdo con una de la SEQ ID NO: 1 a la SEQ ID NO: 2. Se prefiere además que dicha secuencia sea de al menos 20 pares de bases de longitud pero no más de 30 pb de una secuencia de acuerdo con una de la SEQ ID NO: 1 a la SEQ ID NO: 2.

Dicho kit de ensayo preferiblemente comprende además un componente de la enzima de restricción que comprende una o una pluralidad de enzimas de restricción sensibles a la metilación.

10 En una realización adicional, dicho kit de ensayo se caracteriza además porque comprende al menos una enzima de restricción específica de metilación, y en donde los oligonucleótidos comprenden un sitio de restricción de al menos una de dichas enzimas de restricción específicas de metilación.

15 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 de proteína, dicha proteína de unión selectiva a ADN metilado; un componente de ácido nucleico que forma un triplex, uno o una pluralidad de enlazadores, opcionalmente en una solución adecuada; sustancias o soluciones para la realización de una ligación, por ejemplo, ligasas, reguladores; sustancias o soluciones para la realización de una cromatografía en columna; sustancias o soluciones para la realización de un enriquecimiento basado en la inmunología (por ejemplo, inmunoprecipitación); sustancias o soluciones para la realización de una amplificación de ácido nucleico, por ejemplo, PCR; un colorante o varios colorantes, si aplica, con un reactivo de acoplamiento, si aplica, en una solución; sustancias o soluciones para la realización de una hibridación; y/o sustancias o soluciones para la realización de una etapa de lavado.

25 La invención descrita divulga además una composición de materia útil para la detección, diferenciación y distinción entre los trastornos proliferativos celulares del colon. Dicha composición comprende 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: 3 a la SEQ ID NO: 10, y una o más sustancias tomadas del grupo que comprende: cloruro de magnesio 1 - 5 mM, dNTP 100 - 500 μM, 0,5-5 unidades de Taq polimerasa, albúmina de suero bovino, un oligómero, en particular, un oligonucleótido o ácido peptidonucleico (PNA)-oligómero, dicho oligómero que comprende en cada caso al menos una secuencia de base que tiene una longitud de al menos 9 nucleótidos que es complementaria a, o hibrida bajo condiciones moderadamente rigurosas o rigurosas con un ADN genómico pretratado según una de la SEQ ID NO: 3 a la SEQ ID NO: 10 y secuencias complementarias a la misma. 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 con base en polimerasa dentro de dicha solución. Los reguladores adecuados son conocidos en la técnica y están disponibles comercialmente.

35 En otras realizaciones preferidas de la divulgación, al menos dicho ácido nucleico es de 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: 3 a la SEQ ID NO: 10.

40 Aunque la presente invención ha sido descrita con especificidad de acuerdo con algunas de sus realizaciones preferidas, los siguientes ejemplos sirven solamente para ilustrar la invención y no están destinados a limitar la invención dentro de los principios y el alcance de las interpretaciones más amplias y configuraciones equivalentes de la misma.

Ejemplo 1

45 En la siguiente investigación, el rendimiento de los genes seleccionados de acuerdo con la Tabla 2 se analizaron como marcadores de detección de carcinoma colorrectal por medio del ensayo HM (HeavyMethyl). Las regiones objetivo de cada gen fueron convertidas con bisulfito y amplificadas por medio de cebadores que no son MSP, en presencia de oligonucleótidos bloqueadores diseñados para suprimir amplificadas que no habían sido metilados antes del tratamiento con bisulfito. Se detectaron luego los amplificadas por medio de las sondas LightCycler (duales).

Se analizaron las muestras de plasma a partir de las siguientes clases de pacientes:

- carcinoma colorrectal (total 131)

Etapa 0 = 1

50 Etapa I = 13

Etapa II = 32

Etapa III = 27

Etapa IV = 8

Sin clasificación = 50

- colorrectal sana (verificado por colonoscopia) = 169
- 5 - enfermedades no cancerosas (NCD) = 29
- Cánceres de origen no colorrectal (NCC) = 31

En total, se analizaron 360 muestras.

Extracción de ADN y tratamiento con bisulfito

10 Se aisló ADN de todas las muestras por medio del método de Magna Pure (Roche) según las instrucciones del fabricante. A continuación, el eluato resultante de la purificación fue convertido de acuerdo con la siguiente reacción con bisulfito.

15 El eluato se mezcló con 354 µL de solución de bisulfito (5,89 mol/L) y 146 µL de dioxano que contiene un depurador de radicales (ácido 6-hidroxi-2,5,7,8-tetrametilcromano-2-carboxílico, 98,6 mg en 2,5 mL de dioxano). Se desnaturalizó la mezcla de reacción durante 3 min a 99°C y posteriormente se incubó con el siguiente programa de temperatura para un total de 7 h min 50°C; un ThermoSpike (99,9°C) durante 3 min; 1,5 h 50°C; un ThermoSpike (99°C) durante 3 min; 3 h 50°C. Se purificó posteriormente la mezcla de reacción por ultrafiltración usando una columna Millipore Microcon^{MR}. La purificación se llevó a cabo esencialmente de acuerdo con las instrucciones del fabricante. Para este propósito, se mezcló la mezcla de reacción con 300 µL de agua, se cargó sobre una membrana de ultrafiltración, se centrifugó durante 15 min y posteriormente se lavó con regulador 1 x TE. El ADN permanece en la membrana en este tratamiento. 20 Luego, se realizó la desulfonación. Para este propósito, se añadieron 0,2 mol/L de NaOH y se incubó durante 10 min. Luego, se realizó una centrifugación (10 min), seguido por una etapa de lavado con regulador 1 x TE. Después de esto, se eluyó el ADN. Para este propósito, se mezcló la membrana durante 10 minutos con 75 µL de regulador 1 x TE caliente (50°C). Se volteó la membrana de acuerdo con las instrucciones del fabricante. Posteriormente se realizó nuevamente una centrifugación, con lo cual se removió el ADN de la membrana. Se utilizaron 10 µL del eluato para el ensayo PCR en tiempo real LightCycler. 25

Soluciones de reacción y condiciones de ciclos térmicos

Se proporcionan secuencias de componentes para el ensayo de PCR en la Tabla 3. Cada ensayo se realizó dos veces (de forma independiente) en cada muestra.

Las condiciones del ciclo térmico fueron:

| | | | | |
|----|-----------------------------------|--------|-----------|-----------|
| 30 | PCDHGC3 | | | |
| | activación: | 95°C | 10 min | |
| | 50 ciclos: | 95°C | 10 seg | (20°C/s;) |
| | 56°C | 30 seg | (20°C/s) | |
| | 60°C | 3 seg | (20°C/s) | detección |
| 35 | 72°C | 10 seg | (20°C/s) | |
| | curva de fusión: | 95°C | 10 seg | (20°C/s) |
| | 40°C | 10 seg | (20°C/s) | |
| | 95°C | 0 seg | (0,1°C/s) | Continuo |
| | enfriamiento: | 40°C | 5 seg | |
| 40 | Todos los otros ensayos: ensayos: | | | |
| | activación: | 95°C | 10 min | |
| | 55 ciclos: | 95°C | 10 seg | (20°C/s) |
| | 56°C | 30 seg | (20°C/s) | detección |
| | 72°C | 10 seg | (20°C/s) | |
| 45 | curva de fusión: | 95°C | 10 seg | (20°C/s) |
| | 40°C | 10 seg | (20°C/s) | |
| | 95°C | 0 seg | (0,1°C/s) | Continuo |
| | enfriamiento: | 40°C | 5 seg | |
| | Resultados: | | | |

Con el fin de predecir la presencia de ADN tumoral de CRC en las muestras de plasma medidas, se utilizó un modelo de regresión logístico. El modelo de regresión logístico se construye de la siguiente forma. En primer lugar, se codifican los datos de la medición para cada ensayo marcador de forma cualitativa mediante los siguientes 3 niveles:

- Nivel 0 - ninguna de las réplicas de las reacciones de PCR mostraron amplificación
- 5 – Nivel 1 - exactamente una de las dos réplicas de PCR mostró una curva de amplificación
- Nivel 2 - las dos réplicas de PCR mostraron curvas de amplificación

Si cualquiera de las dos repeticiones de PCR no puede ser medida exitosamente, la medición respectiva del marcador se considera como inválida.

10 Se usaron los cinco diferentes marcadores de metilación de ADN como factores independientes con 3 niveles en un modelo de regresión logístico. Se incluyeron en el modelo un factor adicional de interceptación, pero no hay interacciones de los factores. El modelo de regresión logístico fue entrenado y los pesos óptimos para todos los niveles de los factores se determinaron mediante el uso del procedimiento de máxima probabilidad.

15 Las Figuras 1 a 10 proporcionan las gráficas de la media logarítmica medida de metilación de los ensayos individuales. Cada una de las figuras consta de tres gráficas, las gráficas superior e inferior del lado izquierdo proporcionan los análisis binarios y multiclase, respectivamente, la gráfica del lado derecho proporciona una ROC en donde se muestra la sensibilidad en el eje Y y 1-especificidad en el eje X. La Tabla 4 y las figuras 1 a 5 proporcionan una visión general del desempeño del marcador en todos los grupos de muestra. La Tabla 5 y las figuras 6 a 10 proporcionan una visión general del desempeño del marcador en los grupos de carcinoma colorrectal y colorrectal normal.

20 Las Figuras 12 a 21 proporcionan las gráficas de la metilación media logarítmica medida de la mayoría (muestra analizada solamente se cuenta como positiva si ambas repeticiones son positivas, la media de las dos mediciones se toma como la medición cuantitativa de la metilación) de los ensayos individuales. Cada figura consta de tres gráficas, las gráficas superior e inferior del lado izquierdo proporcionan los análisis binarios y multiclase, respectivamente, la gráfica del lado derecho proporciona una ROC en donde se muestra la sensibilidad en el eje Y y 1-especificidad se muestra en el eje X. La Tabla 6 y las figuras 12 a 16 proporcionan una visión general del desempeño del marcador en todos los grupos de muestra. La Tabla 7 y las figuras 18 a 21 proporcionan una visión general del desempeño del marcador en los grupos de carcinoma colorrectal y colorrectal normal

25 La Tabla 8 proporciona una visión general de la AUC y las sensibilidades de los ensayos individuales con 95% de especificidad (todos los valores p eran de menos de 0,00001). En el que dichas clases son:

30 Todas: normal + NCD + NCC vs. CRC etapas I a IV

I-IV: Largo vs CRC etapas I a IV

I-III: Largo vs CRC etapas I a III

35 A partir de la distribución multiclase de la figura 6 (gráfica inferior izquierda) y la tabla 11 se puede determinar que el gen RASSF2 es particularmente efectivo en la detección de la etapa 1 y los carcinomas colorrectales tempranos. Por consiguiente la expresión, lo más preferiblemente la metilación CpG, de dicho gen además de ser un marcador de diagnóstico preferido es particularmente preferido para la detección de poblaciones generales (individuos que no muestran ninguno de los indicadores o síntomas de carcinoma colorrectal) para la detección temprana de carcinomas colorrectales.

Combinaciones de marcadores (paneles)

40 Para identificar el subconjunto de marcadores de metilación de ADN que predice de manera óptima la presencia de CRC, se utilizó el procedimiento de eliminación hacia atrás. En cada etapa de eliminación, se removió el marcador de metilación de ADN con los niveles más bajos del factor del modelo. Se comparó el poder predictivo del modelo reducido con el modelo completo mediante el uso de la prueba de la relación de probabilidad. Para identificar el subconjunto de los marcadores de metilación de ADN que predice de manera óptima la presencia de CRC se utiliza el procedimiento de eliminación hacia atrás. En cada etapa de eliminación, se removió el marcador de metilación de ADN con los niveles más bajos del factor del modelo. Se comparó el poder predictivo del modelo reducido con el modelo completo mediante el uso de la prueba de la relación de probabilidad. La figura 11 muestra que (en la comparación de la muestra normal vs. CRC etapas I a IV) en cada etapa de eliminación se redujo significativamente el poder predictivo del modelo de regresión logístico. Se concluyó que todos los modelos de marcadores de metilación de ADN enumerados dieron lugar

a un desempeño de predicción superior comparado con el marcador solo o los respectivos paneles de marcadores más simples. Se concluyó que los siguientes modelos de marcadores de ADN dieron lugar a un desempeño de predicción superior en comparación con el único marcador o los respectivos paneles de marcadores más simples.

Las figuras 22 a 26 proporcionar una visión general del desempeño de las siguientes combinaciones de marcadores:

5 Septina 9 + TFAP2E + RASSF2 + PCDHGC3 + SND1 (Figura 22)

Todos AUC80

Sens./Espec. 57/95

Todos CRC AUC 80

Sens./Espec. 58/96

10 CRC I-III AUC 76

Sens./Espec. 50/96

Septina 9+TFAP2E+RASSF2+PCDHGC3 (Figura 23)

Todos AUC80

Sens./Espec. 53/95

15 Todos CRC AUC 80

Sens./Espec. 55/96

CRC I-III AUC 77

Sens./Espec. 48/96

20 Septina 9+TFAP2E+RASSF2 (Figura 24)

Todos AUC 77

Sens./Espec. 48/96

Todos CRC AUC 79

Sens./Espec. 52/96

25 CRC I-III AUC 75

Sens./Espec. 42/96

Septina 9+TFAP2E (Figura 25)

Todos AUC 77

Sens./Espec. 45/96

30 Todos CRC AUC 79

Sens./Espec. 51/96

CRC I-III AUC 75

Sens./Espec. 41/96

Septina 9+RASSF2 (Figura 26)

Todos AUC 77

5 Sens./Espec. 43/96

Todos CRC AUC 79

Sens./Espec. 56/95

CRC I-III AUC 74

Sens./Espec. 46/95

10 En cada caso, la gráfica superior muestra todas las muestras (Normales, enfermedades no colorrectal, cánceres no colorrectales y todas las etapas de CRC), la gráfica inferior muestra únicamente muestras normales y de CRC.

Ejemplo 2: Desempeño del marcador en el diagnóstico de cáncer de próstata

15 En la siguiente investigación, se determinó el desempeño de los marcadores seleccionados en la detección de carcinoma de próstata por medio del ensayo HM (HeavyMethyl). Las regiones objetivo de cada gen fueron convertidas con bisulfito y amplificadas por medio de cebadores no MSP, en presencia de un oligonucleótido bloqueador diseñado para suprimir amplificadas que no habían sido metilados antes del tratamiento con bisulfito. Los amplificadas fueron luego detectados por medio de sondas LightCycler (duales) y se determinó el nivel de metilación por referencia a los ensayos de control.

Muestras

20 Para este experimento, se recogieron muestras de orina y de plasma coincidentes de un total de 191 hombres, incluyendo 91 hombres con cáncer de próstata confirmado por biopsia, 51 hombres sin cáncer detectado por biopsia (posteriormente diagnosticados con BPH), y 50 hombres jóvenes sanos. En todos los análisis, la clase positiva se compone de muestras de cáncer de próstata.

25 En el diseño del presente estudio clínico, la principal dificultad estaba en la definición de la clase negativa ya que no hay método de detección que excluya la presencia de cáncer de próstata con 100% de certeza. La biopsia tiene una tasa de diagnóstico de falsos negativos de al menos 10%, mientras que la medición de PSA es propensa tanto falsos negativos como a falsos positivos. Debido a que el objetivo principal del presente estudio fue demostrar la viabilidad de la medición de marcadores metilados de cáncer de próstata en un fluido corporal remoto, nos enfocamos en una clase negativa que reduce al mínimo la probabilidad de falsos positivos. En consecuencia, los hombres jóvenes sanos fueron elegidos como la clase negativa "verdadera". Pensamos que los varones jóvenes sanos sin antecedentes familiares de cáncer de próstata deben ser verdaderos negativos para cáncer de próstata.

30 Con el fin de investigar los marcadores como una continuación de diagnóstico para el PSA, también se incluyó una segunda clase negativa de biopsia negativa, muestras de BPH. Un factor de confusión potencial en esta clase es la probable presencia de biopsias falsos negativos.

35 En cinco casos de cáncer de próstata, únicamente se recolectó una muestra de plasma y en diez casos adicionales sólo se recolectó una muestra de orina. Las muestras se recogieron en múltiples sitios. La orina se recogió después de un masaje prostático. Ambas muestras de plasma y orina se obtuvieron antes de cualquier tratamiento para el cáncer de próstata. Los criterios de inclusión y exclusión fueron diseñados para asegurar que los pacientes analizados reflejaran los pacientes potenciales que usarían las pruebas de detección del cáncer de próstata.

40 Los siguientes criterios de inclusión y exclusión aplican para los pacientes sometidos a biopsia:

Criterios de inclusión:

- Indicación para biopsia (PSA elevado y/o tacto rectal dudoso)

- Biopsia programada dentro de 1 semana después de la recogida de muestras
- edad 40-80

Criterios de exclusión:

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

Los siguientes criterios aplican para los hombres sanos del grupo de control:

Criterios de inclusión:

- Masculino
- 10
- edad 18-30

Criterios de exclusión:

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

15 extracción de ADN

Se extrajo y aisló ADN utilizando protocolos estándar y kits disponibles en el mercado.

Tratamiento con bisulfito

- 20 El eluato se mezcló con 354 μ L de solución de bisulfito (5,89 mol/L) y 146 μ L de dioxano que contiene un depurador de radicales (ácido 6-hidroxi-2,5,7,8-tetrametilcromano-2-carboxílico, 98,6 mg en 2,5 mL de dioxano). Se desnaturalizó la mezcla de reacción durante 3 min a 99°C y posteriormente se incubó con el siguiente programa de temperatura para un total de 7 h min 50°C; un ThermoSpike (99,9°C) durante 3 min; 1,5 h 50°C; un ThermoSpike (99°C) durante 3 min; 3 h 50°C. Se purificó posteriormente la mezcla de reacción por ultrafiltración usando una columna Millipore Microcon^{MR}. La purificación se llevó a cabo esencialmente de acuerdo con las instrucciones del fabricante. Para este propósito, se mezcló la mezcla de reacción con 300 μ L de agua, se cargó sobre una membrana de ultrafiltración, se centrifugó
- 25 durante 15 min y posteriormente se lavó con regulador 1 x TE. El ADN permanece en la membrana en este tratamiento. Luego, se realizó la desulfonación. Para este propósito, se añadieron 0,2 mol/L de NaOH y se incubó durante 10 min. Luego, se realizó una centrifugación (10 min), seguido por una etapa de lavado con regulador 1 x TE. Después de esto, se eluyó el ADN. Para este propósito, se mezcló la membrana durante 10 minutos con 75 μ L de regulador 1 x TE caliente (50°C). Se volteó la membrana de acuerdo con las instrucciones del fabricante. Posteriormente se realizó
- 30 nuevamente una centrifugación, con lo cual se removió el ADN de la membrana.

Ensayo de PCR LightCycler en tiempo real

- 35 Para cada ensayo, se corrió por duplicado un equivalente analito de 1,5 mL, se determinó la metilación por medio de los componentes del ensayo de acuerdo con la Tabla 9. Se utilizaron ensayos de control para el gen de la actina beta y regiones "CFF" para determinar la concentración total de ADN, con el fin de cuantificar la cantidad de ADN metilado como se determinó por medio del ensayo HeavyMethyl.

Resultados

- 40 Uno de los objetivos fue desarrollar marcadores destinados al diagnóstico después de las pruebas de PSA de 4,0 ng/mL o más para hombres mayores de 50 años de edad para discriminar el cáncer de próstata de condiciones no cancerosas. Nos enfocamos en dos indicaciones: una aplicación de detección para identificar hombres mayores de 50 años con un alto riesgo de cáncer de próstata y un diagnóstico después seguido de PSA para informar la decisión de una nueva biopsia próstata en hombres con al menos una biopsia de próstata negativa y un PSA persistentemente elevado. Se

5 analizaron los datos de dos formas diferentes: (i) se utilizaron muestras de cáncer de próstata y biopsia negativa para evaluar el desempeño de los marcadores seguido de una prueba de PSA (aplicación diagnóstica) y (ii), se utilizó las muestras de cáncer de próstata y todas las no cancerosas (biopsias negativas y sanas) para medir el desempeño de los marcadores en la prueba de detección (aplicación de detección). Se reportó el desempeño del marcador para plasma y la orina por separado, también se proporcionó un análisis de datos para los marcadores individuales y paneles de marcadores. Todos los datos se reportan como valores de metilación sin procesamiento de la media logarítmica.

Desempeño del marcador en una aplicación de detección

10 Como prueba de detección primaria, el marcador necesitaría identificar cáncer de próstata en hombres mayores de 50 años de edad con una especificidad mejorada con relación al PSA. Los análisis de aplicación de detección de AU utilizan las muestras de cáncer de próstata como la clase positiva. Para el propósito del presente estudio clínico, se analizaron los datos de nuestra aplicación de detección con dos clases negativas alternativas. La primera clase negativa analizó los 50 hombres jóvenes sanos con mínima probabilidad de cáncer de próstata no detectada. Aunque esta clase negativa representa una "verdadera" prueba negativa, no coincide en edad con la población para la detección de cáncer de próstata y no incluye ninguna clase probable de falsos positivos, por ejemplo, BPH. Por lo tanto, se realizó un segundo análisis en el cual se analizaron 50 controles de hombres jóvenes sanos y todos los 51 controles negativos para la biopsia para un total de 101 de clase negativa.

20 En promedio, se realizan cada año aproximadamente 20.000.000 pruebas de PSA en los EE.UU., con sólo aproximadamente 1.000.000 casos que conducen a una biopsia (de los cuales aproximadamente 750.000 biopsias son innecesarios). Por lo tanto, menos del 5% de los individuos que actualmente son evaluadas por PSA caen en la clase negativa que está representada por un PSA elevado y BPH positivo mientras que la gran mayoría de la población objetivo evaluada cae en la clase negativa con PSA. Mientras que la clase negativa de los varones jóvenes sanos puede representar una sobreestimación de la capacidad discriminadora de nuestros marcadores, la clase negativa combinado de hombre jóvenes sanos, más hombres con biopsia negativa de la misma edad pueden representar una subestimación de la capacidad discriminadora de nuestros marcadores.

25 Desempeño del marcador único en la orina

30 La sensibilidad y especificidad de los marcadores analizados por PCR en tiempo real en orina después de masaje prostático de pacientes con cáncer de próstata, los pacientes biopsia negativa y los individuos de control sanos se muestran en la Tabla 10, y el desempeño del ensayo en orina después de masaje prostático en comparación con la clase I negativa (individuos sanos) se muestra en la Figura 27. El desempeño del ensayo en orina después de masaje prostático comparado con la clase II negativa (individuos sanos más biopsia negativa) se muestra en la Figura 28.

35 La sensibilidad y especificidad de los marcadores analizados por PCR en tiempo real en plasma de pacientes con cáncer de próstata, pacientes biopsia negativa e individuos de control sanos se muestran en la Tabla 11, y el desempeño del ensayo en plasma comparado con la clase I negativa (sana individuos) se muestra en la Figura 29. El desempeño del ensayo en plasma comparado con la clase II negativa (individuos sanos + biopsia negativa) se muestra en la Figura 30.

En todas las comparaciones de clase negativa y para todos los marcadores, la orina fue el analito más sensible como se ilustra en la Tabla 12.

40 Las cantidades crecientes de ADN marcador metilado se correlacionan con puntajes crecientes de Gleason para todos los marcadores en plasma. Esto fue verdad para muestras con altas cantidades de ADN marcador metilado en orina (especialmente los marcadores TFAP2E y RASSF2A), esto fue especialmente cierto en ADN del plasma. El PSA como marcador de cáncer de próstata en pacientes con PSA elevado (> 4 ng/mL) también se correlaciona con puntajes crecientes de Gleason.

45 La Tabla 13 muestra el desempeño de los paneles de marcadores de detección para distinguir el cáncer de próstata de la clase I negativa (varones sanos) en orina. La tabla 14 muestra el desempeño de paneles de marcadores de detección para distinguir el cáncer de próstata de la clase II negativa (hombres sanos más biopsia negativa) en orina. La Tabla 15 muestra el desempeño de los paneles de marcador de detección para distinguir el cáncer de próstata de la clase I negativa (hombres sanos) en el plasma. La tabla 16 muestra el desempeño de los paneles de marcadores de detección para distinguir el cáncer de próstata de la clase II negativa (hombres sanos más biopsia negativa) en plasma.

Desempeño del marcador en aplicación diagnóstico: Seguido por PSA

50 Como aplicación diagnóstica, los marcadores deben identificar cáncer de próstata en hombres mayores de 50 años de edad con PSA persistente elevado (> 4,0 ng/mL) que han sido objeto de al menos una biopsia de próstata negativa. Esta es una aplicación y análisis distinto y requiere una mayor discriminación en comparación con la prueba de detección. Los falsos positivos en esta aplicación se derivan de PSA elevado, biopsias negativo clase HPB. Una vez

más, las muestras de cáncer de próstata representan la clase positiva. Para los propósitos de una aplicación diagnóstica, se analizaron los datos utilizando una clase negativo única compuesta por las 51 muestras de biopsia negativa.

Desempeño de marcador único en orina

5 La Tabla 17 muestra la sensibilidad y especificidad de los marcadores analizados por PCR en tiempo real en orina después de masaje de próstata en pacientes con cáncer de próstata y pacientes con biopsia negativa.

La Figura 31 muestra el desempeño del ensayo único para ensayos por PCR en tiempo real HM en orina después de masaje prostático.

10 Como se muestra en la Tabla 18, para todos los marcadores de metilación analizados la orina fue el analito más sensible. Por lo tanto, se prefiere particularmente para el diagnóstico de cáncer de próstata que el analito sea orina, ya sea evacuada o después de masaje prostático.

La Tabla 19 proporciona el desempeño de los paneles de marcadores de diagnóstico para detectar cáncer de próstata en pacientes con biopsia negativa en orina.

Tabla 1: secuencias de acuerdo con la presente invención

| SEQ ID genómica NO: | Gen | Secuencia metilada convertida con bisulfito (sentido) | Secuencia metilada convertida con bisulfito (antisentido) | Secuencia no metilada convertida con bisulfito (sentido) | Secuencia no metilada convertida con bisulfito (antisentido) |
|---------------------|------------|---|---|--|--|
| 1 | Gen TFAP2E | 3 | 4 | 7 | 8 |

15

Tabla 2: Genes de acuerdo con el Ejemplo 1

| SEQ ID genómico NO: | Gen | Abreviación | Secuencias convertidas con bisulfito |
|--|---|-------------|--------------------------------------|
| | | | |
| | | | |
| | Factor de transcripción AP-2 épsilon (activación de la proteína 2 épsilon que se une al reforzador) | TFAP2E | |
| | | | |
| | | | |
| *A menos que se indique lo contrario todas las ubicaciones establecidas se refieren a la base de datos Ensembl v39 (Junio de 2006) | | | |
| ** Base de datos Ensembl v31.35d (8 Julio de 2005) | | | |

Tabla 3. Secuencias del cebador, bloqueador y sonda de acuerdo con el Ejemplo 1.

| | | | | | |
|--------|----|----|----|----|----|
| TFAP2E | 46 | 51 | 56 | 61 | 66 |
|--------|----|----|----|----|----|

ES 2 597 845 T3

Tabla 4. Desempeño del ensayo HM (Ejemplo 1) en todas las muestras de tejido.

| | RASSF2A | Septina 9 | SND1 | PCDHGC3 | TFAP2E |
|---|-------------------|------------------|------------------|------------------|-------------------|
| AUC (95% intervalo de confianza) | 0,72 (0,67, 0,77) | 0,75 (0,7, 0,79) | 0,66 (0,6, 0,71) | 0,66 (0,6, 0,72) | 0,69 (0,63, 0,74) |
| Sens./Espec. | 0,4/0,95 | 0,47/0,95 | 0,25/0,95 | 0,32/0,95 | 0,29/0,95 |
| Corte de Sens./Espec. | -3,029 | -2,706 | -3,089 | -2,378 | -2,692 |
| Wilcoxon P | 0 | 0 | 0 | 0 | 0 |
| CRC+ Adenoma -(pos) | 131 | 131 | 118 | 119 | 119 |
| Muestra Normal + enfermedades no cancerosas (NCD)+carcinoma diferente al colorrectal (NCC) -(neg) | 228 | 228 | 205 | 206 | 206 |

Tabla 5. Desempeño del ensayo HM (Ejemplo 1) en carcinoma colorrectal y tejido colorrectal normal

| | RASSF2A | Septina 9 | SND1 | PCDHGC3 | TFAP2E |
|----------------------------------|------------------|----------------|------------------|------------------|------------------|
| AUC (95% intervalo de confianza) | 0,73(0,67, 0,78) | 0,76(0,7, 0,8) | 0,67(0,61, 0,73) | 0,68(0,62, 0,73) | 0,71(0,65,0, 76) |
| Sens./Espec. | 0,47/0,95 | 0,48/0,95 | 0,39/0,95 | 0,32/0,95 | 0,39/0,95 |
| Corte de Sens./Espec. | -3,272 | -2,858 | -3,473 | -2,417 | -3,446 |
| Wilcoxon P | 0 | 0 | 0 | 0 | 0 |
| CRC+ Adenoma (pos) | 131 | 131 | 118 | 119 | 119 |
| Normal (neg) | 168 | 169 | 148 | 148 | 148 |

5

Tabla 6. Desempeño del ensayo HM (Ejemplo 1) en todas las muestras de tejido.

| | RASSF2A | Septina 9 | SND1 | PCDHGC3 | TFAP2E |
|---|------------------|------------------|------------------|----------------|----------------|
| AUC (95% intervalo de confianza) | 0,67(0,61, 0,72) | 0,74(0,69, 0,79) | 0,63(0,57, 0,68) | 0,65(0,6, 0,7) | 0,65(0,6, 0,7) |
| Sens./Espec. | 0,37(0,96) | 0,51 | 0,28/0,95 | 0,34/0,95 | 0,34/0,96 |
| Corte de Sens./Espec. | -4 | -4 | -3,45 | -2,523 | -4 |
| Wilcoxon P | 0 | 0 | 0 | 0 | 0 |
| CRC+ Adenoma -(pos) | 121 | 127 | 113 | 127 | 120 |
| Muestra Normal + enfermedades no cancerosas (NCD)+carcinoma diferente al colorrectal (NCC) -(neg) | 206 | 220 | 194 | 224 | 203 |

ES 2 597 845 T3

Tabla 7. Desempeño del ensayo HM (Ejemplo 1) en carcinoma colorrectal y muestras de tejido colorrectal normal.

| | | | | | |
|----------------------------------|------------------|------------------|-----------------|-----------------|-----------------|
| | RASSF2A | Septina 9 | SND1 | PCDHGC3 | TFAP2E |
| AUC (95% intervalo de confianza) | 0,67(0,61, 0,73) | 0,74(0,69, 0,79) | 0,64(0,58, 0,7) | 0,66(0,6, 0,71) | 0,66(0,6, 0,72) |
| Sens./Espec. | 0,37/0,97 | 0,51/0,97 | 0,3/0,97 | 0,35/0,95 | 0,34/0,98 |
| Corte de Sens./Espec. | -4 | -4 | -4 | -2,599 | -4 |
| Wilcoxon P | 0 | 0 | 0 | 0 | 0 |
| CRC+ Adenoma (pos) | 121 | 121 | 113 | 127 | 120 |
| Normal (neg) | 154 | 164 | 146 | 167 | 154 |

5

Tabla 8. AUC y sensibilidad (con especificidad del 95%) para ensayos individuales de marcadores de acuerdo con la clase.*

| | AUC | | | Sensibilidad | | |
|---------------------------------|-----|------|------|--------------|------|------|
| | All | I-IV | I-II | All | I-IV | I-II |
| Septina 9 (media de la mayoría) | 73 | 73 | 67 | 49 | 49 | 37 |
| RASSF2A (Media Log) | 72 | 73 | 70 | 45 | 48 | 41 |
| TFAP2E (Media Log) | 68 | 71 | 67 | 32 | 38 | 30 |
| SND1 (Media Log) | 64 | 65 | 62 | 25 | 35 | 29 |
| PCDHGC3 (Media Log) | 65 | 66 | 64 | 30 | 32 | 29 |

*todos los valores p fueron menores a 0,00001

Tabla 9. Ensayos de acuerdo con el Ejemplo 1.

| | Cebador hacia adelante SEQ ID NO: | Cebador inverso SEQ ID NO: | Bloqueador SEQ ID NO: | Sonda SEQ ID NO: | Sonda SEQ ID NO: |
|----------|--------------------------------------|-------------------------------|--------------------------|---------------------|---------------------|
| Actina B | 77 | 78 | | 79 | |
| TFAP2E | 98 | 99 | 100 | 101 | 102 |

Tabla 10

| Marcador | Clase I negativa: Saludable | | | Clase II negativa: Saludable + Biopsia (-) | | |
|----------|-----------------------------|----------------|------------------|--|----------------|------------------|
| | AUC | Sens. / Espec. | Valor p Wilcoxon | AUC | Sens. / Espec. | Valor p Wilcoxon |
| GSTP1 | 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 |
| HIST1H4J | 0,91 | 0,69 / 0,96 | 0 | 0,78 | 0,36 / 0,96 | 0 |
| TFAP2E | 0,86 | 0,47 / 0,96 | 0 | 0,76 | 0,27 / 0,96 | 0 |

Tabla 11

| Marcador | Clase I negativa: Saludable | | | Clase II negativa: Saludable + Biopsia (-) | | |
|----------|-----------------------------|----------------|------------------|--|----------------|------------------|
| | AUC | Sens. / Espec. | Valor p Wilcoxon | AUC | Sens. / Espec. | Valor p Wilcoxon |
| GSTP1 | 0,61 | 0,17 / 0,96 | 0,0063 | 0,58 | 0,17 / 0,95 | 0,0183 |
| RASSF2A | 0,68 | 0,37 / 1,00 | 0 | 0,64 | 0,20 / 0,95 | 0 |
| HIST1H4J | 0,64 | 0,26 / 0,96 | $5e^{-04}$ | 0,56 | 0,16/0,95 | 0,0572 |
| TFAP2E | 0,61 | 0,22 / 1,00 | $4e^{-04}$ | 0,56 | 0,09 / 0,95 | 0,0128 |

Tabla 12

| Marcador | Clase I negativa: saludable | | Clase II negativa: Saludable + Biopsia (-) | |
|----------|-----------------------------|---------------|--|---------------|
| | AUC de orina | AUC de plasma | AUC de orina | AUC de plasma |
| GESTP1 | 0,89 | 0,61 | 0,79 | 0,58 |
| RASSF2A | 0,90 | 0,68 | 0,79 | 0,64 |
| HIST1H4J | 0,91 | 0,64 | 0,78 | 0,56 |
| TFAP2E | 0,86 | 0,61 | 0,76 | 0,56 |

5

Tabla 13

| Panel del marcador | % Sens. cáncer de próstata | % Espec. saludables |
|---------------------------------|----------------------------|---------------------|
| Marcadores únicos cuantitativos | | |
| RASSF2A | 74 | 96 |
| HIST1H4J | 69 | 96 |
| GSTP1 | 63 | 96 |

| | | |
|--|----|-----|
| TFAP2E | 46 | 100 |
| Paneles cualitativos: | | |
| GSTP1+HIST1H4J | 79 | 98 |
| RASSF2A+HIST1H4J | 94 | 88 |
| Paneles cuantitativos: | | |
| RASSF2A+HIST1H4J | 94 | 88 |
| QuadSVM (todos lo marcadores, sin PSA) | 79 | 98 |

Tabla 14

| Panel del marcador | % Sens. cáncer de próstata | % Espec. saludables + biopsia (-) |
|---------------------------------|----------------------------|-----------------------------------|
| Marcadores únicos cuantitativos | | |
| RASSF2A | 74 | 76 |
| HIST1H4J | 69 | 68 |

(continuación)

| Panel del marcador | % Sens. cáncer de próstata | % Espec. saludables + biopsia (-) |
|--|----------------------------|-----------------------------------|
| GSTP1 | 63 | 80 |
| TFAP2E | 46 | 88 |
| Paneles cualitativos: | | |
| GSTP1+HIST1H4J | 79 | 72 |
| RASSF2A+HIST1H4J | 94 | 54 |
| Paneles cuantitativos: | | |
| RASSF2A+HIST1H4J | 94 | 58 |
| QuadSVM (todos lo marcadores, sin PSA) | 79 | 76 |

5

Tabla 15

| Panel del marcador | % Sens. cáncer de próstata | % Espec. saludables |
|----------------------------------|----------------------------|---------------------|
| Marcadores únicos cuantitativos: | | |
| RASSF2A | 37 | 100 |
| HIST1H4J | 26 | 96 |

ES 2 597 845 T3

| | | |
|---|----|-----|
| GSTP1 | 17 | 94 |
| TFAP2E | 22 | 100 |
| Paneles cualitativos: | | |
| RASSF2A+HIST1H4J | 41 | 98 |
| Paneles cuantitativos: | | |
| RASSF2A+TFAP2E (TFAP2E usado para normalizar) | 32 | 100 |
| QuadSVM (todos los marcadores, sin PSA) | 39 | 96 |

Tabla 16

| Panel del marcador | % Sens. cáncer de próstata | % Espec. saludables + biopsia (-) |
|---|----------------------------|-----------------------------------|
| Marcadores únicos cuantitativos: | | |
| RASSF2A | 37 | 91 |
| HIST1H4J | 26 | 88 |
| GSTP1 | 17 | 95 |
| TFAP2E | 22 | 92 |
| Paneles cualitativos: | | |
| RASSF2A+HIST1H4J | 41 | 88 |
| Paneles cuantitativos: | | |
| RASSF2A+TFAP2E (TFAP2E usado para normalizar) | 32 | 92 |
| QuadSVM (todos los marcadores, sin PSA) | 39 | 94 |

Tabla 17

| Marcador | Cáncer de próstata vs. Biopsia (-) | | |
|----------|------------------------------------|----------------|------------------|
| | AUC | Sens. / Espec. | Valor p Wilcoxon |
| GSTP1 | 0,69 | 0,23 / 0,95 | 6e ⁰⁴ |
| RASSF2A | 0,66 | 0,18 / 0,95 | 0,0043 |
| HIST1H4J | 0,64 | 0,28 / 0,95 | 0,0126 |
| TFAP2E | 0,65 | 0,21 / 0,95 | 0,0062 |
| ***PSA | 0,56 | 0,22 / 0,95 | 0,3098 |

***Pruebas para saber si PSA contiene información adicional más allá de la contribución por la indicación de corte > 4

ng/mL para biopsia de próstata.

Tabla 18

| Marcador | Cáncer de próstata vs. Biopsia (-) | |
|----------|------------------------------------|---------------|
| | AUC de orina | AUC de plasma |
| GSTP1 | 0,69 | 0,55 |
| RASSF2A | 0,66 | 0,60 |
| HIST1H4J | 0,64 | 0,50 |
| TFAP2E | 0,65 | 0,52 |
| ***PSA | na | 0,56 |

***Pruebas para saber si PSA contiene información adicional más allá de la contribución por la indicación de corte > 4 ng/mL para biopsia de próstata.

Tabla 19

| Panel del marcador | % Sens. cáncer de próstata | % Espec. biopsia (-) |
|----------------------------------|----------------------------|----------------------|
| Marcadores únicos cuantitativos: | | |
| RASSF2A | 74 | 55 |

(continuación)

| Panel del marcador | % Sens. cáncer de próstata | % Espec. biopsia (-) |
|---|----------------------------|----------------------|
| HIST1H4J | 69 | 41 |
| GSTP1 | 63 | 64 |
| TFAP2E | 46 | 77 |
| Paneles cualitativos: | | |
| GSTP1+HIST1H4J | 79 | 46 |
| RASSF2A+HIST1H4J | 94 | 21 |
| Paneles cuantitativos: | | |
| RASSF2A+HIST1H4J | 94 | 27 |
| GSTP1+PSA | 83 | 45 |
| QuadSVM (todos los marcadores, sin PSA) | 79 | 55 |

5

<110> EPIGENOMICS AG

<120> MÉTODOS Y ÁCIDOS NUCLEICOS PARA ANÁLISIS DE TRATORNOS PROLIFERATIVOS CELULARES

<130> E30940PCT

<160> 102

5 <210> 1

<211> 24959

<212> ADN

<213> Homo Sapiens

<400> 1

ES 2 597 845 T3

| | | | | | | |
|-------------|-------------|-------------|------------|------------|-------------|------|
| ttgagagcca | ctgatctatg | agtaagagta | aggtagttta | aggtgctgag | acaggggtgcc | 60 |
| atgtgccctc | tgggatacct | gcctcaattc | tccaccacat | accttctgtg | ggtggatggt | 120 |
| taagggttag | aaatggcaac | actccataag | gattagcttg | tgggggttaa | ggaaaggacc | 180 |
| tacgctcttc | tgggatgact | agcaagtttt | gaacttgaaa | aagggtggtg | tcatttatga | 240 |
| aactacagag | caaaggagaa | gcaggttggg | gttggtgctg | gggatagagg | agatgttgag | 300 |
| ttcagtgttg | gacttggaaa | atttcagaca | gctgcaatga | gatgtctgga | gaaaatacgc | 360 |
| agctggcagc | tgtgtctatg | ggtctggatc | tctcactagt | gattcaggcc | agagataaca | 420 |
| cagttgggag | ataacagtag | acaggtagtc | attaagact | tgggccgggc | gtggtgactc | 480 |
| atgcctgtaa | tcccagcact | ttgggaggct | gaggcgagtg | tcacttgaa | gtcaggagtt | 540 |
| cgagatcagc | ctggctgaca | tgatgaaacc | ctgtctctag | taaaaataca | aaagttagcc | 600 |
| gggcgtgggtg | gtgggccctt | gtaatcccag | ttactcagga | aactgagaca | tgagaatcgc | 660 |
| ttgaaccg | gaggtggagg | ttgcagtgag | ccgagattgt | gccactgcac | tccagcccaa | 720 |
| gcaacagagt | gagactctgt | tataaaaaaa | aaaaaaaaaa | agacttgaga | gggatgagc | 780 |
| tcactcatga | agagtgcagg | aagaaaaaga | agcttctctg | ctctgagaaa | gaattaccga | 840 |
| cattaagcat | gtaggtgaaa | agagaaacag | aaaagaagac | agaaatggtc | ggagtaggac | 900 |
| aatgaggcaa | aaaggggcta | tcaaaggtgt | aggaagtaga | ttgagaaaac | ccaaggaagc | 960 |
| cacagacagg | cagcacatcc | cgccactctg | gccttcccaa | gctgctggag | agagcagggg | 1020 |
| catccccgca | gctgggcaca | ggcatcatct | cctttgcggc | tcctagactt | ctcccagagg | 1080 |
| agttcatcaa | tttctattca | cgcaactgaca | cctgtcctgg | agcttgactt | ctgatgccct | 1140 |
| gtggctgggg | ttgagggagg | gaggaacagg | gcattgccac | agactccctc | atatatagat | 1200 |
| acacacaaac | acaatcctga | tctctgacct | atctcgagca | ggaaagttgc | aaagcttccc | 1260 |
| aattccgctt | catcctctgg | aacctcgctg | gcggttccag | acttcagaaa | gttcttaatg | 1320 |
| atgaatcctt | ggcttctagc | ctccttcttt | cccaccact | cccaggctg | agatcatcca | 1380 |
| gccctgagct | ctgcacctgc | catgtgcctt | ctgggatacc | cacttggaac | catatattcc | 1440 |
| atgagttaat | ggtgtcacta | tcacctctc | ttctatatca | ataattttac | cttaatcttc | 1500 |
| tcatttacta | cttcactaaa | aaggaccta | aaaaaacaac | tagtcttctc | tgcttcccc | 1560 |
| tactttttta | ttggtgtttg | cgaagttgca | gaaactgaga | tcacagagga | aacgggatgt | 1620 |
| agctaggatc | acacagttag | gcagagtcag | gactaaatcc | aagcccctga | gtctgaggca | 1680 |
| gtagagaata | gtgattctga | gcagagaggg | gccaggggca | gatcacccca | gctctgccac | 1740 |
| ttagctgctt | gaccttgggc | aagccattta | acctctccat | gcctcacccc | tcccctgaaa | 1800 |
| atggggatga | tatgggtatc | tctttcatag | agctgctgtg | agggttaaag | aacccaatat | 1860 |
| tcgtgaagca | ctttgcacag | aggctggcac | attctgtaat | aagtctgctg | tgattatttc | 1920 |
| tactttgccc | tggctctgcac | atcttgcttt | taaaaaacat | ctttcttctg | ctttcttaat | 1980 |
| tcctgttctc | cttagcattc | ctctcatttg | tccatcttta | atccctgtct | ctgcttgtca | 2040 |
| ggaggggagg | gttgggtttg | gggagaaaga | gagagtgact | acagagactg | gaacttgctt | 2100 |
| gacagccagg | gagtgccctgc | catggttaaca | caggaacatg | tagcctgcgc | atgttagtgt | 2160 |
| aacaatcaca | cggcatccat | gctagtcaat | tacacacagt | cactcaacag | tctctgggta | 2220 |
| tctcatgccc | acactcaatc | agcccaggac | cctgtggctt | cagaaatata | ttggcccctt | 2280 |
| gtcaggtgga | gctcataccc | aaacccatcc | caggcaagtc | tctctctctg | tctggccagc | 2340 |
| ttctgtctcc | aactccctca | gagcctggct | tggcttggaa | gttttcagcc | caggaagcct | 2400 |
| gagagcgaag | tgggcacctt | cgagggattt | tgatccagaa | ccattcagga | aagactgtga | 2460 |
| tggtagacct | cctaaacaca | cgcaactgaca | ctgtgtgcgc | atgagcagat | aggctggagt | 2520 |
| cttcagctca | caaacttctt | ccacgttccc | gattaccac | ctcatccaaa | gtctcttgag | 2580 |
| gtgcttagag | aaaggattga | tacctatctt | accgaggaag | ctgaggccag | cgaagaatag | 2640 |
| tgctttgctt | aaaggcaagg | atgtgataga | gccaggcaaa | tgtagggccc | ttgcgtccca | 2700 |
| gcctcgaata | tttctgtttc | actgcagctg | cgagctgcga | gcccagaagg | cggctctgat | 2760 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| caatgtgggc | tgaattcacg | ccccctaccc | accctaccgg | atccccgcaa | agccactgcc | 2820 |
| tccccctgtgg | cagccgctgg | ggaaaagggg | gagaactgcc | atagcttgtc | tctttaatgc | 2880 |
| gcgcccccg | aggccccggc | cgcccccgcc | actataactg | gagtgcatgg | agcaggctac | 2940 |
| gctcagagga | ggaagggcgg | gcgctgggca | cccgttgacc | gacttttcca | agtgcgatca | 3000 |
| gtgcccgtcc | gtcctgcctc | catggaccgg | cccgggaacg | gccaccgctg | aggacccccac | 3060 |
| gcccactagg | atccccggctg | ggtcgcaccc | agctaccgca | ccgtgacctc | cgcgggctgt | 3120 |
| gcccggctccc | ggcgctctctg | cccgcagcgc | tcgcccgtcg | gctagggctc | cgccgcccgc | 3180 |
| acgcctcgcg | cccggcactc | accgccccat | gctgggtgcac | acctaactccg | ccatggtgag | 3240 |
| tagtctcggg | cccgggacat | attcggccgg | gggatcgggg | cgctgagtg | ctggactttc | 3300 |
| caaccctcct | gtcccgcgt | aacgaaatct | cggaggagg | gggcccagc | gacttcgcca | 3360 |
| gctgagaacg | cgatgcgcaa | gtgaccgctg | tccccagcct | gaggctcctg | cgccccgggg | 3420 |
| tggctcggaa | ataaacctcg | ggccccaaag | aacggggagg | actgcagctg | aaccctcccg | 3480 |
| agctgaggag | gatcctttca | ggctggggtc | ctttcagctg | ccagtgggtc | acctaaggca | 3540 |
| cccctctcct | tccccaggag | cgccccgacg | ggctgggagc | agctgccggc | ggggcccggc | 3600 |
| tgtcgtctct | gccccaggcg | gcctacgggc | cggcgcccc | gctctgccac | acgcccggccg | 3660 |
| ccacagctgc | cgccgaattc | cagccgccc | acttcccggc | gccctaccgg | cagccaccgc | 3720 |
| tgcctacgg | tcaggcgccc | gacgcccggc | cagcctttcc | ccacctggca | ggggacccat | 3780 |
| atggcggcct | ggcgccccctg | gcgcagccgc | agcctcctca | ggccgctgg | gcccgcgccc | 3840 |
| gcgcagccgc | ccgcgcccac | gaggagcctc | ccggcctgct | ggcaccgccc | gcccgcgccc | 3900 |
| tgggccttga | cccgcgccc | gactatgcca | ctgccgtgcc | ccggctcctg | ccggcctcctg | 3960 |
| ccgacggcgc | gcacggcctg | gcagacgcac | ctctcggcct | tccggggctg | gcgggcgccc | 4020 |
| ccggtctgga | ggacctgcag | gtgagaccgg | agggatccgg | gatgggtcgg | gactggccgc | 4080 |
| ggtggtttac | taccatggct | ggaggcagaa | ggtgacaaat | gcaggaagcc | gacttttctc | 4140 |
| ccagctcgcc | acatctcact | ggtgactccg | aggatgtgtc | ccacctcctg | ggttagacgt | 4200 |
| tgcctggcca | tcaaggctgc | ctggcactga | gtccgcccagc | agtgcgtggg | tggcagcaac | 4260 |
| cctcggcagg | gaccgaatca | cttcttttct | ccccgctgcc | acgtgtggca | tccatggggc | 4320 |
| agcctgcacc | tggggcgaag | agaagcataa | tagttaccag | tgtaagattc | aaaattccct | 4380 |
| ttttgactg | cacagtgaga | tgcccagggc | tccagctcag | tgcttgaca | tagcgattcc | 4440 |
| tgggcctgcc | cgtcgccgccc | ccaagcgaag | ctggtgcgcc | ttgggcggag | cagacagaga | 4500 |
| ccctgggtgg | caggggcttg | ggaagacatg | ggcggctagg | gctttatgcg | ccctcaccgc | 4560 |
| tgcctctcgc | tatttgcagg | caatggacga | gcccggaatg | agcctcctag | accagtccgt | 4620 |
| gatcaagaaa | ggtaaggaat | ggtctgtcag | ggcagagccc | ggcgagatgg | tgcaaggcct | 4680 |
| tgggtcacag | atccatthtc | ttcaccggcc | gtgcctcctg | tgtgtcgcca | ggctgggtgt | 4740 |
| ccaccaggca | ctcttcctg | cccagccaga | tgttaggcag | acgtgcgggc | ttggtgagtt | 4800 |
| tgcccagcac | cctgtggcct | ggggtgggccc | tcagcggatc | agcattcact | gggctgcagc | 4860 |
| actgggagcc | tggcctctcc | ccgcccagg | ggagggcact | cttgtggatc | tggagttgat | 4920 |
| ttgcagaacg | agttaaacca | cttccctgtt | tccctaagag | atgggaatgg | aagtgtctgt | 4980 |
| cccacggagt | tggggaaatg | atthttcactt | tacagtgcct | tagcatttcg | gtgcctggcg | 5040 |
| ggcactthct | tcctcttcct | tccaggcagg | gccttgagg | cctctggggg | aatthttcttt | 5100 |
| ctgtgggagt | ctcttgcggc | atthtagactt | aggggagctt | gtgtgtgagt | actgtgtgtt | 5160 |
| agggtctgtg | cacctgagtc | agggcccacc | tgctcctggg | tgtctgtgtc | catgtgagtt | 5220 |
| caggttcctg | tgcatgtctg | aaatgttccc | ttcatgggtg | tcttagtatt | tcttggagtg | 5280 |
| tgagtgtgtc | tgthttctgtg | aatgtgtttg | tgaggtgtgt | ctctgtatgt | tgggtgtgcat | 5340 |
| ttctctgcat | ttgggggatg | tacacatthtc | tcaatatgta | cagtatctct | gthtgtctct | 5400 |
| gcactthgtt | ctthtggatc | tgaggatthtc | caagcatgcg | cgggcccctct | ctgtgtatat | 5460 |
| ataggagtat | ttatgtgact | cctggcatta | gtaaaatcca | gggacacggg | atccaccttt | 5520 |
| tctggcctga | ggaccaagta | ctggccatga | cagggggaag | tgagagacga | caaaaacaga | 5580 |
| gagacagcca | gagaggagca | gagagtca | ggggcccagg | cattgggtag | cagcctcttt | 5640 |
| acatttgggg | caggtgcccc | aaagaattca | gaggtgcaca | tgagcctgag | gtgccccagg | 5700 |
| caggcactgc | tcccacagg | tttggcctga | gthtthtttc | aaacgagtga | attcaagcct | 5760 |
| ggctctatt | tgccctccac | ttgttctcag | ggaggccaa | ggtggaagtg | gtggtagcag | 5820 |
| ggctggggct | ggactccag | gagctggggc | tgagttacca | ggagctgggg | gttgggtgga | 5880 |
| tgacttggag | tgtgtagcag | ggaagatgag | gcaacagggc | aggaagtggg | tggggggagg | 5940 |
| tggaattggg | gctgtgtcct | gtgtcgcttg | gaactgggag | tgtgggaaag | acactaggaa | 6000 |
| cctggttgca | gcgcagctct | gctggtgggg | cttggttggc | ttactgtaca | gagcctthct | 6060 |
| tgaccctga | agaaagagat | ccgtctgcag | tgggcaaaa | cctgcctgga | cttccctggcc | 6120 |
| accagaaata | tgagcatggt | ggtggtcccc | agthccctat | tcatgcttgg | gctcaagaga | 6180 |
| ctgggagtct | aggttcaactg | actccctgag | aaagactaag | accctgcatt | ttagaaagag | 6240 |
| gthtggggat | ctctgcccctg | cgcaagggtg | gaaggatcag | ctgttcctct | gagcacctta | 6300 |
| acccggaacc | ccggtccgaa | gccgagacag | gagactggat | gcgaggccct | cccagagctg | 6360 |
| gthtctctca | aacaacthcc | aaaactccta | gatcctagg | gtacgcccga | atcccccaaa | 6420 |
| gcagtcctaaa | gaacacaacg | agagtcctaa | catcccaggt | ggcggcgcg | tggctccctg | 6480 |
| gagcggggcg | ggacgcggcc | gcgcgggactc | acgtgcacaa | ccgcgcggga | cggggcccacg | 6540 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| cggactcacg | tgacacaaccg | cgggaccccc | gcgccagcgg | gaccccagcg | ccagcgggac | 6600 |
| cccagcgcca | gcgggacccc | agcgcagcgg | ggaccccagc | gccagcggga | ccccagcgcc | 6660 |
| agcgggaccc | cagcgcagc | gggtctgtgg | cccagtgagg | cgagtgaggc | gctggcgacc | 6720 |
| tgagcggaga | ctgcgccttg | gacgccccag | cctagacgtc | aagttacagc | ccgcgcagca | 6780 |
| gcagcaaagg | ggaaggggca | ggagccgggc | acagttggat | ccggaggtcg | tgacccaggg | 6840 |
| gaaagcgtgg | gcggtcgacc | cagggcagct | gcggcggcga | ggcaggtggg | ctccttgctc | 6900 |
| cctggagccg | cccccccc | cacctgccct | cggcgccccc | agcagttttc | accttgcccc | 6960 |
| tccgcggtca | ctgcgggatt | cggcgttgcc | gccagcccag | tggggagtga | attagcggcc | 7020 |
| tccttcgtcc | tcggcccttc | cgacggcacg | aggaactcct | gtcctgcccc | acagaccttc | 7080 |
| ggcctccgcc | gagtgcggtg | ctggagcctg | ccccgcagc | gccctggaat | gcagaaaagt | 7140 |
| cgctctttgg | ccacctgaag | cgtcggatcc | ctacagtgcc | tcccagcctg | ggcgggagcg | 7200 |
| gcggctgctg | cgctgaaggt | tggggtcctt | ggtgcgaaaag | ggaggcagct | gcagcctcag | 7260 |
| ccccaccca | gaagcggcct | tcgcatcgct | gcggtgggcg | ttctcgggct | tcgacttcgc | 7320 |
| cagcgcggcg | gggagaggc | acctggagct | cgcagggccc | agacctgggt | tggaaaagct | 7380 |
| tcgctgactg | caggcaagcg | tccgggaggg | gcggccaggc | gaagccccgg | cgctttacca | 7440 |
| caccttccg | ggtcccagc | cagttgcac | cgcggtattg | ggcaggaaat | ggcagggctg | 7500 |
| aggccgaccc | taggagata | agggagccct | ccatttcctg | cccacatttg | tcacctccag | 7560 |
| ttttgcaacc | tatcccagac | acacagaaaag | caagcaggac | tggtggggag | acggagctta | 7620 |
| acaggaata | tttccagcag | tgagcagggg | ctgtatggga | cgcgggagga | gctcagagga | 7680 |
| ggcgcggaga | gtgcccagg | ttgggtgagt | gcctagaggg | gagatagttg | aaccgggttc | 7740 |
| aagagtgct | tagtgggtgt | ttgttgaatg | aatgagtgat | gggctttgaa | gtctgagtcg | 7800 |
| attgaaagag | ggggtgtgta | aaaagggctc | ctttcatcac | acaggacaca | gcatatgcaa | 7860 |
| atcctctccc | tgtggaaaag | ccagacaggt | taaaaaggtt | acaaacaaat | tagccgggca | 7920 |
| tggtggtgcg | cgtctgtagt | cccagctact | agggaggctg | agccagggga | atcgcttgaa | 7980 |
| cccgggaggc | ggagattgca | gtgagccaag | atcgcgccac | tgactccag | cctggaaaca | 8040 |
| gagcgagact | ccgtctcggg | aaaaaaaaaa | aaaagttaca | aaccgtgtgt | gggtttcagg | 8100 |
| ttatacaatc | agagctggag | gggagtggtc | aaggatgaga | actgagatgg | atccctcgtt | 8160 |
| ccctctggag | gagagtgggt | ggttgcttac | ttgggggtgg | ggaatccctc | tccacgggct | 8220 |
| cagctgtcca | atctcagggg | atctctagga | caggagctga | tgtaaacagt | cgccctattc | 8280 |
| cttctgtct | ttggccctgg | agaaggagga | gggagctggg | gaggtctcc | acttcccaga | 8340 |
| caatctctaa | tgcccaggga | catgggtgag | atgagtgaga | tactgacttc | tgggacagaa | 8400 |
| tttgagaggg | tgcaaaaaaa | ctcagtaatc | aagataaata | ggccgggccc | agtggctcac | 8460 |
| gtctgtaatc | ccagcacttt | gggaggccgg | atcacttgag | gtcaagagtt | cgagaccagc | 8520 |
| ctggccaaga | tggtgaaacc | ccatctctac | taaaaataca | aaaattagcc | cagtgtggtg | 8580 |
| gcgctagcct | gtaatcccag | ccactcagga | ggctgaggca | agagaattgc | ttgaccaggg | 8640 |
| aggcagaggt | tgagtgagc | cgagatcatg | ccactgtact | ccagcctgga | caacagaggg | 8700 |
| agactatctc | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | gaggccgggg | ggcgggtggct | 8760 |
| cacaccatgt | gatcccagca | ctttgggagg | ccgagggcggg | tggatcacct | gaggtctgga | 8820 |
| gttcgagacc | agcctggcca | atatggtgaa | accccgtttc | tactaaaaat | acaaaaatta | 8880 |
| gctgggtggg | gtggcaggca | cctgtaatcc | cagctactcc | ggaggctgag | gcaggagaat | 8940 |
| cccttgaacc | tgccgggccc | aggtgagcgt | gaaccaagat | cacaccattg | cactgagacc | 9000 |
| tggacaacaa | cagcaaaaact | ctgtctcaaa | aaaaaaaaaa | atcttttttt | tcgagacaca | 9060 |
| gttttactct | ctcgcgccag | ttggggtgca | gcaccacgat | ctcagctcac | tgcaacctct | 9120 |
| gcctctcaga | ttctcgtacc | tcagcctccc | aagtagctgg | gattacaggt | acctgtcacc | 9180 |
| acgccagcct | aatttttgta | tttttagtag | ggcgtgggtt | tcaccatggt | ggccaggctg | 9240 |
| gtcctgaaact | cctgacctca | agtgacctgc | ccgcttcagc | cacccaaagt | gctgggatta | 9300 |
| caggcgtgag | ccaccacgct | tggccttttt | aaatgaaaat | agtgcaaaaa | tccacgataa | 9360 |
| acaaaatata | aaaaatttac | tgaacttgca | cttccacaac | cctttctcac | ctgcctccca | 9420 |
| ggctactctc | tgccccagaa | agcaacttaa | aaaatgtgca | gatggagttt | ggactttacc | 9480 |
| tgaaaatggt | gggagctatg | gaaaaccttg | gagcagggga | gtgaaggata | gaaattatat | 9540 |
| gtaaaagaaa | ccctgggccc | ggcgcagtg | cttatgcctg | taatcccagc | actttgggag | 9600 |
| gccgaggcag | gtggattacc | tgaggtcagg | agattgagac | cagcctgacc | aacatgggtga | 9660 |
| aatgtcatct | ctactaaaaa | tacaaaaaaa | attagccagg | catggtgggtg | cacgcctgta | 9720 |
| gtcccagcta | ctccggaggc | tgagacagga | aaatcgcttg | aacccgggag | gcggaggttg | 9780 |
| cagtgagcca | agattgtgcc | attgcactcc | agcctgggca | acaagagcaa | aactccatct | 9840 |
| taaaaaaaaa | gaaagaaaga | aaccctctgg | cagttgatga | gaaggaaact | taatcggcag | 9900 |
| gtcccagcag | gggagatgag | gagactctag | ggagggcaat | tgacacatgct | gtgccccagt | 9960 |
| gtgggcccag | gagcaggtca | ctactcctcc | cgtctacctt | cctcttgctc | caaccccttc | 10020 |
| aagctttgga | ccagtggtac | cctaagtgtg | gtccaaggaa | ccacatgcat | caggaccccc | 10080 |
| agggggtgct | tgtaaaaaat | gcaaattttg | gccaggtgca | gtggctcaca | cctgtaatcc | 10140 |
| cagcactttg | ggaggccgag | gcgggtggat | cacgaggtca | ggagatcgag | accactctgg | 10200 |
| caaacacggt | gaaaccctat | ctctactaaa | aaaacaaaaa | caaaccaaaa | aaaacattag | 10260 |
| ctgggcgtgg | tggcgggccc | ctgtagtccc | agctactcgg | gaggctgagg | caggagaatg | 10320 |

ES 2 597 845 T3

gcgtgaaccc gggaggcgga gcttgacgtg agctgagatt gcgccactgc accccagcct 10380
 gggcgacaga gcgagactct gcctcaaaaa aaaaaaagca aatttcttgg gcaccacccc 10440
 acattgactg catgagaacc cctggaagta tggccctgga acttgcatth tttaacgatt 10500
 ttttgaggta aaatttacat acctaaaatt tacacattgt aagaataata tttgatgatt 10560
 tttagtaaht gtagagttgt gtaaccatta ccacaaaaaa atccattttg aggcatttct 10620
 gccatcccca gaagttccca tttgtagcta ctctccattc tcaccccaag ccttaggcca 10680
 ccactgatct gctttctgtc cataaatttg catattctgg acatgttcta taaatagaat 10740
 atataaatat gcagactttt gcatcaagcg tcttttgctt agcataatat ttttgaggct 10800
 cgccccaaac attatgctaa gtaaaaagaa cagcaataaa aagaaacaaa ctattatcag 10860
 cagtttgtht ctttttattg ctctcatagta tgccattgta cgggtgtaac atattttgth 10920
 tatccattgg tcagttgacg ggcacttggg ttgtttccac tttttagctg tcttaaaataa 10980
 tgttgctatg aacattcaca tacatgcctt actatagaaa tatattttta gcgcttttga 11040
 gtaaactcct aggagtgcaa ttgctggatt atatggtcgg ttttaggtta atttttaaga 11100
 aactgtcaag ctattttcta aagtggctgt attattttgc attctcacca gcaatgtgtg 11160
 agggttccac tttccacat cctacactta gaattgtctt ttttttctt ttttttttt 11220
 gagatggagt ttcactcttg ttgcccaggc tggagtgcaa tgaggatgc tcggctcacc 11280
 gtaacctctg cctcctgggt tcaagtgatt ctctgcctc agcctcctga gtagctggga 11340
 ttacaggthc gtgacaccac gtccggctaa ttttgthttt ttagtagaga cggggtttct 11400
 ccatgttagt caggctagc ctgaaactcc gacctcaggt gatccaccgg ccttggcctc 11460
 ccaaatgct gggcttacag gtagagcca cctgacctg cctttttat ttatttttg 11520
 ttttgatgag gttgtctct ctgtcaccca ggtcggagt cagtggcaca atctcggctc 11580
 actgcaactt gtgcctcctg ggttcaagca attctcctgc ctcagcctcc ccagtagcta 11640
 ggattatagg cacgcaccac cacaccagc taatttttgt atttttagtt gagacggggg 11700
 ttcaccacgt tggacaggct ggctggtctc gaactcctga tctcatgatc catccacctc 11760
 agcctcccaa attgctggga ttacaggcat gagccaccgt gcctggcccc tgtcttttt 11820
 attataacca tcaagtggta tctcagtata gttttaattt gcatttccct aatgactaat 11880
 aatgttaaga acttttcat gtacttattt aaccattcat ctctcatttt tagtgaaatg 11940
 tctatttaca tatttgacc attttctatt taggttgcta tcttattgth gagttataat 12000
 agttccttat ataataat ctgaatacag tcctttatca gatatatgat ttgaaaatat 12060
 tttctccag ttttggttt atcgthttct tttctatttt taaaatttt atttttttt 12120
 tgcattttta aaaattatct ttttttttt tttcatagaa acagggtccc gctatgthc 12180
 ccaggctgggt ctcaaactcc tggcttccag tgatcctcct gtcttggcct ccctaagtgc 12240
 tagggttaca ggagtgagca actgctctct ttttttttt taattaaaa actttttaga 12300
 gataggthc actaagtht tcaagctgga gthtggthc tatttacagg caagatcata 12360
 gcacactgca ccctcgaact cctgaaactca agthgatccc ctgctgagc ctctcgagta 12420
 actgggacta cattctcatg ctaccatacc atgctatttt cttaataata tcttttgagg 12480
 tgcaaaagth ttagacttht atgaagtcca gthtatcatc attttcttt tttcttctt 12540
 tcttttttt ttgagaccag tctcactctg tcgcccaggc tggagtgcag tggcgtatc 12600
 tcagctcact gcaacttccg cctcccgggt tcaagcgatt ctctgcctc agcctcccaa 12660
 agactggga ttacagthg gtgccaccat gctgggctaa tttttgtht ttttagtaga 12720
 acagthtccac catgttggcc aagthgctct taaactcctg acctcagth atctgcccgc 12780
 ttcagcctcc caaagthcgt ggattccagg agthcagcac cacacctggc ctatcatttt 12840
 cthttatggg cctgcttht ggtgthgth ctaggaactc tthgctcaac ctgagthc 12900
 gaagthttt ctctaaagth ttctgthgth gthgthgth ttattthgth tthtagtaga 12960
 gatgggtht tcccatgth tccaggctgg tctcaaaact ctgacctcaa atgatccacc 13020
 cgccttgcct tccaaagth taggattaca gthgthaaagc aggacgctg gccttccatg 13080
 tttcttcta gaaaaactt gcaactthaa caaggcttct tggagagth ctgtgcacact 13140
 tgcctagtht cagaagthc accatatcta ccacctgcgg acagagthc ttgtctgaca 13200
 gthtgaataa ccgggagthc ttgtgthcgt aggagthgth gtgthgthgth gtgthgthgth 13260
 gthgthgth gtgaatggct tctgcatatg tgaccagcag gacaccata gatcacttca 13320
 gcttccaggag tccaccagth gaatcctgth ctgthgthgth gggagacaaa cthtagthc 13380
 tctgctthc cctgthgth agthtaattc ctactccac cactgcccct ctgthgaccga 13440
 cctgcttccc taacctgga ttccatccct ctgthcattt tcaagggag aagctctgag 13500
 aaaaaaggg acagthggg gaagcagcaa gthcagthc tcatatacac aaacacatga 13560
 gcaatgthc ctctgcccct atactagca gcacctatg ttgcataaat ctggcatcaa 13620
 caggagaaag cacttgcaaa tatccatctc taggctacca cattctcca tactccagct 13680
 catcctcatt gctgthtta tcacaagggc atacagthc thgthacagth catccttht 13740
 tathttttt gagatggag thcgtctgth ctgthcaggct ggagthcagth gthgcaatca 13800
 tagctcactg cagcctcaaa ctgggctcag thcattctgth gcctcagcct tccgagcagc 13860
 tgggactatg ggcacatgth accatgctg thtaactthaa aaaaatttt tthttagag 13920
 atgagthc actaagthc ccaggctgth ctcaaaactc tgggctcaag thgctcctc 13980
 gcctcagtht ccaaaagth tgggthtaca gthgthgth accatgthc ggcagcctg 14040
 gccctagtht ttaatagcag thggagthgth aaggctgthc atcctthttc tccaaccct 14100

ES 2 597 845 T3

| | | | | | | |
|------------|------------|-------------|-------------|-------------|-------------|-------|
| acttctggtg | aggggtgatg | ggagggaaaga | cagttacttt | gcccttcttt | ttatttttaa | 14160 |
| tttatcaatt | ttttgttttt | gtttttgttt | ttgtttttga | gacagggctc | caactgtgcca | 14220 |
| cccaggctgg | catggtcagc | tcaactgcaac | ctctgcctct | tgggctcaag | caatcctcct | 14280 |
| acctcagcct | ccagagtagc | tgggactgca | ggcatgtgca | atcacgccgg | gctaattttt | 14340 |
| gtattttttg | tagagacaag | gtctcattgt | attgcccagg | ctggtcttga | actcctgggc | 14400 |
| tcaatcctcc | cacctcggcc | tcccaaaagt | ctgggattac | aggtgtgagc | caccgtgcct | 14460 |
| ggcctcattt | gtagttttat | ttgtttgaag | tcaagctcga | ggtataaatt | ttttatacat | 14520 |
| acagtatagt | tcagttcggg | ttcccagagc | ccttgcatca | ttagcttatt | tcatattcac | 14580 |
| tcatggccac | ctgtgccaa | accctgcctc | ccacacactc | atgcttggtg | gggaagacac | 14640 |
| ccatggaaac | agaccaggaa | ggttcagtgc | ggtaaaggcca | gggatggggg | tgagagggag | 14700 |
| ctggggggca | ggggagagaa | gctctgcaca | ggggttgggt | cctaaactca | gtagagtgga | 14760 |
| gcaggacccc | tgagaaccga | aggagcactg | tggggctcaa | agagctcaca | ccctgggagt | 14820 |
| cactacatcc | caggaggcac | ttcacacttg | ttccatttgg | cctgtgctgg | gtttttaaat | 14880 |
| ctttgaatca | ttgccaatgc | ttaaaactcc | agagatttca | cagaaaaatc | cagaatccta | 14940 |
| gtttcaaaaa | tcagcaaaac | aaaacaaaaa | cacaaacaaa | ccaggagctt | tgtgtcttgc | 15000 |
| ttggcgggac | tttgcataag | gtgcccctct | agcctgccaa | cctccccac | agcccactta | 15060 |
| ctatgttacc | tgttgagccc | ctcctccctt | gaagcctgga | gctttgctgt | ccaccaagtg | 15120 |
| gccactatag | ccacgagtgg | ctcttgagca | cccaaaatgt | gactagtctg | aattgagatg | 15180 |
| cgcttaagca | taaaatgtcc | attggatttt | gaaggctaat | ttcaacaat | gtaaaatc | 15240 |
| tcactaaata | tttttaaaat | attgattaca | tgttgaatag | attctatttt | taaatatttg | 15300 |
| ggctaaataa | aatatactaa | atagtaatat | attatttact | attatattga | ttaaaattaa | 15360 |
| cttcacatgt | ttctttttac | tgtttaaaat | gtgggccagg | cctggtggct | cacccttta | 15420 |
| attccagcac | tttgggaggc | caaggcaggc | aggtggatag | cttgagtcca | ggagttcaag | 15480 |
| accagtctgg | gcaacatggt | gaaaccccat | ctctatgaaa | catacaaaaa | ttaccaggc | 15540 |
| atggtggcac | gcacctgtag | tcccagcaac | ttgggaggct | gaggtgggag | gattgcttga | 15600 |
| gcctgggagg | tggtggttgc | agtgagctgt | gattgtacca | ctgtactcca | gcctgggcaa | 15660 |
| cagagcaaga | ccctgtctca | aaaaaaaaaa | tggtggacac | taggaaatgt | aatgtatac | 15720 |
| atgtggctca | cattatattt | ctcttggtca | gcttgggtct | agagccta | ggacaaatgg | 15780 |
| ggaccagca | agaggggcag | gcagagccag | aacacaaggg | aggggtgaagg | aggccattgt | 15840 |
| aacatcagtg | ttaaagctgt | gaactctgga | gtcagtgctc | agagttgaag | gttcatttct | 15900 |
| gccccttact | tgtgtccca | ctttaaactc | tttgaacttt | cagcacctgg | caactattga | 15960 |
| caggtgaaat | tatagttggc | actcattgac | aggtgacatc | attttgtcta | gctctggagc | 16020 |
| tgggatttga | ttctcagga | actgaccggg | ctctgatgag | ttttaggctg | agtttttggc | 16080 |
| tgcagaatga | tgcattcaga | tggaaagatc | cctctggact | ggggcagggg | gaccagagca | 16140 |
| aggtctgtgg | agaaattcac | atgggaaggg | aacctgtgtc | tgtgtgcatg | cagaggagtg | 16200 |
| gccacataca | aggcacataa | agaaatgaaa | gtcgtaggat | gggcaggggt | gggagaaaca | 16260 |
| aaaaaagaaa | gctgtaggat | ctagtgatgg | atgggatgtg | gtgatggatg | gggtgtgggtg | 16320 |
| atggatggga | tgtgctgggg | gagaggggtg | cagaggggat | cacagaggtt | tctggctggg | 16380 |
| ctgctgggtg | gaggtgggga | ggtttatcca | gagagggaga | gactagacgt | tctgccaggc | 16440 |
| tacggatcag | gctagccttt | cctacttttg | tattcctcag | ccctggcacc | attcttggca | 16500 |
| tgaaagtgtc | cagccaatat | tgccaagtgg | ctggaaggct | aacagcaccg | cctttgagac | 16560 |
| tgcagtcttg | ggtttctac | cggttctccg | tctaccagct | gtgtgatctc | aggcatatta | 16620 |
| cttcacctct | ctggacatca | gtctcctcat | ctgtaaaatg | ggaatagtga | taagtgtagc | 16680 |
| cccttgtggg | ttgttccaag | gactgagtgg | gtttatatac | gggagcacac | agcagtgagt | 16740 |
| gctcagtagc | tgttagtccc | agactccaca | cccagcctc | actcctgatg | tacacacgag | 16800 |
| gccttgtcta | ccaaggctgt | cacctcctta | aaggaaagaa | aggccaaagg | gttccgctaa | 16860 |
| agaagcaagt | gcttcgatat | tctcccttcc | cctgaggcta | agcaggagga | tggagggttag | 16920 |
| actgcaggaa | gaactttcct | tctcttatag | ccaagaatgt | ctaagtgaat | tttagaaaaat | 16980 |
| taacttctgg | ccgggcacgg | tggctcatgc | ctgtaatccc | agcactttgg | gaggccgagg | 17040 |
| cggttgatc | ataggtcag | gagtttgaga | ccagcctggc | caatatgggtg | aaaccccatc | 17100 |
| tctactaaaa | atagaaaaat | tagccagcgg | tgttgggtgg | cacctgtagt | cccagctact | 17160 |
| caggaggctg | aggcagaaga | atcgcttga | cccgggaggg | agaggttgcg | gtgagccgag | 17220 |
| atggtaccac | tgcactccag | cctggtgaca | gagcgagact | ccatctcaaa | aaaaaaaaaa | 17280 |
| aaaaaaagaa | agaaagaaag | aaaattaact | tctcagattc | cttctctggt | cttcagacaa | 17340 |
| ttagtagata | attagggagt | gtttggggag | ggggtatagg | gcctatctcc | tggctcctt | 17400 |
| tagcattcgc | agagtgactt | ggcaggaaga | gagggccagg | gctgaaggca | ttaacagggc | 17460 |
| aataacctcc | ctgtctccag | gatgtggagc | ccacagtctc | caggcccctc | accgatgcag | 17520 |
| gggtgaggtg | gggtgggagg | gcagagagga | aaaacgtggg | aatcacgaat | gggaagaggg | 17580 |
| gatttgggga | ctccctgtga | ttgccagacc | ctggcctgac | cctgctgcag | ggccccaggg | 17640 |
| gtcctgcctc | cagctcacac | tcatatcctt | acatccaggt | ccctctctga | tggcatccgc | 17700 |
| taaaacacca | agattccgct | ggagctgitt | gagtttaaga | taggaccagc | ctggttca | 17760 |
| gagggggaca | cagattcaag | agccagacaa | ctcggagttc | agtctctgag | ttcacatcca | 17820 |
| tcagttgagg | gaacttgggc | gagtcacttt | cacctcactg | tggctcagtt | tctcccttca | 17880 |

ES 2 597 845 T3

| | | | | | | |
|------------|-------------|------------|------------|-------------|-------------|-------|
| taaagcaagg | ataccagacc | ctcccttgag | tggggctgtg | tgcggcagcc | actggctcag | 17940 |
| cgtttccctc | ttctccacag | tgcccatccc | ctccaaagcc | agcagcctct | cagccctctc | 18000 |
| cttggccaaa | gacagcctgg | tgggcggcat | cacaaatcct | ggtgaggtct | tctgctccgt | 18060 |
| gcccggccgg | ctttcactgc | tcagctcaac | gtccaagtac | aaggtgacgg | tgggggaggt | 18120 |
| gcagcggcga | ctctcgcctc | ccgagtgcct | caacgcctcc | ctcctggggg | gtgtcctccg | 18180 |
| caggtaggaa | ggccagccca | caattccccg | cctgattgga | tccttggcct | cttcaggcct | 18240 |
| tctcaaggca | tcagagagga | ggccagtctc | acctaggccc | tctgcctcag | tctccctggg | 18300 |
| aggggagggc | ccggggactc | tggattgtgc | atgttgtggg | ggccggccca | ggtccctggg | 18360 |
| ggctgccaca | tgtctgtggg | tctggggtag | ggaggtggcg | agagaaagaa | gcctgggcac | 18420 |
| agagaaggga | tggagaactc | atccacactg | gtcaggacga | gactgggtct | ggggaagaga | 18480 |
| gagggccatc | ctcaccacct | cttctaggaa | atgaacccca | ggacccgagg | cccattctgc | 18540 |
| gcggcagcat | aggccctgtg | tgttttgggg | tgtggagcat | ggtagtccat | ggttcatctt | 18600 |
| cctagacctg | tgtctctctc | tctctgtgca | tgtctttgtg | cctgggaagg | gtatttctgc | 18660 |
| ttcagggcca | tgagcttggg | tgagtctgag | gatggtgtgg | acatgtgtgt | gcgtgtgcat | 18720 |
| gtgtctatag | tagcggatgg | ccaggagcat | gtgtttccat | gagtgtgagt | atcctgtggt | 18780 |
| tcttctagtg | ggatgtgtgg | cctaggggtg | gtcctgctct | gtgtgtgtgt | gtgtgtgtgt | 18840 |
| gtgtgtgtgt | gtgtgtgtgt | gtgtgtgatt | gcctgtgcat | ctggagctag | tcatatggat | 18900 |
| gtttacatat | ctgtgccact | gtgtgggagt | ggcatgtgtg | cagtcatagg | ggttatctcc | 18960 |
| ttactcctag | cgtcttgagg | gatctctgtg | agtccctgag | tgtctgtaca | tgtgttccctg | 19020 |
| ggaccctgac | ttgggtgtgat | cctgtgtctc | tgttccctg | ggtgtctgag | tctgtgtctc | 19080 |
| tgcatgtctg | tctctattgg | ttctatgtgt | ctctattctt | tcacgtgtgt | gtgtgtgtgt | 19140 |
| gtgtgagaga | gtcaccaggt | ggctgtgtgt | gtacctgtgt | ctccccatgt | gtctctcttt | 19200 |
| ctgtatgtca | ctgcgtgtct | ctacatatag | gtctgtgtat | cccctggtag | gtgtgtgtgg | 19260 |
| ccctctggtg | tgctaaggtc | cccaggatgc | cccaccctgg | ctgcccctgc | tgttgagagg | 19320 |
| attagagtga | actgaggggc | tcctggagct | gggtgctggc | cccctgcct | ctgtcttggc | 19380 |
| tgcaaaacca | tggagattag | ttgcgccctt | gctgaaagag | acctcagagc | atccccagga | 19440 |
| ccaatcccaa | acctcaacag | gggctgggtg | aggcaacaag | gccctttggc | agccacttag | 19500 |
| cttccatgcg | tttctcttgt | cttcatagtt | gtatgtctgt | ctgtctctgt | gcatgtcaag | 19560 |
| ggccaagtcc | aaaaatgggg | gccggtgttt | gcgggaacgg | ttagagaaga | ttgggctcaa | 19620 |
| cctgcagctg | ggccgtcgca | aggccgcca | tgtgacgtgt | ctgacttctc | tagtggaggg | 19680 |
| tgagtgaggc | ctgaaggtgg | gcatgggagt | ggatgtgagg | gcaagtgagt | tgtctgggtg | 19740 |
| gactgtctct | aatgcaggag | gggtgtttaa | gtgggtgtgt | tgtatccgtg | taagtgttgc | 19800 |
| agtatctgtg | tgcgatttct | ctgtcatgta | taaggtgtgt | ttgtgatttc | tgtgtgtagt | 19860 |
| ccttgtacat | gtgtgcatgc | atgtgcaatg | gagacctact | tgcagtgtgc | ggatgtgtat | 19920 |
| gtgggtgtgg | gtgtgggtgt | gtgggtgtgt | gtgttggcgg | ggagtggcca | gtgtttggat | 19980 |
| gtggtaaagg | gatgaagggg | gtgacagccc | cctcccctgg | aggttgggtc | ctctggaata | 20040 |
| catggtgggc | ccagctccca | tctctgagct | cagagcccaa | ccctaattggc | cccaagggtg | 20100 |
| ggcaatggaa | tgggggctgg | agctgggctg | ggaaggaaca | tcagaggggg | catctacaca | 20160 |
| ggcaggatgg | ggcaggatac | ccctgaccag | ggggtcagag | gttcaaagct | gtgttccagg | 20220 |
| cgcacaggta | caccctgcag | tagtgacagc | tcccctcccc | ccaggagagg | ccgtgcacct | 20280 |
| ggcccagac | ttcggttacg | tctgtgagac | ggagtcccc | gccaaggcag | ctgcccagta | 20340 |
| cctgtgcca | cagcacgctg | accgggggga | gctgcacagc | cgcaagagca | tgctcctggc | 20400 |
| tgccaagtga | gtgagggcac | cctgcacagg | cacacgtggg | tgccatgcac | agacagacat | 20460 |
| catgtatggg | cacaatggac | accactgtgt | acatgagcag | tgggcacaca | tgcgtatttg | 20520 |
| cgccacctg | tgtacacgag | cagtgggcac | acacacatac | gtgcgcacca | ctgtgtacat | 20580 |
| cagcagtgag | cacacacgta | tgtgtgcgcc | actgtgtacg | tgagcagtgg | gcacacacac | 20640 |
| gtacgtgctg | gccactgtgt | acacgagcag | tgagcacaca | cacgtacctg | cggccccatgt | 20700 |
| ggcaccacac | atgggcacac | ctaggctccc | tgcttaggct | ttcctgggca | tatgtgtgca | 20760 |
| tgccacctgg | gcacacttgg | gtgccacata | tgggtatatt | tgggcatcat | gtttatcact | 20820 |
| cagacataca | tgaacactct | tcacaagctc | ctggagcaca | catagacact | acatgtgaac | 20880 |
| atatctgggc | tcctgcatca | ggaacagtga | tgtcacatac | tggctcacac | aggctcacgc | 20940 |
| atgggtccta | cactcacata | cactcccagg | acaccaccag | ccctactccc | catcgcaggt | 21000 |
| ctgaggctcc | ctcttctagc | cttgcccctc | aaccctctgc | tccagtgacg | tccatcaggc | 21060 |
| caccaagtca | cacctgactg | ttctgcaaat | cccattgtgc | ctccctcact | accgttcaca | 21120 |
| tgctgccacc | tctactcagg | tcttcatctc | actcctggac | cttcaccca | gagcgtccct | 21180 |
| gtttccaacc | tccttgcacg | tgcacccgcc | ccactgcata | agaatcttca | atggttcccg | 21240 |
| ctgcattgat | tttcaacta | ggtctgcgag | ggcactcag | ggcctgttta | acaaactagt | 21300 |
| attctattaa | gatttagttt | tgtttttgtt | ttttgagatg | gagtttcgct | cttgtttccc | 21360 |
| agactggagt | gcagtggcgc | aatctcagct | cactgcaacc | tccacctcct | gggttcaagc | 21420 |
| aattctcctg | tctcagctc | ccgagttagt | gggatcacag | gcgcctgcca | ccaggcccag | 21480 |
| ctaatttttg | gtatttttaag | tagagatggg | gtttcaccat | gttggccagg | ctggtctcga | 21540 |
| actcctggcc | tcaggtgatc | cacccgcctt | ggactcccaa | agttctggaa | ttacaggtgt | 21600 |
| gaaccaccac | tcccggcccc | aagatttagt | ttttcaggcc | aggcaaggtg | gctcacatct | 21660 |

| | | | | | | |
|------------|-------------|-------------|------------|------------|-------------|-------|
| gtaatcccag | cacttttgag | aggctgaggc | tggtggatcg | cttgagccca | ggaattggag | 21720 |
| actagcctgg | gcaacatggt | gaaaccgggt | ctctagtaaa | aatacaacag | ttaggtgtgg | 21780 |
| tggtgcactc | ctgtagtccc | agctacttgg | gaggctgagg | tgggaggata | acctgagcct | 21840 |
| gggaggtcaa | ggctgcagta | agccatgatt | gcatcgctgt | actctagcct | gcgtgacata | 21900 |
| gtgagacccc | atctcaaaaa | aaaaaaaaag | cgaaagattt | agttcttcca | aaagagttct | 21960 |
| tgtcttagtc | catttgggct | gctataacaa | aataccataa | actgggtagc | ttataaacia | 22020 |
| cagaaatfff | tatttattta | tttattgaga | cagaatcttg | ttctgttgcc | caggccggag | 22080 |
| tgcagtgggt | caatctcagt | gcactgcaac | ctccagctcc | tgggttcaag | caattctcct | 22140 |
| gcctcagcct | cctgagtagc | tgggattaca | ggtgcctgcc | accacatctg | gttaatffff | 22200 |
| gtatffffag | taaagtttc | accatgttgg | ccaggctggg | ctcaaactcc | cgacctcaca | 22260 |
| tgatccacct | gccttgggct | ccctaagtgc | ttggattata | ggtgtgagcc | accacgcccc | 22320 |
| gccaacaaga | gacatttatt | tctcatagtg | ctggaggctg | ggaagttcaa | ggtcaaggca | 22380 |
| gattccgtgt | cagggtgagg | cctgctttct | ggctcataga | cagcaccttg | ctgtgtgtcc | 22440 |
| tcacattgta | gaaggagaa | ggggttactc | tctggcttct | ttataacaa | ggcattaat | 22500 |
| cccattcatg | acggctccac | ctccatgacc | taatcacctc | ccaaagtcct | cacctcctaa | 22560 |
| tactatcacc | tagtggggtg | gaatttcaac | atacgaatff | tggggaacat | aagcattcac | 22620 |
| accatagcag | tcccactgca | agaaaagaca | gagagaggag | ggggatagag | aaagagagaa | 22680 |
| agaaagagag | agagactgag | tgcagtggct | tatgcctgta | atcccagcac | tttgggaagc | 22740 |
| tgaggtgatc | agatcgcttg | agcccaggag | ttcaaggcca | gcctgggcaa | cttggcgaaa | 22800 |
| ccccatctgt | actaaaaaaaa | aattagccag | gagtgatggg | gcgcacctgt | agtcccagct | 22860 |
| actcgggagg | ctgaggttga | gactgcagtg | agctgggatc | acaccactgc | actccagcct | 22920 |
| gggtgacaga | gcaagacctt | gtctcggaaa | aaaaagaaaa | gagagagata | aatctgagaa | 22980 |
| ctaccagttg | gtaggctaaa | gttcaaactc | ttttgagggc | cccacagtct | ggcccacaca | 23040 |
| aactccaccc | acttctgcca | accctgtcct | cacatccttg | ggtttgtttc | cggttagaa | 23100 |
| catccagtc | tcccacctaa | cacagtcctt | caaactgaca | actccacctc | cctgcctttg | 23160 |
| aggctttcct | caagtgccca | aggcagaaga | ggaagcacag | actcctggga | ctgccaaatg | 23220 |
| gagagacctc | tgaggtgct | catactgtgg | ctggagaaac | ctagccagga | cggggctcac | 23280 |
| ctcgttcact | tatcagctca | ttcatcaagc | actttgtggg | tctcagtcac | gtgtaagata | 23340 |
| ctggctggct | ttcatggtga | cacaaagatg | aatgatgagt | gattctcttg | cctgagccac | 23400 |
| acacaggacg | cacaaaggaa | tgaataatga | atggctgcaa | actagagcta | ggttcaaatc | 23460 |
| ccagctcggg | caccttgaca | ggtcatccaa | cctgtccgaa | accagttttc | ttgtccgtat | 23520 |
| gatggggata | ataatagtat | agacttgtta | tgaagattaa | gatgatgcag | acagtgtggg | 23580 |
| tggcacagca | tccagccttc | aattatattc | actgctagtt | attatftccg | gcagattctg | 23640 |
| gattgtttcc | tgatagcttc | aaagacaaag | ttctgtgaag | cattgagtgt | ttgggggtgg | 23700 |
| agagggaaac | caggaaagtt | tcctatataa | ggtagtaagt | agctgagctg | cactttgaa | 23760 |
| gaatatgggc | cagacatcca | aaaaagaaac | attacagggt | gatcgacatg | tgctgtttat | 23820 |
| tagctgtgtg | atactgggct | agtcaaatga | cttctctgaa | ccttagtttc | cttttctgta | 23880 |
| aaaaacagta | tggtcatcat | ttgcagactg | ttgggagggg | cagcaattgt | gtagctaattg | 23940 |
| tctgtaaaat | gcctagcatg | tagcaggctc | ttctctgggc | tcagactttt | gtcctccaac | 24000 |
| ctctgaccct | ccttctcgca | ccaggcagat | ctgcaaggag | tttgcagact | tgatggctca | 24060 |
| ggaccgctca | cgctgggca | acagccgcc | agcactcatc | ctggagcccg | gagtacagag | 24120 |
| ctgcttgaca | cactttagcc | tcatcaccca | tggcttcggg | gggcctgcca | tctgtgctgc | 24180 |
| cctcactgcc | ttccagaact | atftgctgga | gtcactcaag | gggctggaca | agatgtttct | 24240 |
| aagcagtgtg | ggcagtgggc | atggtgaaac | caaggcttcg | gagaaggatg | ccaagcatcg | 24300 |
| gaaataactg | cttctcccac | cccatcccta | aggggctccc | agccctgaa | atagggactt | 24360 |
| agctcttggg | ggtgggcctg | gaaggactga | aagggtggat | tagagtcaag | ccagaaagag | 24420 |
| aacattcatc | cagagatccc | agagttgggg | atctggcttg | gagtaaggga | gggtggcctc | 24480 |
| tctgtgggtg | tgtgttggtg | agttaagggc | ccaggtatft | gtctcatgtg | tgcaatfttc | 24540 |
| tgacctttga | tggttgagaa | gggtttggac | agaaaattga | catgaaaaga | tctggctcat | 24600 |
| ggggcagagc | cctttccatt | agcgtggctg | ggtggccgtg | ggtgcttcta | gaggccaaag | 24660 |
| cctttgtgtt | tttcactggg | ggcaggagga | aaattgataa | atcagaggtg | ctactgagga | 24720 |
| gttgggtgcc | ctcattccag | aatctcctac | ccccagaaaa | ggggtgctgg | aaggaggccc | 24780 |
| cagtggactc | tttgtaccct | tcctactctc | agagagaagt | gggcaggagg | ggtcctcaag | 24840 |
| gaacaaagaa | gataaagcac | aaatcgcaga | acttgaatcc | aggctgctgc | tcatcatagt | 24900 |
| cctgttgctg | tctgtcctat | ttatfttatgt | atgttgtaat | taaatttgaa | atfttaaaa | 24959 |

ES 2 597 845 T3

<210> 2

<211> 6096

<212> ADN

<213> Homo Sapiens

5 <400> 2

taccagtgta agattcaaaa ttcccttttt gcactgcaca gtgagatgcc cagggctcca 60

ES 2 597 845 T3

| | | | | | | |
|------------|------------|-------------|-------------|-------------|-------------|------|
| gctcagtgcc | tggacatagc | gattccttggg | cctgcccgtc | gccgccccaa | gcgaagctgg | 120 |
| tgcgccttgg | gcggagcaga | cagagaccct | gggtggcagg | ggcttgggaa | gacatgggcg | 180 |
| gctagggcct | tatgcgcctt | caccgctgcc | ctctgctatt | tgcaggcaat | ggacgagccg | 240 |
| ggaatgagcc | tcctagacca | gtccgtgatc | aagaaagta | aggaatggtc | tgtcagggca | 300 |
| gagcccggcg | agatggtgca | ggcccttgg | gcacagatcc | atcttcttca | ccggccgtgc | 360 |
| ctcctgtgtg | tcgccaggct | gggtgtccac | caggcaactct | tccctggccca | gccagatgtt | 420 |
| aggcagacgt | gcgggcttgg | tgagtttgcc | cagcaccctg | tggcctgggg | tgggcctcag | 480 |
| cggatcagca | ttcactgggc | tgcagcactg | ggagcctggc | ctctccccgc | cgagggggag | 540 |
| ggcactcttg | tggatctgga | gttgatttgc | agaacgagtt | aaaccacttc | cctgtttccc | 600 |
| taagagatgg | gaatggaagt | gctgttccca | cggagtggg | gaaatgattt | tcactttaca | 660 |
| gtgccttagc | atctcggtgc | ctggcgggca | ctttcttcc | cttccctcca | ggcagggcct | 720 |
| tggaggcctc | tgggggaatt | ttctttctgt | gggagtctct | tgcggcattt | agacttaggg | 780 |
| gagcttgtgt | gtgagtactg | tgtgttaggc | tgtgtgcacc | tgagttaggg | cccacctgct | 840 |
| cctgggtgtc | tgtgtccatg | tgagttcagg | gtcctgtgca | tgtctgaaat | gttcccttca | 900 |
| tgggtgtctt | agtatttctt | ggagtgtgag | tgtgtctggt | tctgtgaatg | tgtttgtgag | 960 |
| gtgtgtctct | gtatgttgg | gtgcatttct | ctgcatttgg | gggatgtaca | catttctcaa | 1020 |
| tatgtacagt | atctctgttg | tgtcctgcac | tttgttcttt | ggtatctgag | gatttccaag | 1080 |
| catgcgcggg | ccctctctgt | gtatataatg | gagtatttat | gtgactcctg | gcattagtaa | 1140 |
| aatccagggg | cacgggatcc | accttttctg | gcctgaggag | caagtactgg | ccatgacagg | 1200 |
| ggaaggtgag | agacgacaaa | aacagagaga | cagccagaga | ggagcagaga | gtcagagggg | 1260 |
| cccaggcatt | gggtagcagc | ctctttacat | ttggggcagg | tgcccgaaag | aattcagagg | 1320 |
| tgcacatgag | cctgaggtgc | cccaggcagg | cactgctccc | acagggtttg | gcttgagttg | 1380 |
| ttttcaaac | gagtgaattc | aagcctgggc | tctatttggc | 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 | ggaaagacac | taggaacctg | gttgacagcg | agctctgctg | gtggggcttg | 1680 |
| gttggttac | tgtacagagc | ctttcttgac | ccctgaagaa | agagatccgt | ctgcagtggg | 1740 |
| caaaagcctg | cctggacttc | ctggccacca | gaaatatgag | catggtggtg | gtcccagttt | 1800 |
| ccctattcat | gcttgggctc | aagagactgg | gagtctaggt | tcactgactc | cctgagaaaag | 1860 |
| actaagaccc | tgcattttag | aaagagttt | gggatctct | gccctgcgca | agggtagaag | 1920 |
| gatcagctgt | tcctctgagc | accttaaccc | ggaaccccg | tccgaagccg | agacaggaga | 1980 |
| ctggatgca | ggccctccca | gagctggttt | ctctcaaaac | acttcaaaaa | ctcctagatc | 2040 |
| ctaggggtac | gccgaaatcc | cccaaagcag | tccaaagaac | acaacgagag | tcctaacatc | 2100 |
| ccaggtggcg | gcgcgctggc | tccctggagc | ggggcgggac | gcggccgcgc | ggactcacgt | 2160 |
| gcacaaccgc | gcgggacggg | gccacgcgga | ctcacgtgca | caaccgcggg | accccagcgc | 2220 |
| cagcgggacc | ccagcgccag | cgggacccca | gcgccagcgg | gaccccagcg | ccagcgggac | 2280 |
| cccagcgcca | gcgggacccc | agcgcagcgg | ggaccccagc | gccagcgggt | ctgtggccca | 2340 |
| gtggagcgag | tggagcgtg | gcgacctgag | cggagactgc | gccctggacg | cccagccta | 2400 |
| gacgtcaagt | tacagcccgc | gcagcagcag | caaaggggaa | ggggcaggag | ccgggcacag | 2460 |
| ttggtaccgg | agtgctgagc | ccaggggaaa | gcgtggcg | tcgacccagg | gcagctcgcg | 2520 |
| cggcgaggca | ggtggctcc | ttgctcctg | gagccgcccc | tccccacacc | tgccctgggc | 2580 |
| gccccagca | gttttcacct | tggccctccg | cggtcactgc | gggattcggc | gttggcccca | 2640 |
| gcccagtggg | gagtgaatta | gcgccctcct | tcgtcctcgg | cccttccgac | ggcacgagga | 2700 |
| actcctgtcc | tgccccacag | accttcggcc | tccgcccag | gcggtaactg | agcctgcccc | 2760 |
| gccagggccc | tggaatcaga | gaaagtcgct | ctttggccac | ctgaagcgtc | ggatccctac | 2820 |
| agtgcctccc | agcctgggcg | ggagcggcgg | ctgctcgct | gaaggttggg | gtccttgggtg | 2880 |
| cgaaagggag | gcagctgcag | cctcagcccc | accccagaag | cggccttcgc | atcgctgcgg | 2940 |
| tgggcgttct | cgggcttca | cttcgccagc | gccgcggggc | agaggcacct | ggagctcgca | 3000 |
| gggccagac | ctgggttga | aaagcttcgc | tgactgcagg | caagcgtccg | ggagggcg | 3060 |
| ccaggcgaag | ccccggcgct | ttaccacaca | cttccgggtc | ccatgccagt | tgcacccg | 3120 |
| gtattgggca | ggaaatggca | gggctgagc | cgacctagg | agtataaggg | agccctccat | 3180 |
| ttcctgccca | catttgcac | ctccagttt | gcaacctatc | ccagacacac | agaaagcaag | 3240 |
| caggactggt | ggggagacgg | agcttaacag | gaatattttc | cagcagtgag | caggggctgt | 3300 |
| atgggacgcg | ggaggagctc | agaggaggcg | cggagagtgc | ccgaggttg | gtgagtgcct | 3360 |
| agaggggaga | tagttgaacc | gggttcaaga | ggtgcttagt | gggtgtttgt | tgaatgaatg | 3420 |
| agtgatgggc | tttgaagtct | gagtgcattg | aaagagggg | tgtgtaaaaa | gggctccttt | 3480 |
| catcacacag | gacacagcat | atgcaaatcc | tctcctgtg | gaaaagccag | acaggttaaa | 3540 |
| aaggttacia | acaaattagc | cgggcatggt | ggtgcgcgtc | tgtagtccca | gctactaggg | 3600 |
| aggctgagcc | aggggaatcg | cttgaacccg | ggaggcggag | attgcagtga | gccaagatcg | 3660 |
| cgccactgca | ctccagcctg | gaaacagagc | gagactccgt | ctcggaaaaa | aaaaaaaaaa | 3720 |
| gttacaacc | gtgtgtgggt | ttcaggttat | acaatcagag | ctggagggga | gtggtcaagg | 3780 |
| atgagaactg | agatggatcc | ctcgttccct | ctggagga | gtgggtggtt | gcctacttgg | 3840 |

ES 2 597 845 T3

| | | | | | | |
|-------------|-------------|-------------|-------------|------------|-------------|------|
| gggtggggaa | tccctctcca | cgggctcagc | tgtccaatct | caggggatct | ctaggacagg | 3900 |
| agctgatgta | aacagtcgcc | ctattccttg | ctgtctttgg | ccctggagaa | ggaggagga | 3960 |
| gctggggagg | gtctccactt | cccagacaat | ctctaagcag | ccaggacatg | ggtgagatga | 4020 |
| gtgagatact | gacttctggg | acagaatttg | agaggggtgcc | aaaaaactca | gtaatcaaga | 4080 |
| taaataggcc | gggcgcagtg | gctcacgtct | gtaatcccag | cactttggga | ggccggatca | 4140 |
| cttgaggtca | agagttcgag | accagcctgg | ccaagatggt | 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 | ggtgggtggg | caggcacctg | taatcccagc | 4560 |
| tactccggag | gctgaggcag | gagaatccct | tgaacctgcg | ggcggagggt | tgcaagtgaac | 4620 |
| caagatcaca | ccattgcact | ccagcctgga | caacaacagc | aaaactctgt | ctcaaaaaaa | 4680 |
| aaaaaaatct | tttttttcga | gacacagttt | tactctctcg | cccaggttgg | ggtgcagcac | 4740 |
| cacgatctca | gctcactgca | acctctgcct | ctcagattct | cgtacctcag | cctcccaagt | 4800 |
| agctgggatt | acaggtacct | gtcaccacgc | ccagctaatt | tttgtatttt | tagtaggggc | 4860 |
| gtggtttcac | catggtggcc | aggctggtct | tgaactcctg | acctcaagtg | acctgcccgc | 4920 |
| ttcagccacc | caaagtgctg | ggattacagg | cgtgagccac | cacgcttggc | ctttttaaat | 4980 |
| gaaaatagtg | caaaaatcca | cgataaacia | aatatcaaaa | atttactgaa | cttgcacttc | 5040 |
| cacaaccctt | tctcacctgc | ctcccaggct | actctctgcc | ccagaaagca | acttaaaaaa | 5100 |
| tgtgcagatg | gagtttgac | tttacctgaa | aatggtggga | gctatggaaa | accttggagc | 5160 |
| aggggagtga | aggatagaaa | ttatatgtaa | aagaaacctt | gggccgggcg | cagtggctta | 5220 |
| tgctgtaat | cccagcactt | tgggaggccg | aggcaggtgg | attacctgag | gtcaggagat | 5280 |
| tgagaccagc | ctgaccaaca | tgggtgaaatg | tcatctctac | taaaaataca | aaaaaaatta | 5340 |
| gccaggcatg | gtggtgcacg | cctgtagtcc | cagctactcc | ggaggctgag | acaggaaaat | 5400 |
| cgcttgaacc | cgggaggcgg | aggttgcagt | gagccaagat | tgtgccattg | cactccagcc | 5460 |
| tgggcaacia | gagcaaaaact | ccatcttaaa | aaaaaagaaa | gaaagaaacc | ctctggcagt | 5520 |
| tgatgagaag | gaaacttaat | cggcaggtcc | cagcagggga | gatgaggaga | ctctagggag | 5580 |
| ggcatttgca | catgctgtgc | cccagtggtg | gccagggagc | aggtcactac | tcctcccgtc | 5640 |
| taccttctc | ttgctccaac | cccttcaagc | tttggaccag | tggtagccta | agtgtagtcc | 5700 |
| aaggaaccac | atgcatcagg | acccccaggg | ggtgcttgtt | aaaaatgcaa | atthttggcca | 5760 |
| ggtgcagtgg | ctcacacctg | taatcccagc | actttgggag | gccgaggcgg | gtggatcacg | 5820 |
| aggctcaggag | atcgagacca | tcctggcaaa | cacggtgaaa | ccccatctct | actaaaaaaa | 5880 |
| caaaaaacia | caaaaaaaaa | cattagctgg | gcgtggtggc | gggcgcctgt | agtcccagct | 5940 |
| actcgggagg | ctgaggcagg | agaatggcgt | gaacccggga | ggcggagctt | gcagtgagct | 6000 |
| gagattgcgc | cactgcaccc | cagcctgggc | gacagagcga | gactctgcct | caaaaaaaaa | 6060 |
| aaagcaaatt | tcttgggcac | caccccat | tgactg | | | 6096 |

<210> 3

<211> 24959

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 3

ES 2 597 845 T3

| | | | | | | |
|-------------|------------|-------------|-------------|------------|------------|-----|
| ttgagagtta | ttgatttatg | agtaagagta | aggtagttta | aggtgttgag | atagggtgtt | 60 |
| atgtgttttt | tgggatattt | gttttaattt | tttattatat | attttttgtg | ggtggatgtt | 120 |
| taagggttag | aaatggtaat | attttataag | gattagtttg | tgggggttaa | ggaaaggatt | 180 |
| tacgtttttt | tgggatgatt | agtaagtttt | gaatttgaaa | aagggtggtg | ttatttatga | 240 |
| aattatagag | taaaggagaa | gtaggttggg | gttgggtgttg | gggatagagg | agatgttgag | 300 |
| tttagtgttg | gatttggaaa | attttagata | gttghtaatga | gatgtttgga | gaaaatacgt | 360 |
| agttggtagt | tgtgtttatg | ggtttggatt | ttttattagt | gatttagggt | agagataata | 420 |
| tagttgggag | ataatagtag | ataggtagtt | attaaagatt | tgggtcgggc | gtggtgattt | 480 |
| atgtttgtaa | ttttagtatt | ttgggagggt | gaggcgagtg | ttattttgaa | gttaggagtt | 540 |
| cgagattagt | ttggttgata | tgatgaaatt | ttgtttttag | taaaaatata | aaagttagtc | 600 |
| gggcgtgggtg | gtgggttttt | gtaatttttag | ttatttagga | aattgagata | tgagaatcgt | 660 |
| ttgaattcgg | gaggtggagg | ttgtagttag | tcgagattgt | gttattgtat | tttagtttaa | 720 |
| gtaatagagt | gagattttgt | tataaaaaaa | aaaaaaaaaa | agatttgaga | gggatgagt | 780 |
| ttatttatga | agagtgtagg | aagaaaaaga | agtttttttg | ttttgagaaa | gaattatcga | 840 |

| | | | | | | |
|-------------|------------|------------|-------------|-------------|-------------|------|
| tattaagtat | gtaggtgaaa | agagaaatag | aaaagaagat | agaaatggtc | ggagtaggat | 900 |
| aatgaggtaa | aaaggggtta | ttaaaggtgt | aggaagtaga | ttgagaaaat | ttaaggaagt | 960 |
| tatagatagg | tagtatat | cgttat | gtttttta | gttggtggag | agagtagggg | 1020 |
| tattttcgta | gttgggtata | ggtattat | tttttgcggt | ttttagattt | tttttagagg | 1080 |
| agtttattaa | tttttattta | cgtattgata | tttgttttg | agtttgattt | ttgatgttt | 1140 |
| gtggttggg | ttgagggagg | gaggaatagg | gtattgttat | agattttttt | atatatagat | 1200 |
| atatataaat | ataattttga | tttttgattt | atctcgagta | ggaaagtgt | aaagtttttt | 1260 |
| aatttcgttt | tattttttg | aatttcggtt | gcggttttag | attttagaaa | gtttttaatg | 1320 |
| atgaattttt | ggtttttagt | ttttttttt | ttttattatt | tttttaggtt | agattattta | 1380 |
| gttttgagtt | ttgatattgt | tatgtgtttt | ttgggatatt | tatttggaa | tatatatttt | 1440 |
| atgagttaat | ggtgttatta | ttattttttt | ttttatatta | ataattttat | tttaattttt | 1500 |
| ttatttatta | ttttattaaa | aagggattta | aaaaaataat | tagttttttt | tgttttttt | 1560 |
| tattttttta | ttggtgttgt | cgaagttgta | gaaattgaga | ttatagagga | aacgggatgt | 1620 |
| agttagatt | atatagttag | gtagagttag | gattaaattt | aagtttttga | gtttgaggta | 1680 |
| gtagagaata | gtgattttga | gtagagaggg | gttaggggta | gattatttta | gttttgttat | 1740 |
| ttagttgttt | gattttgggt | aagtatttta | atttttttat | gttttatttt | tttttgaaa | 1800 |
| atggggatga | tatgggtatt | tttttatag | agttgttgtg | agggttaaa | aatttaatat | 1860 |
| tcgtgaagta | ttttgtatag | aggttggtat | attttgtaat | aagtttggtg | tgattatttt | 1920 |
| tattttgttt | tggtttgtat | attttgtttt | taaaaaat | ttttttttg | ttttttta | 1980 |
| ttttgttttt | tttagtattt | tttttatttg | ttttttttta | atttttgttt | ttgtttgta | 2040 |
| ggagggagg | gttgggtttg | gggagaaa | gagagtgatt | atagagattg | gaatttggtt | 2100 |
| gatagttagg | gagtgtttgt | tatgtaata | taggaatatg | tagtttgcgt | atgtagtgt | 2160 |
| aataattata | cggtatttat | gttagttagt | tatatatagt | tatttaatag | ttttgggta | 2220 |
| ttttatgttt | atatttaatt | agtttaggat | tttggtttt | tagaaatatt | ttggtttttt | 2280 |
| gttaggtgga | gtttatattt | aaatttattt | taggtaaagt | ttttttttg | tttggttagt | 2340 |
| ttttgttttt | aattttttta | gagtttggtt | tggtttgga | gttttttagt | taggaagttt | 2400 |
| gagagcgaag | tgggtatttt | cgagggattt | tgatttagaa | ttatttagga | aagattgtga | 2460 |
| tggtagattt | tttaaata | cgtattgata | ttgtgtgctg | atgagtagat | aggttggagt | 2520 |
| ttttagttta | taaatttttt | ttacgttttc | gattatttat | tttatttaa | gttttttgag | 2580 |
| gtggttagag | aaaggattga | tatttatttt | atcgaggaag | ttgaggttag | cgaagaatag | 2640 |
| tgttttgttt | aaaggtaaag | atgtgataga | gttaggtaaa | tgtagggttt | ttgcgtttta | 2700 |
| gtttcgaata | tttttgtttt | attgtagtgt | cgagttgcga | gtttagaagg | cggttttgat | 2760 |
| taatgtgggt | tgaatttacg | ttttttattt | attttattcg | atctcgtaa | agttattggt | 2820 |
| ttttttgtgg | tagtcggttg | ggaaaaggga | gagaattggt | atagtttgtt | tttttaatgc | 2880 |
| gcgttttcg | aggttcggcg | cgttttcggt | attataattg | gagtgatgg | agtaggttac | 2940 |
| gttagagga | ggaagggcgg | gcgttggtga | ttcgttgatc | gattttttta | agtgcgatta | 3000 |
| gtgttcggtc | gttttgtttt | tatggattcg | ttcgggaacg | gttatcgttg | aggattttac | 3060 |
| gtttattagg | atctcggttg | ggtcgtattt | agttatcgta | tcgtgatttt | cgcggttgt | 3120 |
| gtcggttttc | ggcgtttttg | ttcgtagcgt | tcgtcgtcgg | gttagggttt | cgtcgtcgtt | 3180 |
| acgtttcgg | ttcggatatt | atcgttttat | gttggtgtat | atttatttcg | ttatggtgag | 3240 |
| tagtttcgg | ttcgggat | atcggtcgg | gggatcggg | cgtttgagt | ttggattttt | 3300 |
| taattttttt | gtttcgcgtt | aacgaaattt | cggagggagg | gggtcgtagg | gatttcgtta | 3360 |
| gttgagaacg | cgatgcgtaa | gtgatcgttg | tttttagttt | gaggtttttg | cgttcgcggg | 3420 |
| tggttcggaa | ataaatttcg | ggttttaaga | aacgggaggg | attgtagtgt | aattttttcg | 3480 |
| agttgaggag | gattttttta | ggttggggtt | tttttagttg | ttagtgggtt | atttaaggta | 3540 |
| tttttttttt | tttttaggag | cgtttcgacg | ggttgggagt | agttgtcggc | ggggttcggt | 3600 |
| tgtcgttttt | gttttaggag | gtttacgggt | cggtcgtttt | gttttgttat | acgtcgggtc | 3660 |
| ttatagttgt | cgtcgaattt | tagtcgtttt | atttttcgtc | gttttattcg | tagttatcgt | 3720 |
| tgttttacgg | ttaggcgttc | gacgtcgtcg | tagttttttt | ttatttggtta | ggggatttat | 3780 |
| atggcgggtt | ggcgtttttg | gcgtagtcgt | agttttttta | ggtcgtttgg | gtcgcgtttc | 3840 |
| tcgtagtcgt | tcgcgtttac | gaggagtttt | tcggtttggt | ggtatcgttc | gttcgcgtttt | 3900 |
| gggttttga | ttcgcgtcgt | gattatgta | ttgtcgtggt | tcggtttttg | tcggtttttg | 3960 |
| tcgacggcgc | gtacggtttg | gtagacgtat | ttttcggttt | ttcgggggtg | gcggcgggtt | 4020 |
| tcgggtttgga | ggattttag | gtgagattcg | agggattcgg | gatgggtcgg | gattggtcgc | 4080 |
| ggtggtttat | tattatggtt | ggagtagaa | ggtgataaat | gtaggaagtc | gatttttttt | 4140 |
| ttagttcgtt | atattttatt | ggtgatttcg | aggatgtggt | ttattttttg | ggttagacgt | 4200 |
| tgtttggtta | ttaaggttgt | ttggtattga | gttcgttagt | agtgctggg | tggtagtaat | 4260 |
| tttcggtagg | gatcgaatta | tttttttttt | tttcgttgtt | acgtgtggta | tttatggggt | 4320 |
| agtttgatt | tggggcgaag | agaagtataa | tagttatttag | tgtaagattt | aaaatttttt | 4380 |
| ttttgtattg | tatagtgaga | tgtttaggg | tttagtttag | tgtttgata | tagcgatttt | 4440 |
| tgggtttggt | cgtcgtcgtt | ttaagcgaag | ttggtcgtt | ttgggcggag | tatatagaga | 4500 |
| ttttgggtg | taggggtttg | ggaagatag | ggcgttagg | gttttatg | tttttatcgt | 4560 |
| tgttttttgt | tatttgtagg | taatggacga | gtcgggaatg | agtttttttag | attagttcgt | 4620 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gattaagaaa | ggtaaggaat | ggtttgtag | ggtagagttc | ggcgagatgg | tgtaggtttt | 4680 |
| tggtgtatag | atthattttt | tttatcggtc | gtgttttttg | tgtgtcgtta | ggttgggtgt | 4740 |
| ttattaggta | tttttttttg | tttaggtaga | tgtaggttag | acgtgcgggt | ttggtgagtt | 4800 |
| tgtttagtat | tttggtggtt | gggtggtt | ttagcggatt | agtatttatt | gggtttagt | 4860 |
| attgggagtt | tggtttttt | tcgtcagagg | ggagggtatt | tttgtggatt | tggagttgat | 4920 |
| ttgtagaacg | agttaaatta | tttttttgtt | tttttaagag | atgggaatgg | aagtgttggt | 4980 |
| tttacggagt | tggggaaatg | atthtttatt | tatagtgttt | tagtatttcg | gtgtttggcg | 5040 |
| ggtatttttt | tttttttttt | tttaggttag | gttttgagg | tttttggggg | aatttttttt | 5100 |
| ttgtgggagt | tttttcgggt | atthtagatt | aggggagttt | gtgtgtgagt | attgtgtggt | 5160 |
| aggttgtgtg | tatttgagtt | agggtttatt | tgtttttggg | tgtttgtgtt | tatgtgagtt | 5220 |
| taggtttttg | tgatgattg | aaatgttttt | tttatgggtg | tttttagtatt | ttttggagtg | 5280 |
| tgagtgtggt | tgtttttgtg | aatgtgtttg | tgaggtgtgt | ttttgtatgt | tgggtgtgat | 5340 |
| ttttttgtat | ttgggggatg | tatatatttt | ttaatatgta | tagtattttt | gttgtgtttt | 5400 |
| gtattttgtt | ttttggtatt | tgaggatttt | taagtatgcg | cgggtttttt | ttgtgtatat | 5460 |
| ataggagtat | ttatgtgatt | tttggattta | gtaaaattta | gggatacggg | atthattttt | 5520 |
| tttggtttga | ggattaagta | ttggttatga | taggggaaag | tgagagacga | taaaaataga | 5580 |
| gagatagtta | gagaggagta | gagagttaga | gggttttagg | tattgggtag | tagttttttt | 5640 |
| atatttgggg | taggtgttcg | aaagaattta | gaggtgtata | tgagtttgag | gtgttttagg | 5700 |
| taggtattgt | ttttataggg | tttggtttga | gtttgttttt | aaacgagtga | atthtaagttt | 5760 |
| gggtttttat | tgttttttat | ttgtttttag | gggaggttaa | ggtggaagtg | gtggtagtag | 5820 |
| ggttgggggt | ggttttttag | gagttggggg | tgagttatta | ggagttgggg | gttgggtgga | 5880 |
| tgatttggag | tgtgtagtag | ggaagatgag | gtaaatagggt | aggaagtggg | tggggggagg | 5940 |
| tggaaattggg | gttgtgtttt | gtgtcgtttg | gaattgggag | tgtgggaaag | atattaggaa | 6000 |
| tttgggttga | gcgtagtttt | gttgggtggg | tttggttggt | ttattgtata | gagttttttt | 6060 |
| tgatttttga | agaaagagat | tcgtttgtag | tgggtaaaag | tttgtttgga | ttttttggtt | 6120 |
| attagaaata | tgagtatggt | ggtggttttt | agttttttat | ttatgtttgg | gtttaagaga | 6180 |
| ttgggagttt | aggtttattg | atthttttgag | aaagattaag | atthttgtatt | ttagaaagag | 6240 |
| gtttggggat | ttttgttttg | cgtaagggtta | gaaggattag | ttgttttttt | gagtatthta | 6300 |
| attcggaaatt | tcggttcgaa | gtcagagatag | gagattggat | gcgaggtttt | tttagagttg | 6360 |
| gtttttttta | aataattttt | aaaattttta | gattttaggg | gtacgtcgaa | atthttttaa | 6420 |
| gtagttttaa | gaatataacg | agagttttaa | tattttaggt | ggcggcgcgt | tggttttttg | 6480 |
| gagcggggcg | ggacgcggtc | gcgcggattt | acgtgtataa | tcgcgcggga | cgggggttacg | 6540 |
| cggattttacg | tgtataatcg | cgggattttta | gcgttagcgg | gatttttagcg | ttagcgggat | 6600 |
| tttagcgtta | gcgggatttt | agcgttagcg | ggatttttagc | gtagcggga | tttttagcgtt | 6660 |
| agcgggattt | tagcgttagc | gggtttgtgg | tttagtgagg | cagtgaggac | gttggcagatt | 6720 |
| tgagcggaga | ttgcgttttg | gacgttttag | tttagacgtt | aagttatagt | tcgcgtagta | 6780 |
| gtagtaaagg | ggaaggggtta | ggagtcgggt | atagttggat | tcggaggtcg | tgatttaggg | 6840 |
| gaaagcgtgg | gcggtcagatt | tagggtagtt | gcggcggcga | ggtaggtggg | ttttttgttt | 6900 |
| tttgagtcg | ttttttttta | tattttgttt | cggcgttttt | agtagttttt | atthttggtt | 6960 |
| ttcgcggtta | ttgcgggatt | cggcgttgc | gtagtttag | tggggagtga | attagcgttt | 7020 |
| tttttcgttt | tcggtttttt | cgacggtacg | aggaattttt | gttttgtttt | atagattttc | 7080 |
| ggttttcgtc | gagtcgggtta | ttggagtttg | tttcgttagg | gttttggaat | tagagaaagt | 7140 |
| cgttttttgg | ttatttgaag | cgtcggattt | ttatagtgtt | ttttagtttg | ggcgggagcg | 7200 |
| gcggttgcgt | cgttgaaggt | tgggtttttt | ggtgcgaaag | ggaggtagtt | gtagtttttag | 7260 |
| ttttatttta | gaagcggttt | tcgtatcgtt | gcggtgggcg | ttttcgggtt | tcgatttcgt | 7320 |
| tagcgtcgcg | gggtagaggt | atthggagtt | cgtagggttt | agatttgggt | tggaaaagtt | 7380 |
| tcgttgattg | taggtaaagcg | ttcgggaggg | gcggttaggc | gaagtttcgg | cgttttatta | 7440 |
| tatatthttcg | gtttttatgt | tagttgtatt | cgcggtattg | ggtaggaaat | ggtagggttg | 7500 |
| aggtcagatt | taggagtata | agggagtttt | ttattttttg | tttatatttg | ttatttttag | 7560 |
| ttttgtaatt | tatttttagat | atatagaaaag | taagtaggat | tgggtgggag | acggagttta | 7620 |
| atagcaatat | tattttagtag | tgagtagggt | ttgtatggga | cgcgggagga | gtttagagga | 7680 |
| ggcgcggaga | gtgttcgagg | ttgggtgagt | gttttagaggg | gagatagttg | aatcgggttt | 7740 |
| aagaggtgtt | tagtggtgtg | ttgttgaatg | aatgagtgat | gggttttgaa | gtttgagtg | 7800 |
| attgaaagag | gggtgtgtga | aaaagggttt | tttttattat | ataggatata | gtatatgtaa | 7860 |
| atthtttttt | tgtggaaaag | ttagataggt | taaaaaggtt | ataaataaat | tagtcgggta | 7920 |
| tggtggtgcg | cgtttgtagt | tttagttatt | agggaggttg | agttagggga | atcgtttgaa | 7980 |
| ttcgggaggc | ggagattgta | gtgagttaag | atcgcgttat | tgtatttttag | tttggaaata | 8040 |
| gagcagatt | tcgtttcggg | aaaaaaaaaa | aaaagttata | aatcgtgtgt | gggttttagg | 8100 |
| ttatataaatt | agagttggag | gggagtggtt | aaggatgaga | attgagatgg | atthtttcggt | 8160 |
| ttttttggag | gagagtggtt | ggttgtttat | ttgggggttg | ggaatttttt | tttacgggtt | 8220 |
| tagttgttta | atthtagggg | atthtttagga | taggagttga | tgtaaatagt | cgtttttatt | 8280 |
| tttgtttttt | ttggttttgg | agaaggagga | ggaggttggg | gaggggtttt | atthttttaga | 8340 |
| taattttttaa | gtagtttagga | tatgggtgag | atgagtgaga | tattgatttt | tgggatagaa | 8400 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tttgagaggg | tgttaaaaaa | tttagtaatt | aagataaata | ggtcgggcgt | agtggtttac | 8460 |
| gtttgaatt | ttagtatttt | gggaggtcgg | attattttgag | gttaagagtt | cgagattagt | 8520 |
| ttgggtaaga | tggtgaaatt | ttattttttat | taaaaatata | aaaattagtt | tagtgtgggtg | 8580 |
| gcgtagggtt | gtaatttttag | ttatttagga | ggttgaggtg | agagaattgt | ttgatttagg | 8640 |
| aggtagaggt | tgtagttagt | cgagattatg | ttattgtatt | ttagtttggg | taatagaggg | 8700 |
| agattatttt | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | gaggtcgggc | ggcgggtgggt | 8760 |
| tatattatgt | gatttttagta | ttttgggagg | tcgagggcggg | tggattattt | gaggtttgga | 8820 |
| gttcgagatt | agtttggtta | atatggtgaa | atctcgtttt | tattaaaaat | ataaaaaatta | 8880 |
| gttgggtggg | gtgtaggta | tttgtaattt | gaattatttc | ggaggttgag | gtaggagaat | 8940 |
| ttttgaatt | tcgggggcgg | aggtttagt | gaattaagat | tatattattg | tatttttagtt | 9000 |
| tggataataa | tagtaaaatt | ttgttttaaa | aaaaaaaaaa | atcttttttt | tcgagatata | 9060 |
| gttttatttt | ttcgttttagg | ttgggggtgta | gtattacgat | tttagtttat | tgtaattttt | 9120 |
| gttttttaga | ttttcgtatt | ttagtttttt | aagtagttgg | gattataggt | atctgttatt | 9180 |
| acgttttagt | aattttttgta | tttttagtag | ggcgtgggtt | ttattatggt | ggttagggtg | 9240 |
| gttttgaatt | tttgatttta | agtgatttgt | tcgttttagt | tatttaaagt | gttgggatta | 9300 |
| taggcgtgag | ttattacggt | tgtttttttt | aatgaaaaat | agtgtaaaaa | tttacgataa | 9360 |
| ataaaatatt | aaaaattttat | tgaatttgta | tttttataat | ttttttttat | ttgtttttta | 9420 |
| ggttattttt | tgtttttagaa | agtaatttaa | aaaatgtgta | gatggagttt | ggatttttatt | 9480 |
| tgaaaatggt | gggagttatg | gaaaattttg | gagtagggga | gtgaaggata | gaaattatat | 9540 |
| gtaaaagaaa | ttttgggtcgg | ggcgtagtgg | tttatgtttg | taatttttagt | atcttgggag | 9600 |
| gtcgaggtag | gtggattatt | tgaggttagg | agattgagat | tagtttgatt | aatatgggtga | 9660 |
| aatgttattt | ttattaaaaa | tataaaaaaa | attagttagg | tatggtggtg | tacgtttgta | 9720 |
| gttttagtta | tttcgggaggt | tgagatagga | aaatcgtttg | aattcgggag | gctggaggtt | 9780 |
| tagtgagtta | agattgtggt | attgtatttt | agtttgggta | ataagagtaa | aattttattt | 9840 |
| taaaaaaaaa | gaaagaaaga | aattttttgg | tagttgatga | gaaggaaatt | taatcggtag | 9900 |
| gtttttagtag | gggagatgag | gagattttag | ggagggattt | tgtatatggt | gtgttttagt | 9960 |
| gtgggttagg | gagtaggta | ttattttttt | cgtttatttt | ttttttgttt | taattttttt | 10020 |
| aagttttgga | ttagtggtat | tttaagtgta | gtttaaggaa | ttatatgtat | taggattttt | 10080 |
| agggggtggt | tgtaaaaaat | gtaaattttg | gttaggtgta | gtggtttata | tttgtaattt | 10140 |
| tagtattttg | ggaggtcggg | gcgggtggat | tacgaggtta | ggagatcgag | attattttgg | 10200 |
| taaatcacgt | gaaattttat | ttttattaaa | aaaataaaaa | taaattaaaa | aaaatattag | 10260 |
| ttgggcgtgg | tggcgggcgt | ttgtagtttt | agttattcgg | gaggttgagg | taggagaatg | 10320 |
| gcgtgaattc | gggaggcggg | gttttagtg | agttgagatt | gcgttattgt | atcttagttt | 10380 |
| gggcgataga | gcgagatttt | gttttaaaaa | aaaaaaagta | aattttttgg | gtattatttt | 10440 |
| atattgattg | tatgagaatt | tttggaagta | tggttttgga | atctgtattt | tttaacgatt | 10500 |
| ttttgaggtg | aaatttatat | atttaaaaat | tatatattgt | aagaatataa | tttgatgatt | 10560 |
| tttagtaaat | gtagagttgt | gtaattatta | ttataaaaaa | atcttattttg | aggatttttt | 10620 |
| gttattttta | gaagttttta | ttttagttaa | ttttttattt | ttattttaag | tttttaggta | 10680 |
| ttattgattt | gttttttggt | tataaatttg | tatattttgg | atatgtttta | taaatagaat | 10740 |
| atataaatat | gtagattttt | gtattaagcg | ttttttgttt | agtataaat | ttttgaggtt | 10800 |
| cgtttttaaat | ttatgttaa | gtaaaagaaa | tagtaataaa | aagaaataaa | ttattattag | 10860 |
| tagtttggtt | ttttttattg | ttttatagta | tgttattgta | cgggtgtaat | atattttggt | 10920 |
| tatttattgg | ttagttgacg | ggtatttggt | ttgtttttat | tttttagttg | ttttaaataa | 10980 |
| tgttggttag | aatatttata | tatatgtttt | attatagaaa | tatattttta | gcgtttttga | 11040 |
| gtaaattttt | aggagtgtaa | ttgttggtt | atatggtcgg | tttaggttta | atctttaaga | 11100 |
| aattgttaag | ttatttttta | aagtgttgt | attattttgt | atctttatta | gtaatgtgtg | 11160 |
| agggttttat | ttttttatat | ttataattta | gaattgtttt | tttttttttt | tttttttttt | 11220 |
| gagatggagt | tttatttttg | ttgttttaggt | tggagtgtaa | tgaggtgatt | tcgggtttatc | 11280 |
| gtaatttttg | ttttttgggt | tttaagtatt | tttttgtttt | agttttttga | gtagttggga | 11340 |
| ttataggtgc | gtgatattac | gttcggttaa | ttttgttttt | ttagtagaga | cggggttttt | 11400 |
| ttatttttagt | tagtttggtt | ttgaattttc | gatttttaggt | gatttattcg | ttttggtttt | 11460 |
| ttaaaatggt | gggtttatag | gtatgagtta | ttatgtttgg | ttttttttat | ttattttattg | 11520 |
| ttttgagatg | gattgttttt | ttgttattta | ggttgaggtg | tagtggtata | atctcggttt | 11580 |
| attgtaattt | ttgttttttg | ggtttaagta | atctttttgt | tttagttttt | ttagtagtta | 11640 |
| ggattatagg | tacgtattat | tatatttagt | taattttttg | atcttttagt | gagacggggt | 11700 |
| tttattacgt | tggataggtt | ggttggtttc | gaattttttg | ttttatgatt | tatttatttt | 11760 |
| agttttttaa | attgttggga | ttataggtat | gagttatcgt | gtttgggttt | tgtttttttt | 11820 |
| attataatta | tttaagtggt | tttttagtata | gttttaattt | gtattttttt | aatgattaat | 11880 |
| aatgttaaga | atcttttttat | gtatttattt | aattattttat | tttttatttt | tagtgaaatg | 11940 |
| tttatttata | tattttggatt | atctttttatt | taggttggtta | ttttattggt | gagttataat | 12000 |
| agttttttat | ataatatatt | ttgaatatag | ttttttatta | gatatatgat | ttgaaaaat | 12060 |
| tttttttttag | ttgtgggtt | atcgtttttt | tttttttttt | taaaattttta | atctttttttt | 12120 |
| tgtatttttta | aaaattattt | tttttttttt | tttttatagaa | atagggtttc | gttatgttgt | 12180 |

| | | | | | | |
|-------------|------------|------------|-------------|------------|-------------|-------|
| ttaggttggg | tttaaatttt | tggtttttag | tgattttttt | gttttggttt | tttaaagtgt | 12240 |
| taggggtata | ggagtgagta | attgtttttt | ttttttttt | taattaaaa | attttttaga | 12300 |
| gataggtttt | attaagttgt | ttaagttgga | gtttggtggt | tatttatagg | taagattata | 12360 |
| gtatattgta | ttttcgaatt | tttgaattta | agtgattttt | ttgtttgagt | ttttcgagta | 12420 |
| attgggatta | tatttttatg | ttattatatt | atgttatttt | tttaataata | ttttttgagg | 12480 |
| tgtaaaagtt | ttagattttg | atgaagttta | gtttattatt | attttttttt | tttttttttt | 12540 |
| ttttttttt | ttgagattag | ttttattttg | tcgtttaggt | tggagtgtag | tggcgttatt | 12600 |
| ttagtttatt | gtaattttcg | tttttcgggt | ttaagcggat | tttttggttt | agttttttta | 12660 |
| gtagttggga | ttataggtga | gtgttattat | gttgggttaa | tttttgatt | tttagtagag | 12720 |
| atagttttat | tatgttggtt | aagttgtttt | taaatttttg | attttagtgg | atttgttcgt | 12780 |
| tttagttttt | taaaagtgtg | ggattttagg | agttagttat | tatatttggg | ttattatttt | 12840 |
| tttttatggg | ttttgttttt | ggtgttgtat | ttaggaattt | tttgtttaat | ttgaggttac | 12900 |
| gaagattttt | tttttaagtt | ttttgttgtt | gttgttgttg | ttatttgtat | tttttagtaga | 12960 |
| gatggggttt | ttttatgttg | tttaggttgg | ttttaaattt | ttgattttta | atgatttatt | 13020 |
| cgttttggtt | tttaaagtgg | taggattata | ggtgtaagtt | aggacgtttg | gttttttatg | 13080 |
| ttttttttta | gaaaaaattt | gtaattttta | taaggttttt | tggagagttt | gttgtatatt | 13140 |
| tgtttagggt | tagaagttag | attatattta | ttatttgcgg | atagagtgtt | ttgtttgata | 13200 |
| ggttgaataa | tcgggagttt | ttgtgtttgt | aggagtgtgt | gtgtgtgtgt | gtgtgtgtgt | 13260 |
| gtgtgtgtgt | gtgaatgggt | tttgtatatg | tgattagtag | gatattatta | gattatttta | 13320 |
| gttttaggag | ttattatagg | gaattttgtt | ttgtgtgtgt | gggagataaa | tttagtgtt | 13380 |
| tttgttttgt | tttgtgtgtt | agttaaattt | ttatttttat | tattgttttt | ttgtgatcga | 13440 |
| tttgtttttt | taattatgga | ttttattttt | tttgttattt | ttaaaggaag | aagttttgag | 13500 |
| aaaaaaaggg | ataggtgggg | gaagtagtaa | ggttagttat | ttatataat | aaatataatga | 13560 |
| gtaatgttat | ttttgttttt | atatttagta | gtatttatgt | ttgtataaat | ttggtattaa | 13620 |
| taggagaaa | tatttgtaaa | tatttatatt | taggttatta | tattttttta | tatttttagtt | 13680 |
| tatttttatt | gttgtattta | ttataagggg | atatagtttt | tgtatatagt | tatttttttt | 13740 |
| tatttttttt | gagatggagt | ttcgttttgt | cgtttaggtt | ggagtgtagt | ggtgtaatta | 13800 |
| tagtttattg | tagttttaaa | ttgggtttag | ttattttgtg | gttttagttt | ttcgagtagt | 13860 |
| tgggattatg | ggtatatggt | attatgtttg | gttaatttaa | aaaaattttt | ttttgtagag | 13920 |
| atgaggtttt | attatggtgt | ttaggttggg | tttaaatttt | tgggtttaag | tgattttttt | 13980 |
| gttttagttt | tttaaagtgt | tgggattata | ggtgtgagtt | attatgttta | gttagtttgg | 14040 |
| gttttaggtt | ttaatagtag | ttggatggag | aaggttgttt | attttttttt | tttaattttt | 14100 |
| atttttgggt | aggggtgatg | ggaggaaga | tagttatttt | gttttttttt | ttatttttaa | 14160 |
| tttattaatt | ttttgttttt | gtttttgttt | ttgtttttga | gatagggttt | tattgtgtta | 14220 |
| tttaggttgg | tatggttagt | ttattgtaat | ttttgttttt | tgggtttaag | taattttttt | 14280 |
| attttagttt | ttagagtagt | tgggattgta | ggtatgtgta | attacgtcgg | gttaattttt | 14340 |
| gtattttttg | tagagataag | gttttattgt | attgtttagg | ttggttttga | atttttgggt | 14400 |
| ttaaattttt | tatttcggtt | ttttaaagtg | ttgggattat | aggtgtgagt | tatcgtgttt | 14460 |
| ggttttattt | gtagttttat | ttgtttgaag | ttaagttcga | ggtataaatt | ttttatata | 14520 |
| atagtatagt | ttagttcggg | tttttagagt | ttttgtatta | ttagtttatt | ttatatttat | 14580 |
| ttatggttat | ttgtgtaag | atttgtttt | ttatatattt | atgtttggtg | gggaagatat | 14640 |
| ttatggaaat | agattaggaa | ggttagtgc | ggtaaagttg | gggatggggg | tggaggggag | 14700 |
| ttggggggta | ggggagagaa | gttttgtata | gggttgggtg | tttaaattta | gtagagtgga | 14760 |
| gtaggatttt | tgagaatcga | aggagtattg | tggggtttaa | agagtttata | ttttgggagt | 14820 |
| tattatattt | taggaggtat | tttatatttg | ttttatttgg | tttgtgttgg | gtttttaaat | 14880 |
| ttttgaatta | ttgttaatgt | ttaaaatttt | agagatttta | tagaaaaatt | tagaatttta | 14940 |
| gttttaaaaa | ttagtaaaat | aaaataaaaa | tataaataaa | ttaggagttt | tgtgttttgt | 15000 |
| ttggcgggat | tttgtataag | gtgttttttt | agtttgttaa | ttttttttat | agtttattta | 15060 |
| ttatgttatt | tgttgagttt | tttttttttt | gaagtttggg | gttttgttgt | ttattaagtg | 15120 |
| gttattatag | ttacgagtgg | tttttgagta | tttaaaatgt | gattagtttg | aattgagatg | 15180 |
| cgtttaagta | taaaatgttt | attggatttt | gaagtttaat | tttaaaata | gtaaaatatt | 15240 |
| ttattaataa | tttttaaaat | attgattata | tttgaaatg | attttatttt | taatatattg | 15300 |
| ggttaataaa | aatatattaa | atagtaatat | attatttatt | attatattga | ttaaaattaa | 15360 |
| ttttatattg | ttttttttat | tgtttaaaat | gtgggttagg | tttgggtggt | tattttttta | 15420 |
| attttagtat | tttgggaggt | taaggtaggt | aggtggatag | tttgagttta | ggagttaaag | 15480 |
| attagtttgg | gtaatatggt | gaaattttat | ttttatgaaa | tataaaaaa | ttatttaggt | 15540 |
| atggtggtac | gtatttgtag | ttttagttaa | ttgggaggtt | gaggtgggag | gattgtttga | 15600 |
| gtttgggagg | tgggtggtgt | agtgagttgt | gattgtatta | ttgtatttta | gtttgggtaa | 15660 |
| tagagtaaga | ttttgtttta | aaaaaaaaa | tgggtggatat | taggaaatgt | aaatgtatat | 15720 |
| atgtggttta | tattatattt | tttttggtta | gtttgggttt | agagtttaat | ggataaatgg | 15780 |
| gatttttagta | agagggtag | gtagagttag | aatataaggg | aggggtgaag | aggttattgt | 15840 |
| aatatgatgg | ttaaagtgtg | gaattttgga | gttagtgttt | agagttgaag | gtttattttt | 15900 |
| gttttttatt | tgttgtttta | ttttaatttt | tttgaatttt | tagtattttg | tattttattga | 15960 |

| | | | | | | |
|---------------|---------------|---------------|---------------|-----------------|-----------------|-------|
| taggtgaaat | tatagttggt | at ttattgat | aggtgatatt | at tttgttta | gt tttggagt | 16020 |
| tgggatttga | tt ttttagga | attgatcggg | tt tttgatgag | tt ttttaggtt | ag tttttggt | 16080 |
| tgtagaatga | tgtattttaga | tggaaagatt | tt ttttgatt | ggggtagggt | gattagagta | 16140 |
| aggtttgtgg | agaaat tttat | at ggggaagg | aat tttgtgtt | tgtgtgatg | tagaggagt | 16200 |
| gttatatata | aggtatataa | agaaat gaaa | gt cgtaggat | gggtagggt | gggagaaata | 16260 |
| aaaaaagaaa | gtttaggat | tt agt gatgg | at gggatgtg | gt gatggatg | gggtgtggtg | 16320 |
| atggatggga | tgtgttgggg | gagagggtt | tagagggtat | tatagaggtt | tttggttggg | 16380 |
| ttgttgggtg | gaggtgggga | gg tttat tta | gagagggaga | gattagacgt | tttgttaggt | 16440 |
| tacggattag | gttagt tttt | tttatt tttt | tatt tttttag | tt tttggtatt | at ttttggta | 16500 |
| tgaaagtgtt | tagttaatat | tg ttaagtgg | ttggaaggtt | aat agtatcg | tt tttgagat | 16560 |
| tgtagttttg | gg tttgttat | cg ttttttcg | tttattagtt | gtgtgatttt | aggatatta | 16620 |
| tt tttat tttt | ttggatatta | gt ttttttat | ttgtaaaatg | ggaatagtga | taagtgtagt | 16680 |
| tt tttgtggg | ttgttttaag | gattgagtgg | gt tttatatac | gggagtatat | agt agtgagt | 16740 |
| gtttagtagt | tg ttagt tttt | agatt tttata | ttcagatttt | at tttttgatg | tatatacgag | 16800 |
| gt tttgttta | ttaaggttgt | tatt tttttta | aaggaagaaa | aggttaaagg | gt ttcgttaa | 16860 |
| agaagtaagt | gt ttcgatat | tt tttttttt | tttgaggtta | agt aggagga | tggagggttag | 16920 |
| attgtaggaa | gaat tttttt | tt tttttatag | ttaaagaatgt | ttaaagtgaat | tttagaaaat | 16980 |
| taat tttttgg | tcgggtacgg | tgg ttttatgt | ttgtaatttt | agtatt tttgg | gaggtcgagg | 17040 |
| cgggtggatt | atgaggttag | gag tttgaga | ttag ttttgg | taatatgggt | aaat ttttatt | 17100 |
| tttattaaaa | atatagaaat | tagttaggcg | tggtgggtgg | tattttgagt | tttagttatt | 17160 |
| taggaggttg | aggtagaaga | atcgtttgaa | ttcggaggtt | agaggttgcg | gtgagtcgag | 17220 |
| atggtattat | tgtatt ttttag | tttgggtgata | gagcagagatt | ttatt tttaaa | aaaaaaaaaa | 17280 |
| aaaaaaaaagaa | agaaagaaag | aaaattaatt | tt ttttagattt | tt ttttttgg | tt ttttagataa | 17340 |
| ttagtagata | attagggagt | gt tttggggag | ggggtatagg | gt tttat tttt | tgg tttttttt | 17400 |
| tagtattcgt | agagt gattt | ggt aggaaga | gagggttagg | gttgaaggta | ttaatagggt | 17460 |
| aataat tttt | ttgt tttttag | gatgtggagt | ttatag tttt | taggt tttttt | atc gatgtag | 17520 |
| gggtgaggtg | gggtgggagg | gt agagagga | aaaacgtggg | aattacgaat | gggaagaggg | 17580 |
| gatttgggga | tt tttttgtga | ttgt ttttagtt | ttgg ttttgat | tttgtt gtag | gg ttttaggg | 17640 |
| gt tttgt tttt | tag ttttata | ttatatt tttt | atattttaggt | tt tttttttga | tgg tattcgt | 17700 |
| taaaatatta | agatt ttcgtt | ggag tttgtt | gag ttttaaga | taggattagt | ttgg ttttagt | 17760 |
| gagggggata | tagatt ttaag | agttagataa | ttcggag tttt | ag ttttttgag | tttatattta | 17820 |
| ttagttgagg | gaat tttgggc | gagttatt ttt | tatt tttattg | tgg ttttagtt | tt tttttttta | 17880 |
| taaagtaagg | atattagatt | tt ttttttgag | tggggttgtg | tgcggtagtt | attgg ttttag | 17940 |
| cg tttttttt | tt tttttatag | tg tttat tttt | tt tttaaagtt | agt agt ttttt | tag tttttttt | 18000 |
| tttgg ttaaa | gatag ttttgg | tgggcggtat | tataaatt ttt | ggtgaggttt | tttgt ttcgt | 18060 |
| gttcggtcgg | tt tttattgt | ttag ttttaac | gt ttaagtat | aaggtgacgg | tgggggaggt | 18120 |
| gtagcggcga | tt ttcg tttt | tcgag tgttt | taacgt tttt | tt tttggggg | gtgt tttttcg | 18180 |
| taggtaggaa | gg ttag tttta | taat tttttcg | ttt gattgga | tt ttttgg ttt | tt ttttaggtt | 18240 |
| tt ttaaggta | ttagagagga | gg ttag ttttt | at ttttaggtt | tt tttgt ttttag | tt ttttttggg | 18300 |
| aggggaggtt | tcggggattt | tggattgtgt | atgt tttggg | ggtcgg tttta | gg ttttttgg | 18360 |
| agttgttata | tgtttgtggg | tttgggtag | ggaggtgggt | agagaaagaa | g tttgggtat | 18420 |
| agagaaggga | tggagaattt | at tttatattg | gttaggacga | gattgggttt | ggggaagaga | 18480 |
| gagggttatt | tttatt ttttt | tt ttttaggaa | atgaatt tttta | ggattc gagg | tttatt ttttgc | 18540 |
| gcggtagtat | agg tttttgtg | tg tttttgggg | tgtggagtat | ggtag ttttat | gg tttat tttt | 18600 |
| tttagatttg | tg tttttttt | tt ttttgtgta | tg tttttgtg | tttgggaagg | gtatt tttttgt | 18660 |
| tttaggg tta | tgag ttttga | tgag ttttag | gatggtgtgg | atatgtgtgt | gcgtgtgtat | 18720 |
| gtgtttatag | tagcggatgg | ttaggagtat | gtgt tttttat | gagtggtgagt | at ttttgtgtt | 18780 |
| tt ttttagtg | ggatgtgtgg | tttaggggtg | gt ttttgtttt | gtgtgtgtgt | gtgtgtgtgt | 18840 |
| gtgtgtgtgt | gtgtgtgtgt | gtgtgtgatt | gt tttgtgtat | ttggag ttttag | ttat atggat | 18900 |
| gtttatata | ttgtgttatt | gtgtgggagt | ggatgtgtg | tagttatagg | gg tttat tttt | 18960 |
| ttat ttttttag | cg tttttggag | gatt tttttgt | agt ttttttag | tgtttgtata | tgtgt tttttg | 19020 |
| ggat ttttgat | ttgggtgtgat | tttgtgtttt | tgtgt tttttg | gggtgtttgag | tttgtgt tttt | 19080 |
| tgtatgtttg | tt tttattgg | tt tttatgtgt | tt tttat tttt | ttacgtgtgt | gtgtgtgtgt | 19140 |
| gtgtgagaga | gttatttaggt | gg tttgtgtgt | gtatt tttgtgt | tt tttttatgt | gt ttttttttt | 19200 |
| ttgtatgtta | ttgcgtgttt | ttatataatag | gt tttgtgtat | tt ttttggtag | gtgtgtgtgg | 19260 |
| tt ttttgggtg | tg ttaagg ttt | tttaggatgt | tttatt ttttgg | ttgt tttttgt | tgttgagagg | 19320 |
| attagagtga | attgaggggt | tt tttggagt | gggtgttgg | tt tttttgttt | ttgt ttttgg | 19380 |
| tgtaaaatta | tggagatttag | ttcgt ttttt | gtt gaaagag | at ttttagagt | at ttttttagga | 19440 |
| ttaat ttttaa | at ttttaatag | gggttgggtg | aggtaataag | gt ttttttgg | ag tttat ttttag | 19500 |
| tt ttttatgcg | tt tttttttgt | tt tttatagtt | gtatgt tttgt | ttgt tttttgt | gtatgt taaag | 19560 |
| gg ttaag ttt | aaaaatgggg | gtcgggtgtt | gcgggaacgg | ttagagaaga | ttgggt ttaa | 19620 |
| ttgtttagtt | ggtcgtcgt | aggtcgttaa | tgtgacgttt | ttgatt ttcgt | tagt ggaagg | 19680 |
| tgagtgaggt | ttgaagggtg | gtatgggagt | ggatgtgagg | gt aagt gagt | tgt tttggtgt | 19740 |

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|-------|
| gattgttttt | aatgtaggag | ggtgtgttta | gtggtgtgtg | tgtattcgtg | taagtgttgt | 19800 |
| agtatttgtg | tcgctatttt | ttgttatgta | taaggtgtgt | ttgtgatttt | tgtgtgtagt | 19860 |
| ttttgtatat | gtgtgtatgt | atgtgtaatg | gagatttatt | tgtagtgtgc | ggatgtgtat | 19920 |
| gtgggtgtgg | gtgtgggtgt | gtgggtgtgt | gtgttggcgg | ggagtggta | gtgtttggat | 19980 |
| gtggtaaagg | gatgaagggg | gtgatagttt | ttttttttgg | aggttgggtg | ttttggaata | 20040 |
| tatggtgggt | ttagttttta | tttttgagtt | tagagtttaa | ttttaatggt | tttaaggtgg | 20100 |
| ggaatggaa | tgggggttgg | agttgggttg | ggaaggaata | ttagaggggg | tatttatata | 20160 |
| ggtaggatgg | ggtaggatat | ttttgattag | gggttagag | gtttaaagtt | gtgttttagg | 20220 |
| cgtagaggta | tattttgtag | tagtgatagt | tttttttttt | ttaggagagg | tcgtgtattt | 20280 |
| ggttcgagat | ttcggttacg | tttgtgagac | ggagttttta | gttaaggtag | ttgtcgagta | 20340 |
| tttgtgtcga | tagtacgttg | attcggggga | gttgatagt | cgtaagagta | tgttggtggg | 20400 |
| tgtaagtga | gtgaggggat | tttgtatagg | tatacgtggg | tgttatgtat | agatagatat | 20460 |
| tatgtatggg | tataatggat | attattgtgt | atatgagtag | tgggtatata | tcgctatttg | 20520 |
| cgtattattg | tgtatacgag | tagtgggtat | atatatatac | gtcgtatta | ttgtgtatat | 20580 |
| tagtagtgag | tatatacgta | tgtgtgcggt | attgtgtacg | tgagtgtgg | gtatatatac | 20640 |
| gtacgtgcmc | gttattgtgt | atacgagtag | tgagtatata | tacgtatttg | cggtttatgt | 20700 |
| ggtattatat | atgggtatat | ttaggttttt | tgtttaggtt | tttttgggta | tatgtgtgta | 20760 |
| tgttatttgg | gtatatttgg | gtgttatata | tgggtatatt | tgggtattat | gtttattatt | 20820 |
| tagatatata | tgaatatttt | ttataagttt | ttggagtata | tatagatatt | atatgtgaat | 20880 |
| atatttgggt | ttttgtatta | ggaatagtga | tgttatata | tggtttatat | aggtttacgt | 20940 |
| atgggtttta | tatttatata | tatttttagg | atattattag | ttttattttt | tatcgtaggt | 21000 |
| ttgaggtttt | tttttttagt | tttgtttatt | aattttttgt | tttagtgacg | tttattaggt | 21060 |
| tattaagtta | tatttgattg | ttttgtaaat | tttattgtgt | tttttttatt | atcgtttata | 21120 |
| tgttgttatt | tttatttagg | tttttatttt | atttttggat | ttttatttta | gagcgttttt | 21180 |
| gtttttaaatt | tttttgcacg | tgtattcgtt | ttattgtata | agaattttta | atggttttcg | 21240 |
| ttgtattgat | ttttaaatta | ggtttgcgag | ggtattttag | ggtttgttta | ataaattagt | 21300 |
| attttattaa | gatttagttt | tgtttttgtt | ttttgagatg | gagtttcggt | tttgtttttt | 21360 |
| agattggagt | gtagtggcgt | aatttttagtt | tattgtaatt | tttatttttt | gggtttaagt | 21420 |
| aattttttttg | tttttagttt | tcgagtagtt | gggattatag | gcgtttgtta | ttaggtttag | 21480 |
| ttaattttttg | gtatttttaag | tagagatggg | gttttattat | gttggttagg | ttggtttcga | 21540 |
| atttttgggt | ttaggtgatt | tattcgtttt | ggatttttaa | agttttggaa | ttataggtgt | 21600 |
| gaattattat | tttcggtttt | aagatttagt | tttttaggtt | aggtaagggtg | gtttatattt | 21660 |
| gtaatttttag | tatttttgag | aggttgaggt | tggtggatcg | tttgagttta | ggaattggag | 21720 |
| attagtttgg | gtaatatggt | gaaattcggg | tttttagtaa | aatataatag | ttaggtgtgg | 21780 |
| tggtgtattt | ttgtagtttt | agttatttgg | gaggttgagg | tgggaggata | atttgagttt | 21840 |
| gggaggttaa | ggtttagta | agttatgatt | gtatcgttgt | attttagttt | gcgtgatata | 21900 |
| gtgagatttt | attttaaaaa | aaaaaaaaagg | cgaaagattt | agttttttta | aaagagtttt | 21960 |
| gtttttagtt | tatttgggtt | gttataataa | aatattataa | attgggtagt | ttataaataa | 22020 |
| tagaaatttt | tatttattta | tttattgaga | tagaattttg | ttttgttgtt | taggtcggag | 22080 |
| tgtagtgggt | taatttttagt | gtattgtaat | tttttagttt | tgggtttaag | taattttttt | 22140 |
| gttttagttt | tttagtagt | tgggattata | ggtgttgtt | attatatttg | gttaattttt | 22200 |
| gtatttttag | taaagttttt | attatgttgg | ttaggttggg | tttaaatttt | cgattttata | 22260 |
| tgatttattt | gttttgggtt | ttttaagtgt | ttggattata | ggtgtgagtt | attacgttta | 22320 |
| gttaataaga | gatatttatt | ttttatagtg | ttggaggttg | ggaagtttaa | ggttaaggta | 22380 |
| gatttcggtg | taggtgaggg | tttgtttttt | ggtttataga | tagtattttg | ttgtgtgttt | 22440 |
| ttatattgta | gaagggagaa | gggttatttt | tttggttttt | tttataataa | gggtattaat | 22500 |
| tttatttatg | acgtttttat | ttttatgatt | taattatttt | ttaaagtttt | tattttttta | 22560 |
| tattattatt | tagtggggta | gaattttaat | atacgaattt | tggggaatat | aagtatttat | 22620 |
| attatagtag | ttttattgta | agaaaagata | gagagaggag | ggggatagag | aaagagagaa | 22680 |
| agaaagagag | agagattgag | tgtagtgggt | tatgtttgta | attttagtat | tttgggaagt | 22740 |
| tgaggtgatt | agatcgtttg | agtttaggag | tttaaggtta | gtttgggtaa | tttggcgaaa | 22800 |
| ttttatttgt | attaaaaaaa | aattagttag | gagtgtaggt | gcgtatttgt | agttttagtt | 22860 |
| attcgggagg | ttgaggttga | gattgtagtg | agttgggatt | atattattgt | attttagttt | 22920 |
| gggtgataga | gtaagatttt | gtttcggaaa | aaaaagaaaa | gagagagata | aatttgagaa | 22980 |
| ttattagttg | gtaggttaaa | gtttaaattt | ttttgagggg | tttatagttt | ggtttatata | 23040 |
| aattttattt | atttttgtta | attttgtttt | tatatttttg | ggtttgtttt | cggtttagaa | 23100 |
| tatttagttt | ttttatttaa | tatagttttt | taaattgata | attttatttt | tttgtttttg | 23160 |
| aggttttttt | taagtgttta | aggtagaaga | ggaagtatag | atttttggga | ttgttaaagt | 23220 |
| gagagatttt | tgaggttgtt | tatattgtgg | ttggagaaat | ttagtttagga | cggggtttat | 23280 |
| ttcgtttatt | tattagttta | tttattaagt | attttgtggg | tttttagttat | gtgtaagata | 23340 |
| ttggttgggt | tttatgggtg | tataaagatg | aatgatgagt | gatttttttg | tttgagttat | 23400 |
| atataggacg | tataaaggaa | tgaataatga | atggttgtaa | attagagttta | ggtttaaaatt | 23460 |
| ttagttcggg | tatttttgata | ggttatttaa | tttgttcgaa | attagttttt | ttgttcgtat | 23520 |

ES 2 597 845 T3

| | | | | | | |
|--------------|--------------|--------------|-------------|-------------|--------------|-------|
| gatggggata | ataatagtat | agat ttgtta | tgaagattaa | gatgatgtag | atagtgtggt | 23580 |
| tggatatagta | tttagt tttt | aattata tttt | attgttagtt | attat tttc | gtagat tttg | 23640 |
| gattgt tttt | tgatag tttt | aaagataaag | ttttgtgaag | tattgagtgt | ttgggg tggg | 23700 |
| agaggg aatt | taggaaagtt | ttttatataa | ggtagtaagt | agttgagtgt | tat tttgaat | 23760 |
| gaatatgggt | tagatattta | aaaaagaaat | attatagggt | gatcgatatg | tgttgt tttat | 23820 |
| tagttgtgtg | atattgggtt | agttaaatga | tttttttgaa | ttttag tttt | tttttttgta | 23880 |
| aaaaatagta | tggttattat | ttgttagattg | ttgggagggt | tagtaattgt | gtagttaatg | 23940 |
| tttgtaaaat | gttttagtatg | tagtagg ttt | ttttttgggt | ttagat tttt | gttttttaat | 24000 |
| ttttgat ttt | ttttttcgta | ttaggtagat | ttgtaaggag | tttgtagatt | tgatgg tttta | 24060 |
| ggatcg tttta | tcg tttgggta | atagtcg ttt | agtatttatt | ttggag ttc | gagtatagag | 24120 |
| ttg tttgata | tat ttttagtt | ttattattta | tgg tttcggt | ggg tttgtta | tttgtgtgtg | 24180 |
| ttttattgtt | ttttagaatt | atttgttgga | gttatttaag | gggttg gata | agatgt tttt | 24240 |
| aagtagtgtg | ggtagtgggt | atggtgaaat | taagg tttc | gagaaggatg | ttaagtatcg | 24300 |
| gaaataattg | ttttttttat | tttattttta | aggggt tttt | aggt tttgaa | atagggattt | 24360 |
| agtttttggg | ggtggg tttg | gaaggattga | aaggtgggat | tagagttagg | ttagaaagag | 24420 |
| aatatttatt | tagagat ttt | agagttgggg | atttgg tttg | gagtaaggga | gggtgg tttt | 24480 |
| tttgtggtgg | tgtgttg gta | agttaagggt | ttaggtattt | gttttatgtg | tgtaat tttt | 24540 |
| tgat ttttga | tgg ttgagaa | gggtttggat | agaaaattga | tatgaaaaga | tttgg tttat | 24600 |
| ggggtagagt | tttttttatt | agcgtgg tttg | ggtggtcgtg | ggtgt tttta | gaggttaaag | 24660 |
| tttttgtgtt | ttttattggt | ggtaggagga | aaattgataa | attagagggtg | ttattgagga | 24720 |
| gttggtgttt | tttatttttag | aattttttat | ttttagaaaa | ggggtgttgg | aaggagg ttt | 24780 |
| tagtggattt | tttgtat ttt | ttttat tttt | agagagaagt | gggtaggagg | ggttttt aag | 24840 |
| gaataaagaa | gataaagtat | aaatcgtaga | atttgaattt | aggttgttgt | ttattatagt | 24900 |
| tttgttgttg | tttgtttttat | ttatttatgt | atgttgtaat | taaatttgaa | at ttttaaaa | 24959 |

<210> 4

<211> 24959

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 4

ES 2 597 845 T3

| | | | | | | |
|-------------|------------|------------|------------|-------------|------------|------|
| ttttaaatt | ttaaatttaa | ttataatata | tataaataaa | taggatagat | agtaatagga | 60 |
| ttatgatgag | tagtagtttg | gatttaagtt | ttgcgatttg | tgttttattt | tttttgtttt | 120 |
| ttgaggattt | tttttgttta | tttttttttg | agagtaggaa | gggtataaag | agtttattgg | 180 |
| ggtttttttt | tagtattttt | tttttggggg | taggagattt | tggaatgagg | ggtattaatt | 240 |
| tttttagtagt | atttttgatt | tattaatttt | tttttgttta | ttagtgaaaa | atataaaggt | 300 |
| tttggttttt | agaagtattt | acggttattt | agttacgtta | atggaaaggg | ttttgtttta | 360 |
| tgagttagat | ttttttatgt | taattttttg | tttaaatttt | ttttaattat | taaaggttag | 420 |
| aaaattgtat | atatgagata | aatatttggg | tttttaattt | attaatatat | tattatagag | 480 |
| aggttatttt | tttttatttt | aagttagatt | tttaattttg | ggattttttg | atgaatgttt | 540 |
| tttttttggg | ttgattttaa | ttttattttt | tagttttttt | aggtttattt | ttaagagtta | 600 |
| agtttttatt | ttagggtttg | ggagtttttt | agggatgggg | tgggagaagt | agttattttc | 660 |
| gatgtttggg | attttttttc | gaagttttgg | ttttattatg | tttattgttt | atattgttta | 720 |
| gaaatatttt | gttttagttt | ttgagtgatt | ttagtaaata | gttttggaag | gtagtgaggg | 780 |
| tagtatagat | ggtaggttta | tcgaagttaa | gggtgatgag | gttaaagtgt | gttaagtagt | 840 |
| tttgtatttc | gggttttagg | atgagtgttg | ggcggttggt | gttttagcgg | gagcggtttt | 900 |
| gagttattaa | gtttgtaaat | tttttgtaga | tttgtttggg | gcgagaagga | gggttagagg | 960 |
| ttggaggata | aaagtttgag | tttagagaag | gatttgttat | atgtaggta | ttttatagat | 1020 |
| attagttata | taattgttga | tttttttaat | agtttgtaaa | tgatgattat | attgtttttt | 1080 |
| atagaaaagg | aaattaaggt | ttagagaagt | tatttgatta | gttttagtatt | atatagttaa | 1140 |
| taaatagtat | atgtcgatta | ttttgtaatg | tttttttttt | ggatgtttgg | tttatattta | 1200 |
| tttaaagtgt | agtttagtta | tttattattt | tatataggaa | attttttttg | attttttttt | 1260 |
| ttattttaaa | tatttaatgt | tttatagaat | tttgtttttg | aagttattag | gaaataattt | 1320 |
| agaatttgtc | ggaaataata | attagtagtg | aatataattg | aaggttggat | gttgtgttaa | 1380 |
| ttatattggt | tgtattattt | taatttttat | aataagttta | tattattatt | atttttatta | 1440 |
| tacggataag | aaaattgggt | tcggataggt | tggatgattt | gttaaggtga | tcgagttggg | 1500 |
| atttgaattt | agtttttagt | tgtagttatt | tattatttat | ttttttgtgc | gttttgtgtg | 1560 |
| tggtttaggt | aagagaatta | tttattattt | atttttgtgt | tattatgaaa | gttagttagt | 1620 |
| attttatata | tgattgagat | ttataaagtg | tttgatgaat | gagttgataa | gtgaacgagg | 1680 |

| | | | | | | |
|--------------|-------------|-------------|-------------|-------------|-------------|------|
| tgagtttcgt | tttggttag | tttttttagt | tatagtatga | gtagttttag | aggttttttt | 1740 |
| atttggtagt | tttaggagtt | tgtgtttttt | tttttgtttt | gggtatttga | ggaaagtttt | 1800 |
| aaaggtagga | gggtggagtt | gtagtttga | aggattgtgt | taggtgggag | gattggatgt | 1860 |
| tttaagtcgg | aaataaattt | aaggatgtga | ggatagggtt | ggtagaagt | ggtaggagttt | 1920 |
| gtgtgggtta | gattgtgggg | ttttaaaag | agtttgaatt | ttagtttatt | aattggtagt | 1980 |
| ttttagattt | attttttttt | tttttttttt | tttcgagata | aggttttggt | ttgttattta | 2040 |
| ggttggagtg | tagtgggtgt | attttagttt | attgtagttt | taatttttagt | ttttcgagta | 2100 |
| ggtgggatta | taggtgcgta | ttattatttt | tggtaatttt | tttttttagta | tagatgggggt | 2160 |
| ttcgttaagt | tgtttagggt | ggttttgaat | ttttggggtt | aagcgatttg | attatttttag | 2220 |
| ttttttaaag | tgttgggatt | ataggtataa | gttattgtat | ttagttttttt | tttttttttt | 2280 |
| tttttttttt | tttatttttt | tttttttttt | gttttttttt | gtagtgggat | tgttatgggtg | 2340 |
| tgaatgttta | tgttttttta | aattcgtatg | ttgaaatttt | attttattag | gtgatagtat | 2400 |
| taggaggtga | ggattttggg | agggtattag | gttatggagg | tggagtcggt | atgaatggga | 2460 |
| ttaatgtttt | tgttataaaa | gaagttagag | agtaattttt | tttttttttt | ataatgtgag | 2520 |
| gatatatagt | aagggtgtgt | ttatgagtta | gaaagttagt | ttttatttga | tacggaattt | 2580 |
| gttttgattt | tgaatttttt | agtttttagt | attatgagaa | ataaatgttt | tttgttgggtt | 2640 |
| gggctgggtg | gtttatattt | ataatttaag | tatttaggga | gtttaaggta | ggtggattat | 2700 |
| gtgaggtcgg | gagtttgaga | ttagtttggg | taatattggtg | aaatttttat | taaaaatata | 2760 |
| aaaataaatt | agatgtgggtg | gtaggtattt | gtaatttttag | ttatttagga | ggttgaggta | 2820 |
| ggagaattgt | ttgaatttag | gagttggagg | ttgtagtga | ttgagattgt | attattgtat | 2880 |
| tggggttggg | gtaatagaat | aagattttgt | tttaataaat | aaataaataa | aaatttttgt | 2940 |
| tgtttataag | ttatttagtt | tatgggtattt | tgttatagta | gtttaaatgg | attaagataa | 3000 |
| gaattttttt | ggaagaatta | aatttttcgt | tttttttttt | ttttgagatg | gggttttatt | 3060 |
| atgttacgta | ggtagagta | tagcgatgta | attatggttt | attgtagttt | tgatttttta | 3120 |
| ggtttagggt | atttttttat | tttagttttt | taagtatttg | ggattatagg | agtgtattat | 3180 |
| tatatttaatt | tgttgtattt | ttattagaga | tcgggtttta | ttatgttggt | taggttagtt | 3240 |
| tttaattttt | gggtttaagc | gatttattag | ttttagtttt | ttaaaagtgt | tgggattata | 3300 |
| gatgtgagtt | attttgtttg | gtttgaaaaa | ttaaattttg | gggtcgggag | tggtgggtta | 3360 |
| tatttgtaatt | tttagaattt | tgggagttta | aggcgggtgg | attatttgag | gtaggaggtt | 3420 |
| cgagattagt | ttggttaata | ttgtgaaatt | ttatttttat | ttaaaatatt | aaaaattagt | 3480 |
| tgggtttggg | ggtaggcggt | tgtgatttta | gttattcggg | aggttgagat | aggagaattg | 3540 |
| tttgaattta | ggaggtggag | gttgtagtga | gttgagattg | cgttattgta | ttttagtttg | 3600 |
| ggaaataaga | gcgaaatttt | attttaaaaa | ataaaaaata | aattaaattt | taatagaata | 3660 |
| ttagtttggt | aaatagggtt | tgaggtggtt | tcgtagattt | agtttgaaaa | ttaatgtagc | 3720 |
| gggaattatt | gaagattttt | atgtagtggg | gcgggtgtac | gttaggggag | gtaggaaata | 3780 |
| gggacgtttt | gggtgaagg | tttaggagtg | agatgaagat | ttgagtagag | gtggtagtat | 3840 |
| gtgaacggtg | gtgagggagg | tataatggga | tttgtagaat | agttagggtg | gatttgggtg | 3900 |
| ttgatgggac | gttattggag | tagaggggtg | atgggtaagg | ttagaagagg | gagttttaga | 3960 |
| ttgctgatgg | ggagtagggt | tgggtggtgt | ttgggagtg | atgtgagtg | aggatttatg | 4020 |
| cgtgagtttg | tgtgagttag | tatgtgatag | tattgttttt | gatgtaggag | tttagatatg | 4080 |
| tttatatgta | tggtttatgt | gtgttttagg | agttttgtaa | gagtgtttat | gtatgtttga | 4140 |
| gtgataaata | tgatgtttaa | ataatattat | atgtggtatt | taagtgtggt | taggtgggtat | 4200 |
| gtatatatat | gtttaggaaa | gtttaagtag | ggagtttagg | tgtgtttatg | tgtggtggtta | 4260 |
| tatgggtcgt | aggtacgtgt | gtgtgtttat | tgttcgtgta | tatagtggcg | cgtacgtacg | 4320 |
| tgtgtgtggt | tattgtttac | gtatatagtg | gcgtatatat | acgtgtgtgt | ttattgttga | 4380 |
| tgtatatagt | ggtgcgtacg | tatgtgtgtg | tgtttattgt | tcgtgtatat | agtggtgcgt | 4440 |
| aaatacgtat | gtgtgtttat | tgtttatgta | tatagtgggtg | tttattgtgt | ttatatatga | 4500 |
| tgtttgtttg | tgtatgggtat | ttacgtgtgt | ttgtgtaggg | tgtttttatt | tatttggtag | 4560 |
| ttagtagtat | gtttttgcgg | ttgtgtagtt | ttttcgggtt | agcgtgttgt | cggtataggt | 4620 |
| attcggtagt | tgttttggtt | gggaatttcg | ttttatagac | gtaatcgaag | tttcgggtta | 4680 |
| ggtgtacggt | ttttttggg | gggaggggag | ttgttattat | tgtagggtgt | atttttgcgt | 4740 |
| ttggaatata | gttttgaatt | tttgattttt | tggttagggg | tattttgttt | tattttgttt | 4800 |
| gtgtagatgt | tttttttgat | gttttttttt | agtttagttt | tagtttttat | tttattgttt | 4860 |
| tattttgggg | ttattagggt | tgggttttga | gttttagagat | gggagttggg | tttattatgt | 4920 |
| attttagagg | tattaatttt | taggggaggg | ggttgttatt | ttttttattt | ttttattata | 4980 |
| tttaaatatt | ggttattttt | cgtaaatata | tatatattata | tatttatatt | tatatattata | 5040 |
| tatatattcg | tatattgtaa | gtaggttttt | attgtatatg | tatgtatata | tgtataagga | 5100 |
| ttatatatag | aaattataaa | tatattttat | atatgataga | gaatacgtat | atagatattg | 5160 |
| taatatttat | acggatatat | atatattatt | aaatatattt | ttttgtatta | gagatagttta | 5220 |
| tattagataa | tttatttggt | tttatattta | tttttatgtt | tatttttagg | ttttatttat | 5280 |
| tttttattag | cgaagtttagt | agcgttatat | tggcggtttt | gcgacgggta | gtaggtaggt | 5340 |
| tgggttttaatt | tttttttaatt | cgttttcgtg | aatatcgggt | tttatttttg | gatttgggtt | 5400 |
| ttgatatgta | tagagataga | tagatatata | attatgaaga | taagagaaac | gtatggaagt | 5460 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| taagtgggtg | ttaaaggggtt | ttgttggttt | tattagtttt | tgttgaggtt | tgggattggt | 5520 |
| tttgggggatg | ttttgaggtt | tttttttagta | agggcgtaat | taatttttat | ggtttttag | 5580 |
| ttaagataga | ggtagggggg | ttagtattta | gttttaggag | tttttttagtt | tattttaatt | 5640 |
| tttttaatag | taggggtagt | tagggtgggg | tattttgggg | atttttagtat | attagaggggt | 5700 |
| tatatatatt | tattagggga | tatatagatt | tatatgtaga | gatacgtagt | gatatataga | 5760 |
| aagagagata | tatggggaga | tataggtata | tatatagtta | tttggtgatt | tttttatata | 5820 |
| tatatatata | tatacgtgaa | agaatagaga | tatatagaat | taatagagat | agatatgtag | 5880 |
| agatatagat | ttagatattt | agggatagag | agatatagga | ttatattaag | ttaggggtttt | 5940 |
| aggaatata | gtatagatat | ttagggattt | atagagattt | tttaagacgt | taggagtaag | 6000 |
| gagataaatt | ttatgattgt | atatatgtta | tttttatata | gtggtataga | tatgtaaata | 6060 |
| tttatatgat | tagttttaga | tgtataggtg | attatatata | tatatatata | tatatatata | 6120 |
| tatatatata | tatatatata | gagtaggata | tatttttaggt | tatatatttt | attagaagaa | 6180 |
| atataggata | tttatattta | tggaaatata | tgtttttggt | tattcgttat | tatagatata | 6240 |
| tgtatacgtg | tatatatggt | tatattattt | ttagattttat | tttaagttat | ggttttgaag | 6300 |
| tagaaatatt | tttttttaggt | ataaagatat | gtatagagag | agagagatat | aggtttagga | 6360 |
| agatgaatta | tggattatta | tgttttatat | tttaaaatat | ataggggtta | tgttgtcgcg | 6420 |
| tagaatgggt | ttcgggtttt | ggggtttatt | ttttagaaga | gggggtgagg | atgggttttt | 6480 |
| ttttttttta | gatttagttt | cgttttgatt | agtgtggatg | agtttttttat | tttttttttg | 6540 |
| tgtttaggtt | tttttttttg | tttattttttt | tatttttagat | ttatagatat | gtggtagtta | 6600 |
| ttagggat | gggtcgggtt | ttataatatg | tataattttag | agtttttcggg | gttttttttt | 6660 |
| ttaggagat | tgagtagag | ggtttaggtg | agattggttt | tttttttgat | gttttgagaa | 6720 |
| ggtttgaaga | ggttagggat | tttaattagc | ggggaattgt | gggttggttt | ttttatttgc | 6780 |
| ggaggatatt | tttttaggag | gaggcgttga | ggtattcggg | aggcgagagt | cgtcgttga | 6840 |
| ttttttttat | cgttattttg | tatttgacg | ttgagttgag | tagtgaaagt | cggtcgggta | 6900 |
| cggagtagaa | gattttatta | ggatttgtga | tgtcgtttat | taggttgttt | ttggttaagg | 6960 |
| agagggttga | gaggttgttg | gttttggagg | ggatgggtat | tgtggagaag | agggaaacgt | 7020 |
| tgagttagt | gttgcgtat | atagttttat | tttaaggagg | gtttgggtatt | tttgttttat | 7080 |
| gaaggagaa | attgagttat | agtgaggtga | aagtgattcg | tttaagtttt | tttaattgat | 7140 |
| ggatgtgaat | ttagagattg | aatttcgagt | tgtttggttt | ttgaatttgt | gtttttttta | 7200 |
| ttgaattag | ttggttttat | tttaaattta | aatagtttta | gcggaatttt | ggtgtttag | 7260 |
| cggatgttat | tagagagga | tttggatgta | aggatatgag | tgtgagttgg | aggtaggatt | 7320 |
| tttgggggtt | tgtagtagg | ttaggttag | gttgggtaat | tataggagat | ttttaaattt | 7380 |
| ttttttttta | ttcgtgattt | ttacgttttt | tttttttggt | tttttatttt | attttatttt | 7440 |
| tgtatcggtg | aggggtttg | agattgtggg | ttttatattt | tggagatagg | gaggttattg | 7500 |
| ttttgttaat | gttttttagt | ttggtttttt | ttttttgtta | agttattttg | cgaatgttaa | 7560 |
| aggagattag | gagataggt | ttatattttt | tttttaaata | ttttttaatt | atttattaat | 7620 |
| tgtttgaaga | ttagagaagg | aatttgagaa | gttaattttt | tttttttttt | tttttttttt | 7680 |
| tttttttttt | ttgagatgga | gtttcgtttt | gttattaggt | tggagtgtag | tggattatt | 7740 |
| tcggtttatc | gtaatttttg | tttttcgggt | tttaagcgatt | tttttggttt | agttttttga | 7800 |
| gtagttggga | ttataggtgt | ttattattac | gtttggttaa | tttttatatt | tttagtagag | 7860 |
| atggggtttt | attatattgg | ttaggttgg | tttaaaattt | tgattttatg | atttattcgt | 7920 |
| ttcgggtttt | taaagtgttg | ggattatagg | tatgagttat | cgtgttcggg | tagaagttaa | 7980 |
| tttttaaaaa | tttatttaga | tatttttgg | tataagagaa | agaaagtttt | ttttgtagtt | 8040 |
| taatttttat | ttttttggtt | agtttttagg | gaaggagaa | tatcgaagta | tttgtttttt | 8100 |
| tagcgaatt | ttttggtttt | tttttttttt | aaggaggtga | tagttttggt | agataagggt | 8160 |
| tcgtgtgtat | attaggagt | aggttcgggt | gtggagttt | ggattaatag | ttattgagta | 8220 |
| tttattgttg | tgtgttttcg | tataaaaatt | tatttagttt | ttggaataat | ttataagggg | 8280 |
| ttatatttat | tattattttt | attttataga | tgaggagatt | gatgtttaga | gaggtgaagt | 8340 |
| aatatgttt | agattatata | gttggtagac | ggagaatcgg | tagtaaattt | aagattgtag | 8400 |
| tttaaaagc | ggtgttggtta | gttttttagt | tatttggtaa | tattggttga | gtatttttat | 8460 |
| gtaagaatt | gttttaggt | tgaggaatat | aaaagtagga | aaggttagtt | tgattcgtag | 8520 |
| tttggtagaa | cgtttagttt | tttttttttt | ggataaaatt | ttttattttt | atttagtagt | 8580 |
| ttagttagaa | atttttgtga | tattttttgt | aatttttttt | tttagtatat | tttatttatt | 8640 |
| attatatttt | atttattatt | atattttatt | tattattaga | ttttatagtt | tttttttttt | 8700 |
| gtttttttta | tttttgttta | ttttacgatt | tttatttttt | tatgtgtttt | gtatgtggtt | 8760 |
| atttttttgt | atgtatatag | atataggttt | tttttttatg | tgaatttttt | tatagatttt | 8820 |
| gttttgggtta | ttttgtttta | gttttagagg | atttttttat | ttgaaatgat | tattttgtag | 8880 |
| ttaaaaattt | agtttaaaat | ttattagagt | tcggttagtt | ttttgagaat | taaaattttag | 8940 |
| ttttagagtt | agataaaatg | atgttatttt | ttaatgagtg | tttaattata | ttttatttgt | 9000 |
| taatgagtg | taggtgttga | aagtttaaag | aggttaaagt | gggatagtaa | gtaaggggta | 9060 |
| gaaatgaatt | tttaattttg | agtattgatt | ttagagttta | tagttttaat | tattatgtta | 9120 |
| taattggtttt | ttttgttttt | ttttgtgttt | tggttttgtt | tgtttttttt | gttgggtttt | 9180 |
| tatttgttta | ttaggtttta | gatttaaagt | gattaagaga | aatataatgt | gagttatatg | 9240 |

| | | | | | | |
|----------|---------|---------|---------|----------|---------|-------|
| tatatat | tattttt | tgttttt | ttttttt | gagatag | tttgg | 9300 |
| tttaggt | ggagtat | ggtataa | tagtttt | taattatt | tttttag | 9360 |
| taagta | ttttatt | gttttta | ttgttgg | tataggt | tgttatt | 9420 |
| tttggg | ttttgat | tttatag | tggtttt | ttatggt | tagattg | 9480 |
| ttga | ggattta | tattttt | tttgg | tttttaa | tgttgga | 9540 |
| aaaggg | gtattag | ttgg | tttaaa | taaaaag | tatgtga | 9600 |
| taatttt | taatata | gtaaata | tattatt | tagtat | tatttag | 9660 |
| aatatatt | aataga | atttta | gtaatta | ttttaaa | tattagt | 9720 |
| atatttt | ttg | ttag | aaattta | gatatt | gtttaag | 9780 |
| atttta | agattag | tatttgg | gttaag | tattcgt | tatagt | 9840 |
| atttgg | tagta | ttag | agggag | ggtta | gtaata | 9900 |
| aagtgg | tg | tg | gaggg | ttatg | ttcgt | 9960 |
| taagata | ag | ttgt | ttttg | ttttg | ttttg | 10020 |
| aggatt | at | gaa | gag | tattg | gatttaa | 10080 |
| tttaaa | tagtat | taa | aagt | tg | gatgt | 10140 |
| tttttag | gtgag | tgag | agt | cgt | gggt | 10200 |
| ttatttt | gag | atta | gtg | tttt | gttt | 10260 |
| tttttt | tttt | ggt | tatt | ttttg | tttt | 10320 |
| tg | atta | agt | gg | ttg | tg | 10380 |
| tgaat | aa | gat | ttt | tcga | ttat | 10440 |
| tgata | aa | cgag | ttta | taaa | aat | 10500 |
| ggtac | ttat | ta | at | gtc | gagg | 10560 |
| tttag | ta | ttg | ta | ttg | aaaa | 10620 |
| aaa | cg | gt | gt | tt | ggt | 10680 |
| ggag | ttg | gag | ttg | ttg | tag | 10740 |
| ggtat | g | t | a | aaa | att | 10800 |
| t | aga | a | t | t | att | 10860 |
| ggg | aaa | ag | t | t | g | 10920 |
| tag | g | at | g | t | a | 10980 |
| ggag | ttg | a | ttg | ta | ag | 11040 |
| ttata | aaa | ta | ag | gt | at | 11100 |
| ttg | cg | at | tg | ttg | ag | 11160 |
| gatt | ttg | gt | tag | ttt | aaaa | 11220 |
| aaa | ttg | ga | ttt | aa | at | 11280 |
| g | gg | g | at | ttg | ttt | 11340 |
| tgat | ttat | t | gt | agg | tg | 11400 |
| t | gt | g | gt | tt | ttt | 11460 |
| ttag | ttt | aat | gg | tt | gga | 11520 |
| cgt | agg | tg | gag | a | g | 11580 |
| at | ttg | t | at | ttg | ttt | 11640 |
| ga | ag | t | at | gt | t | 11700 |
| t | t | t | t | ag | t | 11760 |
| g | g | c | g | g | a | 11820 |
| g | g | g | t | g | a | 11880 |
| at | g | g | g | t | g | 11940 |
| g | g | g | t | a | a | 12000 |
| tt | t | a | a | a | a | 12060 |
| tg | g | g | t | t | a | 12120 |
| aa | t | g | g | a | g | 12180 |
| c | g | g | g | t | a | 12240 |
| tt | t | a | t | a | a | 12300 |
| tg | g | a | c | t | g | 12360 |
| a | g | t | t | g | a | 12420 |
| aga | a | a | g | a | a | 12480 |
| tt | a | a | a | a | a | 12540 |
| at | c | g | a | g | a | 12600 |
| a | g | t | t | a | a | 12660 |
| t | t | a | a | a | a | 12720 |
| t | tt | g | g | a | g | 12780 |
| ta | a | t | a | a | a | 12840 |
| a | a | a | a | a | a | 12900 |
| t | tt | a | a | a | a | 12960 |
| tt | a | a | a | a | a | 13020 |

| | | | | | | |
|---------------|---------------|---------------|--------------|---------------|--------------|-------|
| atthttat | aaatgagaga | tgaatggtta | aataagtata | tgaaaaagtt | tttaataatta | 13080 |
| ttagttatta | gggaaatgta | aattaaat | atattgagat | attatttgat | ggttataata | 13140 |
| aaaaagatag | gggttaggta | cggtggttta | tgtttgtaat | tttagtaatt | tgggaggttg | 13200 |
| agggtgatgg | attatgagat | taggagttcg | agattagtta | gtttgtttaa | cgtggtgaaa | 13260 |
| tttcgthttta | attaaaaata | taaaaatag | ttgggtgtgg | tggtgcgtgt | ttataatthtt | 13320 |
| agttattggg | gaggttgagg | taggagaatt | gthttgaatth | aggaggtaga | agttgtagtg | 13380 |
| agtcgagatt | gtgttattgt | atthttagtht | gggtgataga | gagataatth | atthttaaaat | 13440 |
| aataaataaaa | taaaaaaggt | taggtatggt | ggtthtatatt | tgtaagthta | gtatthttggg | 13500 |
| aggthtaaggc | gggtggatta | thttgaggtcg | ggagthtagg | attagthttga | thaatatgga | 13560 |
| gaaatthtcgt | thttatthaaa | aaaataaaat | tagtcggacg | tggtgttacg | thttgtaat | 13620 |
| thtagthtatt | taggaggttg | aggtaggaga | attatthtga | thtaggaggt | agaggttacg | 13680 |
| gtgagtcgag | atthtttht | tgtatthttag | thttgggtaat | aagagtgaaa | thttatthtta | 13740 |
| aaaaaaaaaaa | aagaaaaaaaa | agataatthtt | aagtgtagga | tgtggaaaag | tggaatthttt | 13800 |
| atatattgth | ggtgagaatg | taaaaataata | tagthtatthtt | agaaaaatag | thgatagthtt | 13860 |
| thtaaaaat | aaatthtaaat | cgattatata | atthtagtaat | tgtatthttta | ggagthttatt | 13920 |
| taaaagcgtt | aaaaatata | thttatagta | aggtatgtat | gtgaaatgtht | atagtaaatat | 13980 |
| tatthtaagat | agthtaaaaag | tggaaataat | thaaagtgttc | gthtaattgat | taatggataa | 14040 |
| ataaaaatag | thtatatthtcgt | ataatggtat | atthtagaagt | aataaaaaaga | aataaatthgt | 14100 |
| tgataaaatg | thgtthttthtt | thattgthttg | thttthttatt | tagtataatg | thttggggcga | 14160 |
| gthtttaaaa | tattatgthta | agtaaaaagac | gthttgatgta | aaagtthtga | tatttatata | 14220 |
| thttatthttat | agaatatgth | tagaatatgt | aaatthtagg | atagaaaagta | gatttaggtg | 14280 |
| ggthtaagg | thgggggtgag | aatggagagt | agthtataaat | gggaatthttt | ggggatggta | 14340 |
| gaaatgthtt | aaaatggatt | thttthgtggt | aatggttata | taatthttata | thttatthaaa | 14400 |
| atthattaat | tatthttthta | taatgtgtaa | atthtaggta | tgtaaatthtt | atthttaaaa | 14460 |
| atcgthtaaaa | aatgtaagth | thagggttat | atthtttaggg | gthttthtatgt | agthtaatgtg | 14520 |
| gggtggtgth | taagaaatth | gthttthttt | thttgaggtta | gagthttcgth | thgtcgthttta | 14580 |
| ggttgggggtg | tagtggcgtta | atthttagtht | atthtgaagth | tcgthttthtcg | ggtthtacgth | 14640 |
| atthttthttgt | thtagthttth | cgagtagthttg | ggattatag | cgthtcgthtat | tacgthttag | 14700 |
| taatgthttth | thttgthttgt | thttgthttth | thtagtagaga | tggggthttta | tcgthgtthttg | 14760 |
| taggatgthtt | tcgattthtt | gatttcgtga | thttatthtcgt | tcggtthttth | aaagtthttgg | 14820 |
| gattataggt | tgagthtatt | gthtttggt | aaaatthtga | thtttaataa | gthttthttg | 14880 |
| gggtthttga | tgtatgtggt | thtttggt | atthtttaggg | tattatthggt | thaaagtthttg | 14940 |
| aaggggttg | agtaagagga | aggtagacgg | gaggagtag | gattthgtth | thgtthttata | 15000 |
| thgggggtata | gtatgtgtaa | atgthttthtt | tagagthttth | thattthttth | tgttgggatt | 15060 |
| tgtcgattaa | gthttthttth | tattaaatthgt | tagaggthtt | thttthttth | thttthttthaa | 15120 |
| gatggagthtt | tgtthttgt | gthttaggtg | gagtgtaaatg | gtataatthtt | ggtthattgt | 15180 |
| aaatthtcgth | thtcgggtht | aagcgatthtt | thttgthtttag | thtttcggag | agthtgggatt | 15240 |
| ataggcgtgt | atthattatgt | thggthtaatt | thttthtgtat | thtttagtaga | gatgatatth | 15300 |
| tattatgthtg | gthtaggtthg | thtttaatthtt | thgtatthttag | gtaatthtatt | tgtthtcggt | 15360 |
| thtaaaagt | thgggattat | aggtataagt | tattgcgttc | ggtthtaggg | thttthttata | 15420 |
| tataaattth | atthtttht | thttthgtth | aaagtthttth | atagthttthta | thattthttag | 15480 |
| gtaaaagtthta | aaatthtatt | gtatattthtt | taagtthgtth | thttggggtag | agagtagthtt | 15540 |
| gggaggtagg | tgagaaaggg | thgtggaagt | gtaagtthtag | thaaatthttg | atattthgt | 15600 |
| tatcgthggat | thttthgtatta | thttthattthta | aaaagthta | cggtggtggt | thtacgthttgt | 15660 |
| aaatthtagta | thttthgggtg | thtgaagcggg | taggtthatt | gaggttagga | gthttaaagatt | 15720 |
| agthttgthta | ataggtgaa | atthacgthtt | tattaaaaat | ataaaaaatta | gthtgggcgtg | 15780 |
| gtgataggta | thttgthaaat | tagthtattthg | ggaggtgag | gtacgagaat | thgagaggtta | 15840 |
| gaggtthtag | tgagthtgaga | tcgthggtgth | gtatthttaat | thtgggcgaga | gagthaaaat | 15900 |
| gtgthtcgaa | aaaaaagatt | thttthttthtt | thtgatagata | gthttgthttg | tgtthgtthtag | 15960 |
| gthggagthgt | aatggtgthga | thttthgthttta | thtgaatthtt | cgthttcgtag | gthttaaaggga | 16020 |
| thttthttgth | thagthttthcg | gagtagthttg | gattataggt | gthttgthtatt | thattthtag | 16080 |
| aaatthttgthta | thttthtagtag | aaacgggtht | thattatatt | ggthtaggtg | gthttcgaatt | 16140 |
| thtagatthttta | ggtgatttht | tcgthttcggt | thttthaaagt | gthtgggatta | tatggtgthga | 16200 |
| gthtatcgthcg | thtcggtthtt | thttthttthtt | thttthttthtt | thttthttthttg | agatagthtt | 16260 |
| thttthgtthgt | thtaggtthgga | gtatagthggt | atgattthtcg | thttattgthaa | thttthgtthtt | 16320 |
| thgggtthaa | thaaatthttth | gthtttagtht | thttgagthggt | tgggattata | ggtthacgtht | 16380 |
| atthattthgg | gthtaatthtt | gtatthtttag | tagagatggg | gthtttattat | thttggtthagg | 16440 |
| thggthttcga | atthttthgatt | thtaagtgatt | cggtthttthta | aaagtthttggg | atthataagacg | 16500 |
| tgagthtattg | cgthtcggtth | atthtattthttg | atthattgag | thttthgtthtat | thttthtaaat | 16560 |
| thttgthtttag | aaagtthagth | thttatthttatt | thattthattg | thttggtthgt | tagagattthgt | 16620 |
| thtggaaagt | gagatthttth | thtagthttthtt | thttthttthtt | taggthtaaa | gagatthagg | 16680 |
| aaatgggcga | tgtthttat | tagthttthgt | thtttagagatt | thtttagagatt | ggatagthtga | 16740 |
| gthtcgthggag | agggattthtt | tattthtttaag | taggthaaatta | thttatthttth | thtttagaggga | 16800 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| acgagggatt | tatttttagtt | tttattttttg | attattttttt | tttagttttg | attgtataat | 16860 |
| ttgaaattta | tatacgggtt | gtaattttttt | ttttttttttt | tcgagacgga | gtttcgtttt | 16920 |
| gttttttaggt | tggagtgtag | tggcgcgatt | ttggtttatt | gtaatttttcg | tttttcgggt | 16980 |
| ttaagcgatt | tttttggttt | agttttttta | gtagttggga | ttatagacgc | gtattattat | 17040 |
| gttcggttaa | tttgtttgta | atttttttaa | tttgtttgg | ttttttatag | ggagaggatt | 17100 |
| tgtatatgtt | gtgttttg | tgatgaaagg | agtttttttt | atatattttt | ttttttaatg | 17160 |
| tatttagatt | ttaaagttta | ttatttattt | atttaataaa | tatttattaa | gtattttttg | 17220 |
| aattcgggtt | aattattttt | tttttaggta | tttatttaat | ttcgggtatt | tttcgcggtt | 17280 |
| tttttgagtt | tttttcgct | tttatatagt | ttttgtttat | tgttgaaaa | tatttttggt | 17340 |
| aagtttcggt | tttttattag | tttgtttgt | tttttgtgt | tttgggatag | gttgtaaaat | 17400 |
| tggaggtgat | aaatgtgggt | aggaaatgga | gggttttttt | atatttttag | ggtcgggttt | 17460 |
| agttttgta | tttttggtt | aatatcgcg | atgtaattgg | tatgggattc | ggaagtgtgt | 17520 |
| ggtaaagcgt | cggggtttcg | tttggtcgtt | tttttcggac | gtttgtttgt | agttagcga | 17580 |
| gttttttaa | tttaggtttg | ggttttgcga | gttttaggtg | tttttgtttc | gcggcgttgg | 17640 |
| cgaagtcgaa | gttcgagaac | gtttatcgt | gcgatcga | ggtcgttttt | ggggtgggt | 17700 |
| tgaggttga | gttgttttt | ttcgtatta | aggattttaa | tttttagcga | cgtagtcgtc | 17760 |
| gttttcgtt | aggttgggag | gtattgtagg | gattcgcgct | tttaggtggt | taaagagcga | 17820 |
| tttttttga | ttttagggtt | ttggcgggt | aggttttagt | atcgtattcg | gcggaggtcg | 17880 |
| aaggtttgtg | gggtaggata | ggagtttttc | gtgtcgtcgg | aaggtcagag | gacgaaggag | 17940 |
| ggcgttaatt | tatttttatt | tgggttggcg | gtaacgtcga | atttcgtagt | gatcgcggag | 18000 |
| ggtaaggtg | aaaattgtg | gggctcga | gggtaggtg | ggggagggc | ggttttaggg | 18060 |
| agtaaggagt | ttatttgtt | cgtcgtcga | gtgttttgg | gtcgtcgtt | tacgtttttt | 18120 |
| ttgggttac | gattttcga | tttaattgtg | ttcggttttt | gttttttttt | ttttgttgtt | 18180 |
| gttcgcggg | ttgtaattt | acgttttagt | tgggctgtt | agggcgtagt | tttcgtttag | 18240 |
| gtcgttagcg | ttttattcgt | tttattgggt | tatagattcg | ttggcgttgg | ggtttcgttg | 18300 |
| gcgttgggt | ttcgttggcg | ttgggtttc | gttggcgtt | gggtttcgtt | ggcgttgggg | 18360 |
| ttcgttggc | gttgggttt | cgttggcgtt | gggtttcgc | ggttgtgtac | gtgagttcgc | 18420 |
| gtggtttcgt | ttcgcgcggt | tgtgtacgtg | agttcgcgcg | gtcgcgtttc | gtttcgtttt | 18480 |
| aggaggttag | cgctcgtta | tttgggatgt | taggattttc | gttgtgtttt | ttggattggt | 18540 |
| ttgggggatt | tcggcgtatt | tttaggattt | aggagttttg | gaagttgttt | gagagaaatt | 18600 |
| agttttggga | gggtttcga | tttagttttt | tgtttcgtt | tcggatcggg | gtttcgggtt | 18660 |
| aaggtgttta | gaggaatagt | tgattttttt | atttttgcgt | agggtagaga | tttttaaaatt | 18720 |
| tttttttaa | atgtagggtt | ttagtttttt | ttaggaggtt | agtgaattta | gatttttagt | 18780 |
| tttttgagtt | taagtatgaa | tagggaattg | gggattatta | ttatgtttat | atttttgggt | 18840 |
| gttaggaagt | ttaggtaggt | ttttgtttat | tgtagacgga | tttttttttt | taggggttaa | 18900 |
| gaaaggtttt | gtatagtaag | ttaattaagt | tttattagta | gagttgcgtt | gtaattaggt | 18960 |
| ttttagtgtt | ttttttatat | ttttagtttt | aagcgatata | ggatatagtt | tttaattttat | 19020 |
| ttttttttat | ttattttttg | ttttgttgtt | ttattttttt | tgttatatat | tttaagttat | 19080 |
| ttatttaatt | tttagttttt | ggtaatttag | tttttagttt | tggaagttta | gttttagttt | 19140 |
| tgttattatt | atttttattt | tggttttttt | tgagaataag | tggagggtaa | atagagttta | 19200 |
| ggtttgaatt | tattcgtttg | aaaaataatt | taagttaaat | tttgtgggag | tagtgtttgt | 19260 |
| ttgggttatt | ttaggtttat | gtgtattttt | gaattttttc | gggtatttgt | tttaaatgta | 19320 |
| aagaggttgt | tatttaatgt | ttgggttttt | ttgatttttt | gttttttttt | ggttgttttt | 19380 |
| ttgtttttgt | cgttttttat | ttttttttgt | tatggttagt | atttggtttt | taggttagaa | 19440 |
| aaggtggatt | tcgtgttttt | ggatttttatt | aatgttagga | gttatataaa | tatttttata | 19500 |
| tatatataga | gaggttcgc | gtatgtttgg | aaatttttag | atattaaaga | ataaagtgta | 19560 |
| ggatataata | gagatattgt | atatattgag | aaatgtgtat | attttttaaa | tgtagagaaa | 19620 |
| tgtatattaa | tatatagaga | tatattttat | aaatatattt | atagaaatag | atatattttat | 19680 |
| attttaagaa | atattaagat | atttatgaag | ggaatatttt | agatatgtat | aggattttga | 19740 |
| atttatatgg | atatagatat | ttaggagtag | gtgggttttg | atttaggtgt | atatagttta | 19800 |
| atatatagta | tttatatata | agttttttta | agtttaaatg | tcgtaagaga | tttttataga | 19860 |
| aagaaaattt | tttttagaggt | tttaaggtt | ttgtttggaa | ggaagaggaa | gaaagtgttc | 19920 |
| gttaggtatc | gaaatgttaa | ggtattgtaa | agtgaaaatt | atttttttaa | tttcgtggga | 19980 |
| atagtatttt | tatttttatt | tttttagggaa | atagggaaagt | ggtttaattc | gttttgtaaa | 20040 |
| tttaatttag | atttataaga | gtgttttttt | tttcggcggg | gagaggttag | gttttttagtg | 20100 |
| ttgtagttta | gtgaatgttg | attcgttgag | gtttatttta | ggttataggg | tgttgggtaa | 20160 |
| atttattaag | ttcgtacgtt | tgtttaaat | ttggttgggt | taggaagagt | gtttggtgga | 20220 |
| tatttagttt | ggcgtatat | aggaggtacg | gtcgggtgaag | aaaatggatt | tgtgtattaa | 20280 |
| gggtttgtat | tatttcgtcg | ggttttggtt | tgatagatta | ttttttattt | tttttgatta | 20340 |
| cggattgggt | taggaggttt | attttcgggt | cgtttattgt | ttgtaaatag | tagagggtag | 20400 |
| cggtgagggc | gtataaagtt | ttagtcgttt | atgttttttt | aagtttttgt | tatttagggt | 20460 |
| ttttgtttgt | ttcgtttaag | gcgtattagt | ttcgttttgg | gcggcgcg | gttaggttag | 20520 |
| gaatcgttat | gttttaggtat | tgagttggag | ttttgggtat | tttattgtgt | agtgtaaaaa | 20580 |

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| gggaattttg | aattttatat | tgtaattat | tatgtttttt | ttcgttttag | gtgtaggtg | 20640 |
| ttttatggat | gttatacgtg | gtagcgggga | gaaaagaagt | gattcggttt | ttgtcgaggg | 20700 |
| ttgttgttat | ttacgtattg | ttggcggatt | tagtgtttag | tagttttgat | ggttaggtaa | 20760 |
| cgtttaattt | aggaggtggg | atatattttc | ggagtatta | gtgagatgtg | gcgagttggg | 20820 |
| agaaaagtcg | gttttttcta | tttgttattt | tttgttttta | gttatggtag | taaattatcg | 20880 |
| cggttagttt | cgatttat | cggatttttc | gggttttatt | tgtaggtttt | ttagatcggg | 20940 |
| ggtcgtcgtt | agtttcggaa | ggtcgagagg | tgcgtttgtt | aggtcgtgcg | cgtcgtcgg | 21000 |
| taggtcgtgt | aggagtcggg | gtacggtagt | ggtatagtta | cggcgcgggt | taaggtttag | 21060 |
| ggcgcgggcg | ggcgggtgta | gtaggtcggg | aggtttttcg | tgggcgcggg | cggttgcgcg | 21120 |
| gggcgcgggt | tagcgggtt | gaggaggtg | cggttgcgtt | aggggcgcta | ggtcgttata | 21180 |
| tgggtttttt | gttaggtggg | gaaaggttgc | ggcggcgtcg | ggcgtttgat | cgtagggtag | 21240 |
| cggtggttgc | gggtagggcg | gcgggaagta | ggcgggttg | aattcggcgg | tagttgtggc | 21300 |
| ggtcggcgtg | tggtagagcg | ggggcgtcgg | ttcgtaggtc | gtttggggta | gagacgatag | 21360 |
| gcgggtttcg | tcggtagttg | tttttagttc | gtcggggcgt | ttttggggaa | ggagaggggt | 21420 |
| gttttaggtg | atattattgt | agttgaaagg | attttagttt | gaaaggattt | tttttagttc | 21480 |
| gggagggttt | agttgtagtt | tttttcgttt | tttgggttc | gaggtttatt | ttcagagttat | 21540 |
| tcgcgggctg | aggagtttta | ggttggggat | agcggttatt | tgcgtatcgc | gtttttagtt | 21600 |
| ggcgaagttt | ttgcggtttt | tttttttcga | gatttcgtta | gcgcgggata | ggagggttgg | 21660 |
| aaagtttagt | attaggcgt | ttcgattttt | cggtcgaata | tgtttcgggt | tcgagattat | 21720 |
| ttattatggc | ggagttagtg | tgtattagta | tgggcggtg | agtgtcgggc | gcgagggcgtg | 21780 |
| gcggcggcg | gttttagtt | cgacggcag | cgttgcgggt | agaggcgtcg | gagttcggta | 21840 |
| tagttcgcgg | aggttacggt | gcggtagttg | ggtgcgattt | agtccggatt | ttagtgggcg | 21900 |
| tgggtttttt | agcgggtggtc | gttttcgggc | gggtttatgg | aggtaggacg | gacgggtatt | 21960 |
| gatcgtat | ggaaaagtcg | gttaacgggt | gttttagcgtt | cgtttttttt | tttttgagcg | 22020 |
| tagtttggtt | tatgtat | agttatagtg | gcggggcgc | gtcgggtttt | cgggggcgcg | 22080 |
| tattaagag | ataagttatg | gtagt | ttttttttt | tagcggttgt | tataggggag | 22140 |
| gtagtggttt | tggcgggatc | gggtagggtg | ggtagggggc | gtgaatttag | tttatattga | 22200 |
| ttagagtcgt | tttttgggtt | cgtagttcgt | agttgtagtg | aaatagaaat | attcagaggtt | 22260 |
| gggacgtaag | ggttttata | ttgtttgggt | ttattatatt | tttgttttta | agtaaagtat | 22320 |
| tatttttcgt | tggtttagt | tttttcggta | aaataggtat | taattttttt | tttaagttat | 22380 |
| ttaagagatt | ttggatgagg | tgggtaatcg | ggaacgtgga | agaagtttgt | gagttgaaga | 22440 |
| ttttagttta | ttgtttatg | cgtatatagt | gttagtgctg | gtgtttaggg | ggtttattat | 22500 |
| tatagttttt | tttgaatggt | tttggattaa | aatttttcga | agggtttat | ttcgttttta | 22560 |
| ggttttttgg | ggtgaaaatt | tttaagttaa | gttaggtttt | gagggagttg | gagatagaag | 22620 |
| ttggtagat | agagagagag | atttgtttgg | gatgggtttg | ggtatgagtt | ttatttgata | 22680 |
| aggggttaag | atatttttga | agttataggg | ttttgggttg | attgagtggtg | ggtatgagat | 22740 |
| attagagat | tgttgagtga | ttgtgtgtaa | ttgattagta | tggatgctgt | gtgattgtta | 22800 |
| tattaatag | cgtaggttat | atgtttttgt | gttattatgg | taggtat | ttggttgta | 22860 |
| agtaagttt | agtttttcta | gttatttttt | ttttttttt | aaatttaatt | ttttttttt | 22920 |
| gataagtaga | gatagggatt | aaagatggat | aaatgagagg | aatgtaagg | agaataggaa | 22980 |
| ttaagaaagt | agaagaaga | tgttttttaa | aagtaagatg | tgtagattag | ggtaaagtag | 23040 |
| aaataattat | agtagattta | ttatagaatg | tgttagtttt | tgtgtaaagt | gttttacgaa | 23100 |
| tattgggttt | tttaattttt | atagtagttt | tatgaaagag | atatttata | tatttttatt | 23160 |
| tttaggggag | gggtgaggtta | tggagaggtt | aatgggttg | tttaaggtta | agtagttaag | 23220 |
| tggtagagtt | ggggtgattt | gtttttggtt | ttttttggtt | tagaattatt | attttttatt | 23280 |
| gttttagatt | taggggtttg | gatttagttt | tgattttggt | taattgtgtg | attttagtta | 23340 |
| tatttcgttt | tttttctgat | tttagttttt | gtaatttcga | taatattaat | aaaaaagtag | 23400 |
| gggaaggtag | agaagattag | ttgttttttt | aggttttttt | ttagtgaagt | agtaaatgag | 23460 |
| aagattaagg | taaaattatt | gatatagaag | aggaggtgat | agtgatatta | ttaatattatg | 23520 |
| gaatataagg | ttttaagtgg | gtatttttaga | aggtatatgg | taggtgtaga | gtttagggtt | 23580 |
| gtagatattt | agtttgggga | gtgggtggga | agaagaggag | ttagaagtta | aggatttatt | 23640 |
| attaagaatt | tttgaagtt | tggaaatcgtt | agcagagttt | tagaggatga | agcgggaattg | 23700 |
| ggaagttttg | taattttttt | gttcgagatg | ggttagagat | taggattgtg | tttgtgtgta | 23760 |
| tttatatag | agggagtttg | tggtaatggt | ttgttttttt | tttttttaat | tttagttata | 23820 |
| gggtattaga | agttaagttt | taggataggt | gttagtgctg | gaatagaaat | tgatgaattt | 23880 |
| tttggggaga | agtttaggag | tcgtaaagga | gatgatgttt | gtgtttagtt | gcggggatgt | 23940 |
| ttttgttttt | tttagtagtt | tgggaaggtt | agagtggcgg | gatgtgtgtg | ttgtttgtgg | 24000 |
| tttttttggg | tttttttaat | ttatttttta | tatttttgat | agtttttttt | tgttttattg | 24060 |
| ttttatttcg | attatttttg | tttttttttt | tgtttttttt | tttatttata | tgtttaatgt | 24120 |
| cggtaatttt | tttttagagt | agagaagttt | tttttttttt | ttgtat | tatgagtgag | 24180 |
| tttatttttt | tttaagtttt | tttttttttt | ttttttataa | tagagtttta | ttttgtttgt | 24240 |
| tgggttgag | tgtagtgta | taatttcggt | ttattgtaat | ttttattttt | cgggtttaag | 24300 |
| cgatttttat | gttttagttt | tttagtaaat | tgggattata | ggggtttatt | attacgttcg | 24360 |

ES 2 597 845 T3

| | | | | | | |
|------------|-------------|------------|------------|------------|------------|-------|
| gttaat | gtat | tagagatagg | gttttattat | gttagttagg | ttgatttcga | 24420 |
| at | tttaagtgaat | attcg | gttttttaa | gtg | tataggtatg | 24480 |
| agttattacg | ttcgg | gtttttaatg | attatttg | tattg | ttttaattgt | 24540 |
| gttatt | gtttgaatta | ttagt | at | atagatatag | ttg | 24600 |
| cgtat | ttagat | tattg | gtttgaa | tttta | aatattgaat | 24660 |
| ttaat | ttttat | agtatta | tttaattg | ttttttt | ttt | 24720 |
| tataaatgat | attat | tttaag | aaattg | gttatt | aagagc | 24780 |
| gttttt | taatt | aag | ttatg | ttg | taatt | 24840 |
| atatt | ataga | tgtg | aattg | gg | agg | 24900 |
| gtat | ttagt | aaatt | at | atag | gg | 24959 |

<210> 5

<211> 6096

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 5

ES 2 597 845 T3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tattagtgtg | agatttaaaa | tttttttttt | gtattgtata | gtgagatggt | tagggtttta | 60 |
| gtttagtgtt | tggatatagc | gatttttggg | tttgttcgtc | gtcgttttaa | gcgaagtggg | 120 |
| tgcgttttgg | gcgagtaga | tagagatfff | gggtggtagg | ggtttgggaa | gatatgggcg | 180 |
| gttagggttt | tatgcttttt | tatcgttggt | ttttgttatt | tgtaggtaat | ggacgagtcg | 240 |
| ggaatgagtt | ttttagatta | gttcgtgatt | aagaaaggtg | aggaatgggt | tgttagggta | 300 |
| gagttcggcg | agatgggtga | ggtttttggg | gtatagatff | atftttttta | tcggtcgtgt | 360 |
| tttttgtgtg | tcgttaggft | gggtgtttat | taggtatfff | ttttggttta | gttagatgft | 420 |
| aggtagacgt | gcggttttgg | tgagtttggt | tagtatfff | tggtttgggg | tgggttttag | 480 |
| cggattagta | tttattgggt | tgtagtattg | ggagtttggt | tttttttcgt | cgagggggag | 540 |
| ggatattttg | tggatttggg | gttgatttgt | agaacgagtt | aaattatfff | ttgttttttt | 600 |
| taagagatgg | gaatggaagt | gttgttttta | cggagtggg | gaaatgattt | ttatftttata | 660 |
| gtgttttagt | atftcgggtg | ttggcgggta | tttttttttt | ttttttttta | gtaggggttt | 720 |
| tggaggtttt | tgggggaatt | ttttttttgt | gggagtfttt | tgcggtatft | agatttaggg | 780 |
| gagtttgtgt | gtgagtattg | tgtgttaggt | tgtgtgtatt | tgagttaggg | ttatfttgtt | 840 |
| tttgggtgft | tgtgtttatg | tgagtttagg | gttttgtgta | tgtttgaaat | gtttttttta | 900 |
| tgggtgfttt | agtatftttt | ggagtgtgag | tgtgtttgtt | tttgtgaaat | tgtttgtgag | 960 |
| gtgtgtfttt | gtatgttggg | gtgtatfttt | ttgtatfttg | gggatgtata | tatfttttaa | 1020 |
| tatgtatagt | atftttgttg | tgttttgtat | tttgtftttt | ggtatfttag | gatfttttaag | 1080 |
| tatgctcggg | ttttttttgt | gtatataatg | gagtatftat | gtgatftttg | gtattagtaa | 1140 |
| aatttaggga | tacgggattt | atfttttttt | gtttgaggat | taagtattgg | ttatgatagg | 1200 |
| ggaaggtgag | agacgataaa | aatagagaga | tagttagaga | ggagtagaga | gttagagggg | 1260 |
| tttaggtatt | gggtagtagt | ttttttatat | ttggggtagg | tgttcgaaag | aatttagagg | 1320 |
| tgtatatgag | tttgaggtgt | tttaggttag | tattgtfttt | atagggtttg | gtttgagttg | 1380 |
| ttttttaaac | gagtgaatft | aagtttgggt | ttatfttgtt | ttttatfttgt | ttttagggga | 1440 |
| ggttaaggtg | gaagtgggtg | tagtaggggt | gggttggat | ttttaggagt | tgggtttag | 1500 |
| ttattaggag | ttgggggttg | ggtggatgat | ttggagtgtg | tagtagggaa | gatgaggtaa | 1560 |
| tagggtagga | agtgggtggg | gggaggtgga | attgggttg | tgttttgtgt | cgtttggaat | 1620 |
| tgggagtgtg | ggaaagatat | taggaatftg | gttgtagcgt | agtftttgtg | gtgggtttg | 1680 |
| gttggftttat | tgtatagagt | tttttttgat | ttttgaagaa | agagattcgt | ttgtagtggg | 1740 |
| taaaagtttg | tttggatftt | ttgttatta | gaaatatgag | tatggtggtg | gtttttagtt | 1800 |
| ttttatfttat | gtttgggttt | aagagatttg | gagtfttaggt | ttattgattt | tttgagaaag | 1860 |
| attaagatft | tgtatftttag | aaagaggttt | gggtatfttt | gttttgcgta | agggtagaag | 1920 |
| gattagttgt | ttttttgagt | atftttaattc | ggaatfttcg | ttcgaagtcg | agataggaga | 1980 |
| ttggatgcca | ggttttttta | gagtgggttt | tttttaata | atfttttaaaa | tttttagatt | 2040 |
| ttaggggtac | gtcgaatftt | tttaagtag | tttaagaat | ataacgagag | ttttaatatt | 2100 |
| ttaggtggcg | gctcgttggg | tttttggagc | ggggcgggac | gctcgtcgcg | ggatfttacgt | 2160 |
| gtataatcgc | gctgggacggg | gttacgcgga | tttacgtgta | taatcgcggg | atftttagcgt | 2220 |
| tagcgggatt | ttagcgttag | cgggattftta | gcttagcgg | gatfttttagc | ttagcgggat | 2280 |
| tttagcgtta | gctgggattft | agcgttagc | ggatftttagc | gttagcgggt | ttgtggttta | 2340 |
| gtggagcag | tggagcgttg | gctgatttag | cggagattgc | gttttggacg | tttttagttta | 2400 |
| gacgttaagt | tatagttcgc | gtagtagtag | taaaggggaa | gggttaggag | tcgggtatag | 2460 |
| ttggatcgg | aggtcgtgat | ttaggggaaa | gctgggcg | tcgatfttag | gtagttgcgg | 2520 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| cggcgaggta | ggtgggtttt | ttgttttttg | gagtcgtttt | tttttatatt | tgttttcggc | 2580 |
| gtttttagta | gtttttat | tggtttttcg | cggttattgc | gggattcggc | gttgctgta | 2640 |
| gtttagtggg | gagtgaatta | gcgttttttt | tcgttttcgg | ttttttcgac | ggtacgagga | 2700 |
| atttttgttt | tgttttatag | attttcgggt | ttcgtcagag | gcggtattgg | agtttgtttc | 2760 |
| gttagggttt | tggaattaga | gaaagtcggt | ttttggttat | ttgaagcgtc | ggatttttat | 2820 |
| agtgtttttt | agtttgggcg | ggagcggcg | ttgcgtcgtt | gaaggttggg | gtttttggtg | 2880 |
| cgaaagggag | gtagttgtag | tttttagttt | attttagaag | cggttttcgt | atcgttgagg | 2940 |
| tgggcgtttt | cgggtttcga | tttcgtagc | gtcgcggggt | agaggatatt | ggagttcgt | 3000 |
| gggttttagat | ttgggttgg | aaagtttcgt | tgattgtagg | taagcgttcg | ggagggcg | 3060 |
| ttaggcgaag | tttcggcgtt | ttattatata | ttttcgggtt | ttatgtagt | tgtattcgcg | 3120 |
| gtattgggta | ggaaatggt | gggttgagg | cgatttttag | agtataagg | agttttttat | 3180 |
| tttttgttta | tatttggtat | tttttagttt | gtaatttatt | ttagatata | agaaagtaag | 3240 |
| taggattggt | ggggagacgg | agtttaata | gaatattttt | tagtagtgag | taggggttgt | 3300 |
| atgggacgcg | ggaggagttt | agaggaggcg | cggagagtgt | tcgaggttgg | gtgagtgtt | 3360 |
| agaggggaga | tagttgaatc | gggtttaaga | ggtgttagt | gggtgtttgt | tgaatgaatg | 3420 |
| agtgatgggt | tttgaagttt | gagtgtattg | aaagaggggg | tgtgtaaaaa | gggttttttt | 3480 |
| tattatatag | gatatagtat | atgtaaat | tttttttg | gaaaagttag | ataggttaaa | 3540 |
| aaggttataa | ataaattagt | cgggtatggt | ggtgcgcgtt | tgtagtttta | gttattaggg | 3600 |
| aggttagatt | aggggaatcg | tttgaattcg | ggagcggag | attgtagtga | gttaagatcg | 3660 |
| cgttattgta | ttttagttg | gaaatagagc | gagatttcgt | ttcggaaaaa | aaaaaaaaaa | 3720 |
| gttataaatc | gtgtgtgggt | tttaggttat | ataattagag | ttggagggga | gtggtaagg | 3780 |
| atgagaattg | agatggattt | ttcgtttttt | ttggaggaga | gtgggtgggt | gtttatttgg | 3840 |
| gggtggggaa | ttttttttta | cgggttttag | tgtttaattt | taggggattt | ttaggatagg | 3900 |
| agttgatgta | aatagtcggt | ttattttttg | ttgtttttgg | ttttggagaa | ggaggaggga | 3960 |
| gttggggagg | gtttttat | tttagataat | ttttaagtag | ttaggatatg | ggtgagatga | 4020 |
| gtgagatatt | gatttttggg | atagaatttg | agaggtgtt | aaaaaattta | gtaattaaga | 4080 |
| taaataggtc | gggcgtagtg | gtttacgttt | gtaattttag | tattttggga | ggtcggatta | 4140 |
| tttgaggtta | agagttcag | attagtttg | ttaagatggt | gaaattttat | ttttattaaa | 4200 |
| aatataaaaa | ttagtttagt | gtggtggcgt | tagtttgtaa | ttttagttat | ttaggagggt | 4260 |
| gaggtaaag | aattgtttga | tttaggaggt | agaggttgta | gtgagtcgag | attatgttat | 4320 |
| tgtatttttag | tttgataat | agagggagat | tattttaaaa | aaaaaaaaaa | aaaaaaaaaa | 4380 |
| aaaaaagagg | tcgggcggcg | gtggtttata | ttatgtgatt | ttagttattt | gggaggtcga | 4440 |
| ggcgggtgga | ttatttgagg | tttgagttc | gagattagtt | tgggttaat | ggtgaaattt | 4500 |
| cgtttttatt | aaaaatataa | aaattagttg | ggtgggtg | taggtatttg | taatttttagt | 4560 |
| tatttcggag | gttgaggtag | gagaattttt | tgaatttgcg | gggcggaggt | tgtagtgaat | 4620 |
| taagattata | ttattgtatt | ttagtttgg | taataatagt | aaaattttgt | tttaaaaaaa | 4680 |
| aaaaaaattt | tttttttcga | gatatagttt | tattttttcg | tttaggttgg | ggtgtagtat | 4740 |
| tacgatttta | gtttattgta | atttttgttt | tttagatttt | cgtatttttag | ttttttaagt | 4800 |
| agttgggatt | ataggtattt | gttattacgt | ttagtttaatt | tttgtatttt | tagtaggggc | 4860 |
| gtggttttat | tatgttggtt | aggttggtt | tgaatttttg | attttaagt | atttgttcgt | 4920 |
| tttagttatt | taaagtgtt | ggattatagg | cgtgagttat | tacgtttggt | ttttttaaat | 4980 |
| gaaaatagtg | taaaaattta | cgataataaa | aatataaaa | atttattgaa | tttgtatttt | 5040 |
| tataattttt | ttttatttgt | tttttaggtt | attttttgtt | ttagaaagta | atttaaaaaa | 5100 |
| tgtgtagatg | gagtttgat | tttatttgaa | aatggtggga | gttatggaaa | attttgaggt | 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 | ggaggttgg | ataggaaaat | 5400 |
| cgtttgaatt | cgggagggcg | aggttgtagt | gagttaaagt | tgtgttattg | tatttttagtt | 5460 |
| tgggtaataa | gagtaaaatt | ttattttaaa | aaaaaagaaa | gaaagaaatt | ttttggtagt | 5520 |
| tgatgagaag | gaaatttaatt | cggtaggttt | tagtagggga | gatgaggaga | ttttaggagg | 5580 |
| ggtatttgta | tatgttgtgt | tttagtgtgg | gttagggagt | aggttattat | ttttttcgtt | 5640 |
| tatttttttt | ttgttttaatt | ttttttaagt | tttggttag | tgggtatttt | agtgtagttt | 5700 |
| aaggaattat | atgtatttag | attttttaggg | ggtgtttggt | aaaaatgtaa | attttggtta | 5760 |
| ggtgtagtgg | tttatatttg | taatttttagt | attttgggag | gtcagggcg | gtggattacg | 5820 |
| aggttaggag | atcagagatta | ttttggtaaa | tacggtgaaa | ttttattttt | atataaaaaa | 5880 |
| taaaaataaa | ttaaaaaaa | tattagttgg | gcgtggtggc | gggcgtttgt | agttttagtt | 5940 |
| attcgggagg | ttgaggtagg | agaatggcgt | gaattcggga | ggcggagttt | gtagttagtt | 6000 |
| gagattgcgt | tattgtattt | tagtttgggc | gatagagcga | gattttggtt | taaaaaaa | 6060 |
| aaagtaaat | ttttgggtat | tattttat | tgattg | | | 6096 |

<210> 6

<211> 6096

<212> ADN

<213> Secuencia Artificial

5 <220>

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 6

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tagttaatgt | ggggtggtgt | ttaagaaatt | tgtttttttt | tttttgaggt | agagtttcgt | 60 |
| tttgtcgttt | aggttggggt | gtagtggcgt | aatttttagtt | tattgtaagt | ttcgtttttc | 120 |
| ggggttacgt | tatttttttg | ttttagtttt | tcgagtagtt | gggattatag | gcgttcgtta | 180 |
| ttacgtttag | ttaatgtttt | ttttggtttg | tttttgtttt | tttagtagag | atggggtttt | 240 |
| atcgtgtttg | ttagatggtt | ttcgtttttt | tgatttcgtg | atttattcgt | ttcggttttt | 300 |
| taaagtgttg | ggattatagg | tgtgagttat | tgtatttggt | taaaatttgt | atttttaata | 360 |
| agtatttttt | gggggttttg | atgtatgtgg | ttttttggat | tatatttagg | gtattattgg | 420 |
| tttaaagttt | gaaggggttg | gagtaagagg | aagtagacg | ggaggagtag | tgatttgttt | 480 |
| tttggtttat | attgggggat | agtatgtgta | aatgtttttt | ttagagtttt | ttattttttt | 540 |
| ttggtgggat | ttgtcgatta | agtttttttt | ttattaattg | ttagagggtt | tttttttttt | 600 |
| ttttttttta | agatggagtt | ttgtttttgt | tgtttagggt | ggagtgtaat | ggtataaatt | 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 | ttttggggtta | 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 | tataanaaatt | 1260 |
| agttggcgtg | ggtgataggt | atttgtaatt | ttagttattt | gggaggttga | ggtacgagaa | 1320 |
| tttgagaggt | agaggttgta | gtgagttgag | atcgtggtgt | tgatttttaa | tttgggcgag | 1380 |
| agagtaaaat | tgtgtttcga | aaaaaaagat | tttttttttt | tttgagatag | agttttgttg | 1440 |
| ttgttgttta | ggttgagagt | taatggtgtg | attttgggtt | attgtaattt | tcgtttcgtta | 1500 |
| ggtttaaggg | atttttttgt | tttagttttc | ggagtagttg | ggattatagg | tgtttgttat | 1560 |
| tttatttagt | taattttttgt | attttttagta | gaaacggggt | tttattatat | tggttagggt | 1620 |
| ggtttcgaat | tttagatttt | aggtgattta | ttcgtttcgg | ttttttaaag | tgttgggatt | 1680 |
| atatggtgtg | agttatcgtc | gttcggtttt | tttttttttt | tttttttttt | tttttttttt | 1740 |
| gagatagttt | ttttttgttg | tttaggtttg | agtatagtgg | tatgatttcg | gtttattgta | 1800 |
| atttttgttt | tttgggttaa | gtaatttttt | tgtttttagtt | ttttgagtgg | ttgggattat | 1860 |
| aggttagcgt | tattatattg | ggttaatttt | tgtattttta | gtagagatgg | ggttttatta | 1920 |
| ttttggttag | gttgggtttc | aatttttgat | tttaagtgat | tcggtttttt | aaagtgttgg | 1980 |
| gattatagac | gtgagttatt | gcgttcgggt | tatttatttt | gattattgag | ttttttggta | 2040 |
| tttttttaaa | ttttgtttta | gaagttagta | ttttatttat | tttatttatg | ttttggttgt | 2100 |
| ttagagattg | tttgggaagt | ggagattttt | tttagttttt | tttttttttt | ttagggttaa | 2160 |
| agatagtaag | gaatagggcg | attgtttata | ttagtttttg | tttttagagat | tttttgagat | 2220 |
| tggatagttg | agttcgtgga | gagggatttt | ttatttttaa | gtaggtaatt | atttattttt | 2280 |
| tttttagagg | aacgagggat | ttatttttagt | ttttattttt | gattattttt | tttttagttt | 2340 |
| gattgtataa | tttgaaattt | atatacgggt | tgtaattttt | tttttttttt | ttcagacgag | 2400 |
| agtttcgttt | tgtttttagg | ttggagtgtg | gtggcgcgat | tttggtttat | tgtaattttc | 2460 |
| gtttttcggg | tttaagcgat | ttttttggtt | tagttttttt | agtagttggg | attatagacg | 2520 |
| cgtattatta | tgttcgggtta | atttgttttg | aattttttta | atttgtttgg | tttttttata | 2580 |
| gggagaggat | ttgtatatgt | tgtgtttttg | gtgatgaaag | gagttttttt | tatatatttt | 2640 |
| tttttttaat | gtatttagat | tttaaagttt | attatttatt | tatttaataa | atatttatta | 2700 |
| agtatttttt | gaattcgggt | taattatttt | tttttttaggt | atttatttaa | tttcgggtat | 2760 |
| ttttcgcgtt | ttttttgagt | ttttttcgcg | ttttatatag | tttttgttta | ttgttggaana | 2820 |
| atattttttg | taagtttcgt | ttttttatta | gttttgtttg | ttttttgtgt | gtttgggata | 2880 |
| ggttgtaaaa | ttggaggtga | taaatgtggg | taggaaatgg | agggtttttt | tatattttta | 2940 |
| gggtcggttt | tagttttgtt | atttttttgt | taatatcgcg | gatgtaattg | gtatgggatt | 3000 |
| cggaaagtgtg | tggtaaagcg | tcggggtttc | gtttggtcgt | ttttttcggg | cgtttgtttg | 3060 |
| tagttagcga | agttttttta | attttaggtt | gggttttgcg | agtttttaggt | gtttttgttt | 3120 |
| cgcggcgttg | gcgaagtcga | agttcagaaa | cgtttatcgt | agcagtcgga | aggtcgtttt | 3180 |
| tgggggtggg | ttgaggttgt | agttgttttt | ttttcgtatt | aaggatttta | attttttagcg | 3240 |
| acgtagtcgt | cgtttttcgtt | taggttgggg | ggtattgtag | ggattcagacg | tttttaggtgg | 3300 |

| | | | | | | |
|-------------|-------------|-------------|-------------|------------|-------------|------|
| ttaaagagcg | atTTTTTTTTg | atTTtaggt | ttggcgggg | taggttttag | tatcgtattc | 3360 |
| ggcggaggtc | gaaggtttgt | gggtaggat | aggagtTTTT | cgtgctcgcg | gaagggtcga | 3420 |
| ggacgaagga | gggcgttaat | ttatTTTTta | ttgggttggc | ggtaacgtcg | aatttcgtag | 3480 |
| tgatcgcgga | gggttaaggt | gaaaattggt | ggggcgtcg | agggtaggtg | tggggagggg | 3540 |
| cggTTTTtagg | gagtaaggag | tttatttgggt | tcgtcgtcgt | agttgTTTTg | ggtcgatcgt | 3600 |
| ttacgtTTTT | TTTTgggtta | cgatTTtcgg | atTTaattgt | gttcggTTTT | tgtTTTTTTT | 3660 |
| TTTTgttgt | tgttgcgcg | gttgtaattt | gacgttttagg | ttggggcgtt | tagggcgtag | 3720 |
| TTTTcgTTta | ggtcgttagc | gttttattcg | TTTTattggg | ttatagattc | gttggcgttg | 3780 |
| gggtttcgtt | ggcgttgggg | tttcgttggc | gttggggttt | cgttggcgtt | gggtttcgt | 3840 |
| tggcgttggg | gtttcgttgg | cgttggggtt | tcgttggcgt | tggggtttcg | cggttgtgta | 3900 |
| cgtgagttcg | cgtggtttcg | tttcgcgcgg | ttgtgtacgt | gagttcgcgc | ggtcgcgttt | 3960 |
| cgTTtcgttt | tagggagtta | gcgctcgtt | atTTgggatg | ttaggatttt | cgtttgtgtt | 4020 |
| tttggttgt | tttgggggat | ttcggcgtat | TTTTtaggatt | taggagtttt | ggaagttggt | 4080 |
| tgagagaaat | tagttttggg | agggtttcgt | atTTtagtttt | ttgtttcggg | ttcggatcgg | 4140 |
| ggTTtcgggt | taaggtgttt | agaggaatag | ttgattTTTT | tatttttgcg | tagggtagag | 4200 |
| atTTTTaaat | TTTTTTTTaa | aatgtagggt | tttagttttt | tttagggagt | tagtgaattt | 4260 |
| agattTTTTag | TTTTttgagt | ttaagtatga | atagggaaat | ggggattatt | attatgttta | 4320 |
| tatttttgggt | ggttaggaag | tttaggtagg | TTTTtgTTta | ttgtagacgg | atTTTTTTTT | 4380 |
| ttaggggtta | agaaagggtt | tgtatagtaa | gttaattaag | TTTTattagt | agagttgcgt | 4440 |
| tgtaattagg | TTTTtagtgt | TTTTTTtata | TTTTtagttt | taagcgatat | aggatatagt | 4500 |
| TTtaattTTta | TTTTTTTTta | TTtattTTTT | gttttgTTgt | TTtattTTTT | ttgttatata | 4560 |
| TTTTaagtta | TTtattTaat | TTTTtagtttt | tggtaattta | gttttagttt | ttggaagttt | 4620 |
| agTTTTtagtt | ttgttattat | tatttttatt | ttggTTTTTT | ttgagaataa | gtggagggtta | 4680 |
| aatagagttt | aggtttgaat | ttattcgttt | gaaaaataat | ttaagttaaa | ttttgtggga | 4740 |
| gtagtgTTtg | tttgggggat | tttaggttta | tgtgtatttt | tgaattTTTT | cgggtatttg | 4800 |
| TTTTaaatgt | aaagaggttg | ttattTaatg | tttgggtttt | tttgattttt | tgtTTTTTTT | 4860 |
| tggttgTTTT | tttgtTTTTg | tcgtTTTTta | TTTTTTTTtg | ttatggttag | tatttggttt | 4920 |
| ttaggttaga | aaaggtggat | ttcgtgTTTT | tggattttat | taatgttagg | agttatataa | 4980 |
| atattTTTTat | atatatatag | agagggttcg | cgtatgTTtg | gaaattTTTT | gatattaaag | 5040 |
| aataaagtgt | aggatataat | agagatattg | tatatattga | gaaatgtgta | tattTTTTta | 5100 |
| atgtagagaa | atgtatatta | atatatagag | atatatTTta | taaatatatt | tatagaaata | 5160 |
| gatataTTta | tattTTaaga | aatattaaga | tatttatgaa | gggaatattt | tagatatgta | 5220 |
| taggattTTtg | aatttatatg | gatatagata | tttaggagta | ggtagggttt | gatttaggtg | 5280 |
| tatatagttt | aatatatagt | atTTtatatat | aagtTTTTTT | aagtTTaaat | gtcgtaaagag | 5340 |
| atTTTTatag | aaagaaaatt | TTTTtagagg | TTTTtaaggt | tttgtttgga | aggaagagga | 5400 |
| agaaagtgtt | cgttaggtat | cgaaatgtta | aggtattgta | aagtgaaaat | tattTTTTta | 5460 |
| atTTcgtggg | aatagttatt | ttattTTTTat | TTTTtaggga | aatagggag | tggTTtaatt | 5520 |
| cgTTTTgtaa | attaattTTta | gatttataag | agtgTTTTTT | TTTTcggcgg | ggagaggtta | 5580 |
| ggTTTTtagt | gttgtagttt | agtgaatgtt | gattcgttga | ggtttatttt | aggttatagg | 5640 |
| gtgttgggta | aatttattaa | gttcgtacgt | ttgtttaata | tttggttggg | ttaggaagag | 5700 |
| tgtttggtgg | atatttagtt | tggcgatata | taggaggtac | ggtcggtgaa | gaaaatggat | 5760 |
| ttgtgtatta | agggtttgta | ttatttcgtc | gggttttgtt | ttgatagatt | atTTTTtatt | 5820 |
| TTTTttgatt | acggattggg | ttaggaggtt | tattttcggg | tcgtttattg | tttgtaaata | 5880 |
| gtagagggta | gcggtgaggg | cgtataaagt | tttagtcgtt | tatgtTTTTT | taagTTTTtg | 5940 |
| ttatttaggg | TTTTgttttg | tttcgtTTta | ggcgtattag | tttcgttttg | ggcggcgacg | 6000 |
| ggtaggtTTta | ggaatcgtta | tgtttaggtta | ttgagttgga | gttttgggta | TTTTattgtg | 6060 |
| tagtgtaaaa | agggaaTTTT | gaattTTtata | ttggta | | | 6096 |

<210> 7

<211> 24959

<212> ADN

5 <213> Secuencia Artificial

<220>

ES 2 597 845 T3

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 7

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|-----|
| ttgagagtta | ttgatttatg | agtaagagta | aggtagttta | aggtgttgag | ataggggtgtt | 60 |
| atgtgttttt | tgggatattt | gttttaattt | tttattatat | atTTTTTgtg | ggtggatgtt | 120 |
| taagggttag | aaatggtaat | atTTTataag | gattagtttg | tgggggTtaa | ggaaaggatt | 180 |
| tatgtttttt | tgggatgatt | agtaagtttt | gaatttgaaa | aagggTggtg | ttatttatga | 240 |
| aattatagag | taaaggagaa | gtaggttggg | gttggtgttg | gggatagagg | agatgttgag | 300 |

| | | | | | | |
|--------------|---------------|--------------|---------------|---------------|---------------|------|
| tttagtgttg | gatttgghaaa | atthttagata | gthtgaatga | gatgthttgga | gaaaatatgt | 360 |
| agthtggtagt | tgtgthttatg | ggtthtgatt | thtttattagt | gattthaggth | agagataata | 420 |
| tagthtgggag | ataatagtag | ataggtagth | atthaaagatt | tgggthtgggt | gtggtgattth | 480 |
| atgthttgtaa | thtttagtatt | thtgggaggt | gaggtgagtg | thttatttgaa | gthtaggagth | 540 |
| tgagattagt | thtggthgata | tgatgaaatt | thtghthtttag | thaaaaatata | aaagthtagth | 600 |
| gggtgthggtg | gtgggtthttt | gtaattthtag | thattthtagga | aattgagata | tgagaattgt | 660 |
| thgaattthgg | gaggtggagg | thttagtgag | thtgagattgt | gthattgtat | thtagthttaa | 720 |
| gtaatagagt | gagattthtgt | thataaaaaaa | thaaaaaaaaa | agattthgaga | ggggatgagt | 780 |
| thattthtga | agagtgtagg | aagaaaaaga | agthththttg | thtttgagaaa | gaattattga | 840 |
| tattaagtat | gtaggtgaaa | agagaaatag | thaaagaagat | agaaatggtt | ggagtaggat | 900 |
| aatgaggtaa | aaaggggtta | thaaaggtgt | aggaagtga | thgagaaaat | thaaaggaagt | 960 |
| tatagatagg | tagtatattth | tgttattthtg | gththththta | gthtgtggag | agagtagggg | 1020 |
| tattthttgta | gthtgggtata | ggtattattth | thttthtgggt | thtttagattth | thttthtagagg | 1080 |
| agthttattaa | thttthattta | tgtattgata | thttgthttgg | agthttgattth | thtgatgthttth | 1140 |
| gtggttgggg | thtgagggagg | gaggaatagg | gtattgttat | agattththtt | atataatagat | 1200 |
| atataataaat | ataattthtga | thttthgattth | atthttgagta | ggaaagthtgt | aaagththttth | 1260 |
| aattthtgtht | tattththttg | aattthgttg | gtggtthtttag | atthttagaaa | gththttaaag | 1320 |
| atgaattthtt | ggtththtagt | thththththtt | thtttattatt | thtttaggtht | agattattth | 1380 |
| gthttgagth | thtgaatttht | tatgtgthttth | thtgggatatt | tattthggaat | tataatattth | 1440 |
| atgagthtaat | ggtgthtatta | thattththtt | thttatatta | ataattthtt | thtaattthtt | 1500 |
| thattthatta | thtttattaaa | aagggattth | thaaaaataat | tagthththtt | tgtthththtt | 1560 |
| tattththttta | thtggthttht | tgaagthtga | gaaattgaga | thtatagagga | aatgggatgt | 1620 |
| agthtaggatt | atatagtthtag | gtagagthtag | gattthaaattth | aagthththtga | gthttgaggta | 1680 |
| gtagagaata | gtgattthtga | gtagagaggg | gthtaggggta | gattattthta | gthttgttat | 1740 |
| thtagthttht | gattthtgggt | aagthtattth | atthththttat | gthttatttht | thttthtgaaa | 1800 |
| atggggatga | tatgggtatt | thttthtatag | agthtgtgtg | agggthtaaa | aattthaatat | 1860 |
| thtgaagta | thttgtatag | aggtthggtat | atthttgtaat | aagthttgttg | tgattatttht | 1920 |
| tattthtgtht | tggththtga | atthttgthtt | thaaaaaatat | thttththttg | thttthttaat | 1980 |
| thttgththtt | thtttagtattth | thttthattthg | thttattthtt | atthtttgtht | thtgtthtthta | 2040 |
| ggaggggagg | gthtgggttht | gggagaaaaga | gagagthgatt | atagagattg | gaattthgtth | 2100 |
| gatagthtagg | gagthtthtgt | tatggttaata | taggaatatg | tagthttgtgt | atgthtagtht | 2160 |
| aataattata | tggthattth | gthtagthtagt | tataatagth | tattthaatag | thttthgggta | 2220 |
| thttatgtht | atattthtaatt | agthttaggat | thttgtgthtt | tagaaaatatt | thtggththtt | 2280 |
| gthtaggtgga | gthttatattth | aaattthattth | taggtthaaatt | thttththttg | thttgthtagt | 2340 |
| thttgththtt | aattththttta | gagthttggt | tggthttgga | gthttthtagth | taggaagthtt | 2400 |
| gagagthgaag | tgggtatttht | tgagggattth | tgattthagaa | thattthagga | aagattgtga | 2460 |
| tggtagattth | thttaaata | tgtattgata | thtgtgtgtgt | atgagthagat | aggtthggagth | 2520 |
| thtttagthta | thaaattthtt | thtatgththtt | gattattthtt | thttattthaaa | gthttthtgag | 2580 |
| gtgthtagag | aaaggattga | tattthatttht | atthgaggaag | thtgaggtthtag | tgaagaatga | 2640 |
| tgtthtgatt | aaaggtthaa | atgtgataga | gthtaggtthaaa | tgtaggtthtt | thtgtgthttta | 2700 |
| gtthttgata | thtttgthtt | atgtagtht | tgagthtga | gthttagaag | tggthttgat | 2760 |
| thaatgtgggt | tgaattthtag | thttthtattth | atthttattthg | atthttgthta | agthtattgtht | 2820 |
| thttthtggg | tagthtthtgg | ggaaaaggg | gagaattgtht | atagthttgtht | thttthaatgt | 2880 |
| gtgthttthtgg | aggtthtgggt | tgtthttgtht | atthataattg | gagthgtatgg | agthtaggtth | 2940 |
| gthtagagga | ggaagggthg | gtgthtgggta | thttgthtatt | gattththttta | agthtattgata | 3000 |
| gtgthttgtht | gthttgthtt | tatggattthg | thttgggaatg | gthtattgtht | aggtthttat | 3060 |
| gthtatttagg | atthttgtht | ggtthtattth | agthtattgta | thtgtgatttht | tgtgggttht | 3120 |
| gthtggthtt | ggtgthtttht | thttgtagtht | thtgtthtgg | gthtagggtht | tgtthttht | 3180 |
| atgthttgth | thttggthatt | atthgtthttat | gthtggtht | atthtatttht | thtatggtgag | 3240 |
| tagthttggg | thttgggata | atthtggtht | gggattggg | tgtthtgagth | thtggatttht | 3300 |
| thattththtt | gthttgthtt | aatgaaattth | tggagggagg | gggtthttag | gattthtthta | 3360 |
| gthtgagaatg | tgatgthtga | gtgattgtht | thttthtagtht | gaggtthttg | tgtthtggg | 3420 |
| tggthttgga | ataaatttht | ggtthtthaa | aatgggagg | atthttagtht | aattththttg | 3480 |
| agthtaggag | gattththttta | ggtthgggtht | thttthtagtht | thtagtgggtht | atthtaaggta | 3540 |
| thttththtt | thttthtaggag | tgtthttgat | ggtthgggagth | agthtthtgg | gggthttgtht | 3600 |
| tgtthtthtt | gthtttaggth | gthttatgggth | tggthtthtt | gthttgthttat | atgthtggtht | 3660 |
| thtatagtht | tgtthgaattth | tagthtthtt | atthtttht | gthtttattthg | tagthtattgtht | 3720 |
| tgtthttatgg | thtaggtgtht | gatgthttht | tagthttthtt | thattthggta | gggthattth | 3780 |
| atggtggtht | ggtgthtttht | gtgagthttht | agthttthttta | ggtthtthtgg | gthtthttht | 3840 |
| gtgagthttht | thtgtgthttat | gaggaatttht | thtggthttgtht | ggtattgtht | gthttgthttht | 3900 |
| tgggthttga | thttgtgtht | gattatgthta | thtgtgtht | thtggthtttht | tatggttht | 3960 |
| thttagthttht | gtaggttht | gtagatgtht | thttthtggtht | thttgggtht | gtggtgtht | 4020 |
| thtggthttga | ggattthttag | gtgagatttht | agggatttht | gatgggttht | gattggttht | 4080 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ggtggtttat | tattatgggt | ggaggtagaa | ggtgataaat | gtaggaagtt | gatttttttt | 4140 |
| ttagtttggt | atatttttatt | ggtgattttg | aggatgtggt | ttattttttg | ggtagatgt | 4200 |
| tgtttggtta | ttaaggttgt | ttggatttga | gtttgttagt | agtgtgtggg | tgtagtaat | 4260 |
| ttttggtagg | gattgaatta | tttttttttt | ttttgttgtt | atgtgtggt | tttatggggt | 4320 |
| agtttgtatt | tggggtgaag | agaagtataa | tagttattag | tgtaagattt | aaaatttttt | 4380 |
| ttttgtattg | tatagtgaga | tgtttagggg | tttagtttag | tgtttgata | tagtgatttt | 4440 |
| tgggtttggt | tgttgttgtt | ttaagtgaag | ttggtgtggt | ttgggtggag | tagatagaga | 4500 |
| ttttgggtgg | taggggtttg | ggaagatatg | ggtggttagg | gttttatgtg | tttttattgt | 4560 |
| tgttttttgt | tattttagg | taatggatga | gttgggaatg | agtttttttag | attagtttgt | 4620 |
| gattaagaaa | ggtaaaggaat | ggttgttag | ggtagagttt | ggtgagatgg | tgtaggtttt | 4680 |
| tggtgtatag | atttattttt | tttattgggt | gtgtttttttg | tgtgtgttta | ggttgggtgt | 4740 |
| ttattaggta | ttttttttgg | tttagttaga | tgtaggttag | atgtgtgggt | ttggtgagtt | 4800 |
| tgtttagtat | tttgtggttt | ggggtgggtt | ttagtggatt | agtatttatt | gggtttagt | 4860 |
| attgggagtt | tggttttttt | ttggtgaggg | ggaggtatt | tttgtggatt | tggagttgat | 4920 |
| ttgtagaatg | agttaaatta | tttttttgtt | tttttaagag | atgggaatgg | aagtgttgtt | 4980 |
| tttatggagt | tggggaatg | atttttattt | tatagtgttt | tagtattttg | gtgtttgggtg | 5040 |
| ggtatttttt | tttttttttt | tttaggttag | gttttgagg | tttttggggg | aatttttttt | 5100 |
| ttgtgggagt | tttttgtggt | atttagattt | aggggagttt | gtgtgtgagt | attgtgtggt | 5160 |
| aggttgtgtg | tattttagtt | agggtttatt | tgtttttggg | tgtttgtggt | tatgtgagtt | 5220 |
| tagggttttg | tgtatgtttg | aatgtttttt | tttatgggtg | tttttagtatt | ttttggagtg | 5280 |
| tgagtgtgtt | tgtttttgtg | aatgtgtttg | tgaggtgtgt | ttttgtatgt | tgggtgtgat | 5340 |
| ttttttgtat | ttgggggatg | tatatatttt | ttaatatgta | tagtattttt | gttgtgtttt | 5400 |
| gtattttggt | ttttggtatt | tgaggatttt | taagtatgtg | tgggtttttt | ttgtgtatat | 5460 |
| ataggagtat | ttatgtgatt | tttggtatta | gtaaaattta | gggatatggg | atttattttt | 5520 |
| tttggtttga | ggattaagta | ttggttatga | taggggaagg | tgagagatga | taaaaataga | 5580 |
| gagatagtta | gagaggagta | gagagttaga | gggttttagg | tattgggtag | tagttttttt | 5640 |
| atatttgggg | taggtgtttg | aaagaattta | gaggtgtata | tgagtttgag | gtgttttagg | 5700 |
| taggtattgt | ttttataggg | tttggtttga | gttgtttttt | aaatgagtga | atttaaagtt | 5760 |
| gggttttatt | tgttttttat | ttgttttttag | gggaggttaa | ggtggaagtg | gtggtagtag | 5820 |
| ggttgggggt | ggatttttag | gagttggggg | tgagttatta | ggagttgggg | gttgggtgga | 5880 |
| tgatttggag | tgtgtagtag | ggaagatgag | gtaatagggg | aggaagtggg | tggggggagg | 5940 |
| tggaaattggg | gttgtgtttt | gtgttgtttg | gaattgggag | tgtgggaaag | atattaggaa | 6000 |
| tttgggttga | gtgtagtttt | gttgggtggg | tttgggtggt | ttattgtata | gagttttttt | 6060 |
| tgatttttga | agaaagagat | ttgtttttag | tgggtaaaag | tttgtttgga | ttttttggtt | 6120 |
| attagaaata | tgagtatggt | ggtggttttt | agttttttat | ttatgtttgg | gtttaagaga | 6180 |
| ttgggagttt | aggtttattg | atttttttag | aaagattaag | attttgtatt | ttagaaagag | 6240 |
| gtttggggat | ttttgttttg | tgtaaaggga | gaaggattag | ttgttttttt | gagttttta | 6300 |
| atttggaatt | ttggtttgaa | gttgagatag | gagattggat | gtgaggtttt | tttagagttg | 6360 |
| gtttttttta | aataattttt | aaaattttta | gatttttaggg | gtatgttgaa | atttttttaa | 6420 |
| gtagtttaaa | gaatataatg | agagttttta | tatttttaggt | ggtggtgtgt | tggttttttg | 6480 |
| gagtgggggt | ggatgtgggt | gtgtggattt | atgtgtataa | ttgtgtggga | tggggttatg | 6540 |
| tggatttatg | tgtataattg | tgggatttta | gtgttagtgg | gatttttagtg | ttaggggat | 6600 |
| tttagtgtaa | gtgggatttt | agtgttagtg | ggatttttagt | gttagtggga | tttttagtgt | 6660 |
| agtgggattt | tagtgttagt | gggttttggg | tttagtggag | tgagtggagt | gttggtgatt | 6720 |
| tgagtggaga | ttgtgttttg | gatgttttag | tttagatggt | aagttatagt | ttgtgtagta | 6780 |
| gtagtaaagg | ggaaggggta | ggagtgggt | atagtggat | ttggaggtg | tgatttaggg | 6840 |
| gaaagtgtgg | gtggttgatt | tagggtagtt | gtggtggtga | ggtagggtggg | ttttttggtt | 6900 |
| tttggagttg | ttttttttta | tatttgtttt | tgggtgtttt | agtagttttt | attttgggtt | 6960 |
| tttgtggtaa | ttgtgggatt | tgggtgtggt | gttagtttag | tggggagtg | attagtggtt | 7020 |
| ttttttggtt | ttgggttttt | tgatggatag | aggaattttt | gttttgtttt | atagattttt | 7080 |
| ggtttttggt | gagtgtggta | ttggagtttg | ttttgttagg | gttttggaaat | tagagaaagt | 7140 |
| tgttttttgg | ttatttgaag | tgttggattt | ttatagtgtt | tttttagttg | ggtagggagtg | 7200 |
| gtggttgtgt | tgttgaagg | tggggttttt | ggtgtgaaag | ggaggtagtt | gtagtttttag | 7260 |
| ttttatttta | gaagtgggtt | ttgtattggt | gtggtgggtg | tttttgggtt | ttgattttgt | 7320 |
| tagtgttgtg | gggtagaggt | atttggagtt | tgtagggttt | agatttgggt | tggaaaagtt | 7380 |
| ttggtgattg | taggtaaagt | tttgggaggg | gtggttaggt | gaagttttgg | tgttttatta | 7440 |
| tatatttttg | ggttttatgt | tagttgtatt | tgtggtattg | ggtaggaaat | ggtagggttg | 7500 |
| aggttgattt | taggagtata | agggagtttt | ttattttttg | tttatatttg | ttatttttag | 7560 |
| ttttgtaatt | tatttttagat | atatagaaa | taagtaggat | tgggtggggag | atggagttta | 7620 |
| ataggaatat | tttttagtag | tgagtagggg | ttgtatggga | tgtgggagga | gttttagagga | 7680 |
| ggtgtggaga | gtgtttgagg | ttgggtgagt | gatttagagg | gagatagttg | aattgggttt | 7740 |
| aagaggtggt | tagtgggtgt | ttgttgaatg | aatgagtgat | gggttttgaa | ggttaggtgt | 7800 |
| attgaaagag | ggggtgtgta | aaaagggttt | tttttattat | ataggatata | gtatatgtaa | 7860 |

| | | | | | | |
|-------------|-------------|-------------|--------------|-------------|-------------|-------|
| atTTTTTTTT | tgtggaaaag | ttagataggt | taaaaaggtt | ataaataaat | tagttgggta | 7920 |
| tgggtggtg | tgttttagt | tttagttatt | agggaggtt | agttagggga | attgtttgaa | 7980 |
| tttgggaggt | ggagattgta | gtgagttaag | attgtgttat | tgtatttttag | tttggaaata | 8040 |
| gagtgagatt | ttgttttgg | aaaaaaaaa | aaaagttata | aattgtgtgt | gggttttagg | 8100 |
| ttatataaatt | agagttggag | gggagtggtt | aaggatgaga | attgagatgg | atTTTTtGtt | 8160 |
| ttttttggag | gagagtgggt | ggttgtttat | ttgggggtgg | ggaattttttt | tttatggggtt | 8220 |
| tagttgttta | atTTtagggg | atTTTTtagga | taggagttga | tgtaaatagt | tgTTTTatTT | 8280 |
| tttgttgttt | ttggttttgg | agaaggagga | gggagttggg | gagggttttt | atTTTTtaga | 8340 |
| taatttttaa | gtagttagga | tatgggtgag | atgagtgaga | tattgatttt | tgggatagaa | 8400 |
| tttgagaggg | tgtaaaaaa | tttagtaatt | aagataaata | ggttgggtgt | agtggtttat | 8460 |
| gtttgtaatt | ttagtatttt | gggaggttgg | attatttgag | gttaagagtt | tgagattagt | 8520 |
| ttggttaaga | tggtgaaatt | ttatTTTTat | taaaaatata | aaaattagtt | tagtgtggtg | 8580 |
| gtgtagttt | gtaatttttag | ttatTTtagga | ggttgaggtg | agagaattgt | ttgatttagg | 8640 |
| aggtagaggt | tgtagtgagt | tgagattatg | ttattgtatt | ttagtttgg | taatagaggg | 8700 |
| agattatttt | aaaaaaaaa | aaaaaaaaa | aaaaaaaaa | gaggttgggt | ggtgggtggt | 8760 |
| tatattatgt | gatttttagta | ttttgggagg | ttgaggtggg | tggattattt | gaggtttgga | 8820 |
| gtttgagatt | agtttggtta | atatggtgaa | atTTtGtttt | tattaaaaa | ataaaaaatta | 8880 |
| gttgggtggg | gtggtaggtg | tttGtaattt | tagttatTTT | ggaggttgag | gtaggagaat | 8940 |
| ttttgaatt | tgTggggTgg | aggttGtagt | gaattaaagt | tatattattg | tatttttagtt | 9000 |
| tggataataa | tagtaaaatt | ttgttttaa | aaaaaaaaa | atTTTTttt | ttgagataa | 9060 |
| gTTTTatTTT | ttgttttagg | ttgggtgta | gtattatgat | tttagtttat | tgtaattttt | 9120 |
| gTTTTtaga | ttttgtatt | ttagttttt | aagtagttgg | gattataggt | atTTgttatt | 9180 |
| atgTTtagtt | aatttttGta | tttttagtag | gggtgtggtt | ttattatgtt | ggttaggTtg | 9240 |
| gTTTTgaatt | tttgattttt | agtgatttGt | ttgttttagt | tatttaaagt | gTtgggatta | 9300 |
| taggtgtgag | ttattatgTt | tggTTTTttt | aaatgaaa | agtgtaaaa | tttatgataa | 9360 |
| ataaaatatt | aaaaatttat | tgaatttGta | ttttataat | tttttttat | ttgtttttta | 9420 |
| ggttattttt | tgTTTTtagaa | agtaatttaa | aaaatgtgta | gatggagttt | ggattttatt | 9480 |
| tgaaaatggt | gggagttatg | gaaaattttg | gagtagggga | gtgaaggata | gaaattatat | 9540 |
| gtaaaagaaa | ttttgggttg | ggtgtagtgg | tttatgtttg | taatttttagt | atTTtgggag | 9600 |
| gTtgaggtag | gtgattatt | tgaggttagg | agattgagat | tagtttgatt | aatatggtga | 9660 |
| aatgTtattt | ttattaaaa | tataaaaaa | attagttagg | tatggtggtg | tatgTttgta | 9720 |
| gTTTTagTta | ttttggaggt | tgagatagga | aaattgtttg | aatttgggag | gtggaggttg | 9780 |
| tagtgagTta | agattgtgTt | attgtatttt | agtTtgggta | ataagagtaa | aattttatTT | 9840 |
| taaaaaaaa | gaaagaaaga | aattttttgg | tagttgatga | gaaggaaatt | taattggtag | 9900 |
| gTTTTtagtag | gggagatgag | gagatttttag | ggaggttatt | tgtatatgTt | gtgTtttagt | 9960 |
| gtgggttagg | gagtaggtta | ttatTTTTt | tgtttatttt | ttttttgTtt | taattttttt | 10020 |
| aagTTTTtga | ttagtggTat | tttaagtGta | gtttaaggaa | ttatatgtat | taggattttt | 10080 |
| aggggtgTt | tgTtaaaa | gtaaaTTTTg | gttaggtgta | gtggtttata | tttGtaattt | 10140 |
| tagtattttg | ggaggttgag | gtgggtggat | tatgaggtta | ggagattgag | attattttgg | 10200 |
| taaataatggt | gaaattttat | ttttatttaa | aaaataaaaa | taaattaaaa | aaaatattag | 10260 |
| ttgggtgTgg | tggTgggtg | ttgtagtTtt | agttatttgg | gaggttgagg | taggagaatg | 10320 |
| gtgtgaattt | gggaggtgga | gtttgtagtG | agttgagatt | gtgttattgt | atTTtagttt | 10380 |
| gggtgataga | gtgagatttt | gttttaaaaa | aaaaaaaaagta | aattttttgg | gtattatttt | 10440 |
| atattgattg | tatgagaatt | tttGgaagta | tggTTTTgga | atTTgtattt | tttaattgatt | 10500 |
| ttttgaggta | aaatttatat | atTTaaaatt | tatatattgt | aagaatataa | tttgatgatt | 10560 |
| tttagtaaat | gtagagtTgt | gtaaattata | ttataaaaa | atTTattttg | aggtattttt | 10620 |
| gttattttta | gaagtTttta | tttGtagTta | ttttttattt | ttatttttaag | ttttaggtta | 10680 |
| ttattgattt | gttttttGtt | tataaaTTtG | tatattttgg | atatgtttta | taaatagaat | 10740 |
| atataaatat | gtagattttt | gtattaagtG | ttttttgTtt | agtataaat | ttttgaggtt | 10800 |
| tgTTTTaaat | attatgttaa | gtaaaagaaa | tagtaataaa | aagaaataaa | ttattattag | 10860 |
| tagtttGttt | ttttttattg | ttttatagta | tgtattgta | tgggtgtaat | atattttgTt | 10920 |
| tatttatTgg | ttagtTgatg | ggtatttggg | ttgtttttat | tttttagttg | ttttaaataa | 10980 |
| tgttGttatg | aatatttata | tatatgtTtt | attatagaaa | tatattttta | gtgTttttga | 11040 |
| gtaaattttt | aggagtGtaa | ttgttggatt | atatggttgg | tttaggttta | atTTttaaga | 11100 |
| aattgTtaag | ttatTTTTta | aagtggTtGt | attattttgt | atTTttatta | gtaatgtgtg | 11160 |
| agggTtttat | ttttttatat | tttatattta | gaattgtTtt | ttttttttt | ttttttttt | 11220 |
| gagatggagt | ttatTTTTtG | ttgtttaggt | tggagtGtaa | tgaggtgatt | ttggTttatt | 11280 |
| gtaaTTTTtG | ttttttgggt | ttaaagtatt | tttttGtttt | agTtttttga | gtagttggga | 11340 |
| ttataggtgt | gtgatattat | gtttggttaa | tttttGtttt | ttagtagaga | tggggTtttt | 11400 |
| ttatgttagt | taggttggTt | ttGaaatttt | gatttttagt | gatttattttg | ttttggTttt | 11460 |
| ttaaaatgTt | gggtttatag | gtatgagTta | ttatgTttgg | ttttttttat | ttattttttg | 11520 |
| ttttgagatg | gattttttt | ttgttattta | ggttggagTg | tagtggTata | atTTtggTtt | 11580 |
| attGtaattt | ttgttttttG | ggtttaagta | atTTttttgt | tttagTtttt | tttagatTta | 11640 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ggattatagg | tatgtattat | tatatttagt | taatttttgt | attttttagt | gagatgggg | 11700 |
| tttattatgt | tggatagggt | ggttggtttt | gaatttttga | ttttatgatt | tatttatttt | 11760 |
| agttttttaa | attgttggga | ttatagggtat | gagttattgt | gtttggtttt | tgtttttttt | 11820 |
| attataatta | ttaagtggta | ttttagata | gttttaattt | gtattttttt | aatgattaat | 11880 |
| aatgttaaga | atttttttat | gtatttattt | aattatttat | tttttatttt | tagtgaaatg | 11940 |
| tttatttata | tatttggatt | attttttatt | taggttggtta | ttttattggt | gagttataat | 12000 |
| agttttttat | ataatatatt | ttgaatatag | ttttttatta | gatatatgat | ttgaaaatat | 12060 |
| tttttttttag | tttgtggttt | attgtttttt | tttttatttt | taaaatttta | attttttttt | 12120 |
| tgatttttta | aaaattattt | tttttttttt | ttttatagaa | atagggtttt | gttatggtgt | 12180 |
| ttaggttgg | tttaaatttt | tggtttttag | tgattttttt | gttttggttt | ttttaaggtg | 12240 |
| taggtttata | ggagtggata | attgtttttt | tttttttttt | taattaaaaa | atttttttaga | 12300 |
| gataggtttt | attaagttgt | ttaagttgga | gtttgggtgt | tatttatagg | taagattata | 12360 |
| gtatattgta | tttttgaatt | tttgaattta | agtgattttt | ttgtttgagt | tttttgagta | 12420 |
| attgggatta | tatttttatg | ttattatatt | atgttatttt | tttaataata | ttttttgagg | 12480 |
| tgtaaaagtt | ttagattttg | atgaagtta | gtttattatt | attttttttt | tttttttttt | 12540 |
| tttttttttt | ttgagattag | ttttattttg | ttgtttaggt | tggagtgtag | tgggtgttatt | 12600 |
| ttagtttatt | gtaatttttg | ttttttgggt | ttaagtgttt | tttttgtttt | agttttttaa | 12660 |
| gtagttggga | ttatagggtga | gtgttattat | gttgggttaa | tttttgattt | tttagtagag | 12720 |
| atagttttat | tatgttgggt | aagttgtttt | taaatttttt | attttaggtg | atttgtttgt | 12780 |
| tttagttttt | taaagtgttg | ggatttttag | agttagttat | tatatttgg | ttattttttt | 12840 |
| tttttatggg | ttttgttttt | ggtgttgtat | ttaggaattt | tttgtttaat | ttgaggttat | 12900 |
| gaagattttt | tttttaagtt | ttttgttgtt | gttgttgttg | ttatttgtat | ttttagtaga | 12960 |
| gatggggttt | ttttatgttg | tttaggttgg | ttttaaattt | ttgattttta | atgatttatt | 13020 |
| tgttttgttt | tttaaagtgg | taggattata | ggtgtaagtt | aggatgtttg | gttttttatg | 13080 |
| ttttttttta | gaaaaaattt | gtaattttta | taaggttttt | tggagagttt | gttgtatatt | 13140 |
| tgtttaggtt | tagaagttag | attatattta | ttatttgtgg | atagagtgtt | ttgtttgata | 13200 |
| ggttgaataa | ttgggagttt | ttgtgtttgt | aggagtgtgt | gtgtgtgtgt | gtgtgtgtgt | 13260 |
| gtgtgtgtgt | gtgaatgggt | tttgtatatg | tgattagtag | gatattatta | gattatttta | 13320 |
| gttttaggag | ttattatagg | gaattttgtt | ttgtgtgtgt | gggagataaa | tttttagtgt | 13380 |
| tttgtttttg | tttgtgtgtt | agtttaattt | ttatttttat | tattgttttt | ttgtgattga | 13440 |
| tttgtttttt | taattatgga | ttttattttt | tttgttattt | ttaagggaag | aagttttgag | 13500 |
| aaaaaaaggg | ataggtgggg | gaagtagtaa | ggttagttat | ttatataat | aaatataatga | 13560 |
| gtaatgttat | ttttgttttt | atatttagta | gtatttatgt | ttgtataaat | ttggtattaa | 13620 |
| taggagaaa | tatttgtaaa | tatttatatt | taggttatta | tattttttta | tatttttagtt | 13680 |
| tatttttatt | gttgtattta | ttataagggt | atatagtttt | tgtatatagt | tatttttttt | 13740 |
| tatttttttt | gagatggagt | tttgttttgt | tgtttaggtt | ggagtgtagt | ggtgtaatta | 13800 |
| tagtttattg | tagttttaa | ttgggtttag | ttattttgtg | gttttagttt | tttgagtagt | 13860 |
| tgggattatg | ggtatatgtt | attatgtttg | gttaatttaa | aaaaattttt | ttttgtagag | 13920 |
| atgaggtttt | attatgttgt | ttaggttggt | tttaaatttt | tgggtttaag | tgattttttt | 13980 |
| gttttagttt | tttaaagtgt | tgggattata | ggtgtgagtt | attatgttta | gttagtttgg | 14040 |
| gttttaggtt | ttaatagtag | ttggatggag | aaggttgttt | attttttttt | tttaattttt | 14100 |
| atttttgggt | agggtgatg | ggaggaaga | tagttatttt | gttttttttt | ttatttttaa | 14160 |
| tttattaatt | ttttgttttt | gtttttgttt | ttgtttttga | gatagggttt | tattgtgtta | 14220 |
| tttaggttgg | tatggttagt | ttattgtaat | ttttgttttt | tgggtttaag | taattttttt | 14280 |
| attttagttt | ttagagtagt | tgggattgta | ggtatgtgta | attatgttgg | gttaattttt | 14340 |
| gtattttttg | tagagataag | gttttattgt | attgtttagg | ttggttttga | atttttgggt | 14400 |
| ttaaattttt | tattttggtt | ttttaaagtg | ttgggattat | aggtgtgagt | tattgtgttt | 14460 |
| ggttttattt | gtagttttat | ttgtttgaag | ttaagtttga | ggtataaatt | ttttatata | 14520 |
| atagtatagt | ttagtttggg | tttttagagt | ttttgtatta | ttagtttatt | ttatatttat | 14580 |
| ttatggttat | ttgtgttaag | attttgtttt | ttatataatt | atgtttggta | gggaagatat | 14640 |
| ttatggaaat | agattaggaa | ggtttagtgt | ggtaaagttt | gggatggggg | tgagagggag | 14700 |
| ttggggggta | ggggagagaa | gttttgtata | gggttgggtg | tttaaattta | gtagagtggga | 14760 |
| gtaggatttt | tgagaattga | aggagtattg | tggggtttaa | agagtttata | ttttgggagt | 14820 |
| tattatattt | taggaggtat | tttatatttg | ttttatttgg | tttgtgttgg | gtttttaaat | 14880 |
| ttttgaatta | ttgttaatgt | ttaaaatttt | agagatttta | tagaaaaatt | tagaatttta | 14940 |
| gttttaaaaa | ttagtaaaat | aaaataaaaa | tataaataaa | ttaggagttt | tgtgttttgt | 15000 |
| ttgggtgggat | tttgtataag | gtgttttttt | agtttgttaa | ttttttttat | agtttattta | 15060 |
| ttatgttatt | tgttgagttt | tttttttttt | gaagtttggg | gttttgttgt | ttattaagtg | 15120 |
| gttattatag | ttatgagtgg | tttttgagta | tttaaaatgt | gattagtttg | aattgagatg | 15180 |
| tgtttaagta | taaaatgttt | attggatttt | gaaggttaat | tttaaaataat | gtaaaatatt | 15240 |
| ttattaataa | tttttaaaat | attgattata | tgttgaaatg | attttatttt | taataatttg | 15300 |
| ggttaaataa | aatatattaa | atagtaaat | attatttatt | attatattga | ttaaaattaa | 15360 |
| ttttatagt | ttttttttat | tgtttaaaat | gtgggttagg | tttgggtggt | tattttttta | 15420 |

| | | | | | | |
|--------------|--------------|---------------|--------------|---------------|--------------|-------|
| atthtagtat | tttgggaggt | taaggtaggt | aggtggatag | tttgagttta | ggagtttaag | 15480 |
| attagtttgg | gtaatatggt | gaaatthttat | ttttatgaaa | tatataaaaa | ttatthtaggt | 15540 |
| atgggtggat | gtatthttag | ttthtagtaat | ttgggaggtt | gaggtgggag | gattgtttga | 15600 |
| gthtgggagg | tggtggttgt | agttagttgt | gattgtatta | ttgtatthta | gthtgggtaa | 15660 |
| tagagtaaga | ttthgtthta | aaaaaaaaaa | tggtggatat | taggaaatgt | aaatgtatat | 15720 |
| atgtggttha | tattatattt | ttthtggtha | gthtgggttt | agagthtaat | ggataaatgg | 15780 |
| ggatthtagta | agaggggtag | gtagagttag | aatataaggg | aggggtgaagg | aggttattgt | 15840 |
| aatatgatgg | ttaaagttgt | gaatthttgga | gthtagtthtt | agagttgaag | gthtattthtt | 15900 |
| gthtthttatt | tgthtthttta | thtttaattth | tttgaattth | tagtattthgg | tattthattga | 15960 |
| taggtgaaat | tatagttggg | atthattgat | aggtgatatt | atthtthttta | gthtthggagt | 16020 |
| tgggatttga | ttthtaggga | attgattggg | ttthgatgag | ttthaggttg | agthtthgggt | 16080 |
| tgtagaatga | tgtatthtaga | tggaaagatt | ttthtggatt | ggggtagggt | gattagagta | 16140 |
| aggtthtggg | agaaatthtat | atgggaaggg | aatthtgtgt | tgtgtgtatg | tagaggagtg | 16200 |
| gthtatata | aggtatataa | agaaatgaaa | gthttaggat | gggtagggtt | gggagaaata | 16260 |
| aaaaagaaa | gthttaggat | ttagttaggg | atgggatgtg | gtgatggatg | gggtgtgggtg | 16320 |
| atggatggga | tgtgttgggg | gagagggttg | tagagggat | tatagaggtt | thtgggtggg | 16380 |
| ttgttgggtg | gaggtgggga | ggtthattta | gagagggaga | gattagatgt | thtgttaggt | 16440 |
| tatggattag | gthtagthttt | thtattthttg | tattthtttag | thtthggatt | atthtthggta | 16500 |
| tgaagtgtht | tagthtaaat | tgtthaagtg | thtgaaggtt | aatagtattg | thtthtgagat | 16560 |
| tgtagthttg | ggthtthttat | tggthtttht | thttagttt | gtgtgattth | aggtatatta | 16620 |
| thtthtttht | tgtgatttga | thtthtttat | thtgaatgtg | ggaatagtga | taagtgtagt | 16680 |
| thtthtggg | ttgtthtaag | gattgagtg | gthtatatat | gggagtatat | agtagttagt | 16740 |
| gthtagtagt | tgttagthttt | agattthtata | thttagthttt | atthtthgatg | tatatatgag | 16800 |
| gthtthgtta | thtaaggttgt | tattthtttht | aaggaagaaa | aggtthaaagg | gthtthgtta | 16860 |
| agaagtaagt | gthtthgat | thtthttthtt | thttaggtta | agtaggagga | tggaggttag | 16920 |
| atthtaggaa | gaatthtttht | thtthttatag | thtaagaatgt | thtaagtgaat | thtagaaaaat | 16980 |
| taatthttgg | ttgggtatgg | tggthttatgt | ttgtaattth | agtattthttg | gaggttgagg | 17040 |
| tgggtggatt | atgaggttag | gagthttgaga | ttagthttggt | taatatgggtg | aaatthttatt | 17100 |
| thtattaaaa | atatagaaat | tagthtaggtg | tgggtgggtgg | tattthttagt | thttagttatt | 17160 |
| taggaggttg | aggtagaaga | atthgtthgaa | thtgggaggt | agaggttgtg | gtgaggttag | 17220 |
| atggattht | tgtatthttag | thtgggtgata | gagtgagatt | thtthttaaa | aaaaaaaaaa | 17280 |
| aaaaaaagaa | agaaagaaag | aaaatthaaat | thtthtagatt | thtthtttgggt | thtthtagataa | 17340 |
| ttagtagata | attagggagt | gthtggggag | ggggtatagg | gthtattthtt | tggthttthtt | 17400 |
| tagtatttht | agagtgattt | ggttaggaaga | gagggtttag | gthtgaaggta | thtaatagggt | 17460 |
| aataatthtt | ttgtthtttag | gatgtggagt | thtatagthtt | taggtthtttt | attgatgtag | 17520 |
| gggtgaggtg | gggtgggagg | gtagagagga | aaaatgtggg | aattatgaat | gggaagaggg | 17580 |
| gatttgggga | thtthttgtga | ttgtthtagtt | ttggthttgat | thtgtttag | ggtthttagg | 17640 |
| gthtthgtth | tagthttatat | thtatattthtt | atattthtaggt | thtthttthtga | tggatatttht | 17700 |
| taaaatatta | agattthttht | ggagthtthtt | gagthtthaa | taggatttagt | thtggthttagt | 17760 |
| gagggggata | tagattthaa | agthtagataa | thtggagthtt | agthtthttgag | thtthatttht | 17820 |
| ttagttgagg | gaatthtgggt | gagthtattth | tattthattth | tggthtttagt | thtthttthtt | 17880 |
| taaagtaagg | atattagatt | thtthttthttag | tggggttgtg | tgtggttagt | attggtthtag | 17940 |
| tgtthttthtt | thtthttatag | tgtthattthtt | thtthaaagtt | agttagthtttt | tagthttthtt | 18000 |
| thtggthtaa | gatagthttg | tgggtggtat | tataaattth | ggtgaggtth | thtthttthtgt | 18060 |
| gthtgggttg | thtthatttht | ttagthtthaa | gthtthaaat | aaggtgatgg | tgggggaggt | 18120 |
| gtagtggtga | thtthgtthtt | ttgagthtth | taatgthttth | thtthtggggg | gtgthttthttg | 18180 |
| taggtaggaa | ggttagthtt | taatthtttht | thttagthtga | thtthtggthtt | thtthaggthtt | 18240 |
| thtthaaagta | thtagagagga | ggttagthttt | atthtagthtt | thtthgtthttag | thtthttthggg | 18300 |
| aggggaggtt | ttggggattt | tggattgtgt | atgthtggg | ggtthgtthtt | ggtthttthgg | 18360 |
| ggtgthtata | tgtthtggg | thtthgggttag | ggaggtgggt | agagaaagaa | gthtgggtat | 18420 |
| agagaagga | tggagaattt | atthtatthtt | gttaggtga | gattgggtth | gggaagaga | 18480 |
| gaggttatt | thtthattthtt | thtthtaggaa | atgaaatthtt | ggattthgag | thtthatttht | 18540 |
| gtggtagtat | aggtthttgt | tgtthttggg | tgtggagtat | ggttagthtt | ggtthattthtt | 18600 |
| thtagattth | tgtthttthtt | thtthtgtga | tgtthtttht | thtthgggaagg | gtattthttgt | 18660 |
| thtagggth | ttagthttgga | ttagthttgag | gatggtgtgg | atagthttht | gtgtgtgtat | 18720 |
| gtgthttatag | tagtggatgg | ttaggagtat | gtgthtttht | gagthtthgag | atthtthttht | 18780 |
| thtthtttagt | ggatgtgtgg | thttaggtgt | gthtthgtthtt | gtgtgtgtgt | gtgtgtgtgt | 18840 |
| gtgtgtgtgt | gtgtgtgtgt | gtgtgtgatt | gthtthttht | ttggagthtag | thtatatggat | 18900 |
| gthtatatat | ttgtgttatt | gtgtgggagt | ggtatgtgtg | tagthtatagg | ggttattthtt | 18960 |
| thtthttthtag | tgtthttggag | gattthtttht | agthttthtthg | tgtthttthtata | tgtgtthttth | 19020 |
| ggattthttat | ttgggtgtgat | thtthtthtt | tgtthttthtt | ggtgtthttgag | thtthtthtt | 19080 |
| tgtatgthtt | thtthattthg | thtthatttht | thtthattthtt | thtatgtgtgt | gtgtgtgtgt | 19140 |
| gtgtgagaga | gthtatttagt | ggtgtgtgtgt | gtattthttht | thtthttatgt | gthtthttthtt | 19200 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ttgtatgtta | ttgtgtgttt | ttatatatag | gtttgtgtat | tttttggtag | gtgtgtgtgg | 19260 |
| ttttttggtg | tgtaaggtt | tttaggatgt | tttatttttg | ttgtttttgt | tgttgagagg | 19320 |
| attagagtga | attgaggggt | ttttggagtt | gggtgttgg | ttttttgttt | ttgttttgg | 19380 |
| tgtaaaatta | tggagattag | ttgtgttttt | gttgaaagag | atttttagagt | atttttagga | 19440 |
| ttaattttaa | attttaatag | gggttgggtg | aggtaataag | gttttttgg | agttatttag | 19500 |
| tttttatgtg | ttttttttgt | ttttatagtt | gtatgtttgt | ttgtttttgt | gtatgttaag | 19560 |
| ggtaagttt | aaaaatgggg | gttggtgttt | gtgggaatgg | ttagagaaga | ttgggttaa | 19620 |
| tttgttagtt | ggttgttga | aggttgttaa | tgtgatgttg | ttgattttgt | tagtggaagg | 19680 |
| tgagtgaggt | ttgaaggtgg | gtatgggagt | ggatgtgagg | gtaagtgagt | tgtttgggtg | 19740 |
| gattgttttt | aatgtaggag | gggtgtgtta | gtggtgtgtg | tgattttgtg | taagtgttg | 19800 |
| agtattttgtg | tgtgtatttt | ttgttatgta | taaggtgtgt | ttgtgatttt | tgtgtgtagt | 19860 |
| ttttgtatat | gtgtgtatgt | atgtgtaatg | gagatttatt | tgtagtgtgt | ggatgtgtat | 19920 |
| gtgggtgtgg | gtgtgggtgt | gtgggtgtgt | gtgttgggtg | ggagtggta | gtgtttggat | 19980 |
| gtggtaaagg | gatgaagggg | gtgatagttt | ttttttttgg | aggttgggtg | ttttggaata | 20040 |
| tatggtgggt | ttagttttta | tttttgagtt | tagagttaa | ttttaatggt | tttaaggtgg | 20100 |
| ggtaatggaa | tgggggttgg | agttgggttg | ggaaggaata | ttagaggggg | tatttatata | 20160 |
| ggtaggatgg | ggtaggatat | ttttgattag | ggggttagag | gtttaaagtt | gtgttttagg | 20220 |
| tgtagaggta | tattttgtag | tagtgatagt | tttttttttt | ttaggagagg | ttgtgtattt | 20280 |
| ggtttgagat | tttggttatg | tttgtgagat | ggagttttta | gttaaggtag | ttgttgagta | 20340 |
| tttgtaagta | tagtatgttg | atttggggga | gttgtatagt | tgtaaagata | tgttgttgg | 20400 |
| tgtaagtgta | gtgaggtat | tttgtatagg | tatatgtgg | tgttatgtat | agatagatat | 20460 |
| tatgtatggg | tataatggat | attattgtgt | atatgagtag | tgggtatata | tgtgtatttg | 20520 |
| tgtattattg | tgtatatgag | tagtgggtat | atatatatat | gtgtgtatta | ttgtgtatat | 20580 |
| tagtagtgag | tatatatgta | tgtgtgtgtt | attgtgtatg | tgagtgtgg | gtatatatat | 20640 |
| gtatgtgtgt | gttattgtgt | atatgagtag | tgagtatata | tatgtatttg | tggtttatgt | 20700 |
| ggtattatat | atgggtatat | ttaggttttt | tgtttaggtt | tttttgggta | tatgtgtgta | 20760 |
| tgttatttgg | gtatatttgg | gtgttatata | tgggtatatt | tgggtattat | gtttattatt | 20820 |
| tagatatata | tgaatatttt | ttataagttt | ttggagtata | tatagatatt | atatgtgaat | 20880 |
| atatttgggt | ttttgtatta | ggaatagtga | tgttatatat | tggtttatat | aggtttatgt | 20940 |
| atgggtttta | tatttatata | tatttttagg | atattattag | ttttattttt | tattgtagg | 21000 |
| ttgaggtttt | tttttttagt | tttgtttatt | aattttttgt | tttagtgatg | tttatttagg | 21060 |
| tattaagtta | tatttgattg | ttttgtaaat | tttattgtgt | tttttttatt | attgtttata | 21120 |
| tgttgttatt | tttatttagg | tttttatttt | atttttggat | ttttatttta | gagtgttttt | 21180 |
| gtttttaatt | tttttgtatg | tgtattttgt | ttattgtata | agaattttta | atggtttttg | 21240 |
| ttgtattgat | ttttaaatta | ggtttgtgag | ggtatttttag | ggtttgttta | ataaattagt | 21300 |
| attttattaa | gatttagttt | tgtttttgtt | ttttgagatg | gagttttgtt | tttgtttttt | 21360 |
| agattggagt | gtagtgggtg | aatttttagt | tattgtaatt | tttatttttt | gggtttaagt | 21420 |
| aatttttttg | tttttagttt | ttgagtgtt | gggattatag | gtgtttgtta | ttaggttttag | 21480 |
| ttaatttttg | gtatttttaag | tagagatggg | gttttattat | gttggttagg | ttggttttga | 21540 |
| atttttggtt | ttaggtgatt | tatttgtttt | ggatttttaa | agttttggaa | ttataggtgt | 21600 |
| gaattattat | ttttggtttt | aagatttagt | tttttaggtt | aggtaaggtg | gtttatattt | 21660 |
| gtaatttttag | tatttttgag | aggttgaggt | tgggtgattg | tttgagttta | ggaattggag | 21720 |
| attagtttgg | gtaatatggt | gaaatttgg | ttttagtaaa | aatataatag | ttaggtgtgg | 21780 |
| tgggtgattt | ttgtagtttt | agttatttgg | gaggttgagg | tgggaggata | atttgagttt | 21840 |
| gggaggttaa | ggttgtagta | agttatgatt | gtattgttgt | attttagttt | gtgtgatata | 21900 |
| gtgagatttt | attttaaaaa | aaaaaaaaagg | tgaagatttt | agttttttta | aaagagtttt | 21960 |
| tgttttagtt | tatttgggtt | gtataaataa | aatattataa | attgggtagt | ttataaataa | 22020 |
| tagaaatttt | tatttattta | tttattgaga | tagaattttg | ttttgttgtt | taggttggag | 22080 |
| tgtagtgggtg | taatttttagt | gtattgtaat | tttttagttt | tgggtttaag | taattttttt | 22140 |
| gttttagttt | tttgagtgtg | tgggattata | ggtgtttgtt | attatatttg | gttaattttt | 22200 |
| gtattttttg | taaaggtttt | attatgttgg | ttaggttgg | tttaaatttt | tgtttttata | 22260 |
| tgatttattt | gttttgggtt | ttttaagtgt | ttgagttata | ggtgtgagtt | attatgttta | 22320 |
| gttaataaga | gatatttatt | ttttatagtg | ttggaggttg | ggaagtttaa | ggtaaggta | 22380 |
| gattttgtgt | taggtgaggg | tttgtttttt | ggtttataga | tagtattttg | ttgtgtgttt | 22440 |
| ttatattgta | gaagggagaa | gggttatttt | tttggttttt | tttataataa | gggtattaat | 22500 |
| tttatttatg | atggttttat | ttttatgatt | taattatttt | ttaaagtttt | tattttttaa | 22560 |
| tattattatt | tagtggggtg | gaattttaat | atatgaattt | tggggaatat | aagtatttat | 22620 |
| attatagtag | ttttattgta | agaaaagata | gagagaggag | gggatagag | aaagagagaa | 22680 |
| agaaagagag | agagattgag | tgtagtgggt | tatgtttgta | attttagtat | tttgggaagt | 22740 |
| tgaggtgatt | agattgtttg | agtttaggag | tttaaggtta | gtttgggtaa | tttgggtgaa | 22800 |
| ttttatttgg | attaaaaaaa | aattagttag | gagtgatggt | gtgtatttgt | agttttagtt | 22860 |
| atttgggagg | ttagagttga | gattgtagt | agttgggatt | atattattgt | attttagttt | 22920 |
| gggtgataga | gtaagatttt | gttttggaaa | aaaaagaaaa | gagagagata | aatttgagaa | 22980 |

| | | | | | | |
|-------------|-------------|-------------|------------|------------|------------|-------|
| ttattagttg | gtaggttaa | gtttaaat | ttttgagggt | tttatagttt | ggtttatata | 23040 |
| aattttat | atatttgta | atattgttt | tatattttg | ggttgtttt | tggtttagaa | 23100 |
| tatttagttt | ttttattta | tatagtttt | taaattgata | atatttttt | ttgtttttg | 23160 |
| aggtttttt | taagtgtta | aggtagaaga | ggaagtatag | atatttggga | ttgttaaag | 23220 |
| gagagatttt | tgaggttgt | tatattgtg | ttggagaaat | ttagttagga | tggggtttat | 23280 |
| tttgtttatt | tattagttta | tttattaagt | atattgtggg | tttagttat | gtgtaagata | 23340 |
| ttggttggtt | tttatggtga | tataaagatg | aatgatgagt | gattttttt | tttgagttat | 23400 |
| atataggatg | tataaaggaa | tgaataatga | atggttgtaa | attagagtta | ggtttaaatt | 23460 |
| ttagttgggt | tattttgata | ggttattta | tttgttgaa | attagttttt | ttgtttgtat | 23520 |
| gatggggata | ataatagtat | agatttgta | tgaagattaa | gatgatgtag | atagtgtggt | 23580 |
| tggatatagta | tttagtttt | aattatattt | attgttagtt | attattttt | gtagattttg | 23640 |
| gattgttttt | tgatagtttt | aaagataaag | ttttgtgaag | tattgagtgt | ttgggggtgg | 23700 |
| agaggggaatt | taggaaagtt | ttttatata | ggtagtaagt | agttgagttg | tattttgaat | 23760 |
| gaatatgggt | tagatattta | aaaaagaaat | attatagggt | gattgatatg | tgttgtttat | 23820 |
| tagttgtgtg | atattgggtt | agttaaatga | ttttttgaa | tttagtttt | ttttttgta | 23880 |
| aaaaatagta | tggttattat | ttgtagattg | ttgggagggt | tagtaattgt | gtagttaatg | 23940 |
| tttgtaaaat | gttttagtatg | tagtaggttt | ttttttgggt | ttagattttt | gtttttta | 24000 |
| ttttgatttt | ttttttgta | ttaggtagat | ttgtaaggag | ttttagatt | tgatggttta | 24060 |
| ggattgttta | ttgttgggta | atagttgttt | agtatttatt | ttggagttt | gagtatagag | 24120 |
| ttgtttgata | tattttagtt | ttattattta | tggttttggt | gggtttgta | tttgtgtgt | 24180 |
| ttttattggt | ttttagaatt | atattgttga | gttatttaag | gggttgata | agatgttttt | 24240 |
| aagtagtgtg | ggtagtgggt | atggtgaaat | taaggttttg | gagaaggatg | ttaagtattg | 24300 |
| gaaataattg | ttttttttat | tttattttta | aggggttttt | aggttttgaa | atagggattt | 24360 |
| agtttttggg | ggtaggtttg | gaaggattga | aaggtgggat | tagagttagg | ttagaaagag | 24420 |
| aatatttatt | tagagatttt | agagttgggg | atattgtttg | gagtaaggga | gggtggtttt | 24480 |
| tttgtggtgg | tgtgttggtta | agttaaagggt | ttaggtattt | gttttatgtg | tgtaattttt | 24540 |
| tgatttttga | tggttgagaa | gggtttggat | agaaaattga | tatgaaaaga | tttggtttat | 24600 |
| gggttagagt | tttttttatt | agtgtggtt | gggtgtgtg | gggtttttta | gaggttaaag | 24660 |
| tttttgtgtt | ttttattggt | ggtaggagga | aaattgataa | attagaggtg | ttattgagga | 24720 |
| gttggtgttt | tttatttttag | aattttttat | ttttagaaaa | gggtgtttg | aaggaggttt | 24780 |
| tagtggattt | tttgtatttt | ttttattttt | agagagaagt | gggtaggagg | ggtttttaag | 24840 |
| gaataaagaa | gataaagtat | aaattgtaga | atttgaattt | aggtgtttgt | ttattatagt | 24900 |
| tttgtgtgtg | tttgttttat | ttatttatgt | atgttgta | taaatttgaa | attttaaaa | 24959 |

<210> 8

<211> 24959

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 8

ES 2 597 845 T3

| | | | | | | |
|-------------|------------|------------|------------|-------------|------------|------|
| ttttaaaatt | ttaaatttaa | ttataatata | tataaataaa | taggatagat | agtaatagga | 60 |
| ttatgatgag | tagtagtttg | gatttaagtt | ttgtgatttg | tgttttattt | tttttgtttt | 120 |
| ttgaggattt | tttttgttta | tttttttttg | agagtaggaa | gggtataaag | agtttattgg | 180 |
| ggtttttttt | tagtattttt | tttttggggg | taggagattt | tggaatgagg | ggtattaatt | 240 |
| tttttagtagt | atttttgatt | tattaatttt | ttttttgtta | ttagtgaaaa | atataaaggt | 300 |
| tttggttttt | agaagtattt | atggttattt | agttatgtta | atggaaaggg | ttttgtttta | 360 |
| tgagttagat | ttttttatgt | taattttttg | tttaaatttt | ttttaattat | taaaggttag | 420 |
| aaaattgtat | atatgagata | aatatttggg | tttttaattt | attaatatat | tattatagag | 480 |
| aggttatttt | tttttatttt | aagttagatt | tttaattttg | ggattttttg | atgaatgttt | 540 |
| tttttttggg | ttgattttaa | ttttattttt | tagttttttt | aggtttattt | ttaagagtta | 600 |
| agtttttatt | ttagggtttg | ggagtttttt | agggatgggg | tgggagaagt | agttattttt | 660 |
| gatgtttggg | attttttttt | gaagttttgg | ttttattatg | tttattgttt | atattgttta | 720 |
| gaaatatttt | gttttagttt | ttgagtgatt | ttagtaaata | gttttggaag | gtagtgaggg | 780 |
| tagtatagat | ggtaggttta | ttgaagttaa | gggtgatgag | gttaaagtgt | gttaagtagt | 840 |
| tttgtatttt | gggttttagg | atgagtgttg | gggtggttgt | gttttagtgg | gagtggtttt | 900 |
| gagttattaa | gtttgtaaat | tttttgtaga | tttgtttggg | gtgagaagga | gggttagagg | 960 |
| ttggaggata | aaagtttgag | tttagagaag | gatttgttat | atgttaggta | ttttatagat | 1020 |
| attagttata | taattgttga | tttttttaat | agtttgtaaa | tgatgattat | attgtttttt | 1080 |
| atagaaaagg | aaattaaggt | ttagagaagt | tatttgatta | gttttagtatt | atatagttaa | 1140 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| taaatagtat | atgttgatta | ttttgtaatg | tttttttttt | ggatgtttgg | tttatattta | 1200 |
| tttaaagtgt | agtttagtta | tttattat | tatataggaa | at | tttttttttt | 1260 |
| ttattttaaa | tatttaatgt | tttatagaat | tttgtttttg | aagttattag | gaaataat | 1320 |
| agaatttggt | ggaaataata | attagtagtg | aatataattg | aagtttggat | gttgtgttaa | 1380 |
| ttatattggt | tgtattat | taatttttat | aataagttta | tattattatt | at | 1440 |
| tatggataag | aaaattggtt | ttggataggt | tggatgattt | gttaagggtga | ttgagttggg | 1500 |
| atttgaattt | agttttagtt | tgtagttatt | tattatttat | ttttttgtgt | gttttgtgtg | 1560 |
| tggtttaggt | aagagaatta | tttattat | at | ttttttgtgt | tattatgaaa | 1620 |
| at | ttttatata | tgattgagat | ttataaagtg | tttgatgaat | gagttgataa | 1680 |
| tgagttttgt | tttggtagg | tttttttagt | tatagtatga | gtagtttttag | gtagtttttt | 1740 |
| at | tttggtagt | tttaggagtt | tgtgtttttt | tttttgtttt | gggtatttga | 1800 |
| aaaggtagga | gggtggagtt | gttagtttga | aggattgtgt | taggtgggag | gattggatgt | 1860 |
| tttaagttgg | aaataaattt | aaggatgtga | ggatagggtt | ggtagaagtg | ggtggagttt | 1920 |
| gtgtgggta | gattgtgggg | tttttaaaag | agtttgaatt | ttagtttatt | aattggtagt | 1980 |
| tttagattt | at | ttttttttt | ttttgagata | aggttttgtt | ttgttattta | 2040 |
| ggttggagt | tagtgggtgtg | at | tttagttt | attgtagt | taatttttagt | 2100 |
| gttgggatta | taggtgtgta | ttattat | tggtaatt | tttttttagta | tagatgggggt | 2160 |
| ttgttaagt | tgtttaggtt | ggttttgaat | ttttgggtt | aagtgattt | attat | 2220 |
| ttttttaaag | tgttgggatt | ataggtataa | gttattgtat | ttagtttttt | tttttttttt | 2280 |
| tttttttttt | tttatttttt | tttttttttt | gttttttttt | gtagtgggat | tgtattgggtg | 2340 |
| tgaatgttta | tgttttttaa | aatttgtatg | ttgaaatttt | at | tttttatttag | 2400 |
| taggaggtga | ggattttggg | agggtattag | gttatggagg | tggagtgtt | atgaatggga | 2460 |
| ttaatgtttt | tgttataaaa | gaagttagag | agtaattttt | tttttttttt | ataatgtgag | 2520 |
| gatataat | aaggtgttgt | ttatgagtta | gaaagtaggt | ttttatttga | tatggaattt | 2580 |
| gttttgattt | tgaatttttt | agtttttagt | attatgagaa | ataaatgtt | tttgttgggt | 2640 |
| gggtgtgggtg | gtttatattt | ataatttaag | tatttaggga | gtttaaggta | ggtggattat | 2700 |
| gtgaggttgg | gagtttgaga | ttagtttgg | taataggtg | aaatttttat | taaaaatata | 2760 |
| aaaattaatt | agatgtgggtg | gtaggtattt | gtaatttttag | ttatttagga | ggttgaggta | 2820 |
| ggagaattgt | ttgaatttag | gagttggagg | ttgtagtga | ttgagattgt | attattgtat | 2880 |
| tttggtttgg | gtaatagaat | aagattttgt | tttaataaat | aaataaataa | aaatttttgt | 2940 |
| tgtttataag | ttatttagtt | tatggtat | tgttatagta | gtttaaatgg | attaagataa | 3000 |
| gaattttttt | ggaagaatta | aattttttgt | tttttttttt | ttttgagatg | gggttttatt | 3060 |
| atgttatgta | ggtagagta | tagtgatgta | attatggtt | attgtagt | tgatttttta | 3120 |
| ggtttaggtt | at | tttagtttt | taagtagttg | ggattatagg | agtgtattat | 3180 |
| tatattta | tgttgtat | ttattagaga | ttgggtttta | ttatgttgtt | taggttagtt | 3240 |
| tttaattttt | gggtttaagt | gatttattag | ttttagtttt | ttaaaagtgt | tgggattata | 3300 |
| gatgtgagtt | at | gtttgaaaaa | ttaaattttg | gggttgggag | tgggtggttta | 3360 |
| tatttgaat | tttagaattt | tgggagtta | agggtgggtg | attat | tttaggagtt | 3420 |
| tgagatttagt | ttggttaata | tggtgaaatt | ttatttttat | ttaaaatatt | aaaaattagt | 3480 |
| tgggtttggg | ggtagggtt | tgtgatttta | gttatttggg | aggttgagat | aggagaattg | 3540 |
| ttgaattta | ggaggtggag | gtttagtga | gttagattg | tgttattgta | ttttattt | 3600 |
| ggaaataaga | gtgaaatttt | at | ataaaaataa | aattaaattt | taatagaata | 3660 |
| ttagtttgtt | aaataggtt | tgaggtgtt | ttgtagattt | agtttgaaaa | ttaatgtagt | 3720 |
| gggaattatt | gaagattttt | atgtagtggg | gtgggtgtat | gtgtagggag | gttggaata | 3780 |
| gggatgtttt | gggtgaagg | tttaggagtg | agatgaagat | ttgagtagag | gtggtagtat | 3840 |
| gtgaatggta | gtgagggagg | tataatggga | ttttagaat | agttaggtgt | gatttgggtg | 3900 |
| ttttagggat | gttattggag | tagagggtt | atgggtaagg | ttagaagagg | gagttttaga | 3960 |
| ttttaggtg | ggagtaggtt | tgggtgtt | ttgggagtgt | atgtgagtgt | aggatttatg | 4020 |
| tgtgagttt | tgtgagtttag | tatgtgatat | tattgttttt | gatgtaggag | tttagatatg | 4080 |
| tttatatgta | gtgtttatgt | gtgttttag | agtttgtgaa | gagtgtttat | gtatgtttga | 4140 |
| gtgataaata | tgtatgttaa | atatatttat | atgtggtatt | taagtgtgtt | taggtggtat | 4200 |
| gtatataat | gttaggaaa | gtttaagttag | ggagtttagg | tgtgtttatg | tgtggtgta | 4260 |
| tatgggttgt | aggtagtgt | gtgtgtttat | tgtttgtgta | tatagtgggtg | tgtatgtatg | 4320 |
| tgtgtgtgtt | tattgtttat | gtatatagtg | gtgtatata | atgtgtgtgt | ttattgttga | 4380 |
| tgtatata | gggtgtatg | tatgtgtgtg | tgtttattgt | ttgtgtatat | agtggtgtgt | 4440 |
| aaatagtat | gtgtgtttat | tgtttatgta | tatagtgggtg | tttattgtgt | ttatataatga | 4500 |
| tgtttgtttg | tgtatggtat | ttatgtgtgt | ttgtgtaggg | tgtttttatt | tatttggtag | 4560 |
| ttagtagtat | gtttttgtgg | ttgtgtagt | tttttgggtt | agtgtgttgt | tggatataggt | 4620 |
| at | tttggtagt | tgttttgggt | gggaattttg | ttttatagat | gtaattgaag | 4680 |
| ggtgatggg | tttttttggg | gggaggggag | ttgttattat | tgtaggggtg | at | 4740 |
| ttggaatata | gttttgaatt | tttgattttt | tggttagggg | tattttgttt | ttttttgttt | 4800 |
| gtgtagatgt | tttttttgat | gttttttttt | agtttagttt | tagtttttat | ttttattgtt | 4860 |
| tattttgggg | ttattaggg | tgggttttga | gttttagagat | gggagttggg | tttattatgt | 4920 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| atthtagagg | tattaatttt | taggggaggg | ggttggtatt | ttttttattt | ttttattata | 4980 |
| tttaaatatt | ggttattttt | tgtaaatata | tatatattata | tattttatatt | tatatattata | 5040 |
| tatatatttg | tatattgtaa | gtaggttttt | attgtatatg | tatgtatata | tgtataagga | 5100 |
| ttatatatag | aaattataaa | tatatttttat | atatgataga | gaatatgtat | atagatattg | 5160 |
| taatattttat | atggatatat | atatattatt | aaatatattt | ttttgtatta | gagatagtta | 5220 |
| tattagataa | tttatttggt | tttatattta | tttttatggt | tatttttagg | ttttattttat | 5280 |
| tttttattag | tgaagttagt | agtgttatat | tgggtggtttt | gtgatggtta | gttggtaggt | 5340 |
| tgagtttaaat | tttttttaaat | tgtttttgta | aatattggtt | tttatttttg | gatttggttt | 5400 |
| ttgatatgta | tagagataga | tagatatata | attatgaaga | taagagaaat | gtatggaagt | 5460 |
| taagtgggtg | ttaaaggggt | ttgttggttt | tattagtttt | tgttgagggt | tgggattggt | 5520 |
| ttggggatg | ttttgagggt | ttttttagta | aggggtgaat | taattttttat | ggttttgtag | 5580 |
| ttaaataga | ggtagggggg | ttagtattta | gttttaggag | ttttttagtt | tattttaaatt | 5640 |
| tttttaaatag | taggggtagt | tagggtgggg | tattttgggg | atthtagtat | attagaggggt | 5700 |
| tatatatatt | tattagggga | tatatagatt | tatatgtaga | gatatgtagt | gatatataga | 5760 |
| aagagagata | tatggggaga | tataggtata | tatatagtta | tttgggtgatt | tttttatata | 5820 |
| tatatatata | tatatgtgaa | agaatagaga | tatatagaat | taatagagat | agatatgtag | 5880 |
| agatatagat | ttagatattt | agggatatag | agatatagga | ttatattaag | ttagggtttt | 5940 |
| aggaatatat | gtatagatat | ttagggattt | atagagattt | tttaagatgt | taggagtaag | 6000 |
| gagataaatt | ttatgattgt | atatatgtta | tttttatata | gtggtataga | tatgtaaata | 6060 |
| tttatatgat | tagttttaga | tgtataggta | attatatata | tatatatata | tatatatata | 6120 |
| tatatatata | tatatatata | gagtaggata | tatttttaggt | tatatatttt | attagaagaa | 6180 |
| atataggata | tttatattta | tggaaatata | tgtttttggt | tatttggttat | tatagatata | 6240 |
| tgtatatgta | tatatatggt | tatattattt | ttagattttat | ttaaagtttat | ggttttgaag | 6300 |
| tagaaatatt | tttttttaggt | ataaagatat | gtatagagag | agagagatat | aggttttagga | 6360 |
| agatgaatta | tggattatta | tgtttttatat | tttaaaatat | atagggttta | tgttgttgtg | 6420 |
| tagaatgggt | tttgggtttt | gggttttatt | ttttagaaga | gggggtgagg | atgggttttt | 6480 |
| ttttttttta | gatttagttt | tgttttgatt | agtgtggatg | agttttttat | tttttttttg | 6540 |
| tgtttagggt | tttttttttg | tttatttttt | tatttttagat | ttatagatat | gtggtagtta | 6600 |
| ttagggattt | gggttggttt | ttataaatatg | tataatttag | agttttttggg | gttttttttt | 6660 |
| ttagggagat | tgaggtagag | ggtttaggtg | agatttggtt | tttttttgat | gtttttgagaa | 6720 |
| ggtttgaaga | ggttagggat | ttaaatagg | gggaattgt | gggttggttt | ttttatttgt | 6780 |
| ggaggatatt | ttttaggagg | gaggtgttga | ggtatttggg | aggtgagagt | tgttgttgta | 6840 |
| ttttttttat | tgttattttg | tatttggtatg | ttgagttgag | tagtgaaagt | tggttgggta | 6900 |
| tggagtagaa | gattttatta | ggatttgtga | tgttgtttat | taggttggtt | ttggttaagg | 6960 |
| agaggggtga | gaggttgttg | gttttgagg | ggatgggtat | tgtggagaag | agggaaatgt | 7020 |
| tgagttagtg | gttgttgtat | atagttttat | ttaaaggagg | gtttgggtatt | tttgttttat | 7080 |
| gaagggagaa | attgagttat | agtgaggtga | aagtgatttg | tttaagtttt | tttaattgat | 7140 |
| ggatgtgaat | ttagagattg | aattttgagt | tgtttggttt | ttgaatttgt | gtttttttta | 7200 |
| ttgaattagg | ttggttttat | tttaaattta | aatagttttt | gtggaatttt | ggtgtttttag | 7260 |
| tggatgttat | tagagagga | tttggatgta | aggatatgag | tgtgagttgg | aggtaggatt | 7320 |
| tttggggttt | tgtagtagg | ttagttagg | gttggtaaat | tataggagat | tttaaaattt | 7380 |
| ttttttttta | tttgtgattt | ttatgttttt | tttttttgtt | tttttatttt | atthtatttt | 7440 |
| tgtattgggtg | aggggtttgg | agattgtggg | ttttatattt | tggagatagg | gaggttattg | 7500 |
| ttttgttaaat | gttttttagtt | ttggtttttt | ttttttgta | agttattttg | tgaatgttaa | 7560 |
| aggagattag | gagataggtt | ttatatattt | tttttaaata | ttttttaatt | atthtattat | 7620 |
| tgtttgaaga | ttagagaagg | aatttgagaa | gttaattttt | tttttttttt | tttttttttt | 7680 |
| tttttttttt | ttgagatgga | gttttggttt | gttattaggt | tggagtgtag | tggattattt | 7740 |
| ttggtttatt | gtaatttttg | ttttttgggt | ttaaagtatt | tttttggttt | agttttttga | 7800 |
| gtagttggga | ttataggtgt | ttattattat | gtttggttaa | tttttatatt | tttagtagag | 7860 |
| atggggtttt | attatattgg | ttaggttgg | tttaaatttt | tgattttatg | atthtattgt | 7920 |
| tttggttttt | taaagtgttg | ggattatagg | tatgagttat | tgtgtttggt | tagaagttaa | 7980 |
| tttttaaaaa | tttattttaga | tatttttgg | tataagagaa | agaaagtttt | ttttgtagtt | 8040 |
| taattttttat | ttttttggtt | agttttaggg | gaagggagaa | tattgaagta | tttgtttttt | 8100 |
| tagtggaaatt | ttttgggttt | tttttttttt | aaggaggtga | tagttttggt | agataagggt | 8160 |
| ttgtgtgtat | attaggagtg | aggtttgggt | gtggagtttg | ggattaatag | ttattgagta | 8220 |
| tttattgttg | tgtgtttttg | tatataaatt | tatttagttt | ttggaataat | ttataagggg | 8280 |
| ttatatattat | tattattttt | atthtataga | tgaggagatt | gatgtttaga | gaggtgaaat | 8340 |
| aatatgtttg | agattatata | gttggtagat | ggagaattgg | tagtaaattt | aagattgtag | 8400 |
| tttaaaagg | ggtgttggtta | gttttttagt | tatttggtta | tattggttga | gtattttttat | 8460 |
| gtaagaatg | gtgttaggg | tgaggaatat | aaaagtagga | aaggttagtt | tgattttgtag | 8520 |
| tttggtagaa | tgttttagtt | tttttttttt | ggataaattt | ttttattttt | atthtagtagt | 8580 |
| ttagttgagaa | atthttgtga | tattttttgt | aatttttttt | tttagttatat | ttttattttt | 8640 |
| attatatttt | atthtatttt | atattttatt | tattattaga | ttttatagtt | tttttttttt | 8700 |

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| gtttttttta | tttttgttta | ttttatgatt | tttatttttt | tatgtgtttt | gtatgtgggt | 8760 |
| atttttttgt | atgtatatag | atataggttt | tttttttatg | tgaatttttt | tatagatttt | 8820 |
| gttttgggta | ttttgtttta | gtttagaggg | atttttttat | ttgaatgtat | tattttgtag | 8880 |
| ttaaaaat | agtttaaaat | ttattagagt | ttggttagtt | ttttgagaat | taaatttttag | 8940 |
| ttttagagtt | agataaaaatg | atgttatttg | ttaatgagtg | ttaattataa | ttttatttgt | 9000 |
| taatgagtg | taggtgttga | aagtttaaaag | aggttaaaag | gggatagtaa | gtaaggggta | 9060 |
| gaaatgaatt | tttaattttg | agtattgatt | ttagagttta | tagttttaat | tattatgtta | 9120 |
| taatgggttt | ttttattttt | ttttgtgttt | tggttttgtt | tgtttttttt | gttgggtttt | 9180 |
| tatttgttta | ttaggtttta | gatttaagtt | gattaagaga | aatataatgt | gagttatag | 9240 |
| tatatattta | tatttttttag | tgtttattat | tttttttttt | gagatagggt | tttgttttgt | 9300 |
| tgtttagggt | ggagtatagt | ggtataaatta | tagtttattg | taattattat | tttttagggt | 9360 |
| taagtaattt | ttttattttt | gttttttaag | ttgttgggat | tatagggtgtg | tgttattatg | 9420 |
| tttgggtaat | ttttgtatgt | tttatagaga | tggggtttta | ttatgttgtt | tagattgggt | 9480 |
| ttgaattttt | ggatttaagt | tatttatttg | tttgttttgg | ttttttaaag | tgttgggaatt | 9540 |
| aaaggggtga | gttattaggt | ttggtttata | ttttaaatag | taaaaagaaa | tatgtgaagt | 9600 |
| taattttaat | taatataata | gtaaaataata | tattattatt | tagtatattt | tatttagttt | 9660 |
| aatatattaa | aatagaatt | attttaatat | gtaattaata | ttttaaaaat | tattagttag | 9720 |
| atattttata | ttgtttgaaa | ttagttttta | aaatttaaatg | gatattttat | gtttaagtgt | 9780 |
| attttaattt | agattagtta | tattttgggt | gtttaagagt | tatttgtggt | tatagtgggt | 9840 |
| atttgggtga | tagtaaaagt | ttaggtttta | agggaggag | ggtttaatag | gtaaatag | 9900 |
| aagtgggtg | tgggggaggt | tggtaggtta | gaggggtatt | ttatgtaaag | tttgtttaag | 9960 |
| taagatataa | agtttttgg | ttgtttgtgt | ttttgttttg | ttttgttgat | ttttgaaatt | 10020 |
| aggattttgg | atttttttgt | gaaatttttg | gagttttaag | tattggtaat | gatttaaaga | 10080 |
| tttaaaaatt | tagtataggt | taaatggaat | aagtgtgaag | tgttttttgg | gatgtagtga | 10140 |
| tttttaggg | gtgagttttt | tgagttttat | agtgtttttt | tggtttttag | gggttttgtt | 10200 |
| ttattttatt | gagtttaggt | attaattttt | gtgtagagtt | tttttttttt | gttttttagt | 10260 |
| ttttttttat | ttttattttt | ggttttattg | tattgaattt | ttttggtttg | tttttatggg | 10320 |
| tgtttttttt | attaagtatg | agtgtgtgga | gggtagggtt | ttggtatagg | tggttatgag | 10380 |
| tgaatatgaa | ataagttaat | gatgtaagg | ttttgggaat | ttgaattgaa | ttatattgta | 10440 |
| tgtataaaaa | aattatattt | tgagtttgat | tttaaaaaaa | taaaattata | aatgaggtta | 10500 |
| ggtatgggtg | tttatatttg | taatttttag | attttgggag | gttgagggtg | gaggattgag | 10560 |
| tttaggaggt | taagattagt | ttgggtaata | taatgagatt | ttgtttttat | aaaaaatata | 10620 |
| aaaattagtt | tgggtgtgatt | gtatatgttt | gtagtttttag | ttattttgga | ggttgaggta | 10680 |
| ggaggattgt | ttgagtttaa | gaggtagagg | ttgtagttag | ttgattatgt | tagtttgggt | 10740 |
| ggtatagtga | gattttgttt | taaaaataaa | aataaaaaata | aaaataaaaa | attgataaat | 10800 |
| taaaaataaa | aagaagggtg | aagtaattgt | tttttttttt | attatttttt | attagaagta | 10860 |
| gggttgggag | aaaaaggatg | agtagttttt | tttatttaat | tgttattaaa | atttagggtt | 10920 |
| taggttgggt | gagtatggta | atttatattt | gtaattttta | tattttggga | aatgaggta | 10980 |
| ggaggattat | ttgagtttag | aagtttgaga | ttagtttggg | taatatagtg | agattttatt | 11040 |
| ttataaaaa | aaaatttttt | taagttagtt | aggtatggtg | gtatgtgttt | atagtttag | 11100 |
| ttgtttggaa | ggttgagggt | atagaatgat | tgagtttagt | ttgaggttgt | aggttttat | 11160 |
| gattgtatta | ttgtatttta | gtttggatga | tagagtgaaa | ttttatttta | aaaaaaataa | 11220 |
| aaaaggatga | ttgtatataa | gaattgtatg | tttttgtgat | aaatatagta | atgaggatga | 11280 |
| gttggagtat | ggaggaatgt | ggtagtttag | atatggatat | ttgtaagtgt | ttttttttgt | 11340 |
| tgatgttaga | tttatgtaa | tataggtgtt | gttagatata | agggtagaga | tgatattggt | 11400 |
| tatgtgtttg | tgtatatgag | tgattgattt | tgttgttttt | tttatttgtt | tttttttttt | 11460 |
| ttagagtttt | tttttttgaa | aatggtaaaga | gggatggaat | ttatggtttag | ggaagttagt | 11520 |
| tggttataga | agggtagtgg | tggaggttag | aattaaatta | gtatatagg | taaagttag | 11580 |
| atattaaagt | ttgtttttta | tatatataga | ataggatttt | ttgtgggtgat | ttttgaagtt | 11640 |
| gaagtgattt | aatgggtggt | tgttgggtat | atatgtagaa | gttatttata | tatatatata | 11700 |
| tatatatata | tatatatata | tatatatttg | tagatataag | agtttttgg | tatttaattt | 11760 |
| gtagataaag | gtattttggt | tgtagggtgt | agatatggtt | tgatttttga | atttaggtaa | 11820 |
| gtgtgtaata | gattttttta | gaagttttgt | taaagtgtga | agtttttttt | agaagaaaat | 11880 |
| atggaagggt | aggtgttttg | gtttatattt | gtaattttat | tattttggaa | ggtaagggtg | 11940 |
| gtggattatt | tgaggtttag | agtttgagat | tagtttggat | aatatgggaa | aattttattt | 12000 |
| ttattaaaa | tataaataat | aataataata | ataatagaaa | attaggaga | aaaatttttg | 12060 |
| tgatttttag | ttgagttaaag | agtttttaga | tataatatta | aaagtaagg | ttataaaaga | 12120 |
| aatgatagg | ttaggtgtgg | tgggtgattt | ttggaatttt | agtattttgg | gaggttgaag | 12180 |
| tgggtagatt | atttgagggt | aggagtttaa | gagtaatttg | gttaatatgg | tgaaattggt | 12240 |
| ttattaaaa | atataaaaaat | tagtttagta | tgggtggtatt | tatttgtaat | tttagttatt | 12300 |
| tgggaggttg | aggtaggaga | attgtttgaa | tttgggaggt | ggaagttgta | gtgagttgag | 12360 |
| atagtgttat | tgtatttttag | tttgggtgat | agagttagat | tgggttttaa | aaaaaaagaa | 12420 |
| agaaagaaag | aaagaaaatg | atgataaatt | ggattttatt | aaagttaa | atttttgtat | 12480 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tttaaaagat | attattaaga | aaatagtatg | gtatggtagt | atgagaatgt | agtttttagtt | 12540 |
| atgtgagagg | tttaggtagg | gggattat | gagtttagga | gtttgaggg | gtagtgtgtt | 12600 |
| atgattttgt | ttgtaaatag | ttattaaatt | ttagtttgaa | taatttagtg | agatttat | 12660 |
| ttaaaaagtt | ttttaattaa | aaaaaaaaa | gagagtagtt | gtttattttt | gtaattttag | 12720 |
| tatttaggga | ggtaagata | ggaggattat | tggaagtag | gagtttgaga | ttagttggg | 12780 |
| taatatagtg | ggattttgtt | ttatgaaaa | aaaaaaaaa | gataattttt | aaaaatgtaa | 12840 |
| aaaaaaaaatt | aaaattttaa | aaatagaaaa | gaaaatgata | aattataaat | tggaagaaaa | 12900 |
| tatttttaaa | ttatatattt | gataaaggat | tgtattttaga | atatattata | taaagaatta | 12960 |
| ttataattta | ataataagat | agtaatttaa | atagaaaaat | gtttaaatat | gtaaatagat | 13020 |
| attttattaa | aaatgagaga | tgaatggta | aataagtata | tgaaaaagtt | tttaatatta | 13080 |
| ttagttatta | gggaaatgta | aattaaaaat | atattgagat | attatttgat | ggttataata | 13140 |
| aaaaagatag | gggttaggta | tgggtggtta | tgtttgtaat | tttagtaatt | tgggaggttg | 13200 |
| agggtgatgg | attatgagat | taggagttt | agattagtta | gtttgttta | tgtggtgaaa | 13260 |
| ttttgtttta | attaaaaata | taaaaattag | ttgggtgtgg | tgggtgtgtg | ttataatttt | 13320 |
| agttattggg | gaggttgagg | taggagaatt | gtttgaattt | aggaggtaga | agttgtagt | 13380 |
| agttgagatt | gtgttattgt | attttagttt | gggtgataga | gagataaatt | attttaaaat | 13440 |
| aataaataaa | taaaaaaggt | taggtatggt | ggtttatatt | tgtaagttta | gtattttggg | 13500 |
| aggttaaggt | gggtggatta | tttgaggttg | ggagtttagg | attagtttga | ttaatatgga | 13560 |
| gaaattttgt | ttttattaaa | aaaataaaa | tagttggatg | tgggtttatg | tatttghta | 13620 |
| tttagttatt | taggaggttg | aggtaggaga | attatttga | tttaggaggt | agaggttat | 13680 |
| gtgagtttag | attattttat | tgtattttag | tttggtaaat | aagagtgaaa | ttttatttta | 13740 |
| aaaaaaaaaaa | aagaaaaaaa | agataatttt | aagtgtagga | tgtggaaaag | tggaattttt | 13800 |
| atatattggt | ggtgagaatg | taaaataata | tagttatttt | agaaaatagt | ttgatagt | 13860 |
| tttaaaaatt | aaatttaaat | tgattatata | attttagta | tgtattttta | ggagtttatt | 13920 |
| taaaagtgtt | aaaaatatat | ttttatagta | aggtatgtat | gtgaatg | atagtaatat | 13980 |
| tatttaagat | agttaaaaag | tggaaataat | ttaatgtttt | gttaattgat | taatggataa | 14040 |
| ataaaatag | ttatatttgt | ataatggtat | attatgaa | aataaaaaaga | aataaattgt | 14100 |
| tgataatagt | ttgttttttt | ttattgttgt | tttttttatt | tagtataatg | tttgggggtga | 14160 |
| gttttaaaaa | tattatgtta | agtaaaaagat | gtttgatgta | aaagtttgta | tatttatata | 14220 |
| ttttatttat | agaatattgt | tagaatatgt | aaatttatgg | atagaaaagta | gattagtgg | 14280 |
| ggtttaaggt | ttggggtgag | aatggagagt | agttataaat | gggaattttt | ggggatggta | 14340 |
| gaaatgtttt | aaaatggatt | tttttgtggt | aatggttata | taattttata | tttataaaaa | 14400 |
| attattaaat | tatattttta | taatgtgtaa | attttaggta | tgtaaaatttt | attttaaaaa | 14460 |
| attgttaaaa | aatgtaagtt | ttagggttat | attttttagg | gtttttatgt | agttaatgtg | 14520 |
| gggtggtggt | taagaaattt | gttttttttt | ttttgaggt | gagttttgtt | ttgttgttta | 14580 |
| gggtggggtg | tagtgggtgta | attttagttt | attgtaagtt | ttgttttttg | ggtttatg | 14640 |
| atttttttgt | tttagttttt | tgagtagt | ggattatagg | tgtttgttat | tatgtttagt | 14700 |
| taatgttttt | tttggtttgt | ttttgttttt | ttagtagaga | tggggtttta | ttgtgtttgt | 14760 |
| taggatgggt | ttgatttttt | gattttgtga | tttatttgtt | ttggtttttt | aaagtgttgg | 14820 |
| gattataggt | gtgagttatt | gtatttgggt | aaaatttga | tttttaataa | gtattttttg | 14880 |
| ggggttttga | tgtatgtggt | tttttggtt | atatttaggg | tattattggt | ttaaagtttg | 14940 |
| aagggtttgg | agtaagagg | aggtagatgg | gaggagtagt | gatttgtttt | ttggtttata | 15000 |
| ttggggtata | gtatgtgtaa | atgttttttt | tagagttttt | ttattttttt | tgttgggatt | 15060 |
| tgttgattaa | gttttttttt | tattaattgt | tagagggttt | tttttttttt | ttttttttta | 15120 |
| gatggagttt | tgtttttgtt | gttttaggtt | gagtgtaatg | gtataatttt | ggtttattgt | 15180 |
| aattttttgt | ttttgggttt | aagtgatttt | tttgttttag | tttttgaggt | agttgggatt | 15240 |
| ataggtgtgt | attattatgt | ttggttaatt | ttttttgtat | tttttagtaga | gatgatattt | 15300 |
| tattatgttg | gttaggttgg | ttttaatttt | ttgatttttag | gtaatttatt | tgttttgggt | 15360 |
| tttaaaagtg | ttgggattat | aggtataagt | tattgtgttt | ggtttaggg | tttttttata | 15420 |
| tataattttt | attttttatt | tttttgtttt | aagttttttt | atagttttta | ttatttttag | 15480 |
| gtaaagttta | aatttttatt | gtatattttt | taagtgtttt | tttggggttag | agagtgtttt | 15540 |
| gggaggtagg | tgagaaaggg | ttgtggaagt | gtaagtttag | taaatttttg | atattttg | 15600 |
| tattgtggat | ttttgtatta | tttttattta | aaaaggttta | gtgtgggtgg | ttatgtttgt | 15660 |
| aatttttagta | ttttgggtgg | ttgaagtggg | taggttattt | gaggttagga | gtttaagatt | 15720 |
| agtttggtta | atatgggtgaa | attatgtttt | tattaaaaat | ataaaaaatta | gttgggtgtg | 15780 |
| gtgataggta | tttgaatttt | tagttatttg | ggaggttag | gtatgagaat | ttgagaggta | 15840 |
| gaggtttag | tgagttgaga | ttgtggtgtt | gtattttaat | ttgggtgaga | gagtaaaatt | 15900 |
| gtgttttgaa | aaaaaagatt | tttttttttt | ttgagataga | gttttgttgt | tgttgttttag | 15960 |
| gttgaggtgt | aatggtgtga | ttttggttta | ttgtaatttt | tgtttttag | gtttaaggga | 16020 |
| tttttttgtt | ttagtttttg | gagtagttgg | gattataggt | gtttgttatt | ttatttagtt | 16080 |
| aatttttga | tttttagtag | aaatggggtt | ttattatatt | ggttaggttg | gttttgaatt | 16140 |
| ttagatttta | gggtgattat | ttgttttgg | tttttaaggt | gttgggatta | tatggtgtga | 16200 |
| gttattgttg | tttgggtttt | tttttttttt | tttttttttt | tttttttttg | agatagtttt | 16260 |

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| tttttgttgt | ttaggttggg | gtatagtgg | atgatttttg | tttattgtaa | tttttgtttt | 16320 |
| ttgggttaag | taattttttt | gtttttagtt | tttgagtgg | tgggattata | ggtttagtgt | 16380 |
| attatattgg | gttaattttt | gtatttttag | tagagatggg | gttttattat | tttggttagg | 16440 |
| ttggttttga | atttttgatt | ttaagtgatt | tggtttttta | aagtgttggg | attatagatg | 16500 |
| tgagttattg | tgtttggttt | atttattttg | attattgagt | tttttggat | ttttttaaat | 16560 |
| tttgttttag | aagttagtat | tttatttatt | ttatttatgt | tttggttggt | tagagattgt | 16620 |
| ttgggaagt | gagatttttt | ttagtttttt | tttttttttt | tagggttaaa | gatagtaagg | 16680 |
| aataggggga | ttgtttata | tagttttttg | tttagagatt | ttttgagatt | ggatagttga | 16740 |
| gtttggggag | agggattttt | tatttttaag | taggtaatta | tttatttttt | tttagaggga | 16800 |
| atgagggatt | tatttttagt | tttatttttg | attatttttt | tttagttttg | attgtataat | 16860 |
| tgaaaattta | tatatggttt | gtaatttttt | tttttttttt | ttgagatgga | gttttgtttt | 16920 |
| gtttttaggt | tggagtgtag | tgggtgtgatt | ttggtttatt | gtaatttttg | ttttttgggt | 16980 |
| ttaagtgatt | tttttggttt | agttttttta | gtagttggga | ttatagatgt | gtattattat | 17040 |
| gtttgggtta | tttgtttgta | atttttttta | tttgtttgg | ttttttatag | ggagaggatt | 17100 |
| tgtatatgtt | gtgttttgtg | tgatgaaagg | agtttttttt | ataatatttt | ttttttaatg | 17160 |
| tatttagatt | ttaaagttta | ttatttattt | atttaataaa | tatttattaa | gtattttttg | 17220 |
| aatttggttt | aattattttt | tttttaggta | tttatttaat | tttgggtatt | ttttgtgttt | 17280 |
| tttttgagtt | ttttttgtgt | tttatatagt | ttttgtttat | tgttggaaaa | tatttttgtt | 17340 |
| aagttttgtt | tttttattag | ttttgtttgt | tttttgtgtg | tttgggatag | gttgtaaaat | 17400 |
| tggaggtgat | aaatgtgggt | aggaaatgga | gggttttttt | atattttttg | ggttggtttt | 17460 |
| agttttgtta | ttttttgttt | aatattgtgg | atgtaattgg | tatgggattt | ggaagtgtgt | 17520 |
| ggtaaaagt | tggggttttg | tttggttggt | ttttttggat | gtttgtttgt | agttagttaa | 17580 |
| gtttttttaa | tttaggtttg | ggttttgtga | gtttttagtg | tttttgtttt | gtggtgttgg | 17640 |
| tgaagttgaa | gtttgagaat | gtttattgta | gtgatgtgaa | ggttgttttt | ggggtggggt | 17700 |
| tgaggttgta | gttgtttttt | ttttgtatta | aggattttta | tttttagtga | tgtagttggt | 17760 |
| gtttttgttt | aggttgggag | gtattgtagg | gatttgatgt | tttaggtggt | taaagagtga | 17820 |
| ttttttttga | ttttagggtt | ttggtggggt | aggttttagt | attgtatttg | gtggagggtg | 17880 |
| aaggtttgtg | gggtaggata | ggagtttttt | gtgttgttgg | aagggttgag | gatgaaggag | 17940 |
| ggtgttaatt | tattttttat | tgggttgggt | gtaatgttga | atttttagt | gattgtggag | 18000 |
| ggtaaaagt | aaaattgttt | gggtgttga | ggtaggtgt | gggagggggt | ggttttaggg | 18060 |
| agtaaggagt | tattttgttt | tgttgttga | gttgttttgg | gttgattggt | tatgtttttt | 18120 |
| tttgggttat | gatttttggg | tttaattgtg | tttggttttt | gttttttttt | ttttgttgtt | 18180 |
| gttgtgtggg | ttgtaatttg | atgttttagt | tgggtgtttt | agggtgtagt | ttttgtttag | 18240 |
| gttgttagtg | ttttattttg | tttattgggt | tatagatttg | ttggtgttgg | ggttttgttg | 18300 |
| gtgttgggg | tttgttgggt | ttgggttttt | gttggtgttg | gggttttgtt | ggtgttgggg | 18360 |
| ttttgttgg | gttgggtttt | tgttgggtgt | gggttttgt | ggttgtgtat | gtgagtttgt | 18420 |
| gtggtttgt | tttgtgtggt | tgtgtatgtg | agtttgtgtg | gttgtgtttt | gttttgtttt | 18480 |
| aggaggttag | tgtgttgtta | tttgggatgt | taggattttt | gttgtgtttt | ttggattggt | 18540 |
| ttgggggatt | ttggtgtatt | tttaggattt | aggagttttg | gaagttgttt | gagagaaatt | 18600 |
| agttttggga | gggttttga | tttagttttt | tgttttgggt | ttggattggg | gttttgggtt | 18660 |
| aaggtgttta | gaggaatagt | tgattttttt | atttttgtgt | agggtagaga | tttttaaat | 18720 |
| ttttttaaa | atgtagggtt | ttagtttttt | ttagggagt | agtgaattta | gatttttagt | 18780 |
| tttttgagtt | taagtatgaa | tagggaattg | gggattatta | ttatgtttat | atttttggtg | 18840 |
| gtaggaagt | ttaggtaggt | ttttgtttat | tgtagatgga | tttttttttt | taggggttaa | 18900 |
| gaaaggtttt | gtatagtaa | tttaattaagt | tttattagta | gagttgtgtt | gtaattaggt | 18960 |
| tttttagtgt | ttttttatat | tttttagttt | aagtgatata | ggatagattt | tttaattttat | 19020 |
| ttttttttat | ttattttttg | ttttgttgtt | ttattttttt | tgttatatat | tttaagttat | 19080 |
| ttatttaatt | tttagttttt | ggtaatttag | tttttagttt | tggaaagttta | gtttttagttt | 19140 |
| tgttattatt | atttttattt | tggttttttt | tgagaataag | tggagggtaa | atagagttta | 19200 |
| ggtttgaatt | tatttggttt | aaaaataatt | taagttaaat | tttgtgggag | tagtgtttgt | 19260 |
| ttgggttatt | ttaggtttat | gtgtattttt | gaattttttt | gggtatttgt | tttaaatgta | 19320 |
| aagaggttgt | tatttaagt | ttgggttttt | ttgatttttt | gttttttttt | gtttgttttt | 19380 |
| ttgtttttgt | tgttttttat | ttttttttgt | tatggttagt | atttggtttt | taggttagaa | 19440 |
| aaggtggatt | ttgtgttttt | ggatttttatt | aatgttagga | gttatataaa | tatttttata | 19500 |
| tatatataga | gaggttttgt | gtatgtttgg | aaatttttag | atattaaaga | ataaagtgta | 19560 |
| ggatataata | gagatattgt | atatattgag | aaatgtgtat | attttttaaa | tgtagagaaa | 19620 |
| tgtatattaa | tatatagaga | tatattttat | aaatatattt | atagaaatag | atatattttat | 19680 |
| attttaagaa | atattaagat | atttatgaag | ggaatatttt | agatatgtat | aggattttga | 19740 |
| atttatattg | atatagatat | ttaggagtag | gtgggttttg | attttaggtgt | atatagttta | 19800 |
| atataataga | tttatatata | agttttttta | agttttaaat | ttgtaagaga | tttttataga | 19860 |
| aagaaaattt | tttttagaggt | ttttaaggtt | ttgtttggaa | ggaagaggaa | gaaagtgttt | 19920 |
| gttaggtatt | gaaattgtaa | ggtattgtaa | agtgttaaat | atttttttta | tttttgggga | 19980 |
| atagtatttt | tattttttatt | tttttaggaa | ataggggaagt | ggttttaatt | gttttgtaaa | 20040 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ttaatttttag | atataaaga | gtgttttttt | ttttggtggg | gagaggttag | gttttttagtg | 20100 |
| ttgtagttta | gtgaatggtg | atattggtgag | gtttatttta | ggttataggg | tgttgggtaa | 20160 |
| atataaag | tttgtatggt | tgtttaatat | ttggttgggt | taggaagagt | gtttggtgga | 20220 |
| tatttagttt | ggtgatatat | aggaggtatg | gttgggtgaag | aaaatggatt | tgtgtattaa | 20280 |
| gggtttgfat | tattttgttg | ggttttgttt | tgatagatta | ttttttattt | tttttgatta | 20340 |
| tggattgggt | taggaggttt | atattttggtt | tgtttattgt | ttgtaaatag | tagagggtag | 20400 |
| tgggtagggg | gtataaagtt | ttagttggtt | atgttttttt | aagtttttgt | tatttagggg | 20460 |
| ttttgtttgt | tttgtttaag | gtgtattagt | tttgtttggg | gtggtgatgg | gtaggttttag | 20520 |
| gaattggtat | gtttaggtat | tgagttggag | ttttgggtat | tttattgtgt | agtgtaaaaa | 20580 |
| gggaattttg | aattttatat | tggttaattat | tatgtttttt | tttgtttttag | gttagggttg | 20640 |
| tttatgggat | gttatatgtg | gtagtgggga | gaaaagaagt | gatttgggtt | ttggtgaggg | 20700 |
| ttggtggtat | ttatgtattg | ttggtggatt | tagtgttagg | tagttttgat | ggttaggtaa | 20760 |
| tgtttaattt | aggaggtggg | atataatttt | ggagttatta | gtgagatgtg | gtgagttggg | 20820 |
| agaaaagttg | gttttttgta | tttggtattt | tttgttttta | gttatggtag | taaattattg | 20880 |
| tgggttagttt | tgatttatatt | tggatttttt | gggttttatt | tgtaggtttt | ttagattggg | 20940 |
| ggttgttgtt | agttttggaa | ggttgagagg | tgtgtttgtt | aggttgtgtg | tgttgttggg | 21000 |
| taggttgtgt | aggagtggg | gtatggtagt | ggtatagtta | tgggtgtggg | taaggtttag | 21060 |
| ggtgtgggtg | ggtggtgtta | gtaggttggg | aggttttttg | tgggtgtggg | tggttgtgtg | 21120 |
| gggtgtgggt | taggtgggtt | gaggaggttg | tggttgtgtt | aggggtgtta | ggttgttata | 21180 |
| tgggtttttt | gtaggtggg | gaaaggttg | ggtggtgttg | ggtgtttgat | tgtagggtag | 21240 |
| tgggtggtgt | ggtaggggtg | gtgggaagt | gggtggttgg | aatttgggtg | tagttgtggg | 21300 |
| ggttgggtgtg | tggttagagt | gggtgttgg | ttttaggtt | gtttggggta | gagatgatag | 21360 |
| gtgggttttg | ttggtagtgt | tttttagttt | gttgggggtg | ttttggggaa | ggagaggggt | 21420 |
| gttttaggtg | atattattgt | agttgaaagg | attttagttt | gaaaggattt | tttttagttt | 21480 |
| gggaggggtt | agttgtagt | ttttttgttt | tttgggggtt | gaggtttatt | tttgagttat | 21540 |
| ttgtgggtgt | aggagtttta | ggttggggat | agtgtttatt | tgtgtattgt | gttttttagtt | 21600 |
| ggtgaagttt | ttgtggtttt | ttttttttga | gattttgtta | gtgtgggata | ggaggggttg | 21660 |
| aaagtttagt | atattaggtgt | tttgattttt | tggttgaata | tgttttgggt | ttgagattat | 21720 |
| ttattatggt | ggagttagtg | tgtattagta | tgggtggtg | agtgttgggt | gtgaggtgtg | 21780 |
| tggtggtggg | agtttttagt | tgatggtgag | tgttgtgggt | agaggtgttg | ggagtggta | 21840 |
| tagtttgtgg | aggttatggt | gtggtagtgt | ggtgtgattt | agttgggatt | ttagtgggtg | 21900 |
| tgggtttttt | agtgggtggt | gtttttgggt | gggtttatgg | aggtaggatg | gatgggtatt | 21960 |
| gattgtattt | ggaaaagttg | gttaatgggt | gttttagtgt | tgtttttttt | tttttgagtg | 22020 |
| tagtttgttt | tatgtatttt | agttatagtg | gtgggggtgt | gttgggtttt | tgggggtgtg | 22080 |
| tattaagag | ataagttatg | gtagtttttt | tttttttttt | tagtggttgt | tataggggag | 22140 |
| gtagtgggtt | tgggtgggatt | gggtagggtg | ggtagggggt | gtgaatttag | tttatattga | 22200 |
| ttagagtgtg | tttttgggtt | tgtagtttgt | agttgtagt | aaatagaaat | atttgaggtt | 22260 |
| gggatgtaag | ggttttatat | ttgtttgggt | ttattatatt | tttgttttta | agtaaagtat | 22320 |
| tattttttgt | tggtttttagt | tttttttggt | aaataggtat | taattttttt | tttaagtatt | 22380 |
| taaagagatt | ttgatgagg | tgggtaattg | ggaatgtgga | agaagtttgt | gagttgaaga | 22440 |
| tttagtttta | tttgtttatg | tgtatatagt | gttagtgtgt | gtgtttaggg | ggtttattat | 22500 |
| tatagttttt | tttgaatggt | tttgattaa | aattttttga | aggtgtttat | tttgtttttta | 22560 |
| ggttttttgg | gttgaaaatt | tttaagttaa | gttaggtttt | gagggagttg | gagatagaag | 22620 |
| ttggttagat | agagagagag | atattgtttg | gatgggtttg | ggtatgagtt | ttatttgata | 22680 |
| aggggttaag | atatttttga | agttatagg | ttttgggttg | attgagtgtg | ggtatgagat | 22740 |
| attagagat | tgttgagtga | ttgtgtgtaa | ttgattagta | tggatgttgt | gtgattgtta | 22800 |
| tattaatatg | tgtaggttat | atgtttttgt | gttattatgg | taggtatttt | ttggttgtta | 22860 |
| agtaagttt | agtttttga | gttatttttt | tttttttttt | aaatttaatt | tttttttttt | 22920 |
| gataagtaga | gatagggatt | aaagatggat | aaatgagagg | aatgttaagg | agaataggaa | 22980 |
| taaagaaagt | agaagaaaga | tgttttttta | aagtaagatg | tgtagattag | ggtaaagtag | 23040 |
| aaataattat | agtagattta | ttatagaatg | tgttagattt | tgtgtaaagt | gttttatgaa | 23100 |
| tattgggttt | tttaattttt | atagtatgtt | tatgaaagag | atattttatat | tatttttatt | 23160 |
| tttaggggag | gggtgaggt | tggagaggtt | aaatgggttg | tttaagggtta | agtagtttaag | 23220 |
| tggtagagtt | gggggtgatt | gtttttggtt | tttttttgtt | tagaattatt | attttttatt | 23280 |
| gttttagatt | taggggtttg | gatttagttt | tgattttgtt | taattgtgtg | attttagtta | 23340 |
| tattttgttt | tttttgtgat | tttagttttt | gtaattttga | taatattaat | aaaaagtag | 23400 |
| gggaaggtag | agaagattag | ttgttttttt | aggttttttt | ttagtgaagt | agtaaagttag | 23460 |
| aagattaagg | taaaattatt | gatatagaag | aggaggtgat | agtgatatta | tttaatttatg | 23520 |
| gaatatatgg | ttttaagtgg | gtatttttaga | aggtatatgg | taggtgtaga | gttttaggggt | 23580 |
| ggatgatttt | agtttgggga | gtggtgggga | aagaaggagg | ttagaagtta | aggatttatt | 23640 |
| attaagaatt | ttttgaagtt | tggaattgtt | agtgaggttt | tagaggatga | agtggaattg | 23700 |
| ggaagttttg | taattttttt | gtttgagatg | ggttagagat | taggattgtg | tttgtgtgta | 23760 |
| tttatatatg | agggagtttg | tggtaatggt | ttgttttttt | tttttttaat | tttagttata | 23820 |

ES 2 597 845 T3

```

gggtattaga agttaagttt taggataggt gttagtgtgt gaatagaaat tgatgaattt 23880
ttttgggaga agtttaggag ttgtaaagga gatgatgttt gtgtttagtt gtggggatgt 23940
ttttgttttt ttttagtagtt tgggaagggt agagtgggtg gatgtgttgt ttgtttgtgg 24000
tttttttggg tttttttaat ttatttttta tatttttgat agtttttttt tgttttattg 24060
ttttattttg attatttttg tttttttttt tgtttttttt tttatttata tgtttaatgt 24120
tggtaatttt tttttagagt agagaagttt tttttttttt ttgtattttt tatgagtgag 24180
tttatttttt ttttaagttt tttttttttt ttttttataa tagagtttta ttttgttggt 24240
tgggttggag tgtagtggtg taattttggt ttattgtaat ttttattttt tgggtttaag 24300
tgatttttat gttttagttt tttgagtaat tgggattata ggggtttatt attatgtttg 24360
gttaattttt gtatttttat tagagatagg gttttattat gttagttagg ttgattttga 24420
atttttgatt ttaagtgaat atttgtttta gtttttttaa gtgttgggat tataggtatg 24480
agttattatg tttggtttaa gtttttaatg attatttgtt tattgttatt ttttaattgt 24540
gttatttttg gtttgaatta ttagtgagag atttagattt atagatatag ttgttagttg 24600
tgtatttttt ttagatattt tattgtagtt gtttgaattt ttttaagttt aatattgaat 24660
ttaaatattt ttttattttt agtattaatt ttaatttgtt ttttttttgt tttgtagttt 24720
tataaatgat attatttttt ttttaagttt aaatttgttt gttatttttag aagagtgtag 24780
gttttttttt taatttttat aagttaattt ttatggagtg ttgttatttt taatttttaa 24840
atattttatt atagaaggta tgtggtggag aattgaggta ggtatttttag agggtatatg 24900
gtattttggt ttagtatttt aaattatttt atttttattt atagattagt gttttttaa 24959

```

<210> 9

<211> 6096

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 9

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| tattagtgtg | agatttaaaa | tttttttttt | gtattgtata | gtgagatggt | tagggtttta | 60 |
| gtttagtgtt | tggatagatg | gatttttggg | tttgtttggt | gttgttttaa | gtgaagttgg | 120 |
| tgtgttttgg | gtggagtaga | tagagatfff | gggtggtagg | ggtttgggaa | gatatgggtg | 180 |
| gttagggttt | tatgtgtttt | tattgttggt | ttttgttatt | tgtaggtaat | ggatgagttg | 240 |
| ggaatgagtt | ttttagatta | gtttgtgatt | aagaaaggtg | aggaatgggt | tgttagggta | 300 |
| gagtttgggt | agatgggtga | ggtttttggg | gtatagatff | atftttttta | ttggttgtgt | 360 |
| tttttgtgtg | ttgttagggt | gggtgtttat | taggtatfff | ttttggttta | gttagatggt | 420 |
| aggtagatgt | gtgggtttgg | tgagtttggt | tagtatfttt | tggtttgggg | tgggttttag | 480 |
| tggattagta | tttattgggt | tgtagtattg | ggagtttggg | ttttttttgt | tgagggggag | 540 |
| ggtatftttg | tggatttggg | gttgatttgt | agaatgagtt | aaattatfff | tttgtttttt | 600 |
| taagagatgg | gaatggaagt | gttgftttta | tggagttggg | gaaatgattt | ttatftttata | 660 |
| gtgttttagt | atftttgggt | ttgggtgggt | tttttttttt | ttttttttta | gtaggggttt | 720 |
| tggaggtttt | tgggggaatt | ttttttttgt | gggagttttt | tgtggtatft | agatttaggg | 780 |
| gagtttgtgt | gtgagtattg | tgtgttaggt | tgtgtgtatt | tgagttaggg | ttatfttgggt | 840 |
| tttgggtggt | tgtgtttatg | tgagtttagg | gttttgtgta | tgtttgaaat | gtftttttta | 900 |
| tgggtgfttt | agtatftttt | ggagtgtgag | tgtgtttggt | tttgtgaaat | tgtttgtgag | 960 |
| gtgtgftttt | gtatgttggg | gtgtatfttt | ttgtatfttg | gggatgtata | tatfttttta | 1020 |
| tatgtatagt | atftttgttg | tgttttgtat | tttgtftttt | ggtatfttgag | gatfttttaag | 1080 |
| tatgtgtggg | ttttttttgt | gtatataatg | gagtatfttat | gtgatftttt | gtattagtaa | 1140 |
| aatttaggga | tatgggattt | atfttttttt | gtttgaggat | taagtattgg | ttatgatagg | 1200 |
| ggaaggtgag | agatgataaa | aatagagaga | tagttagaga | ggagtagaga | gttagagggg | 1260 |
| tttaggtatt | gggtagtagt | ttttttatat | ttggggtagg | tgtttgaaag | aatttagagg | 1320 |
| tgtatatgag | tttgaggtgt | tttaggtagg | tattgtftttt | atagggtttt | gtttgagttg | 1380 |
| ttttttaaat | gagtgaattt | aagtfttgggt | ttatfttgggt | ttttatfttgt | tttttagggga | 1440 |
| ggttaaggtg | gaagtgggtg | tagtaggggt | ggggttggat | tttttaggag | tggggttggag | 1500 |
| ttattaggag | ttgggggttg | gggtgatgat | ttggagtgtg | tagtagggaa | gatgaggtaa | 1560 |
| tagggtagga | agtgggtggg | gggaggtgga | attggggttg | tgttttgtgt | tgtttggaat | 1620 |
| tgggagtgtg | ggaaagatat | taggaatftt | gttgtagtgt | agtftttggt | gtggggtttg | 1680 |
| gttggtttat | tgtatagagt | tttttttgat | ttttgaagaa | agagatfttgt | ttgtagtggg | 1740 |
| taaaagtttg | tttggatftt | ttggttatta | gaaatatgag | tatggtgggt | gtfttttagtt | 1800 |
| ttttatfttat | gtttgggttt | aagagattgg | gagtfttaggt | ttatftgatt | tttgagaaag | 1860 |
| attaagattt | tgtatftttg | aaagaggttt | ggggtatfttt | gttttgtgta | agggtagaag | 1920 |
| gattagttgt | ttttttgagt | atftttaatt | ggaatftttg | tttgaagttg | agataggaga | 1980 |

| | | | | | | |
|-------------|------------|------------|-------------|-------------|-------------|------|
| ttggatgtga | ggttttttta | gagttggttt | tttttaaaata | attttttaaaa | tttttagatt | 2040 |
| ttaggggtat | ggtgaaatth | tttaaagtag | tttaaagaat | ataatgagag | ttttaaatatt | 2100 |
| ttaggtgggtg | gtgtgttgg | tttttgag | gggtgggat | gtggttgtgt | ggatttatgt | 2160 |
| gtataattgt | gtgggatggg | gttatgtgga | tttatgtgta | taattgtggg | atttttagtgt | 2220 |
| tagtgggatt | ttagtgttag | tgggatttta | gtgttagtgg | gatttttagt | ttagtgggat | 2280 |
| tttagtgta | gtgggatttt | agtgttagt | ggattttagt | gtagtgggt | ttgtggttta | 2340 |
| gtggagtgag | tggagtgtt | gtgatttgag | tggagattgt | gttttggatg | tttttagtta | 2400 |
| gatgttaagt | tatagtttgt | gtagtagtag | taaaggggaa | gggtaggag | ttgggtatag | 2460 |
| ttggatttg | aggttgtgat | ttaggggaaa | gtgtgggtgg | ttgatttagg | gtagtgtgg | 2520 |
| tggtaggta | ggtgggttt | ttgtttttg | gagttgttt | tttttatatt | ggttttgg | 2580 |
| gtttttagta | gtttttatt | tggttttttg | tggttattgt | gggatttgg | ttgtttgta | 2640 |
| gtttagtggg | gagtgaatta | gtgttttttt | ttgtttttgg | ttttttgat | ggtatgagga | 2700 |
| atttttgth | tgttttatag | atttttggt | tttgttgagt | gtggtattgg | agtttgttt | 2760 |
| gtagggth | tggaattaga | gaaagtgtt | ttttggttat | ttgaagtgtt | ggatttttat | 2820 |
| agtgttttt | agtttgggt | ggagtgggt | ttgtgtgtt | gaaggttgg | gtttttgg | 2880 |
| tgaaagggag | gtagtgttag | tttagtth | attttagaag | tggttttgt | attgttgtgg | 2940 |
| tgggtgttt | tgggttttga | ttttgttag | gttgtgggt | agaggtatt | ggagtttgta | 3000 |
| gggttagat | ttgggttga | aaagtthgt | tgattgtagg | taagtgttt | ggaggggtg | 3060 |
| ttaggtgaag | ttttggtgt | ttattatata | tttttgggt | ttatgttag | tgtatttgt | 3120 |
| gtattggta | gaaaatgta | gggttgagg | tgatttagg | agtataagg | agtttttat | 3180 |
| ttttgttta | tatttgttat | tttagtth | gtaatttatt | ttagatata | agaaagttag | 3240 |
| taggattggt | ggggagatg | agttaaatag | gaatatttt | tagtagtgag | taggggttgt | 3300 |
| atgggatgt | ggaggagth | agaggaggt | tggagagtgt | ttgaggttgg | gtgagtgtt | 3360 |
| agaggggaga | tagttgaatt | gggttaaga | ggtgttagt | gggtgtttgt | tgaatgaatg | 3420 |
| agtgatgggt | tttgaagth | gagtgtatt | aaagagggg | tgtgtaaaaa | gggtttttt | 3480 |
| tattatata | gatatagtat | atgtaaatth | ttttttgtg | gaaaagttag | ataggttaaa | 3540 |
| aaggtataa | ataaattagt | tgggtatggt | ggtgtgtgt | tgtagtthta | gttattaggg | 3600 |
| aggttgagth | aggggaattg | tttgaattt | ggaggtggag | attgtagtga | gtaagattg | 3660 |
| tgttattgta | ttttagth | gaaatagagt | gagattttgt | tttggaaaa | aaaaaaaaa | 3720 |
| gtataaatt | gtgtgtgggt | tttaggttat | ataatagag | ttggagggga | gtggttaagg | 3780 |
| atgagaattg | agatggatt | ttgtttttt | ttggaggaga | gtgggtggt | gtttatttgg | 3840 |
| gggtgggaa | ttttttttta | tgggttagt | tgtttaatt | taggggatt | ttaggatagg | 3900 |
| agttgatgta | aatagttgt | ttattttttg | ttgtttttgg | ttttggagaa | ggaggagggga | 3960 |
| gtaggggag | gthttttatt | tttagataat | ttttaagtag | ttaggatatg | ggtgagatga | 4020 |
| gtgagatatt | gatttttggg | atagaattt | agaggtgtt | aaaaattta | gtaattaaga | 4080 |
| taaataggth | gggtgtagt | gthtatgtt | gtaattttag | tattttggga | ggttggatta | 4140 |
| tttaggth | agagthttag | attagthtgg | ttaagatggt | gaaattttat | ttttattaaa | 4200 |
| aatataaaaa | ttagthttag | gtggtggtgt | tagthtgtaa | ttttagtht | ttaggaggtt | 4260 |
| gaggtaaag | aattgthtga | tttaggaggt | agaggttga | gtgagthttag | attatgtht | 4320 |
| tgtatttag | tttgataat | agagggagat | tattttaaaa | aaaaaaaaa | aaaaaaaaa | 4380 |
| aaaaaagag | tgggtggtg | gtggttata | ttatgtgatt | ttagtattt | gggaggttga | 4440 |
| ggtgggtgga | ttattttag | ttggagtht | gagattagth | tgggttaaat | ggttaaat | 4500 |
| tgtttttatt | aaaaatata | aaattagtht | ggtgggtg | taggtattt | taatttttag | 4560 |
| tattttggag | gthttagtag | gagaatttt | tgaatttgt | gggtggaggt | tgtagtgaat | 4620 |
| taagattata | ttattgtatt | ttagthtgg | taataatagt | aaaattttgt | tttaaaaaa | 4680 |
| aaaaaattt | tttttttga | gatatagtt | tattttttt | tttaggttgg | ggtgtagtat | 4740 |
| tatgattth | gthtattgta | atthttgtt | tttagattt | tgtattttag | tttttaagt | 4800 |
| agttgggatt | ataggtatt | gthtattat | ttagthtaatt | tttgtattt | tagtaggggt | 4860 |
| gtgthttat | tatgthtgg | aggttggth | tgaattttt | atthttaagt | atthttht | 4920 |
| tttagtht | taaggtt | ggattatagg | tgtgagtht | tatgthtgg | tttttaaat | 4980 |
| gaaaatag | taaaaaatt | tgataaata | aatattaaaa | atthtattg | tttatttt | 5040 |
| tataatttt | ttttattgt | tttttaggt | atthttgtt | ttagaaagta | atthtaaaaa | 5100 |
| tgtgtagat | gagthtggat | ttatttga | aatggtgga | gthtatgga | atthtggag | 5160 |
| aggggagtg | aggatagaaa | ttatagtga | aagaaattt | gggttgggt | tagtggtht | 5220 |
| tgtttgta | tttagtatt | tgggaggtt | aggtaggtg | attattttag | gthtaggag | 5280 |
| tgagattagt | ttgattaata | tgggtgaaat | ttattttat | taaaaatata | aaaaaatta | 5340 |
| gthtaggtat | gtggtgtat | ttttagtht | tagthtatt | ggaggtttag | ataggaaaat | 5400 |
| tgtttgaatt | tgggaggtg | aggtttagt | gagthtaag | tgtgttatt | tatttttag | 5460 |
| tgggtaata | gagtaaaatt | ttattttaaa | aaaaagaaa | gaaagaaatt | ttttggtagt | 5520 |
| tgatgagaag | gaaattta | tggtaggtt | tagtagggga | gatgaggaga | ttttagggag | 5580 |
| gthatttga | tatgthtgt | tttagtgg | gthtaggag | aggttatt | ttttttgt | 5640 |
| tattttttt | tgtttta | tttttaagt | ttttagtag | tgggtattt | aggttagtht | 5700 |
| aaggaattat | atgthtag | atthtttag | ggtgtttgt | aaaaatgta | atthtggth | 5760 |

ES 2 597 845 T3

```
ggtgtagtgg tttatatttg taattttagt attttgggag gttgaggtgg gtggattatg 5820
aggtaggag attgagatta ttttggtaaa tatggtgaaa ttttattttt attaaaaaaa 5880
taaaaataaa ttaaaaaaaaa tattagttgg gtgtggtggt ggtggtttgt agttttagtt 5940
at ttggaggagg ttgaggtagg agaatggtgt gaatttggga ggtggagttt gtagtgagtt 6000
gagattgtgt tattgtattt tagtttgggt gatagagtga gattttgttt taaaaaaaaa 6060
aaagtaaatt ttttgggtat tattttatat tgattg 6096
```

<210> 10

<211> 6096

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Constructo sintético: ADN genómico tratado químicamente (Homo sapiens)

<400> 10

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tagttaatgt | ggggtggtgt | ttaagaaatt | tgtttttttt | tttttgaggt | agagttttgt | 60 |
| tttgttgttt | agggtggggt | gtagtgggtg | aatttttagtt | tattgtaagt | tttgtttttt | 120 |
| gggtttatgt | tatttttttg | ttttagtttt | ttgagtagtt | gggattatag | gtgtttgtta | 180 |
| ttatgtttag | ttaatgtttt | ttttggtttg | tttttgtttt | tttagtagag | atggggtttt | 240 |
| attgtgtttg | ttagatgggt | tttgattttt | tgattttgtg | atttatttgt | 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 | ttagagggtt | tttttttttt | 600 |
| ttttttttta | agatggagtt | ttgtttttgt | tgtttagggt | ggagtgtaat | ggtataaatt | 660 |
| tggtttattg | taatttttgt | tttttggggt | taagtgattt | ttttgtttta | gtttttggag | 720 |
| tagttgggat | tataggtgtg | tattattatg | tttggttaat | tttttttgta | tttttagtag | 780 |
| agatgatatt | ttattatggt | ggtagggttg | gttttaattt | tttgatttta | ggtaatttat | 840 |
| ttgttttggg | tttttaaagt | gttgggatta | taggtataag | ttattgtgtt | tggtttaggg | 900 |
| ttttttttat | atataatttt | tattttttat | ttttttgttt | taaggttttt | tatagttttt | 960 |
| attattttta | ggtaaagttt | aaattttatt | tgtatatttt | ttaagttggt | ttttggggtt | 1020 |
| gagagtagtt | tgggaggtag | gtgagaaagg | gttgtggaag | tgtaagttta | gtaaattttt | 1080 |
| gatattttgt | ttattgtgga | tttttgtatt | atttttattt | aaaaaggtta | agtgtggtgg | 1140 |
| tttatgtttg | taatttttagt | attttgggtg | gttgaagtgg | gtaggttatt | tgaggttagg | 1200 |
| agtttaagat | tagtttggtt | aatatggtga | aattatgttt | ttattaaana | tataanaaatt | 1260 |
| agttgggtgt | ggtgataggt | atttgtaatt | ttagttattt | gggaggttga | ggtatgagaa | 1320 |
| tttgagaggt | agaggttgta | gtgagttgag | attgtggtgt | tgtattttta | tttgggtgag | 1380 |
| agagtaaaat | tgtgttttga | aaaaaaagat | tttttttttt | tttgagatag | agttttgttg | 1440 |
| ttgttgttta | ggttggagtg | taatggtgtg | attttgggtt | attgtaattt | ttgttttgta | 1500 |
| ggtttaaggg | atttttttgt | tttagttttt | ggagtagttg | ggattatagg | tgtttgttat | 1560 |
| tttatttagt | taatttttgt | attttttagta | gaaatggggt | tttattatat | tggtttaggtt | 1620 |
| ggttttgaat | tttagatttt | aggtgattta | tttgttttgg | ttttttaaag | tgttgggatt | 1680 |
| atatggtgtg | agttattggt | gtttggtttt | tttttttttt | tttttttttt | tttttttttt | 1740 |
| gagatagttt | ttttttgttg | tttaggttgg | agtatagtgg | tatgattttg | gtttattgta | 1800 |
| atttttgttt | tttgggttaa | gtaatttttt | tgttttagtt | ttttgagtgg | ttgggattat | 1860 |
| aggttagtgt | tattatattg | ggttaatttt | tgtattttta | gtagagatgg | ggttttatta | 1920 |
| ttttggttag | gttggttttg | aatttttgat | tttaagtgat | ttggtttttt | aaagtgttgg | 1980 |
| gattatagat | gtgagttatt | gtgtttgggt | tatttatttt | gattattgag | ttttttggta | 2040 |
| tttttttaaa | ttttgtttta | gaagttagta | ttttatttat | tttatttatg | ttttggttgt | 2100 |
| ttagagattg | tttgggaagt | ggagattttt | tttagttttt | tttttttttt | ttagggttaa | 2160 |
| agatagtaag | gaataggggtg | attgtttata | ttagtttttg | tttttagagat | tttttgagat | 2220 |
| tggatagttg | agtttgtgga | gagggatttt | ttatttttaa | gtaggtaatt | atttattttt | 2280 |
| tttttagaggg | aatgagggat | ttatttttagt | ttttattttt | gattattttt | tttttagtttt | 2340 |
| gattgtataa | tttgaaattt | atataatggtt | tgtaattttt | tttttttttt | tttgagatgg | 2400 |
| agttttgttt | tgtttttagg | ttggagtgtg | gtggtgtgat | tttggtttat | tgtaattttt | 2460 |
| gttttttggg | tttaagtgat | ttttttgggt | tagttttttt | agtagttggg | attatagatg | 2520 |
| tgtattatta | tgttttggta | 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 | ttgttgga | 2820 |
| atatttttgt | taagttttgt | ttttttatta | gttttgtttg | ttttttgtgt | gtttgggata | 2880 |
| ggttgtaaaa | ttggaggtga | taaagtgtgg | taggaaatgg | agggtttttt | tatattttta | 2940 |
| gggttggttt | tagttttggt | attttttgtt | taatattgtg | gatgtaattg | gtatgggatt | 3000 |
| tggaagtgtg | tggtaaagtg | ttggggtttt | gtttggttgt | tttttttggg | tgtttgtttg | 3060 |
| tagttaagtga | agttttttta | atttaggttt | gggttttgtg | agtttttaggt | gtttttgttt | 3120 |
| tgtggtggtg | gtgaagttag | agtttgagaa | tgtttattgt | agtgatgtga | aggttgtttt | 3180 |
| tgggggtggg | ttgaggttgt | agttgttttt | tttttgtatt | aaggatttta | attttttagtg | 3240 |
| atgtagtgtg | tgtttttgtt | taggttggga | ggtattgtag | ggatttgatg | tttttaggtgg | 3300 |
| ttaaagagtg | attttttttg | attttagggt | tttgggtggg | taggttttag | tattgtattt | 3360 |
| ggtggaggtt | gaaggtttgt | ggggtaggat | aggagttttt | tgtggtggtg | gaagggttga | 3420 |
| ggatgaagga | gggtgttaat | ttatttttta | ttgggttggg | ggtaatgttg | aattttgtag | 3480 |
| tgattgtgga | gggttaaggt | gaaaattggt | gggggtggtg | agggtaggtg | tggggagggg | 3540 |
| tggtttttagg | gagtaaggag | tttatttgtt | ttgttgttgt | agttgttttg | ggttgattgt | 3600 |
| ttatgttttt | ttttgggtta | tgatttttgg | atttaattgt | gtttggtttt | tgtttttttt | 3660 |
| tttttgttgt | tgttgtgtgg | gttgtaattt | gatgttttag | ttgggggtgt | taggggtgtag | 3720 |
| tttttgttta | ggttgttagt | gttttatttg | ttttattggg | ttatagattt | gttgggtggtg | 3780 |
| gggttttgtt | ggtgttgggg | ttttgttggg | gttggggttt | tgttgggtgt | ggggttttgt | 3840 |
| tgggtgttggg | gttttgttgg | tgttgggggt | ttgttgggtg | tggggttttg | tggttgtgta | 3900 |
| tgtgagtttg | tgtggttttg | ttttgtgtgg | ttgtgtatgt | gagtttgtgt | ggttgtgttt | 3960 |
| tgttttgttt | tagggagtta | gtgtgttgtt | atttgggatg | ttaggatttt | tgttgtgttt | 4020 |
| tttggattgt | tttgggggat | tttgggtgat | ttttaggatt | taggagtttt | ggaagtgttt | 4080 |
| tgagagaaat | tagttttggg | agggttttgt | atttagtttt | ttgttttggg | tttggattgg | 4140 |
| ggttttgggt | taaggtgttt | agaggaatag | ttgatttttt | tatttttgtg | tagggtagag | 4200 |
| atttttaaat | ttttttttaa | aatgtagggt | tttagttttt | tttagggagt | tagtgaattt | 4260 |
| agatttttag | ttttttgagt | ttaagtatga | atagggaaat | ggggattatt | attatgttta | 4320 |
| tatttttggg | ggttaggaag | tttaggtagg | tttttgttta | ttgtagatgg | attttttttt | 4380 |
| ttaggggtta | agaaaggttt | tgtatagtaa | gttaattaag | ttttattagt | agagttgtgt | 4440 |
| tgtaattagg | tttttagtgt | tttttttata | tttttagttt | taagtgatat | aggatatagt | 4500 |
| tttaatttta | ttttttttta | tttatttttt | gttttgttgt | tttatttttt | ttgttatata | 4560 |
| ttttaagtta | tttatttaaat | tttttagttt | tggttaattta | gttttagttt | ttggaagttt | 4620 |
| agtttttagtt | ttgttattat | tatttttatt | ttggtttttt | ttgagaataa | gtggagggta | 4680 |
| aatagagttt | aggtttgaat | ttatttgttt | gaaaaataat | ttaagttaaa | ttttgtggga | 4740 |
| gtagtgtttg | tttgggggat | tttaggttta | tgtgtatttt | tgaatttttt | tgggtatttg | 4800 |
| ttttaaatgt | aaagaggttg | ttatttaatg | tttgggtttt | tttgattttt | tgtttttttt | 4860 |
| tggttgtttt | tttgtttttg | ttgtttttta | tttttttttg | ttatggttag | tatttggttt | 4920 |
| ttaggttaga | aaaggtggat | tttgtgtttt | tggattttat | taatgttagg | agttatataa | 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 | atttatata | aagttttttt | aagtttaaat | gttgaagag | 5340 |
| atttttatag | aaagaaaatt | tttttagagg | tttttaaggt | tttgttttgg | aggaagagga | 5400 |
| agaaagtgtt | tgtttaggtat | tgaaatgtta | aggatttgtg | aagtgaaaat | tattttttta | 5460 |
| atttttgtgg | aatagttatt | ttatttttat | tttttaggga | aataggggag | tggtttaatt | 5520 |
| tgttttgtaa | attaatttta | gatttataag | agtgtttttt | tttttgggtg | ggagaggtta | 5580 |
| ggtttttagt | gttgtagttt | agtgaatgtt | gatttgttga | ggtttatttt | aggttatagg | 5640 |
| gtgttgggta | aatttattaa | gtttgtatgt | ttgtttaata | tttgggttgg | ttaggaagag | 5700 |
| tgtttgggtg | atatttagtt | tgggtgatata | taggaggtat | ggttgggtgaa | gaaaatggat | 5760 |
| ttgtgtatta | agggtttgta | ttattttgtt | gggttttgtt | ttgatagatt | attttttatt | 5820 |
| ttttttgatt | atggattggg | ttaggaggtt | tatttttggg | ttgtttattg | tttgtaaata | 5880 |
| gtagagggta | gtgggtgagg | tgtataaagt | tttagttgtt | tatgtttttt | taagtttttg | 5940 |
| ttatttaggg | tttttgtttg | ttttgtttta | gggtgtattg | ttttgttttg | ggtgggtgatg | 6000 |
| ggtaggttta | ggaattgtta | tgtttaggta | ttgagttgga | gttttgggta | ttttattgtg | 6060 |
| tagtgtaaaa | agggaaattt | gaattttata | ttggta | | | 6096 |

<210> 11
<400> 11
000
<210> 12
5 <400> 12
000
<210> 13
<400> 13
000
10 <210> 14
<400> 14
000
<210> 15
<400> 15
15 000
<210> 16
<400> 16
000
<210> 17
20 <400> 17
000
<210> 18
<400> 18
000
25 <210> 19
<400> 19
000
<210> 20
<400> 20
30 000
<210> 21

<400> 21
000
<210> 22
<400> 22
5 000
<210> 23
<400> 23
000
<210> 24
10 <400> 24
000
<210> 25
<400> 25
000
15 <210> 26
<400> 26
000
<210> 27
<400> 27
20 000
<210> 28
<400> 28
000
<210> 29
25 <400> 29
000
<210> 30
<400> 30
000
30 <210> 31
<400> 31

ES 2 597 845 T3

000
<210> 32
<400> 32
000
5 <210> 33
<400> 33
000
<210> 34
<400> 34
10 000
<210> 35
<400> 35
000
<210> 36
15 <400> 36
000
<210> 37
<400> 37
000
20 <210> 38
<400> 38
000
<210> 39
<400> 39
25 000
<210> 40
<400> 40
000
<210> 41
30 <400> 41
000

<210> 42
<400> 42
000
<210> 43
5 <400> 43
000
<210> 44
<400> 44
000
10 <210> 45
<400> 45
000
<210> 46
<211> 20
15 <212> ADN
<213> Homo Sapiens
<400> 46
aaacccaaac ctaaattaa 20
<210> 47
20 <400> 47
000
<210> 48
<400> 48
000
25 <210> 49
<400> 49
000
<210> 50
<400> 50
30 000
<210> 51

ES 2 597 845 T3

<211> 17

<212> ADN

<213> Homo Sapiens

<400> 51

5 ggaagtgtgt ggtaaag 17

<210> 52

<400> 52

000

<210> 53

10 <400> 53

000

<210> 54

<400> 54

000

15 <210> 55

<400> 55

000

<210> 56

<211> 28

20 <212> ADN

<213> Homo Sapiens

<400> 56

taaagtgttg gggtttgtt tggttgtt 28

<210> 57

25 <400> 57

000

<210> 58

<400> 58

000

30 <210> 59

<400> 59

<210> 60
<400> 60
000
<210> 61
5 <211> 23
<212> ADN
<213> Homo Sapiens
<400> 61
aaacaaacgt ccgaaaaaaaa cga 23
10 <210> 62
<400> 62
000
<210> 63
<400> 63
15 000
<210> 64
<400> 64
000
<210> 65
20 <400> 65
000
<210> 66
<211> 18
<212> ADN
25 <213> Homo Sapiens
<400> 66
caaacgaaac cccgacgc 18
<210> 67
<400> 67
30 000
<210> 68

<400> 68
000
<210> 69
<400> 69
5 000
<210> 70
<400> 70
000
<210> 71
10 <400> 71
000
<210> 72
<400> 72
000
15 <210> 73
<400> 73
000
<210> 74
<400> 74
20 000
<210> 75
<400> 75
000
<210> 76
25 <400> 76
000
<210> 77
<211> 25
<212> ADN
30 <213> Homo Sapiens
<400> 77

tggatgga ggagtttag taagt 25
<210> 78
<211> 27
<212> ADN
5 <213> Homo Sapiens
<400> 78
aaccaataaa acctactcct cccttaa 27
<210> 79
<211> 30
10 <212> ADN
<213> Homo Sapiens
<400> 79
accaccacc aacacacaat aacaaacaca 30
<210> 80
15 <211> 25
<212> ADN
<213> Homo Sapiens
<400> 80
taagagtaat aatggatgga tgatg 25
20 <210> 81
<211> 17
<212> ADN
<213> Homo Sapiens
<400> 81
25 cctccatct cccttc 17
<210> 82
<211> 25
<212> ADN
<213> Homo Sapiens
30 <400> 82
atggatgaag aaagaaagga tgagt 25

<210> 83
<400> 83
000
<210> 84
5 <400> 84
000
<210> 85
<400> 85
000
10 <210> 86
<400> 86
000
<210> 87
<400> 87
15 000
<210> 88
<400> 88
000
<210> 89
20 <400> 89
000
<210> 90
<400> 90
000
25 <210> 91
<400> 91
000
<210> 92
<400> 92
30 000
<210> 93

<400> 93
000
<210> 94
<400> 94
5 000
<210> 95
<400> 95
000
<210> 96
10 <400> 96
000
<210> 97
<400> 97
000
15 <210> 98
<211> 20
<212> ADN
<213> Homo Sapiens
<400> 98
20 aaacccaaac ctaaattaa 20
<210> 99
<211> 17
<212> ADN
<213> Homo Sapiens
25 <400> 99
ggaagtgtgt ggtaaag 17
<210> 100
<211> 28
<212> ADN
30 <213> Homo Sapiens
<400> 100

taaagtgtg gggtttgtt tggttgtt 28

<210> 101

<211> 23

<212> ADN

5 <213> Homo Sapiens

<400> 101

aaacaaacgt ccgaaaaaaaa cga 23

<210> 102

<211> 18

10 <212> ADN

<213> Homo Sapiens

<400> 102

caaacgaaac cccgacgc 18

Reivindicaciones

1. Un método para detectar carcinoma colorrectal o de próstata en un sujeto, que comprende detectar la presencia de metilación de TFAP2E en una muestra biológica aislada de dicho sujeto; en donde
- 5 (i) en el caso de cáncer de próstata la muestra biológica se selecciona del grupo que consiste de tejido prostático, plasma sanguíneo, suero sanguíneo, sangre entera, células obtenidas de la sangre obtenida del sujeto y en orina; y
- (ii) en el caso de carcinoma de colon, la muestra biológica se selecciona del grupo que consiste de tejido de colon, plasma sanguíneo, suero sanguíneo, sangre entera y células obtenidas de la sangre obtenida del sujeto;
- en donde la presencia de metilación de CpG es indicativo de la presencia de dicho trastorno y la ausencia del mismo es indicativo de la presencia de un trastorno proliferativo celular benigno.
- 10 2. Un método para detectar carcinoma colorrectal o la próstata en un sujeto, que comprende poner en contacto ADN genómico aislado de una muestra biológica obtenida de dicho sujeto; en donde
- (i) en el caso de cáncer de próstata la muestra biológica se selecciona del grupo que consiste de tejido de la prostático, plasma sanguíneo, suero sanguíneo, sangre entera, células obtenidas de la sangre obtenida del sujeto y orina; y
- 15 (ii) en el caso de carcinoma de colon, la muestra biológica se selecciona del grupo que consiste de tejido de colon, plasma sanguíneo, suero sanguíneo, sangre entera y células obtenidas de la sangre obtenida del sujeto;
- con al menos un reactivo, o serie de reactivos que distingue entre dinucleótidos CpG metilado y no metilado dentro de al menos una región objetivo del ADN genómico, en donde la región objetivo comprende al menos una secuencia de dinucleótido CpG del gen TFAP2E e hibrida bajo condiciones rigurosas con una secuencia de al menos 16 nucleótidos contiguos de al menos una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2, respectivamente, y la detección de carcinoma colorrectal o próstata con base en la presencia de metilación de CpG en dicha al menos una secuencia de dinucleótido CpG.
- 20 3. Un método para detectar carcinoma colorrectal o de próstata, que comprende:
- a) extraer o bien el aislar ADN genómico de una muestra biológica obtenida del sujeto; en donde
- 25 (i) en el caso de cáncer de próstata la muestra biológica se selecciona del grupo que consiste de tejido de próstata, plasma sanguíneo, suero sanguíneo, sangre entera, células obtenidas de la sangre obtenida del sujeto y orina; y
- (ii) en el caso de carcinoma de colon, la muestra biológica se selecciona del grupo que consiste de cáncer de colon tejido de colon, plasma sanguíneo, suero sanguíneo y sangre entera, células obtenidas de la sangre obtenida del sujeto;
- b) tratar el ADN genómico de a) con uno o más reactivos, preferiblemente un reactivo seleccionado de entre el grupo que consiste de bisulfito, sulfito de hidrógeno, disulfito, y sus combinaciones, para convertir las bases de citosina que están no metilado en la posición 5 del mismo en uracilo o en otra base que es detectablemente diferente de la citosina en términos de propiedades de hibridación;
- 30 c) poner en contacto el ADN genómico tratado con una enzima de amplificación y al menos un cebador que comprende, una secuencia contigua de al menos 9 nucleótidos que es complementaria a, o se hibrida bajo condiciones rigurosas con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y complementos de las mismas y en dicho cebador comprende al menos un dinucleótido CpG, TpG o CpA, en donde el ADN genómico tratado o el fragmento del mismo, es o bien amplificado para producir al menos un amplificado, o no es amplificado; y
- 35 d) determinar, con base en la presencia o ausencia de, o una propiedad de dicho amplificado, el estado o nivel de metilación de al menos un dinucleótido CpG de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2, o un promedio, o un valor que refleja un estado o nivel de metilación promedio de una pluralidad de dinucleótidos CpG de una secuencia seleccionada de los grupos que consisten de la SEQ ID NO: 1 a la SEQ ID NO: 2, con lo que la detección de carcinoma colorrectal o de próstata es, al menos en parte, permitida.
- 40 4. El método de la reivindicación 3, en donde el contacto o amplificación en c) comprende el uso de al menos un método seleccionado del grupo que comprende: el uso de una ADN polimerasa resistente al calor como la enzima de amplificación; el uso de una polimerasa que carece de actividad exonucleasa 5'-3'; el uso de una reacción en cadena de la polimerasa (PCR); la generación de una molécula de ácido nucleico amplificada que porta una etiqueta detectable.
- 45

5. El método de la reivindicación 3, que comprende además en la etapa d) el uso de al menos una molécula de ácido nucleico o molécula de ácido peptidonucleico que comprende en cada caso una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida bajo condiciones moderadamente rigurosas o rigurosas con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y complementos de las mismas, en donde dicha molécula de ácido nucleico o molécula de ácido peptidonucleico suprime la amplificación del ácido nucleico con el cual se hibrida, y/o en donde la determinación en d) comprende la hibridación de al menos una molécula de ácido nucleico o molécula de ácido peptidonucleico en cada caso comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o se hibrida bajo condiciones moderadamente rigurosas o rigurosas con una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y complementos de las mismas.
6. El método de la reivindicación 5, que comprende además extender al menos una de tales moléculas de ácido nucleico hibridado en al menos una base nucleotídica, y/o en donde el contacto o amplificación en c), comprende el uso de cebadores específicos de metilación.
7. Un método para detectar carcinoma colorrectal o de próstata, que comprende:
- a) extraer o bien aislar ADN genómico de una muestra biológica obtenida del sujeto, en donde
- (i) en el caso de cáncer de próstata la muestra biológica se selecciona del grupo que consiste de tejido prostático, plasma sanguíneo, suero sanguíneo, sangre entera, células obtenidas de la sangre obtenida del sujeto y orina; y
- (ii) en el caso de carcinoma de colon, la muestra biológica se selecciona del grupo que consiste de tejido de colon, plasma sanguíneo, suero sanguíneo y sangre entera, células obtenidas de la sangre obtenida del sujeto;
- b) digerir el ADN genómico de a), o un fragmento del mismo, con una o más enzimas de restricción sensibles a la metilación;
- c) poner en contacto la enzima de restricción de DNA digerida de b), con una enzima de amplificación y al menos dos cebadores adecuados para la amplificación de una secuencia que comprende al menos un dinucleótido CpG de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2; y
- d) determinar, con base en la presencia o ausencia de un amplificado el estado o nivel de metilación de al menos un dinucleótido CpG de una secuencia seleccionada del grupo que consiste de la SEQ ID NO: 1 a la SEQ ID NO: 2, con lo cual la detección de carcinoma colorrectal o próstata es, al menos en parte, permitida.
8. El método de acuerdo con la reivindicación 7, en donde la presencia o ausencia de un amplificado se determina por medio de hibridación con al menos un ácido nucleico o ácido peptidonucleico que es idéntico, complementario, o hibrida bajo condiciones rigurosas o altamente rigurosas con al menos un segmento de 16 bases de longitud de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2.
9. El uso de
- a) un ácido nucleico derivado de la SEQ ID NO: 1 a la SEQ ID NO: 2 genómica por tratamiento con bisulfito, en donde dicho ácido nucleico comprende al menos una secuencia dinucleótido CpG, TpG o CpA,
- b) un ácido nucleico, que comprende al menos 16 nucleótidos contiguos de una secuencia de ADN genómico seleccionado del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y secuencias complementarias de las mismas, en donde la secuencia de bases contigua comprende al menos una secuencia de dinucleótido CpG, TpG o CpA, o
- c) un ácido nucleico, que comprende al menos 50 nucleótidos contiguos de una secuencia de ADN genómico seleccionado del grupo que consiste de la SEQ ID NO: 3 a la SEQ ID NO: 10, y secuencias complementarias de las mismas, en donde la secuencia de bases contigua comprende al menos una secuencia de dinucleótido CpG, TpG o CpA
- para diagnóstico de carcinoma colorrectal o de próstata mediante análisis de metilación de una muestra biológica, en donde
- (i) en el caso de cáncer de próstata la muestra biológica se selecciona del grupo que consiste de tejido prostático, plasma sanguíneo, suero sanguíneo, sangre entera, células obtenidas de la sangre obtenida del sujeto y orina; y

(ii) en el caso de carcinoma de colon, la muestra biológica se selecciona del grupo que consiste de tejido de colon, plasma sanguíneo, suero sanguíneo, sangre entera y células obtenidas de la sangre obtenida del sujeto.

5 10. Uso de un kit para realizar el método según la reivindicación 3, en donde 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
5
10
15
al menos un conjunto de oligonucleótidos que contiene dos oligonucleótidos cuyas secuencias en cada caso son idénticas, son complementarias, o se hibridan bajo condiciones rigurosas o altamente rigurosas con un segmento de 9, o más preferiblemente 18, bases de longitud de una secuencia seleccionada de la SEQ ID NO: 3 a la SEQ ID NO: 10, o un kit adecuado para realizar el método de acuerdo con la reivindicación 9 que 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 de uno o una pluralidad de ácidos nucleicos o ácidos peptidonucleicos que son idénticos, son complementarias, o hibridan bajo condiciones rigurosas o altamente rigurosas con al menos un segmento de 9 bases de longitud de una secuencia seleccionada de la SEQ ID NO: 1 a la SEQ ID NO: 2; y opcionalmente (d) instrucciones para el uso y la interpretación de los resultados del kit.

FIGURA 1

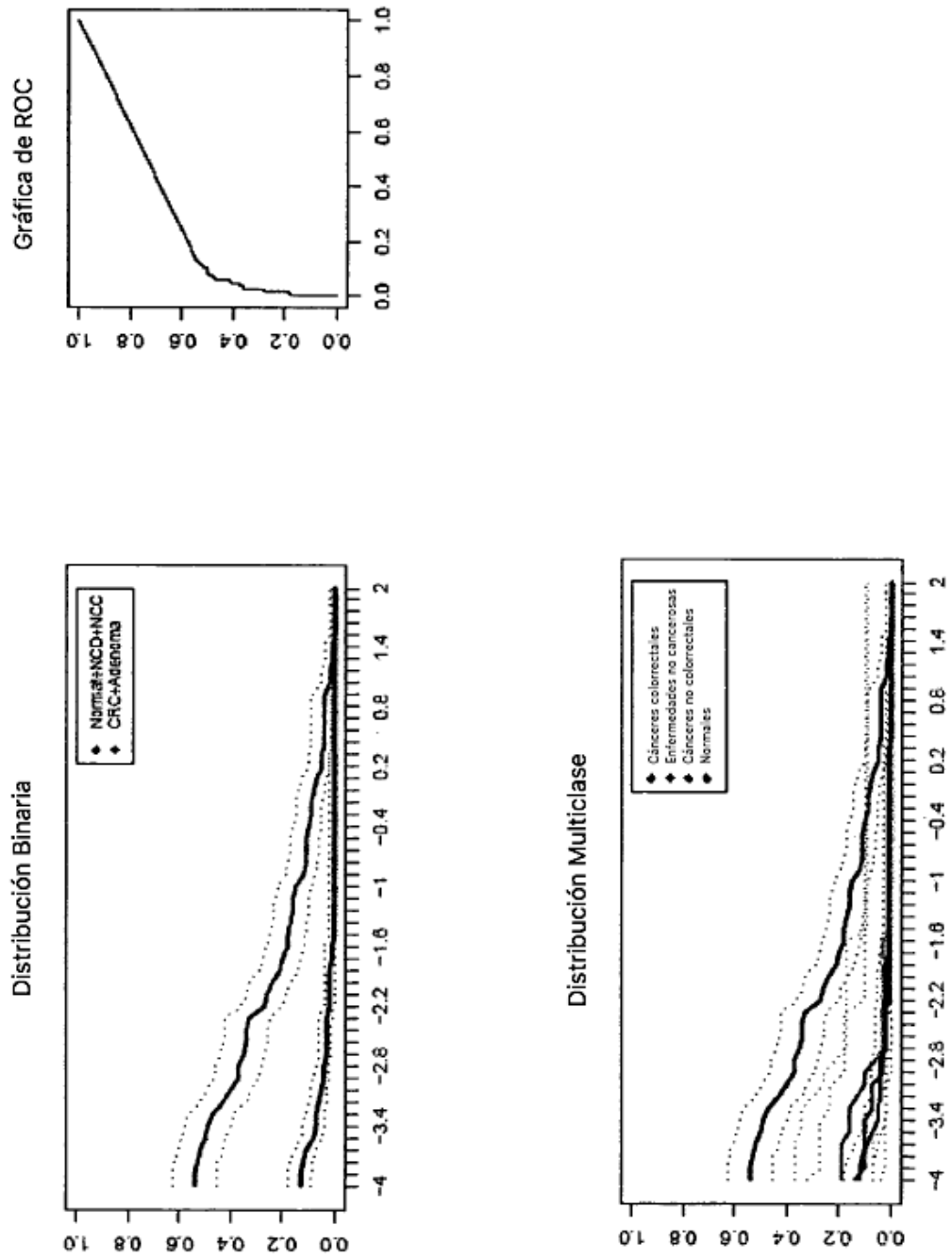


FIGURA 2

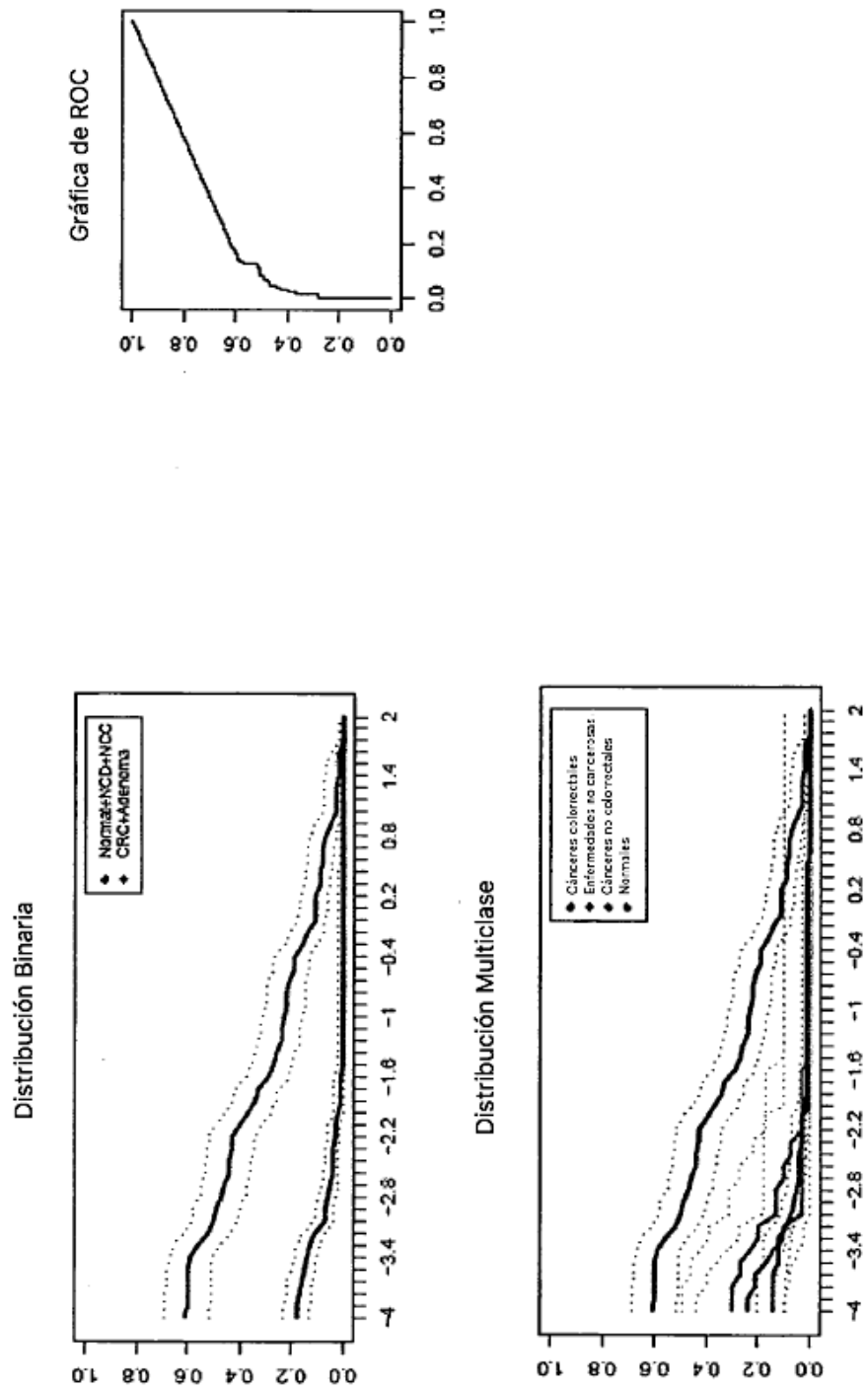


FIGURA 3

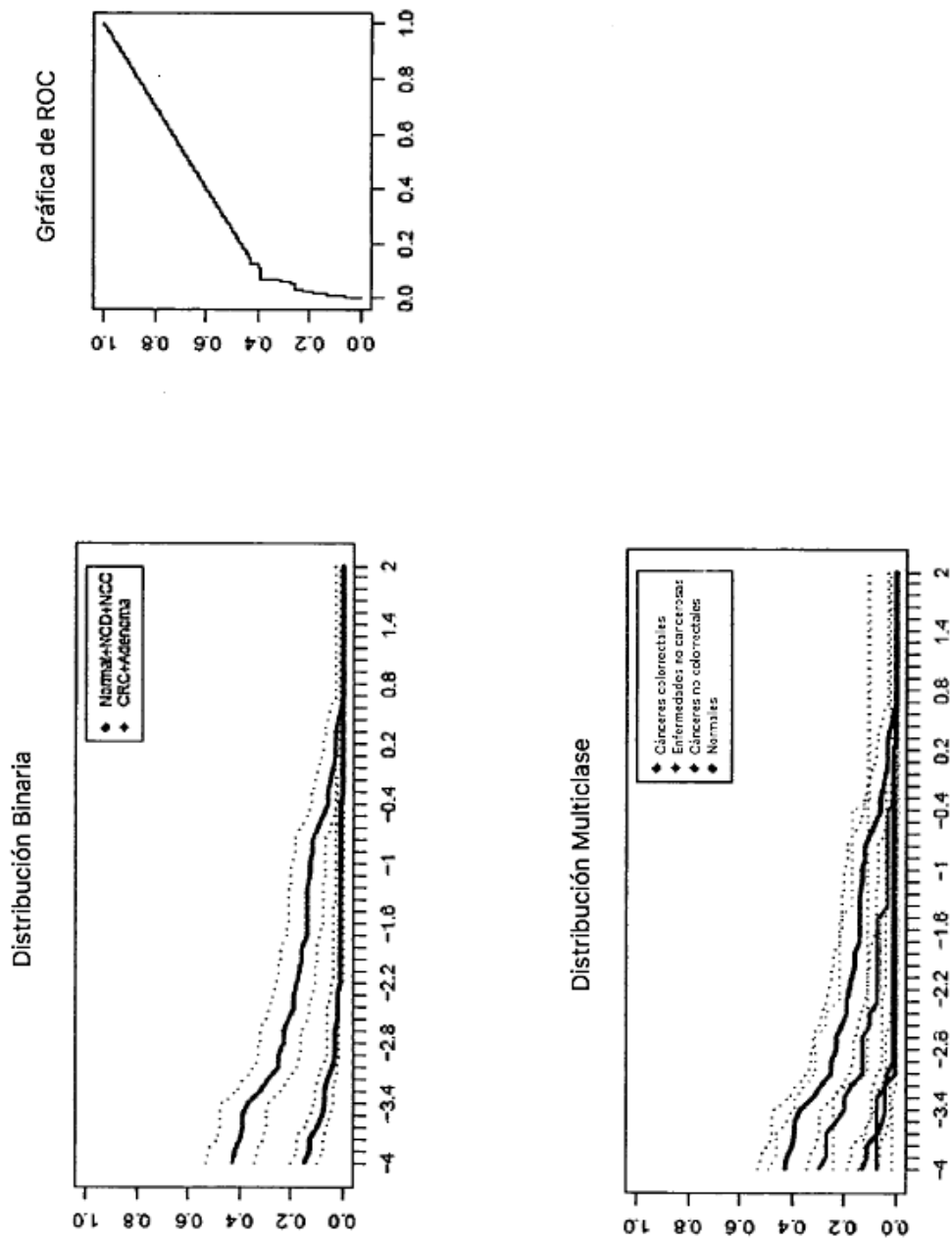


FIGURA 4

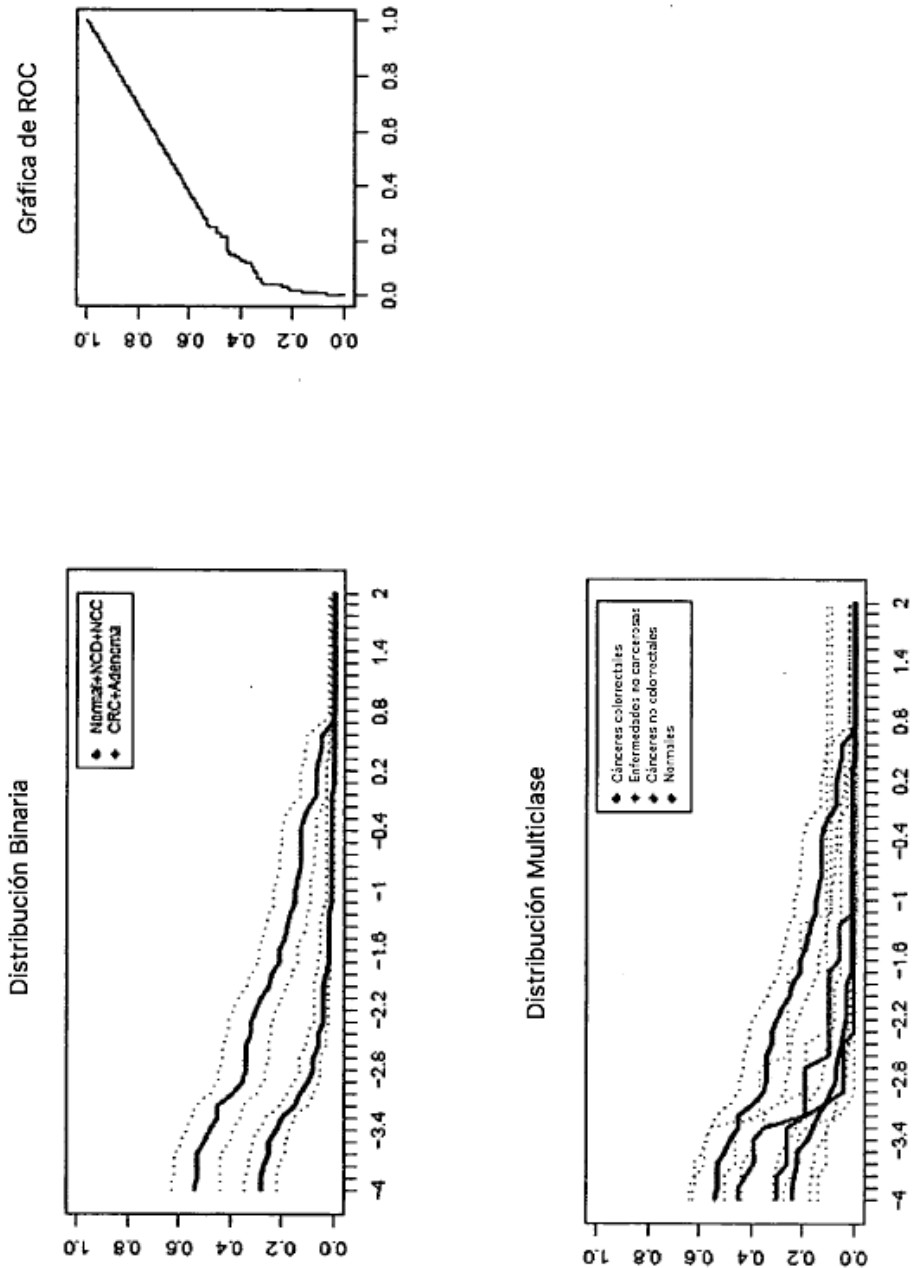


FIGURA 5

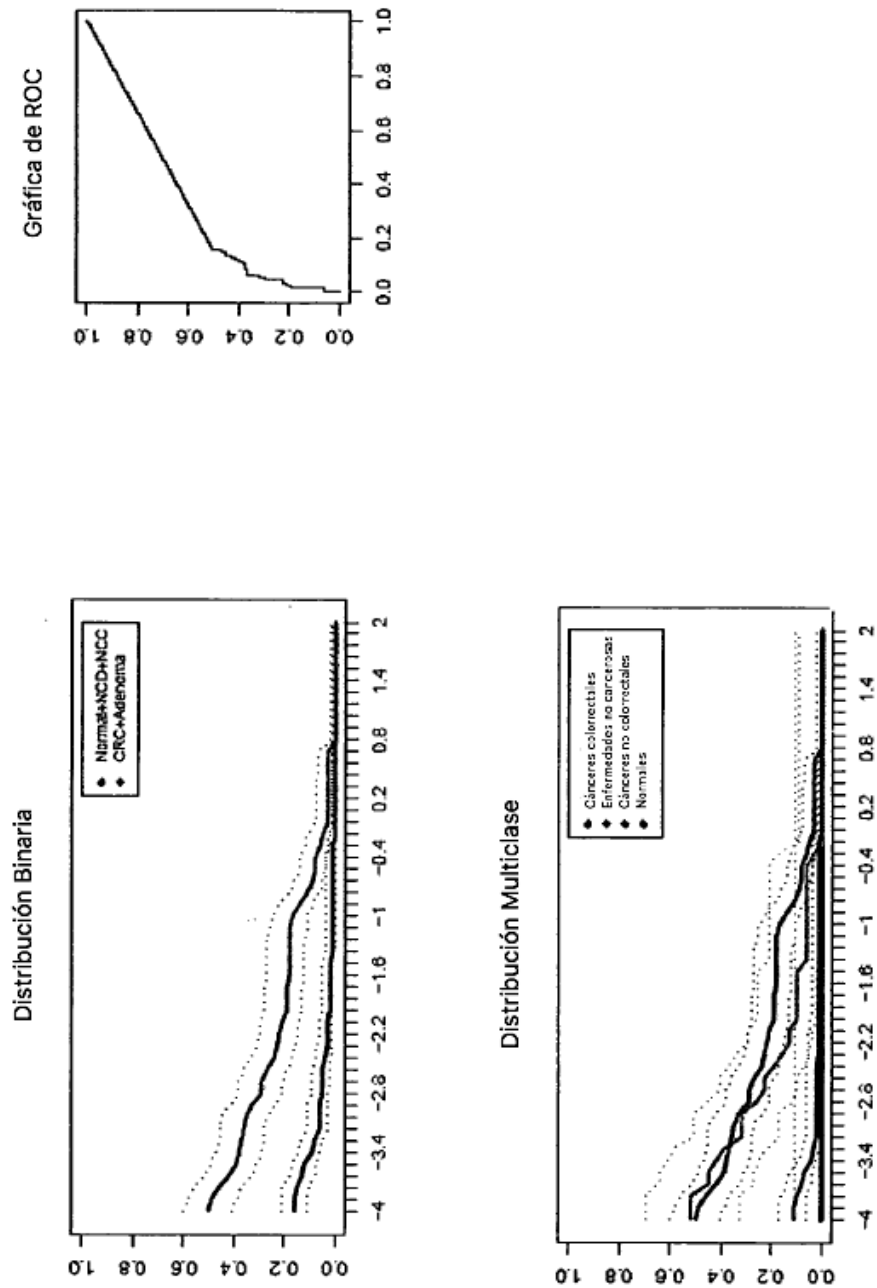


FIGURA 6

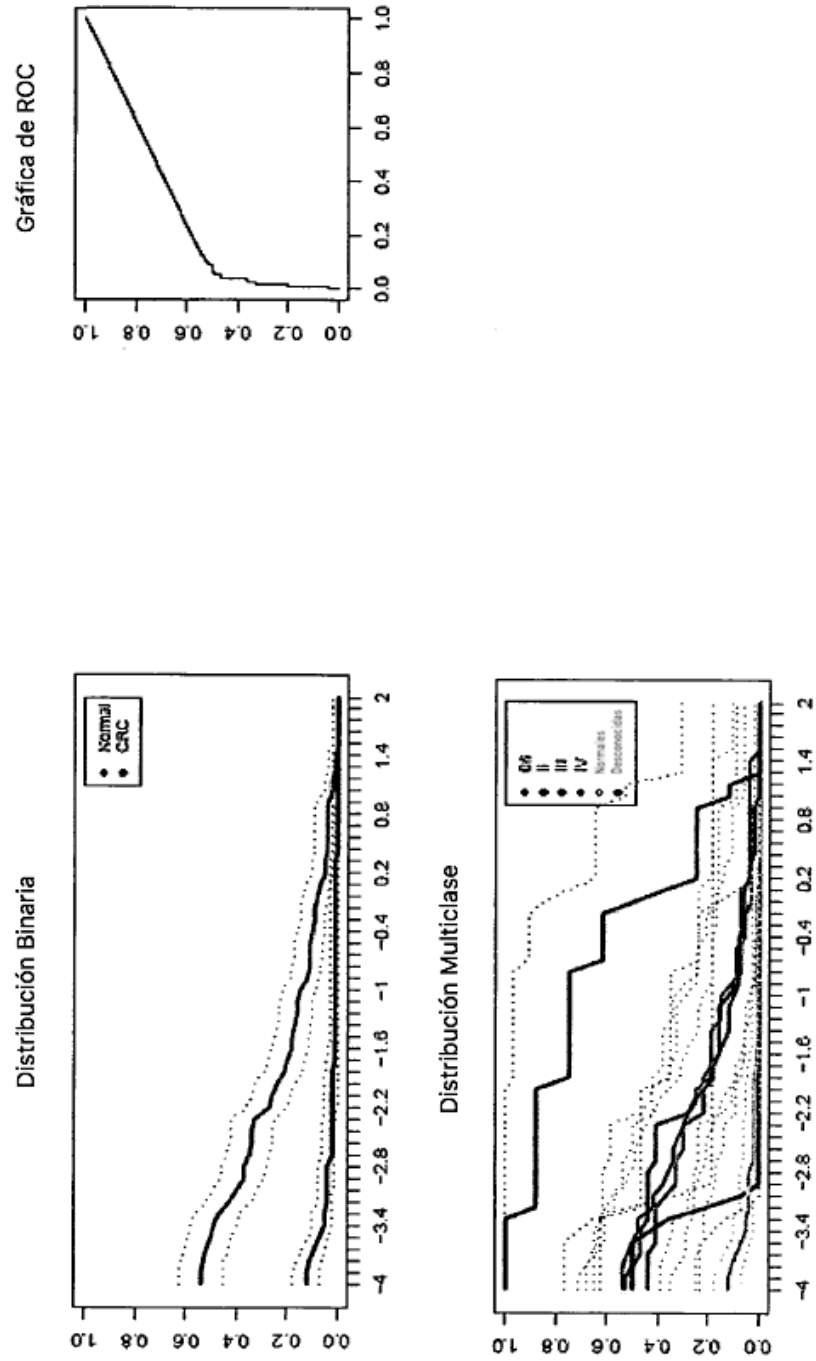


FIGURA 7

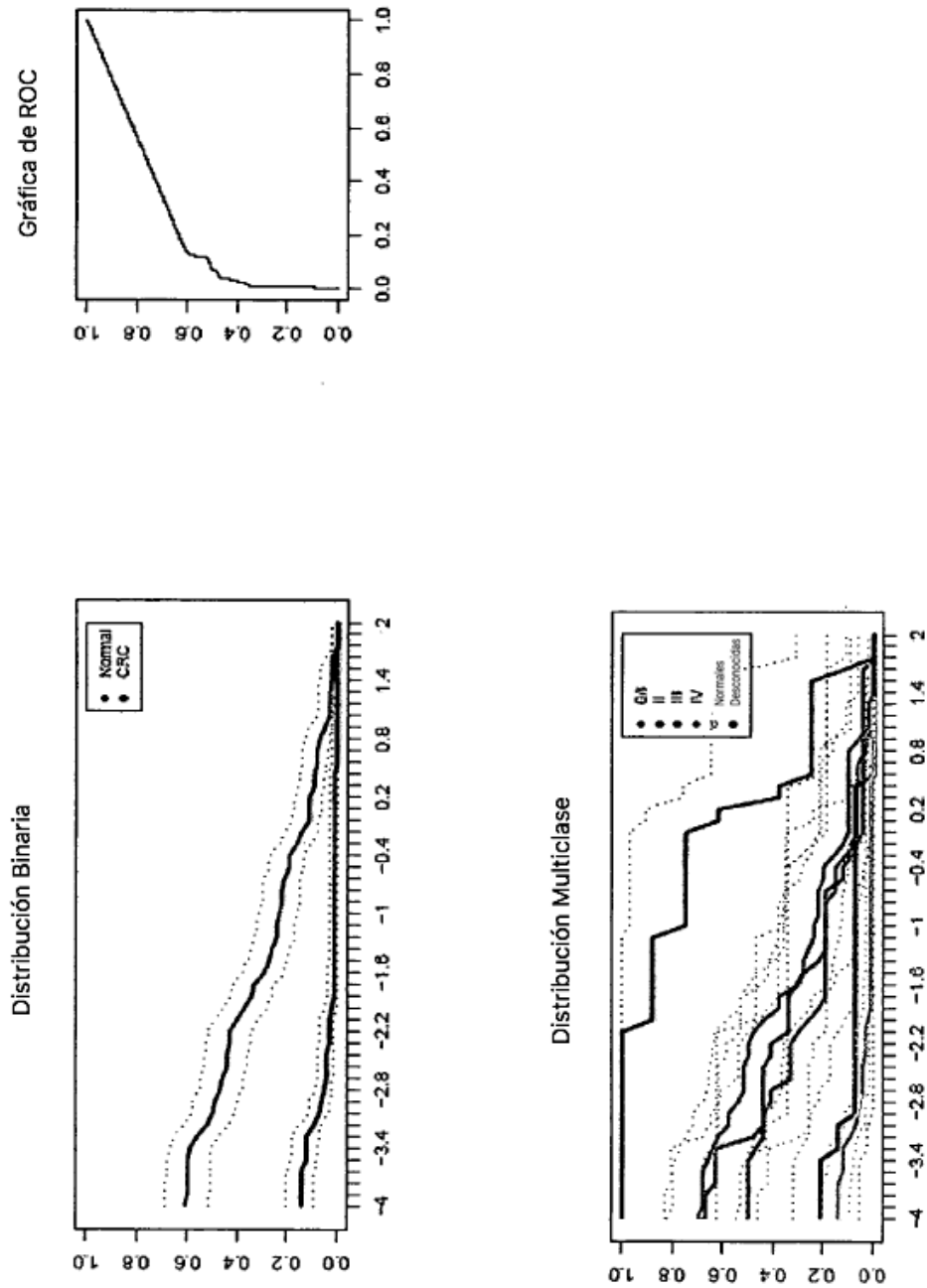


FIGURA 8

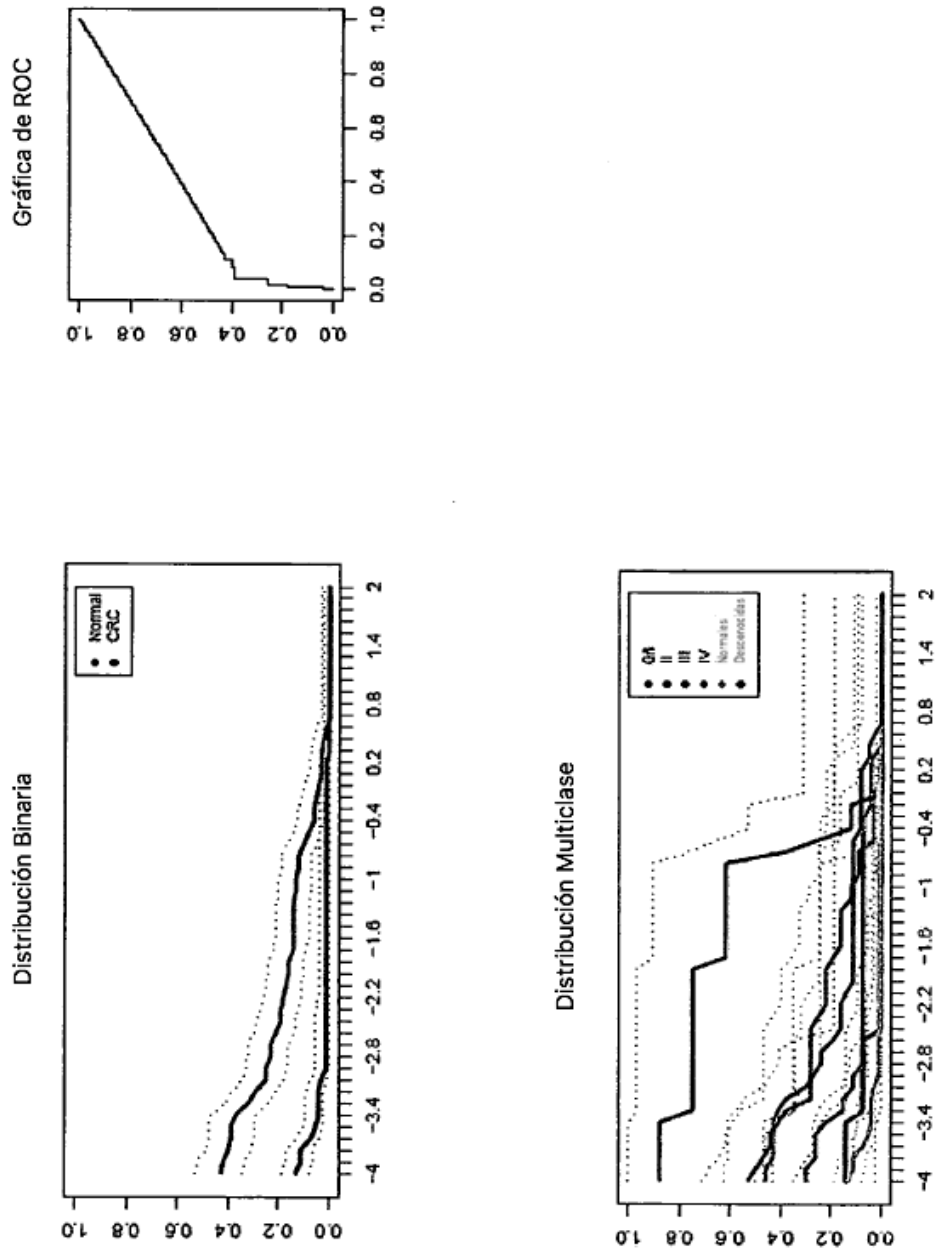


FIGURA 9

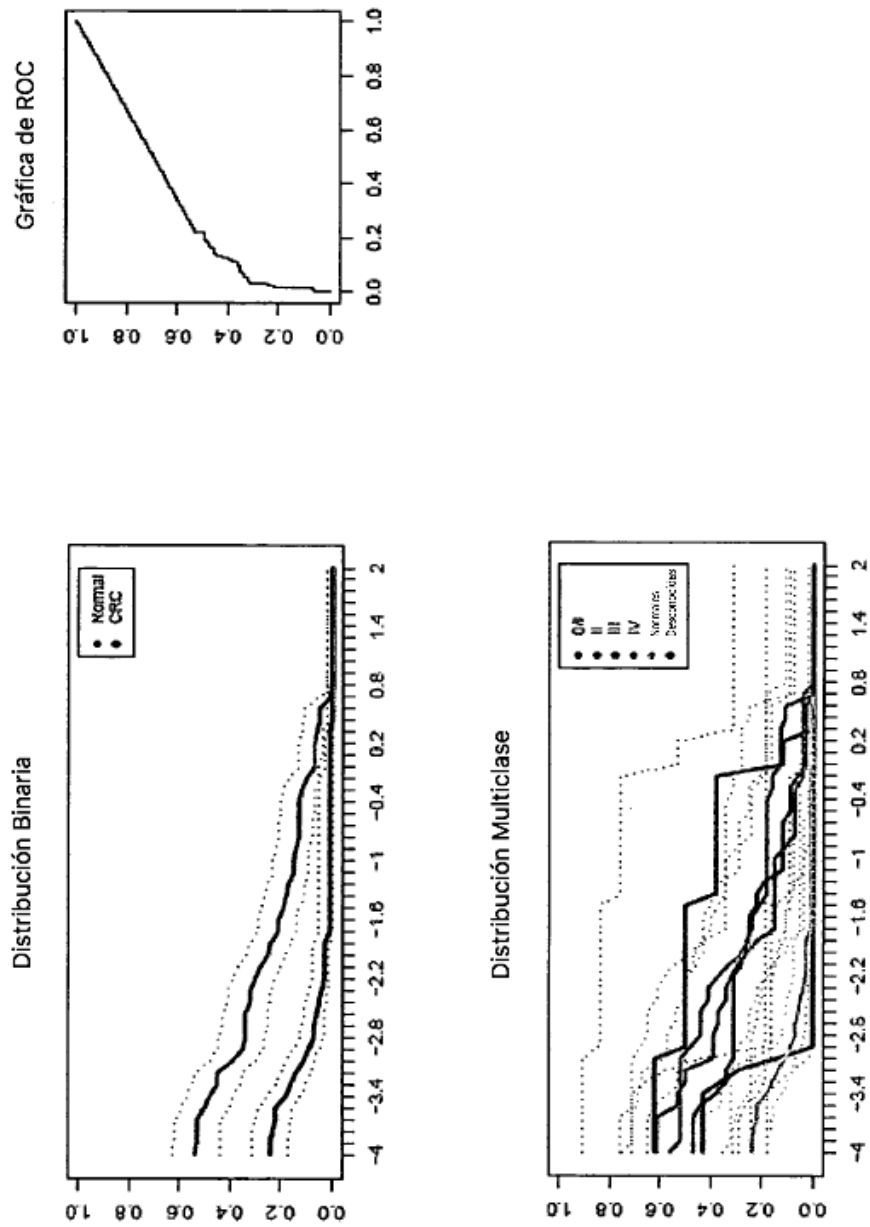


FIGURA 10

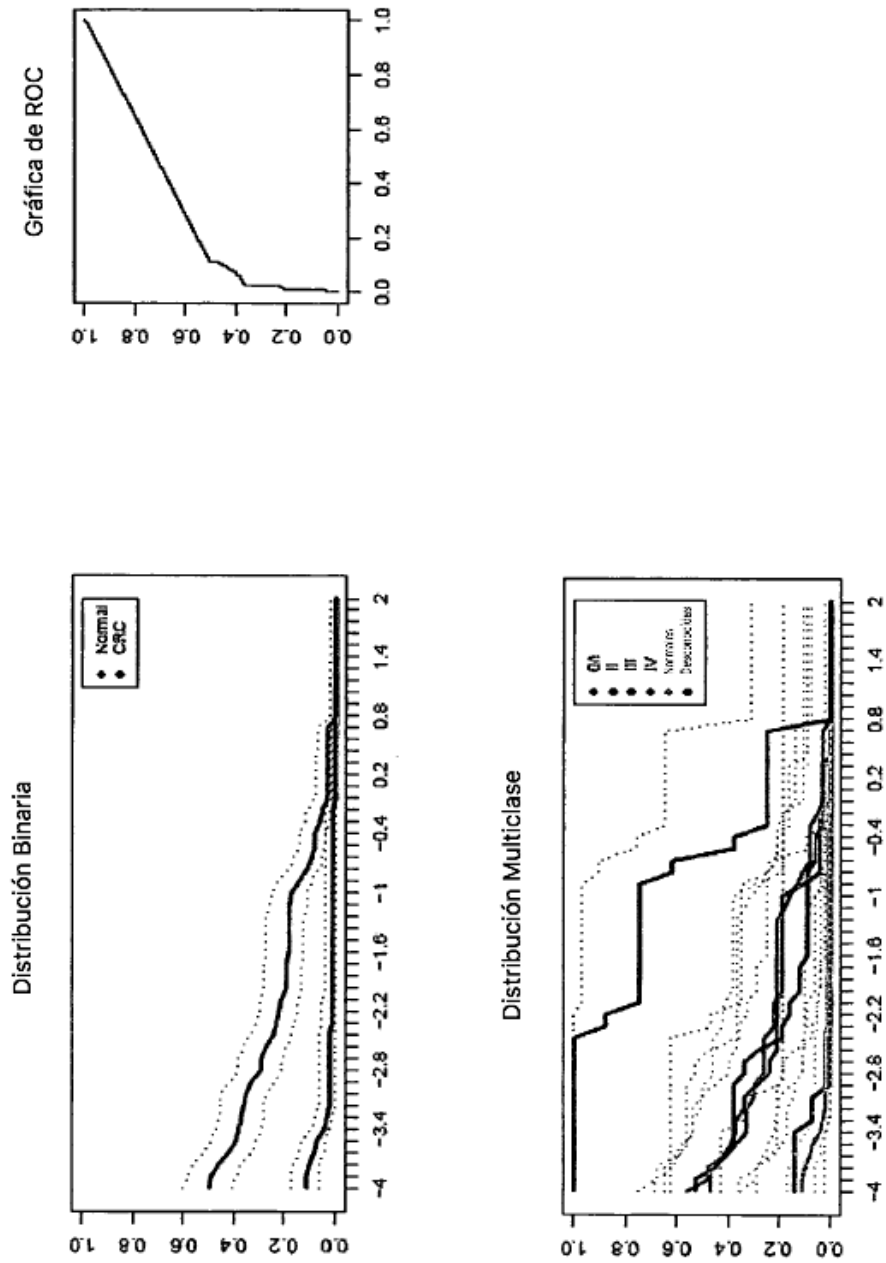


FIGURA 11

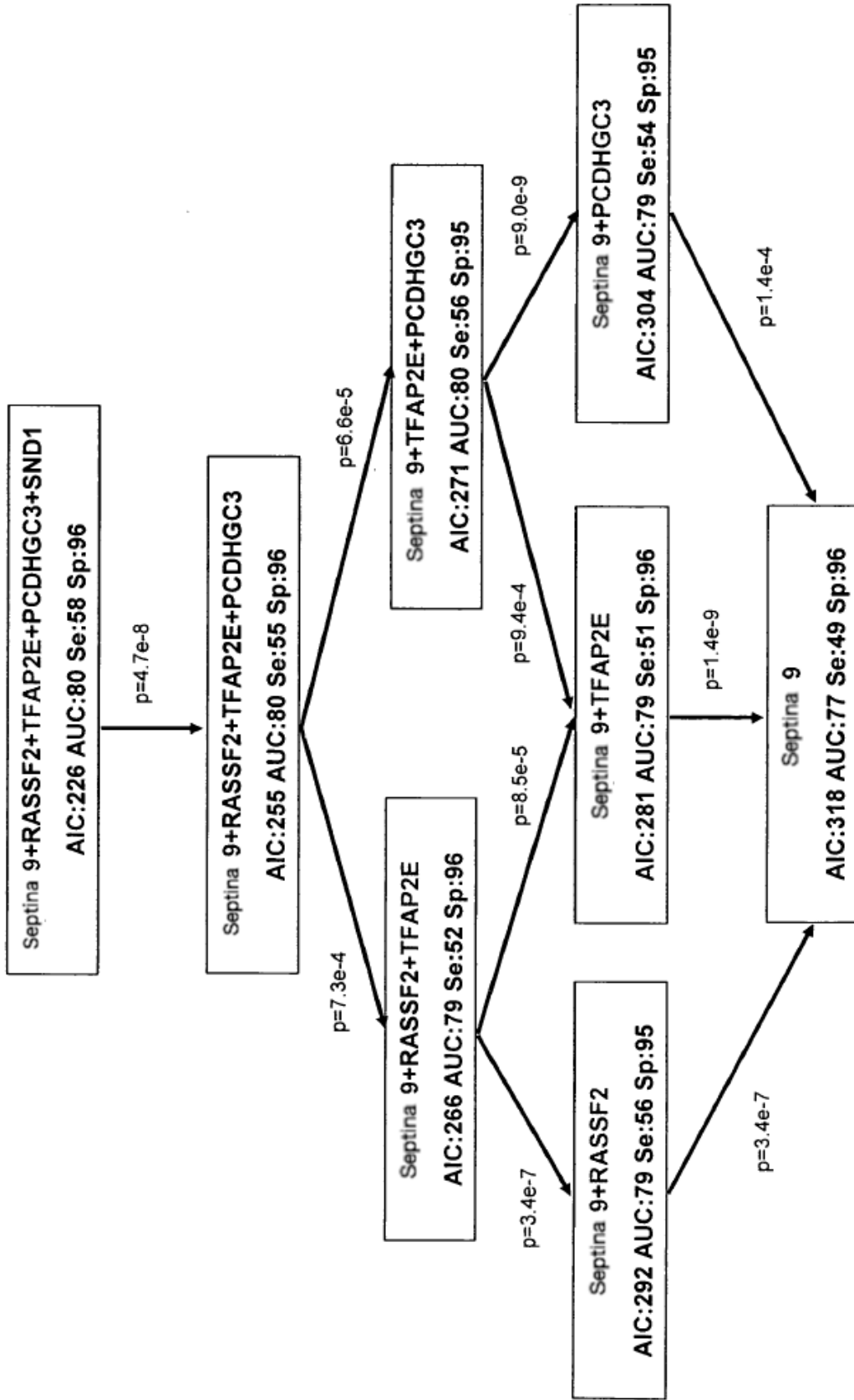


FIGURA 12

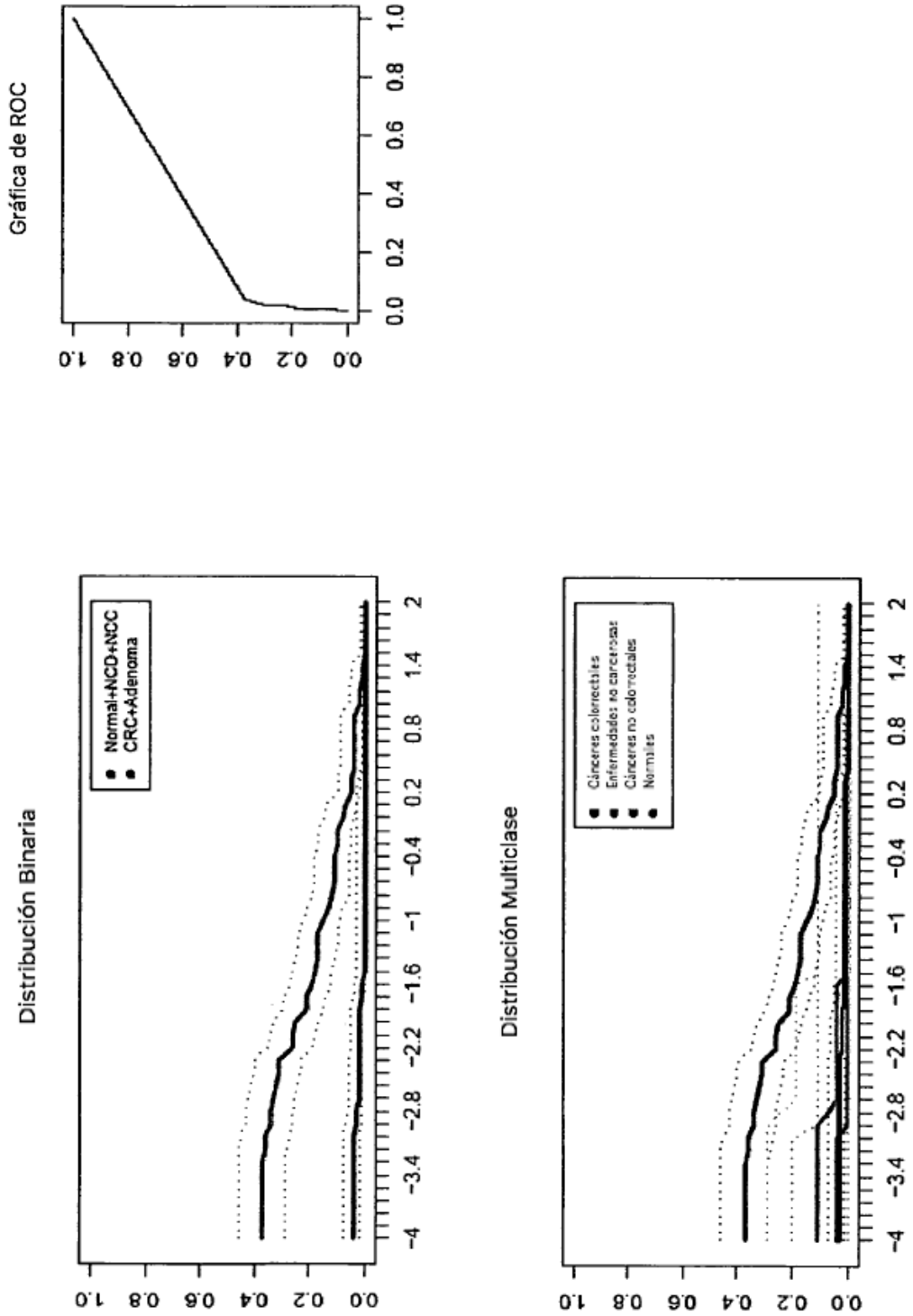


FIGURA 13

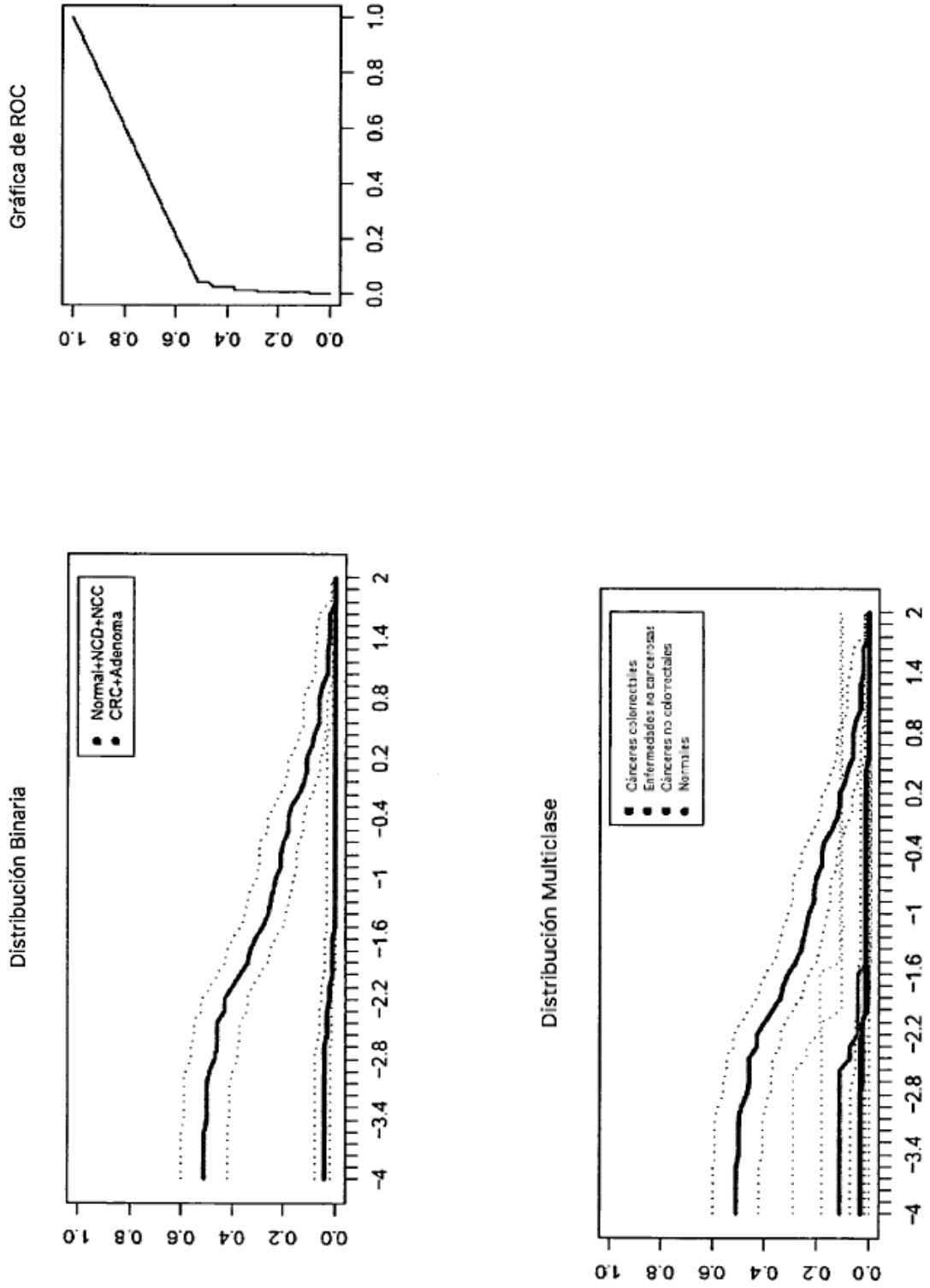
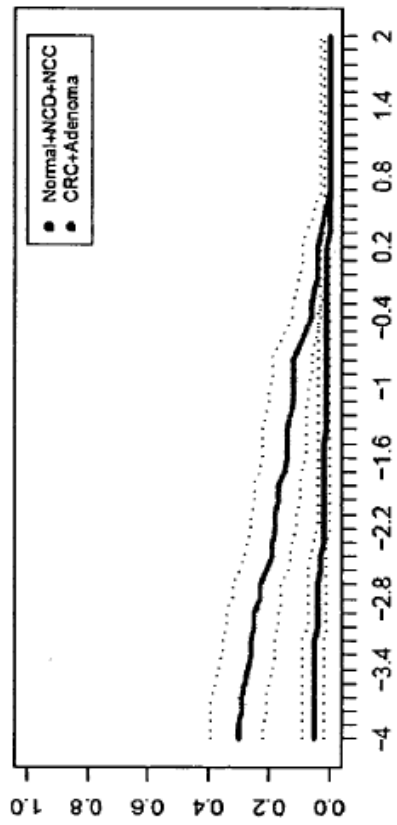
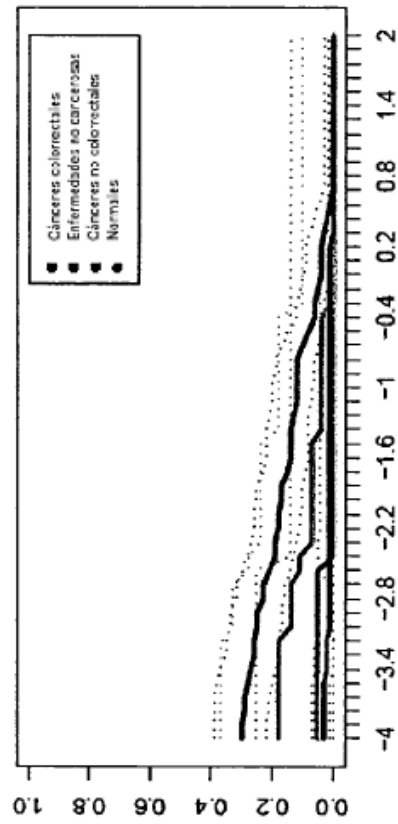


FIGURA 14

Distribución Binaria



Distribución Multiclasa



Gráfica de ROC

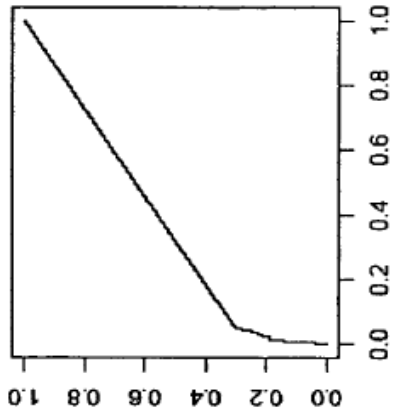


FIGURA 15

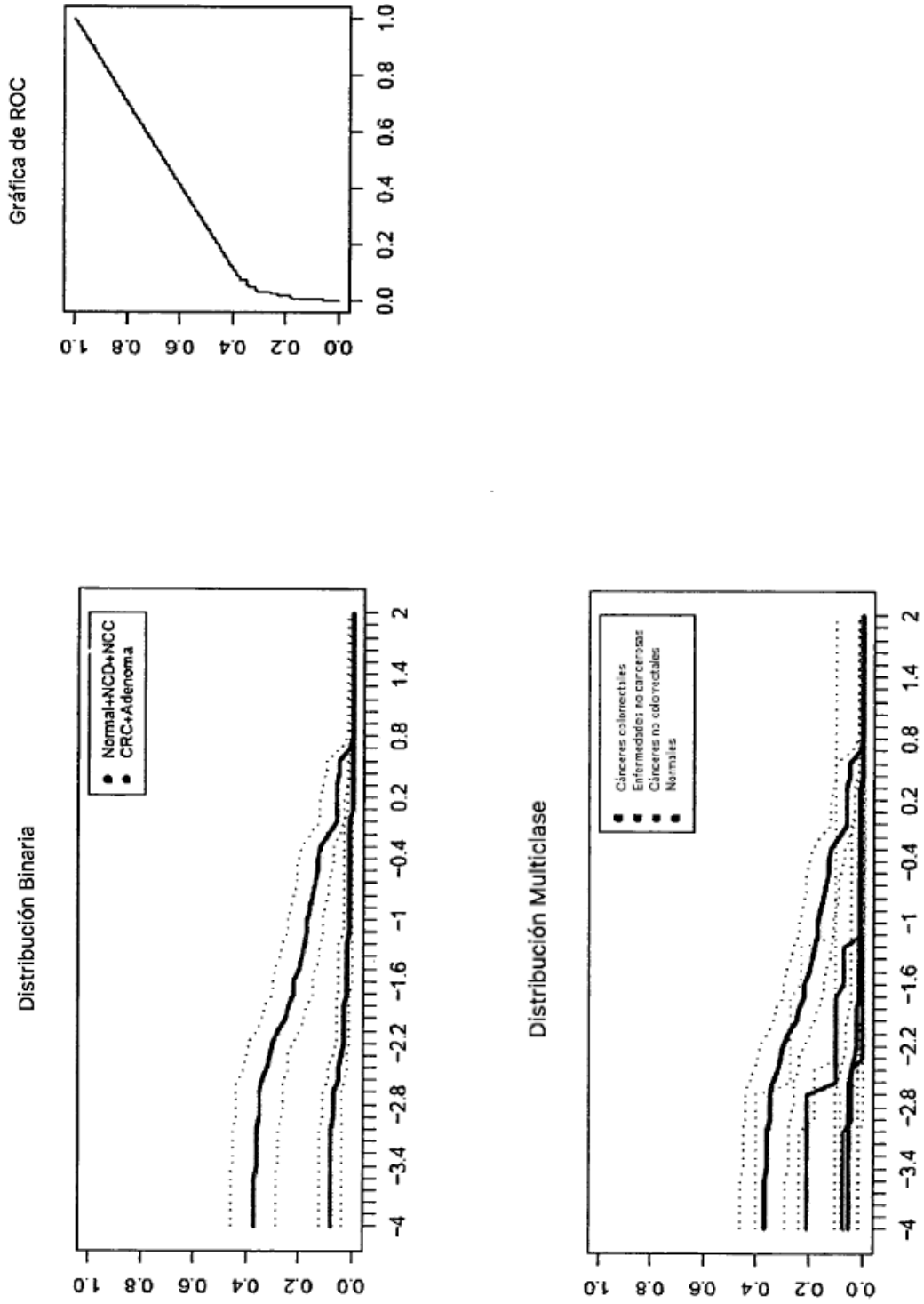


FIGURA 16

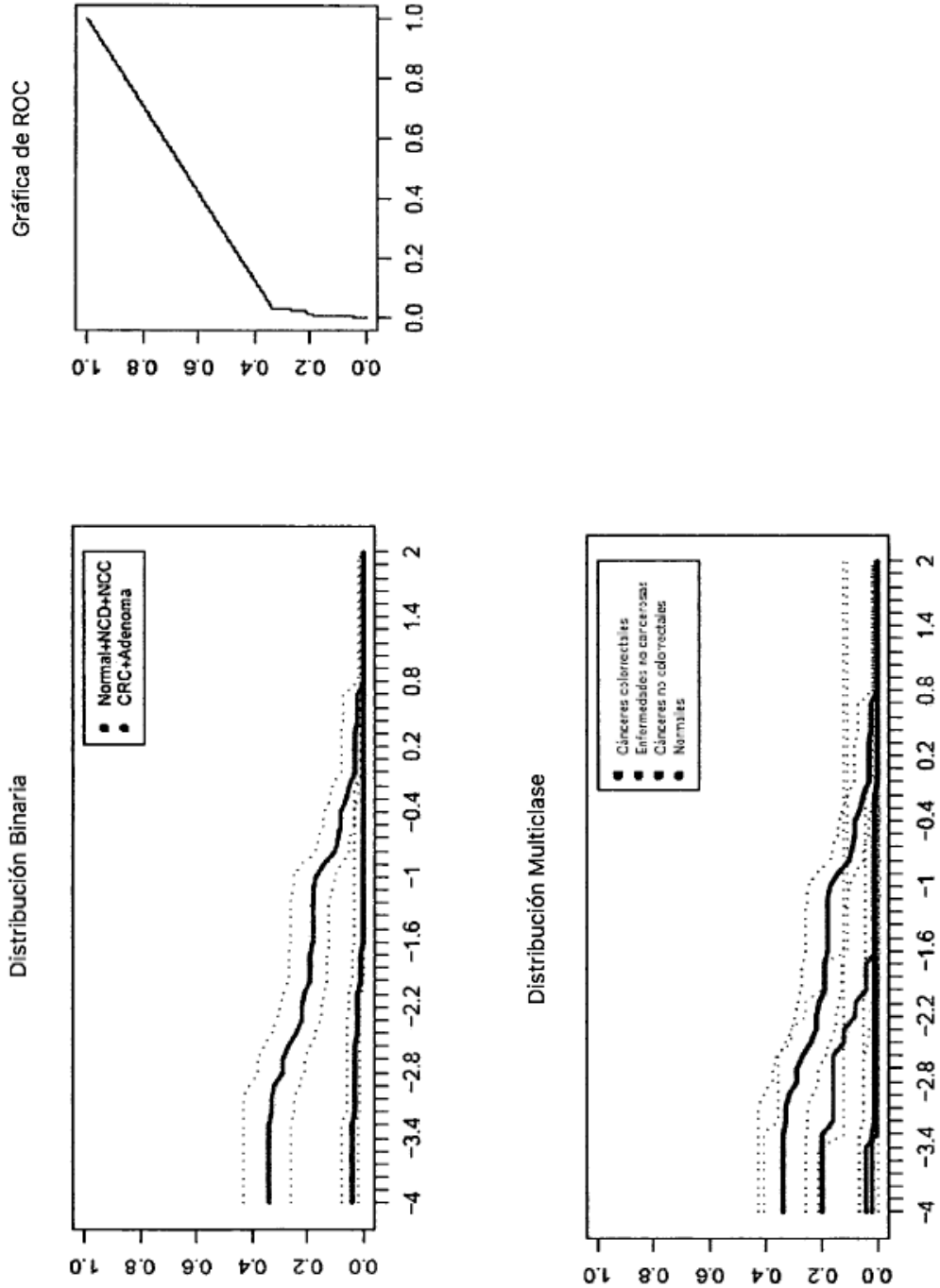


FIGURA 17

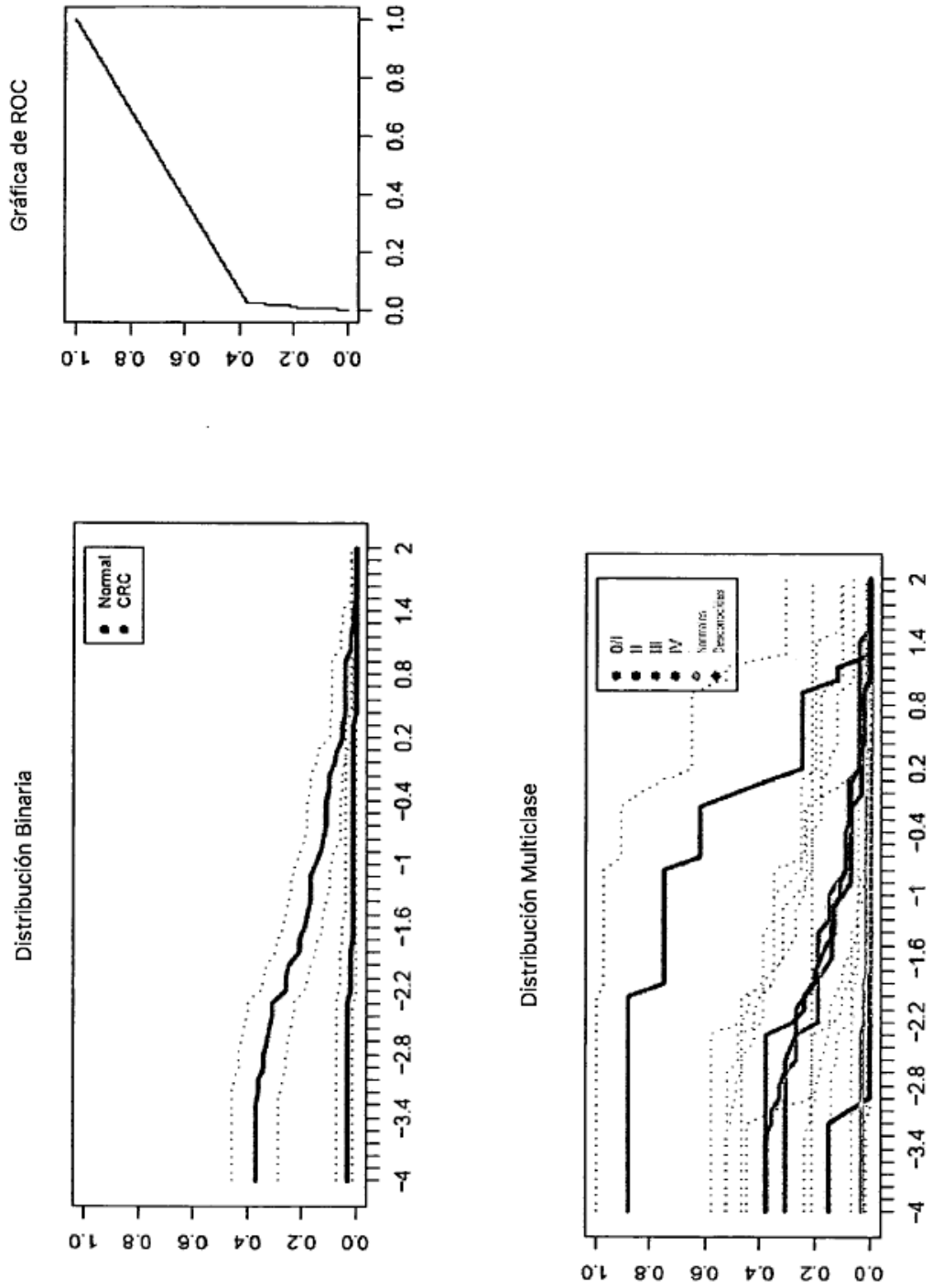


FIGURA 18

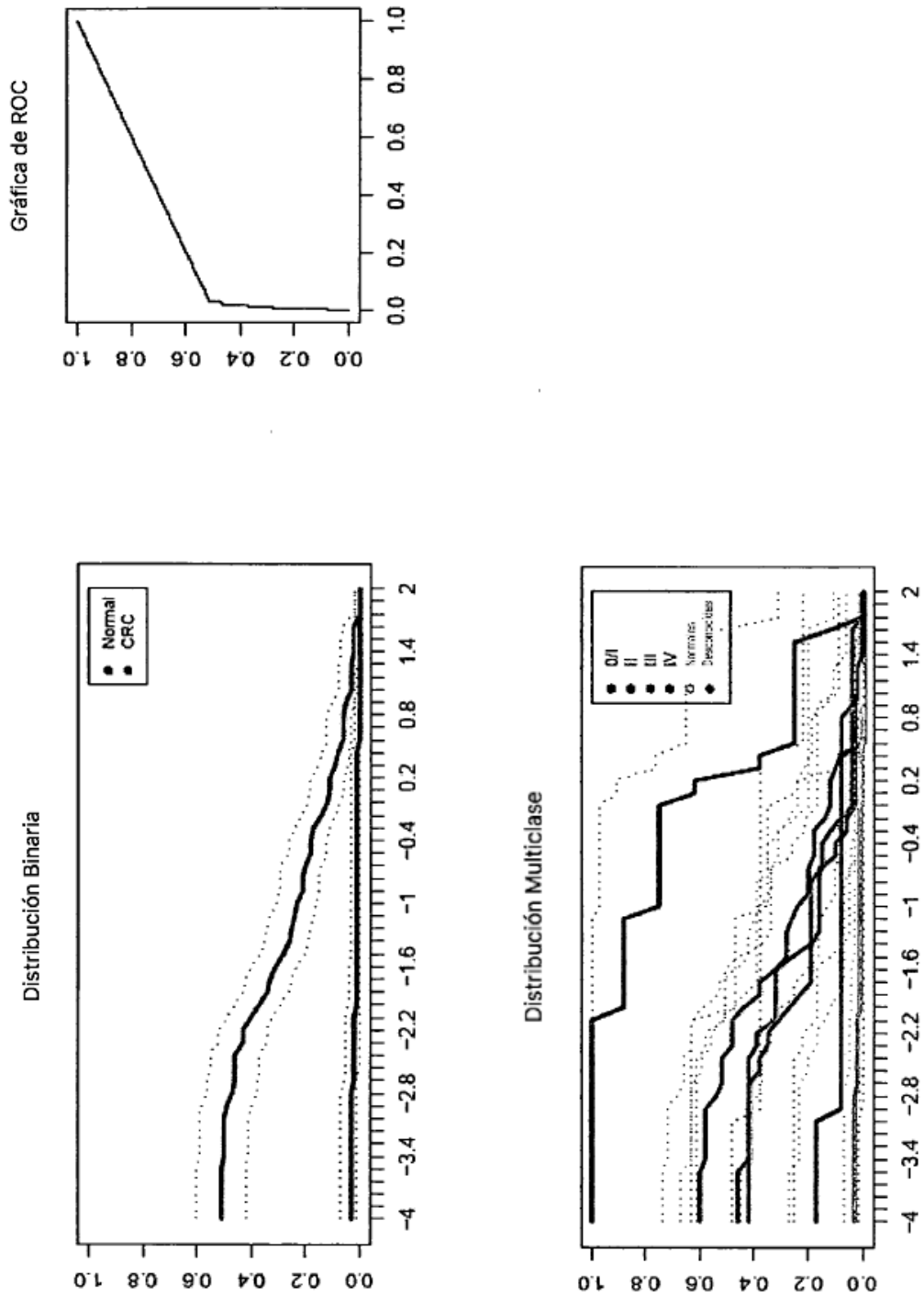


FIGURA 19

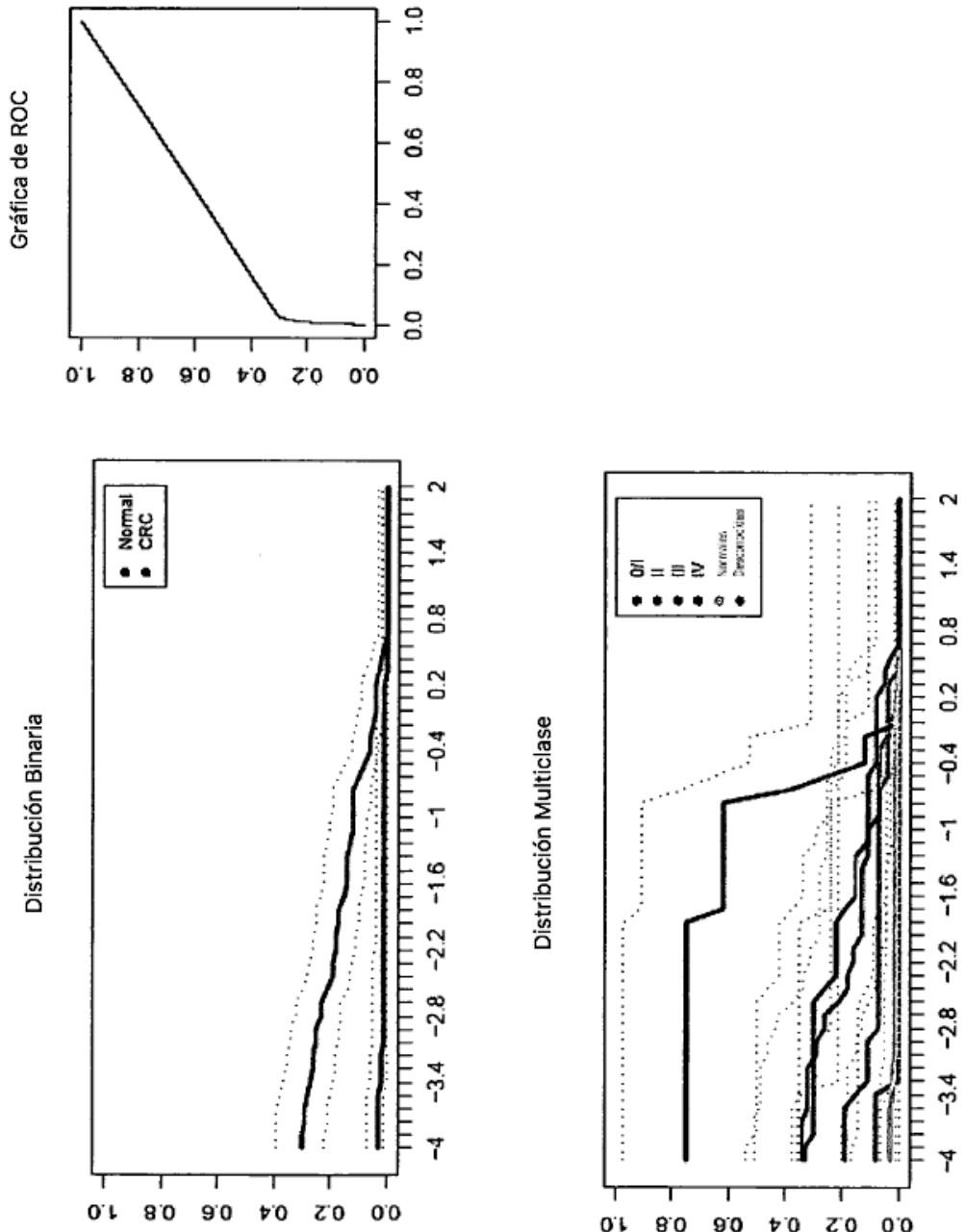


FIGURA 20

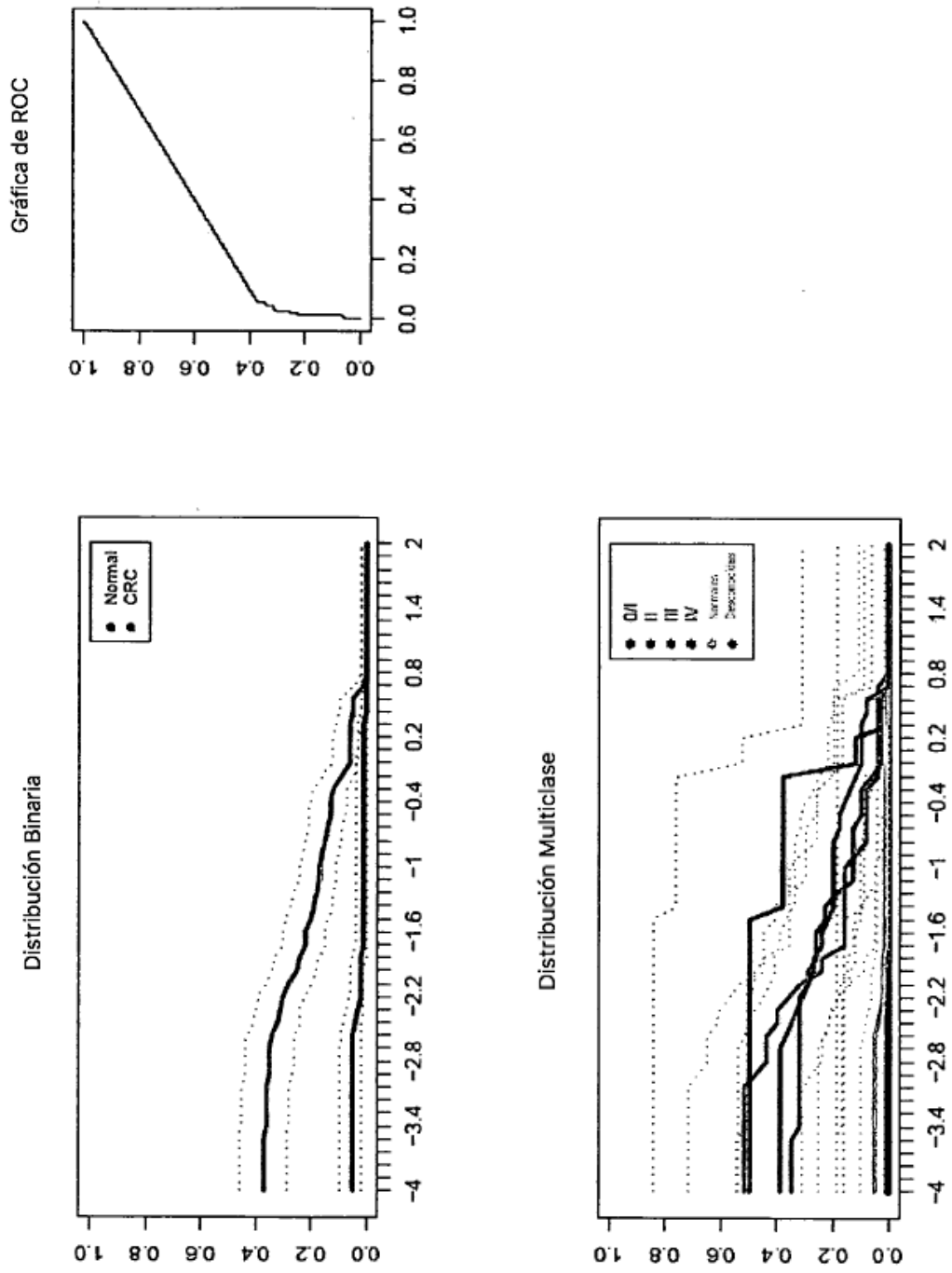


FIGURA 21

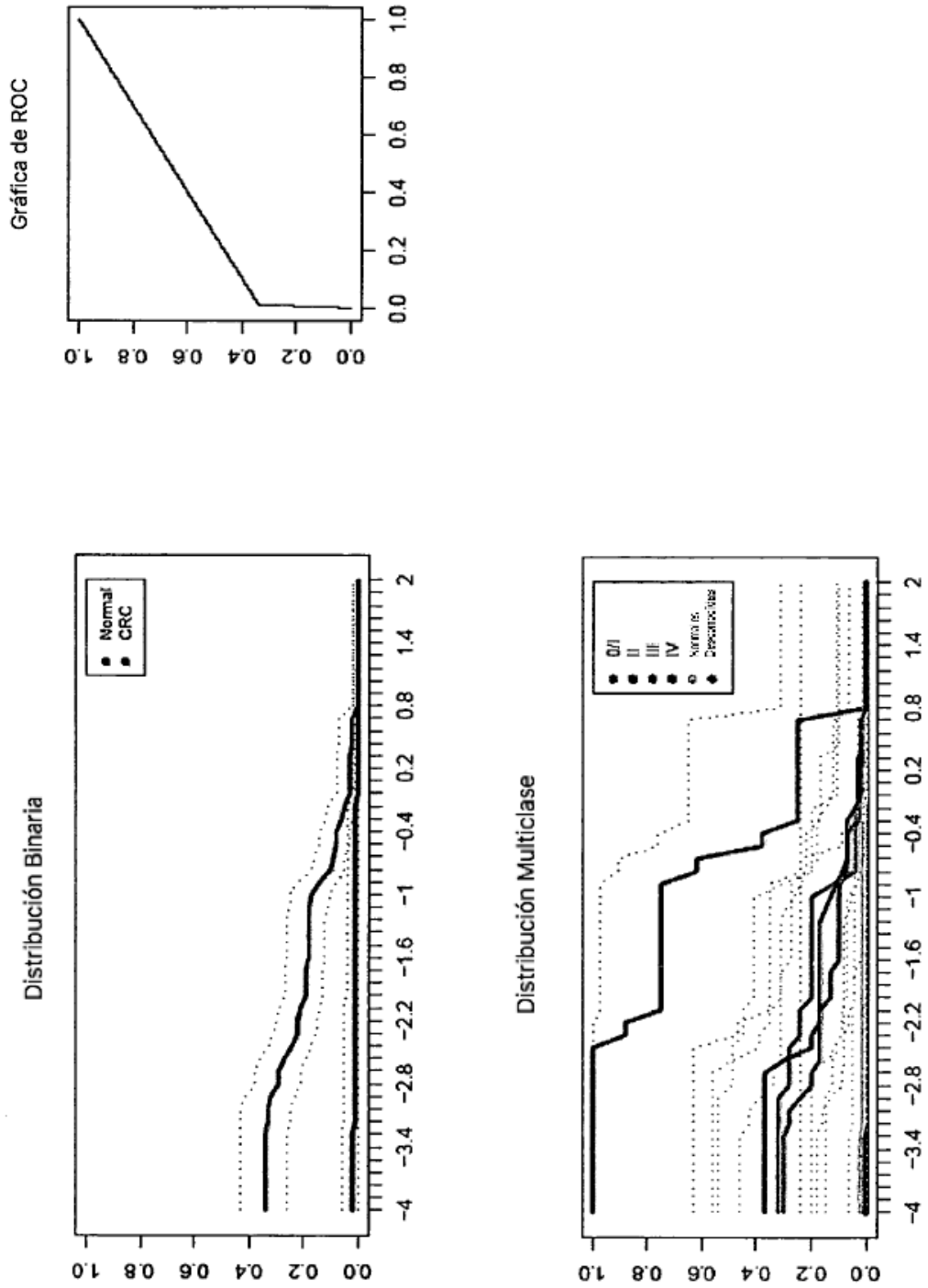


FIGURA 22

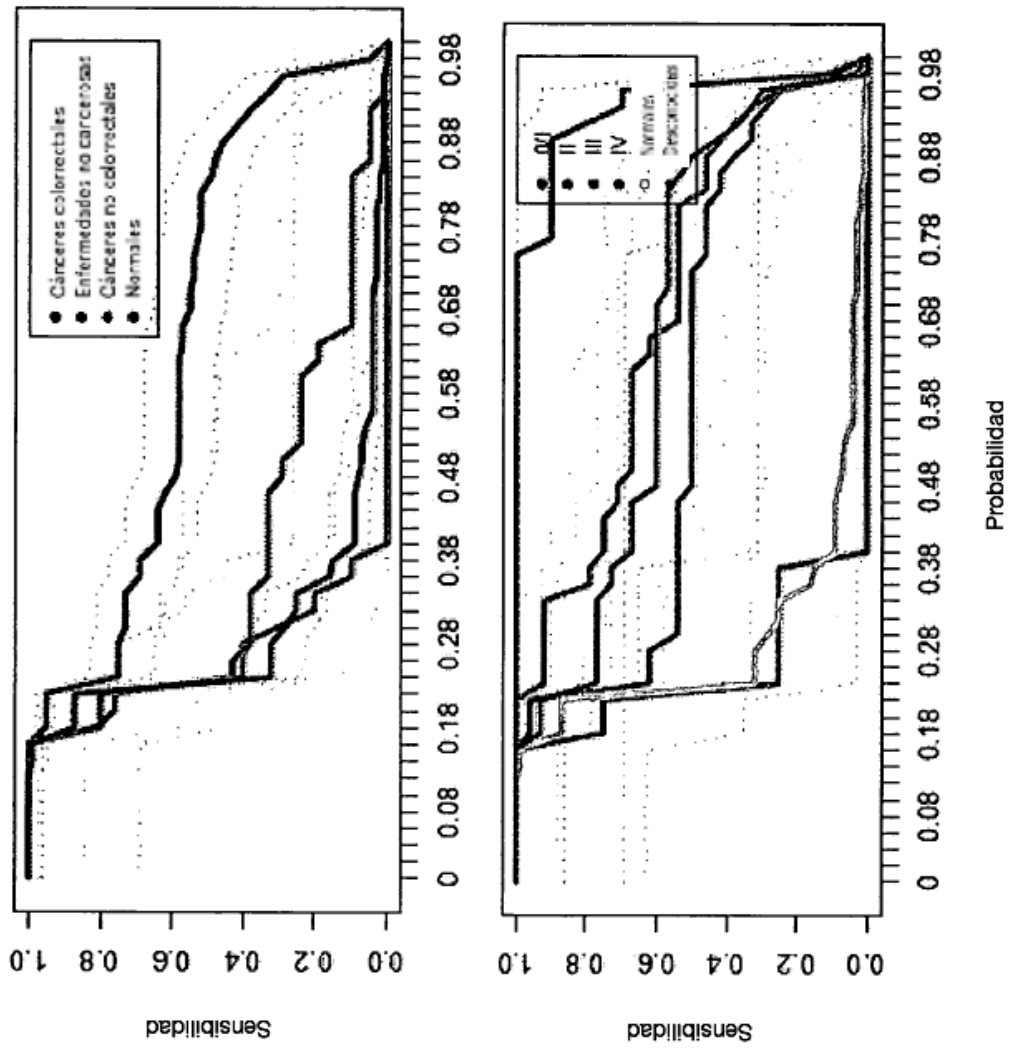


FIGURA 23

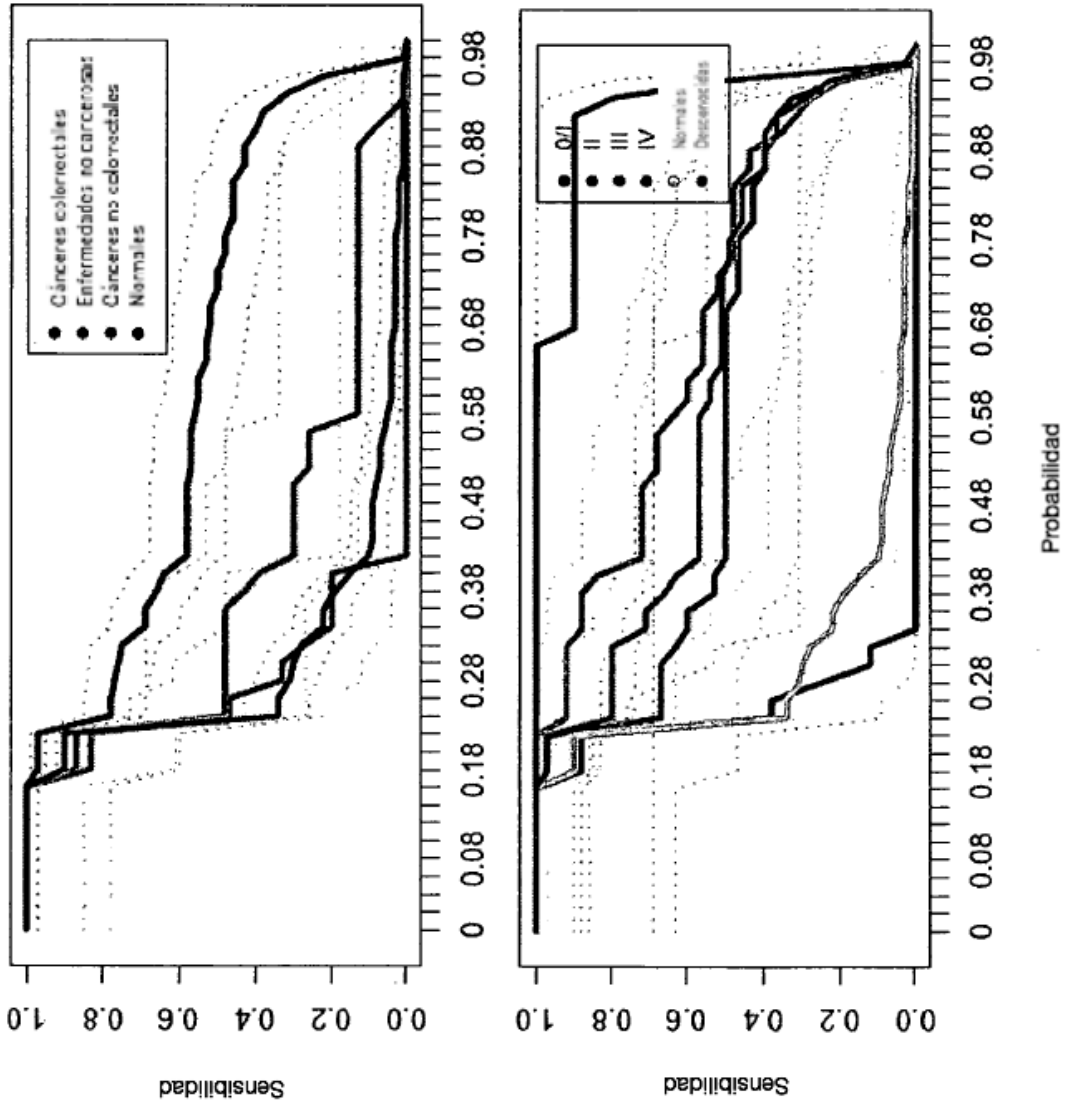


FIGURA 24

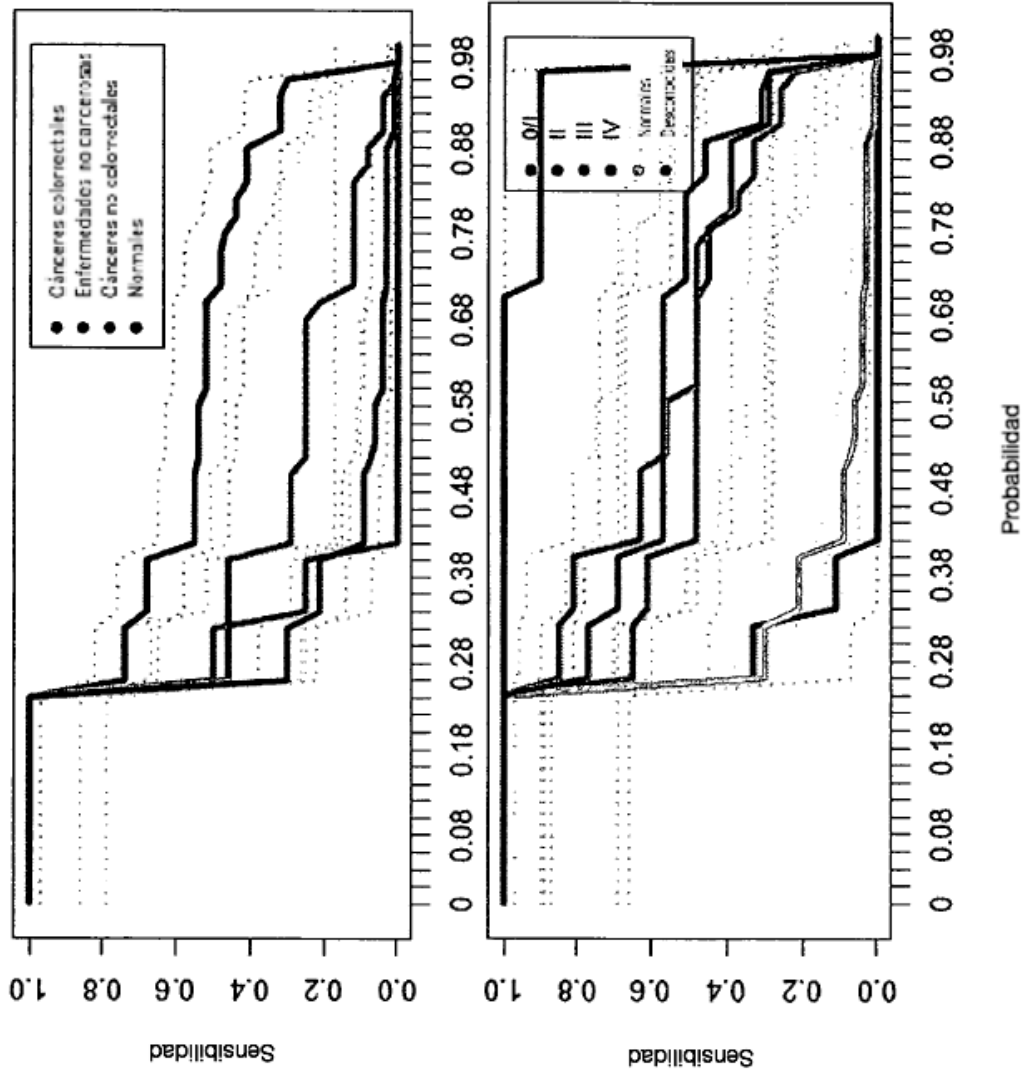


FIGURA 25

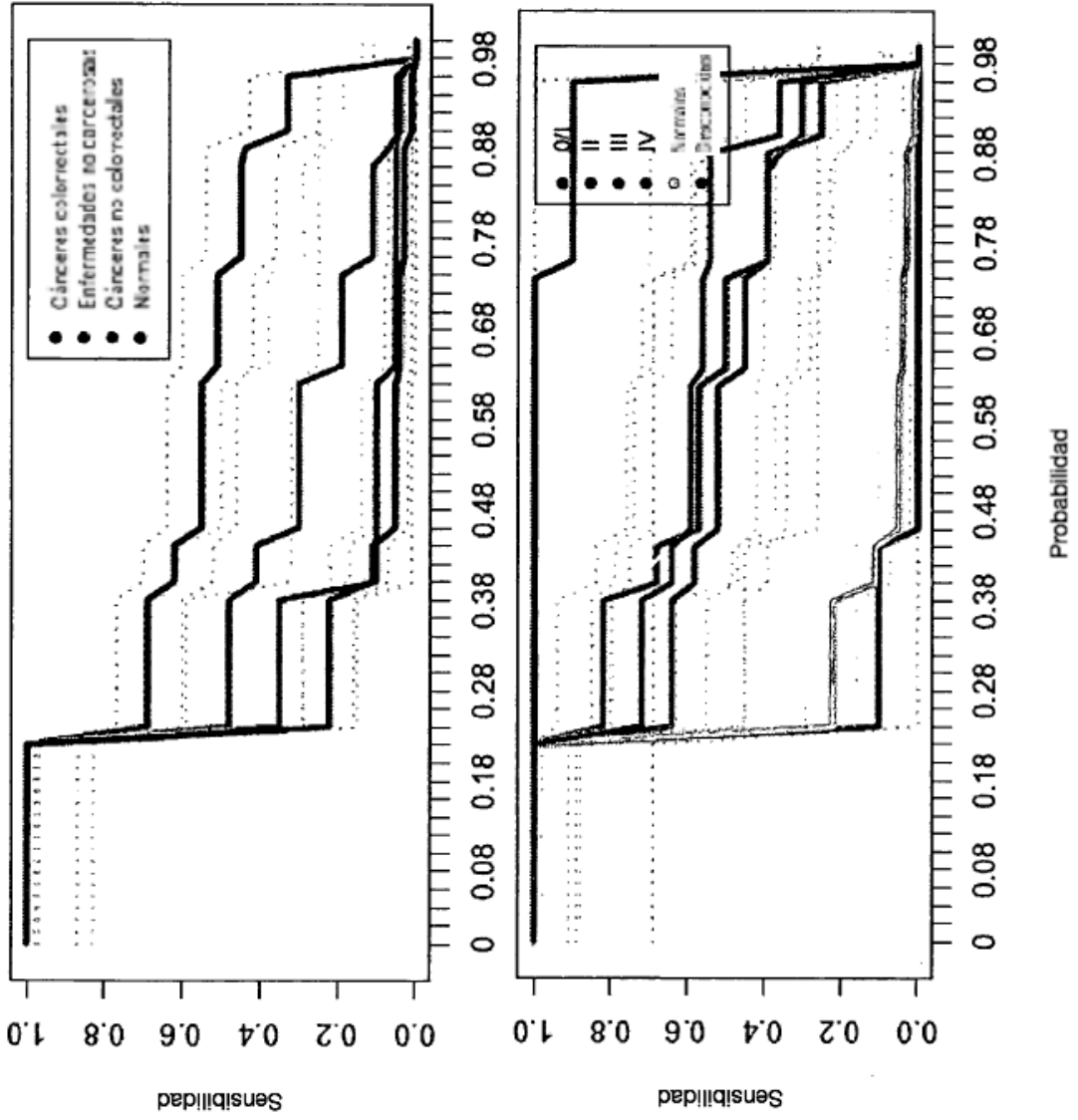


FIGURA 26

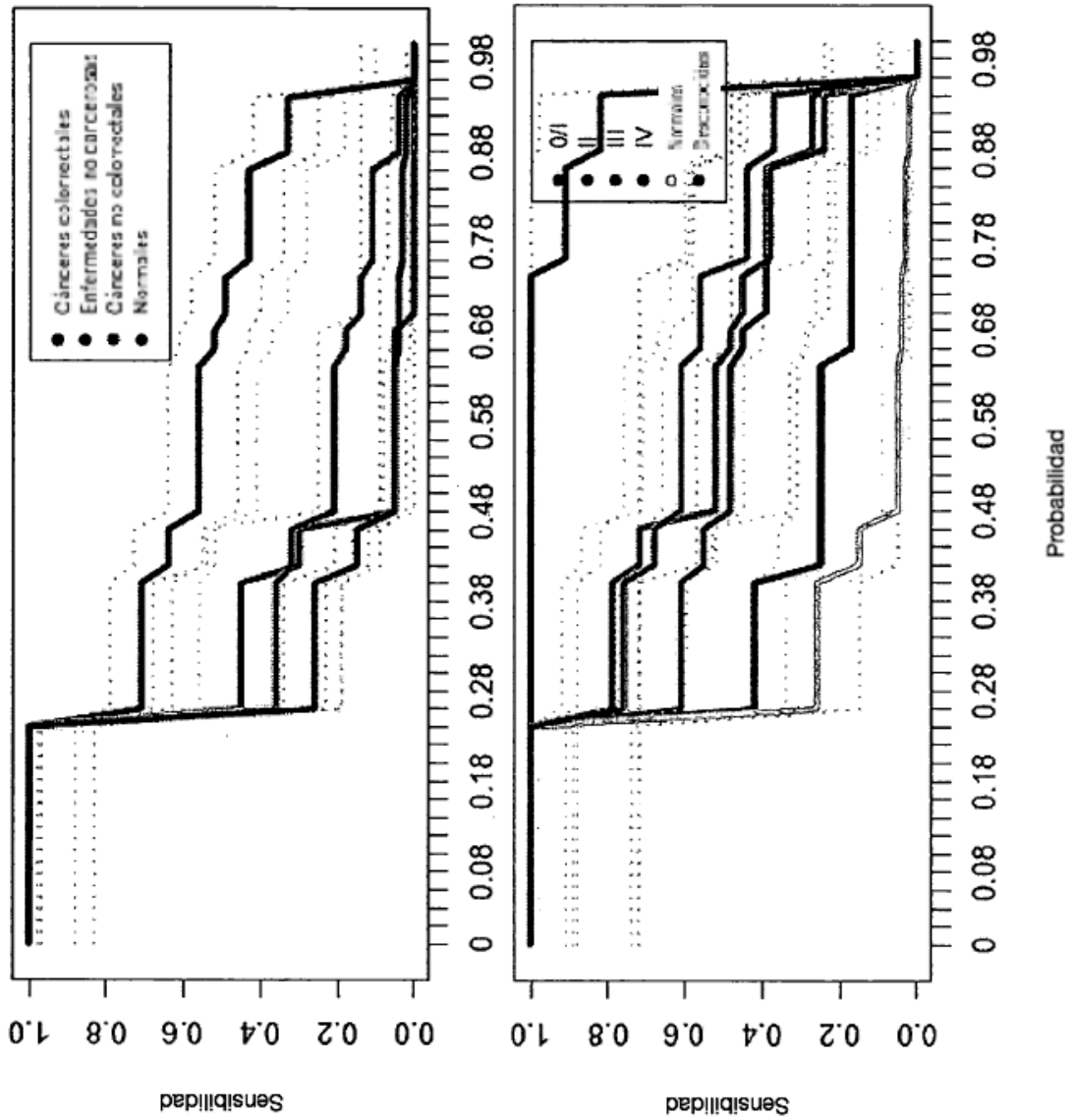


FIGURA 27

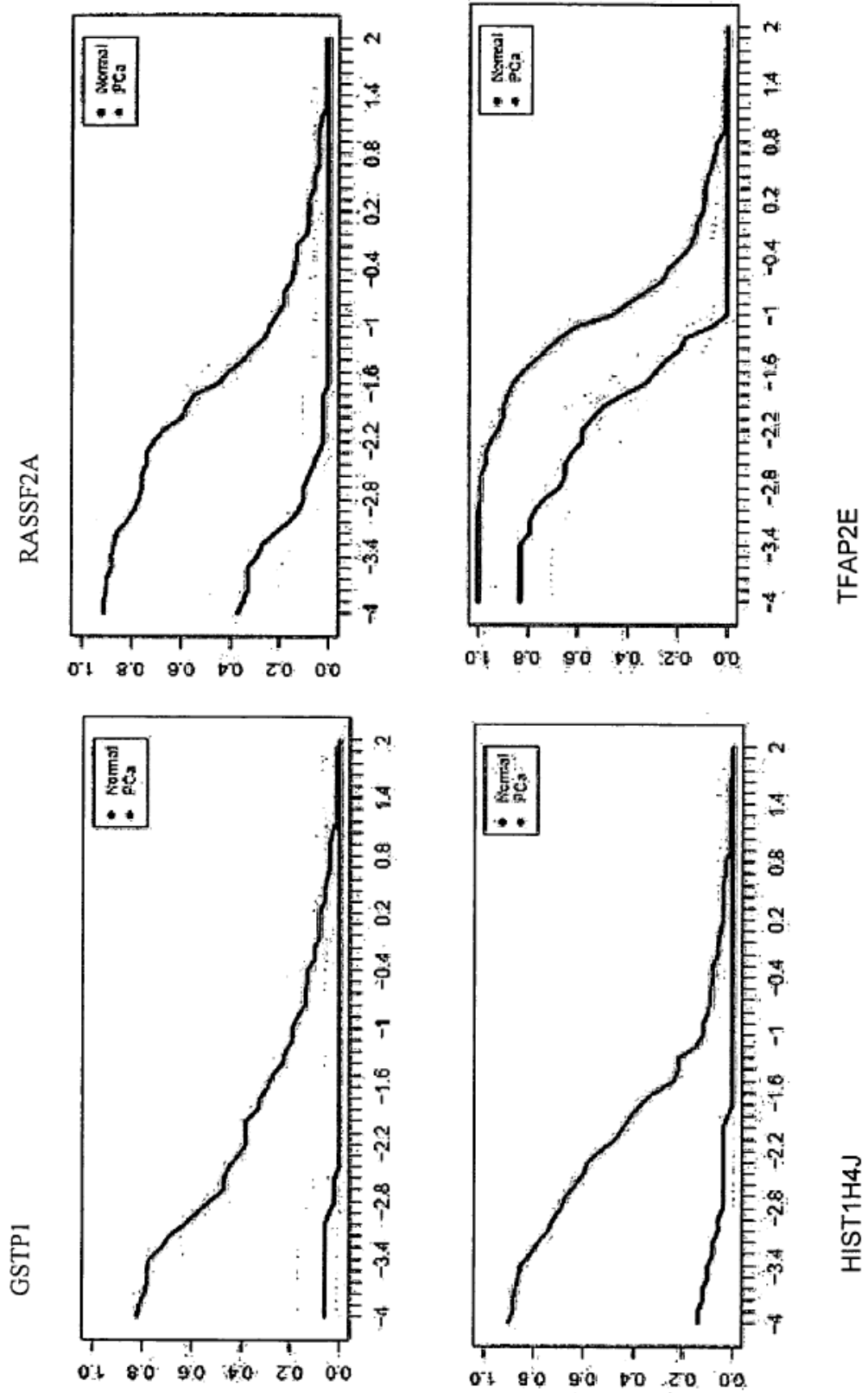


FIGURA 28

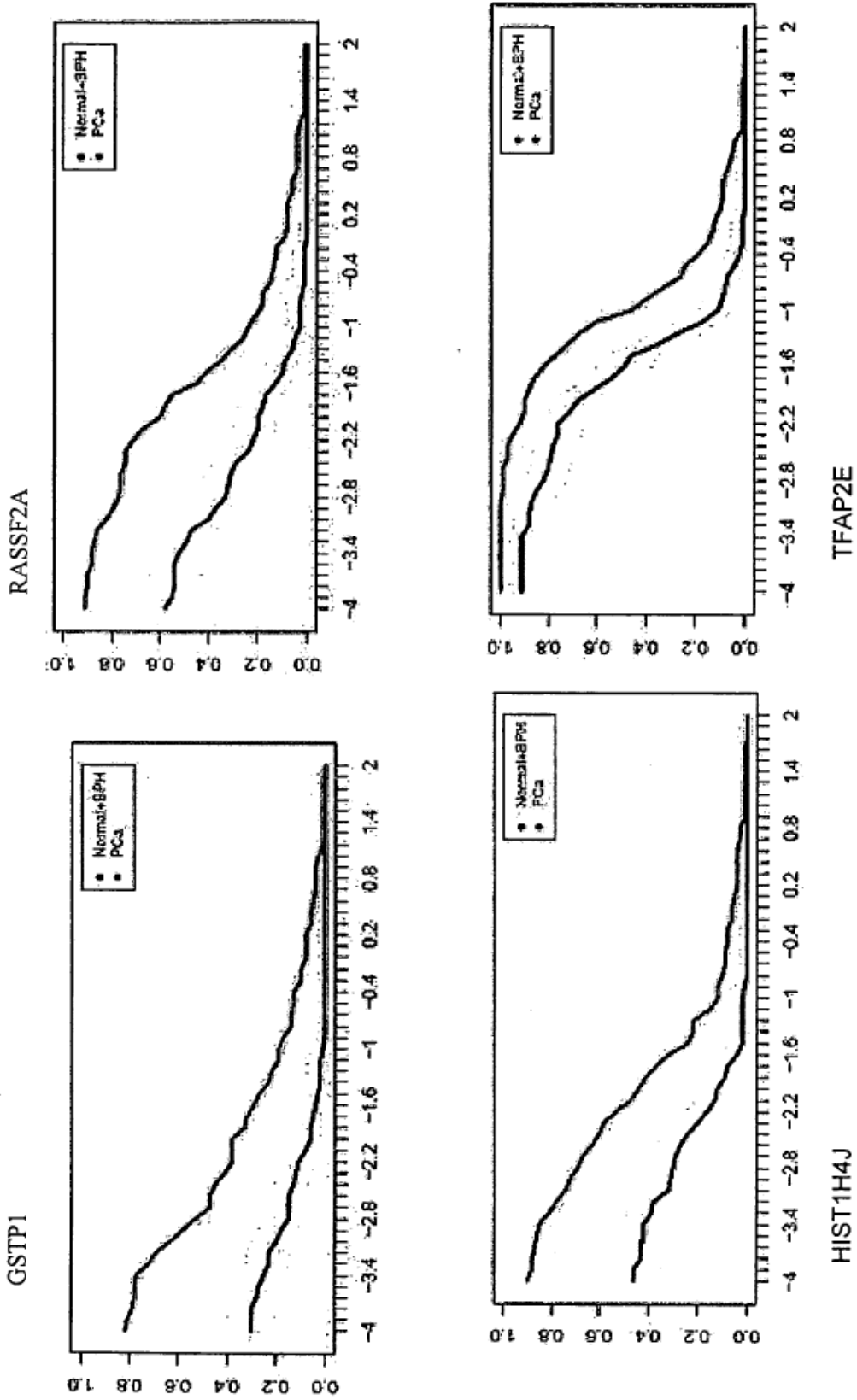
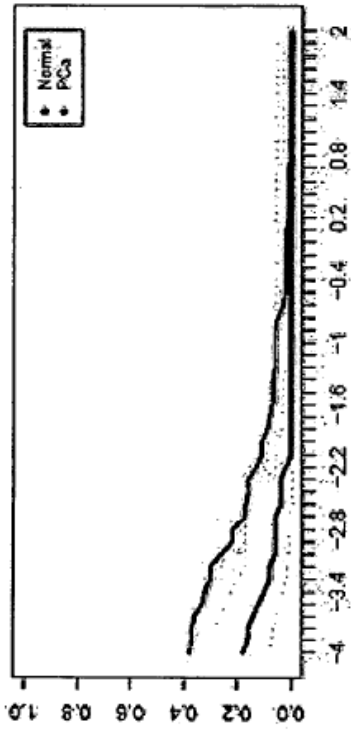
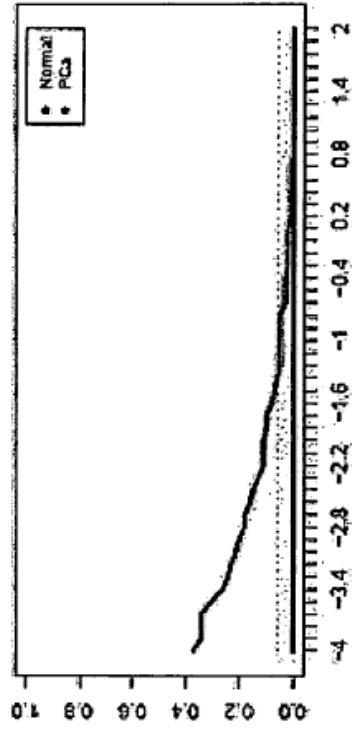


FIGURA 29

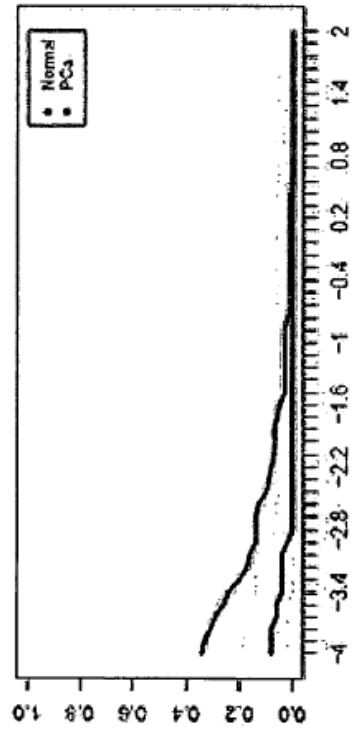
GSTP1



RASSF2A



HIST1H4J



HIST1H4J

TFAP2E

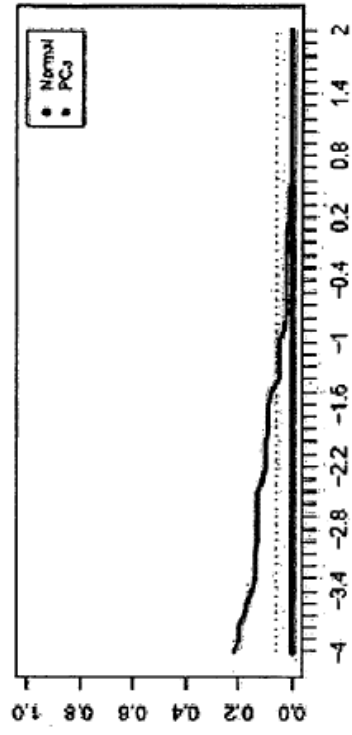


FIGURA 30

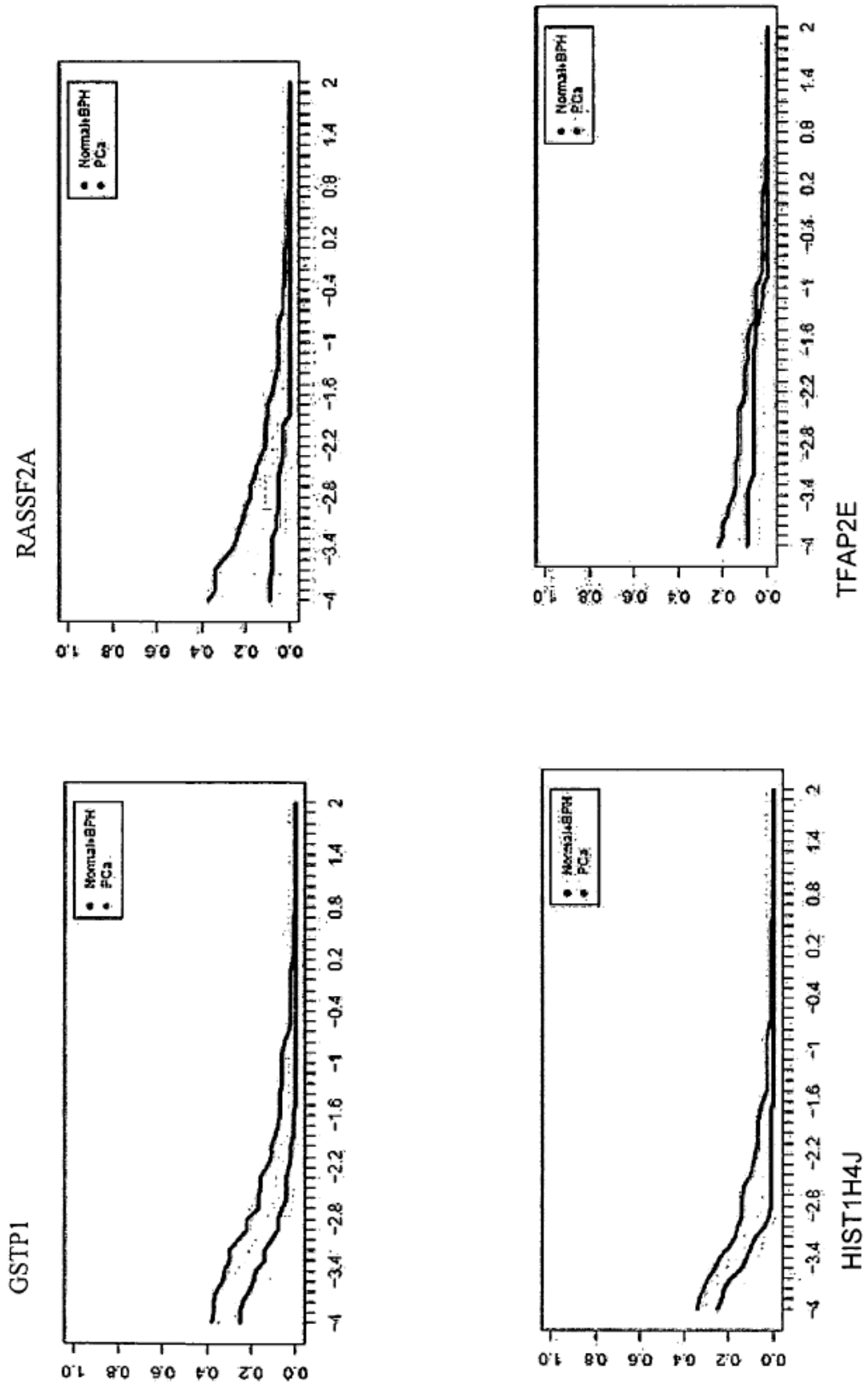


FIGURA 31

