

(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



(11) Número de publicación: **2 994 426**

(51) Int. Cl.:

C12Q 1/68 (2008.01)

C12Q 1/6869 (2008.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **20.09.2013 E 22171427 (2)**

(97) Fecha y número de publicación de la concesión europea: **31.07.2024 EP 4056712**

(54) Título: **Determinación no invasiva del metiloma tumoral a partir de plasma**

(30) Prioridad:

**20.09.2012 US 201261703512 P
15.03.2013 US 201313842209
03.06.2013 US 201361830571 P**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
23.01.2025

(73) Titular/es:

**THE CHINESE UNIVERSITY OF HONG KONG
(100.00%)
Knowledge Transfer Office, Room 301 Pi Ch'iu
Building, Shatin New Territories
Hong Kong, HK**

(72) Inventor/es:

**CHIU, ROSSA WAI KWUN;
CHAN, KWAN CHEE;
LO, YUK-MING DENNIS;
LUN, MIU FAN;
JIANG, PEIYONG y
CHAN, WAI MAN**

(74) Agente/Representante:

PONS ARIÑO, Ángel

ES 2 994 426 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

Determinación no invasiva del metiloma tumoral a partir de plasma

- 5 La presente divulgación se refiere, en general, a la determinación de un patrón de metilación (metiloma) del ADN, y más particularmente al análisis de una muestra biológica (p. ej., plasma) que incluye una mezcla de ADN de diferentes genomas (p. ej., de feto y madre, o de tumor y células normales) para determinar el patrón de metilación (metiloma) del genoma minoritario. También se describen los usos del metiloma determinado.
- 10 El desarrollo embrionario y fetal es un proceso complejo que implica una serie de acontecimientos genéticos y epigenéticos altamente orquestados. El desarrollo del cáncer también es un proceso complejo que suele implicar múltiples etapas genéticas y epigenéticas. Las anomalías en el control epigenético de los procesos evolutivos están implicadas en la infertilidad, aborto espontáneo, anomalías del crecimiento intrauterino y consecuencias postnatales.
- 15 La metilación del ADN es uno de los mecanismos epigenéticos más estudiados. La metilación del ADN se produce principalmente en el contexto de la adición de un grupo metilo al carbono 5' de los restos de citosina entre los dinucleótidos CpG. La metilación de la citosina añade una capa de control a la transcripción de genes y a la función del ADN. Por ejemplo, la hipermetilación de promotores de genes enriquecidos con dinucleótidos CpG, denominados islas CpG, suele asociarse a la inhibición de la función génica.
- 20 A pesar del importante papel de los mecanismos epigenéticos en la mediación de los procesos evolutivos, los tejidos embrionarios y fetales humanos no son fáciles de obtener para su análisis (los tumores pueden ser igualmente inaccesibles). Los estudios de los cambios dinámicos de estos procesos epigenéticos en la salud y la enfermedad durante el período prenatal en seres humanos son prácticamente imposibles. Los tejidos extraembrionarios, especialmente la placenta, que pueden obtenerse como parte de los procedimientos de diagnóstico prenatal o después del nacimiento, han proporcionado una de las principales vías para tales investigaciones. Sin embargo, estos tejidos requieren procedimientos invasivos.

El perfil de metilación del ADN de la placenta humana ha intrigado a los investigadores durante décadas. La placenta humana presenta una pléthora de características fisiológicas peculiares relacionadas con la metilación del ADN. A nivel mundial, los tejidos placentarios están hipometilados en comparación con la mayoría de los tejidos somáticos. A nivel genético, el estado de metilación de determinados locus genómicos es un distintivo específico de los tejidos placentarios. Tanto los perfiles de metilación globales como los específicos de cada locus muestran cambios dependientes de la edad gestacional. Los genes sellados, en concreto, los genes cuya expresión depende del origen parental de los alelos desempeñan funciones clave en la placenta. La placenta se ha descrito como pseudomaligna y se ha observado hipermetilación de varios genes inhibidores de tumores.

Los estudios del perfil de metilación del ADN de los tejidos placentarios han aportado información sobre la fisiopatología de enfermedades asociadas a la gestación o relacionadas con el desarrollo, tales como la preeclampsia y el retraso del crecimiento intrauterino. Los trastornos producidos en el sellado genómico están asociados a trastornos del desarrollo, como el síndrome de Prader-Willi y el síndrome de Angelman. Se han observado perfiles alterados de sellado genómico y metilación global del ADN en tejidos placentarios y fetales en gestaciones resultantes de técnicas de reproducción asistida (H. Hiura *et al.* 2012 *Hum Reprod.* 27: 2541-2548). Una serie de factores ambientales como el tabaquismo materno (K. E. Haworth *et al.* 2013 *Epigenomics*; 5: 37-49), factores dietéticos maternos (X. Jiang *et al.* 2012 *FASEB J.*; 26: 3563-3574) y el estado metabólico materno, como la diabetes (N. Hajj *et al.*, *Diabetes*, doi: 10.2337/db 12-0289) se han asociado a aberraciones epigenéticas de la descendencia.

El documento WO 2011/038507 A1 se refiere a métodos y sistemas para la elaboración de perfiles de metilación del ADN a gran escala de ADN circulante extracelular en humores corporales. Pueden analizarse varios locus (regiones) en paralelo, y la señal del sujeto y de los individuos sanos puede compararse para identificar una región que esté metilada de forma diferencial.

A pesar de décadas de esfuerzos, no se disponía de medios prácticos para estudiar el metiloma fetal o tumoral y controlar los cambios dinámicos producidos a lo largo de la gestación o durante los procesos patológicos, tales como los tumores malignos. Por lo tanto, es deseable proporcionar métodos para analizar todo o porciones de un metiloma fetal y un metiloma tumoral de forma no invasiva.

De acuerdo con la presente invención, se proporciona un método de análisis de una muestra biológica de un organismo como se define en la reivindicación 1.

60 En el contexto del cáncer, la medición de los cambios metilómicos en el plasma puede permitir detectar el cáncer (p. ej., con fines de cribado), para el seguimiento (p. ej., para detectar la respuesta tras un tratamiento contra el cáncer; y para detectar la recaída del cáncer) y para el pronóstico (p. ej., para medir la carga de células cancerosas en el organismo o con fines de estadificación o para evaluar la probabilidad de muerte por enfermedad o progresión de la enfermedad o procesos metastásicos).

65 Se puede obtener una mejor comprensión de la naturaleza y ventajas de las realizaciones de la presente invención

con referencia a la siguiente descripción detallada y los dibujos adjuntos, en las que:

La FIG. 1A muestra una tabla 100 de resultados de secuenciación de sangre materna, placenta y plasma materno de acuerdo con realizaciones de la presente invención.

5 La FIG. 1 muestra la densidad de metilación en ventanas de 1 Mb de muestras secuenciadas de acuerdo con realizaciones de la presente invención.

10 Las FIGS. 2A-2C muestran gráficas de los valores beta frente a los índices de metilación: (A) Células sanguíneas maternas, (B) Muestra de vellosidades coriónicas, (C) Tejido placentario a término.

Las FIGS. 3A y 3B muestran gráficos de barras del porcentaje de sitios CpG metilados en plasma y glóbulos sanguíneos recogidos de un varón adulto y una mujer adulta no gestante: (A) Autosomas, (B) Cromosoma X.

15 Las FIGS. 4A y 4B muestran gráficos de las densidades de metilación de los locus correspondientes en el ADN de glóbulos sanguíneos y en el ADN plasmático: (A) Mujer adulta no gestante, (B) Varón adulto.

20 Las FIGS. 5A y 5B muestran gráficos de barras del porcentaje de sitios CpG metilados entre las muestras recogidas de la gestación: (A) Autosomas, (B) Cromosoma X.

25 La FIG. 6 muestra un gráfico de barras del nivel de metilación de diferentes clases de repetición del genoma humano para la sangre materna, placenta y plasma materno.

La FIG. 7A muestra un diagrama Circos 700 para muestras del primer trimestre. La FIG. 7B muestra un diagrama Circos 750 para muestras del tercer trimestre.

Las FIGS 8A-8D muestran gráficas de comparaciones de las densidades de metilación del ADN del tejido genómico frente al ADN del plasma materno para los sitios CpG que rodean los polimorfismos de un solo nucleótido informativos.

30 La FIG. 9 es un diagrama de flujo que ilustra un método 900 para determinar un primer perfil de metilación a partir de una muestra biológica de un organismo de acuerdo con realizaciones de la presente invención.

La FIG. 10 es un diagrama de flujo que ilustra un método 1000 para determinar un primer perfil de metilación a partir de una muestra biológica de un organismo de acuerdo con realizaciones de la presente invención.

35 35 Las FIGS. 11A y 11B muestran gráficos del rendimiento del algoritmo de predicción utilizando datos de plasma materno y concentración fraccional de ADN fetal de acuerdo con realizaciones de la presente invención.

40 La FIG. 12A es una tabla 1200 que muestra detalles de 15 locus genómicos seleccionados para la predicción de metilación de acuerdo con realizaciones de la presente invención. La FIG. 12B es un gráfico 1250 que muestra las categorías deducidas de los 15 locus genómicos seleccionados y sus correspondientes niveles de metilación en la placenta.

45 La FIG. 13 es un diagrama de flujo de un método 1300 para detectar una anomalía cromosómica fetal a partir de una muestra biológica de una mujer que está gestando al menos un feto.

La FIG. 14 es un diagrama de flujo de un método 1400 para identificar marcadores de metilación comparando un perfil de metilación placentario con un perfil de metilación materno de acuerdo con realizaciones de la presente invención.

50 La FIG. 15A es una tabla 1500 que muestra un rendimiento del algoritmo de identificación de DMR utilizando datos del primer trimestre con referencia a 33 marcadores del primer trimestre previamente reportados. La FIG. 15B es una tabla 1550 que muestra un rendimiento del algoritmo de identificación de DMR utilizando datos del tercer trimestre y comparándolos con la muestra de placenta obtenida en el parto.

55 La FIG. 16 es una tabla 1600 que muestra los números de locus predichos a ser hipermetilados o hipometilados basados en el análisis directo de los datos de secuenciación de plasma materno tratado con bisulfito.

60 La FIG. 17A es un gráfico 1700 que muestra la distribución de tamaños del plasma materno, plasma de control de mujeres no gestantes, ADN placentario y de sangre periférica. La FIG. 17B es un gráfico 1750 de la distribución de tamaños y el perfil de metilación del plasma materno, plasma de control de mujer adulta, tejido placentario y sangre de control de mujer adulta.

Las FIGS. 18A y 18B son gráficas de densidades de metilación y tamaño de moléculas de ADN plasmático de acuerdo con realizaciones de la presente invención.

65 65 La FIG. 19A muestra un gráfico 1900 de las densidades de metilación y los tamaños de las lecturas secuenciadas

- para una mujer adulta no gestante. La FIG. 19B es un gráfico 1950 que muestra la distribución del tamaño y el perfil de metilación de las moléculas de ADN específicas del feto y específicas de la madre en el plasma materno.
- 5 La FIG. 20 es un diagrama de flujo de un método 2000 para estimar un nivel de metilación del ADN en una muestra biológica de un organismo de acuerdo con realizaciones de la presente invención.
- La FIG. 21A es una tabla 2100 que muestra las densidades de metilación del plasma preoperatorio y las muestras de tejido de un paciente con carcinoma hepatocelular (CHC). La FIG. 21B es una tabla 2150 que muestra el número de lecturas de secuencia y la profundidad de secuenciación lograda por muestra.
- 10 La FIG. 22 es una tabla 220 que muestra las densidades de metilación en los autosomas, entre el 71,2 % y el 72,5 %, en las muestras de plasma de los controles sanos.
- 15 Las FIGS. 23A y 23B muestran la densidad de metilación de la capa leucocitaria, tejido tumoral, tejido hepático no tumoral, el plasma preoperatorio y postoperatorio del paciente con CHC.
- Las FIGS. 24A es un gráfico 2400 que muestra las densidades de metilación del plasma preoperatorio del paciente con CHC. Las FIGS. 24B es un gráfico 2450 que muestra las densidades de metilación del plasma postoperatorio del paciente con CHC.
- 20 Las FIGS. 25A y 25B muestran las puntuaciones z de las densidades de metilación del ADN plasmático para las muestras plasmáticas preoperatorias (gráfica 2500) y postoperatorias (gráfica 2550) del paciente con CHC, utilizando los datos del metiloma plasmático de los cuatro sujetos de control sanos como referencia para el cromosoma 1.
- 25 La FIG. 26A es una tabla 2600 que muestra los datos de las puntuaciones z del plasma preoperatorio y postoperatorio. La FIG. 26B es un diagrama Circos 2620 que muestra la puntuación z de las densidades de metilación del ADN plasmático para las muestras de plasma preoperatorio y postoperatorio del paciente con CHC utilizando los cuatro sujetos de control sanos como referencia para bins de 1 Mb analizados de todos los autosomas. La FIG. 26C es una tabla 2640 que muestra una distribución de las puntuaciones z de los bins de 1 Mb para todo el genoma en las muestras de plasma preoperatorio y postoperatorio del paciente con CHC. La FIG. 26D es una tabla 2660 que muestra los niveles de metilación del tejido tumoral y de la muestra de plasma preoperatorio que se solapan con algunas de las muestras de plasma de control cuando se utilizan los contextos de CHH y CHG.
- 30 Las FIG. 27A-H muestran diagramas Circos de la densidad de metilación de 8 pacientes con cáncer de acuerdo con las realizaciones de la presente invención. La FIG. 27I es una tabla 2780 que muestra el número de lecturas de secuencia y la profundidad de secuenciación lograda por muestra. La FIG. 27J es una tabla 2790 que muestra una distribución de las puntuaciones z de los bins de 1 Mb para todo el genoma en plasma de pacientes con diferentes tumores malignos. ACP = adenocarcinoma de pulmón; CNF = carcinoma nasofaríngeo; CCR = carcinoma colorrectal; NE = carcinoma neuroendocrino; SML = sarcoma de músculo liso.
- 40 La FIG. 28 es un diagrama de flujo del método 2800 de análisis de una muestra biológica de un organismo para determinar una clasificación de un nivel de cáncer de acuerdo con realizaciones de la presente invención.
- 45 La FIG. 29A es un gráfico 2900 que muestra la distribución de las densidades de metilación en sujetos de referencia suponiendo que esta distribución sigue una distribución normal. La Fig. 29B es un gráfico 2950 que muestra la distribución de las densidades de metilación en sujetos con cáncer suponiendo que esta distribución sigue una distribución normal y el nivel medio de metilación está 2 desviaciones típicas por debajo del valor de corte.
- 50 La FIG. 30 es un gráfico 3000 que muestra la distribución de las densidades de metilación del ADN plasmático de sujetos sanos y pacientes con cáncer.
- La FIG. 31 es un gráfico 3100 que muestra la distribución de las diferencias en las densidades de metilación entre la media del ADN plasmático de sujetos sanos y el tejido tumoral del paciente con CHC.
- 55 La FIG. 32A es una tabla 3200 que muestra el efecto de reducir la profundidad de secuenciación cuando la muestra de plasma contenía un 5 % o un 2 % de ADN tumoral.
- La FIG. 32B es un gráfico 3250 que muestra las densidades de metilación de los elementos repetidos y de las regiones no repetidas en el plasma de los cuatro sujetos de control sanos, la capa leucoplaquetaria, el tejido hepático normal, el tejido tumoral, las muestras de plasma preoperatorio y postoperatorio del paciente con CHC.
- 60 La FIG. 33 muestra un diagrama de bloques de un sistema informático 3300 ilustrativo que puede utilizarse con el sistema y los métodos de acuerdo con las realizaciones de la presente invención.
- 65 La FIG. 34A muestra una distribución de tamaños del ADN plasmático en el paciente con lupus eritematoso sistémico (LES) LES04. Las FIGS. 34B y 34C muestran análisis de metilación para ADN plasmático de un paciente con LES

LES04 (FIG. 34B) y un paciente con CHC TBR36 (FIG. 34C).

La FIG. 35 es un diagrama de flujo de un método 3500 que determina una clasificación de un nivel de cáncer basado en la hipermetilación de islas CpG de acuerdo con realizaciones de la presente invención.

- 5 La FIG. 36 es un diagrama de flujo que ilustra un método 3600 de análisis de una muestra biológica de un organismo para identificar el origen de una aberración cromosómica de acuerdo con realizaciones de la presente invención.
- 10 La FIG. 37A muestra el análisis de las ANC para tejidos tumorales, ADN plasmático no tratado con bisulfito (BS) y ADN plasmático tratado con bisulfito (de dentro a fuera) para el paciente TBR36. La FIG. 37B es un gráfico de dispersión que muestra la relación entre las puntuaciones z para la detección de ANC utilizando plasma tratado con bisulfito y no tratado con bisulfito de los bins de 1 Mb para el paciente TBR36.
- 15 La FIG. 38A muestra el análisis de las ANC para tejidos tumorales, ADN plasmático no tratado con bisulfito (BS) y ADN plasmático tratado con bisulfito (de dentro a fuera) para el paciente TBR34. La FIG. 38B es un gráfico de dispersión que muestra la relación entre las puntuaciones z para la detección de las ANC utilizando plasma tratado con bisulfito y no tratado con bisulfito de los bins de 1 Mb para el paciente TBR34.
- 20 La FIG. 39A es un diagrama Circos que muestra las ANC (anillo interior) y el análisis de metilación (anillo exterior) para el plasma tratado con bisulfito de un paciente con CHC TBR240. La FIG. 39B es un diagrama Circos que muestra las ANC (anillo interior) y el análisis de metilación (anillo exterior) para el plasma tratado con bisulfito de un paciente con CHC TBR164.
- 25 La FIG. 40A muestra el análisis de las ANC del paciente TBR36 para la muestra previa al tratamiento y la muestra después del tratamiento. La FIG. 40B muestra el análisis de metilación del paciente TBR36 para la muestra previa al tratamiento y la muestra después del tratamiento. La FIG. 41A muestra el análisis de las ANC del paciente TBR34 para la muestra previa al tratamiento y la muestra después del tratamiento. La FIG. 41B muestra el análisis de metilación del paciente TBR34 para la muestra previa al tratamiento y la muestra después del tratamiento.
- 30 La FIG. 42 muestra un diagrama del rendimiento diagnóstico del análisis de hipometilación en todo el genoma con diferente número de lecturas secuenciadas.
- 35 La FIG. 43 es un diagrama que muestra curvas ROC para la detección de cáncer basada en el análisis de hipometilación genómica con diferentes tamaños de bin (50 kb, 100 kb, 200 kb y 1 Mb).
- 40 La FIG. 44A muestra un rendimiento de diagnóstico para la probabilidad acumulada (PA) y el porcentaje de bins con aberraciones. La FIG. 44B muestra los resultados diagnósticos del análisis de plasma para la hipometilación global, hipermetilación de islas CpG y ANC.
- 45 La FIG. 45 muestra una tabla con los resultados de la hipometilación global, hipermetilación de islas CpG y ANC en pacientes con carcinoma hepatocelular.
- 50 La FIG. 46 muestra una tabla con los resultados de la hipometilación global, hipermetilación de islas CpG y ANC en pacientes que padecen cánceres distintos del carcinoma hepatocelular.
- 55 La FIG. 47 muestra un análisis en serie de la metilación plasmática para el caso TBR34.
- 60 La FIG. 48A muestra diagramas Circos que demuestran las ANC (anillo interior) y los cambios de metilación (anillo exterior) en el ADN plasmático tratado con bisulfito para el paciente con CHC TBR36. La FIG. 48B es un gráfico de puntuaciones z de metilación para regiones con ganancias y pérdidas cromosómicas, y regiones sin cambios en el número de copias para el paciente con CHC TBR36.
- 65 La FIG. 49A muestra diagramas Circos que demuestran las ANC (anillo interior) y los cambios de metilación (anillo exterior) en el ADN plasmático tratado con bisulfito para el paciente con CHC TBR34. La FIG. 49B es un gráfico de puntuaciones z de metilación para regiones con ganancias y pérdidas cromosómicas, y regiones sin cambios en el número de copias para el paciente con CHC TBR34.
- 70 Las FIG. 50A y 50B muestran los resultados de los análisis de hipometilación plasmática y ANC para los pacientes con LES LES04 y LES10.
- 75 Las FIGS. 51A y 51B muestran el análisis Z_{met} para regiones con y sin ANC para el plasma de dos pacientes con CHC (TBR34 y TBR36). Las FIGS. 51C y 51D muestran el análisis Z_{met} para regiones con y sin ANC para el plasma de dos pacientes con LES (LES04 y LES10).
- 80 La FIG. 52A muestra el análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las características del grupo A para las ANC, metilación global

- y metilación de islas CpG. La FIG. 52B muestra la agrupación jerárquica utilizando características del grupo B para las ANC, metilación global y metilación de islas CpG.
- 5 La FIG. 53A muestra el análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las características de metilación de las islas CpG del grupo A. La FIG. 53B muestra el análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las densidades de metilación global del grupo A.
- 10 La FIG. 54A muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las ANC globales del grupo A. La FIG. 54B muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las densidades de metilación de las islas CpG del grupo B.
- 15 La FIG. 55A muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las densidades de metilación global del grupo B. La FIG. 55B muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las densidades de metilación global del grupo B.
- 20 La FIG. 56 muestra la densidad de metilación media de bins de 1 Mb (puntos rojos) entre 32 sujetos sanos.
- 25 Un "metiloma" proporciona una medida de la cantidad de metilación del ADN en una pluralidad de sitios o locus en un genoma. El metiloma puede corresponder a todo el genoma, una parte sustancial del genoma o parte o partes relativamente pequeñas del genoma. Un "metiloma fetal" corresponde al metiloma del feto de una mujer gestante. El metiloma fetal puede determinarse utilizando diversos tejidos fetales o fuentes de ADN fetal, incluidos los tejidos placentarios y el ADN fetal extracelular en el plasma materno. Un "metiloma tumoral" corresponde al metiloma de un tumor de un organismo (p. ej., un ser humano). El metiloma tumoral puede determinarse utilizando tejido tumoral o ADN tumoral extracelular en plasma materno. El metiloma fetal y el metiloma tumoral son ejemplos de metiloma de interés. Otros ejemplos de metilomas de interés son los metilomas de órganos (p. ej., metilomas de células cerebrales, huesos, los pulmones, el corazón, los músculos y los riñones, etc.) que pueden aportar ADN a un fluido corporal (p. ej., plasma, suero, sudor, saliva, orina, secreciones genitales, semen, líquido de las heces, líquido diarreico, líquido cefalorraquídeo, secreciones del tubo digestivo, secreciones pancreáticas, secreciones intestinales, esputo, lágrimas, líquidos de aspiración de mama y tiroides, etc.). Los órganos pueden ser órganos trasplantados.
- 30 35 Un "metiloma plasmático" es el metiloma determinado a partir del plasma o suero de un animal (p. ej., un ser humano). El metiloma plasmático es un ejemplo de un metiloma acelular ya que el plasma y el suero incluyen ADN acelular. El metiloma plasmático también es un ejemplo de metiloma mixto, ya que es una mezcla de metiloma fetal/materno o metiloma tumoral/paciente. El "metiloma placentario" puede determinarse a partir de una muestra de vellosidades coriónicas (MVC) o de una muestra de tejido placentario (p. ej., obtenida tras el parto). El "metiloma celular" corresponde al metiloma determinado a partir de células (p. ej., glóbulos sanguíneos) del paciente. El metiloma de los glóbulos sanguíneos se denomina metiloma de los glóbulos sanguíneos (o metiloma sanguíneo).
- 40 45 Un "sitio" corresponde a un único sitio, que puede ser una sola posición de base o un grupo de posiciones de bases correlacionadas, p. ej., un sitio CpG. Un "locus" puede corresponder a una región que incluye múltiples sitios. Un locus puede incluir solo un sitio, lo que haría que el locus fuera equivalente a un sitio en ese contexto.
- 50 55 Un "índice de metilación" de cada sitio genómico (p. ej., un sitio CpG) se refiere a la proporción de lecturas de secuencias que muestran metilación en el sitio sobre el número total de lecturas que cubren ese sitio. La "densidad de metilación" de una región es el número de lecturas en sitios dentro de la región que muestran metilación dividido entre el número total de lecturas que cubren los sitios en la región. Los sitios pueden tener características específicas, p. ej., que son sitios CpG. Por tanto, la "densidad de metilación de CpG" de una región es el número de lecturas que muestran la metilación de CpG dividido entre el número total de lecturas que cubren los sitios CpG en la región (p. ej., un sitio CpG particular, sitios CpG dentro de una isla de CpG o una región más grande). Por ejemplo, la densidad de metilación para cada bin de 100 kb en el genoma humano se puede determinar a partir del número total de citosinas no convertidas después del tratamiento con bisulfito (que corresponde a la citosina metilada) en los sitios CpG como una proporción de todos los sitios CpG cubiertos por las lecturas de secuencia asignadas a la región de 100 kb. Este análisis también se puede realizar para otros tamaños de bins, p. ej., 50 kb o 1 Mb, etc. Una región podría ser el genoma completo o un cromosoma o parte de un cromosoma (p. ej., un brazo cromosómico). El índice de metilación de un sitio CpG es el mismo que la densidad de metilación de una región cuando la región solo incluye ese sitio CpG. La "proporción de citosinas metiladas" se refiere al número de sitios de citosina, "C", que se muestra que están metilados (p. ej., no convertidos después de la conversión con bisulfito) sobre el número total de restos de citosina analizados, es decir, incluyendo citosinas fuera del contexto de CpG, en la región. El índice de metilación, la densidad de metilación y la proporción de citosinas metiladas son ejemplos de "niveles de metilación".
- 60 65 Un "perfil de metilación" (también denominado estado de metilación) incluye información relacionada con la metilación del ADN para una región. La información relacionada con la metilación del ADN puede incluir, aunque no de forma limitativa, un índice de metilación de un sitio CpG, una densidad de metilación de los sitios CpG en una región, una

- distribución de sitios CpG en una región contigua, un patrón o nivel de metilación para cada sitio CpG individual dentro de una región que contiene más de un sitio CpG y metilación distinta de CpG. Un perfil de metilación de una parte sustancial del genoma puede considerarse equivalente al metiloma. La "metilación del ADN" en los genomas de mamíferos normalmente se refiere a la adición de un grupo metilo al carbono 5' de los restos de citosina (es decir, 5-metilcitosinas) entre los dinucleótidos CpG. La metilación del ADN puede producirse en citosinas en otros contextos, por ejemplo, CHG y CHH, donde H es adenina, citosina o timina. La metilación de citosina también puede ser en forma de 5-hidroximetilcitosina. También se ha descrito la metilación no citosínica, tal como la N6-metiladenina.
- 5 Un "tejido" corresponde a cualquier célula. Los distintos tipos de tejidos pueden corresponder a diferentes tipos de células (p. ej., hígado, pulmón o sangre), pero también pueden corresponder a tejidos de organismos diferentes (madre frente a feto) o a células sanas frente a células tumorales. Una "*muestra biológica*" se refiere a cualquier muestra que se toma de un sujeto (p. ej., un ser humano, tal como una mujer gestante, una persona con cáncer o una persona que se sospecha que tiene cáncer, un receptor de un trasplante de órganos o un sujeto que se sospecha que padece una enfermedad que afecta a un órgano (p. ej., el corazón en el infarto de miocardio, o el cerebro en el ictus) y que contiene 10 una o más moléculas de ácido nucleico de interés. La muestra biológica puede ser un líquido corporal, tal como sangre, plasma, suero, orina, fluido vaginal, líquidos de enjuague uterino o vaginal, líquido pleural, líquido ascítico, líquido cefalorraquídeo, saliva, sudor, lágrimas, esputo, líquido de lavado broncoalveolar, etc. También pueden utilizarse muestras de heces.
- 15 20 La expresión "*nivel de cáncer*" puede referirse a si existe cáncer, un estadio de un cáncer, un tamaño del tumor, si hay metástasis, la carga tumoral total del organismo y/u otra medida de la gravedad de un cáncer. El nivel de cáncer podría ser un número u otros caracteres. El nivel podría ser cero. El nivel de cáncer también incluye las afecciones (estados) premalignas o precancerosas asociadas a mutaciones o a un número de mutaciones. El nivel de cáncer puede usarse de diversas maneras. Por ejemplo, el cribado puede comprobar si el cáncer está presente en una persona que no se sabe que haya tenido cáncer con anterioridad. La evaluación puede investigar a una persona a la que se le ha diagnosticado un cáncer para supervisar la evolución del mismo a lo largo del tiempo, estudiar la eficacia de las terapias o determinar el pronóstico. En una realización, el pronóstico puede expresarse como la probabilidad de que un paciente muera de cáncer o la probabilidad de que el cáncer progrese después de una duración o tiempo específicos, o la probabilidad de que el cáncer metastatice. La detección puede significar "cribado" o puede significar comprobar si 25 30 alguien, con características indicativas de cáncer (p. ej., síntomas u otras pruebas positivas), tiene cáncer.
- 35 Los mecanismos epigenéticos desempeñan un papel importante en el desarrollo embrionario y fetal. Sin embargo, los tejidos embrionarios y fetales humanos (incluidos los tejidos placentarios) no son fácilmente accesibles (patente de EE. UU. n.º 6.927.028). Ciertas realizaciones han abordado este problema analizando una muestra que tiene moléculas de ADN fetal libres de células presentes en el sistema circulatorio materno. El metiloma fetal puede deducirse de varias maneras. Por ejemplo, el metiloma plasmático materno puede compararse con un metiloma celular (de glóbulos sanguíneos de la madre) y se demuestra que la diferencia está correlacionada con el metiloma fetal. Como otro ejemplo, los alelos específicos del feto pueden utilizarse para determinar la metilación del metiloma fetal en locus específicos. Adicionalmente, el tamaño de un fragmento puede utilizarse como indicador de un porcentaje de 40 metilación, como se muestra una correlación entre el tamaño y el porcentaje de metilación.
- 45 En una realización, la secuenciación con bisulfito de todo el genoma se utiliza para analizar el perfil de metilación (parte o la totalidad de un metiloma) del ADN plasmático materno con una resolución de un solo nucleótido. Aprovechando las diferencias polimórficas entre la madre y el feto, el metiloma fetal podría ensamblarse a partir de muestras de sangre materna. En otra implementación, no se utilizaron diferencias polimórficas, pero se puede utilizar un diferencial entre el metiloma del plasma y el metiloma de los glóbulos sanguíneos.
- 50 55 En otra realización, explotando las variaciones de un solo nucleótido y/o las aberraciones del número de copias entre un genoma tumoral y un genoma no tumoral, y los datos de secuenciación del plasma (u otra muestra), se puede realizar el perfil de metilación de un tumor en la muestra de un paciente del que se sospecha o se sabe que tiene cáncer. Una diferencia en el nivel de metilación en una muestra de plasma de un individuo de prueba cuando se compara con el nivel de metilación en plasma de un control sano o de un grupo de controles sanos puede permitir la identificación del individuo de prueba como portador de cáncer. Adicionalmente, el perfil de metilación puede actuar como un distintivo que revele el tipo de cáncer, por ejemplo, de qué órgano, que ha desarrollado la persona y si se ha producido metástasis.
- 60 Debido a la naturaleza no invasiva de este enfoque, pudimos evaluar en serie los metilomas plasmáticos fetal y materno a partir de muestras de sangre materna recogidas en el primer trimestre, tercer trimestre y después del parto. Se observaron cambios relacionados con la gestación. El enfoque también puede aplicarse a muestras obtenidas durante el segundo trimestre. El metiloma fetal deducido del plasma materno durante la gestación se parecía al metiloma placentario. Se identificaron genes sellados y regiones metiladas diferencialmente a partir de los datos del plasma materno.
- 65 Por lo tanto, hemos desarrollado un enfoque para estudiar el metiloma fetal de forma no invasiva, en serie y de forma exhaustiva, ofreciendo así la posibilidad de identificar biomarcadores o pruebas directas de patologías relacionadas con la gestación. Las realizaciones también pueden utilizarse para estudiar el metiloma tumoral de forma no invasiva,

en serie y de forma exhaustiva, para cribar o detectar si un sujeto padece cáncer, para el seguimiento de enfermedades malignas en un paciente con cáncer y para el pronóstico. Las realizaciones pueden aplicarse a cualquier tipo de cáncer, incluidos, aunque no de forma limitativa, cáncer de pulmón, cáncer de mama, cáncer colorrectal, cáncer de próstata, cáncer nasofaríngeo, cáncer gástrico, cáncer testicular, cáncer de piel (p. ej., melanoma), cáncer que afecta al sistema nervioso, cáncer de huesos, cáncer de ovario, cáncer de hígado (p. ej., carcinoma hepatocelular), neoplasias malignas hematológicas, cáncer pancreático, endometriocarcinoma, cáncer de riñón, cáncer cervicouterino, cáncer de vejiga, etc.

5 En primer lugar, se describe cómo determinar un metiloma o perfil de metilación y, a continuación, se describen diferentes metilomas (como los metilomas fetales, un metiloma tumoral, metilomas de la madre o de un paciente, y un metiloma mixto, p. ej., de plasma). A continuación, se describe la determinación de un perfil de metilación fetal utilizando marcadores específicos del feto o comparando un perfil de metilación mixto con un perfil de metilación celular. Los marcadores de metilación fetal se determinan comparando los perfiles de metilación. Se analiza la relación entre el tamaño y la metilación. También se proporcionan usos de los perfiles de metilación para detectar el cáncer.

10 15 I. DETERMINACIÓN DE UN METILOMA

Para investigar el metiloma placentario se han utilizado múltiples enfoques, pero cada enfoque tiene sus limitaciones. Por ejemplo, el bisulfito de sodio, una sustancia química que modifica los restos de citosina no metilada a uracilo y 20 deja la citosina metilada sin cambios, convierte las diferencias en la metilación de la citosina en una diferencia de secuencia genética para su posterior exploración. El método de referencia para estudiar la metilación de la citosina se basa en el tratamiento del ADN tisular con bisulfito sódico seguido de la secuenciación directa de clones individuales de moléculas de ADN convertidas con bisulfito. Tras el análisis de múltiples clones de moléculas de ADN, se puede obtener el patrón de metilación de la citosina y el perfil cuantitativo por sitio CpG. Sin embargo, la secuenciación con bisulfito clonada es un procedimiento de bajo rendimiento y laborioso que no puede aplicarse fácilmente a escala de todo el genoma.

25 Las enzimas de restricción sensibles a la metilación, que suelen digerir el ADN no metilado, proporcionan un enfoque de bajo coste para estudiar la metilación del ADN. Sin embargo, los datos generados a partir de estos estudios se limitan a los locus con motivos de reconocimiento enzimático y los resultados no son cuantitativos. La inmunoprecipitación de ADN unido por anticuerpos antimetilados de citosina puede utilizarse para estudiar grandes segmentos del genoma, pero tiende a sesgar hacia locus con metilación densa debido a la mayor fuerza de unión del anticuerpo a tales regiones. Los enfoques basados en micromatrizes dependen del diseño *a priori* de las sondas de exploración y de las eficacias de hibridación entre las sondas y el ADN diana.

30 35 40 Para explorar exhaustivamente un metiloma, en algunas realizaciones, se utiliza la secuenciación masiva en paralelo (MPS, *Massively Parallel Sequencing*) a fin de proporcionar información de todo el genoma y una evaluación cuantitativa del nivel de metilación por nucleótido y por alelo. Recientemente, se ha hecho factible la conversión a bisulfito seguida de la MPS en todo el genoma (R. Lister *et al.*, 2008 *Cell*; 133: 523-536).

45 Entre el reducido número de estudios publicados (R. Lister *et al.* 2009 *Nature*; 462: 315-322; L. Laurent *et al.* 2010 *Genome Res*; 20: 320-331; Y Li *et al.* 2010 *PLoS Biol*; 8: e1000533; y M. Kulis *et al.* 2012 *Nat Genet*; 44: 1236-1242) que aplicó la secuenciación con bisulfito de todo el genoma a la investigación de los metilomas humanos, dos estudios se centraron en células madre embrionarias y fibroblastos fetales (R. Lister *et al.* 2009 *Nature*; 462: 315-322; L. Laurent *et al.* 2010 *Genome Res*; 20: 320-331). Ambos estudios analizaron el ADN derivado de la línea celular.

A. Secuenciación con bisulfito de todo el genoma

50 Ciertas realizaciones pueden superar los retos mencionados y permitir la exploración de un metiloma fetal de forma exhaustiva, de forma no invasiva y en serie. En una realización, se utilizó la secuenciación con bisulfito de todo el genoma para analizar las moléculas de ADN fetal extracelular que se encuentran en el sistema circulatorio de las mujeres gestantes. A pesar de la escasa abundancia y la naturaleza fragmentada de las moléculas de ADN plasmático, pudimos ensamblar un metiloma fetal de alta resolución a partir del plasma materno y observar en serie los cambios con la progresión de la gestación. Dado el intenso interés por las pruebas prenatales no invasivas (NIPT, *NonInvasive Prenatal Testing*), pueden proporcionar una nueva y potente herramienta para el descubrimiento de biomarcadores fetales o servir como plataforma directa para lograr aplicar las NIPT de enfermedades fetales o asociadas a la gestación. Ahora están disponibles los datos de la secuenciación con bisulfito de todo el genoma de varias muestras, a partir de los cuales puede obtenerse el metiloma fetal. En una realización, esta tecnología puede aplicarse para elaborar perfiles de metilación en gestaciones complicadas con preeclampsia, o retraso del crecimiento intrauterino o parto prematuro. Para gestaciones tan complicadas, esta tecnología puede utilizarse en serie gracias a su naturaleza no invasiva, para permitir el seguimiento y/o el pronóstico y/o la respuesta al tratamiento.

60 65 La FIG. 1A muestra una tabla 100 de resultados de secuenciación de sangre materna, placenta y plasma materno de acuerdo con realizaciones de la presente invención. En una realización, la secuenciación del genoma completo se realizó a partir de bibliotecas de ADN convertidas con bisulfito, preparada utilizando adaptadores de bibliotecas de ADN metilado (Illumina) (R. Lister *et al.* 2008 *Cell*; 133: 523-536), de glóbulos sanguíneos de la muestra de sangre

recogida en el primer trimestre, la MVC, el tejido placentario recogido a término, las muestras de plasma materno recogidas durante el primer y tercer trimestres y el período después del parto. También se analizaron muestras de ADN de glóbulos sanguíneos y plasma obtenidas de un varón adulto y una mujer adulta no gestante. En este estudio se generó un total de 9500 millones de pares de lecturas de secuencias sin procesar. La cobertura de secuenciación de cada muestra se muestra en la tabla 100.

Las lecturas de secuencias que se podían asignar de forma exclusiva al genoma humano de referencia alcanzaron coberturas genómicas haploides medias de 50 veces, 34 veces y 28 veces, respectivamente, durante el primer trimestre, muestras de plasma materno del tercer trimestre y posteriores al parto. La cobertura de los sitios CpG en el genoma varió entre el 81 % y el 92 % para las muestras obtenidas de la gestación. Las lecturas de secuencias que abarcaron sitios CpG ascendieron a coberturas haploides promedio de 33 veces por cadena, 23 veces por cadena y 19 veces por cadena, respectivamente, durante el primer trimestre, muestras de plasma materno del tercer trimestre y posteriores al parto. Las eficacias de conversión del bisulfito para todas las muestras fueron >99,9 % (tabla 100).

En la tabla 100, la tasa de ambigüedad (marcada con una "a") se refiere a la proporción de lecturas cartografiadas tanto en la cadena Watson como en la cadena Crick del genoma humano de referencia. La tasa de conversión lambda se refiere a la proporción de citosinas no metiladas en el control interno del ADN lambda que se convierten en restos de "timina" mediante la modificación con bisulfito. H equivale genéricamente a A, C o T. "a" se refiere a lecturas que podrían asignarse a un locus genómico específico, pero que no pueden asignarse a la cadena Watson o Crick, "b" se refiere a lecturas emparejadas con coordenadas iniciales y finales idénticas. Para "c", ADN lambda en cada muestra antes de la conversión a bisulfito. La tasa de conversión lambda se refiere a la proporción de nucleótidos de citosina que permanecen como citosina después de la conversión con bisulfito y se utiliza como una indicación de la tasa de conversión con bisulfito exitosa, "d" se refiere al número de nucleótidos de citosina presentes en el genoma humano de referencia y que permanecen como secuencia de citosina tras la conversión con bisulfito.

Durante la modificación con bisulfito, las citosinas no metiladas se convierten en uracilos y posteriormente en timinas tras las amplificaciones por PCR, mientras que las citosinas metiladas permanecerán intactas (M. Frommer *et al.* 1992 *Proc Natl Acad Sci EE. UU.*;89:1827-31). Tras la secuenciación y la alineación, de este modo, el estado de metilación de un sitio CpG individual podría inferirse a partir del número de lecturas de secuencias metiladas "M" (metilado) y el número de lecturas de secuencias no metiladas "N" (no metilado) en el resto de citosina en el contexto de CpG. Utilizando los datos de secuenciación con bisulfito, se construyeron los metilomas completos de la sangre materna, la placenta y el plasma materno. La densidad media de CpG metilados (también denominada densidad de metilación DM) de locus específicos en el plasma materno puede calcularse mediante la ecuación:

$$DM = \frac{M}{M + N}$$

donde M es el número de lecturas metiladas y N es el número de lecturas no metiladas en los sitios CpG dentro del locus genético. Si hay más de un sitio CpG en un locus, entonces M y N corresponden a los números en todos los sitios.

40 B. Técnicas diferentes

Como se ha descrito anteriormente, el perfil de metilación puede realizarse mediante secuenciación masiva paralela (MPS) de ADN plasmático convertido con bisulfito. La MPS del ADN plasmático convertido con bisulfito puede realizarse de forma aleatoria o escopeteada. La profundidad de la secuenciación puede variar en función del tamaño de la región de interés.

En otra realización, la región o regiones de interés en el ADN plasmático convertido con bisulfito pueden capturarse primero mediante un proceso basado en la hibridación en fase sólida o en fase de solución, seguido por la MPS. La secuenciación masiva en paralelo puede realizarse utilizando una plataforma de secuenciación por síntesis como la Illumina, una plataforma de secuenciación por ligación como la plataforma SOLiD de Life Technologies, un sistema de secuenciación basado en semiconductores como las plataformas Ion Torrent o Ion Proton de Life Technologies, o un sistema de secuenciación de una sola molécula como el sistema Helicos o el sistema Pacific Biosciences o un sistema de secuenciación basado en nanoporos. La secuenciación basada en nanoporos que incluye nanoporos que se construyen mediante, por ejemplo, bicapas lipídicas y nanoporos de proteínas, y nanoporos en estado sólido (tales como los basados en grafeno). Como las plataformas de secuenciación de moléculas individuales seleccionadas permitirían conocer el estado de metilación de las moléculas de ADN (incluida la N6-metiladenina, 5-metilcitosina y 5-hidroximetilcitosina) a dilucidar directamente sin conversión con bisulfito (B. A. Flusberg *et al.* 2010 *Nat Methods*; 7: 461-465; J Shim *et al.* 2013 *Sci Rep*; 3:1389. doi: 10.1038/srep01389), el uso de dichas plataformas permitiría analizar el estado de metilación del ADN de muestras no convertidas con bisulfito (p. ej., el ADN plasmático).

Además de la secuenciación, pueden utilizarse otras técnicas. En una realización, el perfil de metilación puede realizarse mediante PCR específica de metilación o digestión con enzima de restricción sensible a la metilación

seguida de PCR o reacción en cadena de la ligasa seguida de PCR. En otras realizaciones más, la PCR es una forma de PCR de una sola molécula o digital (B. Vogelstein *et al.* 1999 *Proc Natl Acad Sci EE. UU.*; 96: 9236-9241). En otras realizaciones adicionales, la PCR puede ser una PCR en tiempo real. En otras realizaciones, la PCR puede ser PCR multiplex.

5

II. ANÁLISIS DE METILOMAS

- Algunas realizaciones pueden determinar el perfil de metilación del ADN plasmático utilizando la secuenciación del genoma completo con bisulfito. El perfil de metilación de un feto puede determinarse secuenciando muestras de ADN plasmático materno, como se describe a continuación. Por tanto, se accedió a las moléculas de ADN fetal (y al metiloma fetal) de forma no invasiva durante la gestación, y se controlaron los cambios de forma seriada a medida que avanzaba la gestación. Debido a la exhaustividad de los datos de secuenciación, pudimos estudiar los metilomas del plasma materno a escala de todo el genoma con una resolución de un solo nucleótido.
- Dado que se conocían las coordenadas genómicas de las lecturas secuenciadas, estos datos permiten estudiar los niveles globales de metilación del metiloma o de cualquier región de interés del genoma y realizar comparaciones entre distintos elementos genéticos. Adicionalmente, múltiples lecturas de secuencias cubrieron cada sitio CpG o locus. A continuación, se describen algunas de las métricas utilizadas para medir el metiloma.

20 *A. Metilación de las moléculas de ADN plasmático*

Las moléculas de ADN están presentes en el plasma humano a bajas concentraciones y de forma fragmentada, normalmente en longitudes que se asemejan a las unidades mononucleosómicas (YMD Lo *et al.* 2010 *Sci Transl Med*; 2: 61ra91; y YW Zheng *et al.* 2012 *Clin Chem*; 58: 549-558). A pesar de estas limitaciones, una línea de secuenciación con bisulfito de todo el genoma fue capaz de analizar la metilación de las moléculas de ADN plasmático. En otras realizaciones más, ya que las plataformas de secuenciación de una sola molécula seleccionadas permitirían dilucidar directamente el estado de metilación de las moléculas de ADN sin conversión con bisulfito (B. A. Flusberg *et al.* 2010 *Nat Methods*; 7: 461-465; J Shim *et al.* 2013 *Sci Rep*; 3:1389. doi: 10.1038/srep01389), el uso de dichas plataformas permitiría utilizar el ADN plasmático no convertido en bisulfito para determinar los niveles de metilación del ADN plasmático o para determinar el metiloma plasmático. Estas plataformas pueden detectar la N6-metiladenina, 5-metilcitosina y 5-hidroximetilcitosina, que pueden proporcionar mejores resultados (p. ej., mayor sensibilidad o especificidad) relacionados con las diferentes funciones biológicas de las distintas formas de metilación. Dichos resultados mejorados pueden ser útiles al aplicar las realizaciones para la detección o el seguimiento de trastornos específicos, p. ej., preeclampsia o un tipo concreto de cáncer.

La secuenciación con bisulfito también puede diferenciar entre distintas formas de metilación. En una realización, se pueden incluir etapas adicionales que permitan distinguir la 5-metilcitosina de la 5-hidroximetilcitosina. Uno de estos métodos es la secuenciación oxidativa con bisulfito (oxBS-seq), que puede dilucidar la localización de la 5-metilcitosina y la 5-hidroximetilcitosina con resolución de una sola base (M. J. Booth *et al.* 2012 *Science*; 336: 934-937; M. J. Booth *et al.* 2013 *Nature Protocols*; 8: 1841-1851). En la secuenciación con bisulfito, tanto la 5-metilcitosina como la 5-hidroximetilcitosina se leen como citosinas y, por tanto, no se pueden diferenciar. Por otra parte, en oxBS-seq, la oxidación específica de la 5-hidroximetilcitosina a 5-formilcitosina mediante tratamiento con perrutenoato de potasio (KRUO4), seguido de la conversión de la 5-formilcitosina recién formada a uracilo mediante bisulfito permitiría distinguir la 5-hidroximetilcitosina de la 5-metilcitosina. Por consiguiente, se puede obtener una lectura de 5-metilcitosina a partir de una única ejecución de oxBS-seq, y los niveles de 5-hidroximetilcitosina se deducen mediante la comparación con los resultados de la secuenciación con bisulfito. En otra realización, la 5-metilcitosina puede distinguirse de la 5-hidroximetilcitosina mediante secuenciación con bisulfito asistida por Tet (TAB-seq) (M. Yu *et al.* 2012 *Nat Protoc*; 7: 2159-2170). TAB-seq puede identificar la 5-hidroximetilcitosina con una resolución de una sola base, así como determinar su abundancia en cada sitio de modificación. Este método implica la protección de la 5-hidroximetilcitosina mediada por la β-glucosiltransferasa (glucosilación) y la oxidación de la 5-metilcitosina a 5-carboxilcitosina mediada por la Tet1 (mTet1) en ratones recombinantes. Tras el tratamiento posterior con bisulfito y la amplificación por PCR, tanto la citosina como la 5-carboxilcitosina (derivada de la 5-metilcitosina) se convierten en timina (T), mientras que la 5-hidroximetilcitosina se leerá como C.

La FIG. 1B muestra la densidad de metilación en ventanas de 1 Mb de muestras secuenciadas de acuerdo con realizaciones de la presente invención. El gráfico 150 es un diagrama Circos que representa la densidad de metilación en el plasma materno y el ADN genómico en ventanas de 1 Mb a través del genoma. De fuera a dentro: los ideogramas cromosómicos pueden orientarse de pter a qter en el sentido de las agujas del reloj (los centrómeros aparecen en rojo), sangre materna (roja), placenta (amarillo), plasma materno (verde), lecturas compartidas en el plasma materno (azul) y lecturas específicas del feto en el plasma materno (púrpura). Los niveles globales de metilación de CpG (es decir, los niveles de densidad) de los glóbulos sanguíneos maternos, placenta y plasma materno se pueden encontrar en la tabla 100. El nivel de metilación de los glóbulos sanguíneos maternos es, en general, superior al de la placenta en todo el genoma.

65 *B. Comparación de la secuenciación con bisulfito con otras técnicas*

Estudiamos el metiloma placentario mediante secuenciación masiva paralela con bisulfito. Adicionalmente, estudiamos el metiloma placentario utilizando una plataforma de matrices de oligonucleótidos que cubría unos 480.000 sitios CpG del genoma humano (Illumina) (M Kuhs *et al.* 2012 *Nat Genet*; 44: 1236-1242; y C Clark *et al.* 2012 *PLoS One*; 7: e50233). En una realización, utilizando el genotipado basado en chips de microperlas y análisis de metilación, el genotipado se realizó utilizando la matriz de genotipado Illumina HumanOmni2.5-8 de acuerdo con el protocolo del fabricante. Los genotipos se leyeron utilizando el algoritmo GenCall del software Genome Studio (Illumina). Los índices de lectura superaron el 99 %. Para el análisis de metilación basado en micromatrizes, se trató el ADN genómico (500-800 ng) con bisulfito sódico utilizando el kit de metilación de ADN Zymo EZ (Zymo Research, Orange, CA, EE. UU.) de acuerdo con las recomendaciones del fabricante para el ensayo de metilación Illumina Infinium.

El ensayo de metilación se realizó en 4 µl de ADN genómico convertido con bisulfito a 50 ng/µl de acuerdo con el protocolo del ensayo de metilación Infinium HD. El chip de microperla hibridado se escaneó en un instrumento Illumina iScan. Los datos de metilación del ADN se analizaron con el software GenomeStudio (v2011.1) Methylation Module (vl.9.0), con normalización a controles internos y sustracción de fondo. El índice de metilación de cada sitio CpG se representó mediante un valor beta (β), que se calculó utilizando la proporción de intensidades fluorescentes entre los alelos metilados y no metilados:

$$\beta = \frac{\text{Intensidad del alelo metilado}}{\text{Intensidad del alelo no metilado} + \text{Intensidad del alelo metilado} + 100}$$

Para los sitios CpG que estaban representados en la matriz y secuenciados con una cobertura de al menos 10 veces, se comparó el valor beta obtenido por la matriz con el índice de metilación determinado por secuenciación del mismo sitio. Los valores beta representan la intensidad de las sondas metiladas como proporción de la intensidad combinada de las sondas metiladas y no metiladas que cubren el mismo sitio CpG. El índice de metilación para cada sitio CpG se refiere a la proporción de lecturas metiladas sobre el número total de lecturas que cubren ese CpG.

Las FIGS. 2A-2C muestran gráficas de los valores beta determinados por la matriz de chips de microesferas Illumina Infinium HumanMethylation 450K frente a los índices de metilación determinados por secuenciación de todo el genoma con bisulfito de los sitios CpG correspondientes que fueron explorados por ambas plataformas: (A) Células sanguíneas maternas, (B) Muestra de vellosidades coriónicas, (C) Tejido placentario a término. Los datos de ambas plataformas fueron muy concordantes y los coeficientes de correlación de Pearson fueron de 0,972, 0,939 y 0,954, y los valores R² de 0,945, 0,882 y 0,910 para las células de la sangre materna, MVC y tejido placentario a término, respectivamente.

Además, comparamos nuestros datos de secuenciación con los de Chu *et al.*, que investigaron los perfiles de metilación de 12 pares de muestras de ADN de CVS y de glóbulos sanguíneos maternos utilizando una matriz de oligonucleótidos que cubría aproximadamente 27.000 sitios CpG (T Chu *et al.* 2011 *PLoS One*; 6: e14723). Los datos de correlación entre los resultados de la secuenciación de la MVC y el ADN de los glóbulos sanguíneos maternos y cada uno de los 12 pares de muestras del estudio anterior proporcionaron un coeficiente de Pearson medio (0,967) y R² (0,935) para la sangre materna y un coeficiente de Pearson promedio (0,943) y R² (0,888) para la MVC. Entre los sitios CpG representados en ambas matrices, nuestros datos se correlacionaron en gran medida con los datos publicados. Los índices de metilación sin CpG fueron < 1 % para los glóbulos sanguíneos maternos, MVC y tejidos placentarios (tabla 100). Estos resultados eran coherentes con la creencia actual de que cantidades sustanciales de metilación sin CpG estaban restringidas principalmente a las células pluripotentes (R Lister *et al.* 2009 *Nature*; 462: 315-322; L. Laurent *et al.* 2010 *Genome Res*; 20: 320-331).

45 C. Comparación de los metilomas plasmáticos y sanguíneos de las mujeres no gestantes

Las FIGS. 3A y 3B muestran gráficos de barras del porcentaje de sitios CpG metilados en plasma y glóbulos sanguíneos recogidos de un varón adulto y una mujer adulta no gestante: (A) Autosomas, (B) Cromosoma X. Los gráficos muestran una similitud entre los metilomas plasmáticos y sanguíneos de un varón y una mujer no gestante. Las proporciones totales de sitios CpG que estaban metilados en las muestras de plasma de varón y de mujer no gestante eran casi las mismas que las del ADN celular sanguíneo correspondiente (tabla 100 y FIGS. 2A y 2B).

A continuación, estudiamos la correlación de los perfiles de metilación de las muestras de plasma y de glóbulos sanguíneos de forma específica para cada locus. Determinamos la densidad de metilación de cada bin de 100 kb en el genoma humano determinando el número total de citosinas no convertidas en sitios CpG como proporción de todos los sitios CpG cubiertos por lecturas de secuencias cartografiadas en la región de 100 kb. Las densidades de metilación fueron altamente concordantes entre la muestra de plasma y el correspondiente ADN de glóbulos sanguíneos de las muestras de varón y de mujer.

Las FIGS. 4A y 4B muestran gráficos de las densidades de metilación de los loci correspondientes en el ADN de glóbulos sanguíneos y en el ADN plasmático: (A) Mujer adulta no gestante, (B) Varón adulto. El coeficiente de correlación de Pearson y el valor R² para las muestras de mujeres no gestantes fueron respectivamente de 0,963 y 0,927, y el de las muestras de varones fueron respectivamente de 0,953 y 0,908. Estos datos concuerdan con hallazgos anteriores basados en la evaluación de los genotipos de las moléculas de ADN plasmático de receptores de trasplante alogénico de células madre hematopoyéticas que demostraron que las células hematopoyéticas son la

fuente predominante de ADN en el plasma humano (YW Zheng *et al.* 2012 *Clin Chem*; 58: 549-558).

D. Niveles de metilación en los metilomas

5 A continuación, estudiamos los niveles de metilación del ADN del plasma materno, glóbulos sanguíneos maternos y tejido placentario para determinar los niveles de metilación. Los niveles se determinaron para regiones repetidas, regiones no repetidas, y en total.

10 Las FIGS. 5A y 5B muestran gráficos de barras del porcentaje de sitios CpG metilados entre las muestras recogidas de la gestación: (A). Autosomas, (B) Cromosoma X. Las proporciones totales de CpG metilados fueron del 67,0 % y 68,2 % para las muestras de plasma materno del primer y tercer trimestre, respectivamente. A diferencia de los resultados obtenidos en las personas no gestantes, estas proporciones eran inferiores a las de la muestra de glóbulos sanguíneos maternos del primer trimestre, pero superiores a las de las muestras de MVC y de tejido placentario a término (tabla 100). Cabe destacar que, el porcentaje de CpG metilados para la muestra de plasma materno postparto 15 fue del 73,1 %, similar a los datos de los glóbulos sanguíneos (tabla 100). Estas tendencias se observaron en CpG distribuidos por todos los autosomas, así como por el cromosoma X, y abarcaban tanto las regiones no repetidas como múltiples clases de elementos repetidos del genoma humano.

20 Se observó que tanto los elementos repetidos como los no repetidos de la placenta estaban hipometilados en relación con los glóbulos sanguíneos maternos. Los resultados concordaron con los hallazgos publicados de que la placenta está hipometilada en relación con otros tejidos, incluidas las células de sangre periférica.

25 Entre el 71 % y el 72 % de los sitios CpG secuenciados estaban metilados en el ADN de los glóbulos sanguíneos de la mujer gestante, mujer no gestante y varón adulto (tabla 100 de la FIG. 1). Estos datos son comparables con el informe del 68,4 % de sitios CpG de células mononucleares sanguíneas comunicado por Y Li *et al.*

30 2010 *PLoS Biol*; 8: e1000533. En consonancia con los informes anteriores sobre la naturaleza hipometilada de los tejidos placentarios, el 55 % y el 59 % de los sitios CpG estaban metilados en la MVC y en el tejido de placenta a término, respectivamente (tabla 100).

35 30 La FIG. 6 muestra un gráfico de barras del nivel de metilación de diferentes clases de repetición del genoma humano para la sangre materna, placenta y plasma materno. Las clases de repetición son las definidas por el navegador del genoma de la UCSC. Los datos mostrados corresponden a las muestras del primer trimestre. A diferencia de datos anteriores que indicaban que la naturaleza hipometilada de los tejidos placentarios se observaba principalmente en determinadas clases de repeticiones del genoma (B Novakovic *et al.* 2012 *Placenta*; 33: 959-970), aquí demostramos que la placenta estaba de hecho hipometilada en la mayoría de las clases de elementos genómicos con referencia a los glóbulos sanguíneos.

E. Similitud de los metilomas

40 40 Las realizaciones pueden determinar los metilomas de los tejidos placentarios, glóbulos sanguíneos y plasma utilizando la misma plataforma. Por consiguiente, fue posible realizar comparaciones directas de los metilomas de esos tipos de muestras biológicas. El alto nivel de similitud entre los metilomas de los glóbulos sanguíneos y el plasma del varón y la mujer no gestante, así como entre los glóbulos sanguíneos maternos y la muestra de plasma materno posterior al parto, afirmó aún más que las células hematopoyéticas eran las principales fuentes de ADN en el plasma humano (Y. W. Zheng *et al.* 2012 *Clin Chem*; 58: 549-558).

45 50 Las semejanzas son evidentes tanto en términos de la proporción total de CpG metilados en el genoma como de la alta correlación de las densidades de metilación entre los locus correspondientes en el ADN de los glóbulos sanguíneos y el ADN plasmático. Aún así, las proporciones globales de CpG metilados en las muestras de plasma materno del primer y tercer trimestre se redujeron en comparación con los datos de glóbulos sanguíneos maternos o la muestra de plasma materno posterior al parto. Los niveles reducidos de metilación durante la gestación se debieron a la naturaleza hipometilada de las moléculas de ADN fetal presentes en el plasma materno.

55 60 55 La inversión del perfil de metilación en la muestra de plasma materno posterior al parto para asemejarse más al de los glóbulos sanguíneos maternos indica que las moléculas de ADN fetal habían sido eliminadas del sistema circulatorio materno. El cálculo de las concentraciones de ADN fetal basado en los marcadores de SNP (*Single Nucleotide Polymorphism*, polimorfismo de un solo nucleótido) del feto demostró que la concentración había pasado del 33,9 % antes del parto a solo el 4,5 % en la muestra posterior al parto.

F. Otras aplicaciones

65 Se han ensamblado con éxito metilomas de ADN mediante el análisis MPS de ADN plasmático. La capacidad de determinar el metiloma placentario o fetal a partir del plasma materno proporciona un método no invasivo para determinar, detectar y controlar los perfiles de metilación aberrantes asociados a afecciones asociadas a la gestación, tales como la preeclampsia, el retraso del crecimiento intrauterino, el parto prematuro y otros. Por ejemplo, la detección

de un distintivo de metilación aberrante específica de la enfermedad permite el cribado, diagnóstico y seguimiento de estas afecciones asociadas a la gestación. La medición del nivel de metilación del plasma materno permite el cribado, diagnóstico y seguimiento de estas afecciones asociadas a la gestación. Además de las aplicaciones directas en la investigación de afecciones asociadas a la gestación, el enfoque podría aplicarse a otros ámbitos de la medicina en los que el análisis del ADN plasmático resulte de interés. Por ejemplo, los metilomas de los cánceres podrían determinarse a partir del ADN plasmático de pacientes con cáncer. El análisis metilómico del cáncer a partir del plasma, como se describe en el presente documento, es potencialmente una tecnología sinérgica para el análisis genómico del cáncer a partir del plasma (KCA Chan *et al.* 2013 *Clin Chem*; 59: 211-224 y R. J. Leary *et al.* 2012 *Sci Transl Med*; 4:162ral54).

Por ejemplo, la determinación del nivel de metilación de una muestra de plasma podría utilizarse para detectar el cáncer. Cuando el nivel de metilación de la muestra de plasma muestra niveles aberrantes en comparación con los controles sanos, puede sospecharse la presencia de cáncer. A continuación, puede realizarse una confirmación y evaluación adicionales del tipo de cáncer o del origen tisular del cáncer determinando el perfil plasmático de metilación en diferentes locus genómicos o mediante análisis genómicos plasmáticos para detectar aberraciones del número de copias asociadas al tumor, translocaciones cromosómicas y variantes de un solo nucleótido. En efecto, en una realización de la presente invención, el perfil metilómico y genómico del cáncer de plasma puede llevarse a cabo simultáneamente. Como alternativa, se podrían utilizar investigaciones radiológicas y de imagen (p. ej., tomografía computarizada, resonancia magnética, tomografía por emisión de positrones) o endoscopia (p. ej., endoscopia gastrointestinal superior o colonoscopia) para investigar más a fondo a los individuos de los que se sospechaba que tenían cáncer basándose en el análisis del nivel de metilación en plasma.

Para el cribado o la detección del cáncer, la determinación del nivel de metilación de una muestra de plasma (u otra sustancia biológica) puede utilizarse junto con otras modalidades de cribado o detección del cáncer, tales como la medición del antígeno prostático específico (p. ej., para el cáncer de próstata), antígeno carcinoembrionario (p. ej., para carcinoma colorrectal, carcinoma gástrico, carcinoma de páncreas, carcinoma de pulmón, carcinoma de mama, carcinoma medular de tiroides, alfa fetoproteína (p. ej., para el cáncer de hígado o los tumores de células germinales), CA125 (p. ej., para el cáncer de ovario y de mama) y CA19-9 (p. ej., para el carcinoma pancreático).

Adicionalmente, se pueden secuenciar otros tejidos para obtener un metiloma celular. Por ejemplo, se puede analizar el tejido hepático para determinar un patrón de metilación específico del hígado, que puede utilizarse para identificar patologías hepáticas. Otros tejidos que también pueden analizarse son las células cerebrales, los huesos, los pulmones, el corazón, los músculos y los riñones, etc. Los perfiles de metilación de diferentes tejidos pueden cambiar de vez en cuando, p. ej., como resultado del desarrollo, envejecimiento, procesos patológicos (p. ej., inflamación o cirrosis o procesos autoinmunitarios (tales como en el lupus eritematoso sistémico)) o terapéuticos (p. ej., tratamiento con agentes desmetilantes como la 5-azacitidina y la 5-azadeoxicitidina). La naturaleza dinámica de la metilación del ADN hace que este análisis sea potencialmente muy valioso para el seguimiento de procesos fisiológicos y patológicos. Por ejemplo, si se detecta un cambio en el metiloma plasmático de un individuo en comparación con un valor de referencia obtenido cuando estaba sano, se podrían detectar procesos patológicos en los órganos que aportan ADN plasmático.

Del mismo modo, los metilomas de los órganos transplantados podrían determinarse a partir del ADN plasmático de los receptores de trasplantes de órganos. El análisis metilómico de trasplantes a partir de plasma, como se describe en la presente invención, es potencialmente una tecnología sinérgica para trasplantar el análisis genómico del plasma (Y. W. Zheng *et al.*, 2012; YMD Lo *et al.* 1998 *Lancet*; 351: 1329-1330; y T. M. Snyder *et al.* 2011 *Proc Natl Acad Sci EE. UU.*; 108: 6229-6234). El ADN plasmático suele considerarse un marcador de la muerte celular, un aumento del nivel plasmático de ADN liberado por un órgano transplantado podría utilizarse como marcador de un aumento de la muerte celular de dicho órgano, tal como un episodio de rechazo u otros procesos patológicos que afecten a ese órgano (p. ej., infección o absceso). En caso de que se instaure con éxito la terapia antirrechazo, se espera que se reduzca el nivel plasmático de ADN liberado por el órgano transplantado.

III. DETERMINACIÓN DEL METILOMA FETAL O TUMORAL MEDIANTE SNP

Como se ha descrito anteriormente, el metiloma plasmático corresponde al metiloma sanguíneo de una persona normal no gestante. Sin embargo, para una mujer gestante, los metilomas difieren. Las moléculas de ADN fetal circulan en el plasma materno entre un fondo mayoritario de ADN materno (YMD Lo *et al.* 1998 *Am J Hum Genet*; 62: 768-775). Por tanto, para una mujer gestante, el metiloma plasmático es en gran medida un compuesto del metiloma placentario y del metiloma sanguíneo. En consecuencia, se puede extraer el metiloma placentario del plasma.

En una realización, las diferencias de polimorfismo de un solo nucleótido (SNP) entre la madre y el feto se utilizan para identificar las moléculas de ADN fetal en el plasma materno. El objetivo era identificar los locus SNP en los que la madre es homocigótica, pero el feto es heterocigótico; el alelo específico del feto puede utilizarse para determinar qué fragmentos de ADN proceden del feto. El ADN genómico de los glóbulos sanguíneos maternos se analizó mediante una matriz de genotipado de SNP, la Illumina HumanOmni2.5-8. Por otra parte, para los locus de SNP en los que la madre es heterocigótica y el feto homocigótico, entonces el alelo SNP específico de la madre puede utilizarse para determinar qué fragmentos de ADN plasmático proceden de la madre. El nivel de metilación de dichos fragmentos de

ADN reflejaría el nivel de metilación de las regiones genómicas relacionadas en la madre.

A. Correlación de la metilación de las lecturas específicas del feto y el metiloma placentario

- 5 Los locus que tienen dos alelos diferentes, donde la cantidad de un alelo (B) era significativamente menor que la del otro alelo (A), se identificaron a partir de los resultados de secuenciación de una muestra biológica. Las lecturas que cubrían los alelos B se consideraron específicas del feto (lecturas específicas del feto). Se determina que la madre es homocigótica para A y que el feto es heterocigótico para A/B, por lo que las lecturas que cubren el alelo A fueron compartidas por la madre y el feto (lecturas compartidas).
- 10 En un caso de gestación analizado que se utilizó para ilustrar varios de los conceptos de esta invención, la madre gestante resultó ser homocigótica en 1.945.516 locus de los autosomas. Se examinaron las lecturas de secuenciación del ADN del plasma materno que cubrían estos SNP. Se detectaron lecturas portadoras de un alelo no materno en 107.750 locus, que se consideraron los locus informativos. En cada SNP informativo, el alelo que no procedía de la madre se denominó alelo específico del feto, mientras que el otro se denominó alelo compartido.
- 15

Se puede determinar una concentración fraccional de ADN fetal/tumoral (también denominada porcentaje de ADN fetal) en el plasma materno. En una realización, la concentración fraccional de ADN fetal en el plasma materno, f , viene determinada por la ecuación:

20

$$f = \frac{2p}{p + q}$$

donde p es el número de lecturas secuenciadas con el alelo específico del feto y q es el número de lecturas secuenciadas con el alelo compartido entre la madre y el feto (YMD Lo *et al.* 2010 *Sci Transl Med*; 2: 61ra91). Se comprobó que las proporciones de ADN fetal en las muestras de plasma materno del primer trimestre, del tercer trimestre y posteriores al parto eran del 14,4 %, 33,9 % y 4,5 %, respectivamente. Las proporciones de ADN fetal también se calcularon utilizando los números de lecturas que se alinearon con el cromosoma Y. Basándose en los datos del cromosoma Y, los resultados fueron del 14,2 %, 34,9 % y 3,7 %, respectivamente, en las muestras de plasma materno del primer trimestre, del tercer trimestre y después del parto.

30

Analizando por separado las lecturas de secuencias específicas del feto o compartidas, las realizaciones demuestran que las moléculas de ADN fetal circulante estaban mucho más hipometiladas que las moléculas de ADN de fondo. Las comparaciones de las densidades de metilación de los locus correspondientes en las lecturas de plasma materno específicas del feto y los datos de tejido placentario tanto para el primer trimestre como para el tercero revelaron altos niveles de correlación. Estos datos proporcionaron pruebas a nivel del genoma de que la placenta es la fuente predominante de moléculas de ADN derivadas del feto en el plasma materno y representaron un gran paso adelante en comparación con las pruebas anteriores basadas en información derivada de locus seleccionados.

40

Determinamos la densidad de metilación de cada región de 1 Mb en el genoma utilizando las lecturas específicas del feto o compartidas que cubrían los sitios CpG adyacentes a los SNP informativos. Se pueden visualizar los metilomas fetales y no fetales específicos ensamblados a partir de las lecturas de secuencias del plasma materno, por ejemplo, en gráficas Circos (M Krzywinski *et al.* 2009 *Genome Res*; 19: 1639-1645). También se determinaron las densidades de metilación por bin de 1 Mb para los glóbulos sanguíneos maternos y las muestras de tejido placentario.

45

La FIG. 7A muestra un diagrama Circos 700 para muestras del primer trimestre. La FIG. 7B muestra un diagrama Circos 750 para muestras del tercer trimestre. Los gráficos 700 y 750 muestran la densidad de metilación por cada bin de 1 Mb. Los ideogramas de los cromosomas (anillo más externo) están orientados de pter a qter en el sentido de las agujas del reloj (los centrómeros aparecen en rojo). El segundo carril más exterior muestra el número de sitios CpG en las correspondientes regiones de 1 Mb. La escala de las barras rojas mostradas es de hasta 20.000 sitios por cada bin de 1 Mb. Las densidades de metilación de las correspondientes regiones de 1 Mb se muestran en los otros carriles basándose en el esquema de colores mostrado en el centro.

50

Para las muestras del primer trimestre (FIG. 7A), de dentro a fuera, los carriles son: muestra de vellosidades coriónicas, lecturas específicas del feto en el plasma materno, lecturas específicas de la madre en el plasma materno, lecturas fetales y no fetales combinadas en el plasma materno y en los glóbulos sanguíneos maternos. Para las muestras del tercer trimestre (FIG. 7B), los carriles son: tejido de placenta a término, lecturas específicas del feto en el plasma materno, lecturas específicas de la madre en el plasma materno, lecturas fetales y no fetales combinadas en el plasma materno, plasma materno después del parto y glóbulos sanguíneos maternos (de la muestra de sangre del primer trimestre). Se puede apreciar que tanto para las muestras de plasma del primer como del tercer trimestre, los metilomas fetales estaban más hipometilados que los de los metilomas no específicos del feto.

55

El perfil total de metilación de los metilomas fetales se asemejaba más al de las muestras de MVC o de tejido placentario. Por el contrario, el perfil de metilación del ADN de las lecturas compartidas en plasma, que eran

- predominantemente de ADN materno, se asemejó más al de los glóbulos sanguíneos maternos. A continuación, realizamos una comparación sistemática locus por locus de las densidades de metilación de las lecturas de ADN del plasma materno y de los tejidos maternos o fetales. Se determinaron las densidades de metilación de los sitios CpG que estaban presentes en la misma lectura de secuencia que los SNP informativos y que estaban cubiertos por al menos 5 lecturas de secuencia de ADN plasmático materno.
- Las FIGS 8A-8D muestran gráficas de comparaciones de las densidades de metilación del ADN del tejido genómico frente al ADN del plasma materno para los sitios CpG que rodean los polimorfismos de un solo nucleótido informativos. La FIG. 8A muestra las densidades de metilación para lecturas específicas del feto en la muestra de plasma materno del primer trimestre en relación con las densidades de metilación para lecturas en una muestra de MVC. Como puede observarse, los valores específicos del feto se corresponden bien con los valores de las MVC.
- La FIG. 8B muestra las densidades de metilación para lecturas específicas del feto en la muestra de plasma materno del tercer trimestre en relación con las densidades de metilación para lecturas en un tejido placentario a término. De nuevo, los conjuntos de densidades se corresponden bien, lo que indica que el perfil de metilación fetal puede obtenerse analizando lecturas con alelos específicos del feto.
- La FIG. 8C muestra las densidades de metilación de las lecturas compartidas en la muestra de plasma materno del primer trimestre en relación con las densidades de metilación de las lecturas en los glóbulos sanguíneos maternos. Dado que la mayoría de las lecturas compartidas son de la madre, los dos conjuntos de valores se corresponden bien. La FIG. 8D muestra las densidades de metilación de las lecturas compartidas en la muestra de plasma materno del tercer trimestre en relación con las densidades de metilación de las lecturas en los glóbulos sanguíneos maternos.
- Para las lecturas específicas del feto en el plasma materno, el coeficiente de correlación de Spearman entre el plasma materno del primer trimestre y la MVC fue de 0,705 ($P < 2,2 \times 10^{-16}$); y el existente entre el plasma materno del tercer trimestre y el tejido de placenta a término fue de 0,796 ($P < 2,2 \times 10^{-16}$) (FIGS. 8A y 8B). Se realizó una comparación similar para las lecturas compartidas en plasma materno con los datos de glóbulos sanguíneos maternos. El coeficiente de correlación de Pearson fue de 0,653 ($P < 2,2 \times 10^{-16}$) para la muestra de plasma del primer trimestre y de 0,638 ($P < 2,2 \times 10^{-16}$) para la muestra de plasma del tercer trimestre (FIGS. 8C y 8D).
- B. Metiloma fetal**
- En una realización, para ensamblar el metiloma fetal a partir del plasma materno, clasificamos las lecturas de secuencias que abarcaban al menos un sitio SNP fetal informativo y que contenían al menos un sitio CpG dentro de la misma lectura. Las lecturas que mostraban los alelos específicos del feto se incluyeron en el ensamblaje del metiloma fetal. Las lecturas que mostraron el alelo compartido, es decir, el alelo no específico del feto, se incluyeron en el ensamblaje del metiloma no específico del feto, que estaba compuesto predominantemente por moléculas de ADN de origen materno.
- Las lecturas específicas del feto cubrieron 218.010 sitios CpG en los autosomas de las muestras de plasma materno del primer trimestre. Las cifras correspondientes a las muestras de plasma materno del tercer trimestre y posteriores al parto fueron de 263.611 y 74.020, respectivamente. Como término medio, las lecturas compartidas cubrieron esos sitios CpG un promedio de 33,3, 21,7 y 26,3 veces, respectivamente. Las lecturas específicas del feto cubrieron esos sitios CpG 3,0, 4,4 y 1,8 veces, respectivamente, durante el primer trimestre, muestras de plasma materno del tercer trimestre y posteriores al parto.
- El ADN fetal representa una población menor en el plasma materno y, por tanto, la cobertura de esos sitios CpG por lecturas específicas del feto fue proporcional al porcentaje de ADN fetal de la muestra. Para la muestra de plasma materno del primer trimestre, el porcentaje total de CpG metilados entre las lecturas fetales fue del 47,0 %, mientras que el de las lecturas compartidas fue del 68,1 %. Para la muestra de plasma materno del tercer trimestre, el porcentaje de CpG metilados de las lecturas fetales fue del 53,3 %, mientras que el de las lecturas compartidas fue del 68,8 %. Estos datos mostraron que las lecturas específicas del feto en el plasma materno estaban más hipometiladas que las lecturas compartidas en el plasma materno.
- C. Método**
- Las técnicas descritas anteriormente también pueden utilizarse para determinar un perfil de metilación tumoral. A continuación, se describen métodos para determinar los perfiles de metilación fetal y tumoral.
- La FIG. 9 es un diagrama de flujo que ilustra un método 900 para determinar un primer perfil de metilación a partir de una muestra biológica de un organismo de acuerdo con realizaciones de la presente invención. El método 900 puede construir un mapa epigenético del feto a partir del perfil de metilación del plasma materno. La muestra biológica incluye ADN extracelular que comprende una mezcla de ADN extracelular procedente de un primer tejido y de un segundo tejido. Como ejemplos, el primer tejido puede ser de un feto, un tumor o un órgano transplantado.
- En el bloque 910, se analiza una pluralidad de moléculas de ADN de la muestra biológica. El análisis de una molécula

de ADN puede incluir la determinación de una localización de la molécula de ADN en un genoma del organismo, determinar un genotipo de la molécula de ADN, y determinar si la molécula de ADN está metilada en uno o más sitios.

5 En una realización, las moléculas de ADN se analizan utilizando lecturas de secuencias de las moléculas de ADN, donde la secuenciación tiene en cuenta la metilación. Por tanto, las lecturas de secuencia incluyen el estado de metilación de las moléculas de ADN de la muestra biológica. El estado de metilación puede incluir si un determinado resto de citosina es 5-metilcitosina o 5-hidroximetilcitosina. Las lecturas de secuencias pueden obtenerse a partir de diversas técnicas de secuenciación, técnicas de PCR, matrices, y otras técnicas adecuadas para identificar secuencias de fragmentos. El estado de metilación de los sitios de la secuencia leída puede obtenerse como se describe en el 10 presente documento.

En el bloque 920, se identifica una pluralidad de primeros locus en los que un primer genoma del primer tejido es heterocigótico para un primer alelo respectivo y un segundo alelo respectivo y un segundo genoma del segundo tejido es homocigótico para el primer alelo respectivo. Por ejemplo, pueden identificarse lecturas específicas del feto en la 15 pluralidad de primeros locus o pueden identificarse lecturas específicas del tumor en la pluralidad de primeros locus. Las lecturas específicas del tejido pueden identificarse a partir de lecturas de secuenciación en las que el porcentaje de lecturas de secuencia del segundo alelo se encuentre dentro de un intervalo determinado, p. ej., de aproximadamente el 3 %-25 %, indicando así una población minoritaria de fragmento de ADN de un genoma heterocigótico en el locus y una población mayoritaria de un genoma homocigótico en el locus.

20 En el bloque 930, se analizan las moléculas de ADN ubicadas en uno o más sitios de cada uno de los primeros locus. Se determina el número de moléculas de ADN que están metiladas en un sitio y que corresponden al segundo alelo respectivo del locus. Puede haber más de un sitio por locus. Por ejemplo, un SNP puede indicar que un fragmento es 25 específico del feto, y ese fragmento puede tener múltiples sitios cuyo estado de metilación está determinado. Se puede determinar el número de lecturas en cada sitio que están metiladas y el número total de lecturas metiladas para el locus.

El locus puede estar definido por un número específico de sitios, un conjunto específico de sitios o un tamaño particular 30 para una región en torno a una variación que comprende el alelo específico del tejido. Un locus puede tener un solo sitio. Los sitios pueden tener propiedades específicas, p. ej., que son sitios CpG. La determinación de un número de lecturas que no están metiladas es equivalente, y se engloba dentro de la determinación del estado de metilación.

En el bloque 940, para cada uno de los primeros locus, se calcula una densidad de metilación basada en el número 35 de moléculas de ADN metiladas en uno o más sitios del locus y correspondientes al segundo alelo respectivo del locus. Por ejemplo, se puede determinar una densidad de metilación para los sitios CpG correspondientes a un locus.

En el bloque 950, se crea el primer perfil de metilación del primer tejido a partir de las densidades de metilación para 40 los primeros locus. El primer perfil de metilación puede corresponder a sitios concretos, p. ej., sitios CpG. El perfil de metilación puede ser para todos los locus que tengan un alelo específico del feto o solo para algunos de esos locus.

IV. USO DE LA DIFERENCIA DE LOS METILOMAS PLASMÁTICOS Y SANGUÍNEOS

Anteriormente, se demostró que las lecturas específicas del feto procedentes del plasma se correlacionan con el 45 metiloma placentario. Como el componente materno del metiloma del plasma materno es aportado principalmente por los glóbulos sanguíneos, la diferencia entre el metiloma plasmático y el metiloma sanguíneo puede utilizarse para determinar el metiloma placentario para todos los locus, y no solo las localizaciones de los alelos específicos del feto. La diferencia entre el metiloma plasmático y el metiloma sanguíneo también puede utilizarse para determinar el metiloma de un tumor.

A. Método

La FIG. 10 es un diagrama de flujo que ilustra un método 1000 para determinar un primer perfil de metilación a partir 55 de una muestra biológica de un organismo de acuerdo con realizaciones de la presente invención. La muestra biológica (p. ej., plasma) incluye ADN extracelular que comprende una mezcla de ADN extracelular procedente de un primer tejido y de un segundo tejido. El primer perfil de metilación corresponde a un perfil de metilación del primer tejido (p. ej., tejido fetal o tejido tumoral). El método 1200 puede proporcionar una deducción de regiones metiladas diferencialmente a partir del plasma materno.

En el bloque 1010, se recibe una muestra biológica. La muestra biológica podría simplemente recibirse en una máquina 60 (p. ej., una máquina de secuenciación). La muestra biológica puede estar en la forma tomada del organismo o puede estar en una forma procesada, p. ej., la muestra puede ser plasma extraído de una muestra de sangre.

En el bloque 1020, se obtiene un segundo perfil de metilación correspondiente al ADN del segundo tejido. El segundo perfil de metilación podía leerse de memoria, como se haya determinado previamente. El segundo perfil de metilación 65 puede determinarse a partir del segundo tejido, p. ej., una muestra diferente que contenga solo o predominantemente células del segundo tejido. El segundo perfil de metilación puede corresponder a un perfil de metilación celular y

obtenerse a partir de ADN celular. Como otro ejemplo, el segundo perfil puede determinarse a partir de una muestra de plasma recogida antes de la gestación, o antes del desarrollo del cáncer, porque el metiloma plasmático de una persona no gestante sin cáncer es muy similar al metiloma de los glóbulos sanguíneos.

- 5 El segundo perfil de metilación puede proporcionar una densidad de metilación en cada uno de una pluralidad de locus en un genoma del organismo. La densidad de metilación en un locus concreto corresponde a una proporción de ADN del segundo tejido que está metilado. En una realización, la densidad de metilación es una densidad de metilación de CpG, donde los sitios CpG asociados al locus se utilizan para determinar la densidad de metilación. Si hay un sitio para un locus, entonces la densidad de metilación puede ser igual al índice de metilación. La densidad de metilación también corresponde a una densidad de desmetilación, ya que ambos valores son complementarios.

En una realización, el segundo perfil de metilación se obtiene mediante la secuenciación teniendo en cuenta la metilación del ADN celular de una muestra del organismo. Un ejemplo de secuenciación en la que se tiene en cuenta metilación incluye el tratamiento del ADN con bisulfito sódico y la posterior secuenciación del ADN. En otro ejemplo,

- 15 la secuenciación en la que se tiene en cuenta la metilación puede realizarse sin utilizar bisulfito sódico, utilizando una plataforma de secuenciación de una sola molécula que permite conocer el estado de metilación de las moléculas de ADN (incluida la N6-metiladenina, 5-metilcitosina y 5-hidroximetilcitosina) que se han de dilucidar directamente sin conversión a bisulfito (A. B. Flusberg *et al.* 2010 *Nat Methods*; 7: 461-465; J Shim *et al.* 2013 *Sci Rep*; 3:1389. doi: 10.1038/srep01389); o mediante la inmunoprecipitación de la citosina metilada (p. ej., utilizando un anticuerpo contra la metilcitosina o utilizando una proteína o un péptido de unión al ADN metilado (L. G. Acevedo *et al.* 2011 *Epigenomics*; 3: 93-101) seguido de secuenciación; o mediante el uso de enzimas de restricción sensibles a la metilación seguido de la secuenciación. En otra realización, se utilizan técnicas de no secuenciación, tales como las matrices, PCR digital y espectrometría de masas.

- 20 25 En otra realización, la segunda densidad de metilación del segundo tejido podría obtenerse previamente a partir de muestras de control del sujeto o de otros sujetos. La densidad de metilación de otro sujeto puede actuar como un perfil de metilación de referencia que tenga densidades de metilación de referencia. Las densidades de metilación de referencia pueden determinarse a partir de múltiples muestras, donde un nivel medio (u otro valor estadístico) de las diferentes densidades de metilación en un locus puede utilizarse como densidad de metilación de referencia en el

30 locus.

En el bloque 1030, se determina un perfil de metilación extracelular a partir del ADN extracelular de la mezcla. El perfil de metilación extracelular proporciona una densidad de metilación en cada uno de la pluralidad de locus. El perfil de metilación extracelular puede determinarse recibiendo lecturas de secuencia de una secuenciación del ADN extracelular, donde la información de metilación se obtiene con las lecturas de la secuencia. El perfil de metilación extracelular puede determinarse del mismo modo que el metiloma celular.

- 35 40 En el bloque 1040, se determina un porcentaje del ADN extracelular del primer tejido en la muestra biológica. En una realización, el primer tejido es tejido fetal, y el ADN correspondiente es ADN fetal. En otra realización, el primer tejido es tejido tumoral, y el ADN correspondiente es ADN tumoral. El porcentaje puede determinarse de varias maneras, p. ej., utilizando un alelo específico del feto o un alelo específico del tumor. El número de copias también puede utilizarse para determinar el porcentaje, p. ej., como se describe en la solicitud de patente de EE. UU. n.º 13/801.748 titulada "Mutational Analysis Of Plasma DNA For Cancer Detection".

- 45 En el bloque 1050, se identifica una pluralidad de locus para determinar el primer metiloma. Estos locus pueden corresponder a cada uno de los locus utilizados para determinar el perfil de metilación extracelular y el segundo perfil de metilación. Por tanto, la pluralidad de locus puede corresponder. Es posible que se utilicen más locus para determinar el perfil de metilación extracelular y el segundo perfil de metilación.

- 50 En algunas realizaciones, se pueden identificar los locus que estaban hipermetilados o hipometilados en el segundo perfil de metilación, p. ej., utilizando glóbulos sanguíneos maternos. Para identificar los locus hipermetilados en los glóbulos sanguíneos maternos, se puede buscar desde un extremo de un cromosoma un sitio CpG con un índice de metilación $\geq X\%$ (p. ej., 80 %). A continuación, se puede buscar el siguiente sitio CpG dentro de la región en dirección 3' (p. ej., en 200 pb em dirección 3'). Si el sitio CpG inmediatamente posterior en dirección 3' también tenía un índice de metilación $\geq X\%$ (u otra cantidad especificada), los sitios CpG primero y segundo pueden agruparse. La agrupación puede continuar hasta que no haya ningún otro sitio CpG en la siguiente región en dirección 3'; o el sitio CpG inmediatamente posterior en dirección 3' tenía un índice de metilación $< X\%$. La región de los sitios CpG agrupados puede notificarse como hipermetilada en los glóbulos sanguíneos maternos si la región contenía al menos cinco sitios CpG hipermetilados inmediatamente adyacentes. Se puede realizar un análisis similar para buscar locus hipometilados en glóbulos sanguíneos maternos para sitios CpG con índices de metilación $\leq 20\%$. Las densidades de metilación para el segundo perfil de metilación pueden calcularse para los locus preseleccionados y utilizarse para deducir el primer perfil de metilación (p. ej., la densidad de metilación del tejido placentario) de los locus correspondientes, p. ej., a partir de datos de secuenciación con bisulfito del plasma materno.

- 60 65 En el bloque 1060, el primer perfil de metilación del primer tejido se determina calculando un parámetro diferencial que incluye una diferencia entre la densidad de metilación del segundo perfil de metilación y la densidad de metilación del

perfil de metilación extracelular para cada uno de la pluralidad de locus. La diferencia se escala por el porcentaje.

En una realización, la primera densidad de metilación de un locus en el primer tejido (p. ej., la placenta)(D) se dedujo mediante la ecuación:

5

$$D = mbc - \frac{(mbc - mp)}{f * NC} \quad (1)$$

donde mbc denota la densidad de metilación del segundo perfil de metilación en un locus (p. ej., un locus preseleccionado determinado en los datos de la secuenciación con bisulfito de glóbulos sanguíneos maternos); mp denota la densidad de metilación del locus correspondiente en los datos de secuenciación con bisulfito del plasma materno; f representa el porcentaje de ADN extracelular del primer tejido (p. ej., concentración fraccional de ADN fetal) y NC representa el número de copias en el locus (p. ej., un valor más alto para amplificaciones o un número más bajo para delecciones en relación con lo normal). Si no hay amplificación o delección en el primer tejido, entonces el NC puede ser uno. Para la trisomía (o una duplicación de la región en un tumor o un feto), el NC sería de 1,5 (ya que el aumento es de 2 copias a 3 copias) y la monosomía tendría 0,5. Una mayor amplificación puede aumentar en incrementos de 0,5. En este ejemplo, D puede corresponder al parámetro diferencial.

10 En el bloque 1070, la primera densidad de metilación se transforma para obtener una primera densidad de metilación corregida del primer tejido. La transformación puede tener en cuenta las diferencias fijas entre los parámetros diferenciales y el perfil de metilación real del primer tejido. Por ejemplo, los valores pueden diferir en una constante fija o una constante pendiente. La transformación puede ser lineal o no lineal.

15 En una realización, la distribución de los valores deducidos, D , resultó ser inferior al nivel real de metilación del tejido placentario. Por ejemplo, los valores deducidos pueden transformarse linealmente utilizando datos de islas CpG, que eran segmentos genómicos que tenían una sobrerepresentación de sitios CpG. Las posiciones genómicas de las islas CpG utilizadas en este estudio se obtuvieron de la base de datos UCSC Genome Browser (NCBI build 36/hg18) (P. A. Fujita *et al.* 2011 *Nucleic Acids Res*; 39: D876-882). Por ejemplo, una isla CpG puede definirse como un segmento genómico con un contenido de GC $\geq 50\%$, longitud genómica >200 pb y la proporción entre el número de CpG observado/esperado $>0,6$ (M. Gardiner-Garden *et al.*, 1987 *J Mol Biol*; 196: 261-282).

20 En una implementación, para derivar la ecuación de transformación lineal, pueden incluirse las islas CpG con al menos 4 sitios CpG y una profundidad de lectura promedio ≥ 5 por sitio CpG en las muestras secuenciadas. Tras determinar las relaciones lineales entre las densidades de metilación de las islas CpG en las MVC o la placenta a término y los valores deducidos, D , se utilizaron las siguientes ecuaciones para determinar los valores previstos:

25

$$\text{Valores previstos en el primer trimestre} = D \times 1,6 + 0,2$$

$$\text{Valores previstos en el tercer trimestre} = D \times 1,2 + 0,05$$

30 40 B. Ejemplo fetal

Como se ha mencionado anteriormente, el método 1000 puede utilizarse para deducir un panorama de metilación de la placenta a partir del plasma materno. El ADN circulante en plasma procede predominantemente de células hematopoyéticas. Sin embargo, existe una proporción desconocida de ADN extracelular procedente de otros órganos internos. Asimismo, el ADN extracelular derivado de la placenta representa aproximadamente el 5-40 % del ADN total del plasma materno, con una media de aproximada el 15 %. Por tanto, cabe suponer que el nivel de metilación en el plasma materno es equivalente a una metilación basal existente más una contribución placentaria durante la gestación, como se ha descrito anteriormente.

45 50 El nivel de metilación del plasma materno, PM , puede determinarse mediante la siguiente ecuación:

$$PM = BASAL \times (1 - f) + NMP \times f$$

55 donde $BASAL$ es el nivel basal de metilación del ADN en plasma derivado de glóbulos sanguíneos y órganos internos, NMP es el nivel de metilación de la placenta y/es la concentración fraccional de ADN fetal en el plasma materno.

En una realización, el nivel de metilación de la placenta puede deducirse teóricamente mediante:

$$NMP = \frac{PM - BASAL}{f} \times (1 - f) \quad (2)$$

60 Las ecuaciones (1) y (2) son equivalentes cuando NC es igual a uno, D es igual a NMP , y $BASAL$ es igual a mbc . En

otra realización, la concentración fraccionaria de ADN fetal puede suponerse o fijarse en un valor especificado, p. ej., como parte de una suposición de un mínimo/estar presente.

- 5 Se tomó el nivel de metilación de la sangre materna para representar la metilación basal del plasma materno. Además de los locus que estaban hipermetilados o hipometilados en los glóbulos sanguíneos maternos, seguimos explorando el enfoque de la deducción centrándonos en regiones definidas con relevancia clínica, por ejemplo, las islas CpG en el genoma humano.
- 10 Se obtuvo la densidad media de metilación de un total de 27.458 islas CpG (NCBI Build36/hg18) en los autosomas y el cr. X a partir de los datos de secuenciación del plasma materno y la placenta. Solo se seleccionaron aquellos con ≥ 10 sitios CpG cubiertos y una profundidad de lectura promedio ≥ 5 por sitios CpG cubiertos en todas las muestras analizadas, incluida la placenta, sangre materna y plasma materno. Como resultado, 26.698 islas CpG (97,2 %) permanecieron como válidas y su nivel de metilación se dedujó utilizando los datos de metilación plasmática y la concentración fraccional de ADN fetal de acuerdo con la ecuación anterior.
- 15 15 Se observó que la distribución de los valores de *NMP* deducidos era inferior al nivel de metilación real de las islas CpG en el tejido placentario. Por tanto, en una realización, los valores de *NMP* deducidos, o simplemente los valores de (*D*) deducidos, se utilizaron como unidad arbitraria para estimar el nivel de metilación de las islas CpG en la placenta. Después de una transformación, los valores deducidos linealmente y su distribución se asemejó más a la del conjunto 20 de datos real. Los valores deducidos transformados se denominaron valores predictivos de metilación (VPM) y se utilizaron posteriormente para predecir el nivel de metilación de los locus genéticos en la placenta.
- 25 En este ejemplo, las islas CpG se clasificaron en 3 categorías en función de sus densidades de metilación en la placenta: Baja ($\leq 0,4$), Intermedia ($>0,4-<0,8$) y Alta ($\geq 0,8$). Utilizando la ecuación de deducción, se calculó el VPM del mismo conjunto de islas CpG y luego se utilizaron los valores para clasificarlas en 3 categorías con los mismos puntos de corte. Comparando los conjuntos de datos reales y los deducidos, se encontró que el 75,1 % de las islas CpG preseleccionadas se podían emparejar correctamente con las mismas categorías en los datos de tejido de acuerdo con su VPM. Aproximadamente el 22 % de las islas CpG se asignó a grupos con 1 nivel de diferencia (alta frente a intermedia, o intermedia frente a baja) y menos del 3 % se clasificarán completamente mal (alta frente a baja) (FIG. 30 12A). También se determinó el rendimiento total de la clasificación: 86,1 %, 31,4 % y 68,8 % de islas CpG con densidades de metilación $\leq 0,4$, $>0,4-<0,8$ y $\geq 0,8$ en la placenta se dedujeron como "Bajo", "Intermedio" y "Alto" correctamente (FIG. 12B).
- 35 40 Las FIGS. 11A y 11B muestran gráficos del rendimiento del algoritmo de predicción utilizando datos de plasma materno y concentración fraccional de ADN fetal de acuerdo con realizaciones de la presente invención. La FIG. 11A es un gráfico 1100 que muestra la precisión de la clasificación de islas CpG utilizando la clasificación de corrección del VPM (la categoría deducida coincide exactamente con el conjunto de datos real); la diferencia de 1 nivel (la categoría deducida difiere en 1 nivel del conjunto de datos real); y la clasificación errónea (la categoría deducida es opuesta a la del conjunto de datos real). La FIG. 11B es un gráfico 1150 que muestra la proporción de islas CpG clasificadas en cada categoría deducida.
- 45 Siempre que la metilación basal materna sea baja en las respectivas regiones genómicas, la presencia de ADN hipermetilado derivado de la placenta en el sistema circulatorio aumentaría el nivel total de metilación plasmática en un grado que dependería de la concentración fraccional de ADN fetal. Podría observarse un cambio notable cuando el ADN fetal liberado esté totalmente metilado. Por el contrario, cuando la metilación basal materna es elevada, el grado de cambio en el nivel de metilación plasmática sería más significativo si se liberara ADN fetal hipometilado. Por lo tanto, el esquema de deducción puede ser más práctico cuando el nivel de metilación se dedujó para locus genéticos que se sabe que son distintos entre el valor basal materno y la placenta, especialmente para aquellos marcadores hipermetilados e hipometilados en la placenta.
- 50 55 La FIG. 12A es una tabla 1200 que muestra detalles de 15 locus genómicos seleccionados para la predicción de metilación de acuerdo con realizaciones de la presente invención. Para confirmar las técnicas, se seleccionaron 15 locus genómicos metilados diferencialmente que habían sido estudiados previamente. Se dedujeron los niveles de metilación de las regiones seleccionadas y se compararon con los 15 locus genéticos metilados diferencialmente estudiados previamente (R. W. K. Chiu et al. 2007 *Am J Pathol*; 170: 941-950; S.S.C. Chim et al. 2008 *Clin Chem*; 54: 500-511; S. S. C. Chim et al. 2005 *Proc Natl Acad Sci EE. UU.*; 102: 14753-14758; D. W. Y. Tsui et al. 2010 *PLoS One*; 5: e15069).
- 60 65 La FIG. 12B es un gráfico 1250 que muestra las categorías deducidas de los 15 locus genómicos seleccionados y sus correspondientes niveles de metilación en la placenta. Las categorías de metilación deducidas son: Bajo, $\leq 0,4$; Intermedia, $>0,4-<0,8$; Alto, $\geq 0,8$. La tabla 1200 y el gráfico 1300 muestran que sus niveles de metilación en la placenta pudieron deducirse correctamente con varias excepciones: RASSF1A, CGI009, CGI137 y VAPA. De estos 4 marcadores, solo el CGI009 mostró una discrepancia notable con el conjunto de datos real. Los demás solo estaban marginalmente mal clasificados.

$$\frac{PM - BASAL \times (1-f)}{f}$$

- En la tabla 1200, "1" se refiere a los valores deducidos (D) que se calculan con la ecuación: $D =$ donde f es la fracción de concentración de ADN fetal. El número "2" se refiere a los valores predictivos de metilación (VPM) referidos a los valores deducidos transformados linealmente utilizando la ecuación: $VPM = D \times 1,6 + 0,25$. El número "3" se refiere al valor de corte de clasificación de los valores deducidos: Bajo, $\leq 0,4$; Inter(medio), $>0,4- <0,8$; Alto, $\geq 0,8$. El número "4" se refiere al valor de corte de clasificación para el conjunto de datos de placenta real: Bajo, $\leq 0,4$; Inter(medio), $>0,4- <0,8$; Alto, $\geq 0,8$. El número "5" indica que el estado placentario se refiere al estado de metilación de la placenta en relación con el de los glóbulos sanguíneos maternos.

C. Cálculo de las concentraciones fraccionarias de ADN fetal

- En una realización, el porcentaje de ADN fetal del primer tejido puede utilizar un cromosoma Y para un feto masculino. La proporción de secuencias del cromosoma Y (% cr. Y) en una muestra de plasma materno fue una combinación de las lecturas del cromosoma Y derivadas del feto masculino y el número de lecturas maternas (de mujer) que estaban mal alineadas con el cromosoma Y (R. W. K. Chiu *et al.* 2011 BMJ; 342: c7401). Por tanto, la relación entre % de cr. Y y la concentración fraccional de ADN fetal (f) en la muestra puede venir dada por:

$$\% \text{ de cr. Y} = \% \text{ de cr. } Y_{\text{masculino}} \times f + \% \text{ de cr. } Y_{\text{femenino}} \times (1 - f)$$

donde el % de cr. $Y_{\text{masculino}}$ se refiere a la proporción de lecturas alineadas con el cromosoma Y en una muestra de plasma que contiene un 100 % de ADN masculino; y el % de cr. Y_{femenino} se refiere a la proporción de lecturas alineadas con el cromosoma Y en una muestra de plasma que contiene un 100 % de ADN femenino.

El % de cr. Y puede determinarse a partir de lecturas que se alinearon con el cromosoma Y sin desajustes para una muestra de una mujer gestante de un feto masculino, p. ej., donde las lecturas proceden de muestras convertidas con bisulfito. El valor del % de cr. $Y_{\text{masculino}}$ puede obtenerse a partir de la secuenciación con bisulfito de dos muestras de plasma de varones adultos. El valor del % de cr. Y_{femenino} puede obtenerse a partir de la secuenciación con bisulfito de dos muestras de plasma de mujeres adultas no gestantes.

En otras realizaciones, el porcentaje de ADN fetal puede determinarse a partir de alelos específicos del feto en un autosoma. Como otro ejemplo, se pueden utilizar marcadores epigenéticos para determinar el porcentaje de ADN fetal. También pueden utilizarse otras formas de determinar el porcentaje de ADN fetal.

D. Método de utilización de la metilación para determinar el número de copias

- El genoma placentario está más hipometilado que el materno. Como se ha comentado anteriormente, la metilación del plasma de una mujer gestante depende de la concentración fraccional de ADN fetal derivado de la placenta en el plasma materno. Por lo tanto, mediante el análisis de la densidad de metilación de una región cromosómica, es posible detectar la diferencia en la contribución de los tejidos fetales al plasma materno. Por ejemplo, en una mujer gestante portadora de un feto trisómico (p. ej., que padezca trisomía 21 o trisomía 18 o trisomía 13), el feto aportaría una cantidad adicional del ADN del cromosoma trisómico al plasma materno en comparación con los cromosomas disomáticos. En esta situación, la densidad de metilación plasmática del cromosoma trisómico (o de cualquier región cromosómica que presente una amplificación) sería inferior a la de los cromosomas disomáticos. El grado de diferencia puede predecirse mediante cálculo matemático teniendo en cuenta la concentración fraccional de ADN fetal en la muestra de plasma. Cuanto mayor sea la concentración fraccional de ADN fetal en la muestra de plasma, mayor será la diferencia en la densidad de metilación entre los cromosomas trisómicos y disomáticos. Para las regiones que presentan una supresión, la densidad de metilación sería mayor.

Un ejemplo de delección es el síndrome de Turner, cuando un feto femenino solo tiene una copia del cromosoma X. En esta situación, para una mujer gestante portadora de un feto con síndrome de Turner, la densidad de metilación del cromosoma X en su ADN plasmático sería superior a la situación de la misma mujer gestante portadora de un feto femenino con el número normal de cromosomas X. En una realización de esta estrategia, se podría analizar primero el plasma materno para detectar la presencia o ausencia de secuencias del cromosoma Y (p. ej., mediante MPS o una técnica basada en PCR). Si las secuencias del cromosoma Y están presentes, entonces el feto puede clasificarse como masculino y no sería necesario el siguiente análisis. Por otra parte, si las secuencias del cromosoma Y están ausentes en el plasma materno, entonces el feto puede clasificarse como femenino. En esta situación, se puede analizar la densidad de metilación del cromosoma X en el plasma materno. Una densidad de metilación del cromosoma X superior a la normal indicaría que el feto tiene un alto riesgo de padecer el síndrome de Turner. Este enfoque también puede aplicarse a las demás aneuploidías cromosómicas sexuales. Por ejemplo, para un feto afectado por XYY, la densidad de metilación del cromosoma Y en el plasma materno sería inferior a la del feto XY normal con un nivel similar de ADN fetal en el plasma materno. Como otro ejemplo, para un feto con síndrome de Klinefelter (XXY), hay secuencias del cromosoma Y presentes en el plasma materno, pero la densidad de metilación del cromosoma X en el plasma materno será inferior a la de un feto XY normal con un nivel similar de ADN fetal en el plasma materno.

Del análisis anterior, la densidad de metilación plasmática para un cromosoma disómico ($PM_{\text{No aneu}}$) se puede calcular

como: $PM_{No\ aneu} = BASAL \times (1 - f) + NMP \times f$, donde *BASAL* es el nivel basal de metilación del ADN en plasma derivado de glóbulos sanguíneos y órganos internos, *NMP* es el nivel de metilación de la placenta y/es la concentración fraccional de ADN fetal en el plasma materno.

- 5 La densidad de metilación plasmática para un cromosoma trisómico (PM_{Aneu}) puede calcularse como: $PM_{Aneu} = BASAL \times (1 - f) + NMP \times f \times 1,5$, donde el 1,5 corresponde al número de copias NC y la adición de un cromosoma más supone un aumento del 50 %. La diferencia entre un cromosoma trisómico y uno disómico (PM_{Dif}) sería

$$PM_{Dif} = NMP \times f \times 0,5.$$

- 10 En una realización, la comparación de la densidad de metilación del cromosoma (o región cromosómica) potencialmente aneuploide con otro u otros cromosomas presuntamente no aneuploides o con la densidad de metilación total del genoma puede utilizarse para normalizar eficazmente la concentración de ADN fetal en la muestra de plasma. La comparación puede realizarse mediante el cálculo de un parámetro (p. ej., una relación o una diferencia)
- 15 entre las densidades de metilación de las dos regiones para obtener una densidad de metilación normalizada. La comparación puede eliminar una dependencia del nivel de metilación resultante (p. ej., determinado como parámetro a partir de las dos densidades de metilación).

- 20 Si la densidad de metilación del cromosoma potencialmente aneuploide no está normalizada con respecto a la densidad de metilación de otro u otros cromosomas, u otros parámetros que reflejen la concentración fraccional de ADN fetal, la concentración fraccionaria sería un factor importante que afectaría a la densidad de metilación en el plasma. Por ejemplo, la densidad de metilación plasmática del cromosoma 21 de una mujer gestante portadora de un feto con trisomía 21 y una concentración fraccional de ADN fetal del 10 % sería la misma que la de una mujer gestante portadora de un feto euploide y la concentración fraccional de ADN fetal es del 15 %, mientras que una densidad de metilación normalizada mostraría una diferencia.

- 25 En otra realización, la densidad de metilación del cromosoma potencialmente aneuploide puede normalizarse con respecto a la concentración fraccional de ADN fetal. Por ejemplo, se puede aplicar la siguiente ecuación para normalizar la densidad de metilación: $PM_{Normalizada} = + (BASAL - NMP) \times f$, donde $PM_{Normalizada}$ es la densidad de metilación normalizada con la concentración fraccional de ADN fetal en el plasma, $PM_{no\ normalizada}$ es la densidad de metilación medida, *BASAL* es la densidad basal de metilación de las células o tejidos sanguíneos maternos, *NMP* es la densidad de metilación en los tejidos placentarios, y/es la concentración fraccional de ADN fetal. Las densidades de metilación *BASAL* y *NMP* podrían basarse en valores de referencia previamente establecidos a partir de glóbulos sanguíneos maternos y tejidos placentarios obtenidos de gestaciones sanas. Se pueden utilizar diferentes métodos genéticos y epigenéticos para determinar la concentración fraccional de ADN fetal en la muestra de plasma, por ejemplo, mediante la medición del porcentaje de lecturas de secuencias del cromosoma Y utilizando secuenciación masiva paralela o PCR en ADN no convertido en bisulfito.

- 30 En una implementación, se puede comparar la densidad de metilación normalizada para un cromosoma potencialmente aneuploide con un grupo de referencia formado por mujeres gestantes portadoras de fetos euploides. Se puede determinar la media y la DT de la densidad de metilación normalizada del grupo de referencia. A continuación, la densidad de metilación normalizada del caso analizado puede expresarse como una puntuación z que indica el número de DT con respecto a la media del grupo de referencia:

- 35 45 Puntuación z = $\frac{PM_{Normalizada} - Media}{DT}$ donde $PM_{Normalizada}$ es la densidad de metilación normalizada para el caso analizado, la media es la media de la densidad de metilación normalizada de los casos de referencia y DT es la desviación típica de la densidad de metilación normalizada de los casos de referencia. Se puede utilizar un valor de corte, por ejemplo. una puntuación z <-3, para clasificar si un cromosoma está significativamente hipometilado y, por consiguiente, para determinar el estado de aneuploidía de la muestra.

- 40 50 En otra realización, la PM_{Dif} puede utilizarse como la densidad de metilación normalizada. En dicha realización, se puede deducir el *NMP*, p. ej., utilizando el método 1000. En algunas implementaciones, puede determinarse una densidad de metilación de referencia (que puede normalizarse mediante *f*) a partir de un nivel de metilación de una región no aneuploide. Por ejemplo, la Media podría determinarse a partir de una o varias regiones cromosómicas de la misma muestra. El valor de corte ser escalado por *f*, o simplemente fijarse en un nivel suficiente mientras exista una concentración mínima.

- 55 60 En consecuencia, la comparación del nivel de metilación de una región con un valor de corte puede realizarse de varias maneras. La comparación puede implicar una normalización (p. ej., como se ha descrito anteriormente), que puede realizarse de forma equivalente en el nivel de metilación o en el valor de corte, dependiendo de cómo se definan los valores. Por tanto, el hecho de que el nivel de metilación determinado de una región sea estadísticamente diferente de un nivel de referencia (determinado a partir de la misma muestra o de otras muestras) puede determinarse de diversas maneras.

- 65 65 El análisis anterior puede aplicarse al análisis de regiones cromosómicas, que puede incluir un cromosoma entero o

- partes del cromosoma, incluyendo subregiones contiguas o disjuntas de un cromosoma. En una realización, el cromosoma potencialmente aneuploide puede dividirse en varios bins. Los bins pueden ser del mismo tamaño o de tamaños diferentes. La densidad de metilación de cada bin puede normalizarse con respecto a la concentración fraccional de la muestra o a la densidad de metilación de uno o más cromosomas presuntamente no aneuploides o a la densidad de metilación global del genoma. La densidad de metilación normalizada de cada bin puede compararse entonces con un grupo de referencia para determinar si está significativamente hipometilado. A continuación, se puede determinar el porcentaje de bins que están significativamente hipometilados. Se puede utilizar un valor de corte, para ejemplos superiores al 5 %, 10 %, 15 %, 20 % o 30 % de los bins que están significativamente hipometilados para clasificar el estado de aneuploidía del caso.
- Cuando se busca una amplificación o una delección, se puede comparar la densidad de metilación con una densidad de metilación de referencia, que pueden ser específicas para una región concreta sometida a prueba. Cada región puede tener una densidad de metilación de referencia diferente, ya que la metilación puede variar de una región a otra, sobre todo en función del tamaño de las regiones (p. ej., las regiones más pequeñas mostrarán más variación).
- Como se ha mencionado anteriormente, una o más mujeres gestantes, cada una de ellas portadora de un feto euploide, puede utilizarse para definir el intervalo normal de la densidad de metilación para una región de interés o una diferencia en la densidad de metilación entre dos regiones cromosómicas. También puede determinarse un intervalo normal para el NMP (p. ej., por medición directa o deducido por el método 1000). En otras realizaciones, se puede utilizar una relación entre dos densidades de metilación, p. ej., se puede utilizar un cromosoma potencialmente aneuploide y un cromosoma no aneuploide para el análisis en lugar de su diferencia. Este enfoque de análisis de metilación puede combinarse con el enfoque de número de lecturas de secuencias (R. W. K. Chiu *et al.* 2008 *Proc Natl Acad Sci EE. UU.*;105:20458-20463) y enfoques que implican el análisis del tamaño del ADN plasmático (patente de EE. UU. n.º 2011/0276277) para determinar o confirmar una aneuploidía. El enfoque de número de lecturas de secuencias que se utiliza junto con el análisis de metilación puede realizarse utilizando secuenciación aleatoria (R. W. K. Chiu *et al.* 2008 *Proc Natl Acad Sci EE. UU.*;105:20458-20463; D. W. Bianchi D. W. *et al.* 2012 *Obstet Gynecol* 119:890-901) o secuenciación dirigida (A. B. Sparks *et al.* 2012 *Am J Obstet Gynecol* 206:319.e1-9; B. Zimmermann *et al.* 2012 *Prenat Diagn* 32:1233-1241; G. J. Liao *et al.* 2012 *PLoS One*; 7:e38154).
- El uso de BASAL puede tener en cuenta las variaciones del valor basal entre las muestras. Por ejemplo, una mujer podría tener niveles de metilación BASAL diferentes a los de otra mujer, pero la diferencia entre el BASAL y el NMP puede utilizarse entre muestras en tales situaciones. El valor de corte para las distintas regiones cromosómicas puede ser diferente, p. ej., cuando la densidad de metilación de una región del genoma difiere con respecto a otra región del genoma.
- Este enfoque puede generalizarse para detectar cualquier aberración cromosómica, incluida la delección y la amplificación, en el genoma fetal. Adicionalmente, la resolución de este análisis puede ajustarse al nivel deseado, por ejemplo, el genoma puede dividirse en bins de 10 Mb, 5 Mb, 2 Mb, 1 Mb, 500 kb, 100 kb. Por consiguiente, esta tecnología también puede utilizarse para detectar duplicaciones subcromosómicas o delecciones subcromosómicas.
- Esta tecnología permitiría obtener un cariotipo molecular fetal prenatal de forma no invasiva. Cuando se utiliza de esta manera, esta tecnología puede utilizarse junto con los métodos de pruebas prenatales no invasivas que se basan en el número de moléculas (A. Srinivasan *et al.* 2013 *Am J Hum Genet*;92:167-176; S. C. Y. Yu *et al.* 2013 *PLoS One* 8: e60968). En otras realizaciones, el tamaño de los bins no tiene por qué ser idéntico. Por ejemplo, el tamaño de los bins puede ajustarse para que cada bin contenga un número idéntico de dinucleótidos CpG. En este caso, el tamaño físico de los bins sería diferente.
- La ecuación puede reescribirse para aplicarla a distintos tipos de aberraciones cromosómicas como $PM_{Dif} = (BASAL - NMP) \times f \times 0,5 \times NC$. Aquí NC representa el número de cambios en el número de copias en la región afectada. NC igual a 1 para la ganancia de 1 copia de un cromosoma, 2 para la ganancia de 2 copias de un cromosoma y -1 para la pérdida de uno de los dos cromosomas homólogos (p. ej., para detectar el síndrome de Turner fetal en el que un feto femenino ha perdido uno de los cromosomas X, dando lugar a un cariotipo XO). No es necesario modificar esta ecuación cuando se cambia el tamaño de los bins. Sin embargo, se pueden deducir la sensibilidad y la especificidad cuando se utiliza un tamaño de bin más pequeño, ya que un número menor de dinucleótidos CpG (u otras combinaciones de nucleótidos que muestren una metilación diferencial entre el ADN fetal y el ADN materno) estarían presentes en bins más pequeños, dando lugar a una mayor variación estocástica en la medición de las densidades de metilación. En una realización, el número de lecturas necesarias puede determinarse analizando el coeficiente de variación de la densidad de metilación y el nivel de sensibilidad deseado.
- A fin de demostrar la viabilidad de este enfoque, hemos analizado las muestras de plasma de 9 mujeres gestantes. En cinco mujeres gestantes, cada una de ellas era portadora de un feto euploide y las otras cuatro eran portadoras de un feto con trisomía 21 (T21). Se seleccionaron al azar tres de las cinco gestaciones euploides para formar un grupo de referencia. Los dos casos de gestación euploide restantes (Eu1 y Eu2) y los cuatro casos T21 (T21-1, T21-2, T21-3 y T21-4) se analizaron utilizando este enfoque para comprobar un posible estado T21.
- El ADN plasmático se convirtió con bisulfito y se secuenció utilizando la plataforma HiSeq2000 de Illumina. En una realización, se calculó la densidad de metilación de los cromosomas individuales. A continuación, se determinó la

diferencia en la densidad de metilación entre el cromosoma 21 y la media de los otros 21 autosomas para obtener una densidad de metilación normalizada (Tabla 1). Se utilizaron la media y la DT del grupo de referencia para calcular la puntuación z de los seis casos de prueba.

	Eu1	Eu2	T21-1	T21-2	T21-3	T21-4
Puntuación z para PM _{Dif} entre el cr. 21 y otros autosomas	-1,48	1,09	-4,46	-5,30	-8,06	-5,69

- 5 **Tabla 1:** Mediante un valor de corte de <-3 para la puntuación z para clasificar una muestra como T21, la clasificación de todos los casos euploides y T21 era correcta.

En otra realización, el genoma se dividió en bins de 1 Mb y se determinó la densidad de metilación de cada bin de 1 Mb. La densidad de metilación de todos los bins del cromosoma potencialmente aneuploide puede normalizarse con 10 la densidad de metilación media de todos los bins situados en los cromosomas presuntamente no aneuploides. En una implementación, para cada bin, se puede calcular la diferencia en la densidad de metilación con respecto a la mediana de los bins no aneuploides. La puntuación z puede calcularse para estos valores mediante los valores de la media y la DT del grupo de referencia. Se puede determinar el porcentaje de bins que muestran hipometilación (Tabla 15 2) y compararlo con un porcentaje de corte.

	Eu1	Eu2	T21-1	T21-2	T21-3	T21-4
Porcentaje de bins en el cr. 21 tienen una puntuación z de PM _{Dif} <-3	0 %	0 %	33,3 %	58,3 %	19,4 %	52,8 %

Tabla 2: Mediante el 5 % como valor de corte para los bins con un número significativamente mayor de hipometilados en el cromosoma 21, todos los casos se clasificaron correctamente para el estado T21.

Este enfoque basado en la metilación del ADN para detectar aberraciones cromosómicas o subcromosómicas fetales 20 puede utilizarse junto con los basados en el número de moléculas como, por ejemplo, mediante secuenciación (R. W. K. Chiu *et al.* 2008 *Proc Natl Acad Sci* EE. UU.; 105: 20458-20463) o PCR digital (Y. M. D. Uo *et al.* 2007 *Proc Natl Acad Sci* EE. UU.; 104: 13116-13121), o el dimensionamiento de moléculas de ADN (publicación de patente de EE.UU. n.º 2011/0276277). Dicha combinación (p. ej., metilación del ADN más número molecular, o metilación del ADN más 25 determinación del tamaño, o metilación del ADN más número molecular más determinación del tamaño) tendría un efecto sinérgico que sería ventajoso en un contexto clínico, p. ej., mejorando la sensibilidad y/o la especificidad. Por ejemplo, el número de moléculas de ADN que habría que analizar, p. ej., mediante secuenciación, puede reducirse sin afectar negativamente a la precisión del diagnóstico. Esta característica permitiría realizar dichas pruebas de forma más económica. Como otro ejemplo, para un número determinado de moléculas de ADN analizadas, un enfoque combinado permitiría detectar aberraciones cromosómicas o subcromosómicas fetales con una concentración 30 fraccional más baja de ADN fetal.

La FIG. 13 es un diagrama de flujo de un método 1300 para detectar una anomalía cromosómica a partir de una muestra biológica de un organismo. La muestra biológica incluye ADN extracelular que comprende una mezcla de ADN extracelular procedente de un primer tejido y de un segundo tejido. El primer tejido puede proceder de un feto o 35 de un tumor y el segundo tejido puede proceder de una mujer gestante o de un paciente.

En el bloque 1310, se analizan una pluralidad de moléculas de ADN de la muestra biológica. El análisis de una molécula de ADN puede incluir la determinación de una ubicación de la molécula de ADN en un genoma del organismo y la determinación de si la molécula de ADN está metilada en uno o más sitios. El análisis puede realizarse recibiendo 40 lecturas de secuencias de una secuenciación consciente de la metilación, y así el análisis puede realizarse solo sobre datos obtenidos previamente del ADN. En otras realizaciones, el análisis puede incluir la secuenciación real u otras etapas activas de la obtención de los datos.

La determinación de una ubicación puede incluir la cartografía de las moléculas de ADN (p. ej., a través de lecturas 45 de secuencias) a las partes respectivas del genoma humano, p. ej., a regiones específicas. En una implementación, si una lectura no corresponde a una región de interés, entonces la lectura puede ser ignorada.

En el bloque 1320, se determina un número respectivo de moléculas de ADN que están metiladas en el sitio para cada uno de una pluralidad de sitios. En una realización, los sitios son sitios CpG, y pueden ser solo ciertos sitios CpG, 50 seleccionados de acuerdo con uno o varios de los criterios mencionados en el presente documento. El número de ADN metilado equivale a determinar el número de ADN no metilado una vez realizada la normalización utilizando el número total de moléculas de ADN analizadas en un sitio concreto, p. ej., un número total de lecturas de secuencias.

En el bloque 1330, se calcula un primer nivel de metilación de una primera región cromosómica basándose en los 55 números respectivos de moléculas de ADN metiladas en sitios dentro de la primera región cromosómica. La primera región cromosómica puede ser de cualquier tamaño, p. ej., los tamaños mencionados anteriormente. El nivel de metilación puede dar cuenta de un número total de moléculas de ADN alineadas con la primera región cromosómica, p. ej., como parte de un procedimiento de normalización.

La primera región cromosómica puede ser de cualquier tamaño (p. ej., un cromosoma entero) y puede comprender subregiones disjuntas, es decir, subregiones están separadas entre sí. Se pueden determinar los niveles de metilación de cada subregión y combinarlos, p. ej., como un promedio o una mediana, para determinar el nivel de metilación de la primera región cromosómica.

- 5 En el bloque 1340, el primer nivel de metilación se compara con un valor de corte. El valor de corte puede ser un nivel de metilación de referencia o estar relacionado con un nivel de metilación de referencia (p. ej., una distancia especificada de un nivel normal). El valor de corte puede determinarse a partir de otras mujeres gestantes portadoras de fetos sin anomalías cromosómicas en la primera región cromosómica, de muestras de individuos sin cáncer o de locus del organismo que se sabe que no están asociados a una aneuploidía (es decir, regiones que son disómicas).
- 10 En una realización, el valor de corte puede definirse como una diferencia con respecto a un nivel basal de metilación de $(BASAL - NMP) \times f \times 0,5 \times NC$, donde BASAL es el valor basal de la mujer (o un promedio o una mediana de otros sujetos), f es la concentración fraccional de ADN extracelular procedente del primer tejido y NC es un número de copias analizado. NC es un ejemplo de factor de escala correspondiente a un tipo de anomalía (deleción o duplicación).
- 15 Se puede utilizar un valor de corte para un NC de 1 para comprobar inicialmente todas las amplificaciones y, a continuación, se pueden utilizar otros puntos de corte para determinar el grado de amplificación. El valor de corte puede basarse en una concentración fraccional de ADN extracelular procedente del primer tejido para determinar el nivel esperado de metilación de un locus, p. ej., si no hay aberración del número de copias.
- 20 En el bloque 1350, se determina una clasificación de una anomalía para la primera región cromosómica basándose en la comparación. Una diferencia estadísticamente significativa en los niveles puede indicar un mayor riesgo de que el feto presente una anomalía cromosómica. En diversas realizaciones, la anomalía cromosómica puede ser la trisomía 21, trisomía 18, trisomía 13, el síndrome de Turner o el síndrome de Klinefelter. Otros ejemplos son una delección subcromosómica, duplicación subcromosómica o el síndrome de DiGeorge.
- 25

V. DETERMINACIÓN DE MARCADORES

- 30 Como se ha indicado anteriormente, ciertas partes del genoma fetal se metilan de forma diferente al genoma materno. Estas diferencias pueden ser comunes en todas las gestaciones. Las regiones de metilación diferente pueden utilizarse para identificar fragmentos de ADN procedentes del feto.

A. Método para determinar los DMR a partir de tejido placentario y tejido materno

- 35 La placenta tiene distintivos de metilación específicos de cada tejido. Se han desarrollado marcadores de metilación del ADN específicos del feto para la detección en plasma materno y para aplicaciones no invasivas de diagnóstico prenatal basadas en locus que están metilados de forma diferencial entre los tejidos placentarios y los glóbulos sanguíneos maternos (S. S. C. Chim *et al.* 2008 *Clin Chem*; 54: 500-511; E. A. Papageorgiou *et al.*, 2009 *Am J Pathol*; 174: 1609-1618; y T. Chu *et al.* 2011 *PLoS One*; 6: e14723). Se proporcionan realizaciones para la extracción de tales regiones metiladas diferencialmente (DMR) sobre una base de todo el genoma.

- 40 La FIG. 14 es un diagrama de flujo de un método 1400 para identificar marcadores de metilación comparando un perfil de metilación placentario con un perfil de metilación materno (p. ej., determinado a partir de glóbulos sanguíneos) de acuerdo con realizaciones de la presente invención. El método 1400 también puede utilizarse para determinar marcadores de un tumor comparando un perfil de metilación tumoral con un perfil de metilación correspondiente a tejido sano.

- 45 En el bloque 1410, se obtiene un metiloma placentario y un metiloma sanguíneo. El metiloma placentario puede determinarse a partir de una muestra de placenta, p. ej., MVC o una placenta a término. Debe entenderse que el metiloma puede incluir densidades de metilación de solo una parte de un genoma.

- 50 En el bloque 1420, se identifica una región que incluye un número determinado de sitios (p. ej., 5 sitios CpG) y para la que se ha obtenido un número suficiente de lecturas. En una realización, la identificación comenzó desde un extremo de cada cromosoma para localizar la primera región de 500 pb que contuviera al menos cinco sitios CpG cualificados. Un sitio CpG puede considerarse cualificado si está cubierto por al menos cinco lecturas de secuencia.

- 55 En el bloque 1430, se calcula un índice de metilación placentaria y un índice de metilación sanguínea para cada sitio. Por ejemplo, el índice de metilación se calculó individualmente para todos los sitios CpG cualificados dentro de cada región de 500 pb.

- 60 En el bloque 1440, se compararon los índices de metilación entre los glóbulos sanguíneos maternos y la muestra de placenta para determinar si los conjuntos de índices eran diferentes entre sí. Por ejemplo, se compararon los índices de metilación entre las células de la sangre materna y la MVC o la placenta a término utilizando, por ejemplo, la prueba de Mann-Whitney. Un valor de P de, por ejemplo, $\leq 0,01$ se consideró estadísticamente significativamente diferente, aunque pueden utilizarse otros valores, donde un número menor reduciría las regiones de falsos positivos.

En una realización, si el número de sitios CpG calificados era inferior a cinco o la prueba de Mann-Whitney no era significativa, la región de 500 pb desplazada 100 pb en dirección 3'. La región siguió desplazándose en dirección 3' hasta que la prueba de Mann-Whitney se hizo significativa para una región de 500 pb. A continuación, se consideraría la siguiente región de 500 pb. Si la siguiente región presentaba significación estadística mediante la prueba de Mann-Whitney, se añadía a la región actual siempre que la región contigua combinada no superara los 1000 pb.

En el bloque 1450, las regiones adyacentes que eran significativamente diferentes desde el punto de vista estadístico (p. ej., mediante la prueba de Mann-Whitney) pueden fusionarse. Obsérvese la diferencia es entre los índices de metilación de las dos muestras. En una realización, si las regiones adyacentes se encuentran a una distancia determinada (p. ej., 1000 pb) entre sí y si muestran un perfil de metilación similar, se fusionarán. En una implementación, la similitud del perfil de metilación entre regiones adyacentes puede definirse utilizando cualquiera de las siguientes condiciones: (1) que muestre la misma tendencia en el tejido placentario con referencia a los glóbulos sanguíneos maternos, p. ej., ambas regiones estaban más metiladas en los tejidos placentarios que en los glóbulos sanguíneos; (2) que tenga diferencias en las densidades de metilación inferiores al 10 % para las regiones adyacentes en el tejido placentario; y (3) que tenga diferencias en las densidades de metilación inferiores al 10 % para las regiones adyacentes en los glóbulos sanguíneos maternos.

En el bloque 1460, se calcularon las densidades de metilación del metiloma sanguíneo del ADN de los glóbulos sanguíneos maternos y de la muestra placentaria (p. ej., MVC o tejido placentario a término) en las regiones. Las densidades de metilación pueden determinarse como se describe en el presente documento.

En el bloque 1470, se determinan los DMR putativos en los que la densidad total de metilación placentaria y la densidad total de metilación sanguínea para todos los sitios de la región son estadísticamente significativamente diferentes. En una realización, todos los sitios CpG calificados dentro de una región fusionada se someten a una prueba χ^2 . La prueba χ^2 evalúo si el número de citosinas metiladas como proporción de las citosinas metiladas y no metiladas entre todos los sitios CpG calificados dentro de la región fusionada era estadísticamente significativamente diferente entre los glóbulos sanguíneos maternos y el tejido placentario. En una implementación, para la prueba χ^2 , un varlor de $P \leq 0,01$ puede considerarse estadísticamente significativamente diferente. Los segmentos fusionados que mostraron significación mediante la prueba χ^2 se consideraron DMR putativos.

En el bloque 1480, se identificaron los locus en los que las densidades de metilación del ADN de los glóbulos sanguíneos maternos estaban por encima de un umbral alto o por debajo de un umbral bajo. En una realización, se identificaron locus en los que las densidades de metilación del ADN de los glóbulos sanguíneos maternos eran $\leq 20\%$ o $\geq 80\%$. En otras realizaciones, pueden utilizarse humores corporales distintos de la sangre materna, incluidos, pero sin limitarse a la saliva, líquido de lavado uterino o cervical del tracto genital femenino, lágrimas, sudor, saliva y orina.

Una clave para el éxito del desarrollo de marcadores de metilación del ADN que sean específicos del feto en el plasma materno puede ser que el estado de metilación de los glóbulos sanguíneos maternos sea lo más altamente metilado o lo menos metilado posible. Esto puede reducir (p. ej., minimizar) la posibilidad de que las moléculas de ADN materno interfieran con el análisis de las moléculas de ADN fetal derivadas de la placenta que muestran un perfil de metilación opuesto. Por tanto, en una realización, los DMR candidatos se seleccionaron mediante filtrado adicional. Los locus hipometilados candidatos eran aquellos que mostraban densidades de metilación $\leq 20\%$ en los glóbulos sanguíneos maternos y con densidades de metilación de al menos un 20 % superiores en los tejidos placentarios. Los locus hipermetilados candidatos eran aquellos que mostraban densidades de metilación $\geq 80\%$ en los glóbulos sanguíneos maternos y con densidades de metilación inferiores en al menos un 20 % en los tejidos placentarios. Pueden utilizarse otros porcentajes.

En el bloque 1490, a continuación, se identificaron DMR entre el subconjunto de locus en los que las densidades de metilación placentaria son significativamente diferentes de las densidades de metilación sanguínea comparando la diferencia con un umbral. En una realización, el umbral es del 20 %, por lo que las densidades de metilación diferían en al menos un 20 % de las densidades de metilación de los glóbulos sanguíneos maternos. En consecuencia, se puede calcular una diferencia entre las densidades de metilación placentaria y las densidades de metilación sanguínea en cada locus identificado. La diferencia puede ser una simple resta. En otras realizaciones, se pueden utilizar factores de escala y otras funciones para determinar la diferencia (p. ej., la diferencia puede ser el resultado de una función aplicada a la resta simple).

En una implementación, mediante este método, se identificaron 11.729 locus hipermetilados y 239.747 hipometilados a partir de la muestra de placenta del primer trimestre. Los 100 principales locus hipermetilados se enumeran en la tabla S2A del anexo. Los 100 principales locus hipometilados se enumeran en la tabla S2B del anexo. En las tablas S2A y S2B se enumeran el cromosoma, la ubicación inicial y final, el tamaño de la región, la densidad de metilación en la sangre materna, la densidad de metilación en la muestra de placenta, los valores de P (que son todos muy pequeños) y la diferencia de metilación. Las ubicaciones corresponden al genoma de referencia hg18, que puede encontrarse en hgdownload.soe.ucsc.edu/goldenPath/hg18/chromosomes.

Se identificaron 11.920 locus hipermetilados y 204.768 hipometilados a partir de la muestra de placenta del tercer trimestre. Los 100 principales locus hipermetilados para el 3.^{er} trimestre se enumeran en la tabla S2C, y los 100

principales locus hipometilados se enumeran en la tabla S2D. Para validar nuestra lista de candidatos del primer trimestre, se utilizaron treinta y tres locus cuya metilación diferencial entre los glóbulos sanguíneos maternos y los tejidos placentarios del primer trimestre se había descrito anteriormente. El 79 % de los 33 locus se identificaron como DMR mediante nuestro algoritmo.

FIG. 15A es una tabla 1500 que muestra un rendimiento del algoritmo de identificación de DMR utilizando datos del primer trimestre con referencia a 33 marcadores del primer trimestre previamente reportados. En la tabla, "a" indica que los locus 1 a 15 se describieron previamente en (R. W. K. Chiu *et al.* 2007 *Am J Pathol*; 170:941-950 y S. S. C. Chim *et al.* 2008 *Clin Chem*; 54:500-511); los locus 16 a 23 se describieron previamente en (K. C. Yuen, tesis 2007, Universidad China de Hong Kong, Hong Kong); y los locus 24 a 33 se describieron previamente en (E. A. Papageorgiou *et al.* 2009 *Am J Pathol*; 174:1609-1618). "b" indica que estos datos proceden de las publicaciones mencionadas, "c" indica que las densidades de metilación de los glóbulos sanguíneos maternos y de la muestra de vellosidades coriónicas y sus diferencias se observaron a partir de los datos de secuenciación generados en el presente estudio, pero basándose en las coordenadas genómicas proporcionadas por los estudios originales, "d" indica que los datos sobre los locus identificados utilizando las realizaciones del método 1400 sobre los datos de secuenciación con bisulfito sin tomar como referencia las publicaciones citadas anteriormente de Chiu *et al.* (2007), Chim *et al.*, (2008), Yuen (2007) y Papageorgiou *et al.*, (2009). El alcance de los locus incluía las regiones genómicas de las que se había informado anteriormente, pero en general abarcaba regiones más extensas, "e" indica que un DMR candidato se clasificó como verdadero positivo (VP) o falso negativo (FN) en función del requisito de observar una diferencia >0,20 entre las densidades de metilación de las coordenadas genómicas correspondientes de los DMR en los glóbulos sanguíneos maternos y la muestra de vellosidades coriónicas.

La FIG. 15B es una tabla 1550 que muestra un rendimiento del algoritmo de identificación de DMR utilizando datos del tercer trimestre y comparándolos con la muestra de placenta obtenida en el parto, "a" indica que se utilizó la misma lista de 33 locus que la descrita en la FIG. 17A, "b" indica que, dado que los 33 locus se identificaron previamente a partir de muestras de la primera etapa de la gestación, podrían no ser aplicables a los datos del tercer trimestre. Por consiguiente, se revisaron los datos de secuenciación con bisulfito generados en el presente estudio sobre el tejido de placenta a término a partir de las coordenadas genómicas proporcionadas por los estudios originales. Se utilizó una diferencia >0,20 en las densidades de metilación entre el tejido de la célula sanguínea materna y el de la placenta a término para determinar si los locus eran realmente verdaderos DMR en el tercer trimestre, "c" indica que los datos sobre los locus se identificaron utilizando el método 1400 sobre los datos de secuenciación con bisulfito sin tomar como referencia las publicaciones citadas anteriormente de Chiu *et al.*, (2007), Chim *et al.*, (2008), Yuen (2007) y Papageorgiou *et al.*, (2009). El alcance de los locus incluía las regiones genómicas de las que se había informado anteriormente, pero en general abarcaba regiones más extensas, "d" indica que los DMR candidatos que contenían locus calificados como metilados diferencialmente en el tercer trimestre se clasificaron como verdaderos positivos (VP) o falsos negativos (FN) basándose en el requisito de observar una diferencia >0,20 entre las densidades de metilación de las coordenadas genómicas correspondientes de los DMR en glóbulos sanguíneos maternos y tejido placentario a término. Para los locus que no se calificaron como metilados diferencialmente en el tercer trimestre, su ausencia en la lista de DMR o la presencia de un DMR que contuviera los locus, pero mostrara una diferencia de metilación <0,20 se consideraron DMR verdaderos negativos (VN).

B. DMR a partir de los datos de secuenciación del plasma materno

Debería ser posible identificar los DMR de tejido placentario directamente a partir de los datos de secuenciación con bisulfito del ADN del plasma materno, siempre que la concentración fraccional de ADN fetal de la muestra fuera también conocida. Es posible porque la placenta es la fuente predominante de ADN fetal en el plasma materno (S. S. C. Chim *et al.* 2005 *Proc Natl Acad Sci EE.UU.* 102, 14753-14758) y demostramos en este estudio que el estado de metilación del ADN específico del feto en el plasma materno se correlacionaba con el metiloma placentario.

Por lo tanto, los aspectos del método 1400 pueden implementarse utilizando un metiloma plasmático para determinar un metiloma placentario deducido en lugar de utilizar una muestra placentaria. Por tanto, el método 1000 y el método 1400 pueden combinarse para determinar los DMR. El método 1000 puede utilizarse para determinar los valores previstos para el perfil de metilación de la placenta y utilizarlos en el método 1400. Para este análisis, el ejemplo también se centra en los locus que estaban ≤20 % o ≥80 % metilados en los glóbulos sanguíneos maternos.

En una implementación, para deducir los locus que estaban hipermetilados en los tejidos placentarios con respecto a los glóbulos sanguíneos maternos, clasificamos los locus que mostraban ≤20 % de metilación en los glóbulos sanguíneos maternos, y ≥60 % de metilación de acuerdo con el valor predicho con una diferencia de al menos el 50 % entre la densidad de metilación de los glóbulos sanguíneos y el valor predicho. Para deducir los locus que estaban hipometilados en los tejidos placentarios con respecto a los glóbulos sanguíneos maternos, seleccionamos los locus que mostraban ≥80 % de metilación en los glóbulos sanguíneos maternos, y ≤40 % de metilación de acuerdo con el valor predicho con una diferencia de al menos el 50 % entre la densidad de metilación de los glóbulos sanguíneos y el valor predicho.

La FIG. 16 es una tabla 1600 que muestra los números de locus predichos a ser hipermetilados o hipometilados basados en el análisis directo de los datos de secuenciación de plasma materno tratado con bisulfito. "n. a." significa

no aplicable, "a" indica que la búsqueda de locus hipermetilados se inició a partir de la lista de locus que mostraban densidades de metilación <20 % en los glóbulos sanguíneos maternos, "b" indica que la búsqueda de locus hipometilados se inició a partir de la lista de locus que mostraban densidades de metilación >80 % en los glóbulos sanguíneos maternos, "c" indica que los datos de secuenciación con bisulfito de la muestra de vellosidades coriónicas

5 se utilizaron para verificar los datos de plasma materno del primer trimestre, y el tejido de placenta a término se utilizó para verificar los datos de plasma materno del tercer trimestre.

Como se muestra en la Tabla 1600, la mayoría de los locus deducidos de forma no invasiva mostraban el patrón de metilación esperado en los tejidos y coincidían con los DMR extraídos de los datos tisulares y presentados en la 10 sección anterior. El apéndice enumera los DMR identificados a partir del plasma. La Tabla S3A enumera los 100 locus más hipermetilados deducidos a partir de los datos de secuenciación del plasma materno del primer trimestre con bisulfito. La Tabla S3B enumera los 100 locus más hipometilados deducidos a partir de los datos de secuenciación del plasma materno del primer trimestre con bisulfito. La Tabla S3C enumera los 100 locus más hipermetilados deducidos de los datos de secuenciación con bisulfito del plasma materno del tercer trimestre. La Tabla S3D enumera los 100 15 locus más hipometilados deducidos de los datos de secuenciación con bisulfito del plasma materno del tercer trimestre.

C. Variación gestacional en los metilomas placentarios y fetales

La proporción global de CpG metilados en la MVC fue del 55 %, mientras que en la placenta a término fue del 59 % 20 (Tabla 100 de la FIG. 1). Se identificaron más DMR hipometilados en la MVC que en la placenta a término, mientras que el número de DMR hipermetilados fue similar en ambos tejidos. Por tanto, era evidente que la MVC estaba más hipometilada que la placenta a término. Esta tendencia gestacional también se observó en los datos del plasma materno. La proporción de CpG metilados entre las lecturas específicas del feto fue del 47,0 % en el plasma materno del primer trimestre, pero del 53,3 % en el plasma materno del tercer trimestre. El número de locus hipermetilados 25 validados fue similar en las muestras de plasma materno del primer (1457 locus) y tercer trimestre (1279 locus), pero hubo sustancialmente más locus hipometilados en las muestras del primer trimestre (21.812 locus) que en las del tercer trimestre (12.677 locus) (tabla 1600 de la FIG. 16).

D. Uso de marcadores

30 Los marcadores diferencialmente metilados, o DMR, son útiles en varios aspectos. La presencia de estos marcadores en el plasma materno indica y confirma la presencia de ADN fetal o placentario. Esta confirmación puede utilizarse como control de calidad de las pruebas prenatales no invasivas. Los DMR pueden servir como marcadores genéricos del ADN fetal en el plasma materno y presentan ventajas frente a los marcadores que se basan en diferencias genotípicas entre la madre y el feto, tales como los marcadores basados en polimorfismos o los basados en el cromosoma Y. Los DMR son marcadores fetales genéricos útiles para todas las gestaciones. Los marcadores basados en polimorfismos solo son aplicables al subconjunto de gestaciones en las que el feto ha heredado el marcador de su 35 padre y en las que la madre no posee este marcador en su genoma. Adicionalmente, se podría medir la concentración de ADN fetal en una muestra de plasma materno cuantificando las moléculas de ADN procedentes de esos DMR. 40 Conociendo el perfil de DMR esperado para gestaciones normales, las complicaciones asociadas a la gestación, en particular las que implican cambios en el tejido placentario, se podrían detectar observando una desviación en el perfil de DMR o de metilación del plasma materno con respecto al esperado para gestaciones normales. Las complicaciones asociadas a la gestación que implican cambios en el tejido placentario incluyen, entre otras, las aneuploidías cromosómicas fetales. Algunos ejemplos son la trisomía 21, preeclampsia, retraso del crecimiento intrauterino y parto 45 prematuro.

E. Kits con marcadores

50 Las realizaciones pueden proporcionar composiciones y kits para poner en práctica los métodos descritos en el presente documento y otros métodos aplicables. Los kits pueden utilizarse para realizar ensayos de análisis del ADN fetal, p. ej., ADN fetal extracelular en el plasma materno. En una realización, un kit puede incluir al menos un oligonucleótido útil para la hibridación específica con uno o más locus identificados en el presente documento. Un kit también puede incluir al menos un oligonucleótido útil para la hibridación específica con uno o más locus de referencia. En una realización, se miden los marcadores de hipermetilación placentaria. El locus de prueba puede ser el ADN 55 metilado del plasma materno y el locus de referencia puede ser el ADN metilado del plasma materno. Se podría componer un kit similar para analizar el ADN tumoral en plasma.

En algunos casos, los kits pueden incluir al menos dos cebadores oligonucleotídicos que pueden utilizarse en la amplificación de al menos una sección de un locus diana (p. ej., un locus del apéndice) y un locus de referencia. En 60 lugar o además de los cebadores, un kit puede incluir sondas marcadas para detectar un fragmento de ADN correspondiente a un locus diana y a un locus de referencia. En diversas realizaciones, uno o más oligonucleótidos del kit corresponden a un locus de las tablas del apéndice. Habitualmente, los kits también proporcionan manuales de instrucciones para guiar a los usuarios en el análisis de las muestras de prueba y la evaluación del estado de la fisiología o patología en un sujeto de prueba.

65 En diversas realizaciones, se proporciona un kit para analizar el ADN fetal en una muestra biológica que contiene una

mezcla de ADN fetal y ADN de una mujer gestante de un feto. El kit puede comprender uno o más oligonucleótidos para hibridar específicamente con al menos una sección de una región genómica enumerada en las tablas S2A, S2B, S2C, S2D, S3A, S3B, S3C y S3D. Por tanto, se puede utilizar cualquier número de oligonucleótidos de todas las tablas o solo de una tabla. Los oligonucleótidos pueden actuar como cebadores, y pueden organizarse como pares de cebadores, donde un par corresponde a una región concreta de las tablas.

5

VI. RELACIÓN ENTRE EL TAMAÑO Y LA DENSIDAD DE METILACIÓN

Se sabe que las moléculas de ADN plasmático existen en circulación en forma de moléculas cortas, siendo la mayoría 10 de las moléculas de aproximadamente 160 pb de longitud (YMD Lo *et al.* 2010 *Sci Transl Med*; 2: 61ra91, Y. W. Zheng *et al.* 2012 *Clin Chem*; 58: 549-558). De manera interesante, nuestros datos revelaron una relación entre el estado de metilación y el tamaño de las moléculas de ADN plasmático. Por tanto, la longitud del fragmento de ADN plasmático 15 está vinculada al nivel de metilación del ADN. Los perfiles de tamaño característico de las moléculas de ADN plasmático indican que la mayoría están asociadas a mononucleosomas, posiblemente derivados de la degradación enzimática durante la apoptosis.

15

El ADN circulante está fragmentado por naturaleza. En concreto, el ADN fetal circulante es más corto que el ADN 20 derivado de la madre en muestras de plasma materno (K. C. A. Chan *et al.* 2004 *Clin Chem*; 50: 88-92). La alineación de ambos extremos permite analizar el tamaño del ADN tratado con bisulfito, se podría evaluar directamente si existe 25 alguna correlación entre el tamaño de las moléculas de ADN plasmático y sus respectivos niveles de metilación. Esto se estudió en el plasma materno y en una muestra de plasma de control de una mujer adulta no gestante.

25

Para analizar cada muestra de este estudio se utilizó la secuenciación de ambos extremos (que incluye la secuenciación de una molécula completa) de ambos extremos de cada molécula de ADN. Alineando el par de secuencias finales de cada molécula de ADN con el genoma humano de referencia y anotando las coordenadas 30 genómicas de los extremos de las lecturas secuenciadas, se pueden determinar las longitudes de las moléculas de ADN secuenciadas. Las moléculas de ADN plasmático se fragmentan de forma natural en pequeñas moléculas y las bibliotecas de secuenciación del ADN plasmático se preparan normalmente sin ninguna etapa de fragmentación. Por consiguiente, las longitudes deducidas por la secuenciación representaban los tamaños de las moléculas de ADN plasmáticas originales.

35

En un estudio anterior, determinamos los perfiles de tamaño de las moléculas de ADN fetal y materno en el plasma materno (Y. M. D. Lo *et al.* 2010 *Sci Transl Med*; 2: 61ra91). Demostramos que las moléculas de ADN plasmático tenían tamaños parecidos a los mononucleosomas y que las moléculas de ADN fetal eran más cortas que las maternas. En este estudio, hemos determinado la relación del estado de metilación de las moléculas de ADN 40 plasmático con su tamaño.

A. Resultados

40

La FIG. 17A es un gráfico 1700 que muestra la distribución de tamaños del plasma materno, plasma de control de mujeres no gestantes, ADN placentario y de sangre periférica. Para la muestra materna y el plasma de control de mujer no gestante, las dos muestras de plasma tratadas con bisulfito mostraron la misma distribución de tamaños 45 característica de la que se había informado anteriormente (Y. M. D. Lo *et al.* 2010 *Sci Transl Med*; 2: 61ra91) con las secuencias totales más abundantes de 166-167 pb de longitud y una periodicidad de 10 pb de las moléculas de ADN inferiores a 143 pb.

45

La FIG. 17B es un gráfico 1750 de la distribución de tamaños y el perfil de metilación del plasma materno, plasma de control de mujer adulta, tejido placentario y sangre de control de mujer adulta. Para moléculas de ADN del mismo tamaño y que contengan al menos un sitio CpG, se calculó su densidad de metilación media. A continuación, trazamos 50 la relación entre los tamaños de las moléculas de ADN y sus densidades de metilación. Específicamente, se determinó la densidad media de metilación para cada longitud de fragmento, desde 50 pb hasta 180 pb, para las lecturas secuenciadas que cubrían al menos 1 sitio CpG. De manera interesante, la densidad de metilación aumentó con el tamaño del ADN plasmático y alcanzó un máximo en torno a 166-167 pb. Este patrón, sin embargo, no se observó en las muestras de ADN de la placenta y de la sangre de control, que se fragmentaron mediante un sistema ultrasónico.

55

La FIG. 18 muestra gráficas de las densidades de metilación y el tamaño de las moléculas de ADN plasmático. La FIG. 18A es un gráfico 1800 para el plasma materno del primer trimestre. La FIG. 18B es un gráfico 1850 para el plasma materno del tercer trimestre. Los datos de todas las lecturas secuenciadas que cubrían al menos un sitio CpG 60 están representados por la curva azul 1805. Los datos de las lecturas que también contenían un alelo SNP específico del feto están representados por la curva roja 1810. Los datos de las lecturas que también contenían un alelo SNP específico de la madre están representados por la curva verde 1815.

65

Se consideró que las lecturas que contenían un alelo SNP específico del feto procedían de moléculas de ADN fetal. Se consideró que las lecturas que contenían un alelo SNP específico de la madre procedían de moléculas de ADN materno. En general, las moléculas de ADN con alta densidad de metilación eran más largas. Esta tendencia se observó tanto en las moléculas de ADN fetal como en las de ADN materno en el primer y el tercer trimestre. Los

tamaños globales de las moléculas de ADN fetal eran más cortos que los de las moléculas maternas, como se había señalado anteriormente.

- 5 La FIG. 19A muestra un gráfico 1900 de las densidades de metilación y los tamaños de las lecturas secuenciadas para una mujer adulta no gestante. La muestra de ADN plasmático de la mujer adulta no gestante también mostró la misma relación entre los tamaños y el estado de metilación de las moléculas de ADN. Por otra parte, las muestras de ADN genómico se fragmentaron mediante una etapa de ultrasonidos antes del análisis MPS. Como se muestra en la gráfica 1900, los datos de las muestras de glóbulos sanguíneos y tejido placentario no revelaron la misma tendencia. Dado que la fragmentación de las células es artificial, uno esperaría que no hubiera relación de tamaño y densidad.
- 10 Dado que las moléculas de ADN fragmentadas de forma natural en el plasma sí muestran una dependencia del tamaño, cabe suponer que las menores densidades de metilación hacen más probable que las moléculas se rompan en fragmentos más pequeños.
- 15 La FIG. 19B es un gráfico 1950 que muestra la distribución del tamaño y el perfil de metilación de las moléculas de ADN específicas del feto y específicas de la madre en el plasma materno. Las moléculas de ADN plasmático específicas del feto y de la madre también mostraron la misma correlación entre el tamaño del fragmento y el nivel de metilación. Tanto la longitud del fragmento de ADN extracelular derivado de la placenta como el ADN extracelular circulante materno aumentaron con el nivel de metilación. Asimismo, la distribución de su estado de metilación no se solapaba entre sí, lo que indica que el fenómeno existe independientemente de la longitud del fragmento original de
- 20 las fuentes de moléculas de ADN circulantes.

B. Método

- 25 En consecuencia, se puede utilizar una distribución de tamaños para estimar el porcentaje total de metilación de una muestra de plasma. Esta medición de la metilación puede seguirse durante la gestación, durante el seguimiento del cáncer o durante el tratamiento mediante la medición en serie de las distribuciones de tamaño del ADN plasmático de acuerdo con la relación mostrada en las FIGS. 18A y 18B. La medición de la metilación también puede utilizarse para buscar una mayor o menor liberación de ADN de un órgano o un tejido de interés. Por ejemplo, se pueden buscar específicamente distintivos de metilación del ADN propias de un órgano concreto (p. ej., el hígado) y medir las concentraciones de estas distintivos en el plasma. Como el ADN se libera en el plasma cuando las células mueren, un aumento de los niveles podría significar un aumento de la muerte celular o de los daños en ese órgano o tejido concreto. Una disminución del nivel de un órgano concreto puede significar que el tratamiento para contrarrestar daños o procesos patológicos en ese órgano está bajo control.
- 35 La FIG. 20 es un diagrama de flujo de un método 2000 para estimar un nivel de metilación del ADN en una muestra biológica de un organismo de acuerdo con realizaciones de la presente invención. El nivel de metilación puede estimarse para una región concreta de un genoma o para todo el genoma. Si se desea una región específica, entonces solo podrán utilizarse fragmentos de ADN de esa región específica.
- 40 En el bloque 2010, se miden las cantidades de fragmentos de ADN correspondientes a distintos tamaños. Para cada tamaño de una pluralidad de tamaños, puede medirse una cantidad de una pluralidad de fragmentos de ADN de la muestra biológica correspondiente al tamaño. A modo de ejemplo, puede medirse el número de fragmentos de ADN con una longitud de 140 bases. Las cantidades pueden guardarse como un histograma. En una realización, se mide el tamaño de cada uno de la pluralidad de ácidos nucleicos de la muestra biológica, que puede hacerse de forma individual (p. ej., mediante la secuenciación de una molécula entera o solo de sus extremos) o en grupo (p. ej., mediante electroforesis). Los tamaños pueden corresponder a un intervalo. Por tanto, una cantidad puede ser para fragmentos de ADN que tengan un tamaño dentro de un intervalo determinado. Cuando se realiza la secuenciación de ambos extremos, los fragmentos de ADN (determinados por las lecturas de secuencias emparejadas) asignados (alineados) a una región concreta pueden utilizarse para determinar el nivel de metilación de la región.
- 45 50 En el bloque 2020, se calcula un primer valor de un primer parámetro basándose en las cantidades de fragmentos de ADN de múltiples tamaños. En un aspecto, el primer parámetro proporciona una medida estadística de un perfil de tamaño (p. ej., un histograma) de fragmentos de ADN en la muestra biológica. El parámetro puede denominarse parámetro de tamaño, ya que se determina a partir de los tamaños de la pluralidad de fragmentos de ADN.
- 55 60 65 El primer parámetro puede ser de varias formas. Un parámetro es el porcentaje de fragmentos de ADN de un determinado tamaño o intervalo de tamaños en relación con todos los fragmentos de ADN o en relación con fragmentos de ADN de otro tamaño o intervalo. Dicho parámetro es un número de fragmentos de ADN de un tamaño determinado dividido entre el número total de fragmentos, que puede obtenerse a partir de un histograma (cualquier estructura de datos que proporcione números absolutos o relativos de fragmentos de determinados tamaños). Como otro ejemplo, un parámetro podría ser un número de fragmentos de un tamaño concreto o dentro de un intervalo concreto dividido entre un número de fragmentos de otro tamaño o intervalo. La división puede actuar como una normalización para tener en cuenta un número diferente de fragmentos de ADN que se analizan para diferentes muestras. Se puede realizar una normalización analizando el mismo número de fragmentos de ADN para cada muestra, que proporciona de manera eficaz el mismo resultado que dividir entre el número total de fragmentos analizados. En la solicitud de patente de EE. UU. n.º 13/789.553 se pueden encontrar ejemplos adicionales de parámetros y sobre el análisis del

tamaño.

En el bloque 2030, el primer valor de tamaño se compara con un valor de tamaño de referencia. El valor del tamaño de referencia puede calcularse a partir de fragmentos de ADN de una muestra de referencia. Para determinar los

5 valores del tamaño de referencia, el perfil de metilación puede calcularse y cuantificarse para una muestra de referencia, así como un valor del primer parámetro de tamaño. Por tanto, cuando el primer valor de tamaño se compara con el valor de tamaño de referencia, se puede determinar un nivel de metilación.

10 En el bloque 2040, el nivel de metilación se estima a partir de la comparación. En una realización, se puede determinar si el primer valor del primer parámetro está por encima o por debajo del valor del tamaño de referencia, y determinar así si el nivel de metilación de la muestra instantánea está por encima o por debajo del nivel de metilación respecto al valor del tamaño de referencia. En otra realización, la comparación se realiza introduciendo el primer valor en una función de calibración. La función de calibración puede comparar eficazmente el primer valor con los valores de

15 calibración (un conjunto de valores de tamaño de referencia) identificando el punto de una curva correspondiente al primer valor. El nivel de metilación estimado se proporciona entonces como valor de salida de la función de calibración.

En consecuencia, se puede calibrar un parámetro de tamaño a un nivel de metilación. Por ejemplo, se puede medir un nivel de metilación y asociarlo a un parámetro de tamaño concreto para esa muestra. A continuación, los puntos

20 de datos de varias muestras pueden ajustarse a una función de calibración. En una implementación, se pueden utilizar diferentes funciones de calibración para diferentes subconjuntos de ADN. Por tanto, puede haber algún tipo de calibración basada en conocimientos previos sobre la relación entre metilación y tamaño para un subconjunto concreto de ADN. Por ejemplo, la calibración del ADN fetal y materno podría ser diferente.

25 Como se ha mostrado anteriormente, la placenta está más hipometilada que la sangre materna, por lo que el ADN fetal es más pequeño debido a la menor metilación. En consecuencia, se puede utilizar un tamaño medio de los fragmentos de una muestra (u otro valor estadístico) para estimar la densidad de metilación. Como los tamaños de los fragmentos pueden medirse mediante secuenciación de ambos extremos, en lugar de la secuenciación consciente de la metilación, potencialmente más compleja desde el punto de vista técnico, este enfoque sería potencialmente rentable si se utilizara clínicamente. Este enfoque puede utilizarse para controlar los cambios de metilación asociados

30 al progreso de la gestación o a trastornos asociados a la gestación, como la preeclampsia, parto prematuro y trastornos fetales (como los provocados por anomalías cromosómicas o genéticas o retraso del crecimiento intrauterino).

35 En otra realización, este enfoque puede utilizarse para detectar y controlar el cáncer. Por ejemplo, con el éxito del tratamiento del cáncer, el perfil de metilación en plasma u otro fluido corporal medido mediante este enfoque basado en el tamaño cambiaría hacia el de individuos sanos sin cáncer. Por el contrario, en caso de que el cáncer progrese, entonces el perfil de metilación en plasma u otro fluido corporal divergiría del de individuos sanos sin cáncer.

40 Resumiendo, las moléculas hipometiladas eran más cortas que las hipermetiladas en plasma. Se observó la misma tendencia en las moléculas de ADN fetal y materno. Como se sabe que la metilación del ADN influye en el empaquetamiento de los nucleosomas, nuestros datos indican que quizás las moléculas de ADN hipometilado estaban menos densamente empaquetadas con histonas y, por tanto, eran más susceptibles a la degradación enzimática. Por otra parte, los datos presentados en las FIGS. 18A y 18B también mostraron que a pesar de que el ADN fetal estaba mucho más hipometilado que las lecturas maternas, la distribución de tamaños del ADN fetal y materno no se separa completamente. En la FIG. 19B, se puede ver que incluso para la misma categoría de tamaño, el nivel de metilación de las lecturas específicas del feto y de la madre difieren entre sí. Esta observación indica que el estado hipometilado del ADN fetal no es el único factor que explica su relativa escasez con respecto al ADN materno.

VII. ESTADO DE SELLADO DE LOS LOCUS GÉNICOS

50 Se pueden detectar moléculas de ADN derivadas del feto que comparten el mismo genotipo, pero con distintivos epigenéticos diferentes a los de la madre en el plasma materno (L. L. M. Poon *et al.* 2002 *Clin Chem*; 48: 35-41). Demostrar que el método de secuenciación es sensible para detectar moléculas de ADN fetal en el plasma materno, aplicamos la misma estrategia para detectar los alelos fetales sellados en muestras de plasma materno. Se identificaron dos regiones genómicas selladas: H19 (cr. 11: 1.977.419-1.977.821, NCBI Build36/hg18) y MEST (cr. 7: 129.917.976-129.920.347, NCBI Build36/hg18). Ambos contienen SNP informativos para la diferenciación entre las secuencias materna y fetal. Para H19, un gen de expresión materna, la madre era homocigótica (A/A) y el feto heterocigótico (A/C) para el SNP rs2071094 (cr. 11: 1.977.740) en la región. Uno de los alelos A maternos estaba totalmente metilado y el otro no lo estaba. En la placenta, sin embargo, el alelo A no estaba metilado, mientras que el alelo C heredado del padre estaba totalmente metilado. Detectamos dos lecturas metiladas con el genotipo C, correspondientes a los alelos paternos sellados derivados de la placenta, en el plasma materno.

60 MEST, también conocido como PEG1, es un gen de expresión paterna. Tanto la madre como el feto eran heterocigóticos (A/G) para el SNP rs2301335 (cr. 7: 129.920.062) dentro del locus sellado. El alelo G estaba metilado, mientras que el alelo A no estaba metilado en la sangre materna. El patrón de metilación se invirtió en la placenta, con el alelo A materno metilado y el alelo G paterno sin metilar. Tres alelos G no metilados, que se derivaban paternamente, eran detectables en el plasma materno. Por el contrario, VAV1, un locus génico no sellado en el

cromosoma 19 (cr. 19: 6.723.621-6.724.121), no mostró ningún patrón de metilación alélica en las muestras de tejido ni en las de ADN plasmático.

- Por tanto, el estado de metilación puede utilizarse para determinar qué fragmentos de ADN proceden del feto. Por ejemplo, la simple detección del alelo A en el plasma materno no puede utilizarse como marcador fetal cuando la madre es heterocigótica GA. Sin embargo, si se distingue el estado de metilación de las moléculas A en plasma, las moléculas A metiladas son específicas del feto, mientras que las moléculas A no metiladas son específicas de la madre, o viceversa.
- A continuación, nos centramos en los locus de los que se ha informado que demuestran el sellado genómico en los tejidos placentarios. Basándose en la lista de locus reportada por Woodfine *et al.* (2011 *Epigenetics Chromatin*; 4:1), clasificamos además los que contenían SNP dentro de la región de control del sellado. Cuatro locus cumplían los criterios y eran H19, KCNQ10T1, MEST y NESP.
- En cuanto a las lecturas de la muestra de glóbulos sanguíneos maternos para H19 y KCNQ10T1, las lecturas maternas eran homocigóticas para el SNP y había proporciones aproximadamente iguales de lecturas metiladas y no metiladas. El MVC y la muestra de tejido placentario a término revelaron que el feto era heterocigótico para ambos locus y que cada alelo estaba exclusivamente metilado o no metilado, es decir, que muestran metilación monoalélica. En las muestras de plasma materno, las moléculas de ADN fetal heredadas paternalmente se detectaron para ambos locus.
- Para H19, las moléculas heredadas paternalmente estaban representadas por las lecturas secuenciadas que contenían el alelo específico del feto y estaban metiladas. Para KCNQ10T1, las moléculas heredadas paternalmente estaban representadas por las lecturas secuenciadas que contenían el alelo específico del feto y no estaban metiladas.
- Por otra parte, la madre era heterocigótica tanto para MEST como para NESP. Para MEST, tanto la madre como el feto eran heterocigóticos GA para el SNP. Sin embargo, como se desprende de los datos de la cadena Watson para los glóbulos sanguíneos maternos y el tejido placentario, el estado de metilación de los CpG adyacentes al SNP era opuesto en la madre y el feto. El alelo A no estaba metilado en el ADN de la madre, pero sí en el del feto. Para MEST, el alelo materno estaba metilado. Por consiguiente, se podría precisar que el feto había heredado el alelo A de su madre (metilado en la MVC) y que la madre había heredado el alelo A de su padre (no metilado en los glóbulos sanguíneos maternos). De manera interesante, en las muestras de plasma materno, los cuatro grupos de moléculas podían distinguirse fácilmente, incluyendo cada uno de los dos alelos de la madre y cada uno de los dos alelos del feto. Por tanto, combinando la información del genotipo con el estado de metilación en los locus sellados, pudimos distinguir fácilmente las moléculas de ADN fetal heredadas maternalmente de las moléculas de ADN materno de fondo (L. L. M. Poon *et al.* 2002 *Clin Chem*; 48: 35-41).
- Este enfoque podría utilizarse para detectar la disomía uniparental. Por ejemplo, si se sabe que el padre de este feto es homocigótico para el alelo G, la no detección del alelo G no metilado en el plasma materno significa la ausencia de contribución del alelo paterno. Adicionalmente, en tal circunstancia, cuando se detectaron tanto el alelo G metilado como el alelo A metilado en el plasma de esta gestación, esto indicaría que el feto tiene heterodisomía de la madre, es decir, que había heredado dos alelos diferentes de la madre sin herencia del padre. Como alternativa, si se detectaban en el plasma materno tanto el alelo A metilado (alelo fetal heredado de la madre) como el alelo A no metilado (alelo materno heredado del abuelo materno) sin el alelo G no metilado (alelo paterno que debería haber heredado el feto), esto indicaría que el feto tiene isodisomía de la madre, es decir, que había heredado dos alelos idénticos de la madre sin herencia del padre.
- Para NESP, la madre era heterocigótica GA en el SNP mientras que el feto era homocigótico para el alelo G. El alelo paterno estaba metilado para NESP. En las muestras de plasma materno, los alelos G fetales heredados paternalmente que estaban metilados podían distinguirse fácilmente de los alelos G maternos de fondo que no estaban metilados.
- ## VIII. CÁNCER/DONANTES
- Algunas realizaciones pueden utilizarse para la detección, cribado, seguimiento (p. ej., de las recaídas, remisión o respuesta (p. ej., presencia o ausencia) al tratamiento), estadificación, clasificación (p. ej., para ayudar a elegir la modalidad de tratamiento más adecuada) y pronóstico del cáncer mediante el análisis de metilación del ADN circulante en plasma/suero.
- Se sabe que el ADN del cáncer muestra una metilación aberrante del ADN (J. G. Herman *et al.* 2003 *N Engl J Med*; 349: 2042-2054). Por ejemplo, los promotores de las islas CpG de los genes, p. ej., genes inhibidores de tumores, están hipermetilados, mientras que los sitios CpG del cuerpo del gen están hipometilados en comparación con las células no cancerosas. Siempre que el perfil de metilación de las células cancerosas pueda reflejarse mediante el perfil de metilación de las moléculas de ADN plasmático derivadas del tumor por medio de los métodos descritos en el presente documento, esperamos que el perfil general de metilación en plasma sea diferente entre los individuos con cáncer en comparación con los individuos sanos sin cáncer o en comparación con aquellos cuyo cáncer se ha curado.
- Los tipos de diferencias en el perfil de metilación podrían ser en términos de diferencias cuantitativas en las densidades de metilación del genoma y/o densidades de metilación de segmentos de los genomas. Por ejemplo, debido a la

naturaleza hipometilada general del ADN de los tejidos cancerosos (Gama-Sosa M. A. et al. 1983 *Nucleic Acids Res*; 11: 6883-6894), se observaría una reducción de las densidades de metilación en el metiloma plasmático o en segmentos del genoma en el plasma de pacientes con cáncer.

- 5 Los cambios cualitativos en el perfil de metilación también deberían reflejarse entre los datos del metiloma plasmático. Por ejemplo, las moléculas de ADN plasmático procedentes de genes hipermetilados solo en células cancerosas mostrarian hipermetilación en el plasma de un paciente con cáncer en comparación con las moléculas de ADN plasmático procedentes de los mismos genes, pero en una muestra de un control sano. Dado que, en la mayoría de los cánceres, se produce la metilación aberrante, los métodos descritos en el presente documento podrían aplicarse
10 a la detección de todas las formas de neoplasias malignas con metilación aberrante, por ejemplo, tumores malignos en, aunque no de forma limitativa, el pulmón, mama, colorrectal, próstata, nasofaringe, estómago, testículos, piel, sistema nervioso, hueso, ovario, hígado, tejidos hemáticos, páncreas, útero, riñón, vejiga, tejido linfoide, etc. Las neoplasias malignas pueden ser de una variedad de subtipos histológicos, por ejemplo, carcinomas, adenocarcinomas, sarcomas, fibroadenocarcinomas, neuroendocrino e indiferenciado, etc.
15
- Por otra parte, esperamos que las moléculas de ADN derivadas de tumores puedan distinguirse de las moléculas de ADN de fondo no derivadas de tumores, ya que el perfil general de tamaño corto del ADN derivado de tumores se acentúa en el caso de las moléculas de ADN procedentes de locus con hipometilación aberrante asociada a tumores, lo que tendría un efecto adicional sobre el tamaño de la molécula de ADN. Del mismo modo, las moléculas de ADN
20 plasmático derivadas de tumores pueden distinguirse de las moléculas de ADN plasmático de fondo no derivadas de tumores utilizando múltiples rasgos característicos que se asocian al ADN tumoral, incluidas, entre otras, las variantes de un solo nucleótido, ganancias y pérdidas del número de copias, translocaciones, inversiones, hiper- o hipometilación aberrante y perfil de tamaños. Como todos estos cambios podrían producirse de forma independiente, el uso combinado de estas características puede proporcionar una ventaja aditiva para la detección sensible y específica
25 de ADN canceroso en el plasma.

A. El tamaño y el cáncer

- 30 El tamaño de las moléculas de ADN derivadas de tumores en plasma también se asemeja a los tamaños de las unidades mononucleosómicas y son más cortas que las moléculas de ADN de fondo no derivadas de tumores, que coexiste en el plasma de pacientes con cáncer. Se ha demostrado que los parámetros de tamaño están correlacionados con el cáncer, como se describe en la solicitud de patente de EE. UU. n.º 13/789.553.
35 Dado que tanto el ADN derivado del feto como el derivado de la madre en el plasma mostraban una relación entre el tamaño y el estado de metilación de la molécula, se espera que las moléculas de ADN derivadas de tumores muestren la misma tendencia. Por ejemplo, las moléculas hipometiladas serían más cortas que las moléculas hipermetiladas en el plasma de pacientes con cáncer o en sujetos sometidos a cribado de cáncer.

B. Densidades de metilación de diferentes tejidos en un paciente con cáncer

- 40 En este ejemplo, analizamos las muestras de plasma y tejido de un paciente con carcinoma hepatocelular (CHC). Se recogieron muestras de sangre del paciente con CHC antes y 1 semana después de la resección quirúrgica del tumor. El plasma y la capa leucocitaria se recogieron tras centrifugar las muestras de sangre. Se recogió el tumor resecado y el tejido hepático adyacente no tumoral. Las muestras de ADN extraídas de las muestras de plasma y tejido se analizaron mediante secuenciación masiva paralela con y sin tratamiento previo con bisulfito. También se analizó como control el ADN plasmático de cuatro individuos sanos sin cáncer. El tratamiento con bisulfito de una muestra de ADN convertiría los restos de citosina no metilados en uracilo. En la reacción en cadena de la polimerasa posterior y la secuenciación, estos restos de uracilo se comportarían como timidina. Por otra parte, el tratamiento con bisulfito no convertiría los restos de citosina metilados en uracilo. Tras la secuenciación masiva en paralelo, las lecturas de secuenciación se analizaron mediante el programa Methyl-Pipe (P Jiang, et al. Methyl-Pipe: An integrated bioinformatics data analysis pipeline for whole genome methylome analysis, ponencia presentada en la IEEE International Conference on Bioinformatics and Biomedicine Workshops, Hong Kong, 18 a 21 de diciembre de 2010), para determinar el estado de metilación de los restos de citosina en todas las posiciones del dinucleótido CG, es decir, sitios CpG.
45
50
55 La FIG. 21A es una tabla 2100 que muestra las densidades de metilación del plasma preoperatorio y las muestras de tejido de un paciente con CHC. La densidad de metilación CpG para las regiones de interés (p. ej., sitios CpG, promotor o regiones de repetición, etc.) se refiere a la proporción de lecturas que muestran metilación CpG sobre el número total de lecturas que cubren dinucleótidos CpG genómicos. Las densidades de metilación de la capa leucocitaria y del tejido hepático no tumoral son similares. La densidad de metilación global del tejido tumoral, basado en datos de todos los autosomas, era un 25 % inferior a las de la capa leucocitaria y el tejido hepático no tumoral. La hipometilación era consistente en cada cromosoma individual. La densidad de metilación del plasma se situaba entre los valores de los tejidos no malignos y los tejidos cancerosos. Esta observación es coherente con el hecho de que tanto los tejidos cancerosos como los no cancerosos contribuirían al ADN circulante de un paciente con cáncer. Se ha demostrado que el sistema hematopoyético es la principal fuente de ADN circulante en individuos sin una enfermedad maligna activa (Y. Y. N. Lui, et al. 2002 *Clin Chem*; 48: 421-7). Por lo tanto, también analizamos muestras de plasma obtenidas de
60
65

cuatro controles sanos. El número de lecturas de secuencia y la profundidad de secuenciación conseguida por muestra se muestran en la tabla 2150 de la FIG. 21B.

- 5 La FIG. 22 es una tabla 220 que muestra que las densidades de metilación en los autosomas oscilaban entre el 71,2 % y el 72,5 % en las muestras de plasma de los controles sanos. Estos datos mostraron el nivel esperado de metilación del ADN en muestras de plasma obtenidas de individuos sin una fuente de ADN tumoral. En un paciente con cáncer, el tejido tumoral también liberaría ADN en el sistema circulatorio (K. C. A. Chan *et al.* 2013 *Clin Chem*; 59: 211-224); R. J. Leary *et al.* 2012 *Sci Transl Med*; 4: 162ral54). Debido a la naturaleza hipometilada del tumor de CHC, la presencia de ADN tanto tumoral como no tumoral en el plasma preoperatorio del paciente provocó una reducción de la densidad de metilación en comparación con los niveles plasmáticos de los controles sanos. De hecho, la densidad de metilación de la muestra de plasma preoperatoria se encontraba entre las densidades de metilación del tejido tumoral y el plasma de los controles sanos. La razón es que el nivel de metilación del ADN plasmático de los pacientes con cáncer estaría influido por el grado de metilación aberrante, hipometilación en este caso, del tejido tumoral y la concentración fraccional del ADN derivado del tumor en el sistema circulatorio. Una menor densidad de metilación del tejido tumoral y una mayor concentración fraccional de ADN derivado del tumor en el sistema circulatorio conducirían a una menor densidad de metilación del ADN plasmático en un paciente con cáncer. La mayoría de los tumores muestran hipometilación global (J. G. Herman *et al.* 2003 *N Engl J Med*; 349: 2042-2054; M. A. Gama-Sosa *et al.* 1983 *Nucleic Acids Res*; 11: 6883-6894). Por tanto, las observaciones actuales observadas en las muestras de CHC también deberían ser aplicables a otros tipos de tumores.
- 10 20 En una realización, la densidad de metilación del ADN plasmático puede utilizarse para determinar la concentración fraccional de ADN derivado del tumor en una muestra de plasma/suero cuando se conoce el nivel de metilación del tejido tumoral. El nivel de metilación, p. ej., la densidad de metilación, del tejido tumoral puede obtenerse si se dispone de la muestra del tumor o de una biopsia del mismo. En otra realización, la información relativa al nivel de metilación del tejido tumoral puede obtenerse a partir del estudio del nivel de metilación en un grupo de tumores de un tipo similar y esta información (p. ej., un nivel medio o un nivel mediano) se aplica al paciente que se va a analizar utilizando la tecnología descrita en esta invención.
- 15 25 30 35 40 45 50 55 60 65 El nivel de metilación del tejido tumoral puede determinarse mediante el análisis del tejido tumoral del paciente o inferirse del análisis de los tejidos tumorales de otros pacientes con el mismo tipo de cáncer o uno similar. La metilación de los tejidos tumorales puede determinarse utilizando una serie de plataformas que tienen en cuenta la metilación, incluida, entre otras, la secuenciación paralela masiva, la secuenciación de una sola molécula, las micromatrices (p. ej., las matrices de oligonucleótidos) o espectrometría de masas (tales como el Epityper, Sequenom, Inc., análisis). En algunas realizaciones, estos análisis pueden ir precedidos de procedimientos sensibles al estado de metilación de las moléculas de ADN, incluidos, aunque no de forma limitativa, la inmunoprecipitación de citosina y digestión con enzimas de restricción que tienen en cuenta la metilación. Cuando se conoce el nivel de metilación de un tumor, la concentración fraccional de ADN tumoral en el plasma de los pacientes con cáncer podría calcularse tras el análisis del metiloma plasmático.
- 40 La relación entre el nivel de metilación plasmática, P, con la concentración fraccional de ADN tumoral, f y el nivel de metilación del tejido tumoral, TUM, puede describirse como: $P = BASAL \times (1-f) + TUM \times f$ donde BASAL es el nivel basal de metilación del ADN en plasma derivado de glóbulos sanguíneos y otros órganos internos. Por ejemplo, la densidad de metilación global de todos los autosomas resultó ser del 42,9 % en el tejido de la biopsia tumoral obtenida de este paciente con CHC, es decir, el valor TUM para este caso. La densidad media de metilación de las muestras de plasma de los cuatro controles sanos fue del 71,6 %, es decir, el valor BASAL de este caso. La densidad de metilación plasmática del plasma preoperatorio fue del 59,7 %. Utilizando estos valores, f se estima en un 41,5 %.
- 45 En otra realización, el nivel de metilación del tejido tumoral puede estimarse de forma no invasiva basándose en los datos del metiloma plasmático cuando se conoce la concentración fraccional del ADN derivado del tumor en la muestra de plasma. La concentración fraccional del ADN derivado del tumor en la muestra de plasma puede determinarse mediante otros análisis genéticos, por ejemplo, el análisis genómico de pérdidas alélicas (GAAU) y el análisis de mutaciones de un solo nucleótido, como se ha descrito anteriormente (solicitud de patente de EE. UU. n.º 13/308.473; K. C. A. Chan *et al.* 2013 *Clin Chem*; 59: 211-224). El cálculo se basa en la misma relación descrita anteriormente, salvo que, en esta realización, el valor de f es conocido y el valor de TUM se convierte en la incógnita. La deducción puede realizarse para todo el genoma o para partes del genoma, de manera similar a los datos observados en el contexto de la determinación del nivel de metilación del tejido placentario a partir de los datos del plasma materno.
- 50 55 60 65 En otra realización, se puede utilizar la variación o el perfil entre bins de las densidades de metilación para diferenciar a los sujetos con cáncer de los que no lo tienen. La resolución del análisis de metilación puede aumentarse aún más dividiendo el genoma en bins de un tamaño determinado, p. ej., 1 Mb. En dicha realización, se calculó la densidad de metilación de cada bin de 1 Mb para las muestras recogidas, p. ej., capa leucoplaquetaria, el tejido resecado del CHC, el tejido hepático no tumoral adyacente al tumor y el plasma recogido antes y después de la resección tumoral. En otra realización, no es necesario mantener constante el tamaño de los bins. En una implementación, el número de sitios CpG se mantiene constante dentro de cada bin, mientras que el propio bin puede variar de tamaño.
- Las FIGS. 23 A y 23B muestran la densidad de metilación de la capa leucocitaria, tejido tumoral, tejido hepático no

tumoral, el plasma preoperatorio y postoperatorio del paciente con CHC. La FIG. 23A es un gráfico 2300 de resultados para el cromosoma 1. La FIG. 23B es un gráfico 2350 de resultados para el cromosoma 2.

Para la mayoría de los intervalos de 1 Mb, las densidades de metilación de la capa leucocitaria y del tejido hepático no tumoral adyacente al tumor eran similares, mientras que las de los tejidos tumorales eran inferiores.

Las densidades de metilación del plasma preoperatorio se situaron entre las del tumor y las de los tejidos no malignos. Las densidades de metilación de las regiones genómicas exploradas en los tejidos tumorales pudieron deducirse utilizando los datos de metilación del plasma preoperatorio y la concentración fraccional de ADN tumoral. El método es el mismo que el descrito anteriormente, utilizando los valores de densidad de metilación de todos los autosomas. La deducción de la metilación tumoral descrita también puede realizarse utilizando estos datos de metilación de mayor resolución del ADN plasmático. También pueden utilizarse otros tamaños de bins, tales como 300 kb, 500 kb, 2 Mb, 3 Mb, 5 Mb o más de 5 Mb. En una realización, no es necesario mantener constante el tamaño de los bins. En una implementación, el número de sitios CpG se mantiene constante dentro de cada bin, mientras que el propio bin puede variar de tamaño.

C. Comparación de la densidad de metilación plasmática entre el paciente con cáncer y los individuos sanos

Como se muestra en 2100, las densidades de metilación del ADN plasmático preoperatorio eran inferiores a las de los tejidos no malignos del paciente con cáncer. Es probable que esto se deba a la presencia de ADN del tejido tumoral que estaba hipometilado. Esta menor densidad de metilación del ADN en plasma puede utilizarse como biomarcador para la detección y el seguimiento del cáncer. Para el seguimiento del cáncer, si un cáncer está progresando, entonces, con el tiempo, habrá una mayor cantidad de ADN derivado del cáncer en el plasma. En este ejemplo, una mayor cantidad de ADN circulante derivado del cáncer en el plasma conducirá a una mayor reducción de la densidad de metilación del ADN plasmático a nivel de todo el genoma.

Por el contrario, si un cáncer responde al tratamiento, entonces la cantidad de ADN derivado del cáncer en el plasma disminuirá con el tiempo. En este ejemplo, una disminución de la cantidad de ADN derivado del cáncer en el plasma conllevará un aumento de la densidad de metilación del ADN plasmático. Por ejemplo, si un paciente con cáncer de pulmón con mutación del receptor del factor de crecimiento epidérmico ha sido tratado con una terapia dirigida, p. ej., la inhibición de la tirosina cinasa, entonces un aumento de la densidad de metilación del ADN en plasma significaría una respuesta. Posteriormente, la aparición de un clon tumoral resistente a la inhibición de la tirosina cinasa se asociaría a una disminución de la densidad de metilación del ADN en plasma, lo que indicaría una recaída.

Las mediciones de la densidad de metilación plasmática pueden realizarse en serie y la tasa de cambio de dichas mediciones puede calcularse y utilizarse para predecir o correlacionar con la progresión clínica o la remisión o el pronóstico. Para los locus genómicos seleccionados que están hipermetilados en tejidos cancerosos, pero hipometilados en tejidos normales, p. ej., las regiones promotoras de varios genes inhibidores de tumores, la relación entre la progresión del cáncer y la respuesta favorable al tratamiento será opuesta a los patrones descritos anteriormente.

A fin de demostrar la viabilidad de este enfoque, comparamos las densidades de metilación del ADN de muestras de plasma recogidas del paciente con cáncer antes y después de la extirpación quirúrgica del tumor con ADN plasmático obtenido de cuatro sujetos de control sanos.

La Tabla 2200 muestra las densidades de metilación del ADN de cada autosoma y los valores combinados de todos los autosomas de las muestras de plasma preoperatorias y postoperatorias del paciente con cáncer y de los cuatro sujetos de control sanos. Para todos los cromosomas, las densidades de metilación de la muestra de ADN plasmático preoperatorio eran inferiores a las de la muestra postoperatoria y a las de las muestras de plasma de los cuatro sujetos sanos. La diferencia en las densidades de metilación del ADN plasmático entre las muestras preoperatorias y postoperatorias proporcionó pruebas de apoyo de que las densidades de metilación más bajas en la muestra de plasma preoperatoria se debían a la presencia de ADN procedente del tumor de CHC.

La inversión de las densidades de metilación del ADN en la muestra de plasma postoperatoria hasta niveles similares a los de las muestras de plasma de los controles sanos indicaba que gran parte del ADN derivado del tumor había desaparecido debido a la extirpación quirúrgica del origen, es decir, el tumor. Estos datos indican que la densidad de metilación del plasma preoperatorio determinada utilizando los datos disponibles de una gran cantidad de regiones genómicas, tales como todos los autosomas o cromosomas individuales, era de un nivel de metilación inferior al de los controles sanos para permitir la identificación, es decir, diagnóstico o cribado, del caso de prueba de enfermo de cáncer.

Los datos del plasma preoperatorio también mostraron un nivel de metilación mucho más bajo que el del plasma postoperatorio, lo que indica que el nivel de metilación del plasma también podría utilizarse para controlar la carga tumoral, por tanto, para pronosticar y controlar la evolución del cáncer en el paciente. Los valores de referencia pueden determinarse a partir del plasma de controles sanos o de personas con riesgo de padecer el cáncer pero que actualmente no lo padecen. Entre las personas con riesgo de sufrir un CHC se encuentran las que padecen una

infección crónica por hepatitis B o hepatitis C, las que padecen hemocromatosis y las que padecen cirrosis hepática.

Los valores de densidad de metilación en plasma más allá, por ejemplo, inferiores a, un valor de corte definido basado en los valores de referencia se puede utilizar para evaluar si el plasma de una persona no gestante tiene ADN tumoral o no. Para detectar la presencia de ADN tumoral circulante hipometilado, el valor de corte puede definirse como inferior al percentil 5 o 1 de los valores de la población de control, o basarse en un número de desviaciones típicas, por ejemplo, 2 o 3 desviaciones típicas (DT), por debajo de los valores medios de densidad de metilación de los controles, o basándose en la determinación de un múltiplo de la mediana (MdM). Para el ADN tumoral hipermetilado, el valor de corte puede definirse como superior al percentil 95 o 99 de los valores de la población de control, o basarse en un número de desviaciones típicas, por ejemplo, 2 o 3 DT, por encima de los valores medios de densidad de metilación de los controles, o basándose en la determinación de un múltiplo de la mediana (MdM). En una realización, la población de control tiene la misma edad que el sujeto de la prueba. El ajuste de la edad no tiene por qué ser exacto y puede realizarse en franjas de edad (p. ej., de 30 a 40 años, para un sujeto de 35 años).

15 A continuación, comparamos las densidades de metilación de bins de 1 Mb entre las muestras de plasma del paciente con cáncer y los cuatro sujetos de control. Con fines ilustrativos, se muestran los resultados del cromosoma 1.

Las FIGS. 24A es un gráfico 2400 que muestra las densidades de metilación del plasma preoperatorio del paciente con CHC. Las FIGS. 24B es un gráfico 2450 que muestra las densidades de metilación del plasma postoperatorio del paciente con CHC. Los puntos azules representan los resultados de los sujetos de control, los puntos rojos representan los resultados de la muestra de plasma del paciente con CHC.

Como se muestra en la figura 24A, las densidades de metilación del plasma preoperatorio del paciente con CHC eran inferiores a las de los sujetos de control en la mayoría de los bins. Se observaron patrones similares en otros cromosomas. Tal como se muestra en la Figura 24B, las densidades de metilación del plasma postoperatorio del paciente con CHC fueron similares a las de los sujetos de control en la mayoría de los bins. Se observaron patrones similares en otros cromosomas.

30 Para evaluar si un sujeto examinado padece cáncer, se compararía el resultado del sujeto sometido a la prueba con los valores de un grupo de referencia. En una realización, el grupo de referencia puede estar formado por una serie de sujetos sanos. En otra realización, el grupo de referencia puede estar formado por sujetos con enfermedades no malignas, por ejemplo, infección crónica por hepatitis B o cirrosis. A continuación, se puede cuantificar la diferencia en las densidades de metilación entre el sujeto analizado y el grupo de referencia.

35 En una realización, puede obtenerse un intervalo de referencia a partir de los valores del grupo de control.

Entonces, se pueden utilizar las desviaciones en el resultado del sujeto analizado con respecto a los límites superior o inferior del grupo de referencia para determinar si el sujeto tiene un tumor. Esta cantidad se vería afectada por la concentración fraccional de ADN derivado del tumor en el plasma y la diferencia en el nivel de metilación entre tejidos malignos y no malignos. Una mayor concentración fraccional de ADN derivado de tumores en el plasma daría lugar a mayores diferencias de densidad de metilación entre la muestra de plasma de prueba y los controles. Un mayor grado de diferencia en el nivel de metilación de los tejidos malignos y no malignos también se asocia con mayores diferencias de densidad de metilación entre la muestra de plasma de prueba y los controles.

45 En otra realización más, se eligen diferentes grupos de referencia para los sujetos de ensayo de distintos intervalos de edad.

En otra realización, se calcularon la media y la DT de las densidades de metilación de los cuatro sujetos de control para cada bin de 1 Mb. A continuación, para las ubicaciones correspondientes, se calculó la diferencia entre las densidades de metilación del paciente con CHC y el valor medio de los sujetos de control. En una realización, a continuación, esta diferencia se dividió entre la DT del bin correspondiente para determinar la puntuación z. Dicho de otro modo, la puntuación z representa la diferencia en las densidades de metilación entre las muestras de plasma de prueba y de control, expresada como un número de DT a partir de la media de los sujetos de control. Una puntuación z >3 de un bin indica que el ADN plasmático del paciente con CHC está más hipermetilado que el de los sujetos de control en más de 3 DT en ese bin, mientras que una puntuación z <-3 de un bin indica que el ADN plasmático del paciente con CHC está más hipometilado que el de los sujetos de control en más de 3 DT en ese bin.

Las FIGS. 25A y 25B muestran las puntuaciones z de las densidades de metilación del ADN plasmático para las muestras plasmáticas preoperatorias (gráfica 2500) y postoperatorias (gráfica 2550) del paciente con CHC, utilizando los datos del metiloma plasmático de los cuatro sujetos de control sanos como referencia para el cromosoma 1. Cada punto representa el resultado de un bin de 1 Mb. Los puntos negros representan los bins con puntuación z de entre -3 y 3. Los puntos rojos representan bins con puntuación z <-3.

65 La FIG. 26A es una tabla 2600 que muestra los datos de las puntuaciones z del plasma preoperatorio y postoperatorio. La mayoría de los bins del cromosoma 1 (80,9 %) de la muestra de plasma preoperatoria tenían una puntuación z <-3, lo que indica que el ADN plasmático preoperatorio del paciente con CHC estaba significativamente más hipometilado

que el de los sujetos de control. Por el contrario, el número de puntos rojos disminuyó sustancialmente en la muestra de plasma postoperatoria (8,3 % de los bins del cromosoma 1), lo que indica que la mayor parte del ADN tumoral se había eliminado del sistema circulatorio debido a la resección quirúrgica de la fuente de ADN tumoral circulante.

- 5 La FIG. 26B es un diagrama Circos 2620 que muestra la puntuación z de las densidades de metilación del ADN plasmático para las muestras de plasma preoperatorio y postoperatorio del paciente con CHC utilizando los cuatro sujetos de control sanos como referencia para bins de 1 Mb analizados de todos los autosomas. El anillo exterior muestra los ideogramas de los autosomas humanos. El anillo central muestra los datos de la muestra de plasma preoperatorio. El anillo más interior muestra los datos de la muestra de plasma postoperatorio. Cada punto representa el resultado de un bin de 1 Mb. Los puntos negros representan los bins con puntuaciones z de entre -3 y 3. Los puntos rojos representan los bins con puntuaciones z <-3. Los puntos verdes representan los bins con puntuaciones z >3.
- 10 15 20 25

La FIG. 26C es una tabla 2640 que muestra una distribución de las puntuaciones z de los bins de 1 Mb para todo el genoma en las muestras de plasma preoperatorio y postoperatorias del paciente con CHC. Los resultados indican que el ADN plasmático preoperatorio del paciente con CHC estaba más hipometilado que el de los controles en la mayoría de las regiones (85,2 % de los bins de 1 Mb) de todo el genoma. Por el contrario, la mayoría de las regiones (93,5 % de los bins de 1 Mb) en la muestra de plasma postoperatorio no mostraron hipermetilación o hipometilación significativa en comparación con los controles. Estos datos indican que gran parte del ADN tumoral, principalmente de naturaleza hipometilada para este CHC, ya no estaba presente en la muestra de plasma postoperatorio.

20 En una realización, el número, el porcentaje o la proporción de bins con puntuaciones z <-3 puede utilizarse para indicar si hay un cáncer presente. Por ejemplo, como se muestra en la tabla 2640, 2330 de los 2734 bins analizados (85,2 %) mostraron puntuaciones z <-3 en el plasma preoperatorio, mientras que solo 171 de los 2734 bins analizados (6,3 %) mostraron puntuaciones z <-3 en el plasma postoperatorio. Los datos indicaron que la carga de ADN tumoral en el plasma preoperatorio era mucho mayor que en el plasma postoperatorio.

30 35 40 45 50

Los valores de corte del número de bins pueden determinarse utilizando métodos estadísticos. Por ejemplo, se esperaría que aproximadamente el 0,15 % de los bins tuvieran una puntuación z de <-3 basándose en una distribución normal. Por lo tanto, el número de corte de los bins puede ser el 0,15 % del número total de bins analizados. Dicho de otro modo, si una muestra de plasma de un individuo no gestante presenta más de un 0,15 % de bins con puntuación z <-3, existe una fuente de ADN hipometilado en el plasma, en concreto, cáncer. Por ejemplo, el 0,15 % de los 2734 bins de 1 Mb que hemos analizado en este ejemplo son aproximadamente 4 bins. Utilizando este valor como valor de corte, tanto las muestras de plasma preoperatorio como las de plasma postoperatorio contenían ADN hipometilado derivado del tumor, aunque la cantidad es mucho mayor en la muestra de plasma preoperatorio que en el plasma postoperatorio. Para los cuatro sujetos de control sanos, ninguno de los bins mostró una hipermetilación o hipometilación significativa. Pueden utilizarse otros valores de corte (p. ej., 1,1 %), que pueden variar en función de los requisitos del ensayo utilizado. Como otros ejemplos, el porcentaje de corte puede variar en función de la distribución estadística, así como la sensibilidad deseada y una especificidad aceptable.

40 45 50

En otra realización, el número de corte puede determinarse mediante el análisis de la curva de eficacia diagnóstica (ROC) analizando una serie de pacientes con cáncer e individuos sin cáncer. Para validar aún más la especificidad de este enfoque, se analizó una muestra de plasma de un paciente que acudió a consulta médica por una afección no maligna (C06). El 1,1 % de los bins tenía una puntuación z <-3. En una realización, se pueden utilizar diferentes umbrales para clasificar distintos niveles de estado de la enfermedad. Se puede utilizar un umbral porcentual más bajo para diferenciar el estado sano de las afecciones benignas y un umbral porcentual más alto para diferenciar las afecciones benignas de las malignas.

55 60 65

El rendimiento diagnóstico del análisis de la hipometilación plasmática mediante secuenciación masiva en paralelo parece ser superior al que se obtiene utilizando la amplificación basada en la reacción en cadena de la polimerasa (PCR) de clases específicas de elementos repetitivos, p. ej., el elemento nuclear intercalado largo-1 (LINE-1) (P Tangkijvanich *et al.* 2007 *Clin Chim Acta*; 379:127-133). Una posible explicación de esta observación es que, aunque la hipometilación es omnipresente en el genoma tumoral, presenta cierto grado de heterogeneidad de una región genómica a otra.

De hecho, observamos que las densidades medias de metilación en plasma de los sujetos de referencia variaban a lo largo del genoma (Fig. 56). Cada punto rojo de la Fig. 56 muestra la densidad de metilación media de un bin de 1 Mb entre 32 sujetos sanos. El gráfico muestra todos los bins de 1 Mb analizados en el genoma. El número dentro de cada recuadro representa el número de cromosoma. Observamos que las densidades medias de metilación variaban de un bin a otro.

Un simple ensayo basado en la PCR no podría tener en cuenta esta heterogeneidad entre regiones en su algoritmo de diagnóstico. Esta heterogeneidad ampliaría el intervalo de densidades de metilación observadas entre los individuos sanos. Se necesitaría entonces una mayor magnitud de reducción de la densidad de metilación para considerar que una muestra presenta hipometilación. Esto reduciría la sensibilidad de la prueba.

Por el contrario, un enfoque basado en la secuenciación masiva en paralelo divide el genoma en bins de 1 Mb (o de

- otro tamaño) y mide las densidades de metilación de cada uno de esos bins. Este enfoque reduce el impacto de las variaciones en las densidades de metilación de referencia a través de diferentes regiones genómicas, ya que cada región se compara entre una muestra de prueba y los controles. En efecto, dentro del mismo bin, la variación interindividual entre los 32 controles sanos fue relativamente pequeña. El 95 % de los bins tuvo un coeficiente de variación (CV) en los 32 controles sanos de ≤1,8 %. Aún así, para aumentar aún más la sensibilidad en la detección de la hipometilación asociada al cáncer, la comparación puede realizarse a través de múltiples regiones genómicas. La sensibilidad aumentaría si se analizaran varias regiones genómicas, ya que se evitaría el efecto de la variación biológica cuando la muestra de cáncer no mostrara hipometilación en una región concreta al analizar solo una región.
- 5 10 15 20 25 30 35 40 45 50 55 60 65
- El enfoque de comparar las densidades de metilación de regiones genómicas equivalentes entre los controles y las muestras de prueba (p. ej., analizando cada región genómica por separado y luego, posiblemente, combinando dichos resultados) y realizar esta comparación para múltiples regiones genómicas tiene una mayor relación señal-ruido para la detección de la hipometilación asociada con el cáncer. Este enfoque de secuenciación paralela masiva se muestra a modo de ilustración. Sería previsible que otras metodologías que pudieran determinar las densidades de metilación de múltiples regiones genómicas y permitir la comparación de las densidades de metilación de las regiones correspondientes entre los controles y las muestras de prueba lograran un efecto similar. Por ejemplo, para lograr el efecto deseado, podrían diseñarse sondas de hibridación o sondas de inversión molecular que pudieran dirigirse a moléculas de ADN plasmático procedentes de regiones genómicas específicas, así como determinar el nivel de metilación de la región.
- En otra realización más, la suma de las puntuaciones z de todos los bins puede utilizarse para determinar si hay cáncer o para controlar los cambios en serie del nivel de metilación del ADN en plasma. Debido a la naturaleza hipometilada general del ADN tumoral, la suma de las puntuaciones z sería menor en el plasma recogido de un individuo con cáncer que en los controles sanos. La suma de las puntuaciones z de la muestra de plasma preoperatorio y postoperatorio del paciente con CHC fue de -49843,8 y -3132,13, respectivamente.
- En otras realizaciones, se pueden utilizar otros métodos para estudiar el nivel de metilación del ADN plasmático. Por ejemplo, la proporción de restos de citosina metilados sobre el contenido total de restos de citosina puede determinarse mediante espectrometría de masas (ML Chen *et al.* 2013 *Clin Chem*; 59: 824-832) o la secuenciación paralela masiva. Sin embargo, ya que la mayoría de los restos de citosina no se encuentran en el contexto del dinucleótido CpG, la proporción de citosina metilada entre el total de restos de citosina sería relativamente pequeña en comparación con los niveles de metilación estimados en el contexto de los dinucleótidos CpG. Se determinó el nivel de metilación de las muestras de tejido y plasma obtenidas del paciente con CHC, así como de las cuatro muestras de plasma obtenidas de los controles sanos. Los niveles de metilación se midieron en el contexto de CpG, cualquier citosina, en los contextos de CHG y CHH utilizando los datos de secuenciación paralela masiva de todo el genoma. H se refiere a los restos de adenina, timina o citosina.
- La FIG. 26D es una tabla 2660 que muestra los niveles de metilación del tejido tumoral y de la muestra de plasma preoperatorio que se solapan con algunas de las muestras de plasma de control cuando se utilizan los contextos de CHH y CHG. Los niveles de metilación del tejido tumoral y de la muestra de plasma preoperatorio fueron sistemáticamente inferiores en comparación con la capa leucocitaria, tejido hepático no tumoral, muestra de plasma postoperatorio y muestras de plasma de control sano tanto en los CpG como en las citosinas no especificadas. Sin embargo, los datos basados en los CpG metilados, es decir, las densidades de metilación, mostraron un intervalo dinámico más amplio que los datos basados en las citosinas metiladas.
- En otras realizaciones, el estado de metilación del ADN plasmático puede determinarse mediante métodos que utilizan anticuerpos contra la citosina metilada, por ejemplo, inmunoprecipitación del ADN metilado (MeDIP). Sin embargo, se espera que la precisión de estos métodos sea inferior a la de los métodos basados en la secuenciación debido a la variabilidad en la unión de anticuerpos. En otra realización más, se puede determinar el nivel de 5-hidroximetilcitosina en el ADN plasmático. A este respecto, se ha descubierto que la reducción del nivel de 5-hidroximetilcitosina es una característica epigenética de ciertos tipos de cáncer, p. ej., melanoma (C. G. Lian, *et al.* 2012 *Cell*; 150: 1135-1146).
- Además de los CHC, también investigamos si este enfoque podía aplicarse a otros tipos de cáncer. Analizamos las muestras de plasma de 2 pacientes con adenocarcinoma de pulmón (CP1 y CP2), 2 pacientes con carcinoma nasofaríngeo (CNF1 y CNF2), 2 pacientes con cáncer colorrectal (CCR1 y CCR2), 1 paciente con tumor neuroendocrino metastásico (NE1) y 1 paciente con sarcoma de músculo liso metastásico (SML1). El ADN plasmático de estos sujetos se convirtió en bisulfito y se secuenció utilizando la plataforma HiSeq2000 de Illumina para 50 pb en un extremo. Los cuatro sujetos de control sanos mencionados anteriormente se utilizaron como grupo de referencia para el análisis de estos 8 pacientes. Se utilizaron 50 pb de la secuencia leída en un extremo. El genoma completo se dividió en bins de 1 Mb. Se calcularon la media y la DT de la densidad de metilación para cada bin utilizando los datos del grupo de referencia. A continuación, los resultados de los 8 pacientes con cáncer se expresaron como puntuaciones z, que representan el número de DT respecto a la media del grupo de referencia. Un valor positivo indica que la densidad de metilación del caso de prueba es inferior a la media del grupo de referencia, y viceversa. El número de lecturas de secuencia y la profundidad de secuenciación conseguida por muestra se muestran en la tabla 2780 de la FIG. 271.

Las FIG. 27A-H muestran diagramas Circos de la densidad de metilación de 8 pacientes con cáncer de acuerdo con las realizaciones de la presente invención. Cada punto representa el resultado de un bin de 1 Mb. Los puntos negros representan los bins con puntuaciones z de entre -3 y 3. Los puntos rojos representan los bins con puntuaciones z < -3. Los puntos verdes representan los bins con puntuaciones z >3. El intervalo entre dos líneas consecutivas representa una diferencia de puntuación z de 20.

5 Se observó una hipometilación significativa en múltiples regiones de los genomas de pacientes con la mayoría de los tipos de cáncer, incluido el cáncer de pulmón, carcinoma nasofaríngeo, cáncer colorrectal y tumor neuroendocrino metastásico. De manera interesante, además de la hipometilación, se observó una hipermetilación significativa en 10 múltiples regiones de todo el genoma en el caso del sarcoma de músculo liso metastásico. El origen embrionario del sarcoma de músculo liso es el mesodermo, mientras que el origen embrionario de los otros tipos de cáncer en los 7 pacientes restantes es el ectodermo. Por lo tanto, es posible que el patrón de metilación del ADN del sarcoma sea diferente de aquel del carcinoma.

15 Como puede observarse a partir de este caso, el patrón de metilación del ADN plasmático también puede ser útil para diferenciar distintos tipos de cáncer, que, en este ejemplo, es una diferenciación de carcinoma y sarcoma. Estos datos también indican que el enfoque podría utilizarse para detectar hipermetilación aberrante asociada a la malignidad. Para todos estos 8 casos, solo se disponía de muestras de plasma y no se había analizado tejido tumoral. Esto demostró que incluso sin el perfil de metilación previo o los niveles de metilación del tejido tumoral, el ADN derivado 20 de tumores puede detectarse fácilmente en plasma utilizando los métodos descritos.

25 La FIG. 27J es una tabla 2790 es una tabla que muestra una distribución de las puntuaciones z de los bins de 1 Mb para todo el genoma en plasma de pacientes con diferentes tumores malignos. Los porcentajes de bins con puntuación z <-3, de -3 a 3 y >3 se muestran para cada caso. Más del 5 % de los bins tenían una puntuación z <-3 en todos los casos. Por lo tanto, si utilizamos un umbral del 5 % de los bins significativamente hipometilados para clasificar una muestra como positiva para cáncer, entonces todos estos casos se clasificarían como positivos para cáncer. Nuestros resultados muestran que es probable que la hipometilación sea un fenómeno general para diferentes tipos de cáncer y el análisis del metiloma plasmático sería útil para detectar diferentes tipos de cáncer.

30 D. Método

La FIG. 28 es un diagrama de flujo del método 2800 de análisis de una muestra biológica de un organismo para determinar una clasificación de un nivel de cáncer de acuerdo con realizaciones de la presente invención. La muestra biológica incluye ADN procedente de células normales y puede incluir ADN de células asociadas al cáncer. Al menos 35 parte del ADN puede ser extracelular en la muestra biológica.

En el bloque 2810, se analizan una pluralidad de moléculas de ADN de la muestra biológica. El análisis de una molécula de ADN puede incluir la determinación de una ubicación de la molécula de ADN en un genoma del organismo y la determinación de si la molécula de ADN está metilada en uno o más sitios. El análisis puede realizarse recibiendo 40 lecturas de secuencias de una secuenciación consciente de la metilación, y así el análisis puede realizarse solo sobre datos obtenidos previamente del ADN. En otras realizaciones, el análisis puede incluir la secuenciación real u otras etapas activas de la obtención de los datos.

En el bloque 2820, se determina un número respectivo de moléculas de ADN que están metiladas en el sitio para cada 45 uno de una pluralidad de sitios. En una realización, los sitios son sitios CpG, y pueden ser solo ciertos sitios CpG, seleccionados de acuerdo con uno o varios de los criterios mencionados en el presente documento. El número de moléculas de ADN metilado equivale a determinar el número de ADN no metilado una vez realizada la normalización utilizando el número total de moléculas de ADN analizadas en un sitio concreto, p. ej., un número total de lecturas de secuencias. Por ejemplo, un aumento de la densidad de metilación de CpG de una región equivale a una disminución 50 de la densidad de CpG no metilados de la misma región.

En el bloque 2830, se calcula un primer nivel de metilación basado en los números respectivos de moléculas de ADN metiladas en la pluralidad de sitios. El primer nivel de metilación puede corresponder a una densidad de metilación que se determina en función del número de moléculas de ADN correspondientes a la pluralidad de sitios. Los sitios 55 pueden corresponder a una pluralidad de locus o a un solo locus.

En el bloque 2840, el primer nivel de metilación se compara con un primer valor de corte. El primer valor de corte puede ser un nivel de metilación de referencia o estar relacionado con un nivel de metilación de referencia (p. ej., una distancia especificada de un nivel normal). El nivel de metilación de referencia puede determinarse a partir de muestras 60 de individuos sin cáncer o de locus o del organismo que se sabe que no están asociados con un cáncer del organismo. El primer valor de corte puede establecerse a partir de un nivel de metilación de referencia determinado a partir de una muestra biológica previa del organismo obtenida antes de la muestra biológica analizada.

En una realización, el primer valor de corte es una distancia especificada (p. ej., un número especificado de 65 deviaciones típicas) de un nivel de metilación de referencia establecido a partir de una muestra biológica obtenida de un organismo sano. La comparación puede realizarse determinando una diferencia entre el primer nivel de metilación

y un nivel de metilación de referencia, y luego comparando la diferencia con un umbral correspondiente al primer valor de corte (p. ej., para determinar si el nivel de metilación es estadísticamente diferente del nivel de metilación de referencia).

- 5 En el bloque 2850, a partir de la comparación se determina la clasificación de un nivel de cáncer. Algunos ejemplos de un nivel de cáncer incluyen si el sujeto tiene cáncer o una afección premaligna, o una mayor probabilidad de desarrollar cáncer. En una realización, el primer valor de corte puede determinarse a partir de una muestra obtenida previamente del sujeto (p. ej., puede determinarse un nivel de metilación de referencia a partir de la muestra anterior).
- 10 En algunas realizaciones, el primer nivel de metilación puede corresponder a un número de regiones cuyos niveles de metilación superan un valor umbral. Por ejemplo, puede identificarse una pluralidad de regiones de un genoma del organismo. Las regiones pueden identificarse utilizando los criterios aquí mencionados, p. ej., de ciertas longitudes o cierto número de sitios. Se pueden identificar uno o más sitios (p. ej., sitios CpG) dentro de cada una de las regiones. Se puede calcular el nivel de metilación de cada región. El primer nivel de metilación corresponde a una primera región.
- 15 Cada uno de los niveles de metilación de la región se compara con un valor de corte de la región respectiva, que pueden ser los mismos o variar de acuerdo con las regiones. El valor de corte de la región para la primera región es el primer valor de corte. Los valores de corte de la región respectiva pueden ser una cantidad especificada (p. ej., 0,5) a partir de un nivel de metilación de referencia, contando así solo las regiones que presentan una diferencia significativa con respecto a una referencia, que pueden determinarse a partir de sujetos no cancerosos.
- 20
- 25 Se puede determinar un primer número de regiones cuyo nivel de metilación de la región supere el valor de corte de la región respectiva, y compararlo con un valor umbral para determinar la clasificación. En una implementación, el valor umbral es un porcentaje. Comparar el primer número con un valor umbral puede incluir dividir el primer número de regiones entre un segundo número de regiones (p. ej., todas las regiones) antes de comparar con el valor umbral, p. ej., como parte de un proceso de normalización.

Como se ha descrito anteriormente, una concentración fraccionaria de ADN tumoral en la muestra biológica puede utilizarse para calcular el primer valor de corte. La concentración fraccionaria puede estimarse simplemente para que sea superior a un valor mínimo, mientras que una muestra con una concentración fraccionaria inferior al valor mínimo puede marcarse, p. ej., como no aptos para el análisis. El valor mínimo puede determinarse basándose en una diferencia esperada en los niveles de metilación de un tumor en relación con un nivel de metilación de referencia. Por ejemplo, si la diferencia es de 0,5 (p. ej., como valor de corte), entonces se requeriría una cierta concentración tumoral lo suficientemente alta como para ver esta diferencia.

- 30
- 35 Para el método 2800 pueden aplicarse técnicas específicas del método 1300. En el método 1300, pueden determinarse variaciones en el número de copias de un tumor (p. ej., cuando puede comprobarse si la primera región cromosómica de un tumor presenta un cambio en el número de copias en relación con una segunda región cromosómica del tumor). Por tanto, el método 1300 puede suponer que existe un tumor. En el método 2800, se puede analizar una muestra para ver si hay indicios de que exista algún tumor, independientemente de las características del número de copias.
- 40 Algunas técnicas de los dos métodos pueden ser similares. Sin embargo, los valores de corte y los parámetros de metilación (p. ej., niveles de metilación normalizados) para el método 2800 pueden detectar una diferencia estadística con respecto a un nivel de metilación de referencia para el ADN no canceroso, en contraposición a una diferencia con respecto a un nivel de metilación de referencia para una mezcla de ADN canceroso y ADN no canceroso con algunas regiones que posiblemente tengan variaciones en el número de copias. Por tanto, los valores de referencia para el
- 45 método 2800 pueden determinarse a partir de muestras sin cáncer, como de organismos sin cáncer o de tejido no canceroso del mismo paciente (p. ej., plasma tomado previamente o de muestras adquiridas contemporáneamente que se sabe que no tienen cáncer, que puede determinarse a partir del ADN celular).

E. Predicción de la concentración fraccional mínima de ADN tumoral que debe detectarse mediante el análisis de metilación del ADN plasmático

Una forma de medir la sensibilidad del enfoque para detectar el cáncer utilizando el nivel de metilación del ADN plasmático está relacionada con la concentración fraccional mínima de ADN derivado del tumor que se requiere para revelar un cambio en el nivel de metilación del ADN plasmático cuando se compara con el de los controles. La sensibilidad de la prueba también depende del grado de diferencia en la metilación del ADN entre el tejido tumoral y los niveles plasmáticos de referencia de metilación del ADN en controles sanos o ADN de glóbulos sanguíneos. Los glóbulos sanguíneos son la fuente predominante de ADN en el plasma de los individuos sanos. Cuanto mayor sea la diferencia, más fácilmente se puede discriminar a los pacientes con cáncer de los individuos sin cáncer y se reflejaría como un límite de detección más bajo de derivados tumorales en plasma y una mayor sensibilidad clínica en la detección de los pacientes con cáncer. Adicionalmente, las variaciones en la metilación del ADN plasmático en los sujetos sanos o en sujetos con diferentes edades (G Hannum et al. 2013 Mol Cell; 49: 359-367) también afectaría a la sensibilidad de la detección de los cambios de metilación asociados a la presencia de un cáncer. Una variación menor de la metilación del ADN plasmático en los sujetos sanos facilitaría la detección del cambio provocado por la presencia de una pequeña cantidad de ADN derivado del cáncer.

65 La FIG. 29A es un gráfico 2900 que muestra la distribución de las densidades de metilación en sujetos de referencia

- suponiendo que esta distribución sigue una distribución normal. Este análisis se basa en que cada muestra de plasma solo proporciona un valor de densidad de metilación, por ejemplo, la densidad de metilación de todos los autosomas o de un cromosoma en particular. Ilustra cómo se vería afectada la especificidad del análisis. En una realización, para determinar si una muestra analizada está significativamente más hipometilada que las muestras de los sujetos de referencia se utiliza un valor de corte de 3 DT por debajo de la densidad media de metilación del ADN de los sujetos de referencia. Cuando se utiliza este corte, se espera que aproximadamente el 0,15 % de los sujetos no cancerosos tengan resultados falsos positivos de ser clasificados como enfermos de cáncer, lo que resulta en una especificidad del 99,85 %.
- La FIG. 29B es un gráfico 2950 que muestra las distribuciones de las densidades de metilación en sujetos de referencia y pacientes con cáncer. El valor de corte es 3 DT por debajo de la media de las densidades de metilación de los sujetos de referencia. Si la media de las densidades de metilación de los pacientes con cáncer es 2 DT por debajo del valor de corte (es decir, 5 DT por debajo de la media de los sujetos de referencia), cabría esperar que el 97,5 % de los sujetos con cáncer tuvieran una densidad de metilación por debajo del valor de corte. Dicho de otro modo, la sensibilidad esperada sería del 97,5 % si se proporciona un valor de densidad de metilación para cada sujeto, por ejemplo, cuando se analiza la densidad de metilación total de todo el genoma, de todos los autosomas o de un cromosoma en particular. La diferencia entre las densidades medias de metilación de las dos poblaciones se ve afectada por dos factores, en concreto, el grado de diferencia en el nivel de metilación entre los tejidos cancerosos y no cancerosos y la concentración fraccional de ADN derivado de tumores en la muestra de plasma. Cuanto más altos sean los valores de estos dos parámetros, mayor será la diferencia en el valor de las densidades de metilación de estas dos poblaciones. Adicionalmente, cuanto menor sea la DT de las distribuciones de las densidades de metilación de las dos poblaciones, menor es el solapamiento de las distribuciones de las densidades de metilación de las dos poblaciones.
- En el presente caso se utilizó un ejemplo hipotético para ilustrar este concepto. Supongamos que la densidad de metilación del tejido tumoral es de aproximadamente 0,45 y la del ADN plasmático de los sujetos sanos es de aproximadamente 0,7. Estos valores supuestos son similares a los obtenidos en nuestro paciente con CHC, donde la densidad de metilación global de los autosomas es del 42,9 % y la densidad de metilación media de los autosomas para las muestras de plasma de controles sanos fue del 71,6 %. Suponiendo que el CV de la medición de la densidad de metilación del ADN plasmático para todo el genoma fuera del 1 %, el valor de corte sería $0,7 \times (100\% - 3 \times 1\%) = 0,679$. Para alcanzar una sensibilidad del 97,5 %, la densidad de metilación media del ADN plasmático de los pacientes con cáncer debe ser aproximadamente $0,679 - 0,7 \times (2 \times 1\%) = 0,665$. Si f es la concentración fraccional de ADN derivado del tumor en la muestra de plasma. Entonces f puede calcularse como $(0,7 - 0,45) \times f = 0,7 - 0,665$. Por lo tanto, f es de aproximadamente el 14 %. A partir de este cálculo, se estima que la concentración fraccional mínima que puede detectarse en el plasma es del 14 % para alcanzar una sensibilidad de diagnóstico del 97,5 % si se utiliza la densidad de metilación total de todo el genoma como parámetro de diagnóstico.

A continuación, realizamos este análisis con los datos obtenidos del paciente con CHC. Para esta ilustración, solo se realizó una medición de la densidad de metilación basada en el valor estimado de todos los autosomas para cada muestra. La densidad media de metilación fue del 71,6 % en las muestras de plasma obtenidas de los sujetos sanos. La DT de las densidades de metilación de estas cuatro muestras fue del 0,631 %. Por lo tanto, el valor de corte para la densidad de metilación plasmática tendría que ser $71,6\% - 3 \times 0,631\% = 69,7\%$ para alcanzar una puntuación z <-3 y una especificidad del 99,85 %. Para alcanzar una sensibilidad del 97,5%, la densidad de metilación plasmática media de los pacientes con cáncer tendría que ser 2 DT por debajo del valor de corte, es decir, del 68,4 %. Dado que la densidad de metilación del tejido tumoral era del 42,9% y utilizando la fórmula: $P = BASAL \times (1-f) + TUM \times f$, tendría que ser al menos del 11,1 %.

En otra realización, las densidades de metilación de diferentes regiones genómicas pueden analizarse por separado, p. ej., como se muestra en las FIGGS. 25A o 26B. Dicho de otro modo, se realizaron múltiples mediciones del nivel de metilación para cada muestra. Como se muestra a continuación, podría detectarse una hipometilación significativa con una concentración fraccional de ADN tumoral en plasma mucho más baja y, por tanto, mejoraría el rendimiento diagnóstico del análisis de metilación del ADN en plasma para la detección del cáncer. Se puede contar el número de regiones genómicas que muestran una desviación significativa en las densidades de metilación con respecto a la población de referencia. A continuación, se puede comparar el número de regiones genómicas con un valor de corte para determinar si existe una hipometilación general significativa del ADN plasmático en toda la población de regiones genómicas estudiadas, por ejemplo, los bins de 1 Mb de todo el genoma. El valor de corte puede establecerse mediante el análisis de un grupo de sujetos de referencia sin cáncer o derivarse matemáticamente, por ejemplo, de acuerdo con la función de distribución normal.

La FIG. 30 es un gráfico 3000 que muestra la distribución de las densidades de metilación del ADN plasmático de sujetos sanos y pacientes con cáncer. La densidad de metilación de cada bin de 1 Mb se compara con los valores correspondientes del grupo de referencia. Se determinó el porcentaje de bins que mostraban una hipometilación significativa (3 DT por debajo de la media del grupo de referencia). Para determinar la presencia de ADN derivado de tumores en la muestra de plasma se utilizó un umbral del 10 % de hipometilación significativa. Otros valores de corte, tales como el 5 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 60 %, 70 %, 80 % o 90 % también se pueden utilizar en función de la sensibilidad y especificidad deseadas de la prueba.

Por ejemplo, para clasificar que una muestra contiene ADN derivado de un tumor, podemos utilizar el 10 % de los bins de 1 Mb que muestran una hipometilación significativa (puntuación z <-3) como valor de corte. Si más del 10 % de los bins están significativamente más hipometilados que el grupo de referencia, entonces la muestra se clasifica como positiva para la prueba del cáncer. Para cada bin de 1 Mb, se utiliza un punto de corte de 3 DT por debajo de la densidad de metilación media del grupo de referencia para definir una muestra como significativamente más hipometilada. Para cada uno de los bins de 1 Mb, si la densidad plasmática media de metilación del ADN de los pacientes con cáncer es 1,72 DT inferior a las densidades plasmáticas medias de metilación del ADN de los sujetos de referencia, entonces hay un 10 % de posibilidades de que el valor de la densidad de metilación de cualquier bin particular de un paciente con cáncer sea inferior al valor de corte (es decir, puntuación z <-3) y dé un resultado positivo. A continuación, si observamos todos los bins de 1 Mb de todo el genoma, entonces cabría esperar que aproximadamente el 10 % de los bins mostraran resultados positivos de tener densidades de metilación significativamente más bajas (es decir, puntuaciones z <-3). Suponiendo que la densidad de metilación total del ADN plasmático de los sujetos sanos es de aproximadamente 0,7 y que el coeficiente de variación (CV) de la medición de la densidad de metilación del ADN plasmático para cada bin de 1 Mb es del 1 %, la densidad de metilación media del ADN plasmático de los pacientes con cáncer tendría que ser de $0,7 \times (100\% - 1,72 \times 1\%) = 0,68796$. Sea/la concentración fraccional de ADN derivado del tumor en plasma para alcanzar esta densidad media de metilación del ADN en plasma. Suponiendo que la densidad de metilación del tejido tumoral sea de 0,45, entonces / puede calcularse mediante la ecuación

$$20 \quad (\bar{M}_{P_{ref}} - \bar{M}_{tumor}) \times f = \bar{M}_{P_{ref}} - \bar{M}_{P_{cancer}}$$

donde $\bar{M}_{P_{ref}}$ representa la densidad media de metilación del ADN plasmático en los individuos de referencia; \bar{M}_{tumor} representa la densidad de metilación del tejido tumoral del paciente con cáncer; y $\bar{M}_{P_{cancer}}$ representa la densidad media de metilación del ADN plasmático en los pacientes con cáncer.

Por medio de esta ecuación, $(0,7-0,45) \times f = 0,7 - 0,68796$. Por tanto, la concentración fraccionaria mínima que puede detectarse utilizando este enfoque se deduciría como del 4,8 %. La sensibilidad puede aumentarse aún más disminuyendo el porcentaje de corte de los bins que están significativamente más hipometilados, por ejemplo, del 10 % al 5 %.

Tal como se muestra en los ejemplos anteriores, la sensibilidad de este método viene determinada por el grado de diferencia en el nivel de metilación entre los tejidos cancerosos y no cancerosos, por ejemplo, glóbulos sanguíneos. En una realización, solo se seleccionan las regiones cromosómicas que muestran una gran diferencia en las densidades de metilación entre el ADN plasmático de los sujetos sin cáncer y el tejido tumoral. En una realización, solo se seleccionan las regiones con una diferencia de densidad de metilación >0,5. En otras realizaciones, puede utilizarse una diferencia de 0,4, 0,6, 0,7, 0,8 o 0,9 para seleccionar las regiones adecuadas. En otras realizaciones más, el tamaño físico de las regiones genómicas no es fijo. En lugar de ello, se definen las regiones genómicas, por ejemplo, basándose en una profundidad de lectura fija o en un número fijo de sitios CpG. Se evalúan los niveles de metilación en varias de estas regiones genómicas para cada muestra.

La FIG. 31 es un gráfico 3100 que muestra la distribución de las diferencias en las densidades de metilación entre la media del ADN plasmático de sujetos sanos y el tejido tumoral del paciente con CHC. Un valor positivo significa que la densidad de metilación es mayor en el ADN plasmático de los sujetos sanos y un valor negativo significa que la densidad de metilación es mayor en el tejido tumoral.

En una realización, se pueden seleccionar los bins con la mayor diferencia entre la densidad de metilación de los tejidos cancerosos y no cancerosos, por ejemplo, aquellos con una diferencia >0,5, independientemente de si el tumor está hipometilado o hipermetilado para estos bins. El límite de detección de la concentración fraccional de ADN derivado de tumores en plasma puede reducirse centrándose en estos bins debido a las mayores diferencias entre las distribuciones de los niveles de metilación del ADN en plasma entre sujetos con cáncer y sin cáncer dada la misma concentración fraccional de ADN derivado de tumores en plasma. Por ejemplo, si solo se utilizan bins con diferencias >0,5 y se adopta un valor de corte del 10 % de los bins significativamente más hipometilados para determinar si un individuo analizado padece un cáncer, la concentración fraccional mínima (f) de ADN derivado del tumor detectado puede calcularse mediante la siguiente ecuación: $(\bar{M}_{P_{ref}} - \bar{M}_{tumor}) \times f = \bar{M}_{P_{ref}} - \bar{M}_{P_{cancer}}$ donde $\bar{M}_{P_{ref}}$ representa la densidad media de metilación del ADN plasmático en los individuos de referencia; \bar{M}_{tumor} representa la densidad de metilación del tejido tumoral del paciente con cáncer; y $\bar{M}_{P_{cancer}}$ representa la densidad media de metilación del ADN plasmático en los pacientes con cáncer.

60 Mientras que la diferencia en la densidad de metilación entre el plasma de los sujetos de referencia y los tejidos tumorales es de al menos 0,5. A continuación, tenemos $0,5 \times f = 0,7 - 0,68796$ y $f = 2,4\%$. Por lo tanto, centrándose en los bins con una mayor diferencia en la densidad de metilación entre los tejidos cancerosos y no cancerosos, el límite inferior de la fracción de ADN derivado del tumor puede reducirse del 4,8 % al 2,4 %. La información relativa a qué bins mostrarían mayores grados de diferencias de metilación entre los tejidos cancerosos y no cancerosos, por

ejemplo, glóbulos sanguíneos, podría determinarse a partir de tejidos tumorales del mismo órgano o del mismo tipo histológico obtenidos de otros individuos.

- 5 En otra realización, un parámetro puede derivarse de la densidad de metilación del ADN plasmático de todos los bins y teniendo en cuenta la diferencia en las densidades de metilación entre tejidos cancerosos y no cancerosos. Los bins con mayor diferencia pueden recibir un peso mayor. En una realización, la diferencia en la densidad de metilación entre el tejido canceroso y no canceroso de cada bin puede utilizarse directamente como el peso del bin concreto para calcular el parámetro final.
- 10 En otra realización más, los distintos tipos de cáncer pueden presentar diferentes patrones de metilación en el tejido tumoral. Del grado de metilación del tipo específico de cáncer puede derivarse un perfil ponderal específico.
- 15 En otra realización más, se puede determinar la relación entre bins de la densidad de metilación en sujetos con y sin cáncer. En la Figura 8, podemos observar que en un pequeño número de bins, los tejidos tumorales estaban más metilados que el ADN plasmático de los sujetos de referencia. Por tanto, se pueden seleccionar los bins con los valores más extremos de diferencia, p. ej., diferencia $>0,5$ y diferencia <0 . La relación de la densidad de metilación de estos bins puede utilizarse entonces para indicar si el individuo analizado tiene cáncer. En otras realizaciones, la diferencia y el cociente de la densidad de metilación de distintos bins pueden utilizarse como parámetros para indicar la relación entre bins.
- 20 Además, evaluamos la sensibilidad de detección del enfoque para detectar o evaluar el tumor utilizando las densidades de metilación de múltiples regiones genómicas, como ilustran los datos obtenidos del paciente con CHC. En primer lugar, mezclamos lecturas del plasma preoperatorio con las obtenidas de las muestras de plasma de los controles sanos para simular muestras de plasma que contuvieran una concentración fraccional de ADN tumoral que oscilara entre el 20 % y el 0,5 %. A continuación, puntuamos el porcentaje de bins de 1 Mb (de 2734 bins en todo el genoma) con densidades de metilación equivalentes a puntuaciones $z <-3$. Cuando la concentración fraccional de ADN tumoral en plasma era del 20 %, el 80,0 % de los bins mostraron una hipometilación significativa. Los datos correspondientes a la concentración fraccional de ADN tumoral en plasma del 10 %, 5 %, 2 %, 1 % y 0,5 % fueron el 67,6 %, 49,7 %, 18,9 %, 3,8 % y 0,77 % de los bins que muestran hipometilación, respectivamente. Dado que el límite teórico del número de bins que presentan puntuaciones $z <-3$ en las muestras de control es del 0,15 %, nuestros datos muestran que todavía había más bins (0,77 %) más allá del valor de corte de corte teórico incluso cuando la concentración fraccional del tumor era de solo el 0,5 %.
- 35 La FIG. 32A es una tabla 3200 que muestra el efecto de reducir la profundidad de secuenciación cuando la muestra de plasma contenía un 5 % o un 2 % de ADN tumoral. Aún se podía detectar una elevada proporción de bins ($>0,15\%$) que mostraban una hipometilación significativa cuando la profundidad media de secuenciación era de solo 0,022 veces el genoma haploide.
- 40 La FIG. 32B es un gráfico 3250 que muestra las densidades de metilación de los elementos repetidos y de las regiones no repetidas en el plasma de los cuatro sujetos de control sanos, la capa leucocitaria, el tejido hepático normal, el tejido tumoral, las muestras de plasma preoperatorio y postoperatorio del paciente con CHC. Puede observarse que los elementos repetidos estaban más metilados (mayor densidad de metilación) que las regiones no repetidas, tanto en los tejidos cancerosos como en los no cancerosos. Sin embargo, la diferencia de metilación entre los elementos repetidos y las regiones no repetidas era mayor en los tejidos no cancerosos y en el ADN plasmático de los sujetos sanos en comparación con los tejidos tumorales.
- 45 Como resultado, el ADN plasmático del paciente con cáncer presentaba una mayor reducción de la densidad de metilación en los elementos repetidos que en las regiones no repetidas. La diferencia en la densidad de metilación del ADN plasmático entre la media de los cuatro controles sanos y el paciente con CHC fue de 0,163 y 0,088 para los elementos repetidos y las regiones no repetidas, respectivamente. Los datos de las muestras de plasma preoperatorias y postoperatorias también mostraron que el intervalo dinámico en el cambio de la densidad de metilación era mayor en las regiones con repeticiones que en las que no las tenían. En una realización, la densidad de metilación del ADN plasmático de los elementos repetidos puede utilizarse para determinar si un paciente está afectado de cáncer o para el seguimiento de la progresión de la enfermedad.
- 55 Como se ha analizado anteriormente, la variación de las densidades de metilación en el plasma de los sujetos de referencia también afectaría a la precisión para diferenciar a los pacientes con cáncer de los individuos sin cáncer. Cuanto más ajustada sea la distribución de las densidades de metilación (es decir, menor desviación típica), más preciso sería diferenciar los sujetos con cáncer de los que no lo tienen. En otra realización, el coeficiente de variación (CV) de las densidades de metilación de los bins de 1 Mb puede utilizarse como criterio para seleccionar los bins con baja variabilidad de las densidades de metilación del ADN plasmático en el grupo de referencia. Por ejemplo, solo se seleccionan los bins con $CV <1\%$. Otros valores, por ejemplo, del 0,5 %, 0,75 %, 1,25 % y 1,5 % también pueden utilizarse como criterios para seleccionar los bins con baja variabilidad en la densidad de metilación. En otra realización más, los criterios de selección pueden incluir tanto el CV del bin como la diferencia en la densidad de metilación entre tejidos cancerosos y no cancerosos.

- La densidad de metilación también puede utilizarse para estimar la concentración fraccional de ADN derivado del tumor en una muestra de plasma cuando se conoce la densidad de metilación del tejido tumoral. Esta información puede obtenerse mediante el análisis del tumor del paciente o del estudio de los tumores de varios pacientes con el mismo tipo de cáncer. Como se ha analizado anteriormente, la densidad de metilación plasmática (P) puede expresarse mediante la siguiente ecuación: $P = BASAL \times (1 - f) + TUM \times f$ donde $BASAL$ es la densidad basal de metilación de los glóbulos sanguíneos y otros órganos, TUM es la densidad de metilación en el tejido tumoral, y f es la concentración fraccional de ADN derivado del tumor en la muestra de plasma. Esto se puede reescribir como:

$$f = \frac{BASAL - P}{BASAL - TUM}$$

- 10 Los valores basales pueden determinarse analizando la muestra de plasma del paciente en un momento en que el cáncer no está presente o a partir del estudio de un grupo de referencia de individuos sin cáncer. Por lo tanto, tras medir la densidad de metilación plasmática, se puede determinar f .

F. Combinación con otros métodos

- 15 Los enfoques de análisis de metilación descritos en el presente documento pueden utilizarse junto con otros métodos basados en los cambios genéticos del ADN derivado de tumores en plasma. Algunos ejemplos de estos métodos son el análisis de aberraciones cromosómicas asociadas al cáncer (K. C. A. Chan *et al.* 2013 *Clin Chem*; 59:211-224; R. J. Leary *et al.* 2012 *Sci Transl Med*; 4:162ra154) y variaciones de nucleótido único asociadas al cáncer en plasma (K. C. A. Chan *et al.* 2013 *Clin Chem*; 59:211-224). El enfoque del análisis de metilación tiene ventajas sobre esos enfoques genéticos.

- 25 Como se muestra en la FIG. 21 A, la hipometilación del ADN tumoral es un fenómeno global que afecta a regiones distribuidas por casi todo el genoma. Por lo tanto, los fragmentos de ADN de todas las regiones cromosómicas serían informativos con respecto a la posible contribución del ADN hipometilado derivado del tumor al ADN del plasma/suero del paciente. Por el contrario, las aberraciones cromosómicas (ya sea amplificación o delección de una región cromosómica) solo están presentes en algunas regiones cromosómicas y los fragmentos de ADN de las regiones sin aberración cromosómica en el tejido tumoral no serían informativos en el análisis (K. C. A. Chan *et al.* 2013 *Clin Chem*; 59: 211-224). Del mismo modo, solo se observan unos pocos miles de alteraciones de un solo nucleótido en cada genoma del cáncer (K. C. A. Chan *et al.* 2013 *Clin Chem*; 59: 211-224). Los fragmentos de ADN que no se solapan con estos cambios de un solo nucleótido no serían informativos para determinar si el ADN derivado del tumor está presente en el plasma. Por lo tanto, este enfoque de análisis de metilación es potencialmente más rentable que esos enfoques genéticos para detectar cambios asociados al cáncer en el sistema circulatorio.

- 30 35 En una realización, la rentabilidad del análisis de metilación del ADN plasmático puede mejorarse aún más enriqueciendo los fragmentos de ADN de las regiones más informativas, por ejemplo, las regiones con mayor diferencia de metilación diferencial entre tejidos cancerosos y no cancerosos. Algunos ejemplos de métodos de enriquecimiento para estas regiones incluyen el uso de sondas de hibridación (p. ej., el sistema Nimblegen SeqCap y el sistema Agilent SureSelect Target Enrichment), Amplificación PCR e hibridación en fase sólida.

G. Análisis de tejidos específicos/donantes

- 40 Las células derivadas del tumor invaden y metastatizan en órganos adyacentes o distantes. Los tejidos invadidos o los focos metastásicos aportan ADN al plasma como resultado de la muerte celular. Al analizar el perfil de metilación del ADN en el plasma de pacientes con cáncer y detectar la presencia de distintivos de metilación específicos de cada tejido, se podrían detectar los tipos de tejidos implicados en el proceso de la enfermedad. Este enfoque proporciona una exploración anatómica no invasiva de los tejidos implicados en el proceso canceroso para ayudar en la identificación de los órganos implicados como sitios primarios y metastásicos. El control de las concentraciones relativas de los distintivos de metilación de los órganos implicados en el plasma también permitiría evaluar la carga tumoral de esos órganos y determinar si el proceso del cáncer en ese órgano se está deteriorando o mejorando o si se había curado. Por ejemplo, si un gen X está específicamente metilado en el hígado. A continuación, se espera que la afectación metastásica del hígado por un cáncer (p. ej., cáncer colorrectal) aumente la concentración de secuencias metiladas del gen X en el plasma. También habría otra secuencia o grupos de secuencias con características de metilación similares a las del gen X. A continuación, se podrían combinar los resultados de dichas secuencias. 55 Consideraciones similares son aplicables a otros tejidos, p. ej., al cerebro, los huesos, pulmones y riñones, etc.

- 50 Por otra parte, Se sabe que el ADN de distintos órganos presenta distintivos de metilación específicos de cada tejido (B. W. Futscher *et al.* 2002 *Nat Genet*; 31:175-179; S. S. C. Chim *et al.* 2008 *Clin Chem*; 54: 500-511). Por tanto, el perfil de metilación en plasma puede utilizarse para dilucidar la contribución de los tejidos de diversos órganos al plasma. La dilucidación de dicha contribución puede utilizarse para evaluar el daño orgánico, ya que se cree que el ADN plasmático se libera cuando las células mueren. Por ejemplo, la patología hepática como la hepatitis (p. ej., por virus, procesos autoinmunitarios, etc.) o la hepatotoxicidad (p. ej., sobredosis de fármacos (tal como por paracetamol) o toxinas (tal como el alcohol) causada por fármacos se asocia con daño celular hepático y se esperará que se asocie

con un mayor nivel de ADN derivado del hígado en plasma. Por ejemplo, si un gen X está específicamente metilado en el hígado. A continuación, se espera que la patología hepática aumente la concentración de secuencias metiladas del gen X en el plasma. Por el contrario, si un gen Y está específicamente hipometilado en el hígado. A continuación, se espera que la patología hepática disminuya la concentración de secuencias metiladas del gen Y en el plasma. En otra realización más, el gen X o Y puede ser sustituido por cualquier secuencia genómica que no sea un gen y que presente una metilación diferencial en los distintos tejidos del organismo.

Las técnicas descritas en el presente documento también podrían aplicarse a la evaluación del ADN derivado del donante en el plasma de los receptores de trasplantes de órganos (Y. M. D. Lo *et al.* 1998 *Lancet*; 351:1329-1330).

Se han utilizado diferencias polimórficas entre el donante y el receptor para distinguir el ADN derivado del donante del ADN derivado del receptor en plasma (Y. W. Zheng *et al.* 2012 *Clin Chem*; 58: 549-558). Proponemos que los distintivos de metilación específicos del tejido del órgano transplantado también podrían utilizarse como método para detectar el ADN del donante en el plasma del receptor.

Al controlar la concentración de ADN del donante, se podría evaluar de forma no invasiva el estado del órgano transplantado. Por ejemplo, el rechazo del trasplante se asocia a una mayor tasa de muerte celular y, por tanto, de concentración de ADN del donante en el plasma (o suero) del receptor, como refleja el distintivo de metilación del órgano transplantado, aumentaría en comparación con el tiempo en que el paciente se encuentra en situación estable o en comparación con otros receptores de trasplantes estables o controles sanos sin trasplante. De manera similar a lo descrito para el cáncer, el ADN derivado del donante podría identificarse en el plasma de los receptores de trasplantes mediante la detección de todos o algunos de los rasgos característicos, incluidas las diferencias polimórficas, el ADN de menor tamaño para los órganos sólidos transplantados (Y. W. Zheng *et al.* 2012 *Clin Chem*, 58: 549-558) y el perfil de metilación específico de cada tejido.

H. Normalización de la metilación en función del tamaño

Como se ha descrito anteriormente y por Lun *et al.* (F. M. F. Lun *et al.* *Clin. Chem.* 2013; doi: 10.1373/clinchem.2013.212274), la densidad de metilación (p. ej., del ADN plasmático) está correlacionada con el tamaño de los fragmentos de ADN. La distribución de las densidades de metilación de los fragmentos de ADN plasmático más cortos fue significativamente inferior a la de los fragmentos más largos. Proponemos que algunas afecciones no cancerosas (p. ej., el lupus eritematoso sistémico [LES]) con patrones de fragmentación anormales del ADN plasmático pueden mostrar una aparente hipometilación del ADN plasmático debido a la presencia de fragmentos cortos de ADN plasmático más abundantes, que están menos metilados. Dicho de otro modo, la distribución del tamaño del ADN plasmático puede ser un factor de confusión para la densidad de metilación del ADN plasmático.

La FIG. 34A muestra una distribución de tamaños del ADN plasmático en el paciente con LES LES04. Las distribuciones de tamaño de nueve sujetos de control sanos se muestran como líneas grises punteadas y la de LES04 se muestra como línea sólida negra. Los fragmentos cortos de ADN plasmático fueron más abundantes en LES04 que en los nueve sujetos de control sanos. Dado que los fragmentos de ADN más cortos suelen estar menos metilados, este patrón de distribución de tamaños puede confundir el análisis de metilación para el ADN plasmático y dar lugar a una hipometilación más aparente.

En algunas realizaciones, un nivel de metilación medido puede normalizarse para reducir el efecto de confusión de la distribución de tamaños en el análisis de metilación del ADN plasmático. Por ejemplo, se puede medir el tamaño de las moléculas de ADN en la pluralidad de sitios. En diversas implementaciones, la medición puede proporcionar un tamaño específico (p. ej., la longitud) a una molécula de ADN o simplemente determinar que el tamaño se encuentra dentro de un intervalo específico, que también puede corresponder a un tamaño. El nivel de metilación normalizado puede entonces compararse con un valor de corte. Existen varias formas de realizar la normalización para reducir el efecto de confusión de la distribución de tamaños en el análisis de metilación del ADN plasmático.

En una realización, puede realizarse el fraccionamiento por tamaño del ADN (p. ej., ADN plasmático). El fraccionamiento por tamaño puede garantizar que se utilicen fragmentos de ADN de un tamaño similar para determinar el nivel de metilación de forma coherente con el valor de corte. Como parte del fraccionamiento por tamaño, pueden seleccionarse fragmentos de ADN que tengan un primer tamaño (p. ej., un primer intervalo de longitudes), donde el primer valor de corte corresponde al primer tamaño. La normalización puede lograrse calculando el nivel de metilación utilizando únicamente los fragmentos de ADN seleccionados.

El fraccionamiento por tamaños puede lograrse de varias maneras, p. ej., bien mediante la separación física de moléculas de ADN de diferentes tamaños (p. ej., mediante electroforesis o tecnologías basadas en la microfluídica, o tecnologías basadas en la centrifugación), bien mediante análisis informáticos. Para los análisis informáticos, en una realización, se puede realizar la secuenciación masiva en paralelo de ambos extremos de las moléculas de ADN plasmático. A continuación, se puede deducir el tamaño de las moléculas secuenciadas por comparación con la localización de cada uno de los dos extremos de una molécula de ADN plasmático con un genoma humano de referencia. A continuación, se puede realizar un análisis posterior mediante la selección de moléculas de ADN secuenciadas que coincidan con uno o más criterios de selección de tamaño (p. ej., el criterio de que el tamaño esté

dentro de un intervalo especificado). Por tanto, en una realización, la densidad de metilación puede analizarse para fragmentos con un tamaño similar (p. ej., dentro de un intervalo especificado). El valor de corte (p. ej., en el bloque 2840 del método 2800) puede determinarse basándose en fragmentos dentro del mismo intervalo de tamaño. A modo de ejemplo, los niveles de metilación pueden determinarse a partir de muestras que se sabe que tienen cáncer o que no lo tienen, y los valores de corte pueden determinarse a partir de estos niveles de metilación.

En otra realización, se puede determinar una relación funcional entre la densidad de metilación y el tamaño del ADN circulante. La relación funcional puede definirse mediante puntos de datos o coeficientes de una función. La relación funcional puede proporcionar valores de escala correspondientes a los tamaños respectivos (p. ej., los tamaños más cortos pueden tener aumentos correspondientes a la metilación). En diversas implementaciones, el valor de escala puede estar comprendido entre 0 y 1 o ser superior a 1.

La normalización puede hacerse a partir de un tamaño promedio. Por ejemplo, puede calcularse un tamaño promedio correspondiente a las moléculas de ADN utilizadas para calcular el primer nivel de metilación, y el primer nivel de metilación puede multiplicarse por el valor de escala correspondiente (es decir, correspondiente al tamaño promedio). Como otro ejemplo, la densidad de metilación de cada molécula de ADN puede normalizarse en función del tamaño de la molécula de ADN y de la relación entre el tamaño del ADN y la metilación.

En otra implementación, la normalización puede hacerse por molécula. Por ejemplo, se puede obtener un tamaño respectivo de una molécula de ADN en un sitio particular (p. ej., como se ha descrito anteriormente), y se puede identificar un valor de escala correspondiente al tamaño respectivo a partir de la relación funcional. Para un cálculo no normalizado, cada molécula se contaría por igual a la hora de determinar un índice de metilación en el sitio. Para el cálculo normalizado, la contribución de una molécula al índice de metilación puede ponderarse mediante el factor de escala correspondiente al tamaño de la molécula.

Las FIGS. 34B y 34C muestran análisis de metilación para ADN plasmático de un paciente con LES LES04 (FIG. 34B) y un paciente con CHC TBR36 (FIG. 34C). Los círculos exteriores muestran los resultados de Z_{met} para el ADN plasmático sin fraccionamiento de tamaños informático. Los círculos interiores muestran los resultados de Z_{met} para el ADN plasmático de 130 pb o más. Para el paciente con LES LES04, el 84 % de los bins mostraron hipometilación sin fraccionamiento de tamaños informático. El porcentaje de los bins que mostraban hipometilación se redujo al 15 % cuando solo se analizaron fragmentos de 130 pb o más. Para el paciente con CHC TBR36, el 98,5 % y el 98,6 % de los bins mostraron hipometilación para el ADN plasmático con y sin fraccionamiento de tamaños informático, respectivamente. Estos resultados indican que el fraccionamiento de tamaños informático puede reducir eficazmente los resultados falsos positivos de hipometilación relacionados con el aumento de la fragmentación del ADN plasmático, p. ej., en pacientes con LES o en otras afecciones inflamatorias.

En una realización, los resultados de los análisis con y sin fraccionamiento por tamaño pueden compararse para indicar si existe algún efecto de confusión del tamaño en los resultados de metilación. Por tanto, además o en lugar de la normalización, el cálculo de un nivel de metilación en un tamaño concreto puede utilizarse para determinar si existe la probabilidad de un falso positivo cuando el porcentaje de bins por encima de un valor de corte difiere con y sin fraccionamiento por tamaño, o si solo difiere un nivel de metilación concreto.

Por ejemplo, la presencia de una diferencia significativa entre los resultados de las muestras con y sin fraccionamiento por tamaño puede utilizarse para indicar la posibilidad de un resultado falso positivo debido a un patrón de fragmentación anormal. El umbral para determinar si la diferencia es significativa puede establecerse mediante el análisis de una cohorte de pacientes con cáncer y una cohorte de sujetos de control sin cáncer.

I. Análisis de la hipometilación de las islas CpG del genoma en plasma

Además de la hipometilación general, también se observa con frecuencia la hipometilación de las islas CpG en los cánceres (S. B. Baylin *et al.* 2011 *Nat Rev Cancer*, 11: 726-734; P. A. Jones *et al.* 2007, *Cell*, 128: 683-692; M. Esteller *et al.* 2007 *Nat Rev Genet* 2007; 8: 286-298; M. Ehrlich *et al.* 2002 *Oncogene* 2002; 21: 5400-5413). En esta sección, describimos el uso del análisis genómico de la hipometilación de islas CpG para la detección y seguimiento de cánceres.

La FIG. 35 es un diagrama de flujo de un método 3500 que determina una clasificación de un nivel de cáncer basado en la hipometilación de islas CpG de acuerdo con realizaciones de la presente invención. La pluralidad de sitios del método 2800 puede incluir sitios CpG, en donde los sitios CpG se organizan en una pluralidad de islas CpG, cada isla CpG incluye uno o más sitios CpG. Los niveles de metilación de cada isla CpG pueden utilizarse para determinar la clasificación del nivel de cáncer.

En el bloque 3510, se identifican las islas CpG que se van a analizar. En este análisis, como ejemplo, primero definimos un conjunto de islas CpG para analizar, que se caracterizan por densidades de metilación relativamente bajas en el plasma de los sujetos sanos de referencia. En un aspecto, la variación de las densidades de metilación en el grupo de referencia puede ser relativamente pequeña para permitir detectar más fácilmente la hipometilación asociada al cáncer. En una realización, las islas CpG tienen una densidad de metilación media inferior a un primer porcentaje en

un grupo de referencia, y un coeficiente de variación de la densidad de metilación en el grupo de referencia es inferior a un segundo porcentaje.

Como ejemplo, con fines ilustrativos, para la identificación de las islas CpG útiles se utilizan los siguientes criterios:

- 5 i. La densidad media de metilación para la isla CpG en el grupo de referencia (p. ej., sujetos sanos) <5 %
- ii. El coeficiente de variación para el análisis de la densidad de metilación en plasma para el grupo de referencia (p. ej., sujetos sanos) <30 %.

Estos parámetros pueden ajustarse para una aplicación específica. De nuestro conjunto de datos, 454 islas CpG del genoma cumplían estos criterios.

10 En el bloque 3520, se calcula la densidad de metilación para cada isla CpG. Las densidades de metilación se pueden determinar, como se describe en el presente documento.

15 En el bloque 3530, se determina si cada una de las islas CpG está hipermetilada. Por ejemplo, para el análisis de la hipermetilación de la isla CpG de un caso analizado, se comparó la densidad de metilación de cada isla CpG con los datos correspondientes de un grupo de referencia. La densidad de metilación (un ejemplo de nivel de metilación) puede compararse con uno o más valores de corte para determinar si una isla concreta está hipermetilada.

20 En una realización, un primer valor de corte puede corresponder a una media de densidades de metilación para el grupo de referencia más un porcentaje especificado. Otro valor de corte puede corresponder a la media de las densidades de metilación del grupo de referencia más un número determinado de desviaciones típicas. En una implementación, se calculó una puntuación z (Z_{met}) y se comparó con los valores de corte. Como ejemplo, una isla CpG en un sujeto de ensayo (p. ej., un sujeto sometido a un cribado de cáncer) se consideró significativamente hipermetilada si cumplía los siguientes criterios:

- 25 i. su densidad de metilación era superior a la media del grupo de referencia en un 2 %, y
- ii. $Z_{met} > 3$

Estos parámetros también pueden ajustarse para una aplicación específica.

30 En el bloque 3540, las densidades de metilación (p. ej., como puntuaciones z) de las islas CpG hipermetiladas se utilizan para determinar una puntuación acumulada. Por ejemplo, tras la identificación de todas las islas CpG significativamente hipermetiladas, se puede calcular una puntuación que implique una suma de puntuaciones z o funciones de puntuaciones z de todas las islas CpG hipermetiladas. Un ejemplo de puntuación es la puntuación de probabilidad acumulada (PA), como se describe en otra sección. La puntuación de probabilidad acumulada utiliza Z_{met} para determinar la probabilidad de que se produzca dicha observación por azar de acuerdo con una distribución de probabilidad (p. ej., la distribución de probabilidad t de Student con 3 grados de libertad).

35 En el bloque 3550, la puntuación acumulada se compara con un umbral acumulado para determinar una clasificación de un nivel de cáncer. Por ejemplo, si la hipermetilación total en las islas CpG identificadas es lo suficientemente grande, entonces el organismo puede ser identificado como enfermo de cáncer. En una realización, el umbral acumulado corresponde a la puntuación acumulada más alta del grupo de referencia.

IX. METILACIÓN Y ANC

40 Como se ha mencionado anteriormente, los enfoques de análisis de metilación descritos en el presente documento pueden utilizarse junto con otros métodos basados en los cambios genéticos del ADN derivado de tumores en plasma. Algunos ejemplos de estos métodos son el análisis de aberraciones cromosómicas asociadas al cáncer (K. C. A. Chan *et al.* 2013 *Clin Chem*; 59: 211-224; R. J. Leary *et al.* 2012 *Sci Transl Med*; 4: 162ral54). En la solicitud de patente de EE. UU. n.º 13/308.473 se describen algunos aspectos de las aberraciones del número de copias (ANC).

50 A. ANC

Las aberraciones del número de copias pueden detectarse contando los fragmentos de ADN que se alinean con una parte concreta del genoma, normalizando el número y comparando el número con un valor de corte. En diversas realizaciones, la normalización puede realizarse mediante un número de fragmentos de ADN alineados con otro haplotipo de la misma parte del genoma (dosificación relativa de haplotipos (RHDO, *Relative Haplotype DOsage*)) o mediante un número de fragmentos de ADN alineados con otra parte del genoma.

55 El método RHDO se basa en el uso de locus heterocigóticos. Las realizaciones descritas en esta sección también pueden utilizarse para locus que son homocigóticos comparando dos regiones y no dos haplotipos de la misma región, y por lo tanto no son específicos del haplotipo. En un método de dosificación relativa de regiones cromosómicas, se compara el número de fragmentos de una región cromosómica (p. ej., determinado por el número de las lecturas de secuencia alineadas con esa región) con un valor esperado (que puede ser de una región cromosómica de referencia o de la misma región en otra muestra que se sabe que está sana). De esta manera, un fragmento se contabilizaría para una región cromosómica independientemente del haplotipo al que pertenezca la etiqueta secuenciada. Por tanto, que no contengan locus heterocigóticos. Para realizar la comparación, una realización puede normalizar el número de etiquetas antes de la comparación. Cada región está definida por al menos dos locus (que están separados entre sí),

y los fragmentos de estos locus pueden utilizarse para obtener un valor colectivo sobre la región.

Se puede calcular un valor normalizado de las lecturas secuenciadas (etiquetas) para una región concreta dividiendo el número de lecturas secuenciadas que se alinean con esa región entre el número total de lecturas secuenciadas

- 5 alineables con todo el genoma. Este número normalizado de etiquetas permite comparar los resultados de una muestra con los de otra. Por ejemplo, el valor normalizado puede ser la proporción (p. ej., porcentaje o fracción) de lecturas secuenciadas que se espera que procedan de la región concreta, como se ha indicado anteriormente. En otras realizaciones, son posibles otros métodos de normalización. Por ejemplo, se puede normalizar dividiendo el número de recuentos de una región entre el número de recuentos de una región de referencia (en el caso anterior, la región de referencia es todo el genoma). A continuación, este número normalizado de etiquetas puede compararse con un valor umbral, que puede determinarse a partir de una o más muestras de referencia que no presenten cáncer.

A continuación, el número normalizado de etiquetas del caso analizado se compararía con el número normalizado de etiquetas de uno o varios sujetos de referencia, p. ej., los que no tienen cáncer. En una realización, la comparación se

- 15 realiza calculando la puntuación z del caso para la región cromosómica concreta. La puntuación z puede calcularse mediante la siguiente ecuación: puntuación z = (número normalizado de etiquetas del caso - media)/DT, donde "media" es el número medio normalizado de etiquetas alineadas con la región cromosómica concreta para las muestras de referencia; y DT es la desviación típica del número de números normalizados de etiquetas que se alinean con la región concreta para las muestras de referencia. Por consiguiente, la puntuación z es el número de desviaciones típicas que

20 el número normalizado de etiquetas de una región cromosómica para el caso analizado se aleja de la media del número normalizado de etiquetas para la misma región cromosómica de uno o más sujetos de referencia.

En la situación en la que el organismo analizado tiene cáncer, las regiones cromosómicas amplificadas en los tejidos tumorales estarían sobrerepresentadas en el ADN plasmático. Esto daría lugar a un valor positivo de la puntuación z.

- 25 Por otra parte, las regiones cromosómicas suprimidas en los tejidos tumorales estarían infrarrepresentadas en el ADN plasmático. Esto daría lugar a un valor negativo de la puntuación z. La magnitud de la puntuación z viene determinada por varios factores.

30 Uno de los factores es la concentración fraccional de ADN derivado del tumor en la muestra biológica (p. ej., plasma).

- Cuanto mayor sea la concentración fraccional de ADN derivado del tumor en la muestra (p. ej., plasma), mayor será la diferencia entre el número normalizado de etiquetas del caso probado y los casos de referencia. Por consiguiente, se obtendría una mayor magnitud de la puntuación z.

35 Otro factor es la variación del número normalizado de etiquetas en uno o varios casos de referencia. Con el mismo grado de sobrerepresentación de la región cromosómica en la muestra biológica (p. ej., plasma) del caso analizado, una variación menor (es decir, una desviación típica menor) del número normalizado de marcas en el grupo de referencia daría lugar a una puntuación z mayor. De manera similar, con el mismo grado de infrarrepresentación de la región cromosómica en la muestra biológica (p. ej., plasma) del caso analizado, una desviación típica menor del número normalizado de marcas en el grupo de referencia daría lugar a una puntuación z más negativa.

- 40 40 Otro factor es la magnitud de la aberración cromosómica en los tejidos tumorales. La magnitud de la aberración cromosómica se refiere a los cambios en el número de copias de la región cromosómica concreta (ganancia o pérdida). Cuanto mayores sean los cambios en el número de copias en los tejidos tumorales, mayor será el grado de sobrerepresentación o infrarrepresentación de la región cromosómica concreta en el ADN plasmático. Por ejemplo,

- 45 la pérdida de ambas copias del cromosoma daría lugar a una mayor infrarrepresentación de la región cromosómica en el ADN plasmático que la pérdida de una de las dos copias del cromosoma y, por consiguiente, dio lugar a una puntuación z más negativa. Habitualmente, existen múltiples aberraciones cromosómicas en los cánceres. Las aberraciones cromosómicas de cada cáncer pueden variar además por su naturaleza (es decir, amplificación o delección), su grado (ganancia o pérdida de una o varias copias) y su extensión (tamaño de la aberración en términos de longitud cromosómica).

50 La precisión de la medición del número normalizado de etiquetas se ve afectada por el número de moléculas analizadas. Prevemos que sería necesario analizar 15.000, 60.000 y 240.000 moléculas para detectar aberraciones cromosómicas con un cambio de copia (ya sea ganancia o pérdida) cuando la concentración fraccional es de

- 55 aproximadamente el 12,5 %, 6,3 % y 3,2 %, respectivamente. En la publicación de patente de EE. UU. n.º 2009/0029377 titulada "Diagnosing Fetal Chromosomal Aneuploidy Using Massively Parallel Genomic Sequencing" de Lo *et al.* se describen más detalles sobre el número de etiquetas para la detección del cáncer en diferentes regiones cromosómicas.

- 60 60 Las realizaciones también pueden utilizar el análisis de tamaño, en lugar del método de número de etiquetas. También puede utilizarse el análisis de tamaño, en lugar de un número normalizado de etiquetas. El análisis del tamaño puede utilizar varios parámetros, como se menciona en el presente documento, y en la solicitud de patente de EE. UU. n.º 12/940.992. Por ejemplo, pueden utilizarse los valores Q o F anteriores. Estos valores de tamaño no necesitan una normalización por números de otras regiones, ya que estos valores no escalan con el número de lecturas. También se pueden utilizar técnicas de los métodos haplotipo-específicos, tales como el método RHDO descrito anteriormente y con más detalle en la solicitud de patente de EE.UU. n.º 13/308.473, para los métodos no

específicos. Por ejemplo, se pueden utilizar técnicas que implican la profundidad y el perfeccionamiento de una región. En algunas realizaciones, al comparar dos regiones, puede tenerse en cuenta el sesgo de GC de una región concreta. Dado que el método RHDO utiliza la misma región, tal corrección no es necesaria.

- 5 Aunque algunos tipos de cáncer pueden presentar aberraciones en determinadas regiones cromosómicas, estos cánceres no siempre presentan exclusivamente aberraciones en dichas regiones. Por ejemplo, otras regiones cromosómicas podrían mostrar aberraciones, y la localización de dichas regiones adicionales puede ser desconocida. De manera adicional, al examinar a los pacientes para identificar las primeras fases del cáncer, se puede querer identificar una amplia selección de cánceres, que podría mostrar aberraciones presentes en todo el genoma. Para hacer frente a estas situaciones, pueden analizar una pluralidad de regiones de forma sistemática para determinar qué regiones muestran aberraciones. Se puede utilizar el número de aberraciones y su localización (p. ej., si son contiguas), por ejemplo, para confirmar las aberraciones, determinar el estadio del cáncer, proporcionar un diagnóstico de cáncer (p. ej., si el número es superior a un valor umbral), y proporcionar un pronóstico basado en el número y la ubicación de varias regiones que presentan una aberración.
- 10 15 En consecuencia, algunas realizaciones pueden identificar si un organismo tiene cáncer basándose en el número de regiones que muestran una aberración. Por tanto, se puede probar una pluralidad de regiones (p. ej., 3000) para identificar un número de regiones que presentan una aberración. Las regiones pueden abarcar todo el genoma o solo partes de él, p. ej., región no repetida.
- 20 La FIG. 36 es un diagrama de flujo que ilustra un método 3600 de análisis de una muestra biológica de un organismo para identificar el origen de una aberración cromosómica de acuerdo con realizaciones de la presente invención. La muestra biológica incluye moléculas de ácido nucleico (también llamadas fragmentos).
- 25 30 35 40 45 50 55 60 65
- En el bloque 3610, se identifica una pluralidad de regiones (p. ej., regiones no solapadas) del genoma del organismo. Cada región cromosómica incluye una pluralidad de locus. Una región puede tener un tamaño de 1 Mb o cualquier otro de igual tamaño. Para la situación de una región de 1 Mb de tamaño, el genoma completo puede incluir entonces aproximadamente 3000 regiones, cada una de tamaño y ubicación predeterminados. Dichas regiones predeterminadas pueden variar para adaptarse a la longitud de un cromosoma concreto o a un número especificado de regiones que se vayan a utilizar, así como a cualquier otro criterio mencionado en el presente documento. Si las regiones tienen longitudes diferentes, estas longitudes pueden utilizarse para normalizar los resultados, p. ej., como se describe en el presente documento. Las regiones pueden seleccionarse específicamente en función de determinados criterios del organismo específico y/o en función del conocimiento del cáncer que se está analizando. Las regiones también pueden seleccionarse arbitrariamente.
- En el bloque 3620, se identifica una localización de la molécula de ácido nucleico en un genoma de referencia del organismo para cada una de una pluralidad de moléculas de ácido nucleico. La ubicación puede determinarse de cualquiera de las formas mencionadas, p. ej., secuenciando los fragmentos para obtener etiquetas secuenciadas y alineando las etiquetas secuenciadas con el genoma de referencia. También se puede determinar el haplotipo particular de una molécula para los métodos específicos de haplotipo.
- Los bloques 3630-3650 se realizan para cada una de las regiones cromosómicas. En el bloque 3630, se identifica un grupo respectivo de moléculas de ácido nucleico como perteneciente a la región cromosómica basándose en las ubicaciones identificadas. El grupo respectivo puede incluir al menos una molécula de ácido nucleico ubicada en cada uno de la pluralidad de locus de la región cromosómica. En una realización, el grupo pueden ser fragmentos que se alinean con un haplotipo concreto de la región cromosómica, p. ej., como en el método RHDO anterior. En otra realización, el grupo puede ser de cualquier fragmento que se alinee con la región cromosómica.
- En el bloque 3640, un sistema informático calcula un valor respectivo del grupo respectivo de moléculas de ácido nucleico. El valor respectivo define una propiedad de las moléculas de ácido nucleico del grupo respectivo. El valor respectivo puede ser cualquiera de los valores aquí mencionados. Por ejemplo, el valor puede ser el número de fragmentos del grupo o un valor estadístico de una distribución de tamaños de los fragmentos del grupo. El valor respectivo también puede ser un valor normalizado, p. ej., un número de etiquetas de la región dividido entre el número total de números de etiquetas de la muestra o el número de números de etiquetas de una región de referencia. El valor respectivo también puede ser una diferencia o una proporción con respecto a otro valor (p. ej., en RHDO), de la región.
- En el bloque 3650, el valor respectivo se compara con un valor de referencia para determinar una clasificación de si la primera región cromosómica presenta una delección o una amplificación. Este valor de referencia puede ser cualquier umbral o valor de referencia descrito en el presente documento. Por ejemplo, el valor de referencia podría ser un valor umbral determinado para muestras normales. Para RHDO, el valor respectivo podría ser la diferencia o la proporción de números de marcas para los dos haplotipos, y el valor de referencia puede ser un umbral para determinar que existe una desviación estadísticamente significativa. Como otro ejemplo, el valor de referencia podría ser el número de etiquetas o el valor de tamaño para otro haplotipo o región, y la comparación puede incluir tomar una diferencia o una proporción (o función de los mismos) y luego determinar si la diferencia o proporción es mayor que un valor umbral.
- El valor de referencia puede variar en función de los resultados de otras regiones. Por ejemplo, si las regiones vecinas

- también muestran una desviación (aunque sea pequeña en comparación con un umbral, p. ej., una puntuación z de 3), puede utilizarse un umbral más bajo. Por ejemplo, si tres regiones consecutivas superan un primer umbral, entonces el cáncer puede ser más probable. Por tanto, este primer umbral puede ser inferior a otro umbral necesario para identificar el cáncer de regiones no consecutivas. Tener tres regiones (o más de tres) que presenten incluso una pequeña desviación puede tener una probabilidad de efecto fortuito lo suficientemente baja como para que se conserven la sensibilidad y la especificidad.
- En el bloque 3660, se determina una cantidad de regiones genómicas clasificadas como que presentan una delección o una amplificación. Las regiones cromosómicas que se cuentan pueden tener restricciones. Por ejemplo, solo pueden contarse las regiones contiguas con al menos otra región (o puede exigirse que las regiones contiguas tengan un tamaño determinado, p. ej., 4 o más regiones). Para las realizaciones en las que las regiones no son iguales, el número también puede representar las longitudes respectivas (p. ej., el número podría ser una longitud total de las regiones aberrantes).
- En el bloque 3670, la cantidad se compara con un valor umbral de cantidad para determinar una clasificación de la muestra. Como ejemplos, la clasificación puede ser si el organismo tiene cáncer, un estadio del cáncer y un pronóstico del mismo. En una realización, se cuentan todas las regiones aberrantes y se utiliza un único valor umbral independientemente de dónde aparezcan las regiones. En otra realización, un valor umbral puede variar en función de las ubicaciones y del tamaño de las regiones que se cuentan. Por ejemplo, se puede comparar la cantidad de regiones en un cromosoma o brazo de un cromosoma en particular con un umbral para ese cromosoma (o brazo) en particular. Se pueden utilizar varios umbrales. A modo de ejemplo, la cantidad de regiones aberrantes en un cromosoma (o brazo) concreto debe ser superior a un primer valor umbral, y la cantidad total de regiones aberrantes en el genoma debe ser superior a un segundo valor umbral. El valor umbral puede ser un porcentaje de las regiones que se determina que presentan una delección o una amplificación.
- Este valor umbral para la cantidad de regiones también puede depender de lo fuerte que sea el desequilibrio de las regiones contadas. Por ejemplo, la cantidad de regiones que se utilizan como umbral para determinar una clasificación de cáncer puede depender de la especificidad y la sensibilidad (umbral aberrante) utilizadas para detectar una aberración en cada región. Por ejemplo, si el umbral aberrante es bajo (p. ej., puntuación z de 2), entonces el umbral de cantidad puede seleccionarse para que sea alto (p. ej., 150). Sin embargo, si el umbral aberrante es alto (p. ej., una puntuación z de 3), el umbral de cantidad puede ser más bajo (p. ej., 50). La cantidad de regiones que muestran una aberración también puede ser un valor ponderado, p. ej., una región que muestre un gran desequilibrio puede tener una ponderación más alta que una región que solo muestre un pequeño desequilibrio (es decir, hay más clasificaciones que solo positiva y negativa para la aberración). Como ejemplo, se puede utilizar una suma de puntuaciones z, utilizando así los valores ponderados.
- En consecuencia, la cantidad (que puede incluir el número y/o el tamaño) de regiones cromosómicas que muestren una sobrerepresentación o infrarrepresentación significativa de un número normalizado de etiquetas (u otro valor respectivo para la propiedad del grupo) puede utilizarse para reflejar la gravedad de la enfermedad. La cantidad de regiones cromosómicas con un número de etiquetas normalizado aberrante puede determinarse mediante dos factores, en concreto, el número (o tamaño) de aberraciones cromosómicas en los tejidos tumorales y la concentración fraccional de ADN derivado del tumor en la muestra biológica (p. ej., plasma). Los cánceres más avanzados tienden a presentar más aberraciones cromosómicas (y de mayor tamaño). Por consiguiente, más aberraciones cromosómicas asociadas al cáncer serían potencialmente detectables en la muestra (p. ej., plasma). En pacientes con cáncer más avanzado, la mayor carga tumoral conduciría a una mayor concentración fraccional de ADN derivado del tumor en el plasma. Como resultado, las aberraciones cromosómicas asociadas al tumor se detectarían más fácilmente en la muestra de plasma.
- Un posible enfoque para mejorar la sensibilidad sin sacrificar la especificidad es tener en cuenta el resultado del segmento cromosómico adyacente. En una realización, el valor de corte para la puntuación z sigue siendo >2 y <-2 . Sin embargo, una región cromosómica se clasificaría como potencialmente aberrante solo cuando dos segmentos consecutivos mostraran el mismo tipo de aberraciones, p. ej., ambos segmentos tuvieran una puntuación z >2 . En otras realizaciones, la puntuación z de los segmentos vecinos puede sumarse utilizando un valor de corte más alto. Por ejemplo, se pueden sumar las puntuaciones z de tres segmentos consecutivos y utilizar un valor de corte de 5. Este concepto puede ampliarse a más de tres segmentos consecutivos.
- La combinación de umbrales de cantidad y aberración también puede depender de la finalidad del análisis y de cualquier conocimiento previo del organismo (o falta del mismo). Por ejemplo, si se realiza un cribado de cáncer a una población sana normal, entonces uno normalmente usaría alta especificidad, potencialmente tanto en la cantidad de regiones (es decir, un umbral alto para el número de regiones) como un umbral aberrante para cuando se identifica que una región tiene una aberración. Sin embargo, en un paciente con mayor riesgo (p. ej., un paciente que se queja de un bulto o de antecedentes familiares, fumador, portador crónico del virus del papiloma humano (VPH), portador del virus de la hepatitis o portador de otro virus) entonces los umbrales podrían ser más bajos para tener más sensibilidad (menos falsos negativos).
- En una realización, si se utiliza una resolución de 1 Mb y un límite de detección inferior del 6,3 % del ADN derivado

del tumor para detectar una aberración cromosómica, el número de moléculas en cada segmento de 1 Mb tendría que ser de 60.000. Esto se traduciría en aproximadamente 180 millones (60.000 lecturas/Mb x 3000 Mb) de lecturas alineables para todo el genoma.

- 5 Un tamaño de segmento más pequeño proporcionaría una mayor resolución para detectar aberraciones cromosómicas más pequeñas. Sin embargo, esto aumentaría la exigencia del número de moléculas que se van a analizar en total. Un mayor tamaño del segmento reduciría el número de moléculas necesarias para el análisis a expensas de la resolución. Por lo tanto, solo pueden detectarse las aberraciones de mayor tamaño. En una implementación, podrían utilizarse regiones más extensas, los segmentos que muestran una aberración podrían subdividirse y estas 10 subregiones analizarse para obtener una mejor resolución (p. ej., como se ha descrito anteriormente). Si se dispone de una estimación del tamaño de la delección o amplificación que debe detectarse (o de la concentración mínima que debe detectarse), se puede determinar el número de moléculas que hay que analizar.

B. ANC basadas en la secuenciación del ADN plasmático tratado con bisulfito

- 15 La hipometilación genómica y las ANC pueden observarse con frecuencia en los tejidos tumorales. En el presente documento, demostramos que la información de las ANC y los cambios de metilación asociados al cáncer pueden obtenerse simultáneamente a partir de la secuenciación con bisulfito del ADN plasmático. Como los dos tipos de análisis pueden realizarse con el mismo conjunto de datos, prácticamente no hay coste adicional para el análisis de 20 las ANC. Otras realizaciones pueden utilizar procedimientos diferentes para obtener la información de metilación y la información genética. En otras realizaciones, se puede realizar un análisis similar para la hipermetilación asociada al cáncer junto con el análisis de las ANC.

- 25 La FIG. 37A muestra el análisis de las ANC para tejidos tumorales, ADN plasmático no tratado con bisulfito (BS) y ADN plasmático tratado con bisulfito (de dentro a fuera) para el paciente TBR36. La FIG. 37A muestra el análisis de las ANC para tejidos tumorales, ADN plasmático no tratado con bisulfito (BS) y ADN plasmático tratado con bisulfito (de dentro a fuera) para el paciente TBR36. El anillo más externo muestra el ideograma cromosómico. Cada punto representa el resultado de una región de 1 Mb. Los puntos verdes, rojos y grises representan regiones con ganancia de número de copias, pérdida de número de copias y sin cambio en el número de copias, respectivamente. Para el 30 análisis de plasma, se muestran las puntuaciones z. Entre dos líneas concéntricas hay una diferencia de 5. Para el análisis de tejido tumoral, se muestra el número de copia. Hay una diferencia de copias entre dos líneas concéntricas. La FIG. 38A muestra el análisis de las ANC para tejidos tumorales, ADN plasmático no tratado con bisulfito (BS) y ADN plasmático tratado con bisulfito (de dentro a fuera) para el paciente TBR34. Los patrones de ANC detectados en las muestras de plasma tratadas con bisulfito y no tratadas con bisulfito fueron concordantes.

- 35 Los patrones de ANC detectados en los tejidos tumorales, el plasma no tratado con bisulfito y el plasma tratado con bisulfito fueron concordantes. Para seguir evaluando la concordancia entre los resultados del plasma tratado con bisulfito y el no tratado con bisulfito, se construye un gráfico de dispersión. La FIG. 37B es un gráfico de dispersión que muestra la relación entre las puntuaciones z para la detección de ANC utilizando plasma tratado con bisulfito y no 40 tratado con bisulfito de los bins de 1 Mb para el paciente TBR36. Se observó una correlación positiva entre las puntuaciones z de los dos análisis ($r = 0,89$, $p < 0,001$, correlación de Pearson). La FIG. 38B es un gráfico de dispersión que muestra la relación entre las puntuaciones z para la detección de las ANC utilizando plasma tratado con bisulfito y no tratado con bisulfito de los bins de 1 Mb para el paciente TBR34. Se observó una correlación positiva entre las puntuaciones z de los dos análisis ($r = 0,81$, $p < 0,001$, correlación de Pearson).

45 *C. Análisis sinérgico de los cambios de las ANC y metilación asociados al cáncer*

- Como se ha descrito anteriormente, el análisis de las ANC puede implicar el recuento del número de lecturas de 50 secuencias en cada región de 1 Mb, mientras que el análisis de la densidad de metilación puede implicar la detección de la proporción de restos de citosina en dinucleótidos CpG metilados. La combinación de estos dos análisis puede aportar información sinérgica para la detección del cáncer. Por ejemplo, la clasificación de metilación y la clasificación de las ANC pueden utilizarse para determinar una tercera clasificación de un nivel de cáncer.

- 55 En una realización, la presencia de ANC asociadas al cáncer o de un cambio en la metilación puede utilizarse para indicar la posible presencia de un cáncer. En tal realización, la sensibilidad de la detección del cáncer puede aumentar cuando las ANC o los cambios de metilación están presentes en el plasma de un sujeto analizado. En otra realización, la presencia de ambos cambios puede utilizarse para indicar la presencia de un cáncer. En tal realización, la especificidad de la prueba puede mejorarse porque cualquiera de los dos tipos de cambios puede detectarse potencialmente en algunos sujetos que no padecen cáncer. Por tanto, la tercera clasificación puede ser positiva para el cáncer solo cuando tanto la primera clasificación como la segunda indican cáncer.

- 60 Se reclutaron 26 pacientes con CHC y 22 sujetos sanos. Se recogió una muestra de sangre de cada sujeto y se secuenció el ADN plasmático tras el tratamiento con bisulfito. Para los pacientes con CHC, las muestras de sangre se recogieron en el momento del diagnóstico. La presencia de cantidades significativas de ANC fue, por ejemplo, definida como la que tiene $>5\%$ de los BINS que muestran una puntuación z de <-3 o >3 . La presencia de cantidades significativas de hipometilación asociada al cáncer se definió como la existencia de $>3\%$ de los BINS que mostraban

una puntuación z de <-3. Como ejemplos, la cantidad de regiones (bins) puede expresarse como un número bruto de bins, un porcentaje y una longitud de los bins.

La Tabla 3 muestra la detección de cantidades significativas de ANC y cambios de metilación en el plasma de 26 pacientes con CHC mediante secuenciación masiva paralela en ADN plasmático tratado con bisulfito.

		ANC	
		Presencia	Ausencia
Cambio de metilación	Presencia	12	6
	Ausencia	1	7

Tabla 3

Las tasas de detección de cambios de metilación asociados al cáncer y de ANC fueron del 69 % y del 50 %, respectivamente. La tasa de detección (es decir, la sensibilidad diagnóstica) mejoró hasta el 73 % si se utilizaba la presencia de cualquiera de los dos criterios para indicar la posible presencia de un cáncer.

Se muestran los resultados de dos pacientes con presencia de ANC (FIG. 39A) o cambios de metilación (FIG. 39B). La FIG. 39A es un diagrama Circos que muestra las ANC (anillo interior) y el análisis de metilación (anillo exterior) para el plasma tratado con bisulfito de un paciente con CHC TBR240. Para el análisis de las ANC, los puntos verdes, rojos y grises representan regiones con ganancia o pérdida cromosómica y sin cambio en el número de copias, respectivamente. Para el análisis de metilación, los puntos verdes, rojos y grises representan regiones con hipometilación, hipometilación y metilación normal, respectivamente. En este paciente, se detectó ANC asociadas al cáncer en el plasma, mientras que el análisis de metilación no reveló cantidades significativas de hipometilación asociada al cáncer. La FIG. 39B es un diagrama Circos que muestra las ANC (anillo interior) y el análisis de metilación (anillo exterior) para el plasma tratado con bisulfito de un paciente con CHC TBR164. En este paciente, se detectó hipometilación asociada al cáncer en el plasma. Sin embargo, no se observaron cantidades significativas de ANC. En las FIGS. 48A (TBR36) y 49A (TBR34) se muestran los resultados de dos pacientes que mostraban la presencia de cambios tanto en las ANC como en la metilación.

La Tabla 4 muestra la detección de cantidades significativas de ANC y cambios de metilación en el plasma de 22 sujetos de control mediante secuenciación paralela masiva en ADN plasmático tratado con bisulfito. Se utilizó un enfoque de bootstrapping (es decir, *leave-one-out*) para la evaluación de cada uno de los sujetos de control. Por tanto, cuando se evaluó un sujeto en concreto, los otros 21 sujetos se utilizaron para el cálculo de la media y la DT del grupo de control.

		ANC	
		Presencia	Ausencia
Cambio de metilación	Presencia	1	2
	Ausencia	1	18

Tabla 4

La especificidad de la detección de cantidades significativas de cambio de metilación y ANC fueron del 86 % y el 91 %, respectivamente. La especificidad mejoró hasta el 95 % si se exigía la presencia de ambos criterios para indicar la posible presencia de un cáncer.

En una realización, las muestras positivas para las ANC y/o hipometilación se consideran positivas para cáncer, y las muestras en las que ambos son indetectables se consideran negativas. El uso de la lógica "o" proporciona una mayor sensibilidad. En otra realización, solo las muestras que son positivas tanto para las ANC como para hipometilación se consideran positivas para cáncer, proporcionando así una mayor especificidad. En otra realización más, pueden utilizarse tres niveles de clasificación. Los sujetos se clasifican en i. ambos normales; ii. un anormal; iii. ambos anormales.

Se pueden utilizar diferentes estrategias de seguimiento para estas tres clasificaciones. Por ejemplo, los sujetos de (iii) pueden someterse al protocolo de seguimiento más intensivo, p. ej., con imágenes de todo el cuerpo; los sujetos de (ii) pueden someterse a un protocolo de seguimiento menos intensivo, p. ej., repetir la secuenciación del ADN plasmático tras un intervalo de tiempo relativamente corto de varias semanas; y los sujetos de (i) pueden someterse al protocolo de seguimiento menos intensivo, como repetir las pruebas al cabo de varios años. En otras realizaciones, las mediciones de metilación y ANC pueden utilizarse junto con otros parámetros clínicos (p. ej., resultados de imágenes o bioquímica sérica) para afinar aún más la clasificación.

D. Valor pronóstico del análisis del ADN plasmático tras un tratamiento con intención curativa

- La presencia de ANC y/o cambios de metilación asociados al cáncer en plasma indicaría la presencia de ADN derivado del tumor en el sistema circulatorio del paciente con cáncer. Se esperaría una reducción o desaparición de estos cambios asociados al cáncer tras el tratamiento (p. ej., cirugía). Por otra parte, la persistencia de estos cambios en el plasma después del tratamiento podría indicar la eliminación incompleta de todas las células tumorales del organismo y puede ser un pronóstico útil de la recurrencia de la enfermedad.
- Se recogieron muestras de sangre de los dos pacientes con CHC TBR34 y TBR36 una semana después de la resección quirúrgica con intención curativa de los tumores. Se realizaron análisis de las ANC y la metilación en las muestras de plasma tras el tratamiento tratadas con bisulfito.
- La FIG. 40A muestra el análisis de ANC en ADN plasmático tratado con bisulfito recogido antes (anillo interior) y después (anillo exterior) de la resección quirúrgica del tumor del paciente con CHC TBR36. Cada punto representa el resultado de una región de 1 Mb. Los puntos verdes, rojos y grises representan regiones con ganancia de número de copias, pérdida de número de copias y sin cambio en el número de copias, respectivamente. La mayor parte de las ANC observada antes del tratamiento desapareció tras la resección del tumor. La proporción de bins que mostraban una puntuación z de <-3 o >3 disminuyó del 25 % al 6,6 %.
- La FIG. 40B muestra el análisis de metilación en ADN plasmático tratado con bisulfito recogido antes (anillo interior) y después (anillo exterior) de la resección quirúrgica del tumor del paciente con CHC TBR36. Los puntos verdes, rojos y grises representan regiones con hipermetilación, hipometilación y metilación normal, respectivamente. Se produjo una marcada reducción en la proporción de bins que mostraban una hipometilación significativa, del 90 % al 7,9 %, y el grado de hipometilación también mostró una marcada reducción. Este paciente tuvo una remisión clínica completa 22 meses después de la resección del tumor.
- La FIG. 41A muestra el análisis de ANC en ADN plasmático tratado con bisulfito recogido antes (anillo interior) y después (anillo exterior) de la resección quirúrgica del tumor del paciente con CHC TBR34. Aunque hay una reducción tanto en el número de bins que muestran ANC como en la magnitud de ANC en los bins afectados tras la resección quirúrgica del tumor, pudo observarse ANC residuales en la muestra de plasma postoperatoria. El círculo rojo destaca la región en la que las ANC residuales eran más evidentes. La proporción de bins que mostraban una puntuación z de <-3 o >3 disminuyó del 57 % al 12 %.
- La FIG. 41B muestra el análisis de metilación en ADN plasmático tratado con bisulfito recogido antes (anillo interior) y después (anillo exterior) de la resección quirúrgica del tumor del paciente con CHC TBR34. La magnitud de la hipometilación disminuyó tras la resección del tumor, con una puntuación z media para los bins hipometilados que se redujo de -7,9 a -4,0. Sin embargo, la proporción de bins con una puntuación z <-3 mostró un cambio opuesto, con un aumento del 41 % al 85 %. Esta observación indica potencialmente la presencia de células cancerosas residuales tras el tratamiento. Desde el punto de vista clínico, se detectaron múltiples focos de nódulos tumorales en el resto del hígado no resecado a los 3 meses de la resección tumoral. Se observaron metástasis pulmonares a partir del 4.º mes tras la cirugía. El paciente falleció de recidiva local y enfermedad metastásica 8 meses después de la operación.
- Las observaciones en estos dos pacientes (TBR34 y TBR36) indican que la presencia de cambios residuales de ANC e hipometilación asociados al cáncer puede utilizarse para el seguimiento y pronóstico de pacientes con cáncer tras tratamientos con intención curativa. Los datos también mostraron que el grado de cambio en la cantidad de ANC plasmáticas detectada puede utilizarse de forma sinérgica con la evaluación del grado de cambio en la extensión de la hipometilación del ADN plasmático para el pronóstico y la monitorización de la eficacia del tratamiento.
- En consecuencia, en algunas realizaciones, una muestra biológica se obtiene antes del tratamiento y una segunda muestra biológica se obtiene después del tratamiento (p. ej., cirugía). Los primeros valores se obtienen para la primera muestra, tal como las puntuaciones z de las regiones (p. ej., los niveles de metilación de las regiones y los valores normalizados de ANC) y el número de regiones que muestran hipometilación y ANC (p. ej., amplificación o delección). Los segundos valores pueden obtenerse para la segunda muestra. En otra realización, una tercera, o incluso muestras adicionales, se pueden obtener después del tratamiento. El número de regiones que muestran hipometilación y ANC (p. ej., amplificación o delección) puede obtenerse a partir de la tercera muestra o incluso de muestras adicionales.
- Como se ha descrito anteriormente para las FIGS. 40A y 41A, el primer número de regiones que muestran hipometilación para la primera muestra puede compararse con la segunda cantidad de regiones que muestran hipometilación para la segunda muestra. Como se ha descrito anteriormente para las FIGS. 40B y 41B, la primera cantidad de regiones que muestran hipometilación para la primera muestra puede compararse con la segunda cantidad de regiones que muestran hipometilación para la segunda muestra. La comparación de la primera cantidad con la segunda cantidad y del primer número con el segundo número puede utilizarse para determinar un pronóstico del tratamiento. En diversas realizaciones, solo una de las comparaciones puede ser determinante para el pronóstico o pueden utilizarse ambas. En las realizaciones en las que se obtiene la tercera muestra o incluso muestras adicionales, una o más de estas muestras pueden utilizarse para determinar el pronóstico del tratamiento, por sí solas o junto con la segunda muestra.

- En una implementación, se predice que el pronóstico es peor cuando una primera diferencia entre la primera cantidad y la segunda cantidad está por debajo de un primer umbral de diferencia. En otra implementación, se predice que el pronóstico es peor cuando una segunda diferencia entre el primer número y el segundo número está por debajo de un segundo umbral de diferencia. El umbral puede ser el mismo o diferente. En una realización, el primer umbral de diferencia y el segundo umbral de diferencia son cero. Por tanto, para el ejemplo anterior, la diferencia entre los valores de metilación indicaría un peor pronóstico para el paciente TBR34.
- Un pronóstico puede ser mejor si la primera diferencia y/o la segunda diferencia están por encima de un mismo umbral o de umbrales respectivos. La clasificación para el pronóstico puede depender de lo por encima o por debajo del umbral que estén las diferencias. Se pueden utilizar varios umbrales para obtener diversas clasificaciones. Las diferencias mayores pueden predecir mejores resultados y las diferencias menores (e incluso valores negativos) pueden predecir peores resultados.
- En algunas realizaciones, también se anotan los momentos en que se toman las distintas muestras. Con tales parámetros temporales, se podría determinar la cinética o la tasa de cambio de la cantidad. En una realización, una reducción rápida de la hipometilación asociada al tumor en plasma y/o una reducción rápida de las ANC asociadas al tumor en plasma serán predictoras de un buen pronóstico. Por el contrario, un aumento estático o rápido de la hipometilación asociada al tumor en plasma y/o un aumento estático o rápido del ANC asociado al tumor serán predictivos de mal pronóstico. Las mediciones de metilación y ANC pueden utilizarse junto con otros parámetros clínicos (p. ej., resultados de imágenes o marcadores de bioquímica o proteínas séricas) para predecir el resultado clínico.
- Las realizaciones pueden utilizar otras muestras además de plasma. Por ejemplo, las aberraciones de metilación asociadas a tumores (p. ej., hipometilación) y/o las ANC asociadas a tumores pueden medirse a partir de células tumorales que circulan en la sangre de pacientes con cáncer, a partir de ADN extracelular o de células tumorales en la orina, heces, saliva, esputo, líquido biliar, líquido pancreático, hisopos cervicales, secreciones del aparato genital (p. ej., de la vagina), líquido ascítico, líquido pleural, semen, sudor y lágrimas.
- En diversas realizaciones, las aberraciones de metilación asociadas a tumores (p. ej., hipometilación) y/o las ANC asociadas a tumores pueden detectarse a partir de la sangre o el plasma de pacientes con cáncer de mama, cáncer de pulmón, cáncer colorrectal, cáncer pancreático, cáncer de ovario, carcinoma nasofaríngeo, cáncer cervicouterino, melanoma, tumores cerebrales, etc. En efecto, como la metilación y las alteraciones genéticas como las ANC son fenómenos universales en el cáncer, los enfoques descritos pueden utilizarse para todos los tipos de cáncer. Las mediciones de metilación y ANC pueden utilizarse junto con otros parámetros clínicos (p. ej., resultados de imágenes) para predecir el resultado clínico. Las realizaciones también pueden utilizarse para el cribado y seguimiento de pacientes con lesiones preneoplásicas, p. ej., adenomas.
- En consecuencia, en una realización, la muestra biológica se toma antes del tratamiento, y las mediciones de ANC y metilación se repiten después del tratamiento. Las mediciones pueden producir una primera cantidad posterior de regiones que se determina que presentan una delección o una amplificación y pueden producir una segunda cantidad posterior de regiones que se determina que tienen un nivel de metilación de región que excede el valor de corte de región respectivo. La primera cantidad puede compararse con la primera cantidad posterior, y la segunda cantidad puede compararse con la segunda cantidad posterior para determinar un pronóstico del organismo.
- La comparación para determinar el pronóstico del organismo puede incluir la determinación de una primera diferencia entre la primera cantidad y la primera cantidad posterior, y la primera diferencia puede compararse con uno o más umbrales de primera diferencia para determinar un pronóstico. La comparación para determinar el pronóstico del organismo también puede incluir la determinación de una segunda diferencia entre la segunda cantidad y la segunda cantidad posterior, y la segunda diferencia puede compararse con uno o más umbrales de segunda diferencia. Los umbrales pueden ser cero u otro número.
- Se puede predecir que el pronóstico es peor cuando la primera diferencia está por debajo de un primer umbral de diferencia que cuando la primera diferencia está por encima del primer umbral de diferencia. Se puede predecir que el pronóstico es peor cuando la segunda diferencia está por debajo de un segundo umbral de diferencia que cuando la segunda diferencia está por encima del segundo umbral de diferencia. Algunos ejemplos de tratamientos son la inmunoterapia, cirugía, radioterapia, quimioterapia, terapia basada en anticuerpos, genoterapia, terapia epigenética o terapia dirigida.
- E. Rendimiento*
- A continuación, se describe el rendimiento del diagnóstico para diferentes números de lecturas de secuencias y de tamaño de bin para el análisis de las ANC y la metilación.
- 1. Número de secuencias leídas**

- De acuerdo con una realización, analizamos el ADN plasmático de 32 sujetos de control sanos, 26 pacientes con carcinoma hepatocelular y 20 pacientes con otros tipos de cáncer, incluido el carcinoma nasofaríngeo, cáncer de mama, cáncer de pulmón, cáncer neuroendocrino y sarcoma de músculo liso. Veintidós de los 32 sujetos sanos fueron seleccionados aleatoriamente como grupo de referencia. La media y la desviación típica (DT) de estos 22 individuos de referencia se utilizaron para determinar el intervalo normal de densidad de metilación y representación genómica. El ADN extraído de la muestra de plasma de cada individuo se utilizó para la construcción de bibliotecas de secuenciación utilizando el kit de secuenciación de ambos extremos de Illumina. A continuación, las bibliotecas de secuenciación se sometieron a un tratamiento con bisulfito que convirtió los restos de citosina no metilados en uracilo. La biblioteca de secuenciación convertida con bisulfito de cada muestra de plasma se secuenció utilizando un carril de un secuenciador Illumina HiSeq2000.
- Después la lectura de bases, se eliminaron las secuencias adaptadoras y las bases de baja calidad (es decir, puntuación de calidad <5) en los extremos de los fragmentos. A continuación, las lecturas recortadas en formato FASTQ fueron procesadas por un conducto de análisis de datos de metilación llamado Methy-Pipe (P. Jiang *et al.* 2010, IEEE International Conference on Bioinformatics and Biomedicine, doi: 10.1109/BIBMW.2010.5703866). Para alinear las lecturas de secuenciación convertidas con bisulfito, primero realizamos la conversión informática de todos los restos de citosina a timina en las cadenas Watson y Crick por separado utilizando el genoma humano de referencia (NCBI build 36/hg19). A continuación, realizamos la conversión informática de cada citosina a timina en todas las lecturas procesadas y conservamos la información posicional de cada resto convertido. Se utilizó SOAP2 para alinear las lecturas convertidas con los dos genomas humanos de referencia preconvertidos (R. Li *et al.* 2009 Bioinformatics 25:1966-1967), con un máximo de dos desajustes permitidos para cada lectura alineada. Solo se utilizaron para el análisis posterior las lecturas asignables a una ubicación genómica única. Se eliminaron las lecturas ambiguas asignadas a las cadenas Watson y Crick y las lecturas duplicadas (cloniales). Los restos de citosina en el contexto del dinucleótido CpG se utilizaron para el análisis de metilación posterior. Después de la alineación, las citosinas presentes originalmente en las lecturas secuenciadas se recuperaron basándose en la información posicional conservada durante la conversión informática. Las citosinas recuperadas entre los dinucleótidos CpG se puntuaron como metiladas. Las timinas entre los dinucleótidos CpG se puntuaron como no metiladas.
- Para el análisis de la formetilación, el genoma se dividió en bins de igual tamaño. El tamaño de los bins probados incluye 50 kb, 100 kb, 200 kb y 1 Mb. La densidad de metilación para cada bin se calculó como el número de citosinas metiladas en el contexto del dinucleótido CpG dividido entre el número total de citosinas en posiciones CpG. En otras realizaciones, el tamaño de los bins puede no ser igual en todo el genoma. En una realización, cada bin entre tales bins de tamaños no iguales se compara a través de múltiples sujetos.
- Determinar si la densidad de metilación plasmática de un caso analizado era normal, se comparó la densidad de metilación con los resultados del grupo de referencia. Veintidós de los 32 sujetos sanos fueron seleccionados aleatoriamente como grupo de referencia para el cálculo de la puntuación z de metilación Z (Z_{met}):
- $$Z_{met} = \frac{DM_{prueba} - \overline{DM}_{ref}}{DM_{DT}}$$
- donde DM_{prueba} era la densidad de metilación del caso analizado para un bin concreto de 1 Mb; \overline{DM}_{ref} era la densidad de metilación media del grupo de referencia para el bin correspondiente; y DM_{DT} fue la DT de la densidad de metilación del grupo de referencia para el bin correspondiente.
- Para el análisis de las ANC, se determinó el número de lecturas secuenciadas asignadas a cada bin de 1 Mb (K. C. A. Chan *et al.* 2013 *Clin Chem* 59:211-24). Se determinó la densidad de lecturas secuenciadas para cada ubicación tras la corrección del sesgo de GC mediante la regresión de suavizado de dispersión ponderada localmente, tal y como se ha descrito anteriormente (E. Z. Chen *et al.* 2011 *PLoS One* 6: e21791). Para el análisis de plasma, la densidad de lecturas secuenciadas del caso analizado se comparó con el grupo de referencia para calcular la puntuación z de las ANC (Z_{ANC}):
- $$Z_{ANC} = \frac{DL_{prueba} - \overline{DL}_{ref}}{DL_{DT}}$$
- donde DL_{prueba} era la densidad de lecturas secuenciadas del caso analizado para un bin concreto de 1 Mb; \overline{DL}_{ref} era la densidad media de lecturas secuenciadas del grupo de referencia para el bin correspondiente; y DL_{DT} era la DT de la densidad de lecturas secuenciadas del grupo de referencia para el bin correspondiente. Se definió que un bin presentaba ANC si el Z_{ANC} del bin era <-3 o >3.
- Se obtuvo una media de 93 millones de lecturas alineadas (intervalo: 39 millones a 142 millones) por caso. Para evaluar el efecto de la reducción del número de lecturas secuenciadas en el rendimiento del diagnóstico, seleccionamos aleatoriamente 10 millones de lecturas alineadas de cada caso. Se utilizó el mismo conjunto de individuos de referencia para establecer el intervalo de referencia de cada bin de 1 Mb para el conjunto de datos con lecturas secuenciadas reducidas. El porcentaje de bins que muestran una hipometilación significativa, es decir, $Z_{met} <$

3 y el porcentaje de bins con ANC, es decir, $Z_{ANC} <-3$ o >3 , se determinaron para cada caso. Se utilizaron curvas de eficacia diagnóstica (ROC) para ilustrar el rendimiento diagnóstico de los análisis de hipometilación genómica y ANC para los conjuntos de datos con todas las lecturas secuenciadas de 1 carril y 10 millones de lecturas por caso. En el análisis de ROC, para el análisis se utilizaron los 32 sujetos sanos.

- 5 La FIG. 42 muestra un diagrama del rendimiento diagnóstico del análisis de hipometilación en todo el genoma con diferente número de lecturas secuenciadas. Para el análisis de la hipometilación, las áreas bajo la curva para las curvas de ROC no fueron significativamente diferentes entre los dos conjuntos de datos que analizaron todas las lecturas secuenciadas de un carril y 10 millones de lecturas por caso ($P = 0,761$). Para el análisis de las ANC, el 10 rendimiento del diagnóstico se deterioró con una reducción significativa de las áreas bajo la curva cuando el número de lecturas secuenciadas se redujo de utilizar los datos de un carril a 10 millones ($P <0,001$).

2. Efecto de utilizar distintos tamaños de bins

- 15 Además de dividir el genoma en bins de 1 Mb, también hemos estudiado la posibilidad de utilizar bins más pequeños. En teoría, el uso de bins más pequeños puede reducir potencialmente la variabilidad en la densidad de metilación dentro de un bin. Esto se debe a que la densidad de metilación entre diferentes regiones genómicas puede variar ampliamente. Cuando un bin es más grande, la posibilidad de incluir regiones con diferentes densidades de metilación aumentaría y, por consiguiente, conduciría a un aumento global de la variabilidad de la densidad de metilación de los 20 bins.

Aunque el uso de un tamaño de bin más pequeño puede reducir potencialmente la variabilidad en la densidad de metilación relacionada con la diferencia interregional, por otro lado, esto reduciría el número de lecturas secuenciadas asignadas a un bin concreto. La reducción del número de lecturas asignadas a bins individuales aumentaría la variabilidad debida a la variación del muestreo. El tamaño óptimo del bin que puede dar lugar a la menor variabilidad global en la densidad de metilación puede determinarse experimentalmente para los requisitos de una aplicación de diagnóstico concreta, p. ej., el número total de lecturas secuenciadas por muestra y el tipo de secuenciador de ADN utilizado.

- 30 La FIG. 43 es un diagrama que muestra curvas ROC para la detección de cáncer basada en el análisis de hipometilación genómica con diferentes tamaños de bin (50 kb, 100 kb, 200 kb y 1 Mb). Los valores de P mostrados corresponden a la comparación del área bajo la curva con un tamaño de bin de 1 Mb. Se observa una tendencia a la mejora cuando el tamaño de los bins se reduce de 1 Mb a 200 kb.

35 F. Puntuación de probabilidad acumulada

La cantidad de regiones para la metilación y las ANC puede ser de varios valores. Los ejemplos anteriores describían un número de regiones que superaban un valor de corte o un porcentaje de dichas regiones que mostraban una hipometilación o ANC significativa como parámetro para clasificar si una muestra estaba asociada con el cáncer. Estos 40 enfoques no tienen en cuenta la magnitud de la aberración en cada uno de los bins. Por ejemplo, un bin con una Z_{met} de -3,5 sería lo mismo que un bin con una Z_{met} de -30, ya que ambos se clasificarían como con una hipometilación significativa. Sin embargo, el grado de hipometilación cambia en el plasma, es decir, la magnitud del valor de Z_{met} , se ve afectado por la cantidad de ADN asociado al cáncer en la muestra y, por consiguiente, puede complementar la información del porcentaje de bins que muestran aberraciones para reflejar la carga tumoral. Una mayor concentración 45 fraccional de ADN tumoral en la muestra de plasma conduciría a una menor densidad de metilación y esto se traduciría en un valor de Z_{met} menor.

1. La puntuación de probabilidad acumulada como parámetro de diagnóstico

- 50 Para aprovechar la información de la magnitud de las aberraciones, desarrollamos un enfoque denominado puntuación de probabilidad acumulada (PA). Basándonos en la función de probabilidad de distribución normal, cada valor Z_{met} se tradujo en una probabilidad de tener tal observación por azar.

La puntuación de PA se calculó como:

$$55 \quad \text{Puntuación PA} = \sum -\log(\text{Prob}_i) \text{ para bin}(i) \text{ con } Z_{met} < -3$$

donde Prob_i es la probabilidad para la Z_{met} del bin(i) de acuerdo con la distribución t de Student con 3 grados de libertad, y log es la función logaritmo natural. En otra realización, se puede utilizar un logaritmo de base 10 (u otro número). En otras realizaciones, otras distribuciones, por ejemplo, pero sin limitarse a la distribución normal y la distribución gamma, puede aplicarse para transformar la puntuación z en la PA.

Una puntuación de PA mayor indica una menor probabilidad de tener esa densidad de metilación desviada en una población normal por azar. Por lo tanto, una puntuación de PA alta indicaría una mayor probabilidad de tener ADN 65 anormalmente hipometilado en la muestra, p. ej., la presencia de ADN asociado al cáncer.

Comparado con el porcentaje de bins que muestran aberración, la medición de la puntuación de PA tiene un mayor intervalo dinámico. Aunque la carga tumoral entre distintos pacientes puede variar mucho, el mayor intervalo de valores de PA sería útil para reflejar las cargas tumorales de pacientes con cargas tumorales relativamente altas y relativamente bajas. Adicionalmente, el uso de puntuaciones de PA puede ser potencialmente más sensible para detectar los cambios en la concentración de ADN asociado a tumores en plasma. Esto es ventajoso para el seguimiento de la respuesta al tratamiento y el pronóstico. Por consiguiente, una reducción de las puntuaciones de PA durante el tratamiento es indicativa de una buena respuesta al mismo. La falta de reducción o incluso el aumento de las puntuaciones de PA durante el tratamiento indicaría una respuesta escasa o nula. Para pronosticar, una puntuación de PA alta es indicativa de una carga tumoral elevada e indica un mal pronóstico (p. ej., mayor probabilidad de muerte o de progresión tumoral).

La FIG. 44A muestra un rendimiento de diagnóstico para la probabilidad acumulada (PA) y el porcentaje de bins con aberraciones. No hubo diferencias significativas entre las áreas bajo la curva de los dos tipos de algoritmo diagnóstico ($P = 0,791$).

La FIG. 44B muestra los resultados diagnósticos del análisis de plasma para la hipometilación global, la hipermetilación de islas CpG y las ANC. Con un carril de secuenciación por muestra (200 kb de tamaño de bin para el análisis de hipometilación y 1 Mb de tamaño de bin para las ANC, y las islas CpG definidas de acuerdo con la base de datos alojada en la Universidad de California, Santa Cruz (UCSC)), las áreas bajo la curva de los tres tipos de análisis fueron superiores a 0,90.

En los análisis posteriores, la puntuación más alta de PA en los sujetos de control se utilizó como valor de corte para cada uno de los tres tipos de análisis. La selección de estos puntos de corte proporcionó una especificidad diagnóstica del 100 %. Las sensibilidades diagnósticas para la hipometilación general, Los análisis de hipermetilación de islas CpG y ANC fueron del 78 %, 89 % y 52 %, respectivamente. En 43 de los 46 pacientes con cáncer, se detectó al menos uno de los tres tipos de aberraciones, por tanto, dando lugar a una sensibilidad del 93,4 % y una especificidad del 100 %. Nuestros resultados indican que los tres tipos de análisis pueden utilizarse de forma sinérgica para la detección del cáncer.

La FIG. 45 muestra una tabla con los resultados de la hipometilación global, hipermetilación de islas CpG y ANC en pacientes con carcinoma hepatocelular. Los valores de corte de la puntuación de PA para los tres tipos de análisis fueron 960, 2,9 y 211, respectivamente. Los resultados positivos de la puntuación de PA aparecen en negrita y subrayados.

La FIG. 46 muestra una tabla con los resultados de la hipometilación global, la hipermetilación de islas CpG y las ANC en pacientes que padecen cánceres distintos del carcinoma hepatocelular. Los valores de corte de la puntuación de PA para los tres tipos de análisis fueron 960, 2,9 y 211, respectivamente. Los resultados positivos de la puntuación de PA aparecen en negrita y subrayados.

40 2. Aplicación de la puntuación de PA al seguimiento del cáncer

Se recogieron muestras en serie de un paciente con CHC TBR34 antes y después del tratamiento. Las muestras se analizaron en busca de hipometilación global.

La FIG. 47 muestra un análisis en serie de la metilación plasmática para el caso TBR34. El anillo más interior muestra la densidad de metilación de la capa leucocitaria (negro) y de los tejidos tumorales (púrpura). Para las muestras de plasma, la Z_{met} se muestra para cada intervalo de 1 Mb. La diferencia entre dos líneas representa una diferencia de Z_{met} de 5. Los puntos rojos y grises representan bins con hipometilación y sin cambios en la densidad de metilación en comparación con el grupo de referencia. Desde el 2.º anillo interior hacia fuera son las muestras de plasma tomadas antes del tratamiento, a los 3 días y a los 2 meses de la resección del tumor, respectivamente. Antes del tratamiento, se observó un alto grado de hipometilación en el plasma y más del 18,5 % de los bins tenían una Z_{met} de <-10. A los 3 días de la resección del tumor, se pudo observar que el grado de hipometilación se redujo en el plasma sin que ninguno de los bins tuviera una Z_{met} de <-10.

N.º de caso	Punto temporal	Análisis de metilación		
		Porcentaje de bins que muestran hipometilación significativa	Probabilidad acumulada (PA)	Puntuación z sumativa
TBR34	Antes de la OT	62,6 %	37.573	14.285
	3 días después de la OT	80,5 %	17.777	9195
	2 meses después de la OT	40,1 %	15.087	5201

55 Tabla 5

La Tabla 5 muestra que, aunque la magnitud de los cambios de hipometilación se redujo a los 3 días de la resección quirúrgica del tumor, el porcentaje de bins que presentaban aberraciones mostró un aumento paradójico. Por otra parte, la puntuación de PA reveló con mayor precisión la reducción del grado de hipometilación en plasma y puede reflejar mejor los cambios en la carga tumoral.

- 5 A los 2 meses después de la OT, todavía había un porcentaje significativo de bins que mostraban cambios de hipometilación. La puntuación de PA también permaneció estática en torno a los 15.000 puntos. A este paciente se le diagnosticaron posteriormente depósitos tumorales multifocales (previamente desconocidos en el momento de la intervención quirúrgica) en el resto del hígado no resecado a los 3 meses y se le observaron múltiples metástasis pulmonares a los 4 meses de la operación. El paciente falleció de enfermedad metastásica a los 8 meses de la operación. Estos resultados indican que la puntuación de PA podría ser más potente que el porcentaje de bins con aberración para reflejar la carga tumoral.

- 10 En términos generales, la PA puede ser útil para aplicaciones que requieran medir la cantidad de ADN tumoral en plasma. Algunos ejemplos de estas aplicaciones son: pronóstico y seguimiento de pacientes con cáncer (p. ej., para observar la respuesta al tratamiento o la progresión del tumor).

- 15 La puntuación z sumativa es una suma directa de las puntuaciones z, es decir, sin convertir a probabilidad. En este ejemplo, la puntuación z sumativa muestra el mismo comportamiento que la puntuación de PA. En otros casos, la PA puede ser más sensible que la puntuación z sumativa para el seguimiento de la enfermedad residual debido al mayor intervalo dinámico de la puntuación de PA.

X. IMPACTO DE LAS ANC EN LA METILACIÓN

- 20 25 El uso de las ANC y la metilación para determinar las clasificaciones respectivas de un nivel de cáncer, donde las clasificaciones se combinan para obtener una tercera clasificación, se ha descrito anteriormente. Además de tal combinación, las ANC se pueden utilizar para cambiar los valores de corte del análisis de metilación y para identificar falsos positivos comparando los niveles de metilación de grupos de regiones con diferentes características de las ANC. A modo de ejemplo, el nivel de metilación para la sobreabundancia (p. ej., $Z_{ANC} > 3$) puede compararse con el nivel de metilación para la abundancia normal (p. ej., $-3 < Z_{ANC} < 3$). En primer lugar, se describe el impacto del ANC en los niveles de metilación.

A. Alteración de la densidad de metilación en regiones con ganancias y pérdidas cromosómicas

- 30 35 Dado que los tejidos tumorales muestran generalmente una hipometilación global, la presencia de ADN derivado de tumores en el plasma de pacientes con cáncer conduciría a la reducción de la densidad de metilación en comparación con sujetos sin cáncer. El grado de hipometilación en el plasma de pacientes con cáncer es teóricamente proporcional a la concentración fraccional de ADN derivado del tumor en la muestra de plasma.
- 40 45 Para las regiones que muestran una ganancia cromosómica en los tejidos tumorales, se liberaría una dosis adicional de ADN tumoral de los segmentos de ADN amplificados al plasma. Esta mayor contribución de ADN tumoral al plasma conduciría teóricamente a un mayor grado de hipometilación en el ADN plasmático de la región afectada. Un factor adicional es que se espera que las regiones genómicas que muestran amplificación confieran una ventaja de crecimiento a las células tumorales y, por tanto, se espera que se expresen. Estas regiones suelen estar hipometiladas.

- 50 55 Por el contrario, para las regiones que muestran pérdida cromosómica en el tejido tumoral, la menor contribución del ADN tumoral al plasma conduciría a un menor grado de hipometilación en comparación con las regiones sin cambios en el número de copias. Un factor adicional es que las regiones genómicas que se eliminan en las células tumorales podrían contener genes supresores de tumores y podría ser ventajoso para las células tumorales tener silenciadas dichas regiones. Por tanto, se espera que estas regiones tengan una mayor probabilidad de estar hipermetiladas.
- En el presente documento, utilizamos los resultados de dos pacientes con CHC (TBR34 y TBR36) para ilustrar este efecto. Las FIGS. 48A (TBR36) y 49A (TBR34) tienen círculos que resaltan las regiones con ganancias o pérdidas cromosómicas y el correspondiente análisis de metilación. Las FIGS. 48B y 49B muestran gráficas de puntuaciones z de metilación para pérdidas, normal y ganancias para los pacientes TBR36 y TBR34, respectivamente.

- 60 65 La FIG. 48A muestra diagramas Circos que demuestran las ANC (anillo interior) y los cambios de metilación (anillo exterior) en el ADN plasmático tratado con bisulfito para el paciente con CHC TBR36. Los círculos rojos destacan las regiones con ganancias o pérdidas cromosómicas. Las regiones que mostraban ganancias cromosómicas estaban más hipometiladas que las regiones sin cambios en el número de copias. Las regiones que presentaban pérdidas cromosómicas estaban menos hipometiladas que las regiones sin cambios en el número de copias. La FIG. 48B es un gráfico de puntuaciones z de metilación para regiones con ganancias y pérdidas cromosómicas, y regiones sin cambios en el número de copias para el paciente con CHC TBR36. En comparación con las regiones sin cambios de copia, las regiones con ganancias cromosómicas tenían puntuaciones z más negativas (más hipometilación) y las regiones con pérdidas cromosómicas tenían puntuaciones z menos negativas (menos hipometilación).

La FIG. 49A muestra diagramas Circos que demuestran las ANC (anillo interior) y los cambios de metilación (anillo exterior) en el ADN plasmático tratado con bisulfito para el paciente con CHC TBR34. La FIG. 49B es un gráfico de puntuaciones z de metilación para regiones con ganancias y pérdidas cromosómicas, y regiones sin cambios en el número de copias para el paciente con CHC TBR34. La diferencia en las densidades de metilación entre las regiones con ganancias y pérdidas cromosómicas fue mayor en el paciente TBR36 que en el paciente TBR34 porque la concentración fraccional de ADN derivado del tumor en el primer paciente era mayor.

En este ejemplo, las regiones utilizadas para determinar las ANC son las mismas que las utilizadas para determinar la metilación. En una realización, los valores de corte de la región respectiva dependen de si la región respectiva presenta una delección o una amplificación. En una implementación, un valor de corte de la región respectiva (p. ej., el corte de la puntuación z utilizado para determinar la hipometilación) tiene una magnitud mayor cuando la región respectiva presenta una amplificación que cuando no presenta amplificación (p. ej., la magnitud podría ser superior a 3, y puede utilizarse un corte inferior a -3). Por tanto, para comprobar la hipometilación, el valor de corte de una región respectiva puede tener un valor negativo mayor cuando la región respectiva presenta una amplificación que cuando no presenta amplificación. Se espera que esta implementación mejore la especificidad de la prueba para detectar el cáncer.

En otra implementación, un valor de corte de la región respectiva tiene una magnitud menor (p. ej., inferior a 3) cuando la región respectiva presenta una delección que cuando no presenta ninguna delección. Por tanto, para comprobar la hipometilación, el valor de corte de una región respectiva puede tener un valor menos negativo cuando la región respectiva presenta una delección que cuando no presenta ninguna delección. Se espera que esta implementación mejore la sensibilidad de la prueba para detectar el cáncer. El ajuste de los valores de corte en las implementaciones anteriores puede modificarse en función de la sensibilidad y especificidad deseadas para un escenario de diagnóstico concreto. En otras realizaciones, las mediciones de metilación y ANC pueden utilizarse junto con otros parámetros clínicos (p. ej., resultados de imágenes o bioquímica sérica) para la predicción del cáncer.

B. Utilización de las ANC para seleccionar regiones

Como se ha descrito anteriormente, hemos demostrado que la densidad de metilación plasmática estaría alterada en regiones con aberraciones en el número de copias de los tejidos tumorales. En las regiones con aumento del número de copias en el tejido tumoral, una mayor contribución de ADN tumoral hipometilado al plasma conduciría a un mayor grado de hipometilación del ADN plasmático en comparación con las regiones sin aberración del número de copias. Por el contrario, en regiones con pérdida de número de copias en el tejido tumoral, la menor contribución del ADN hipometilado derivado del cáncer al plasma conduciría a un menor grado de hipometilación del ADN plasmático. Esta relación entre la densidad de metilación del ADN plasmático y la representación relativa puede utilizarse para diferenciar los resultados de hipometilación asociados a la presencia de ADN asociado a cáncer y otras causas no cancerosas (p. ej., LES) de hipometilación en el ADN plasmático.

Para ilustrar este enfoque, analizamos las muestras de plasma de dos pacientes con carcinoma hepatocelular (CHC) y dos pacientes con LES sin cáncer. Estos dos pacientes con LES (LES04 y LES10) mostraron la presencia evidente de hipometilación y ANC en plasma. Para el paciente LES04, el 84 % de los bins mostraron hipometilación y el 11,2 % mostraron ANC. Para el paciente LES10, el 10,3 % de los bins mostraron hipometilación y el 5,7 % mostraron ANC.

Las FIG. 50A y 50B muestran los resultados de los análisis de hipometilación plasmática y ANC para los pacientes con LES LES04 y LES10. El círculo exterior muestra las puntuaciones z de metilación (Z_{met}) con una resolución de 1 Mb. Los bins con metilación $Z_{met} < -3$ estaban en rojo y los que tenían $Z_{met} > +3$ estaban en gris. El círculo interior muestra las puntuaciones z del ANC (Z_{ANC}). Los puntos verdes, rojos y grises representan bins con $Z_{ANC} > 3$, < -3 y entre -3 y 3, respectivamente. En estos dos pacientes con LES, se observaron cambios en la hipometilación y las ANC en el plasma.

Para determinar si los cambios en la metilación y las ANC eran coherentes con la presencia de ADN derivado del cáncer en el plasma, comparamos el valor de Z_{met} para regiones con $Z_{ANC} > 3$, < -3 y entre -3 y 3. Para los cambios de metilación y ANC aportados por el ADN derivado del cáncer en plasma, cabría esperar que las regiones con $Z_{ANC} < -3$ estuvieran menos hipometiladas y tuvieran menos Z_{met} negativa. Por el contrario, cabría esperar que las regiones con $Z_{ANC} > 3$ estuvieran más hipometiladas y tuvieran Z_{met} más negativas. Con fines ilustrativos, aplicamos la prueba de suma de rangos unilateral para comparar la Z_{met} de las regiones con ANC (es decir, las regiones con $Z_{ANC} < -3$ o > 3) con las regiones sin ANC (es decir, las regiones con Z_{ANC} entre -3 y 3). En otras realizaciones, se pueden utilizar otras pruebas estadísticas, por ejemplo, pero sin limitarse a la prueba de la t de Student, la prueba de análisis de la varianza (ANOVA) y la prueba de Kruskal-Wallis.

Las FIGS. 51A y 51B muestran el análisis Z_{met} para regiones con y sin ANC para el plasma de dos pacientes con CHC (TBR34 y TBR36). Las regiones con $Z_{ANC} < -3$ and > 3 representan regiones con infrarepresentación y sobrerepresentación en plasma, respectivamente. Tanto en TBR34 como en TBR36, las regiones que estaban infrarepresentadas en el plasma (es decir, las regiones con $Z_{ANC} < -3$) tenían una Z_{met} significativamente mayor (valor de $P < 10^{-5}$, prueba de suma de rangos unilateral) que las regiones con representación normal en plasma (es decir, regiones con Z_{ANC} entre -3 y 3). Una representación normal corresponde a la esperada para un genoma euploide. En

el caso de las regiones con sobrerepresentación en plasma (es decir, las regiones con $Z_{ANC} > 3$), presentaban una Z_{met} significativamente inferior al de las regiones con representación normal en plasma (valor de $P < 10^{-5}$, prueba de suma de rangos unilateral). Todos estos cambios fueron

5 consistentes con la presencia de ADN tumoral hipometilado en las muestras de plasma.

Las FIGS. 51C y 51D muestran el análisis Z_{met} para regiones con y sin ANC para el plasma de dos pacientes con LES (LES04 y LES10). Las regiones con $Z_{ANC} < -3$ and > 3 representan regiones con infrarepresentación y sobrerepresentación en plasma, respectivamente. Para LES04, las regiones que estaban infrarepresentadas en el plasma (es decir, las regiones con $Z_{ANC} < -3$) no presentaban una Z_{met} significativamente superior (valor de $P = 0,99$, prueba de suma de rangos unilateral) que las regiones con representación normal en plasma (es decir, regiones con Z_{ANC} entre -3 y 3) y las regiones con sobrerepresentación en plasma (es decir, regiones con $Z_{ANC} > 3$) no presentaron una Z_{met} significativamente inferior que las regiones con representación normal en plasma (valor de $P = 0,68$, prueba de suma de rangos unilateral). Estos resultados fueron diferentes de los cambios esperados debido a la presencia de 10 ADN hipometilado derivado del tumor en el plasma. De manera similar, para LES10, las regiones con una $Z_{ANC} < -3$ no presentaron una Z_{met} significativamente más elevada que las regiones con una Z_{ANC} de entre -3 y 3 (valor de $P = 0,99$, prueba de suma de rangos unilateral).

Una razón de no tener el patrón típico asociado al cáncer entre Z_{met} y Z_{ANC} en los pacientes con LES es que, en los 20 pacientes con LES, las ANC no están presentes en un tipo celular específico que también presenta hipometilación. En lugar de ello, la presencia evidente observada de ANC e hipometilación se debe a la distribución alterada del tamaño del ADN circulante en pacientes con LES. La distribución alterada del tamaño podría alterar potencialmente las densidades de lectura secuenciadas para diferentes regiones genómicas, dando lugar a ANC evidentes, ya que las referencias procedían de sujetos sanos. Como se ha descrito en secciones anteriores, existe una correlación entre el 25 tamaño de un fragmento de ADN circulante y su densidad de metilación. Por lo tanto, la distribución alterada del tamaño también puede conducir a una metilación aberrante.

Aunque las regiones con $Z_{ANC} > 3$ tenían niveles de metilación ligeramente inferiores a las regiones con Z_{ANC} entre -3 y 3, el valor de p de la comparación fue muy superior a los observados en dos pacientes con cáncer. En una realización, 30 el valor de p puede utilizarse como parámetro para determinar la probabilidad de que un caso analizado padezca un cáncer. En otra realización, la diferencia de Z_{met} entre regiones con representación normal y aberrante puede utilizarse como parámetro para indicar la probabilidad de presencia de cáncer. En una realización, se puede utilizar un grupo de pacientes con cáncer para establecer la correlación entre Z_{met} y Z_{ANC} y para determinar los umbrales de los distintos 35 parámetros a fin de indicar que los cambios son coherentes con la presencia de ADN hipometilado derivado del cáncer en la muestra de plasma analizada.

En consecuencia, en una realización, se puede realizar un análisis de ANC para determinar un primer conjunto de regiones en el que todas presentan una de: una delección, una amplificación o una representación normal. Por ejemplo, 40 en el primer conjunto de regiones todas pueden presentar una supresión, o todas una amplificación, o todas una representación normal (p. ej., tener una primera cantidad normal de regiones, tal como una Z_{met} normal). Se puede determinar un nivel de metilación para este primer conjunto de regiones (p. ej., el primer nivel de metilación del método 2800 puede corresponder al primer conjunto de regiones).

45 El análisis de ANC puede determinar un segundo conjunto de regiones que presentan todas una segunda de: una delección, una amplificación o una representación normal. El segundo conjunto de regiones se comportaría de forma diferente al primero. Por ejemplo, si el primer conjunto de regiones fuera normal, entonces el segundo conjunto de regiones puede presentar una delección o una amplificación. Se puede calcular un segundo nivel de metilación basado en los números respectivos de moléculas de ADN metiladas en sitios del segundo conjunto de regiones.

50 A continuación, se puede calcular un parámetro entre el primer y el segundo nivel de metilación. Por ejemplo, se puede calcular una diferencia o proporción y compararla con un valor de corte. La diferencia o proporción también puede someterse a una distribución de probabilidad (p. ej., como parte de una prueba estadística) para determinar la probabilidad de obtener el valor, y esta probabilidad puede compararse con un valor de corte para determinar un nivel 55 de cáncer basado en los niveles de metilación. Dicho valor corte puede elegirse para diferenciar las muestras que tienen cáncer de las que no lo tienen (p. ej., LES).

En una realización, se puede determinar un nivel de metilación para el primer conjunto de regiones o una mezcla de regiones (es decir, mezcla de regiones que muestran amplificación, supresión y normalidad). Este nivel de metilación 60 puede entonces compararse con un primer valor de corte como parte de una primera etapa de análisis. Si se supera el valor de corte, indicando así la posibilidad de cáncer, entonces se puede realizar el análisis anterior para determinar si la indicación era un falso positivo. Así pues, la clasificación final del nivel de cáncer puede incluir la comparación del parámetro de los dos niveles de metilación con un segundo valor de corte.

65 El primer nivel de metilación puede ser un valor estadístico (p. ej., promedio o mediana) de los niveles de metilación de la región calculados para cada región del primer conjunto de regiones. El segundo nivel de metilación también puede ser un valor estadístico de los niveles de metilación de la región calculados para cada región del segundo

conjunto de regiones. Como ejemplos, los valores estadísticos pueden determinarse mediante la prueba de suma de rangos unilateral, la prueba de la *t* de Student, prueba de análisis de la varianza (ANOVA) o prueba de Kruskal-Wallis.

XI. CLASIFICACIÓN DEL TIPO DE CÁNCER

- Además de determinar si un organismo tiene cáncer o no, algunas realizaciones pueden identificar un tipo de cáncer asociado a la muestra. Esta identificación del tipo de cáncer puede utilizar patrones de hipometilación global, hipermetilación de islas CpG y/o ANC. Los patrones pueden implicar la agrupación de pacientes con un diagnóstico conocido utilizando los niveles de metilación de regiones medidos, valores respectivos de ANC para las regiones y nivel de metilación para las islas CpG. Los siguientes resultados muestran que los organismos con un tipo similar de cáncer tienen valores similares para las regiones e islas CpG, así como los pacientes sin cáncer que presentan valores similares. En la agrupación, cada uno de los valores de una región o isla puede ser una dimensión independiente en el proceso de agrupación.
- Se sabe que el mismo tipo de cánceres compartiría cambios genéticos y epigenéticos similares (E Gebhart *et al.* 2004 *Cytogenet Genome Res.* 104: 352-358; P. A. Jones *et al.* 2007 *Cell*; 128: 683-692). A continuación, describimos cómo los patrones de cambios de ANC y metilación detectados en el plasma son útiles para inferir el origen o tipo de cáncer. Las muestras de ADN plasmático de los pacientes con CHC, los pacientes sin CHC y los sujetos de control sanos se clasificaron mediante, por ejemplo, análisis de agrupación jerárquica. El análisis se realizó utilizando, por ejemplo, la función heatmap.2 del paquete de scripts de R (cran.r-project.org/web/packages/gplots/gplots.pdf).

A fin de ilustrar el potencial de este enfoque, utilizamos dos conjuntos de criterios (grupo A y grupo B) como ejemplos para identificar características útiles para la clasificación de las muestras de plasma (véase la Tabla 6). En otras realizaciones, se pueden utilizar otros criterios para identificar las características. Las características utilizadas incluyeron ANC global a 1 Mb de resolución, densidad de metilación global a 1 Mb de resolución y metilación de islas CpG.

	Metilación global a 1 Mb de resolución	
	Criterios del Grupo A	Criterios del Grupo B
Criterios	>20 casos de cáncer con una puntuación z >3 o <-3	>20 casos de cáncer con una puntuación z >2,5 o <-2,5
N.º de elementos identificados	584	1911

	Características de las ANC	
	Criterios del Grupo A	Criterios del Grupo B
Criterios	>10 casos de cáncer con una puntuación z >3 o <-3	>10 casos de cáncer con una puntuación z >2,5 <-2,5
N.º de elementos identificados	355	759

	Metilación de islas CpG	
	Criterios del Grupo A	Criterios del Grupo B
Criterios	>5 casos de cáncer con una densidad de metilación que difiere de la media de la referencia en un 2 % en las islas CpG concretas	>1 caso de cáncer con una densidad de metilación que difiere de la media de la referencia en un 2 % en las islas CpG concretas
N.º de elementos identificados	110	191

Tabla 6

- En los dos primeros ejemplos, usamos todas las ANC, metilación global a 1 Mb de resolución y características de metilación de islas CpG para la clasificación. En otras realizaciones, se pueden utilizar otros criterios, por ejemplo, aunque no exclusivamente, la precisión de la medición de la característica en el plasma del grupo de referencia.
- La FIG. 52A muestra el análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando el total de las 1130 características del grupo A, que incluía 355 ANC, 584 características de metilación global a 1 Mb de resolución y el estado de metilación de 110 islas CpG. La barra de color de la parte superior representa los grupos de muestra: el verde, el azul y el rojo representan a los sujetos sanos, y los pacientes con y sin cáncer CHC, respectivamente. En general, los tres grupos de sujetos tendían

a agruparse. El eje vertical representa las características de clasificación. Se agruparon las características con patrones similares en diferentes sujetos. Estos resultados indican que los patrones de metilación de las islas CpG cambian, los cambios de metilación en el genoma a 1 Mb de resolución y las ANC en plasma pueden utilizarse potencialmente para determinar el origen del cáncer en pacientes con primarios desconocidos.

- 5 La FIG. 52B muestra el análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando el total de las 2780 características del grupo B, incluidas 759 ANC, 1911 características de metilación global a 1 Mb de resolución y el estado de metilación de 191 islas CpG. La barra de color de la parte superior representa los grupos de muestra: el verde, el azul y el rojo representan a los sujetos sanos, y los pacientes con y sin cáncer CHC, respectivamente. En general, los tres grupos de sujetos tendían a agruparse. El eje vertical representa las características de clasificación. Se agruparon las características con patrones similares en diferentes sujetos. Estos resultados indican que los patrones de los distintos conjuntos de islas CpG de metilación cambian, los cambios de metilación en el genoma a 1 Mb de resolución y las ANC en plasma pueden utilizarse para determinar el origen del cáncer en pacientes con primarios desconocidos. La selección de las 10 características de clasificación puede ajustarse para aplicaciones específicas. Adicionalmente, se puede dar peso a la predicción del tipo de cáncer en función de las probabilidades a priori de los sujetos para los distintos tipos de cáncer. Por ejemplo, los pacientes con hepatitis vírica crónica son propensos al desarrollo de carcinoma hepatocelular y los fumadores crónicos son propensos al desarrollo de cáncer de pulmón. Por tanto, se puede calcular una 15 probabilidad ponderada del tipo de cáncer mediante, por ejemplo, pero sin limitación, regresión logística, múltiple o de agrupaciones.
- 20

En otras realizaciones, se puede utilizar un único tipo de características para el análisis de clasificación. Por ejemplo, en las siguientes ejemplos, solo se utilizaron la metilación global a 1 Mb de resolución, la hipermetilación de las islas CpG o las ANC con una resolución de 1 Mb para el análisis de agrupación jerárquica. El poder de diferenciación puede 25 ser distinto cuando se utilizan características diferentes. Un mayor perfeccionamiento de las características de clasificación puede mejorar las precisiones de clasificación.

La FIG. 53A muestra el análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las características de metilación de islas CpG del grupo A. 30 Generalmente, los pacientes con cáncer se agruparon y los sujetos sin cáncer estaban en otro grupo. Sin embargo, los pacientes con CHC y sin CHC se separaron menos en comparación con el uso de los tres tipos de características.

La FIG. 53B muestra el análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las densidades de metilación global del grupo A a 1 Mb de 35 resolución como características de clasificación. Se observó una agrupación preferencial de pacientes con CHC y sin CHC.

La FIG. 54A muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las ANC globales del grupo A a 1 Mb de resolución como 40 características de clasificación. Se vio una agrupación preferencial de pacientes con CHC y sin CHC.

La FIG. 54B muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las densidades de metilación de las islas CpG del grupo B como 45 características de clasificación. Se observó una agrupación preferencial de pacientes con cáncer con y sin CHC.

La FIG. 55A muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las densidades de metilación global del grupo B a 1 Mb de resolución como 50 características de clasificación. Se observó una agrupación preferencial de pacientes con cáncer con y sin CHC.

La FIG. 55B muestra un análisis de agrupación jerárquica para muestras de plasma de pacientes con CHC, pacientes con cáncer no CHC y sujetos de control sanos utilizando las ANC globales del grupo B a 1 Mb de resolución como 55 características de clasificación. Se observó una agrupación preferencial de pacientes con cáncer con y sin CHC.

Estos resultados de agrupación jerárquica para muestras de plasma indican que la combinación de diferentes características puede utilizarse para la identificación de los tipos primarios de cáncer. El perfeccionamiento de los criterios de selección puede mejorar aún más la precisión de la clasificación.

En consecuencia, en una realización, cuando una clasificación de metilación indica que existe cáncer para el 60 organismo, se puede identificar un tipo de cáncer asociado con el organismo comparando un nivel de metilación (p. ej., la primera metilación del método 2800 o cualquier nivel de metilación de la región) con un valor correspondiente determinado a partir de otros organismos (es decir, otros organismos del mismo tipo, tales como seres humanos). El valor correspondiente podría ser para una misma región o conjunto de sitios para el que se calculó el nivel de metilación. Al menos dos de los otros organismos están identificados con diferentes tipos de cáncer. Por ejemplo, los 65 valores correspondientes pueden organizarse en agrupaciones, donde dos grupos se asocian a cánceres diferentes.

- Además, cuando las ANC y la metilación se utilizan conjuntamente para obtener una tercera clasificación del nivel de cáncer, las características de ANC y metilación pueden compararse con los valores correspondientes de otros organismos. Por ejemplo, la primera cantidad de regiones (p. ej., de la FIG. 36) que presentan una delección o amplificación puede compararse con los valores correspondientes determinados a partir de los otros organismos para identificar el tipo de cáncer asociado con el organismo.
- En algunas realizaciones, las características de metilación son los niveles de metilación de regiones de una pluralidad de regiones del genoma. Pueden utilizarse las regiones cuyo nivel de metilación supere el valor de corte de la región correspondiente, p. ej., los niveles de metilación de regiones del organismo pueden compararse con los niveles de metilación de regiones de otros organismos para las mismas regiones del genoma. La comparación puede permitir diferenciar tipos de cáncer, o simplemente proporcionar un filtro adicional para confirmar el cáncer (p. ej., para identificar falsos positivos). Por tanto, se puede determinar si el organismo padece el primer tipo de cáncer, la ausencia de cáncer o el segundo tipo de cáncer basándose en la comparación.
- Los demás organismos (junto con el que se está analizando) pueden agruparse utilizando los niveles de metilación de la región. Por tanto, la comparación de los niveles de metilación de las regiones puede utilizarse para determinar a qué grupo pertenece el organismo. La agrupación también puede utilizar valores normalizados de ANC para las regiones en las que se determine que presentan una delección o una amplificación, como se ha descrito anteriormente. Asimismo, la agrupación puede utilizar las respectivas densidades de metilación de las islas CpG hipermetiladas.
- Para ilustrar el principio de este método, mostramos un ejemplo de uso de la regresión logística para la clasificación de dos muestras desconocidas. El objetivo de esta clasificación era determinar si estas dos muestras eran cánceres CHC o no CHC. Se recopiló un conjunto de muestras de preparación que incluía 23 muestras de plasma recogidas de pacientes con CHC y 18 muestras de pacientes que padecían un cáncer distinto del CHC. Por tanto, había un total de 41 casos en el conjunto de preparación. En este ejemplo, se seleccionaron 13 características, incluyendo cinco características sobre la metilación de islas CpG (X1-X5), seis características sobre la metilación de regiones de 1-Mb (X6-X11) y 2 características sobre las ANC de regiones de 1 Mb (X12-X13). Las características de metilación CpG se seleccionaron basándose en el criterio de que al menos 15 casos del conjunto de entrenamiento tuvieran una puntuación z de >3 o <-3 . Las características de metilación de 1 Mb se seleccionaron basándose en el criterio de que al menos 39 casos del conjunto de entrenamiento tuvieran una puntuación z de >3 o <-3 . Las características de ANC se seleccionaron basándose en el criterio de que al menos 20 casos tuvieran una puntuación z >3 o <-3 . Se realizó una regresión logística con las muestras de este conjunto de entrenamiento para determinar el coeficiente de regresión de cada una de las características (X1-X13). Las características con coeficientes de regresión de las magnitudes más grandes (independientemente de que sea en sentido positivo o negativo) ofrecen una mejor discriminación entre las muestras de CHC y las que no lo son. Las puntuaciones z de cada caso para las características respectivas se utilizaron como valores de entrada de las variables independientes. Luego dos muestras de plasma, una de un paciente con CHC (TBR36) y otra de un paciente con cáncer de pulmón (TBR177) se analizaron en busca de las 13 características.
- En este análisis de clasificación del tipo de cáncer, se supuso que estas dos muestras procedían de pacientes con cánceres de origen desconocido. Para cada muestra, las puntuaciones z de las características respectivas se introdujeron en la ecuación de regresión logística para determinar el logaritmo natural de la razón de posibilidades ($\ln(\text{razón de posibilidades})$), en donde la razón de posibilidades representaba la proporción entre las probabilidades de tener un CHC y de no tenerlo (CHC/no CHC).
- La Tabla 7 muestra los coeficientes de regresión para las 13 características de la ecuación de regresión logística. También se muestran las puntuaciones z de las características respectivas de los dos casos analizados (TBR36 y TBR177).
- El $\ln(\text{razón de posibilidades})$ de CHC para TBR36 y TBR177 fue de 37,03 y -4,37, respectivamente. A partir de estas razones de posibilidades, la probabilidad de que las muestras de plasma procedieran de pacientes con CHC se calculó en >99,9 % y 1 %, respectivamente. En resumen, TBR36 tenía una alta probabilidad de ser una muestra de un paciente con CHC, mientras que TBR177 tenía una baja probabilidad de ser una muestra de un paciente con CHC.

Característica	Coeficiente de regresión	Puntuación z de la característica respectiva	
		TBR36	TBR177
X1	-2,9575	14,8	0
X2	2,2534	21,3	0
X3	-1,5099	6,1	0
X4	-0,236	34,0	0

(continuación)

Característica	Coeficiente de regresión	Puntuación z de la característica respectiva	
		TBR36	TBR177
X5	0,7426	17,3	0
X6	-0,6682	-26,3	-1,5
X7	-0,2828	-13,9	-2,6
X8	-0,7281	-9,4	-4,4
X9	1,0581	-7,8	-3,7
X10	0,3877	-20,8	-4,3
X11	0,3534	-15,5	-3,1
X12	-1,1826	4,8	3,3
X13	-0,3805	-11,7	-1,4
ln(razón de posibilidades)		37,03	-4,37463

Tabla 7

En otras realizaciones, se pueden utilizar la regresión jerárquica de agrupamientos, el análisis de árbol de clasificación y otros modelos de regresión para determinar el probable origen primario del cáncer.

XII. MATERIALES Y MÉTODOS

A. Preparación de bibliotecas de ADN y secuenciación tratadas con bisulfito

Se fragmentó el ADN genómico (5 µg) añadido con 0,5 % (p/p) de ADN lambda no metilado (Promega) mediante un sistema Covaris S220 (Covaris) hasta alcanzar una longitud aproximada de 200 pb. Las bibliotecas de ADN se prepararon utilizando el kit de preparación de muestras para la secuenciación de ambos extremos (Illumina) de acuerdo con las instrucciones del fabricante, excepto que los adaptadores metilados (Illumina) se ligaron a los fragmentos de ADN. Tras dos series de purificación utilizando perlas magnéticas AMPure XP (Beckman Coulter), se dividieron los productos de ligación en 2 porciones, una de las cuales se sometió a 2 series de modificación con bisulfito con un kit EpiTect Bisulfite (Qiagen). Las citosinas no metiladas en los sitios CpG de las inserciones se convirtieron en uracilos, mientras que las citosinas metiladas permanecieron inalteradas. Las moléculas de ADN ligadas al adaptador, tratadas o no tratadas con bisulfito sódico, se enriquecieron mediante 10 ciclos de PCR mediante al siguiente protocolo: 2,5 U de ADN polimerasa PfuTurboCx hotstart (Agilent Technologies), 1 x tampón de reacción PfuTurboCx, dNTP 25 µM, 1 µl de cebador de PCR PE 1.0 y 1 µl de cebador de PCR PE 2.0 (Illumina) en una reacción de 50 µl. El perfil de termociclado fue: 95 °C durante 2 min, 98 °C durante 30 s, luego 10 ciclos a 98 °C durante 15 s, 60 °C durante 30 s y 72 °C durante 4 min, con una etapa final de 72 °C durante 10 min (R Lister, *et al.* 2009 *Nature*; 462: 315-322). Los productos de la PCR se purificaron utilizando perlas magnéticas AMPure XP.

El ADN plasmático extraído de 3,2-4 ml de muestras de plasma materno se enriqueció con ADN lambda fragmentado (25 pg por ml de plasma) y se sometió a la construcción de bibliotecas como se ha descrito anteriormente (R. W. K. Chiu *et al.* 2011 *BMJ*; 342: c7401). Despues de ligar a los adaptadores metilados, se dividieron los productos de ligación en 2 mitades y una porción se sometió a 2 series de modificación con bisulfito. A continuación, los productos de ligación tratados con bisulfito o no tratados se enriquecieron mediante 10 ciclos de PCR como se ha descrito anteriormente.

Las bibliotecas de ADN tratadas con bisulfito o no tratadas se secuenciaron para 75 pb en formato de ambos extremos en instrumentos HiSeq2000 (Illumina). Los grupos de ADN se generaron con un kit de generación de grupos de ambos extremos v3 en un instrumento cBot (Illumina). El análisis de imágenes en tiempo real y la lectura de bases se realizaron utilizando el software HiSeq Control (HCS) v1.4 y el software Real Time Analysis (RTA) v1.13 (Illumina), mediante el que los cálculos automatizados de la matriz y el desfase se basaron en el control PhiX v3 secuenciado con las bibliotecas de ADN.

B. Alineación de secuencias e identificación de citosinas metiladas

Después la lectura de bases, se eliminaron las secuencias adaptadoras y las bases de baja calidad (es decir, puntuación de calidad <20) en los extremos de los fragmentos. A continuación, se procesaron las lecturas recortadas en formato FASTQ mediante una línea de análisis de datos de metilación denominada Methyl-Pipe (P Jiang, *et al.* Methyl-Pipe: An integrated bioinformatics data analysis pipeline for whole genome methylome analysis, ponencia

presentada en la IEEE International Conference on Bioinformatics and Biomedicine Workshops, Hong Kong, del 18 al 21 de diciembre de 2010). Para alinear las lecturas de secuenciación convertidas con bisulfito, primero realizamos una conversión informática de todos los restos de citosina a timina, en las cadenas de Watson y Crick por separado, utilizando el genoma humano de referencia (NCBI build 36/hg18). A continuación, realizamos la conversión informática de cada citosina en timina en todas las lecturas procesadas y conservamos la información posicional de cada resto convertido. SOAP2 (R Li, *et al.* 2009 *Bioinformatics*; 25: 1966-1967) para alinear las lecturas convertidas con los dos genomas humanos de referencia preconvertidos, con un máximo de dos desajustes permitidos para cada lectura alineada. Solo se seleccionaron las lecturas asignables a una ubicación genómica única. Se eliminaron las lecturas ambiguas que correspondían tanto a la cadena de Watson como a la de Crick y las lecturas duplicadas (clonales) que tenían las mismas posiciones genómicas inicial y final. Las lecturas secuenciadas con un tamaño de inserción ≤600 pb se retuvieron para los análisis de metilación y tamaño.

Los restos de citosina en el contexto del dinucleótido CpG fueron los principales objetivos de los estudios posteriores de metilación del ADN. Después de la alineación, las citosinas presentes originalmente en las lecturas secuenciadas se recuperaron basándose en la información posicional conservada durante la conversión informática. Las citosinas recuperadas entre los dinucleótidos CpG se puntuaron como metiladas. Las timinas entre los dinucleótidos CpG se puntuaron como no metiladas. El ADN lambda no metilado incluido durante la preparación de la biblioteca sirvió como control interno para estimar la eficacia de la modificación con bisulfito sódico.

Todas las citosinas del ADN lambda deberían haberse convertido en timinas si la eficacia de conversión del bisulfito fuera del 100 %.

XIII. SUMARIO

Con el uso de las realizaciones descritas, en el presente documento, se podría cribar, detectar, hacer un seguimiento o pronosticar el cáncer de forma no invasiva utilizando, por ejemplo, el plasma de un sujeto. También se podría realizar un cribado prenatal, diagnóstico, investigación o seguimiento de un feto mediante la deducción del perfil de metilación del ADN fetal a partir del plasma materno. Para ilustrar la potencia del enfoque, demostramos que la información que convencionalmente se obtenía mediante el estudio de los tejidos placentarios podía evaluarse directamente a partir del plasma materno. Por ejemplo, el estado de sellado de los locus génicos, la identificación de los locus con metilación diferencial entre el ADN fetal y materno, y la variación gestacional en el perfil de metilación de los locus génicos se lograron mediante el análisis directo del ADN plasmático materno. La principal ventaja de nuestro enfoque es que el metiloma fetal podría evaluarse de forma exhaustiva durante la gestación sin interrupción del mismo ni necesidad de tomar muestras invasivas de tejidos fetales. Dada la conocida asociación entre el estado alterado de metilación del ADN y las numerosas afecciones asociadas a la gestación, el enfoque descrito en este estudio puede servir como una herramienta importante para investigar la fisiopatología y la identificación de biomarcadores para esas condiciones. Centrándonos en los locus sellados, demostramos que tanto los perfiles de metilación fetal transmitidos por el padre como los transmitidos por la madre podían evaluarse a partir del plasma materno. Este enfoque puede ser potencialmente útil para la investigación de enfermedades relacionadas con el sellado genómico. También se pueden aplicar algunas realizaciones directamente a la evaluación prenatal de enfermedades fetales o asociadas a la gestación.

Hemos demostrado que es posible aplicar la secuenciación con bisulfito de todo el genoma para investigar el perfil de metilación del ADN de los tejidos placentarios. Hay aproximadamente 28 millones de sitios CpG en el genoma humano (C Clark *et al.* 2012 *PLoS One*; 7: e50233). Nuestros datos de secuenciación con bisulfito de la muestra de tejido de la placenta a término y de las MVC cubrieron más del 80 % de los CpG. Esto representa una cobertura sustancialmente más amplia que la que se puede conseguir utilizando otras plataformas de alto rendimiento. Por ejemplo, la micromatriz de microesferas Infinium HumanMethylation 27K de Illumina que se utilizó en un estudio anterior sobre tejidos placentarios (T Chu *et al.* 2011 *PLoS One*; 6: e14723) solo cubría el 0,1 % de los CpG del genoma. La micromatriz de microesferas Infinium HumanMethylation 450K de Illumina que estaba disponible más recientemente solo cubría el 1,7 % de los CpG (C. Clark *et al.* 2012 *PLoS One*; 7: e50233). Dado que el enfoque MPS está exento de restricciones relacionadas con el diseño de la sonda, la eficiencia de hibridación o la fuerza de captura de anticuerpos, fue posible evaluar los CpG dentro o fuera de las islas CpG y en la mayoría de los contextos de secuencia.

XIV. SISTEMA INFORMÁTICO

Cualquiera de los sistemas informáticos mencionados en el presente documento puede utilizar cualquier número adecuado de subsistemas.

En la FIG. 33, en el aparato informático 3300 se muestran ejemplos de dichos subsistemas. En algunas realizaciones, un sistema informático incluye un único aparato informático, donde los subsistemas pueden ser los componentes del aparato informático. En otras realizaciones, un sistema informático puede incluir varios aparatos informáticos, siendo cada uno de ellos un subsistema, con componentes internos.

Los subsistemas mostrados en la FIG. 33 están interconectados a través de un bus de sistema 3375. Se muestran subsistemas adicionales tales como una impresora 3374, un teclado 3378, dispositivo(s) de almacenamiento 3379, un

monitor 3376, que se acopla al adaptador de visualización 3382, y otros. Los periféricos y los dispositivos de entrada/salida (E/S), que se acoplan al controlador de E/S 3371, pueden conectarse al sistema informático por cualquier número de medios conocidos en la técnica, tal como el puerto serie 3377. Por ejemplo, puerto serie 3377 o la interfaz externa 3381 (p. ej. Ethernet, Wi-Fi, etc.) pueden utilizarse para conectar el sistema informático 3300 a una red de área amplia, tal como Internet, un dispositivo de entrada de ratón o un escáner. La interconexión mediante el bus de sistema 3375 permite que el procesador central 3373 se comunique con cada subsistema y controle la ejecución de instrucciones procedentes de la memoria del sistema 3372 o de los dispositivos de almacenamiento 3379 (p. ej., un disco fijo), así como el intercambio de información entre subsistemas. La memoria del sistema 3372 y/o los dispositivos de almacenamiento 3379 pueden ser realizaciones de un medio legible por ordenador. Cualquiera de los valores mencionados en el presente documento puede ser la salida de un componente a otro componente y puede ser la salida al usuario.

Un sistema informático puede incluir una pluralidad de los mismos componentes o subsistemas, p. ej., conectados entre sí por la interfaz externa 3381 o por una interfaz interna. En algunas realizaciones, los sistemas informáticos, subsistema o aparatos pueden comunicarse a través de una red. En tales casos, un ordenador puede considerarse como cliente y otro ordenador como servidor, donde cada uno puede formar parte de un mismo sistema informático. Un cliente y un servidor pueden incluir cada uno múltiples sistemas, subsistemas o componentes.

Debe entenderse que cualquiera de las realizaciones de la presente invención puede implementarse en forma de lógica de control utilizando un hardware (p. ej., un circuito integrado específico de la aplicación o una matriz de selección programable por campo) y/o utilizando un software de ordenador con un procesador generalmente programable de manera modular o integrada. Como se utiliza en el presente documento, un procesador incluye un procesador multinúcleo en un mismo chip integrado, o varias unidades de procesamiento en una única placa de circuito o conectadas en red. Basándose en la divulgación y las enseñanzas proporcionadas en el presente documento, un experto habitual en la materia conocerá y percibirá otras formas y/o métodos para implementar las realizaciones de la presente invención utilizando equipos físicos y una combinación de equipos físicos y programas informáticos.

Puede implementarse cualquiera de los componentes o las funciones de programas informáticos descritos en la presente solicitud como código de programación para su ejecución mediante un procesador utilizando cualquier lenguaje informático adecuado, tal como, por ejemplo, Java, C++ o Perl utilizando, por ejemplo, técnicas convencionales u orientadas por objetos. El código de software puede almacenarse como una serie de instrucciones o comandos en un medio legible por ordenador para su almacenamiento y/o transmisión, los medios adecuados incluyen la memoria de acceso aleatorio (RAM), una memoria de solo lectura (ROM), un medio magnético tal como un disco duro o disquete o un medio óptico tal como un disco compacto (CD) o un DVD (disco digital versátil), memoria flash y similares. El medio legible por ordenador puede ser cualquier combinación de tales dispositivos de almacenamiento o transmisión.

Dichos programas también pueden codificarse y transmitirse utilizando señales portadoras adaptadas para la transmisión a través de redes cableadas, ópticas y/o inalámbricas que se ajusten a una variedad de protocolos, incluida Internet. Así pues, se puede crear un medio legible por ordenador de acuerdo con una realización de la presente invención utilizando una señal de datos codificada con dichos programas. Los medios legibles por ordenador codificados con el código del programa pueden empaquetarse con un dispositivo compatible o proporcionarse por separado de otros dispositivos (p. ej., a través de descarga de Internet). Cualquier medio legible por ordenador puede residir en o dentro de un solo producto de programa de ordenador (p. ej., un disco duro, un CD o un sistema informático completo) y puede estar presente en o dentro de diferentes productos de programa de ordenador dentro de un sistema o red. Un sistema informático puede incluir un monitor, una impresora u otro dispositivo de presentación adecuado para proporcionar a un usuario cualquiera de los resultados mencionados en el presente documento.

Cualquiera de los métodos descritos en el presente documento puede realizarse total o parcialmente con un sistema informático que incluya uno o más procesadores, que pueden configurarse para realizar las etapas. Por tanto, las realizaciones pueden dirigirse a sistemas informáticos configurados para realizar las etapas de cualquiera de los métodos descritos en el presente documento, potencialmente con diferentes componentes que realizan una etapa o un grupo de etapas respectivas. Aunque se presentan como etapas numeradas, las etapas de los métodos descritos en el presente documento pueden realizarse al mismo tiempo o en un orden diferente. Adicionalmente, pueden usarse partes de estas etapas con partes de otras etapas de otros métodos. Del mismo modo, la totalidad de una etapa o partes de la misma pueden ser opcionales. Adicionalmente, cualquiera de las etapas de cualquiera de los métodos puede realizarse con módulos, circuitos u otros medios para realizar estas etapas.

Los detalles específicos de las realizaciones particulares pueden combinarse de cualquier manera adecuada sin apartarse del espíritu ni del alcance de las realizaciones de la invención. Sin embargo, otras realizaciones de la invención pueden dirigirse a realizaciones específicas relativas a cada aspecto individual, o a combinaciones específicas de estos aspectos individuales.

La descripción anterior de las realizaciones ilustrativas de la invención se ha presentado con fines ilustrativos y descriptivos. No pretenden ser exhaustivos ni limitar la invención a la forma precisa descrita y son posibles muchas modificaciones y variaciones a la luz de las enseñanzas anteriores. Las realizaciones se eligieron y se describieron

para explicar mejor los principios de la invención y su aplicación práctica para permitir así que otros expertos en la materia utilicen mejor la invención en diversas realizaciones y con diversas modificaciones que se adapten al uso particular contemplado.

- 5 Una cita de "un", "una" o "el/la" se entiende como "uno o más" a menos que se indique específicamente lo contrario.
de cualquiera de los métodos puede realizarse con módulos, circuitos u otros medios para realizar estas etapas.
- 10 Los detalles específicos de las realizaciones particulares pueden combinarse de cualquier manera adecuada sin apartarse del alcance de las realizaciones de la invención. Sin embargo, otras realizaciones de la invención pueden dirigirse a realizaciones específicas relativas a cada aspecto individual, o a combinaciones específicas de estos aspectos individuales.
- 15 La descripción anterior de las realizaciones ilustrativas de la invención se ha presentado con fines ilustrativos y descriptivos. No pretenden ser exhaustivos ni limitar la invención a la forma precisa descrita y son posibles muchas modificaciones y variaciones a la luz de las enseñanzas anteriores. Las realizaciones se eligieron y se describieron para explicar mejor los principios de la invención y su aplicación práctica para permitir así que otros expertos en la materia utilicen mejor la invención en diversas realizaciones y con diversas modificaciones que se adapten al uso particular contemplado.
- 20 Una cita de "un", "una" o "el/la" se entiende como "uno o más" a menos que se indique específicamente lo contrario.

Tabla S2A. Lista de las 100 regiones más hipermetiladas identificadas a partir de muestras de vellosidades coriónicas del primer trimestre y de glóbulos sanguíneos maternos.

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	MVC	Valores de P	Diferencia de metilación
cr. 13	113063600	113064100	500	0,009	0,9	3,67E-15	0,891
cr. 6	36279700	36280200	500	0,0068.	0,8957	2,39E-22	0,8889
cr. 16	66876000	66876500	500	0,0327	0,9211	3,82E-21	0,8884
cr. 10	163500	164000	500	0,0195	0,9034	3,60E-35	0,8839
cr. 9	3518300	3518800	500	0,0263	0,9045	1,32E-26	0,8782
cr. 12	31877100	31877600	500	0,007	0,8784	3,08E-22	0,8714
cr. 22	37477400	37478400	1000	0,0152	0,8848	0,00E+00	0,8696
cr. 4	148940500	148941000	500	0,0055	0,8717	4,40E-29	0,8662
cr. 5	131836300	131836800	500	0,075	0,9403	1,54E-10	0,8653
cr. 17	26661700	26663600	1900	0,0187	0,875	2,95E-38	0,8563
cr. 2	105758600	105759600	1000	0,0305	0,8828	1,19E-53	0,8523
cr. 22	39188800	39189800	1000	0	0,8514	2,05E-46	0,8514
cr. 3	153443900	153444900	1000	0,0436	0,8945	5,43E-34	0,8509
cr. 6	25149600	25150600	1000	0,0135	0,8632	0,00E+00	0,8497
cr. 5	98296800	98297300	500	0,0432	0,8925	4,97E-23	0,8493
cr. 7	150679900	150680400	500	0,0496	0,8944	6,50E-17	0,8448
cr. 7	107563100	107563600	500	0,0495	0,8895	9,58E-26	0,84
cr. 7	37348300	37349300	1000	0,0012	0,8409	0,00E+00	0,8397
cr. 14	58837800	58838300	500	0,0097	0,848	3,35E-16	0,8383
cr. 6	119238100	119238600	500	0,0899	0,928	2,38E-19	0,8381
cr. 16	93669900	93670400	500	0,0753	0,913	2,19E-10	0,8377
cr. 17	26669200	26670200	1000	0,0221	0,859	1,44E-29	0,8369

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	MVC	Valores de P	Diferencia de metilación
cr. 2	88108100	88108600	500	0,075	0,9109	3,55E-17	0,8359
cr. 13	98363800	98364300	500	0,11	0,9457	1,28E-11	0,8357
cr. 16	66948000	66948500	500	0,0331	0,8685	0,00E+00	0,8354
cr. 6	42098000	42098500	500	0,0484	0,8835	3,73E-16	0,8351
cr. 3	129876000	129876500	500	0,0565	0,8897	8,81E-17	0,8332
cr. 3	142700300	142700800	500	0,0063	0,8393	2,59E-22	0,833
cr. 8	145883800	145884300	500	0,0392	0,872	0,00E+00	0,8328
cr. 10	8320700	8321200	500	0,0566	0,8871	9,40E-09	0,8305
cr. 3	120438100	120438600	500	0,102	0,9292	7,09E-16	0,8272
cr. 3	173792600	173793100	500	0,0182	0,8453	2,84E-39	0,8271
cr. 17	40320700	40321200	500	0,0539	0,8788	6,50E-30	0,8249
cr. 15	72076200	72076700	500	0,0299	0,8525	6,48E-10	0,8226
cr. 16	29663900	29665400	1500	0,0081	0,8305	0,00E+00	0,8224
cr. 11	66961100	66962100	1000	0,0489	0,8712	0,00E+00	0,8223
cr. 9	27083100	27084100	1000	0,097	0,9177	2,37E-55	0,8207
cr. 9	111249600	111250100	500	0,0613	0,8795	1,99E-20	0,8182
cr. 14	101412400	101412900	500	0	0,8167	8,26E-32	0,8167
cr. 1	242549200	242549700	500	0	0,8155	3,50E-21	0,8155
cr. 8	38642800	38643300	500	0,0191	0,8346	3,22E-41	0,8155
cr. 4	85893600	85894100	500	0,0394	0,8533	1,45E-15	0,8139
cr. 5	142368600	142369100	500	0,0385	0,8523	1,18E-18	0,8138
cr. 8	130969500	130970000	500	0,069	0,8824	2,42E-24	0,8134
cr. 2	196783900	196784400	500	0	0,8123	3,63E-40	0,8123
cr. 16	49258100	49258600	500	0,0733	0,8851	4,29E-18	0,8118
cr. 1	232601200	232601700	500	0,0594	0,8707	1,73E-13	0,8113
cr. 1	109039500	109040000	500	0,0366	0,8471	1,07E-11	0,8105
cr. 17	59491300	59491800	500	0,0662	0,8758	2,15E-17	0,8096
cr. 21	42194100	42194600	500	0,11	0,9182	1,61E-12	0,8082
cr. 9	116174500	116175500	1000	0,0062	0,8132	1,98E-60	0,807
cr. 15	73429200	73429700	500	0	0,8066	9,81E-33	0,8066
cr. β	157462800	157463300	500	0,0758	0,8819	7,94E-16	0,8061
cr. 3	16858500	16859500	1000	0,0021	0,8068	4,76E-68	0,8047
cr. 9	96662800	96663300	500	0,0614	0,8651	6,79E-28	0,8037
cr. 9	88143000	88143500	500	0,1538	0,9559	7,43E-09	0,8021
cr. 19	16090000	16091000	1000	0,0899	0,8904	1,60E-53	0,8005
cr. 15	29436300	29437300	1000	0,0553	0,8556	1,18E-80	0,8003
cr. 11	77816100	77816600	500	0,1069	0,9068	2,31E-17	0,7999

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	MVC	Valores de P	Diferencia de metilación
cr. 10	30346800	30347300	500	0,1212	0,9211	7,48E-07	0,7999
cr. 1	89510300	89511300	1000	0,0203	0,8191	3,53E-77	0,7988
cr. 3	125986100	125986600	500	0,1686	0,9674	5,24E-22	0,7988
cr. 19	60162800	60163300	500	0,0127	0,8113	9,99E-19	0,7986
cr. 16	73655900	73656900	1000	0,082	0,8806	4,48E-41	0,7986
cr. 16	30104300	30105800	1500	0,0298	0,8282	0,00E+00	0,7984
cr. 10	118642400	118642900	500	0,0588	0,8571	8,63E-11	0,7983
cr. 16	4495000	4496000	1000	0,0632	0,8615	2,27E-44	0,7983
cr. 1	2048300	2048800	500	0,0309	0,8289	1,19E-80	0,798
cr. 2	136481800	136482800	1000	0,0554	0,8533	8,50E-48	0,7979
cr. 10	29959200	29959700	500	0,1429	0,94	2,60E-08	0,7971
cr. 6	139642400	139642900	500	0,0618	0,8585	2,16E-29	0,7967
cr. 14	69825300	69825800	500	0,0654	0,8615	6,85E-14	0,7961
cr. 8	49739700	49740200	500	0,0324	0,828	2,88E-30	0,7956
cr. 17	42205700	42206200	500	0,057	0,852	2,11E-30	0,795
cr. 4	77445300	77445800	500	0,0442	0,8377	1,79E-35	0,7935
cr. 17	53762700	53766300	3600	0,0003	0,7926	0,00E+00	0,7923
cr. 17	44269900	44270400	500	0,026	0,8182	3,49E-21	0,7922
cr. 6	42462700	42463200	500	0,0761	0,8678	4,74E-22	0,7917
cr. 2	23396200	23396700	500	0,0333	0,8235	1,25E-14	0,7902
cr. 9	100921100	100921600	500	0,0244	0,814	3,32E-21	0,7896
cr. 7	74016100	74016600	500	0,1442	0,9333	6,74E-10	0,7891
cr. 6	157879000	157879500	500	0,133	0,9219	6,36E-17	0,7889
cr. 3	3189400	3190400	1000	0,0693	0,8571	1,38E-24	0,7878
cr. 16	29581500	29584500	3000	0,0081	0,7956	0,00E+00	0,7875
cr. 17	42201800	42202800	1000	0,0884	0,8751	0,00E+00	0,7867
cr. 11	94257000	94257500	500	0,1122	0,8986	4,29E-10	0,7864
cr. 10	14741600	14742100	500	0,0139	0,8	1,73E-20	0,7861
cr. 21	33826900	33827400	500	0,0879	0,8739	2,81E-11	0,786
cr. 4	130057200	130057700	500	0,0893	0,875	1,76E-13	0,7857
cr. 21	35343400	35343900	500	0	0,7853	7,43E-18	0,7853
cr. 12	105372800	105373300	500	0,0923	0,8767	8,67E-22	0,7844
cr. 6	10799800	10800300	500	0,1429	0,9263	8,21E-17	0,7834
cr. 5	16753100	16753600	500	0,041	0,8241	1,40E-15	0,7831
cr. 3	135746000	135746500	500	0,1429	0,9259	2,86E-09	0,783
cr. 6	53708300	53708800	500	0,0412	0,8235	2,74E-31	0,7823
cr. 2	128122900	128123400	500	0,0634	0,8455	4,82E-21	0,7821
cr. 6	150574200	150574700	500	0,0876	0,8696	1,56E-20	0,782

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	MVC	Valores de P	Diferencia de metilación
cr. 16	84326000	84327000	1000	0,1071	0,8891	3,58E-61	0,782
cr. 1	26744500	26745500	1000	0,0336	0,8152	0,00E+00	0,7816
cr. 2	234882000	234882500	500	0,0392	0,819	7,63E-14	0,7798

Tabla S2B. Lista de las 100 regiones más hipometiladas identificadas a partir de muestras de vellosidades coriónicas del primer trimestre y de glóbulos sanguíneos maternos.

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	MVC	Valores de P	Diferencia de metilación
cr. 18	12217500	12218500	1000	0,9873	0	3,05E-25	0,9873
cr. 17	22885400	22885900	500	0,9714	0,0161	8,92E-12	0,9553
cr. 3	184827100	184827600	500	0,9875	0,033	4,79E-16	0,9545
cr. 5	148968300	148968800	500	0,98	0,0426	6,70E-09	0,9374
cr. 10	104794500	104795000	500	0,973	0,0385	9,33E-10	0,9345
cr. 4	84977900	84978400	500	0,9643	0,0417	2,98E-08	0,9226
cr. 3	180395300	180395800	500	0,9877	0,0667	6,72E-08	0,921
cr. 2	138908300	138908800	500	0,939	0,0208	1,10E-16	0,9182
cr. 6	139873100	139873600	500	0,9667	0,0526	1,29E-07	0,914
cr. 8	59604700	59605200	500	0,9468	0,033	2,88E-14	0,9138
cr. 6	167622300	167622800	500	0,9452	0,0316	3,86E-14	0,9136
cr. 3	175701300	175701800	500	0,9846	0,0735	7,43E-10	0,9111
cr. 13	59246400	59246900	500	0,9402	0,0313	2,3E-11	0,9089
cr. 12	71263600	71264100	500	0,9296	0,0213	1,08E-08	0,9083
cr. 5	39459400	39459900	500	0,9219	0,014	5,01E-22	0,9079
cr. 17	24904700	24905200	500	0,9161	0,0092	5,04E-35	0,9069
cr. 12	31889900	31890400	500	0,9524	0,0465	6,78E-13	0,9059
cr. 3	152897800	152898300	500	0,9402	0,0345	1,70E-17	0,9057
cr. 1	40378700	40379200	500	0,9565	0,0526	3,31E-09	0,9039
cr. 12	43979300	43979800	500	0,952	0,05	6,68E-13	0,902
cr. 18	1395900	1397400	1500	0,9308	0,0293	0,00E+00	0,9015
cr. 1	223482900	223483400	500	0,9579	0,0575	3,36E-24	0,9004
cr. 9	130357000	130357500	500	0,9282	0,0286	9,19E-13	0,8996
cr. 3	72878300	72878800	500	0,9612	0,0625	8,20E-14	0,8987
cr. 7	84347200	84348700	1500	0,9401	0,0418	0,00E+00	0,8983
cr. 15	37317500	37318000	500	0,9358	0,0385	7,58E-14	0,8973
cr. 8	42528600	42529100	500	0,9302	0,0337	1,73E-14	0,8965
cr. 6	134914000	134914500	500	0,9037	0,0076	4,84E-21	0,8961
cr. 13	56207100	56208100	1000	0,9184	0,0245	0,00E+00	0,894
cr. 2	209074000	209074500	500	0,9309	0,037	6,13E-27	0,8938

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	MVC	Valores de P	Diferencia de metilación
cr. 12	74021100	74022100	1000	0,9513	0,058	0,00E+00	0,8933
cr. 4	118939300	118939800	500	0,9192	0,0276	6,58E-27	0,8916
cr. 5	12626600	12628600	2000	0,9266	0,0355	0,00E+00	0,8911
cr. 5	105517300	105518300	1000	0,927	0,0359	0,00E+00	0,891
cr. 12	70056300	70057300	1000	0,9488	0,0609	0,00E+00	0,888
cr. 6	153238200	153239200	1000	0,9123	0,0244	0,00E+00	0,8879
cr. 17	60374800	60375300	500	0,9655	0,0777	3,64E-14	0,8878
cr. 14	68272700	68273200	500	0,9389	0,0523	1,23E-22	0,8866
cr. 19	54533800	54534800	1000	0,9117	0,0262	0,00E+00	0,8855
cr. 12	15392200	15393200	1000	0,9307	0,0457	0,00E+00	0,885
cr. 1	212517400	212517900	500	0,9266	0,0417	9,81E-12	0,8849
cr. 10	49344400	49345400	1000	0,9422	0,0579	0,00E+00	0,8844
cr. 3	47410400	47410900	500	0,9213	0,0381	5,59E-16	0,8832
cr. 3	879500	880000	500	0,9455	0,0625	8,06E-06	0,883
cr. 2	31572400	31573400	1000	0,9176	0,0357	0,00E+00	0,8819
cr. 1	89131200	89131700	500	0,9314	0,0498	5,15E-70	0,8816
cr. 8	94832000	94832500	500	0,9156	0,0351	2,16E-65	0,8805
cr. 7	14008300	14009800	1500	0,9349	0,0545	0,00E+00	0,8804
cr. 12	12971300	12972300	1000	0,9361	0,0559	0,00E+00	0,8802
cr. 5	43114700	43115200	500	0,9638	0,0842	1,79E-13	0,8796
cr. 11	107872400	107872900	500	0,9472	0,0677	2,31E-32	0,8794
cr. 8	49757600	49758100	500	0,9048	0,0269	3,15E-52	0,8779
cr. 13	33106400	33106900	500	0,9384	0,0606	9,54E-15	0,8778
cr. 3	190658800	190659300	500	0,9388	0,0617	2,71E-22	0,877
cr. 1	181508000	181508500	500	0,9259	0,0495	3,78E-15	0,8764
cr. 1	180436900	180437400	500	0,9412	0,0652	2,36E-13	0,876
cr. 6	122642800	122643800	1000	0,9218	0,0458	0,00E+00	0,8759
cr. 5	166429300	166429800	500	0,9551	0,08	5,26E-05	0,8751
cr. 12	14972900	14973400	500	0,9483	0,0733	2,10E-18	0,8749
cr. 5	123933900	123934400	500	0,943	0,0683	1,12E-39	0,8746
cr. 2	15969400	15970400	1000	0,8939	0,0196	9,43E-46	0,8743
cr. 3	167635200	167636200	1000	0,9363	0,0625	9,03E-41	0,8738
cr. 5	159442700	159443200	500	0,9174	0,044	6,27E-14	0,8734
cr. 4	48027200	48027700	500	0,9839	0,1111	8,89E-06'	0,8728
cr. 6	140071500	140072000	500	0,9234	0,0506	4,39E-33	0,8728
cr. 10	22356300	22356800	500	0,9548	0,0822	1,04E-18	0,8726

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	MVC	Valores de P	Diferencia de metilación
cr. 8	61007300	61007800	500	0,9197	0,0476	1,24E-15	0,8721
cr. 11	95463500	95464000	500	0,9348	0,0629	1,16E-20	0,8718
cr. 2	216399800	216400300	500	0,938	0,0667	5,98E-06	0,8713
cr. 18	57359700	57360200	500	0,9293	0,0584	1,77E-19	0,871
cr. 3	102734400	102734900	500	0,8917	0,0207	7,94E-22	0,871
cr. 1	173605700	173606200	500	0,96	0,0891	5,46E-13	0,8709
cr. 2	86993700	86995700	2000	0,8965	0,0261	0,00E+00	0,8704
cr. 3	162621100	162621600	500	0,9226	0,0526	7,89E-38	0,8699
cr. 12	10144800	10145300	500	0,929	0,0598	3,45E-17	0,8691
cr. 3	113855100	113855600	500	0,9667	0,0982	3,97E-14	0,8685
cr. 2	156958200	156959200	1000	0,9252	0,0571	8,89E-50	0,8681
cr. 2	55775000	55776000	1000	0,9159	0,0483	0,00E+00	0,8676
cr. 6	124898400	124898900	500	0,8987	0,0313	1,91E-15	0,8675
cr. 5	420Q3700	42004700	1000	0,9262	0,0588	0,00E+00	0,8674
cr. 3	24162200	24162700	500	0,883	0,0161	1,75E-27	0,8668
cr. 6	35394000	35395000	1000	0,9204	0,0539	0,00E+00	0,8665
cr. 17	8451800	8453300	1500	0,9376	0,0714	0,00E+00	0,8662
cr. 14	53487700	53488700	1000	0,9013	0,0353	0,00E+00	0,866
cr. 7	98572800	98573300	500	0,9651	0,0995	8,37E-26	0,8656
cr. 6	52298700	52299200	500	0,9427	0,0772	1,31E-28	0,8655
cr. 6	159047900	159048400	500	0,908	0,0426	3,34E-08	0,8655
cr. 14	22152600	22153100	500	0,9085	0,0435	4,43E-17	0,865
cr. 12	103285000	103285500	500	0,9321	0,0674	0,00E+00	0,8647
cr. 7	43302200	43302700	500	0,968	0,1037	6,40E-16	0,8643
cr. 14	22247400	22247900	500	0,9804	0,1163	3,83E-06	0,8641
cr. 2	66780900	66781400	500	0,9355	0,0714	8,37E-09	0,8641
cr. 12	97393000	97393500	500	0,9045.	0,0408	3,46E-21	0,8637
cr. 5	162797900	162798900	1000	0,9271	0,0635	1,75E-57	0,8636
cr. 2	83598400	83599400	1000	0,9354	0,0719	0,00E+00	0,8635
cr. 11	111358800	111359300	500	0,9156	0,0523	4,15E-24	0,8632
cr. 11	104891100	104892600	1500	. 0,9164	0,0533	0,00E+00	0,863
cr. 1	184583600	184584100	500	0,9647	0,1026	3,02E-13	0,8621
cr. 5	132350500	132351500	1000	0,9042	0,0426	1,86E-36	0,8616
cr. 5	53268300	53268800	500	0,972	0,1111	4,76E-16	0,8609

Tabla S2C. Lista de las 100 regiones más hipermetiladas identificadas a partir de tejido placentario y glóbulos sanguíneos maternos del tercer trimestre.

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	Placenta a término	Valores de P	Diferencia de metilación
cr. 4	78129700	78130200	500	0,0488	0,9747	3,97E-33	0,926
cr. 5	131467400	131467900	500	0,0213	0,9275	7,10E-27	0,9063
cr. 17	26661700	26663600	1900	0,0187	0,9226	1,79E-41	0,9039
cr. 4	148940500	148941000	500	0,0055	0,9079	1,82E-29	0,9024
cr. 9	100921100	100921600	500	0,0244	0,9242	1,38E-25	0,8998
cr. 6	137114200	137114700	500	0	0,8934	8,87E-14	0,8934
cr. 3	173792600	173793100	500	0,0182	0,9091	1,70E-42	0,8908
cr. 5	98296800	98297300	500	0,0432	0,9333	2,58E-23	0,8901
cr. 12	44898000	44898500	500	0	0,8889	4,47E-11	0,8889
cr. 3	197328900	197329400	500	0,0169	0,9048	5,55E-10	0,8878
cr. 8	49739700	49740200	500	0,0324	0,9194	5,71E-34	0,887
cr. 12	122279300	122279800	500	0,0135	0,8969	3,46E-21	0,8834
cr. 17	43092200	43092700	500	0	0,8824	4,34E-10	0,8824
cr. 7	107563100	107563600	500	0,0495	0,931	1,05E-28	0,8815
cr. 11	72543200	72543700	500	0,0377	0,9167	2,94E-09	0,8789
cr. 14	58837800	58838300	500	0,0097	0,886	9,16E-18	0,8763
cr. 3	153443900	153444900	1000	0,0436	0,9197	6,24E-39	0,876
cr. 3	16953200	16953700	500	0,0896	0,9655	6,78E-09	0,876
cr. 17	42205700	42206200	500	0,057	0,933	1,13E-31	0,8759
cr. 6	53217600	53218100	500	0,0818	0,9571	1,54E-19	0,8754
cr. 3	112749000	112749500	500	0,0403	0,9154	4,11E-22	0,8752
cr. 8	22453700	22454200	500	0,003	0,8765	1,64E-50	0,8735
cr. 1	162860900	162861400	500	0,023	0,8932	8,37E-14	0,8702
cr. 6	36279700	36280200	500	0,0068	0,8762	1,14E-21	0,8694
cr. 5	80962500	80963000	500	0	0,8679	2,08E-15	0,8679
cr. 16	11312500	11313000	500	0	0,8679	2,14E-10	0,8679
cr. 16	29663900	29665400	1500	0,0081	0,8759	0,00E+00	0,8678
cr. 3	120438100	120438600	500	0,102	0,9639	2,98E-15	0,8618
cr. 8	134157000	134157500	500	0,0625	0,9219	6,10E-20	0,8594
cr. 6	42620900	42621400	500	0	0,8571	5,68E-08	0,8571
cr. 5	131836300	131836800	500	0,075	0,9315	1,26E-10	0,8565
cr. 14	60290000	60290500	500	0	0,8544	5,63E-14	0,8544
cr. 6	42850300	42851300	1000	0,0676	0,9211	2,38E-24	0,8534
cr. 8	28974100	28974600	500	0,0394	0,8927	2,03E-51	0,8533
cr. 22	22368500	22369000	500	0,0248	0,8778	1,18E-70	0,8529
cr. 14	69825300	69825800	500	0,0654	0,9174	2,73E-14	0,852

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	Placenta a término	Valores de P	Diferencia de metilación
cr. 3	142700300	142700800	500	0,0063	0,8582	2,56E-23	0,8519
cr. 17	59491300	59491800	500	0,0662	0,9175	1,81E-16	0,8513
cr. 15	30881700	30882200	500	0,0493	0,8995	2,38E-26	0,8502
cr. 15	91496300	91496800	500	0	0,85	3,13E-17	0,85
cr. 17	18745300	18745800	500	0,0294	0,8775	3,47E-51	0,848
cr. 15	29436500	29437000	500	0,0336	0,8811	2,62E-66	0,8476
cr. 2	217795300	217795800	500	0	0,8472	1,78E-22	0,8472
cr. 11	16328100	16328600	500	0,0278	0,875	3,43E-11	0,8472
cr. 13	113063500	113064000	500	0,0102	0,8571	1,82E-15	0,8469
cr. 5	40472400	40472900	500	0,0197	0,8664	7,54E-35	0,8467
cr. 1	242549200	242549700	500	0	0,8462	8,53E-23	0,8462
cr. 11	58099100	58099600	500	0,0162	0,8612	4,45E-35	0,845
cr. 9	16020400	16020900	500	0,0132	0,8555	8,05E-23	0,8423
cr. 8	37550700	37551200	500	0,0093	0,8512	1,11E-16	0,8419
cr. 5	75722400	75722900	500	0,1215	0,9627	5,97E-23	0,8411
cr. 19	60454700	60455200	500	0,0316	0,8722	2,44E-62	0,8405
cr. 4	99587100	99587600	500	0,0128	0,8526	1,49E-12	0,8398
cr. 6	25149600	25150600	1000	0,0135	0,8514	0,00E+00	0,8379
cr. 1	32065200	32065700	500	0	0,8371	1,09E-44'	0,8371
cr. 7	5337200	5337700	500	0,0727	0,9098	2,18E-14	0,8371
cr. 17	44269900	44270400	500	0,026	0,8621	3,94E-22	0,8361
cr. 1	36180800	36181300	500	0,0714	0,9067	1,23E-09	0,8352
cr. 18	10472700	10473700	1000	0,0713	0,9064	9,35E-70	0,8351
cr. 5	350000	350500	500	0,0297	0,8643	1,49E-16	0,8346
cr. 2	136481800	136482800	1000	0,0554	0,8887	1,87E-52	0,8332
cr. 4	89241100	89241600	500	0,1091	0,9423	1,05E-12	0,8332
cr. 17	40320700	40321200	500	0,0539	0,8859	6,94E-31	0,832
cr. 7	133897200	133897700	500	0,0769	0,9077	1,64E-24	0,8308
cr. 8	98060600	98061100	500	0,0741	0,9048	3,11E-07	0,8307
cr. 8	134141500	134142000	500	0	0,829	2,77E-58	0,829
cr. 14	80250600	80251100	500-	0,0839	0,9122	2,05E-18	0,8283
cr. 2	100730900	100731400	500	0,0787	0,9067	4,85E-11	0,828
cr. 2	88108100	88108600	500	0,075	0,901	2,10E-16	0,826
cr. 19	16338500	16339500	1000	0,0011	0,8259	0,00E+00	0,8247
cr. 5	141791900	141792900	1000	0,0225	0,8467	0,00E+00	0,8243
cr. 11	116227400	116227900	500	0	0,8242	1,01E-17	0,8242

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	Placenta a término	Valores de P	Diferencia de metilación
cr. 22	48705500	48706000	500	0,0649	0,8891	4,00E-76	0,8242
cr. 9	3518300	3518800	500	0,0263	0,8493	6,66E-25	0,823
cr. 11	16791000	16791500	500	0,1095	0,9322	1,05E-22	0,8228
cr. 3	135746000	135746500	500	0,1429	0,9651	7,63E-10	0,8223
cr. 1	19323400	19323900	500	0,0411	0,8624	2,36E-20	0,8213
cr. 9	96662800	96663300	500	0,0614	0,8826	1,07E-28	0,8212
cr. 7	37348300	37349300	1000	0,0012	0,821	0,00E+00	0,8198
cr. 2	234882000	234882500	500	0,0392	0,8591	5,36E-15	0,8198
cr. 6	44694000	44694500	500	0,1024	0,9222	5,68E-19	0,8198
cr. 17	18320500	18321000	500	0	0,8197	2,78E-39	0,8197
cr. 22	28992000	28994000	2000	0,0012	0,8195	0,00E+00	0,8183
cr. 17	53762700	53766300	3600	0,0003	0,8179	0,00E+00	0,8176
cr. 1	114215500	114216000	500	0	0,8169	2,24E-20	0,8169
cr. 6	13381700	13382700	1000	0,0037	0,8206	1,03E-40	0,8169
cr. 5	17045400	17045900	500	0,0235	0,84	3,24E-13	0,8165
cr. 12	110924300	110924800	500	0,0855	0,9016	1,01E-18	0,816
cr. 1	200499800	200500300	500	0,011	0,8269	9,73E-24	0,8159
cr. 4	8311000	8311500	500	0,053	0,8687	7,82E-18	0,8157
cr. 8	6535300	6535800	500	0,0667	0,8824	5,25E-09	0,8157
cr. 6	42462700	42463200	500	0,0761	0,8919	5,78E-23	0,8157
cr. 1	91969900	91970400	500	0,0172	0,8325	5,23E-18	0,8152
cr. 2	105758600	105759600	1000	0,0305	0,8455	5,15E-52	0,815
cr. 21	37538500	37539000	500	0,1595	0,9745	8,81E-16	0,815
cr. 9	92953000	92953500	500	0,0189	0,8333	1,88E-15	0,8145
cr. 16	30104400	30105900	1500	0,0505	0,8636	0,00E+00	0,8131
cr. 1	234184400	234185400	1000	0,0346	0,8477	9,66E-31	0,813
cr. 8	19116400	19116900	500	0	0,8125	9,33E-11	0,8125
cr. 4	141194300	141195300	1000	0,0865	0,899	5,46E-30	0,8125

Tabla S2D. Lista de las 100 regiones más hipometiladas identificadas a partir de tejido placentario y glóbulos sanguíneos maternos del tercer trimestre.

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	Placenta a término	Valores de P	Diferencia de metilación
cr. 9	40380300	40380800	500	0,9667	0	1,13E-06	0,9667
cr. 1	31769200	31769700	500	0,9548	0,0256	5,57E-25	0,9291
cr. 18	12217600	12218100	500	0,9873	0,0602	1,63E-19	0,9271
cr. 20	19704400	19704900	500	0,9426	0,018	4,34E-18	0,9246

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	Placenta a término	Valores de P	Diferencia de metilación
cr. 15	37317500	37318000	500	0,9358	0,0132	1,90E-25	0,9226
cr. X	83368400	83368900	500	0,913	0	3,15E-07	0,913
cr. 11	27549100	27549600	500	0,9224	0,0123	3,92E-24	0,9101
cr. 18	58141500	58142000	500	0,9737	0,0645	1,07E-09	0,9092
cr. 1	159897000	159897500	500	0,9067	0	2,53E-16	0,9067
cr. 7	84347200	84348700	1500	0,9401	0,0407	0,00E+00	0,8994
cr. 2	216916100	216916600	500	0,9695	0,0714	2,13E-16	0,8981
cr. 7	144200000	144200500	500	0,9294	0,0317	1,24E-10	0,8977
cr. 1	241331600	241332100	500	0,9198	0,0227	0,00E+00	0,8971
cr. 7	123190000	123191000	1000	0,9341	0,0384	0,00E+00	0,8957
cr. 5	12626600	12628600	2000	0,9266	0,0321	0,00E+00	0,8944
cr. 12	12971300	12972300	1000	0,9361	0,0438	0,00E+00	0,8923
cr. 22	20936500	20937000	500	0,9528	0,0606	1,87E-06	0,8922
cr. 13	31321900	31322400	500	0,9231	0,0313	1,43E-06	0,8918
cr. 22	21701500	21702000	500	0,9579	0,0667	1,30E-09	0,8912
cr. 10	104794400	104794900	500	1	0,1111	6,10E-09	0,8889
cr. 7	21835800	21836300	500	0,9156	0,0267	3,85E-13	0,8889
cr. 10	16134800	16135300	500	0,95	0,0635	4,79E-10	0,8865
cr. 3	47410400	47410900	500	0,9213	0,0357	6,63E-17	0,8855
cr. 10	49344400	49345400	1000	0,9422	0,0571	0,00E+00	0,8851
cr. 2	209073900	209074400	500	0,9196	0,0353	1,63E-22	0,8843
cr. 1	89131200	89131700	500	0,9314	0,0472	1,05E-75	0,8842
cr. 3	167118500	167119500	1000	0,9365	0,0527	0,00E+00	0,8838
cr. 18	1395900	1397400	1500	0,9308	0,0472	0,00E+00	0,8836
cr. 2	59670300	59670800	500	0,9433	0,0599	5,09E-23	0,8834
cr. 14	28368900	28369400	500	0,9446	0,0619	5,03E-64	0,8827
cr. 3	126028800	126029300	500	0,9379	0,0556	1,83E-20	0,8823
cr. 9	69378900	69379900	1000	0,8816	0	6,02E-51	0,8816
cr. 5	105517300	105518300	1000	0,927	0,0461	0,00E+00	0,8808
cr. 2	31572400	31573400	1000	0,9176	0,037	0,00E+00	0,8806
cr. 5	42003700	42004700	1000	0,9262	0,0462	0,00E+00	0,88
cr. 14	94718300	94718800	500	0,9548	0,0764	6,67E-19	0,8784
cr. 19	56417800	56418300	500	0,9615	0,0833	1,71E-06	0,8782
cr. 2	70183000	70183500	500	0,9694	0,0914	9,49E-39	0,878
cr. 4	118939300	118939800	500	0,9192	0,0412	2,20E-34	0,878
cr. 13	59246400	59246900	500	0,9402	0,0633	5,40E-12	0,8769

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	Placenta a término	Valores de P	Diferencia de metilación
cr. 12	74021100	74022100	1000	0,9513	0,0752	0,00E+00	0,8761
cr. 2	173432500	173433000	500	0,9529	0,0778	5,39E-12	0,8752
cr. 16	24004400	24004900	500	0,9239	0,0488	3,25E-23	0,8751
cr. 13	27596300	27597300	1000	0,9538	0,0795	0,00E+00	. 0,8743
cr. 15	88904300	88904800	500	0,9212	0,0481	7,69E-27	0,8731
cr. 18	12720200	12721200	1000	0,9346	0,0618	0,00E+00	0,8728
cr. 15	60975900	60976900	1000	0,9311	0,0587	0,00E+00	0,8724
cr. 21	39630100	39631100	1000	0,9423	0,07	4,68E-43	0,8723
cr. 5	123933900	123934400	500	0,943	0,0707	2,60E-38	0,8722
cr. 8	77382600	77383600	1000	0,9117	0,0395	0,00E+00	0,8722
cr. 21	32238800	32239300	500	0,9396	0,0677	1,28E-18	0,8719
cr. 5	175019600	175020100	500	0,9542	0,0828	4,34E-20	0,8714
cr. 8	134437400	134438400	1000	0,9083	0,037	4,79E-29	0,8713
cr. 5	69668800	69669300	500	0,9194	0,0492	2,88E-09	0,8702
cr. 1	60877900	60878900	1000	0,9378	0,068	0,00E+00	0,8698
cr. 16	80650400	80650900	500	0,9309	0,0611	2,49E-32	0,8698
cr. 18	59388800	59389300	500	0,9706	0,1008	4,84E-15	0,8697
cr. 2	15969400	15970400	1000	0,8939	0,0244	3,07E-52.	0,8695
cr. 13	56207100	56208100	1000	0,9184	0,0505	0,00E+00	0,868
cr. 3	180395300	180395800	500	0,9877,	0,12	1,73E-09	0,8677
cr. 6	153238200	153239200	1000	0,9123	0,0452	0,00E+00	0,8671
cr. 18	61635100	61635600	500	0,9268	0,06	2,77E-13	0,8668
cr. 3	177562200	177563200	1000	0,9121	0,0455	0,00E+00	0,8666
cr. 4	160368300	160370200	1900	0,9272	0,0606	0,00E+00	0,8665
cr. 6	144626900	144627400	500	0,9114	0,046	2,10E-12	0,8654
cr. 16	59885500	59886500	1000	0,9407	0,0757	1,12E-62	0,865
cr. 1	55667100	55667600	500	0,9095	0,0446	8,62E-39	0,8649
cr. 2	83598300	83599300	1000	0,9366	0,0718	0,00E+00	0,8648
cr. 4	105135200	105136200	1000	0,913	0,0486	0,00E+00	0,8644
cr. 14	32048400	32048900	500	0,9142	0,0499	5,43E-53	0,8643
cr. 1	223482700	223483700	1000	0,9636	0,0997	2,69E-34	0,864
cr. 14	47487700	47488200	500	0,915	0,0514	5,45E-33	0,8636
cr. 3	104515000	104515500	500	1	0,1373	1,08E-06	0,8627
cr. 7	14008300	14009800	1500	0,9349	0,0725	0,00E+00	0,8624
cr. 1	243134000	243135500	1500	0,9208	0,0588	0,00E+00	0,8619
cr. 10	14156400	14156900	500	0,9105	0,0489	0,00E+00	0,8616
cr. 2	118616200	118617200	1000	0,9178	0,0565	0,00E+00	0,8613

(continuación)

Cromosoma	Inicio	Final	Tamaño (pb)	Glóbulos sanguíneos maternos	Placenta a término	Valores de P	Diferencia de metilación
cr. 17	8455500	8456000	500	0,8941	0,0331	1,94E-18	0,8611
cr. 12	15392200	15393200	1000	0,9307	0,0697	0,00E+00	0,861
cr. 8	81275900	81276900	1000	0,9291	0,0684	0,00E+00	0,8606
cr. 1	234269300	234269800	500	0,9471	0,087	2,25E-25	0,8602
cr. 1	181970300	181970800	500	0,9167	0,0566	2,15E-08	0,8601
cr. 2	55775000	55776000	1000	0,9159	0,0559	0,00E+00	0,8599
cr. 3	88338000	88339000	1000	0,8909	0,0311	0,00E+00	0,8598
cr. 5	140078700	140079200	500	0,8852	0,0253	0,00E+00	0,8598
cr. 21	16720900	16721400	500	0,9317	0,0721	1,38E-15	0,8596
cr. 11	104891100	104892600	1500	0,9164	0,0569	0,00E+00	0,8595
cr. 1	184204700	184205200	500	0,9194	0,0603	8,16E-16	0,859
cr. 6	160732500	160733000	500	0,9191	0,0606	1,31E-10	0,8585
cr. 8	37134300	37134800	500	0,9151	0,0567	1,21E-26	0,8584
cr. 18	5869800	5870300	500	0,913	0,0548	1,21E-09	0,8582
cr. 1	98448100	98448600	500	0,9574	0,1	2,08E-05	0,8574
cr. 3	152897800	152898300	500	0,9402	0,0828	3,28E-18	0,8574
cr. 1	110304000	110304500	500	0,9783	0,121	2,60E-18	0,8572
cr. 2	86993600	86995600	2000	0,8965	0,0395	0,00E+00	0,857
cr. 19	15428100	15430600	2500	0,9424	0,0862	0,00E+00	0,8563
cr. 13	75176800	75177800	1000	0,9258	0,0697	3,09E-47	0,8561
cr. 13	24126700	24127200	500	0,9498	0,0938	1,57E-17	0,856
cr. 16	28238500	28240500	2000	0,9427	0,0868	0,00E+00	0,8559
cr. 2	158079500	158080500	1000	0,9199	0,0642	0,00E+00	0,8557

Tabla S3A. Lista de los 100 locus más hipermetilados deducidos a partir de los datos de secuenciación del plasma materno del primer trimestre con bisulfito.

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 22	39189067	39189863	0	0,8444	0,8444
cr. 17	53763065	53764027	0	0,7922	0,7922
cr. 7	41887694	41888212	0	0,7614	0,7614
cr. 2	1,14E+08	1,14E+08	0	0,751	0,751
cr. 12	25096242	25097206	0	0,7098	0,7098
cr. 1	66574104	66574793	0	0,7025	0,7025
cr. 6	11489985	11490755	0	0,7004	0,7004
cr. 6	1,07E+08	1,07E+08	0	0,6978	0,6978
cr. 10	30858286	30858871	0	0,6693	0,6693

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 17	21131574	21132167	0	0,6496	0,6496
cr. 18	13454740	13455292	0	0,5468	0,5468
cr. 16	11298755	11299326	0	0,5373	0,5373
cr. 2	1,75E+08	1,75E+08	0	0,5196	0,5196
cr. 19	44060511	44061036	0	0,5128	0,5128
cr. 6	1,08E+08	1,08E+08	0	0,5	0,5
cr. 3	71261611	71262501	0	0,4587	0,4587
cr. 9	36247847	36248885	0	0,447	0,447
cr. 19	17819240	17820082	0	0,4279	0,4279
cr. 17	53769900	53770731	0	0,4102	0,4102
cr. 1	1,12E+08	1,12E+08	0,0002	0,6167	0,6166
cr. 7	1,34E+08	1,34E+08	0,0003	0,4351	0,4348
cr. 3	11658550	11659929	0,0004	0,4299	0,4295
cr. 17	53764417	53765963	0,0005	0,7967	0,7961
cr. 10	11246762	11249052	0,0005	0,4002	0,3997
cr. 22	28992647	28993434	0,0006	0,8092	0,8087
cr. 15	62460278	62461007	0,0006	0,4334	0,4328
cr. 1	31002038	31003474	0,0007	0,5926	0,5919
cr. 19	3129246	3132159	0,0008	0,7725	0,7717
cr. 12	1,21E+08	1,21E+08	0,0008	0,7303	0,7295
cr. 19	12304446	12305741	0,0009	0,6986	0,6978
cr. 3	67788734	67789395	0,001	0,9131	0,9121
cr. 9	1,32E+08	1,32E+08	0,001	0,7047	0,7037
cr. 19	6723370	6724479	0,001	0,689	0,688
cr. 3	1,84E+08	1,84E+08	0,001	0,4384	0,4374
cr. 2	53848089	53849214	0,001	0,4368	0,4358
cr. 17	59450886	59452113	0,0012	0,469	0,4678
cr. 5	1,72E+08	1,72E+08	0,0014	0,578	0,5766
cr. 21	35342527	35343373	0,0014	0,5392	0,5378
cr. 21	45164804	45165437	0,0015	0,4251	0,4236
cr. X	3742417	3744601	0,0016	0,4486	0,447
cr. 21	45158293	45159003	0,0017	0,7799	0,7782
cr. 7	39839340	39839876	0,0017	0,4074	0,4057
cr. 2	1,75E+08	1,75E+08	0,0018	0,4816	0,4797
cr. 12	1,24E+08	1,24E+08	0,0019	0,6306	0,6287
cr. 3	50352688	50353823	0,002	0,624	0,622

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 9	97264382	97265523	0,0021	0,5008	0,4987
cr. 7	64178628	64179354	0,0021	0,4088	0,4066
cr. 9	94767202	94767802	0,0023	0,7568	0,7544
cr. 5	42986308	42988304	0,0023	0,4882	0,4859
cr. 17	63854127	63854693	0,0024	0,8266	0,8242
cr. 12	1,22E+08	1,22E+08	0,0024	0,4869	0,4844
cr. 17	16260170	16260909	0,0026	0,6404	0,6378
cr. 4	39874787	39875456	0,0027	0,7233	0,7206
cr. 12	6441080	6441608	0,0027	0,6228	0,6201
cr. 19	45015653	45016886	0,0028	0,5444	0,5416
cr. 6	30757752	30758823	0,0028	0,4783	0,4755
cr. 6	41636176	41637112	0,0028	0,4254	0,4226
cr. 12	6315199	6315765	0,0029	0,4613	0,4584
cr. 14	76576283	76577070	0,0029	0,4365	0,4336
cr. 16	48857790	48858300	0,0031	0,5625	0,5594
cr. 5	1,7E+08	1,7E+08	0,0031	0,4752	0,4721
cr. 13	26897813	26898557	0,0032	0,4354	0,4322
cr. 14	52753948	52754571	0,0032	0,4221	0,4189
cr. 1	1,66E+08	1,66E+08	0,0033	0,5579	0,5545
cr. 12	56157424	56158348	0,0033	0,47	0,4667
cr. 22	16079971	16080532	0,0034	0,6226	0,6193
cr. 7	1946410	1946975	0,0036	0,6826	0,6789
cr. 11	258799	259749	0,0036	0,5072	0,5037
cr. 6	13381944	13382477	0,0037	0,5945	0,5908
cr. 7	1,27E+08	1,27E+08	0,0037	0,5096	0,5058
cr. 13	23743886	23744467	0,0037	0,4534	0,4497
cr. 2	1,21E+08	1,21E+08	0,0038	0,7175	0,7137
cr. 21	25855853	25857105	0,0039	0,4661	0,4622
cr. 2	43211724	43212565	0,0039	0,4345	0,4306
cr. 12	1,08E+08	1,08E+08	0,0041	0,6024	0,5983
cr. 15	92928924	92929575	0,0041	0,4074	0,4033
cr. 19	10731043	10731636	0,0042	0,5868	0,5826
cr. 6	1,45E+08	-1,45E+08	0,0043	0,5783	0,574
cr. 1	52875323	52875907	0,0044	0,4145	0,4101
cr. 14	75058186	75058956	0,0045	0,602	0,5975
cr. 12	1,21E+08	1,21E+08	0,0045	0,4821	0,4776
cr. 17	76873737	76874417	0,0046	0,601-2	0,5966

(continuación)

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 2	2,38E+08	2,38E+08	0,0049	0,7654	0,7604
cr. 2	1,98E+08	1,98E+08	0,0049	0,7228	0,7179
cr. 6	1,47E+08	1,47E+08	0,0049	0,4967	0,4918
cr. 9	1,36E+08	1,36E+08	0,0049	0,4584	0,4535
cr. 1	67545402	67546771	0,005	0,4971	0,4921
cr. 6	1,58E+08	1,58E+08	0,0052	0,6145	0,6093
cr. 3	1,7E+08	1,7E+08	0,0052	0,5845	0,5794
cr. 1	2,34E+08	2,34E+08	0,0053	0,7033	0,6979
cr. 10	80715722	80716751	0,0053	0,6515	0,6462
cr. 4	48602901	48603736	0,0053	0,6315	0,6262
cr. 19	13957965	13958580	0,0053	0,599	0,5937
cr. 1	90081114	90082367	0,0053	0,4574	0,4521
cr. 2	1,06E+08	1,06E+08	0,0054	0,8858	0,8804
cr. 16	29664213	29665369	0,0054	0,8339	0,8285
cr. 1	1,59E+08	1,59E+08	0,0054	0,7663	0,7608
cr. 13	97926489	97927025	0,0054	0,6229	0,6175
cr. 1	41604452	41605277	0,0054	0,6011	0,5956
cr. 9	1,28E+08	1,28E+08	0,0054	0,5871	0,5818

Tabla S3B. Lista de los 100 locus más hipometilados deducidos a partir de los datos de secuenciación del plasma materno del primer trimestre con bisulfito

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 1	235771917	235772426	0,9868	0,549	0,4377
cr. 1	97357972	97358622	0,9835	0,4805	0,503
cr. 1	4490516	4491074	0,9826	0,4793	0,5032
cr. 4	181124168	181124671	0,9825	0,4725	0,5099
cr. 16	71908694	71909213	0,982	0,5581	0,4239
cr. 3	182727915	182728477	0,981	0,3577	0,6233
cr. 5	115339535	115340038	0,9802	0,5455	0,4347
cr. 3	195855575	195856122	0,9801	0,3793	0,6008
cr. 6	155437621	155438161	0,9799	0,5991	0,3808
cr. 9	20468093	20468904	0,9798	0,4271	0,5527
cr. 10	90702298	90702987	0,9787	0,3324	0,6463
cr. 1	170581654	170582162	0,9785	0,4817	0,4968
cr. 3	108816849	108817794	0,9783	0,4793	0,4989
cr. 20	36912749	36913319	0,9783	0,5	0,4783
cr. 13	72517281	72517839	0,9782	0,4855	0,4927

(continuación)

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 12	103553001	103553677	0,9774	0,492	0,4854
cr. 22	27638905	27639408	0,9766	0,5385	0,4382
cr. 7	17290850	17291462	0,9763	0,59	0,3863
cr. 6	17227866	17228510	0,976	0,4058	0,5703
cr. 15	56998547	56999107	0,9754	0,3766	0,5988
cr. 7	70965945	70966842	0,9753	0,5893	0,386
cr. 3	32159338	32160065	0,9752	0,5379	0,4372
cr. 16	17043258	17043854	0,9752	0,5521	0,4231
cr. 16	22776223	22776850	0,9752	0,5735	0,4017
cr. 5	169344029	169344869	0,9751	0,4211	0,5541
cr. 11	34324955	34325722	0,975	0,5561	0,4189
cr. 8	58554745	58555376	0,9747	0,5784	0,3964
cr. 1	153933389	153934121	0,9746	0,463	0,5116
cr. 14	88003983	88004485	0,9745	0,5379	0,4366
cr. 3	151738501	151739120	0,9741	0,4901	0,484
cr. 14	105618699	105619606	0,974	0,3457	0,6283
cr. 16	24060085	24060702	0,9738	0,3991	0,5747
cr. 8	68941792	68942711	0,9738	0,5449	0,429
cr. 12	53208707	53209304	0,9737	0,4847	0,489
cr. 7	76892564	76893249	0,9736	0,5664	0,4072
cr. 3	69464294	69464971	0,9736	0,5893	0,3843
cr. 19	61401137	61401745	0,9732	0,4933	0,4799
cr. 11	124569867	124570490	0,9732	0,5136	0,4595
cr. 18	42618440	42619096	0,9732	0,5942	0,379
cr. 5	169398896	169399637	0,9731	0,498	0,4751
cr. 5	169328124	169328983	0,9731	0,572	0,401
cr. 20	34679880	34680448	0,9731	0,5922	0,3809
cr. 16	9042198	9042702	0,973	0,4286	0,5444
cr. 10	90205044	90205701	0,973	0,4407	0,5323
cr. 13	33236454	33236997	0,973	0,5906	0,3824
cr. 16	73284579	73285087	0,9729	0,5602	0,4127
cr. 8	29100691	29101428	0,9728	0,505	0,4678
cr. 2	202383851	202384447	0,9727	0,5461	0,4267
cr. 3	179501620	179502300	0,9722	0,5766	0,3956
cr. 6	107674976	107675906	0,9719	0,4434	0,5285
cr. 6	107880632	107881161	0,9718	0,5623	0,4095
cr. 12	56350283	56350933	0,9718	0,5909	0,3809
cr. 19	40636458	40637339	0,9717	0,4941	0,4776

ES 2 994 426 T3

(continuación)

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 2	223472599	223473287	0,9714	0,1824	0,7891
cr. 22	20709067	20709787	0,9714	0,5149	0,4565
cr. 19	46095583	46096190	0,9713	0,5385	0,4328
cr. 6	90258338	90259318	0,9712	0,3415	0,6297
cr. 2	54598347	54598933	0,9712	0,5894	0,3819
cr. 3	114810453	114811493	0,9711	0,5166	0,4545
cr. 19	15851125	15851654	0,9711	0,5236	0,4476
cr. 8	42889138	42890084	0,9711	0,5652	0,4059
cr. 18	52354390	52355064	0,971	0,598	0,373
cr. 15	38206236	38207010	0,9709	0,4186	0,5523
cr. 7	99700554	99701110	0,9708	0,305	0,6658
cr. 12	19487336	19487855	0,9708	0,4105	0,5603
cr. 7	87996908	87997437	0,9708	0,5462	0,4246
cr. 6	63628653	63629378	0,9707	0,529	0,4417
cr. 15	38209108	38209618	0,9706	0,5882	0,3824
cr. 19	6623769	6624450	0,9704	0,5179	0,4526
cr. 2	10794513	10795242	0,9704	0,5976	0,3728
cr. 2	118472785	118474454	0,9704	0,5992	0,3712
cr. 5	57820209	57820801	0,9701	0,5815	0,3886
cr. 10	100183380	100184702	0,9701	0,5826	0,3875
cr. 2	8151989	8152646	0,97	0,4701	0,4999
cr. 10	3938374	3938914	0,9699	0,1741	0,7958
cr. 9	123724524	123725439	0,9697	0,57	0,3997
cr. 14	89085469	89086097	0,9696	0,3278	0,6418
cr. 16	14129437	14130133	0,9695	0,5304	0,4392
cr. 5	60746367	60747191	0,9695	0,5571	0,4124
cr. 1	92002953	92003729	0,9694	0,52	0,4494
cr. 6	31264677	31265413	0,9693	0,5135	0,4558
cr. 7	99317013	99318281	0,9692	0,5117	0,4574
cr. 8	8808867	8809422	0,9692	0,5691	0,4002
cr. 19	20052165	20052720	0,969	0,2792	0,6898
cr. 8	129139026	129139573	0,969	0,3458	0,6232
cr. 11	122314929	122315458	0,969	0,4232	0,5458
cr. 13	98377663	98378165	0,9688	0,3319	0,6369
cr. 9	107606194	107606872	0,9688	0,449	0,5198
cr. 8	56096904	56097736	0,9688	0,5267	0,4422
cr. 7	128093836	128094339	0,9688	0,5929	0,3758
cr. 2	103109370	103109916	0,9686	0,3333	0,6352

(continuación)

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	MVC	Diferencia de metilación
cr. 3	101803534	101804063	0,9686	0,5027	0,4659
cr. 10	69505720	69506278	0,9684	0,2515	0,7169
cr. 13	26608225	26608754	0,9683	0,3614	0,6069
cr. 1	90993315	90993828	0,9683	0,5519	0,4164
cr. 6	11361243	11361801	0,9681	0,2578	0,7103
cr. 21	36529300	36529981	0,968	0,1944	0,7736
cr. 21	37813953	37814521	0,9679	0,2175	0,7505
cr. 2	15226273	15227211	0,9679	0,5134	0,4545
cr. 19	4102809	4103443	0,9679	0,5646	0,4034

Tabla S3C. Lista de los 100 locus más hipometilados deducidos a partir de los datos de secuenciación del plasma materno del tercer trimestre con bisulfito

Cromosoma	Inicio	Final	Glóbulos sanguíneos maternos	Placenta a término	Diferencia de metilación
cr. 17	53763065	53764027	0,0000	0,8680	0,8680
cr. 22	39189067	39189863	0,0000	0,8233	0,8233
cr. 10	30858286	30858871	0,0000	0,7713	0,7713
cr. 7	41887694	41888212	0,0000	0,7578	0,7578
cr. 2	1,14E+08	1,14E+08	0,0000	0,7500	0,7500
cr. 12	25096242	25097206	0,0000	0,7332	0,7332
cr. 6	1,07E+08	1,07E+08	0,0000	0,7229	0,7229
cr. 1	66574104	66574793	0,0000	0,7136	0,7136
cr. 16	11298755	11299326	0,0000	0,7005	0,7005
cr. 6	11489985	11490755	0,0000	0,6935	0,6935
cr. 18	13454740	13455292	0,0000	0,6594	0,6594
cr. 6	1,08E+08	1,08E+08	0,0000	0,6231	0,6231
cr. X	3627885	3628549	0,0000	0,6133	0,6133
cr. 12	7979754	7980413	0,0000	0,6118	0,6118
cr. 3	71261611	71262501	0,0000	0,5938	0,5938
cr. 17	53769900	53770731	0,0000	0,5586	0,5586
cr. 11	1,18E+08	1,18E+08	0,0000	0,5558	0,5558
cr. 19	44060511	44061036	0,0000	0,5464	0,5464
cr. 2	2,38E+08	2,38E+08	0,0000	0,5330	0,5330
cr. 1	1,91E+08	1,91E+08	0,0000	0,5294	0,5294
cr. 1	1,44E+08	1,44E+08	0,0000	0,4857	0,4857
cr. 2	1,75E+08	1,75E+08	0,0000	0,4785	0,4785
cr. 4	15366889	15367646	0,0000	0,4729	0,4729
cr. 2	19537237	19537737	0,0000	0,4599	0,4599
cr. 1	1,15E+08	1,15E+08	0,0000	0,4351	0,4351

ES 2 994 426 T3

Cromosoma	Inicio	Final	(continuación)	Placenta a término	Diferencia de metilación
			Glóbulos sanguíneos maternos		
cr. 1	1,54E+08	1,54E+08	0,0000	0,4299	0,4299
cr. 14	51383387	51384149	0,0000	0,4186	0,4186
cr. 1	1,12E+08	1,12E+08	0,0002	0,5350	0,5348
cr. 3	11658550	11659929	0,0004	0,5579	0,5575
cr. 17	53764417	53765963	0,0005	0,7894	0,7889
cr. 22	28992647	28993434	0,0006	0,8053	0,8047
cr. 6	27214981	27215823	0,0006	0,4593	0,4587
cr. 1	31002038	31003474	0,0007	0,6309	0,6302
cr. 12	1,12E+08	1,12E+08	0,0008	0,7360	0,7352
cr. 19	3129246	3132159	0,0008	0,7257	0,7249
cr. 19	12304446	12305741	0,0009	0,6397	0,6388
cr. 6	28723918	28724965	0,0009	0,4344	0,4335
cr. 19	6723370	6724479	0,0010	0,7280	0,7270
cr. 2	53848089	53849214	0,0010	0,4060	0,4050
cr. 9	1,32E+08	1,32E+08	0,0010	0,7558	0,7548
cr. 3	67788734	67789395	0,0010	0,9219	0,9209
cr. 19	18276368	18277132	0,0011	0,5136	0,5125
cr. 17	59450886	59452113	0,0012	0,5196	0,5184
cr. 17	74243852	74244670	0,0012	0,4117	0,4105
cr. 3	1,85E+08	1,85E+08	0,0014	0,4961	0,4948
cr. 21	35342527	35343373	0,0014	0,5126	0,5112
cr. 5	1,72E+08	1,72E+08	0,0014	0,6531	0,6516
cr. 21	45164804	45165437	0,0015	0,4364	0,4349
cr. X	3742417	3744601	0,0016	0,7517	0,7501
cr. 21	45158293	45159003	0,0017	0,8180	0,8163
cr. 2	1,75E+08	1,75E+08	0,0018	0,6214	0,6196
cr. 12	1,24E+08	1,24E+08	0,0019	0,5906	0,5887
cr. 3	50352688	50353823	0,0020	0,6082	0,6062
cr. 9	94767202	94767802	0,0023	0,8327	0,8304
cr. 17	63854127	63854693	0,0024	0,7886	0,7862
cr. 12	1,22E+08	1,22E+08	0,0024	0,5021	0,4997
cr. 17	16260170	16260909	0,0026	0,5780	0,5754
cr. 12	6441080	6441608	0,0027	0,7471	0,7444
cr. 4	39874787	39875456	0,0027	0,7962	0,7935
cr. 18	50536032	50536649	0,0027	0,4920	0,4893
cr. 6	30757752	30758823	0,0028	0,4029	0,4001
cr. 19	2571763	2572292	0,0031	0,4200	0,4169
cr. 5	1,7E+08	1,7E+08	0,0031	0,4218	0,4187

ES 2 994 426 T3

Cromosoma	Inicio	Final	(continuación)	Placenta a término	Diferencia de metilación
			Glóbulos sanguíneos maternos		
cr. 13	26897813	26898557	0,0032	0,6485	0,6453
cr. 12	56157424	56158348	0,0033	0,5541	0,5508
cr. 1	1,66E+08	1,66E+08	0,0033	0,5147	0,5113
cr. 22	16079971	16080532	0,0034	0,6265	0,6231
cr. 6	16820551	16821134	0,0035	0,4800	0,4765
cr. 11	258799	259749	0,0036	0,5475	0,5439
cr. 7	1946410	1946975	0,0036	0,8251	0,8215
cr. 6	13381944	13382477	0,0037	0,8221	0,8183
cr. 7	1,27E+08	1,27E+08	0,0037	0,4767	0,4730
cr. 2	1,21E+08	1,21E+08	0,0038	0,6734	0,6697
cr. 2	43211724	43212565	0,0039	0,4256	0,4217
cr. 15	92928924	92929575	0,0041	0,5605	0,5564
cr. 12	1,08E+08	1,08E+08	0,0041	0,7313	07271
cr. 19	10731043	10731636	0,0042	0,5668	0,5626
cr. 6	1,45E+08	1,45E+08	0,0043	0,5910	0,5867
cr. 1	52875323	52875907	0,0044	0,6115	0,6071
cr. 12	1,21E+08	1,21E+08	0,0045	0,5884	0,5839
cr. 14	75058186	75058956	0,0045	0,6534	0,6489
cr. 17	76873737	76874417	0,0046	0,5658	0,5612
cr. 6	1,47E+08	1,47E+08	0,0049	0,5826	0,5777
cr. 2	1,98E+08	1,98E+08	0,0049	0,7944	0,7895
cr. 2	2,38E+08	2,38E+08	0,0049	0,7328	0,7278
cr. 8	1,42E+08	1,42E+08	0,0050	0,7728	0,7679
cr. 3	1,7E+08	1,7E+08	0,0052	0,7227	0,7176
cr. 6	1,58E+08	1,58E+08	0,0052	0,6389	0,6337
cr. 2	2,38E+08	2,38E+08	0,0052	0,4238	0,4185
cr. 11	56948854	56949496	0,0053	0,4484	0,4431
cr. 4	48602901	48603736	0,0053	0,5920	0,5867
cr. 5	1,31E+08	1,31E+08	0,0053	0,4858	0,4805
cr. 10	80715722	80716751	0,0053	0,5249	0,5196
cr. 19	13957965	13958580	0,0053	0,4379	0,4326
cr. 1	2,34E+08	2,34E+08	0,0053	0,8440	0,8387
cr. 13	97926489	97927025	0,0054	0,7233	0,7179
cr. 9	1,28E+08	1,28E+08	0,0054	0,7312	0,7258
cr. 2	1,06E+08	1,06E+08	0,0054	0,8513	0,8459
cr. 2	96556705	96557637	0,0054	0,4095	0,4041
cr. 16	29664213	29665369	0,0054	0,8837	0,8783

Cromosoma	Tabla S3D. Lista de los 100 locus más hipometilados deducidos a partir de los Glóbulos sanguíneos maternos					Diferencia de metilación
	Inicio	Final		Placenta a término		
cr. 10	7548948	7549483	0,9866	0,5685	0,4181	
cr. 1	4490516	4491074	0,9826	0,5015	0,4810	
cr. 4	1,81E+08	1,81E+08	0,9825	0,5981	0,3843	
cr. 3	1,83E+08	1,83E+08	0,9810	0,2925	0,6886	
cr. 3	1,96E+08	1,96E+08	0,9801	0,4643	0,5158	
cr. 6	1,55E+08	1,55E+08	0,9799	0,4610	0,5189	
cr. 1	1,71E+08	1,71E+08	0,9785	0,5122	0,4662	
cr. 20	36912749	36913319	0,9783	0,4513	0,5269	
cr. 22	38583100	38583616	0,9783	0,5428	0,4355	
cr. 1	19391314	19392207	0,9778	0,5273	0,4505	
cr. 5	1,74E+08	1,74E+08	0,9770	0,5852	0,3918	
cr. 19	13678906	13679531	0,9760	0,5812	0,3949	
cr. 14	83650790	83651395	0,9760	0,5378	0,4382	
cr. 15	56998547	56999107	0,9754	0,4691	0,5063	
cr. 16	22776223	22776850	0,9752	0,5114	0,4638	
cr. 5	1,69E+08	1,69E+08	0,9751	0,4809	0,4943	
cr. 8	58554745	58555376	0,9747	0,5977	0,3770	
cr. 14	1,06E+08	1,06E+08	0,9740	0,2069	0,7671	
cr. 8	68941792	68942711	0,9738	0,5872	0,3866	
cr. 16	24060085	24060702	0,9738	0,3470	0,6268	
cr. 12	53208707	53209304	0,9737	0,5278	0,4459	
cr. 5	1,69E+08	1,69E+08	0,9731	0,5057	0,4673	
cr. 16	9042198	9042702	0,9730	0,1860	0,7869	
cr. 10	90205044	90205701	0,9730	0,5922	0,3808	
cr. 3	1,89E+08	1,89E+08	0,9720	0,4949	0,4771	
cr. 6	1,08E+08	1,08E+08	0,9719	0,5825	0,3894	
cr. 2	2,23E+08	2,23E+08	0,9714	0,3333	0,6381	
cr. 19	46095583	46096190	0,9713	0,5065	0,4648	
cr. 8	1,41E+08	1,41E+08	0,9713	0,5753	0,3959	
cr. 6	90258338	90259318	0,9712	0,4357	0,5355	
cr. 13	51403556	51404069	0,9710	0,3980	0,5731	
cr. 18	66875048	66875726	0,9710	0,5259	0,4451	
cr. 7	99700554	99701110	0,9708	0,3757	0,5951	
cr. 7	87996908	87997437	0,9708	0,5720	0,3988	
cr. 19	6623769	6624450	0,9704	0,4774	0,4930	
cr. 1	97639047	97639749	0,9701	0,4148	0,5553	
cr. 16	23892096	23892772	0,9701	0,5000	0,4701	

Cromosoma	Inicio	Final	(continuación)	Placenta a término	Diferencia de metilación
			Glóbulos sanguíneos maternos		
cr. 10	3938374	3938914	0,9699	0,1148	0,8551
cr. 14	89085469	89086097	0,9696	0,2964	0,6732
cr. 8	1,29E+08	1,29E+08	0,9690	0,3565	0,6126
cr. 13	98377663	98378165	0,9688	0,3123	0,6566
cr. 8	56096904	56097736	0,9688	0,4562	0,5126
cr. 2	1,03E+08	1,03E+08	0,9686	0,3459	0,6227
cr. 13	26608225	26608754	0,9683	0,4562	0,5121
cr. 2	22738157	22738760	0,9682	0,5122	0,4560
cr. 6	11361243	11361801	0,9681	0,2646	0,7035
cr. 21	36529300	36529981	0,9680	0,1829	0,7852
cr. 21	37813953	37814521	0,9679	0,3061	0,6619
cr. 2	2,43E+08	2,43E+08	0,9679	0,5750	0,3929
cr. 4	12413543	12414103	0,9679	0,5944	0,3735
cr. 3	1,27E+08	1,27E+08	0,9677	0,4030	0,5648
cr. 7	33509047	33509556	0,9676	0,4627	0,5048
cr. 14	59284846	59285553	0,9674	0,5254	0,4420
cr. 17	42623453	42624024	0,9673	0,4318	0,5355
cr. 19	6778363	6779377	0,9671	0,4416	0,5255
cr. 4	41798250	41798788	0,9670	0,5000	0,4670
cr. 5	88054080	88054588	0,9669	0,2238	0,7431
cr. 16	24109379	24110289	0,9669	0,5062	0,4607
cr. 10	13847159	13847895	0,9667	0,3188	0,6479
cr. 10	1,27E+08	1,27E+08	0,9667	0,5423	0,4244
cr. 12	1,12E+08	1,12E+08	0,9663	0,3722	0,5941
cr. 10	17220886	17221845	0,9662	0,4455	0,5207
cr. 8	5947355	5947862	0,9662	0,5171	0,4491
cr. 3	73740840	73741439	0,9659	0,3657	0,6002
cr. 14	57945953	57946875	0,9658	0,5357	0,4301
cr. 14	50905777	50906333	0,9658	0,3008	0,6650
cr. 15	90275374	90276000	0,9657	0,5409	0,4248
cr. 22	24717299	24718197	0,9657	0,5160	0,4497
cr. 7	36530128	36530987	0,9656	0,5194	0,4462
cr. 2	1,31E+08	1,31E+08	0,9655	0,4384	0,5271
cr. 4	42116988	42117788	0,9654	0,5195	0,4459
cr. 12	1,16E+08	1,16E+08	0,9653	0,5594	0,4059
cr. 2	7491785	7492736	0,9652	0,4556	0,5097
cr. 19	6599638,	6600187	0,9652	0,5488	0,4163
cr. 6	25326803	25327398	0,9651	0,3974	0,5677

ES 2 994 426 T3

Cromosoma	Inicio	Final	(continuación) Glóbulos sanguíneos maternos	Placenta a término	Diferencia de metilación
cr. 4	1,7E+08	1,7E+08	0,9651	0,4933	0,4718
cr. 7	99875338	99876155	0,9650	0,2696	0,6953
cr. 14	97144328	97145208	0,9649	0,5377	0,4272
cr. 3	11718596	11719163	0,9649	0,5521	0,4128
cr. 14	1E+08	1E+08	0,9649	0,3794	0,5855
cr. 7	1,5E+08	1,5E+08	0,9648	0,3327	0,6322
cr. 12	56357827	56358328	0,9648	0,4217	0,5430
cr. 10	8275750	8276276	0,9647	0,3100	0,6547
cr. 11	16999685	17000209	0,9647	0,2765	0,6882
cr. 22	34419356	34419861	0,9646	0,4245	0,5401
cr. 18	72453151	72453725	0,9646	0,4700	0,4946
cr. 5	49919879	49920699	0,9645	0,3169	0,6476
cr. 1	24580891	24581805	0,9643	0,3565	0,6078
cr. 22	18233774	18234492	0,9641	0,5205	0,4436
cr. 14	45356178	45356903	0,9640	0,3934	0,5706
cr. 3	53007193	53008661	0,9638	0,4902	0,4737
cr. 4	55027912	55028539	0,9637	0,5254	0,4384
cr. 5	1,37E+08	1,37E+08	0,9637	0,5290	0,4347
cr. 1	2,23E+08	2,23E+08	0,9636	0,0997	0,8640
cr. 7	1,35E+08	1,35E+08	0,9636	0,2959	0,6677
cr. 5	80350438	80351169	0,9636	0,4969	0,4667
cr. 12	31889600	31890343	0,9636	0,1745	0,7891
cr. 12	8365395	8366096	0,9636	0,5721	0,3914
cr. 19	15424819	15425355	0,9635	0,2836	0,6799
cr. 10	10985469	10986409	0,9635	0,4877	0,4759

REIVINDICACIONES

1. Un método de análisis de una muestra biológica de un organismo, comprendiendo el método:
identificar una pluralidad de regiones de un genoma del organismo, incluyendo cada una de las regiones uno o varios sitios;
5 analizar una pluralidad de moléculas de ADN de la muestra biológica del organismo que comprende moléculas de ácido nucleico, en donde al menos algunas de las moléculas de ácido nucleico son extracelulares en la muestra biológica, y en donde el análisis de una molécula de ADN incluye:
determinar una ubicación de la molécula de ADN en el genoma del organismo; y
10 determinar si la molécula de ADN está metilada en uno o más sitios;
para cada una de la pluralidad de regiones:
determinar un número respectivo de moléculas de ADN que están metiladas en sitios de la región; y
calcular un nivel de metilación de la región respectiva basándose en el número respectivo de moléculas de ADN
que están metiladas en sitios de la región;
15 comparar cada uno de los niveles de metilación de las regiones respectivas con un nivel de metilación de referencia correspondiente de un grupo de referencia;
determinar una cantidad de regiones que tienen un nivel de metilación de la región superior al nivel de metilación de referencia correspondiente;
comparar la cantidad de regiones con un valor umbral; y
20 determinar una clasificación de la presencia de cáncer basándose en la comparación.
2. El método de la reivindicación 1, en donde las regiones son islas CpG y la identificación de la pluralidad de regiones comprende identificar regiones que tengan un nivel medio de metilación superior o inferior a un primer porcentaje en el grupo de referencia.
25
3. El método de la reivindicación 2, en donde la identificación de la pluralidad de regiones comprende además identificar regiones que tengan un coeficiente de variación del nivel de metilación en el grupo de referencia inferior a un segundo porcentaje.
30
4. El método de la reivindicación 2, en donde la determinación de la cantidad de regiones incluye calcular, por medio de un ordenador, una puntuación acumulada de los respectivos niveles de metilación de las regiones en las que se ha determinado que el nivel de metilación de la región respectiva supera el nivel de metilación de referencia correspondiente.
35
5. El método de la reivindicación 4, en donde la comparación de la cantidad de regiones con el valor umbral incluye comparar la puntuación acumulada con un valor de corte acumulado para determinar la clasificación.
6. El método de la reivindicación 4 o la reivindicación 5, en donde el cálculo de la puntuación acumulada a partir de los niveles de metilación de las regiones respectivas incluye:
40 para cada una de las regiones en las que se determine que el nivel de metilación de la región respectiva supera el nivel de metilación de referencia correspondiente:
calcular una diferencia respectiva entre el nivel de metilación de la región y el nivel de metilación de referencia correspondiente; y
calcular una probabilidad respectiva correspondiente a la diferencia respectiva; y
45 determinar la puntuación acumulada utilizando las probabilidades respectivas.
7. El método de la reivindicación 1, en donde las regiones son islas CpG y la determinación de la cantidad de regiones incluye calcular, por medio de un ordenador, una puntuación acumulada de los respectivos niveles de metilación de las regiones en las que se ha determinado que el nivel de metilación de la región respectiva supera el nivel de metilación de referencia correspondiente.
50
8. El método de la reivindicación 7, en donde la comparación de la cantidad de regiones con el valor umbral incluye comparar la puntuación acumulada con un valor de corte acumulado para determinar la clasificación.
9. El método de la reivindicación 7 o la reivindicación 8, en donde el cálculo de la puntuación acumulada a partir de los niveles de metilación de las regiones respectivas incluye:
55 para cada una de las regiones en las que se determine que el nivel de metilación de la región respectiva supera el nivel de metilación de referencia correspondiente:
calcular una diferencia respectiva entre el nivel de metilación de la región y el nivel de metilación de referencia correspondiente; y
calcular una probabilidad respectiva correspondiente a la diferencia respectiva; y
determinar la puntuación acumulada utilizando las probabilidades respectivas.
60
10. El método de la reivindicación 1, en donde las regiones son islas CpG.
65
11. Un programa informático que comprende instrucciones que, cuando son ejecutadas por un ordenador, hacen que

el ordenador realice el método de cualquiera de las reivindicaciones 1 a 10.

12. Un medio legible por ordenador que almacena el programa informático de la reivindicación 11.

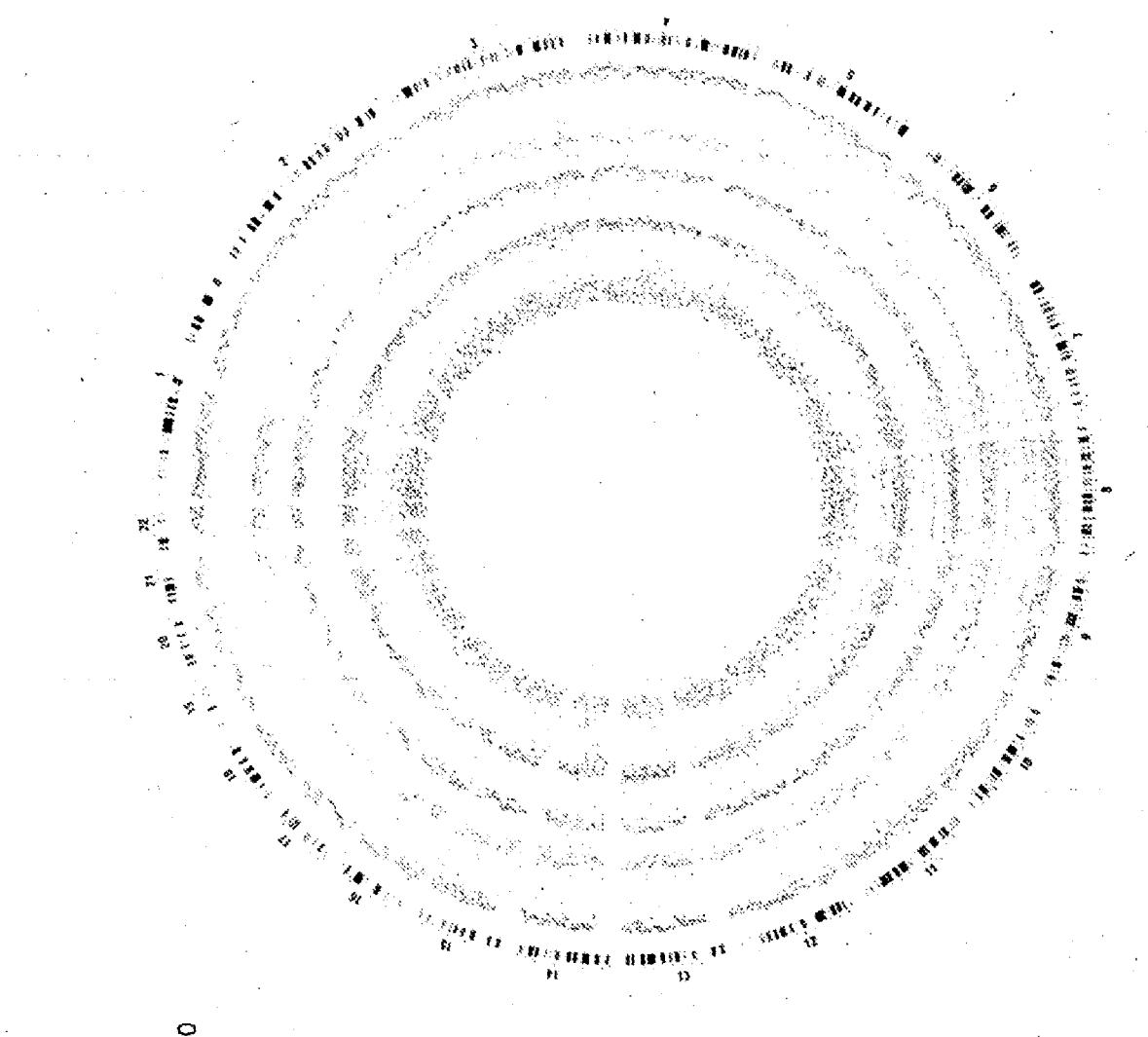
100

	Sangre materna (1er trimestre)		MVC		Placenta a término		Placenta materno (3er trimestre)		Placenta materno (tras el parto)	
Número de lecturas	watson	crick	watson	crick	watson	crick	watson	crick	watson	crick
Congruibilidad total (%)	547522712	422873853	422873463	637357170	637357170	4567228854	1289337674	1277615253	1277615253	1277615253
Lecturas congruentes totales	45,08	44,39	44,12	43,83	43,70	43,40	25,82	25,81	31,85	37,40
Tasa de ambigüedad (%) ^a	246508452	243028086	186956998	183460726	297758773	283850215	1361958304	1346252524	461639066	463867803
Tasa de duplicación (%) ^b	6,72	5,83	5,64	5,89	5,25	5,54	6,36	6,43	5,93	6,05
Tasa de duplicación (%) ^b	6,80	5,86	5,52	9,27	32,94	30,39	64,86	64,41	25,70	4,01
Profundidad del genoma completo tras la eliminación de la tasa de ambigüedad y de conversión falso(a) (%) ^c	11,29	11,19	8,38	8,31	9,43	9,25	25,40	25,26	16,95	14,88
Número de M de la sec C genómica ^d	95,98	99,98	91,93	95,94	99,95	99,96	99,96	99,97	99,96	99,96
Número de N de la sec C genómica ^d	268495905	265563985	199033545	188456370	226318386	223718285	460694054	635295115	453569856	394672185
Cobertura de C genómica (%)	4384534136	4449849557	423238985	4201566735	4983445023	4908658234	1,16365E+10	1,1641E+10	855105686	39515396
Densidad de metilación de C genómica (%)	66,45	66,39	62,86	62,56	67,49	67,49	80,48	80,48	69,76	67,88
Profundidad de la sec C genómica	8,14	8,07	7,57	7,51	8,89	8,78	21,12	21,01	15,39	12,43
Densidad de metilación de C genómica (%)	5,65	5,63	4,31	4,29	4,36	4,36	5,19	5,17	5,05	5,04
Número de M de la sec de CpG	265033983	262504347	183242316	181245307	213757505	211325313	627866879	622753476	440725023	388810638
Número de N de la sec de CpG	104463657	10393611	1452523817	1454598340	147288261	146811702	310000575	308133594	205477042	1422711
Cobertura de CpG (%)	85,17	85,12	81,92	80,92	84,23	84,15	92,05	92,02	86,19	84,93
Profundidad de la sec de CpG	13,12	13,01	11,68	11,71	12,82	12,71	33,30	33,06	22,94	16,91
Densidad de metilación sin CpG (%)	71,80	71,63	55,72	54,97	59,21	58,04	65,95	65,90	68,23	73,14
Número de M de la sec sin CpG	33759043	3145047	7236715	7162226	12502445	12301418	128118574	125332928	12971639	12843542
Número de N de la sec sin CpG	4302383851	4296948441	3375838287	3971182968	4560807406	4537334784	1,1358E+10	1,1368E+10	8367572669	7859475
Cobertura de no CpG (%)	0,13	0,13	0,12	0,12	0,12	0,13	0,16	0,16	0,14	0,13
Profundidad de la sec sin CpG	1,52	1,52	1,41	1,41	1,62	1,61	4,02	4,00	2,95	2,39
Densidad de metilación sin CpG (%)	0,08	0,07	0,68	0,18	0,27	0,27	0,11	0,11	0,15	0,12
Número de M de la sec de CHG	1029734	1030287	2160577	2071016	32233342	3176751	3765562	3891198	3626562	3542617
Número de N de la sec de CHG	1345439602	1341393696	1217932855	1216509371	136263765	1357569840	3416559203	34042259868	2519517702	2347392
Cobertura de CHG (%)	77,90	77,82	73,36	73,13	77,56	77,38	83,16	86,12	79,68	77,75
Profundidad de la sec de CHG	10,98	10,85	9,86	9,85	11,04	11,00	27,66	27,55	20,40	16,48
Densidad de metilación de CHG (%)	0,08	0,07	0,17	0,17	0,24	0,24	0,11	0,11	0,14	0,12
Número de M de la sec sin CHH	224513049	2148760	5185738	5091111	5278803	91246867	9052012	8841730	93536307	92513256
Número de N de la sec sin CHH	2956545249	2943709345	2761813432	27565733597	3198124541	3180265944	7393543368	7303564596	58380454967	582102651
Cobertura de CHH (%)	61,96	61,90	58,43	59,32	63,54	63,45	77,53	77,52	65,96	65,89
Profundidad de metilación de CHH (%)	6,84	6,81	6,40	6,38	7,42	7,38	16,38	16,30	13,55	10,91
Densidad de metilación de CHH (%)	0,08	0,07	0,19	0,18	0,29	0,29	0,11	0,11	0,16	0,12

	Sangre de varón adulto	Sangre de mujer adulta	Plasma de varón adulto	Plasma de mujer adulta
Watson	Crick	Watson	Crick	Watson
196544792	196544792	197387755	179688600	141199439
43,52	42,97	44,94	22,47	18,98
85526774	84462496	88893242	4036938	26795639
6,05	6,13	6,29	6,37	6,09
1,99	1,73	1,88	1,53	7,89
Tasa de ambigüedades (%) ^a				
Tasa de duplicación (%) ^b				
Profundidad del genoma completo tras la eliminación de la tasa de ambigüedad y de conversión lambda (%) ^c	4,13	4,09	4,28	4,24
Número de M de la sec C genómica ^d	99,96	99,97	99,97	99,97
Número de N de la sec C genómica ^d	127154246	125487793	124983452	123238608
Cobertura de la C genómica (%)	2047971831	2026731827	2058153525	203988644
Profundidad de la sec C genómica	51,56	51,55	55,17	56,06
Densidad de metilación de C genómica (%)	3,72	3,68	3,74	3,70
Número de M de la sec de CpG	5,85	5,83	5,72	5,70
Número de N de la sec de CpG	125161042	123578850	1223029828	121374309
Cobertura de CpG (%)	49564325	49936003	47562383	47812716
Profundidad de la sec de CpG	73,37	73,25	76,24	76,12
Densidad de metilación sin CpG (%)	6,20	6,16	6,06	6,01
Número de M de la sec sin CpG	71,63	71,22	72,12	71,74
Número de N de la sec sin CpG	1979410	1899750	1951160	1856642
Cobertura de no CpG (%)	1960312392	195030795	1976031265	1966332257
Profundidad de la sec sin CpG	0,10	0,10	0,11	0,11
Densidad de metilación sin CpG (%)	0,69	0,69	0,70	0,70
Número de M de la sec de CHG	0,10	0,10	0,10	0,09
Número de N de la sec de CHG	618974	598717	5943336	573203
Cobertura de CHG (%)	612387336	610126996	607312325	605093927
Profundidad de la sec de CHG	53,24	63,11	66,75	66,62
Densidad de metilación de CHG (%)	4,96	4,94	4,91	4,90
Número de M de la sec sin CHH	0,10	0,10	0,10	0,09
Número de N de la sec sin CHH	1360436	1301033	12358824	1283439
Cobertura de CHH (%)	1347925055	1340180699	1368718940	1361138330
Profundidad de la sec de CHH	46,93	46,83	50,48	50,38
Densidad de metilación de CHH (%)	3,12	3,10	3,17	3,16

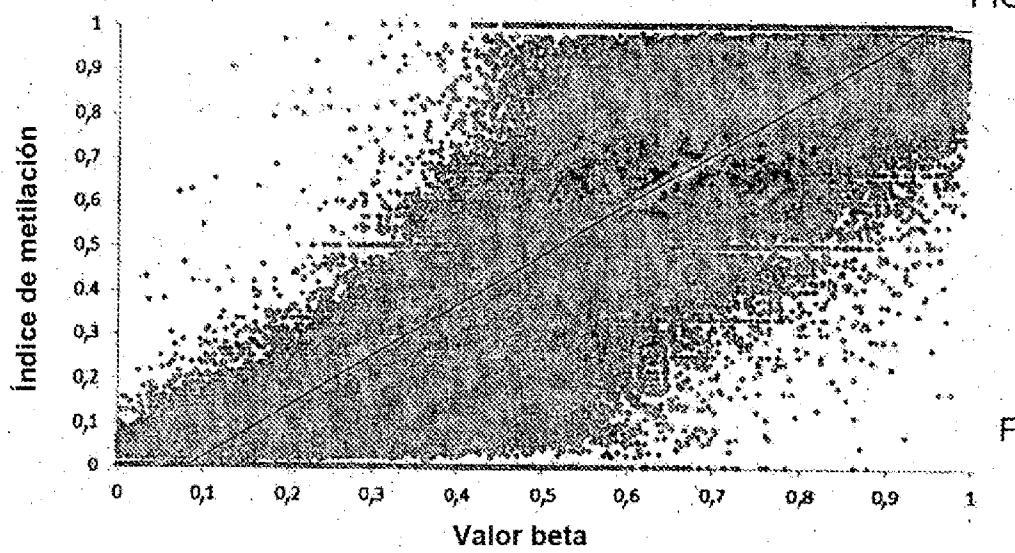
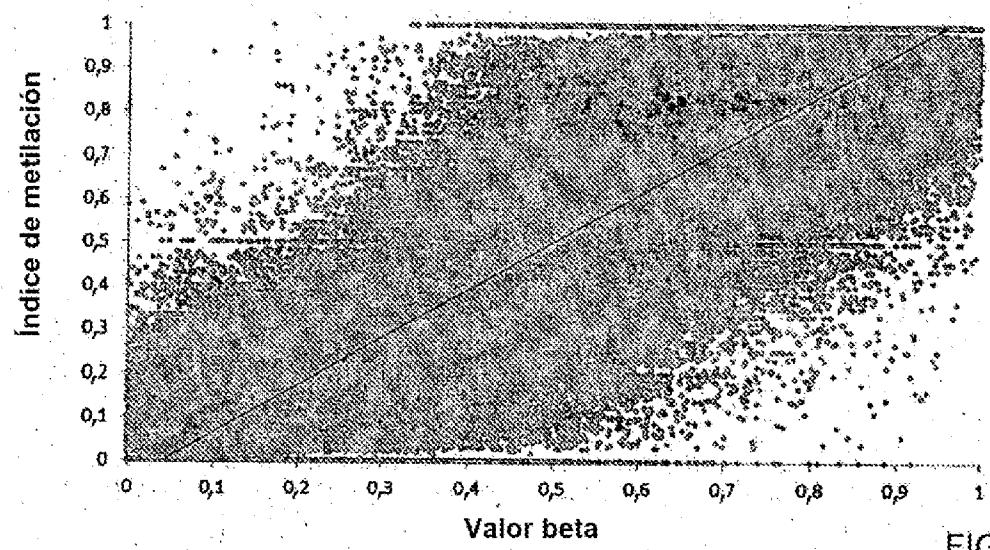
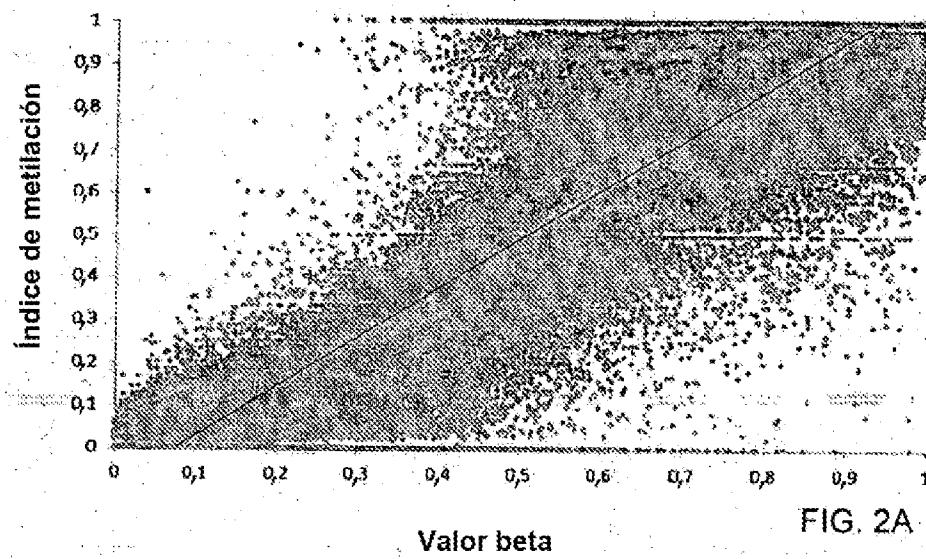
FIG. 1A (Cont.)

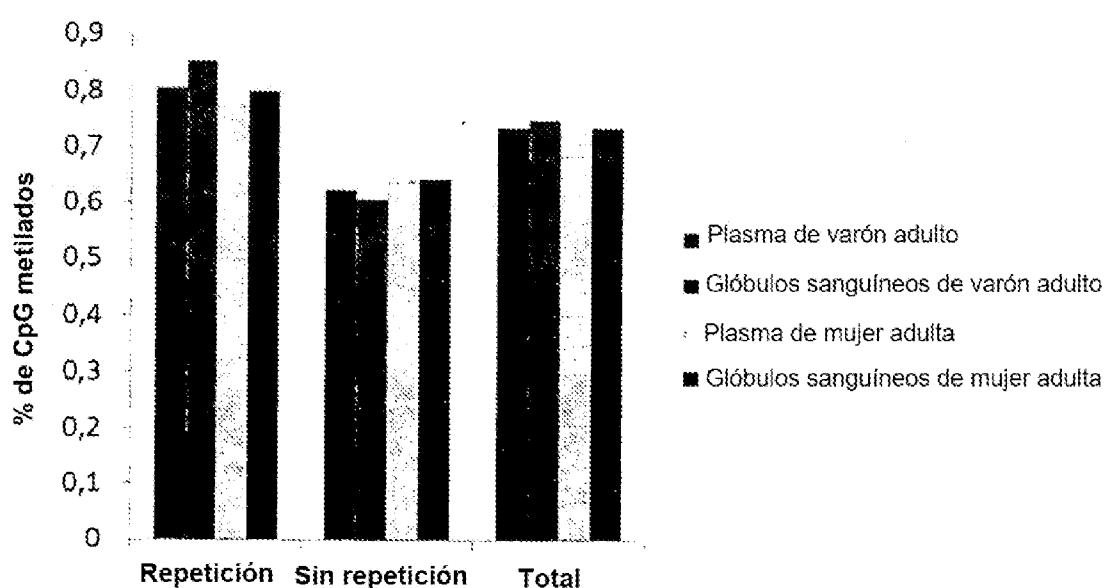
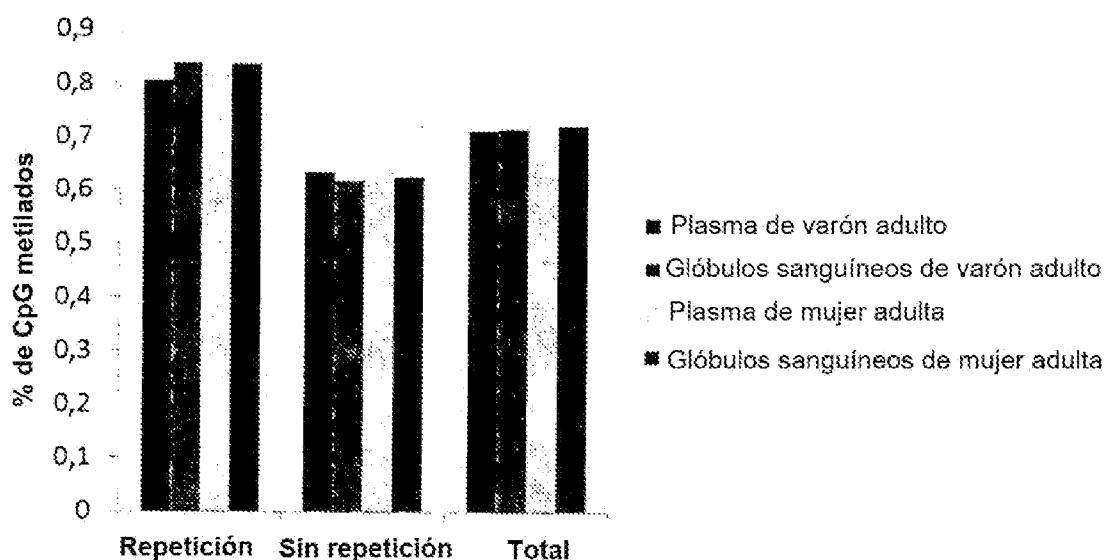
ES 2 994 426 T3



150

FIG. 1B





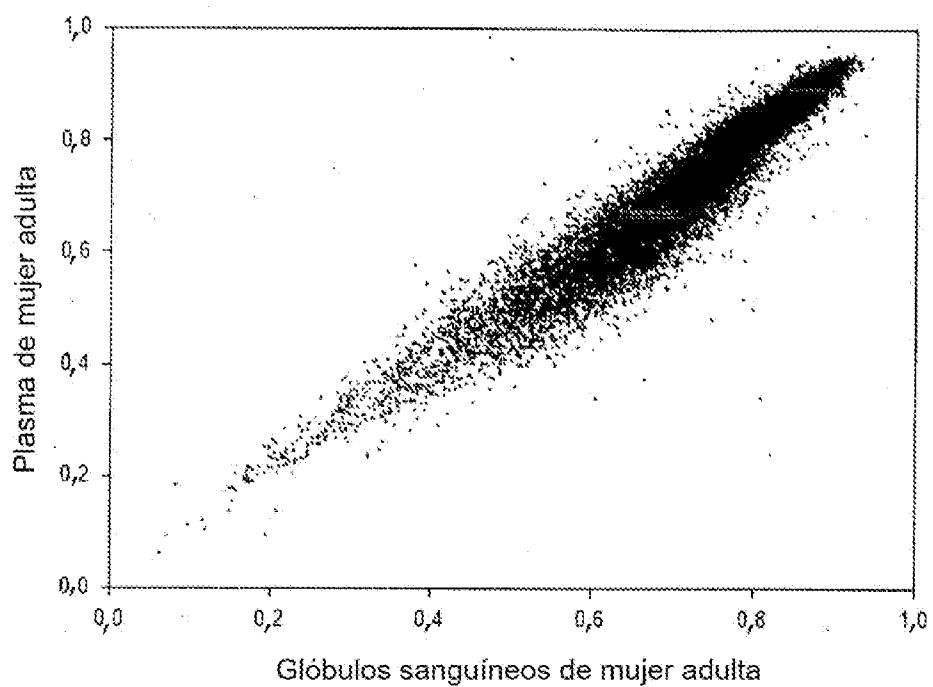


FIG. 4A

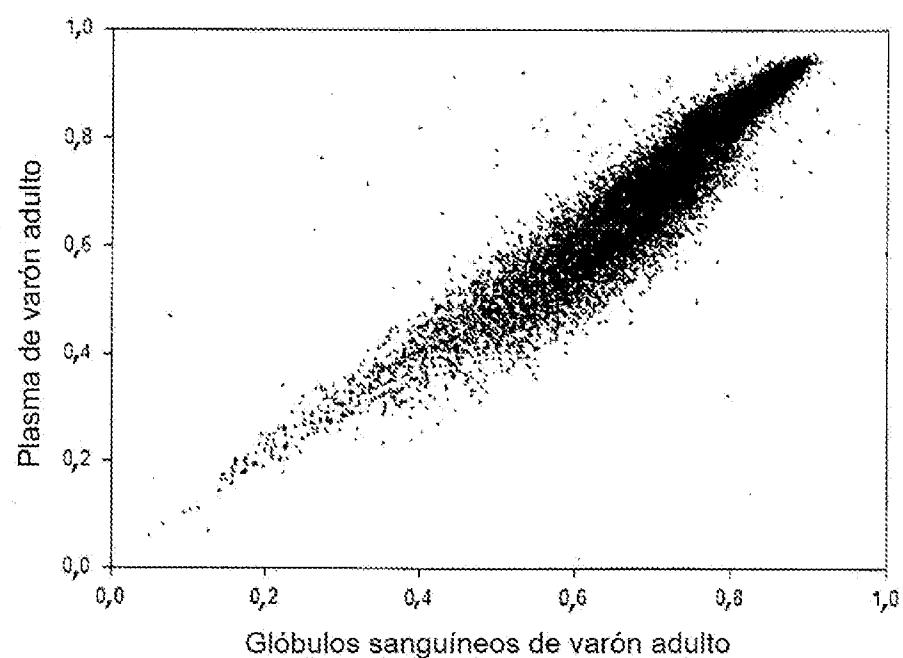


FIG. 4B

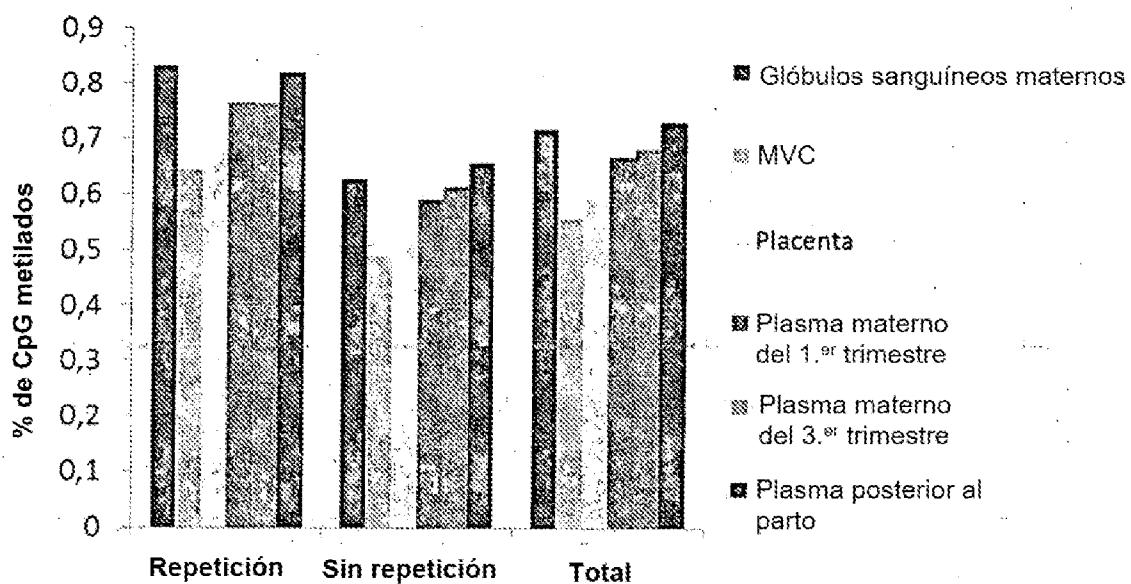


FIG. 5A

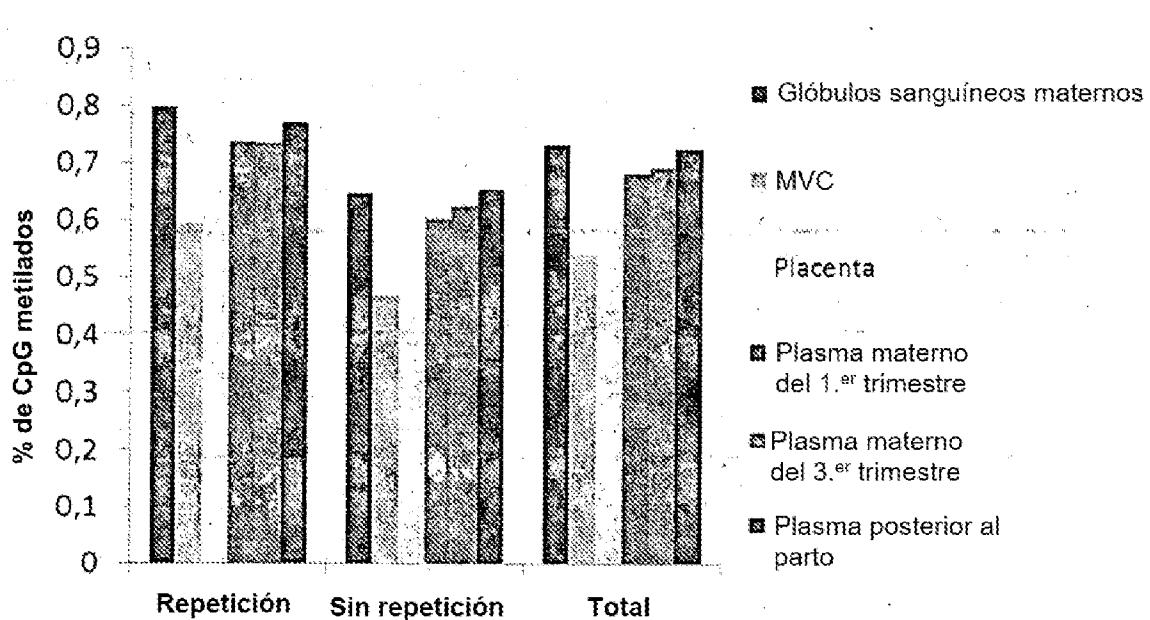


FIG. 5B

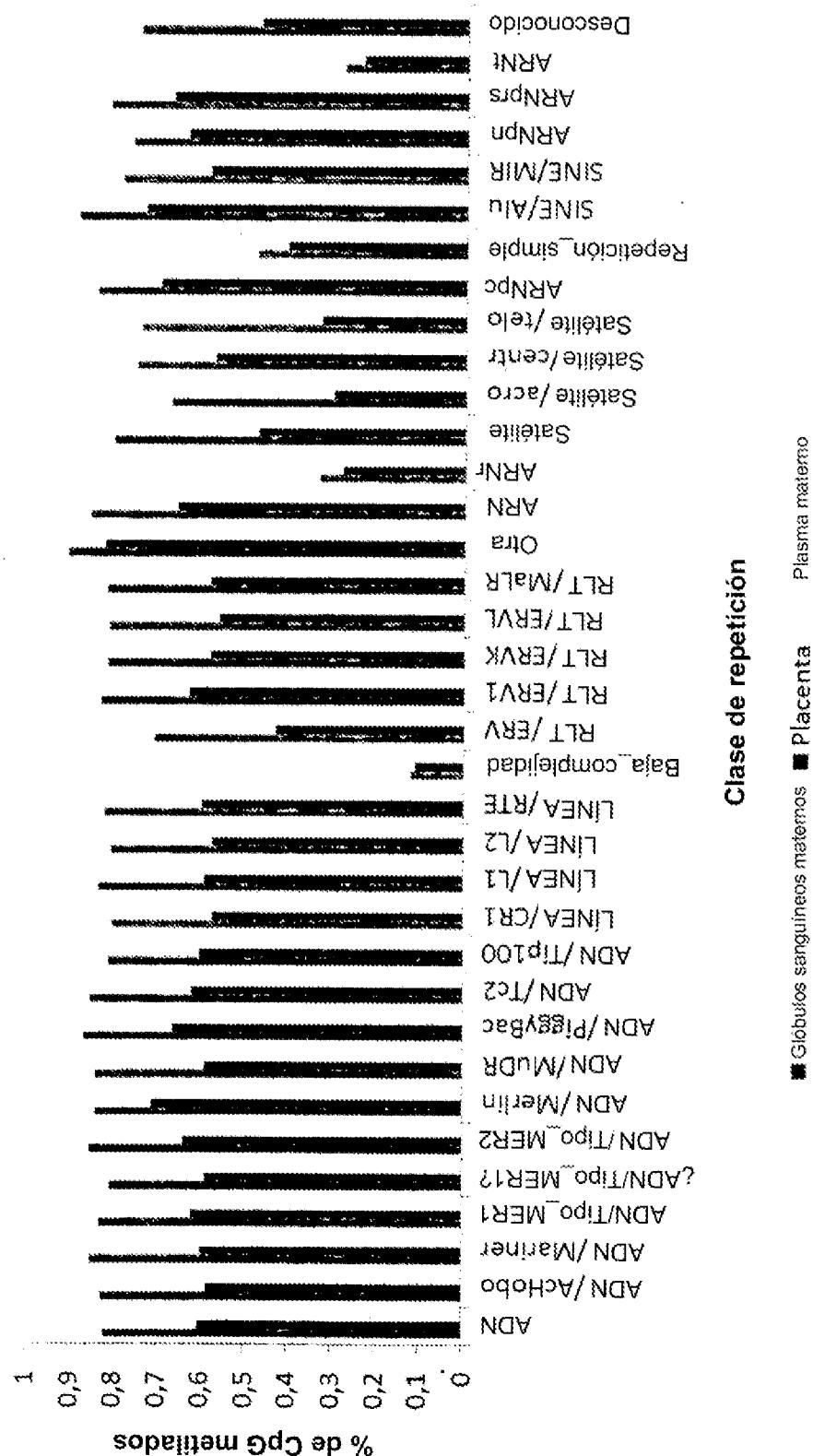


FIG. 6

700

Desde fuera hacia adentro:

Densidad de CpG

Densidad de CPO
Globulos sanguíneos maternos

Plasma materno (1.º trimestre)

Lecturas de plasma compartidas

Lecturas de plasma específico de MVC

MVC

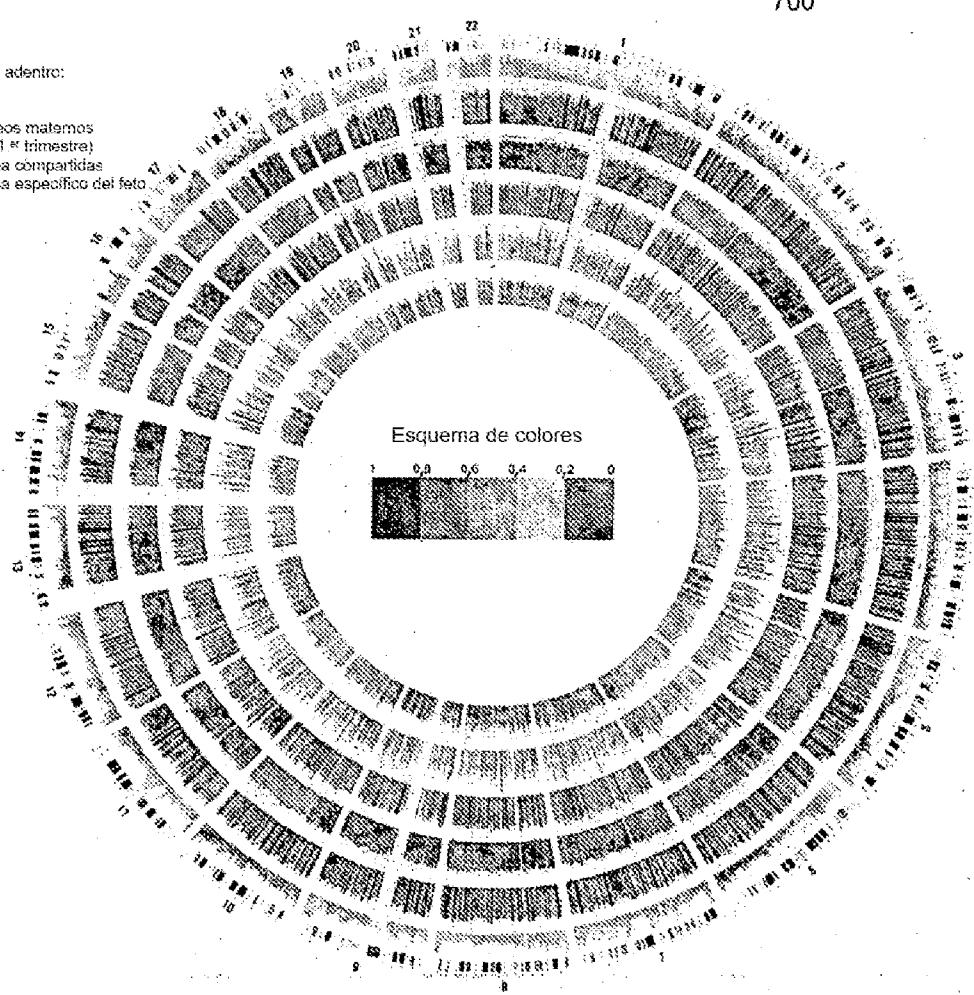


FIG. 7A

750

Desde fuera hacia adentro:
Densidad de CpG
Globulos sanguíneos maternos
Plasma tras el parto
Plasma materno (3.^{er} trimestre)
Lecturas de plasma compartidas
Lecturas de plasma específico del feto
Tejido de placenta

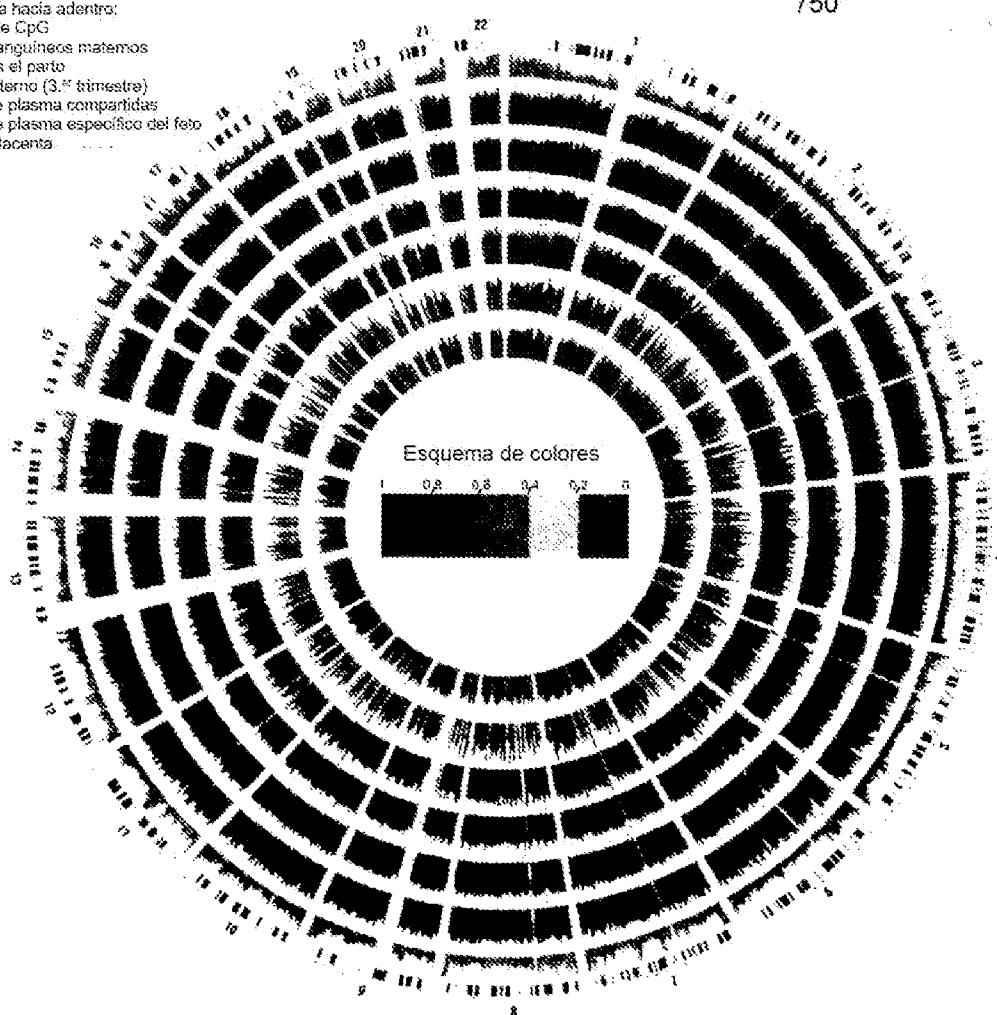
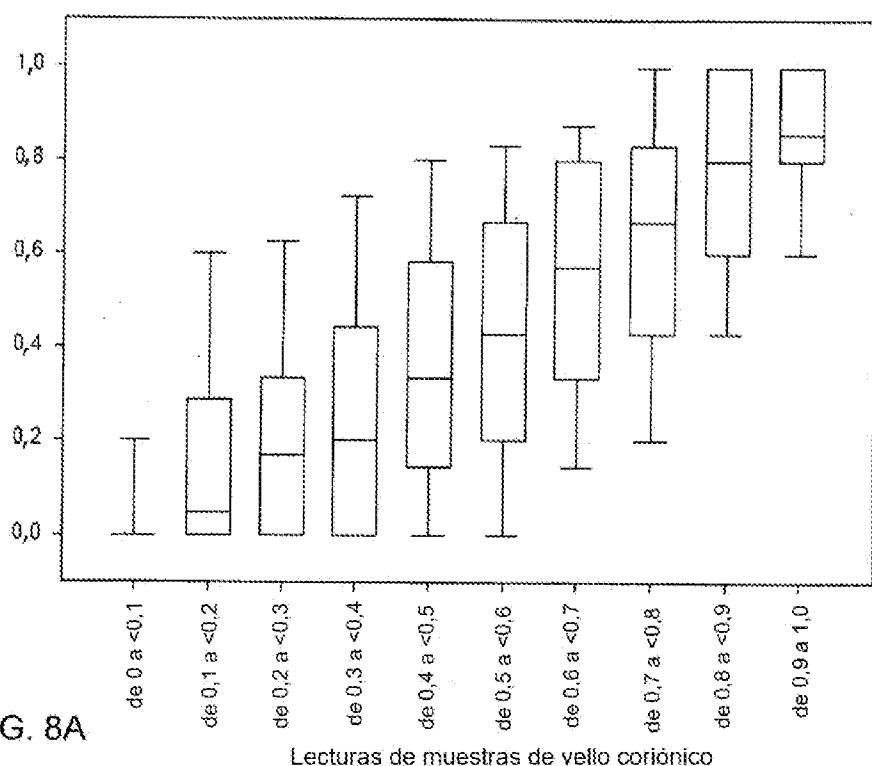
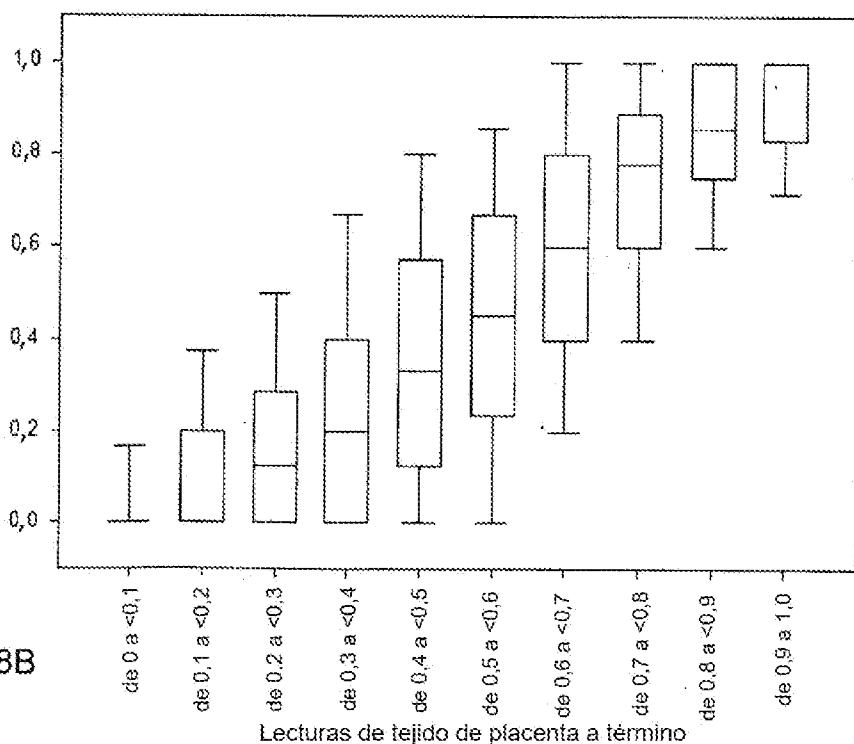


FIG. 7B

Lecturas específicas del feto en plasma materno del 1.^o trimestre



Lecturas específicas del feto en plasma materno del 3.^o trimestre



Lecturas compartidas en plasma materno del 1.^{er} trimestre

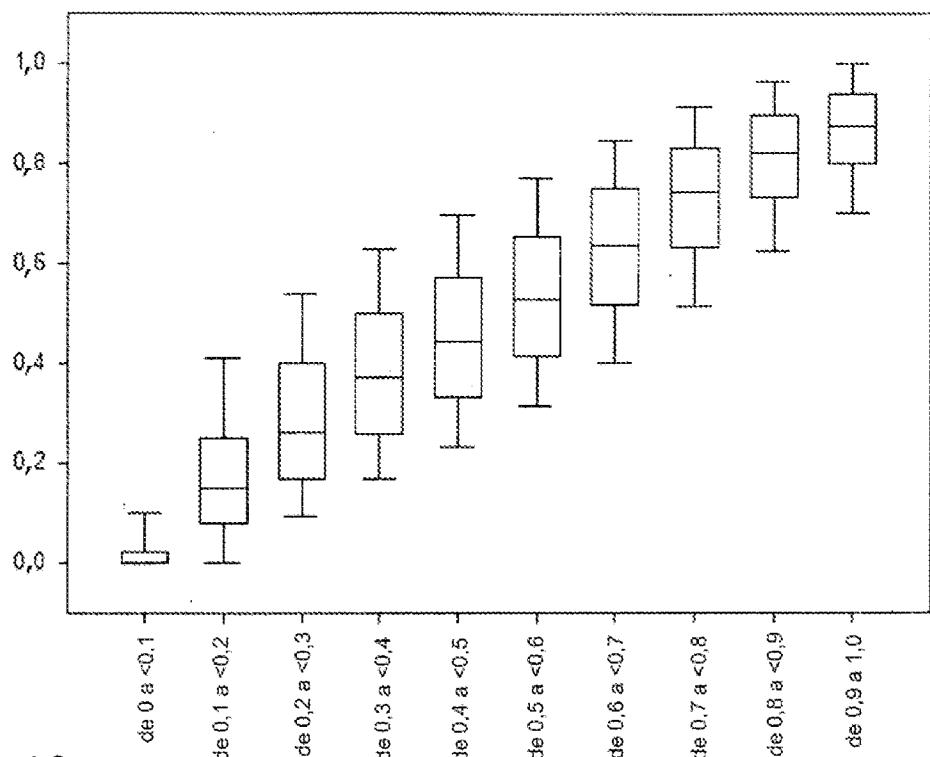


FIG. 8C

Lecturas de glóbulos sanguíneos maternos

Lecturas compartidas en plasma materno del 3.^{er} trimestre

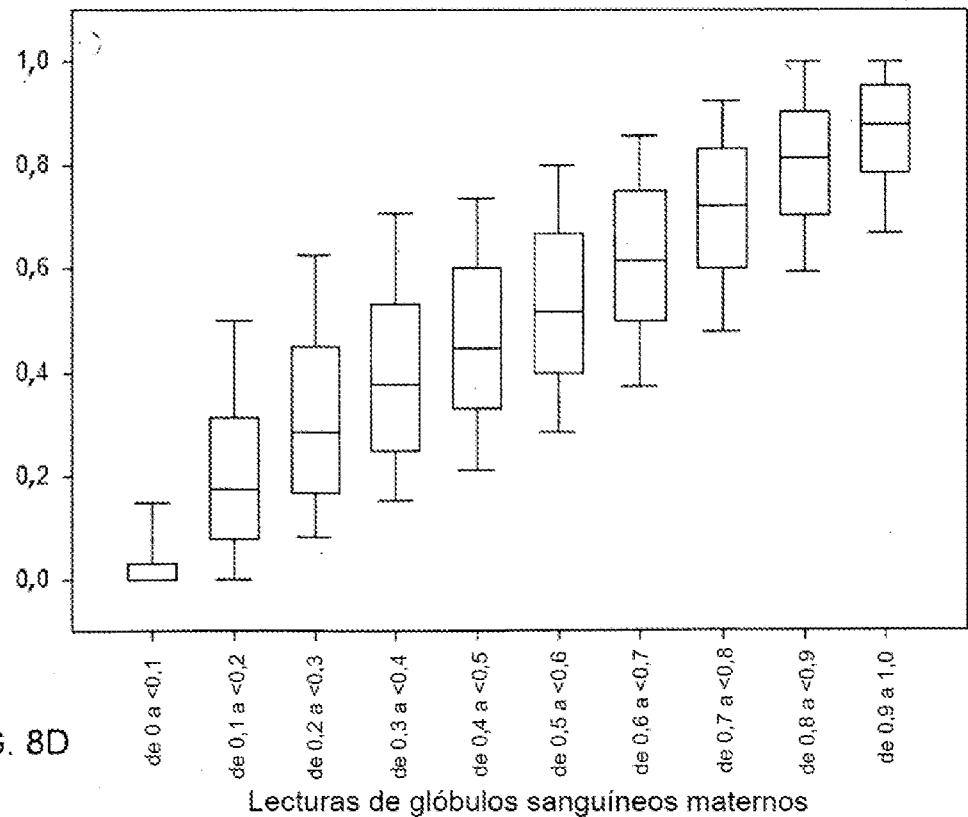


FIG. 8D

Lecturas de glóbulos sanguíneos maternos

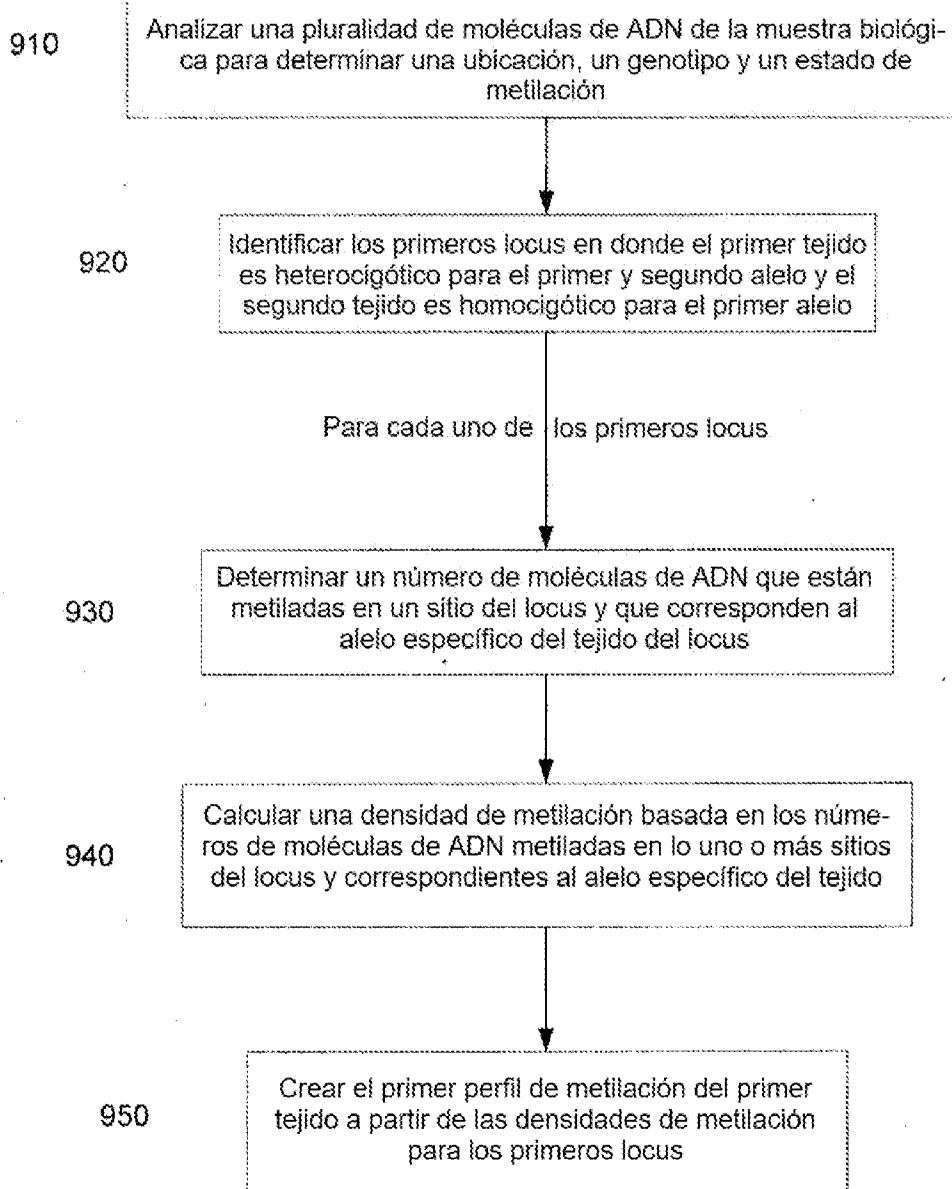


FIG. 9

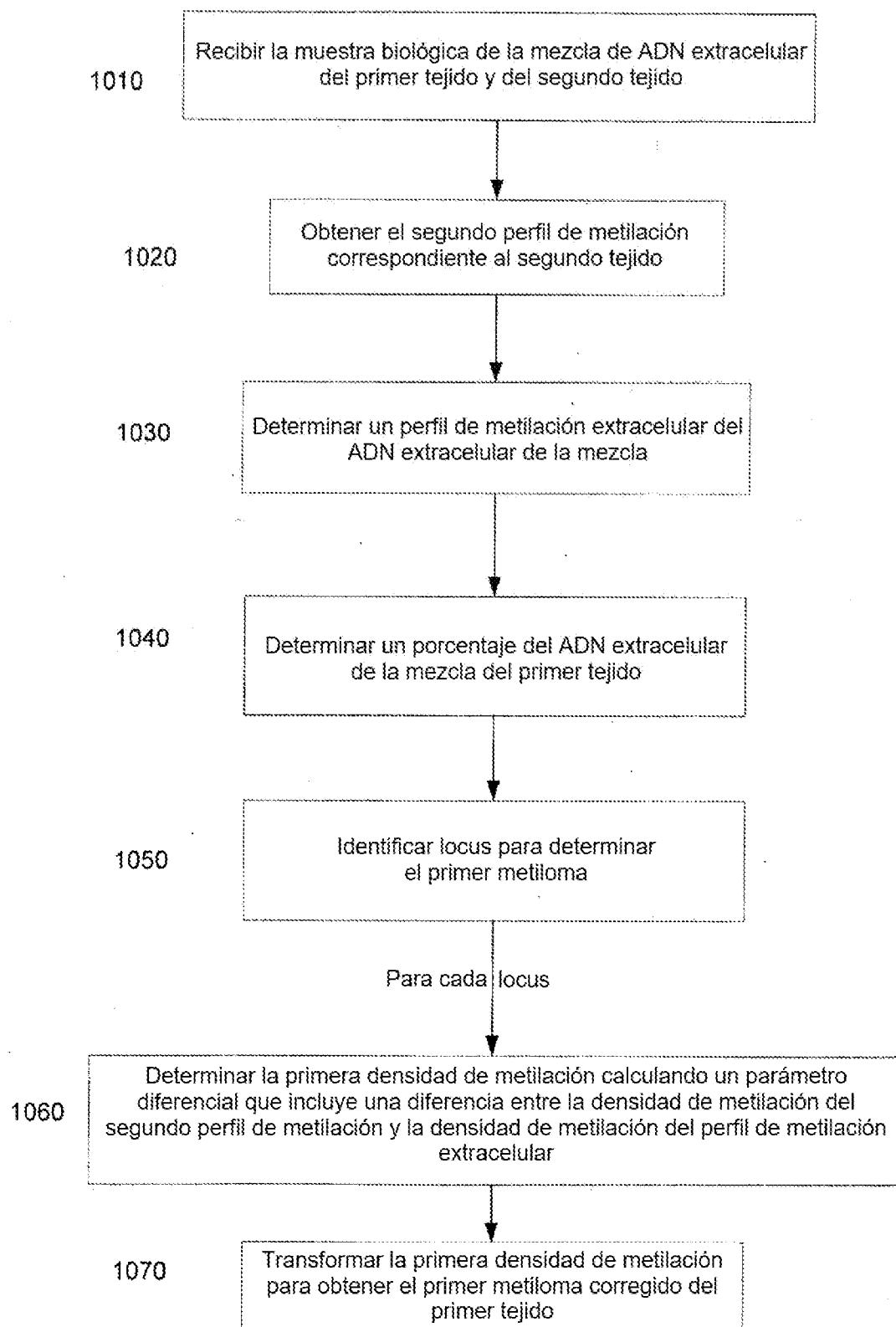


FIG. 10

1000

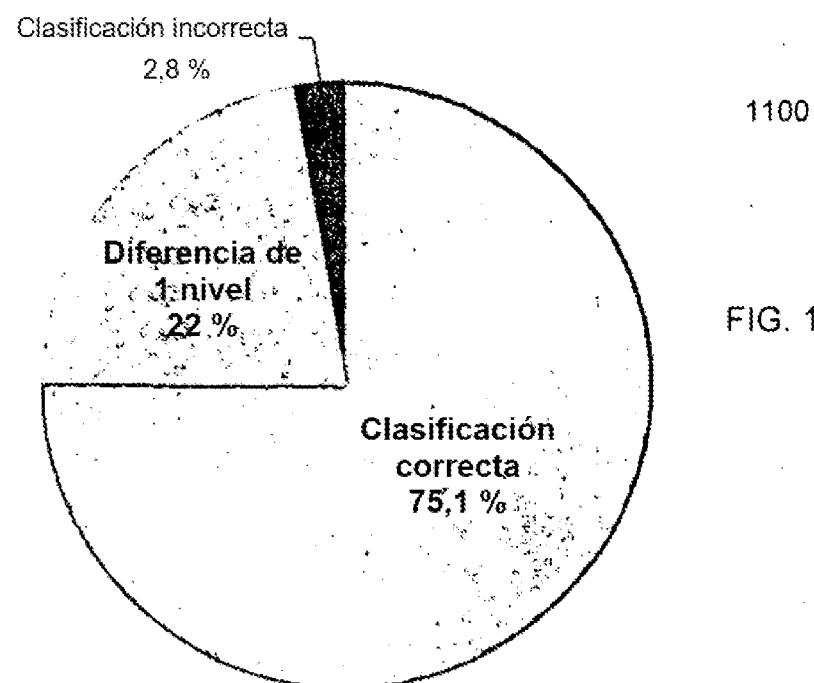


FIG. 11A

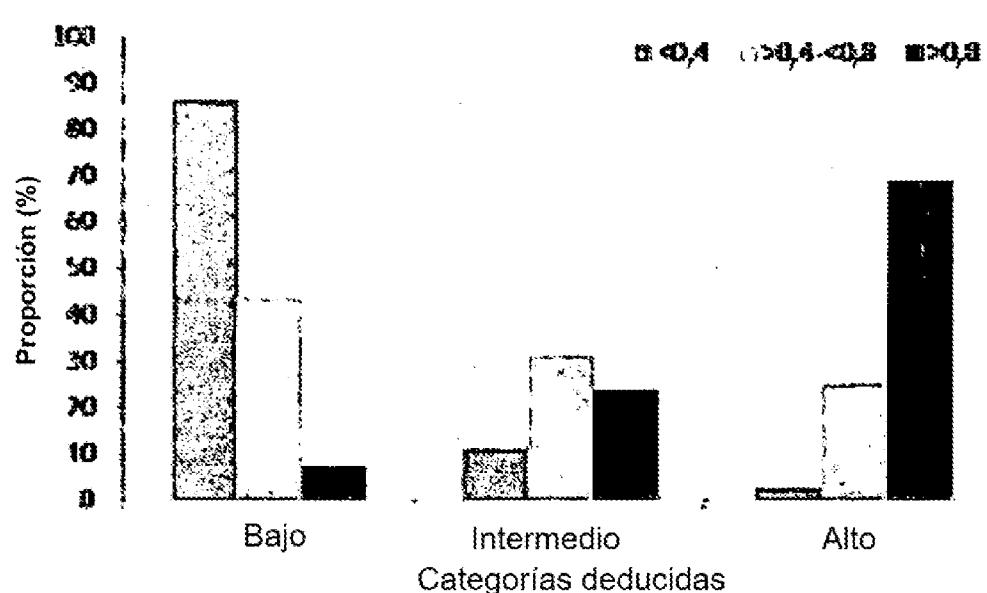


FIG. 11B

N.º Locus genético	Ubicación genómica	Densidad de metilación (datos de secuenciación)		Deducción		Estado de metilación de las iteraciones			
		Sangre materna (BASAL)	Placenta	Plasma materno dedducido (PM)	Valores dedducido (D) ¹	Punturación predictiva de la metilación deducida ³ (PPM) ²	Categorías metilación de la placenta ⁴	Nivel de metilación de la placenta ⁵	Referencias
1 RASSF1A	cr. 3 :50352938-50353234	0,000	0,637	0,067	0,506	1,060	Alto	Inter	Hipometilado [1]
2 CG1009	cr. 21 :25856489-25856756	0,001	0,432	0,050	0,374	0,849	Alto	Inter	Hipometilado [2]
3 CG1028	cr. 21 :33321342-33322182	0,035	0,331	0,040	0,073	0,367	Bajo	Bajo	Hipometilado [2]
4 CG1051	cr. 21 :36995232-36995509	0,043	0,303	0,045	0,072	0,365	Bajo	Bajo	Hipometilado [2]
5 CG1071	cr. 21 :41140466-41140898	0,031	0,531	0,046	0,146	0,483	Inter	Inter	Hipometilado [2]
6 CG1113	cr. 21 :44953820-44953972	0,013	0,547	0,051	0,306	0,739	Inter	Inter	Hipometilado [2]
7 CG1337	cr. 21 :46243872-46250196	0,940	0,413	0,220	0,030	0,299	Bajo	Inter	Hipometilado [2]
8 CG1149	cr. 21 :46911575-46912533	0,004	0,440	0,030	0,202	0,573	Inter	Inter	Hipometilado [2]
9 HLC5	cr. 21 :372774945-372775031	0,038	0,576	0,059	0,203	0,575	Inter	Inter	Hipometilado [2]
10 PDE9A	cr. 21 :42979270-42979525	0,950	0,303	0,823	-0,018	0,222	Bajo	Bajo	Hipometilado [2]
11	cr. 21 :36589647-36589995	0,952	0,149	0,827	0,007	0,262	Bajo	Bajo	Hipometilado Sin publicar
12 SERPINB5	cr. 18 :59294739-59295038	0,936	0,530	0,856	0,329	0,777	Inter	Inter	Hipometilado [3]
13 VAPA	cr. 18 :10025563-10025186	0,036	0,752	0,082	0,382	0,861	Alto	Inter	Hipometilado [4]
14 CABLES1	cr. 18 :19035954-19036069	0,335	0,801	0,491	1,519	2,680	Alto	Alto	Hipometilado [4]
15 B4GALT6	cr. 18 :27486040-27486932	0,113	0,827	0,331	1,764	3,072	Alto	Alto	Hipometilado [4]

1200

FIG. 12A

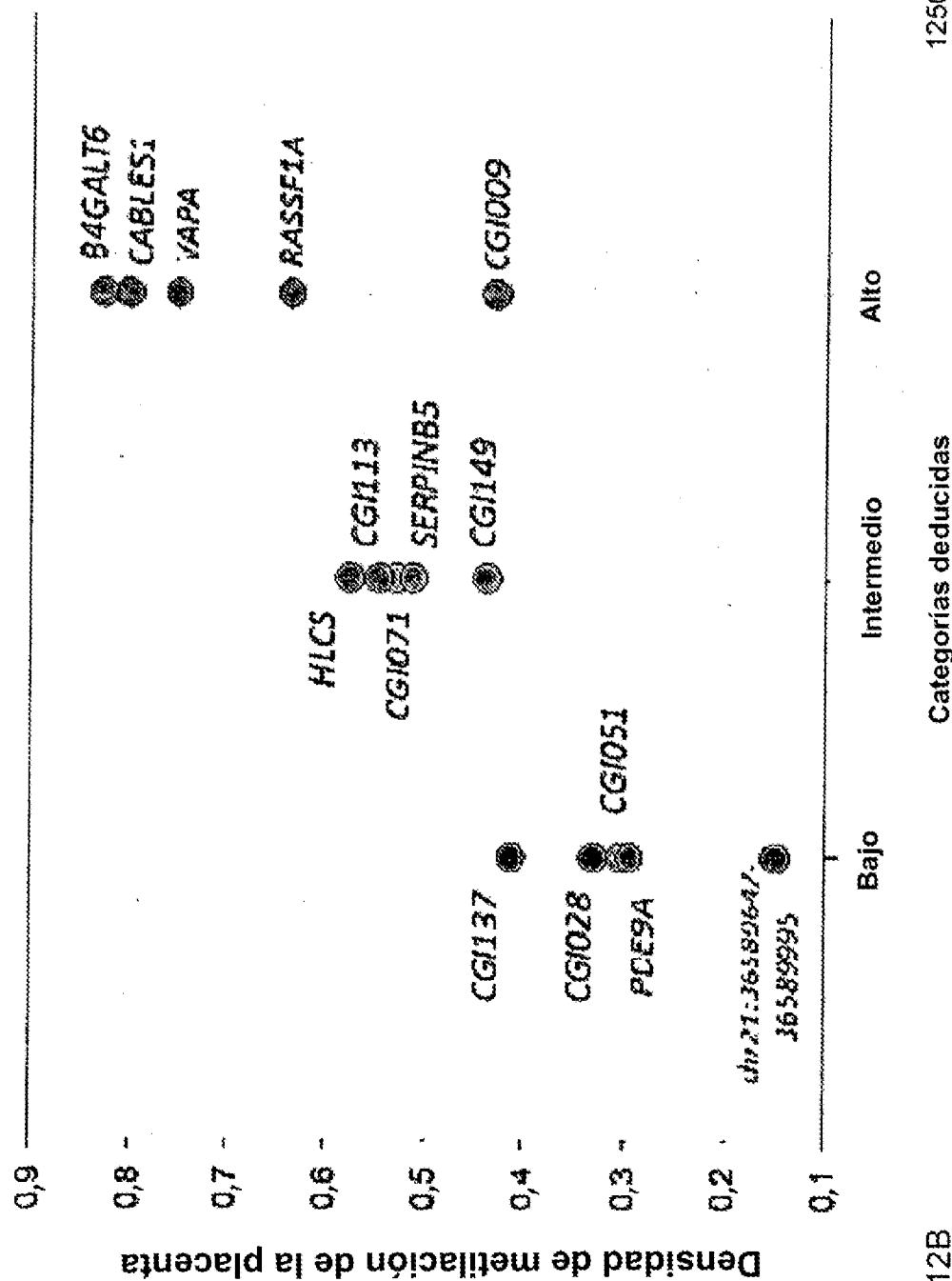


FIG. 12B

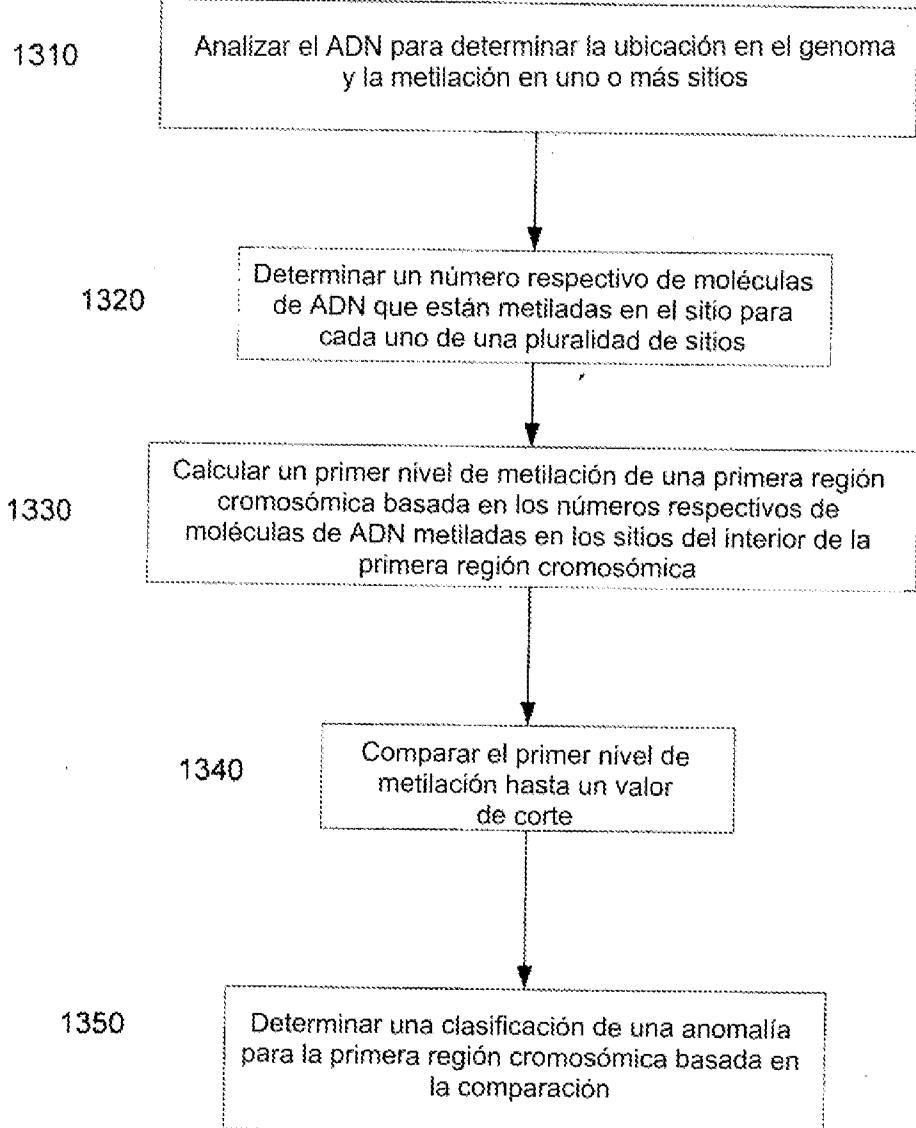


FIG. 13

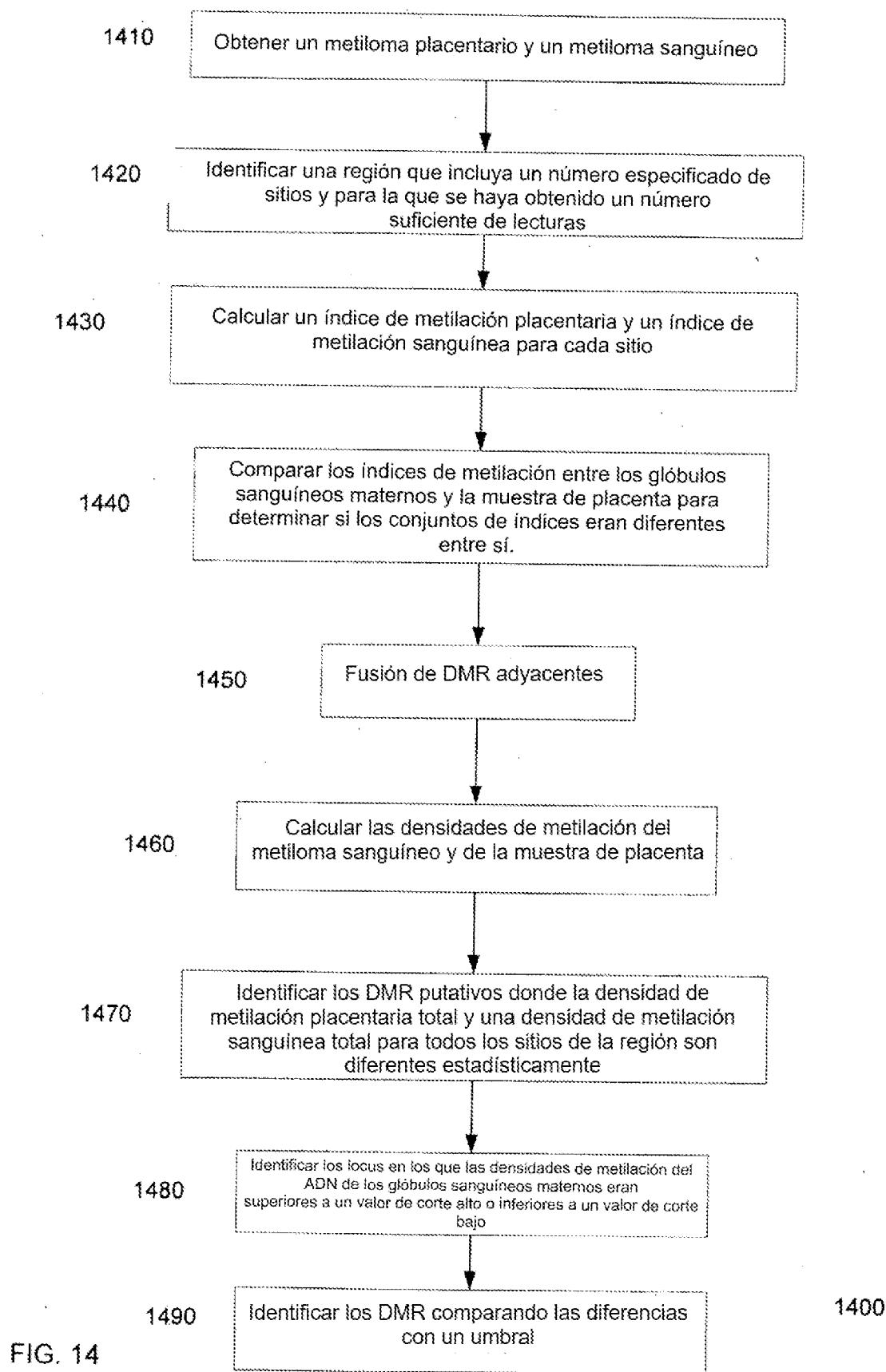


FIG. 14

N. ^a	Locus genético	Ubicación	Tamaño ^b	Estado de la placentas ^c	BS-Sq ^d			DMR ^e								
					Globulos sanguíneos maternos	MVC	Dif. de mat.	cr.	Inicio	Final	Tamaño	Globulos sanguíneos maternos	MVC	valor de p	Dif. de mat.	Closer ^f
1	RASSF1A	cr.3 50352938- 90353234	297	Hipermetilado	0,00	83,73 63,73		cr. 3	50351000	50354000	3000	0,26	65,68 0,00E+00	80,42		VP
2	CGI039	cr.21 25856485- 25856736	268	Hipermetilado	0,08	43,26 43,73		cr. 21	25855500	25857500	2000	4,97	48,70 0,00E+00	43,73		VP
3	CGI028	cr.21 3321942- 3322182	241	Hipermetilado	3,51	33,10 29,56		cr. 21	33316400	33323900	7500	4,52	23,85 0,00E+00	19,33		FN
4	CGI051	cr.21 38992322- 36985539	270	Hipermetilado	4,26	30,33 26,07		cr. 21	38992800	37003800	11100	3,12	21,91 0,00E+00	18,76		FN
5	CGI071	cr.21 41140486- 41140885	433	Hipermetilado	3,09	53,06 46,97		cr. 21	41140000	41141000	1000	3,40	52,00 0,00E+00	46,80		VP
6	CGI113	cr.21 44553620- 44553672	153	Hipermetilado	1,27	54,59 53,42		cr. 21	44548500	44554700	8200	7,18	57,23 0,00E+00	50,95		VP
7	CGI137	cr.21 46249872- 46250195	325	Hipometilado	94,00	41,34 32,65		cr. 21	46227800	46250400	23600	89,74	52,67 0,00E+00	37,07		VP
8	CGI148	cr.21 48011575- 48912533	959	Hipermetilado	0,40	43,26 43,56		cr. 21	48011200	48912700	1500	6,31	46,17 0,00E+00	40,86		VP
9	HLCS	cr.21 37274945- 37275031	87	Hipermetilado	3,77	57,91 53,85		cr. 21	37274200	37275200	1000	7,03	54,24 0,00E+00	46,81		VP
10	POEBA	cr.21 42679270- 42979525	258	Hipometilado	85,01	30,33 64,68		cr. 21	42978100	43004800	29700	89,97	51,44 0,00E+00	38,53		VP
11	-	cr.21 36585647- 36585955	349	Hipometilado	95,17	14,89 80,29		cr. 21	36586200	36586200	1000	94,49	13,27 0,00E+00	75,22		VP
12	SERPINB8	cr.18 59294739- 59295038	300	Hipometilado	93,64	53,00 40,54		cr. 18	59294800	59296300	1500	88,47	40,78 0,33E-27	47,71		VP
13	VAFA	cr.18 10027743- 10028142	625	Hipermetilado	0,60	75,17 71,54		cr. 18	10021800	10023300	1600	24,08	80,33 0,00E+00	58,26		VP
14	CABLES1	cr.18 19535964- 19536072	117	Hipermetilado	33,47	80,09 48,82		-	-	-	-	-	-	-		FN
15	B4GALT8	cr.18 27488604- 27488633	894	Hipermetilado	15,34	82,75 71,41		cr. 18	27488100	27487100	1000	11,34	82,75 0,6408E- 71,41			VP
16	CABP8	cr.22 201630342- 201631869	1544	Hipermetilado	0,52	47,84 47,03		cr. 2	201630200	201632200	2000	3,77	49,13 4,22E-42	45,38		VP
17	RARB	cr.3 252444840- 25445090	251	Hipermetilado	2,28	28,93 24,65		cr. 3	25443600	25445400	1500	3,44	22,97 0,65E-50	19,53		FN
18	APC	cr.6 11210099- 112101943	945	Hipermetilado	0,71	45,78 45,06		cr. 9	112100300	112101800	1500	3,59	46,09 0,00E+00	44,30		VP
19	DAB2IP_ex	cr. 9 121541221- 121541480	240	Hipometilado	84,87	52,83 32,24		cr. 9	121540400	121541800	1100	85,21	38,86 5,2805E- 48,35			VP
20	DAB2IP_ex	cr. 9 123128251- 123129077	817	Hipometilado	85,09	32,99 62,10		cr. 9	123127700	123129200	1500	80,72	50,61 1,4700E- 40,11			VP
21	THY1	cr.11 118798952- 116798382	431	Hipermetilado	0,88	39,32 39,44		cr. 11	118797600	118799600	2000	2,25	27,26 0,00E+00	25,03		VP
22	PTPN6	cr.12 68393211- 6833772	3562	Hipermetilado	1,52	78,89 78,29		cr. 12	68360000	68340000	4000	4,26	80,13 0,00E+00	75,87		VP
23	PYCARGO_ex	cr.19 31121808- 31122146	339	Hipermetilado	4,90	46,51 41,61		-	-	-	-	-	-	-		FN
24	cr. 18A	cr.18 550908284- 550908306	323	Hipermetilado	8,14	31,79 25,65		cr. 18	55089700	55094200	4500	4,78	23,91 0,00E+00	18,13		FN
25	cr. 25A	cr.21 39279723- 39280304	282	Hipermetilado	12,80	84,52 51,21		cr. 21	39276300	39279600	500	4,74	82,07 7,17E-62	57,33		VP
26	cr. 21B	cr.21 44161027- 44181371	345	Hipermetilado	0,60	83,98 83,04		cr. 21	44160700	44161700	1000	11,05	77,32 0,00E+00	89,20		VP
27	cr. 21C	cr.21 33320644- 33320829	286	Hipermetilado	5,49	37,92 32,43		cr. 21	33316400	33323900	7500	4,52	23,85 0,00E+00	19,33		FN
28	cr. 21D	cr.21 42189223- 42189683	461	Hipermetilado	13,02	85,84 72,83		cr. 21	42188700	42189700	1000	7,78	81,67 3,82E-77	73,89		VP
29	cr. 21E	cr.21 42355712- 42356906	197	Hipermetilado	0,00	97,85 97,85		cr. 21	42055000	42357500	2500	5,21	62,23 0,00E+00	57,02		VP
30	cr. 21F	cr.21 42357215- 42357341	127	Hipermetilado	0,57	48,29 44,88		cr. 21	42355000	42357500	2500	5,21	62,23 0,00E+00	57,02		VP
31	cr. 21H	cr.21 32236803- 32285943	141	Hipermetilado	8,49	76,84 78,35		cr. 21	32268300	32269300	1000	20,49	85,26 4,9563E- 64,76			VP
32	cr. 21I	cr.21 44079235- 44079838	301	Hipermetilado	1,23	57,44 58,21		cr. 21	44079400	44080400	1000	15,19	86,03 4,9E-58	52,87		VP
33	SERPINB8_P	cr.18 59294811- 4865	Hipometilado	91,98	39,71 52,25		cr. 18	59294800	59296300	1500	88,47	40,76 3,3284E- 47,71			VP	

FIG. 15A

1500

ES 2 994 426 T3

N. ^a	Locus genético	BS-Seq			DMR ^b								
		Globulos sanguíneos maternos	Placenta	Dif. de met.	Cr.	Inicio	Final	Tamano	Globulos sanguíneos maternos	Placenta	valor de p	Dif. de met.	Clase ^c
1	RASSF1A	0,00	60,84	60,84	cr. 3	50352100	50354100	2000	2,15	62,21	0,00E+00	60,05	VP
2	CGI009	0,08	20,50	20,43	cr. 21	25855600	25857100	1500	2,04	24,37	0,00E+00	22,33	VP
3	CGI026	3,51	17,38	13,86	-	-	-	-	-	-	-	-	VN
4	CGI051	4,26	13,93	9,67	-	-	-	-	-	-	-	-	VN
5	CGI071	3,09	24,85	21,76	cr. 21	41139900	41141400	1500	6,83	29,28	0,00E+00	22,45	VP
6	CGI113	1,27	40,24	38,97	cr. 21	44948500	44954700	6200	7,18	45,00	0,00E+00	37,82	VP
7	CGI137	94,00	37,10	56,90	cr. 21	46246400	46250400	2000	92,28	43,85	1,66E-76	48,43	VP
8	CGI149	0,40	22,19	21,79	cr. 21	46911200	46912700	1500	5,31	25,73	0,00E+00	20,42	VP
9	HLCS	3,77	58,49	54,72	cr. 21	37274200	37275200	1000	7,33	56,83	0,00E+00	49,60	VP
10	PDE9A	95,01	32,41	82,60	cr. 21	42978900	42990100	11200	90,89	44,38	0,00E+00	46,51	VP
11	-	95,17	19,92	76,24	cr. 21	36589200	36590200	1000	94,49	22,66	0,00E+00	71,83	VP
12	SERPINE5	93,64	28,06	65,57	cr. 18	59294500	59295600	1000	92,11	36,53	3,85E-31	55,58	VP
13	VAPA	3,63	73,88	70,25	cr. 18	10022100	10023100	1000	11,67	75,76	0,00E+00	64,09	VP
14	CABLES1	33,47	81,40	47,93	-	-	-	-	-	-	-	-	FN
15	B4GALT6	11,34	88,57	77,23	cr. 18	27485800	27486800	1000	28,97	90,00	4,3779E-59	61,03	VP
16	CASP8	0,52	56,27	55,75	cr. 2	201830200	201832200	2000	3,77	57,60	1,33E-52	53,83	VP
17	RARB	2,28	38,18	36,90	cr. 3	25444000	25445000	1000	2,36	24,42	8,00E-54	22,06	VP
18	APC	0,71	34,89	34,18	cr. 6	112100700	112101700	1000	0,38	32,97	0,00E+00	32,59	VP
19	DAB2IP_ori	84,87	40,28	44,60	cr. 8	121540400	121541400	1000	84,79	30,63	1,7338E-14	54,16	VP
20	DAB2IP_ex	95,09	16,14	78,95	cr. 9	123128600	123129100	500	95,14	17,42	9,6247E-47	77,72	VP
21	THY1	0,88	22,28	21,40	cr. 11	118797800	118799300	1500	1,22	20,50	0,00E+00	19,28	FN
22	PTPN6	1,52	81,33	79,82	cr. 12	6930000	6934000	4000	4,26	81,58	0,00E+00	77,32	VP
23	PYCARD_ex	4,90	48,73	43,83	cr. 16	31121700	31122200	500	4,65	32,87	1,1138E-34	28,02	VP
24	cr. 18A	6,14	16,25	10,10	-	-	-	-	-	-	-	-	VN
25	cr. 21A	12,80	74,65	61,85	cr. 21	39278300	39260300	2000	9,96	75,07	0,00E+00	65,11	VP
26	cr. 21B	0,86	81,66	80,70	cr. 21	44160700	44161700	1000	11,06	78,6	0,00E+00	67,54	VP
27	cr. 21C	5,49	14,91	9,42	-	-	-	-	-	-	-	-	VN
28	cr. 21D	13,02	88,71	75,70	cr. 21	42188700	42189700	1000	7,78	84,42	3,5806E-77	76,64	VP
29	cr. 21E1	0,00	73,21	73,21	cr. 21	42355300	42357800	2500	4,66	79,19	0,00E+00	74,53	VP
30	cr. 21E2	3,57	84,00	80,43	cr. 21	42355300	42357800	2500	4,66	79,19	0,00E+00	74,53	VP
31	cr. 21H	0,49	77,49	77,00	cr. 21	32268400	32268900	500	18,4	83,85	2,1972E-30	65,45	VP
32	cr. 21I	1,23	58,62	57,39	cr. 21	44079300	44080300	1000	12,18	72,55	2,3559E-60	60,39	VP
33	SERPINE5_P	91,96	28,61	63,35	cr. 18	59294500	59295500	1000	92,11	36,53	3,85E-31	55,58	VP

FIG. 15B

1550

	Plasma materno del primer trimestre		Plasma materno del tercer trimestre	
	Hipometilado ^a	Hipometilado ^b	Hipometilado ^a	Hipometilado ^b
N.º de locus predichos	3081	44455	1746	14930
N.º de locus con densidades de metilación >40 % en los datos de tejido de placenta ^c	1678	n. a.	1525	n. a.
N.º de locus con densidades de metilación <60 % en los datos de tejido de placenta ^c	n. a.	23468	n. a.	13475
N.º de locus solapados con DMR extraídos de los datos de tejido de placenta ^c y de los glóbulos sanguíneos maternos	1457	21812	1279	12677

FIG. 16

1600

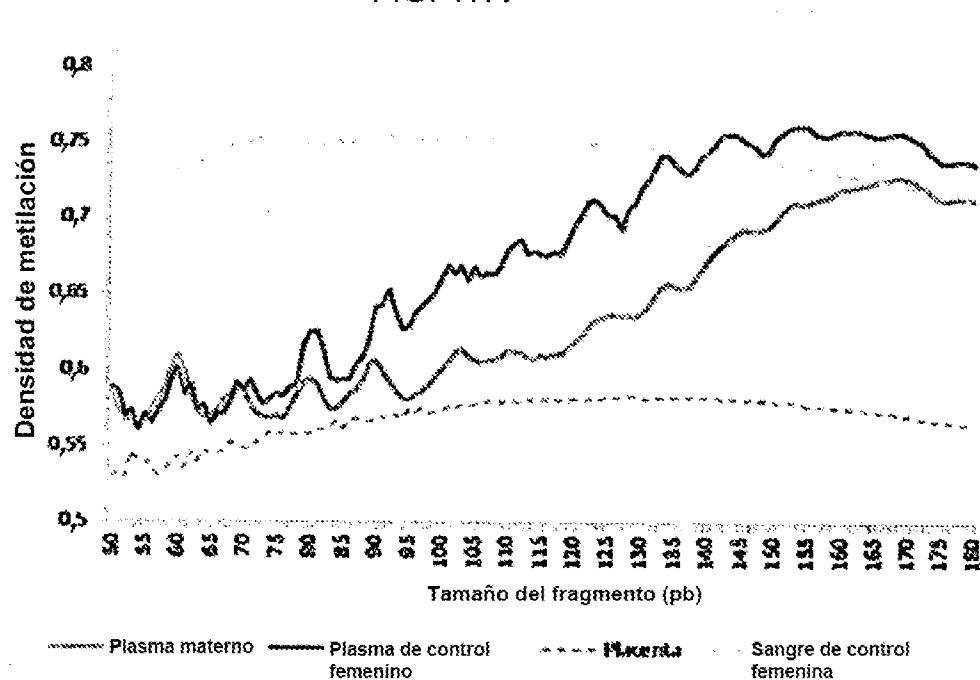
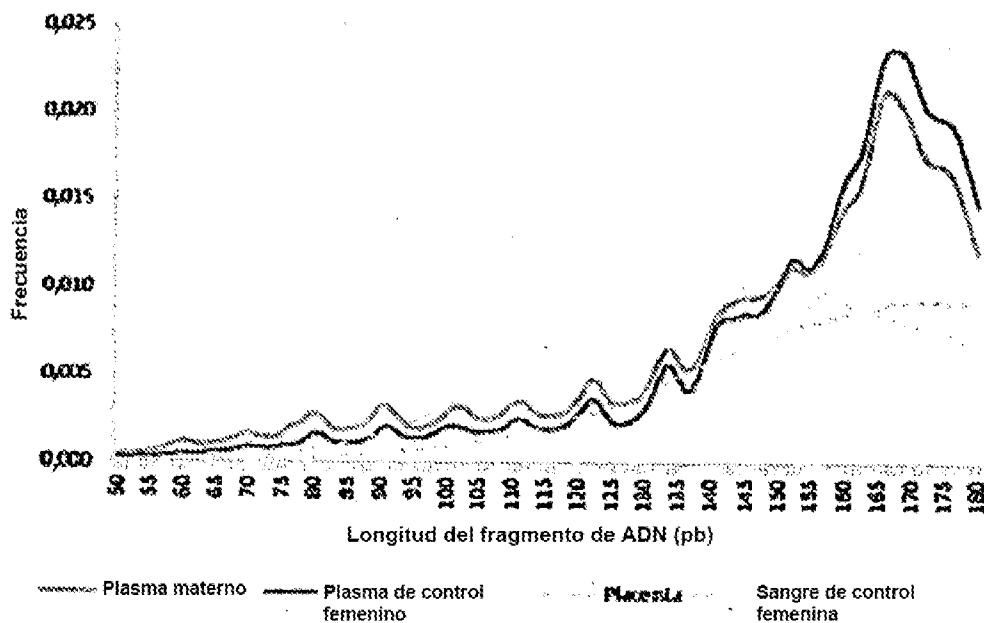


FIG. 17B

ES 2 994 426 T3

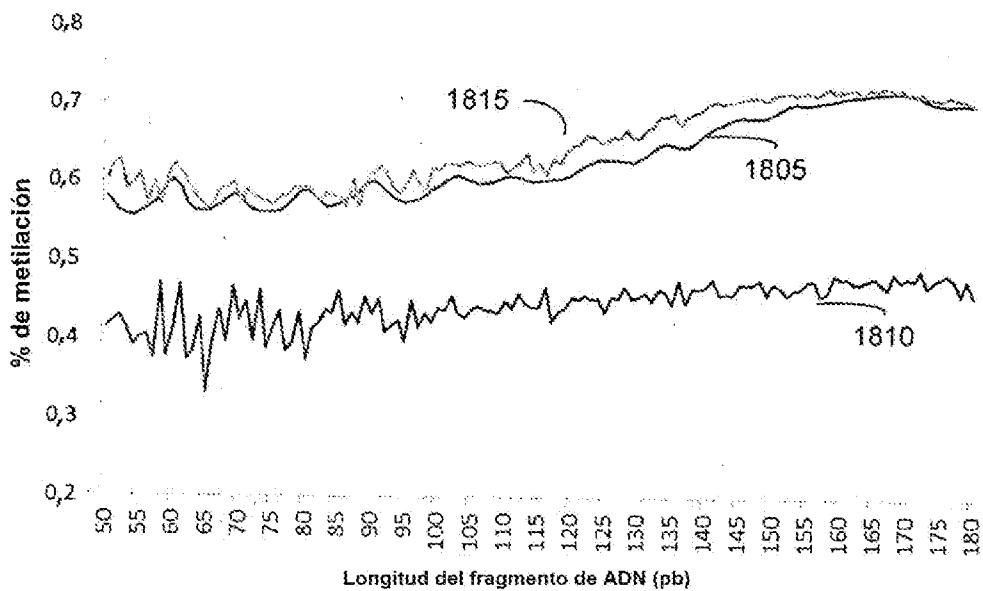


FIG. 18A
1800

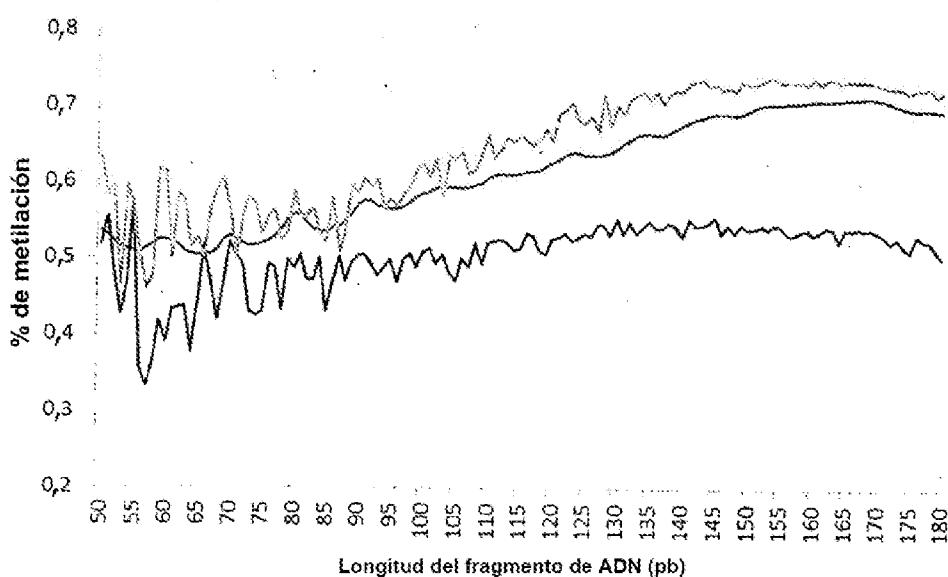


FIG. 18B
1850

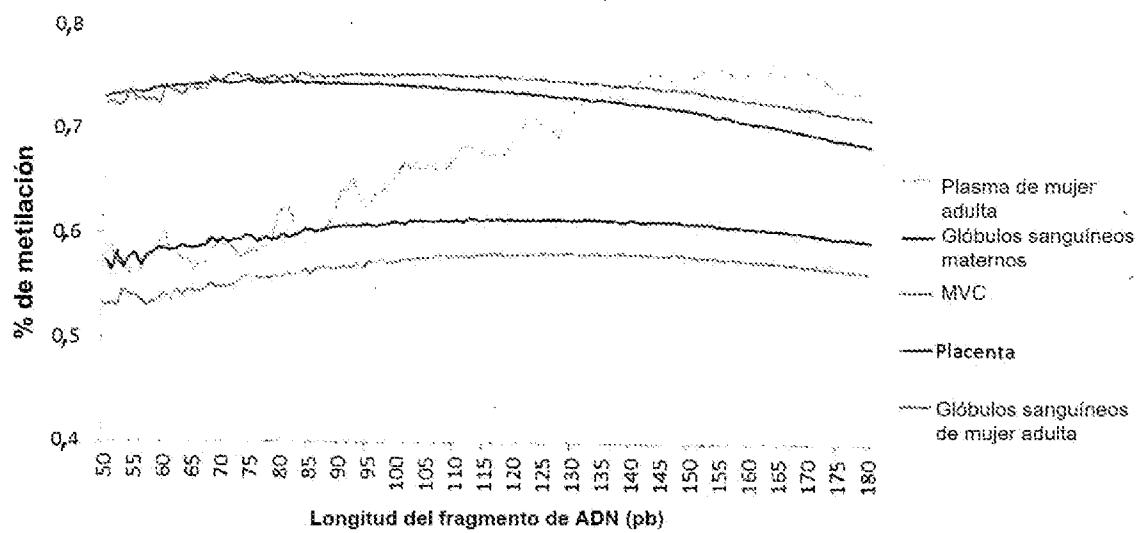


FIG. 19A

1900

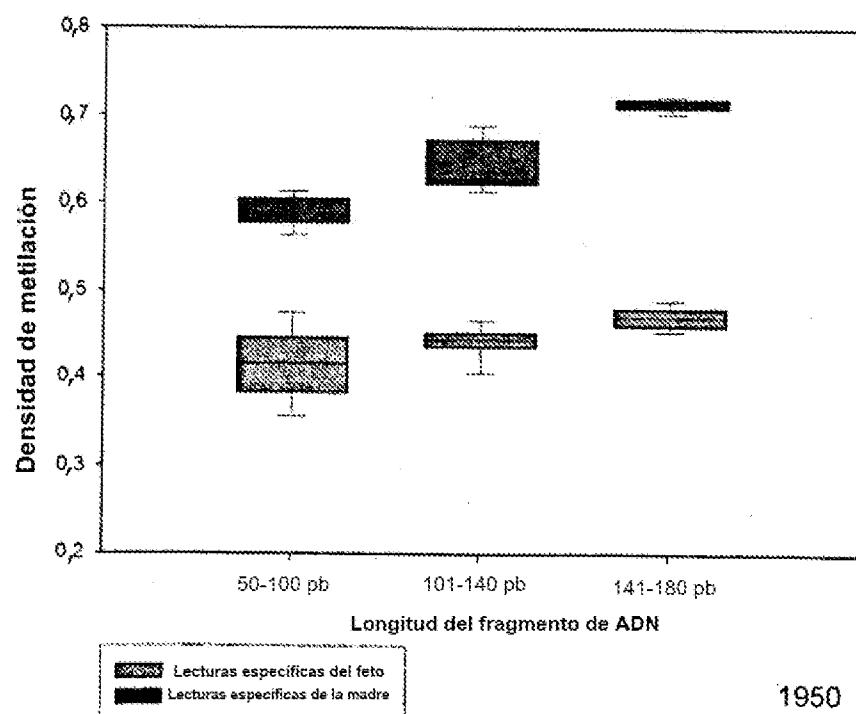


FIG. 19B

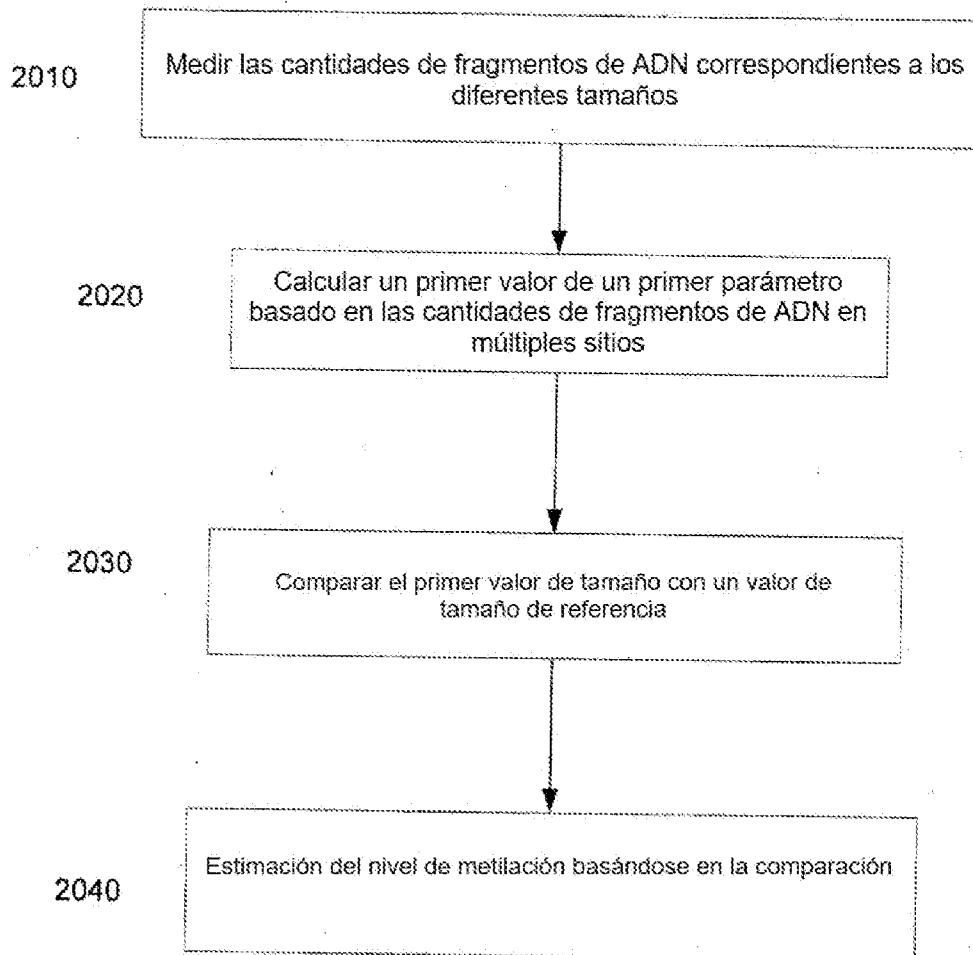


FIG. 20

2000

Cromosoma	Paciente con CHC			
	Capa leucocitaria	Tejido hepático no tumoral	Tejido tumoral	Plasma preoperatorio
1	65,4 %	65,7 %	44,1 %	60,2 %
2	68,3 %	68,2 %	42,0 %	58,6 %
3	68,2 %	68,5 %	44,6 %	60,9 %
4	68,9 %	67,7 %	37,8 %	56,9 %
5	67,4 %	66,7 %	37,8 %	56,9 %
6	67,4 %	67,7 %	43,2 %	59,9 %
7	69,1 %	68,8 %	43,0 %	60,0 %
8	69,1 %	68,0 %	40,4 %	60,0 %
9	66,4 %	66,2 %	44,9 %	60,6 %
10	68,9 %	68,2 %	40,8 %	58,5 %
11	65,5 %	65,6 %	41,6 %	58,5 %
12	67,6 %	67,4 %	45,9 %	62,0 %
13	69,8 %	68,8 %	40,6 %	60,2 %
14	66,4 %	66,2 %	41,4 %	59,2 %
15	66,1 %	66,8 %	44,6 %	61,0 %
16	68,9 %	67,6 %	45,6 %	60,7 %
17	64,0 %	64,9 %	45,8 %	60,2 %
18	69,9 %	69,1 %	42,9 %	60,0 %
19	62,3 %	62,7 %	47,5 %	61,0 %
20	66,0 %	65,3 %	39,1 %	57,8 %
21	61,8 %	61,1 %	39,9 %	57,7 %
22	67,9 %	67,6 %	49,1 %	63,8 %
Todos los autosomas	67,1 %	66,9 %	42,9 %	59,7 %

FIG. 21A

2100

	Número bruto de lecturas	Número de lecturas cartografiadas	Fragmentos no duplicados (solo autosomas)	Proporción de lecturas utilizables (%)	Profundidad por cadena	% de cobertura de CpG	Profundidad en los sitios CpG con lecturas cubiertas por cadena
Resumen							
Control 1	159554637	56323206	41653680	26 %	1,12	54,86 %	3,56
Control 2	118056735	49232079	45741385	39 %	1,23	60,26 %	3,24
Control 3	181913746	75753907	68081010	37 %	1,82	65,74 %	4,37
Control 4	142782870	50195168	42370350	30 %	1,13	59,15 %	3,14
C06	262763161	197344994	882660988	33 %	2,11	64,31 %	3,65
Capa leucocitaria	579329822	439663895	375814248	65 %	10,07	78,49 %	21,67
Tejido hepático no canceroso	549773334	417736311	355154964	65 %	9,51	84,89 %	18,21
Tejido tumoral	568004640	487494291	176985129	31 %	4,74	83,48 %	8,32
Plasma preoperatorio	796183998	637735722	474751039	60 %	12,72	87,18 %	18,69
Plasma postoperatorio	921833848	624851729	550104801	60 %	14,73	89,94 %	24,81

FIG. 21B

2150

Cr.	Densidad de metilación de ADN plasmático					
	Paciente con CHC		Control 1	Control 2	Control 3	Control 4
	Preoperatorio	Postoperatorio				
1	60,2 %	69,6 %	70,0 %	70,3 %	70,1 %	71,3 %
2	58,6 %	71,9 %	72,2 %	72,4 %	72,3 %	73,5 %
3	60,9 %	72,3 %	72,7 %	73,1 %	72,9 %	73,9 %
4	56,9 %	71,8 %	71,9 %	72,2 %	72,2 %	73,4 %
5	56,9 %	70,9 %	71,0 %	71,3 %	71,3 %	72,6 %
6	59,9 %	71,2 %	71,6 %	71,9 %	71,7 %	72,8 %
7	60,0 %	72,2 %	72,8 %	72,8 %	72,5 %	73,9 %
8	60,0 %	72,0 %	72,6 %	72,6 %	72,4 %	73,7 %
9	60,6 %	70,4 %	71,0 %	71,1 %	71,0 %	72,2 %
10	58,5 %	71,7 %	72,0 %	72,1 %	72,0 %	73,2 %
11	58,5 %	69,6 %	70,2 %	70,4 %	70,2 %	71,6 %
12	62,0 %	71,0 %	71,6 %	71,7 %	71,5 %	72,7 %
13	60,2 %	72,8 %	73,2 %	73,4 %	73,2 %	74,3 %
14	59,2 %	70,2 %	70,7 %	70,9 %	70,8 %	72,0 %
15	61,0 %	70,5 %	70,6 %	71,0 %	70,9 %	72,0 %
16	60,7 %	71,3 %	72,4 %	71,8 %	71,6 %	73,3 %
17	60,2 %	68,6 %	69,3 %	69,4 %	69,3 %	70,8 %
18	60,0 %	72,7 %	72,9 %	73,1 %	73,0 %	74,2 %
19	61,0 %	66,9 %	68,1 %	68,1 %	67,6 %	69,3 %
20	57,8 %	69,6 %	70,3 %	70,3 %	70,3 %	71,6 %
21	57,7 %	67,0 %	64,0 %	64,5 %	65,7 %	67,4 %
22	63,8 %	71,4 %	72,5 %	72,0 %	71,9 %	73,3 %
Todos los autosomas	59,7 %	70,8 %	71,2 %	71,3 %	71,2 %	72,5 %

FIG. 22

2200

ES 2 994 426 T3

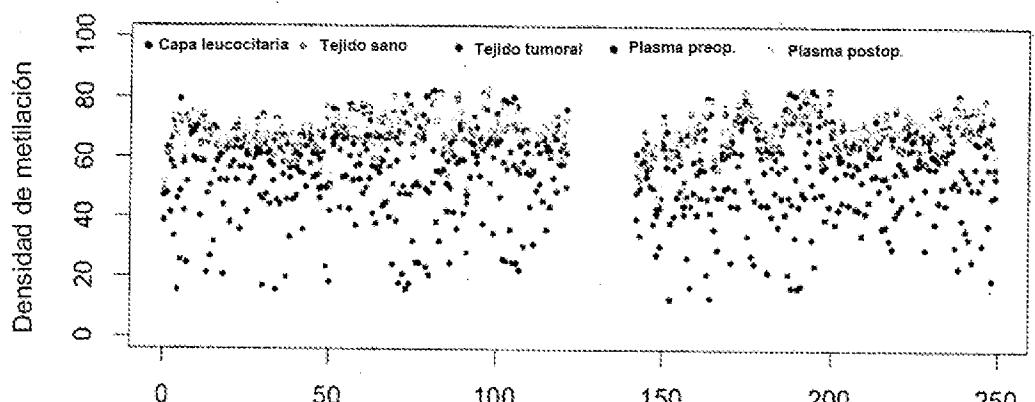


FIG. 23A

2300

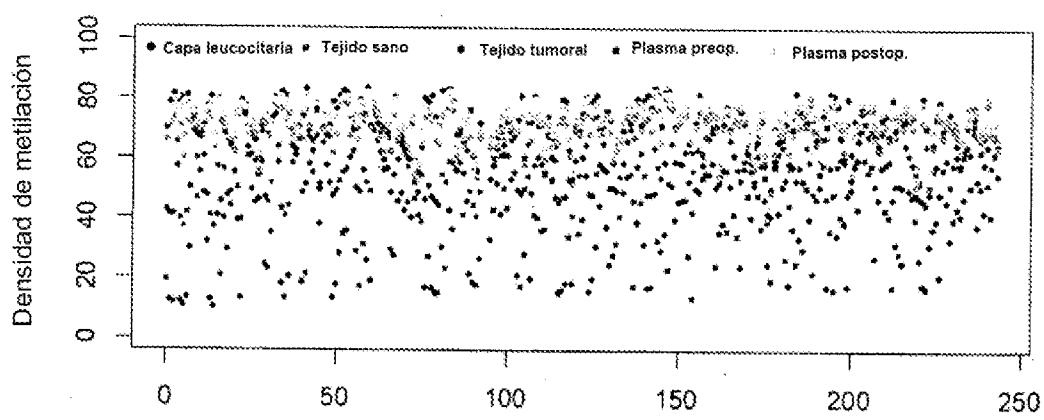


FIG. 23B

2350

ES 2 994 426 T3

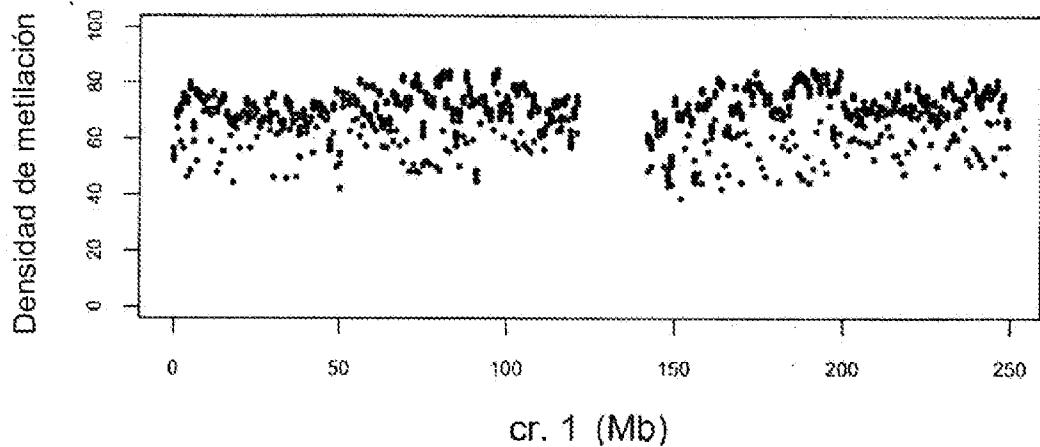


FIG. 24A

2400

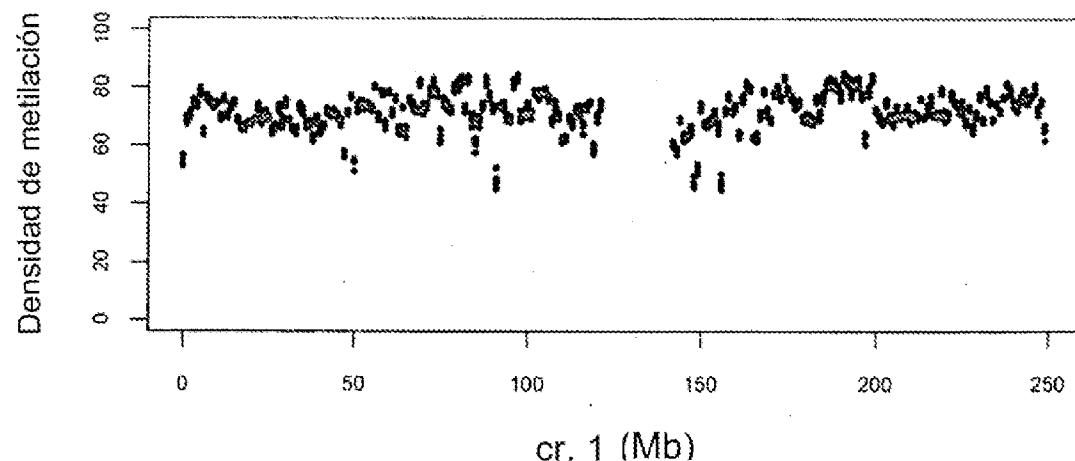


FIG. 24B

2450

ES 2 994 426 T3

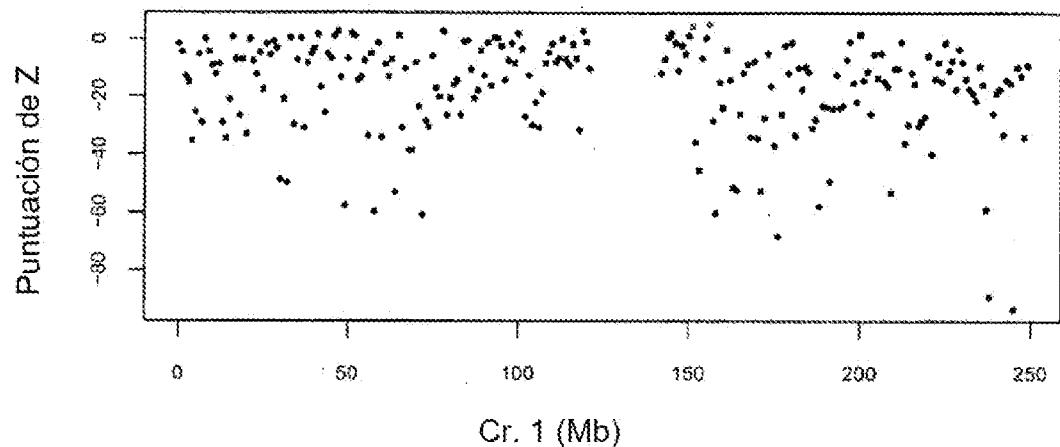


FIG. 25A

2500

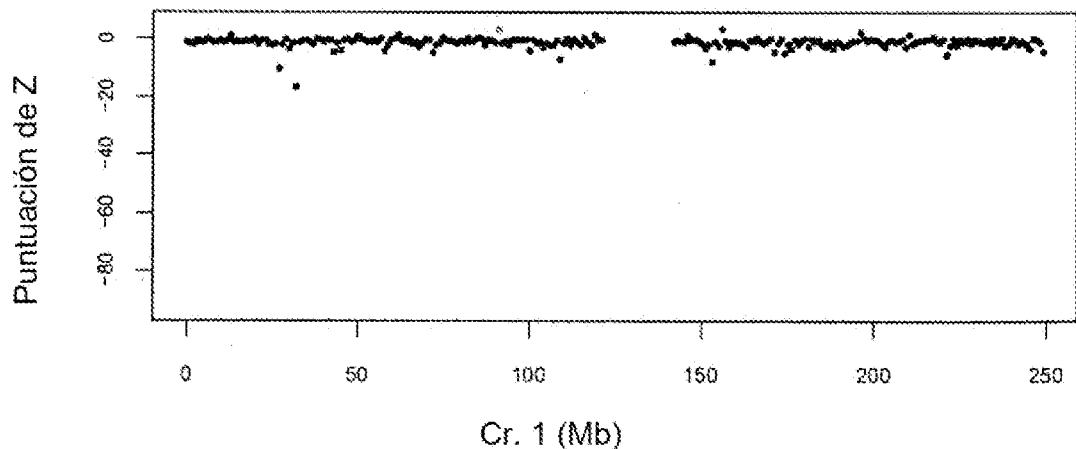


FIG. 25B

2550

puntuaciones de z	Plasma preoperatorio (n.º de bins)	Plasma postoperatorio (n.º de bins)
<-3	186 (80,9 %)	19 (8,3 %)
-3 to 3	42 (18,3%)	209 (90,3 %)
>3	2 (0,9 %)	2 (0,9 %)
Total	230 (100 %)	230 (100 %)

FIG. 26A

2600



FIG. 26B

2620

ES 2 994 426 T3

Puntuación de z	Plasma pre-operatorio (n.º de bins)	Plasma post-operatorio (n.º de bins)	C06	Control 1 (n.º de bins)	Control 2 (n.º de bins)	Control 3 (n.º de bins)	Control 4 (n.º de bins)
<-3	2330 (85,2 %)	171 (6,3 %)	29 (1,1 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
de -3 a 3	362 (13,3 %)	2556 (93,5 %)	2688 (98,3 %)	2734 (100 %)	2734 (100 %)	2734 (100 %)	2734 (100 %)
>3	42 (1,6 %)	7 (0,2 %)	17 (0,62 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
Total	2734 (100 %)	2734 (100 %)	2734 (100 %)	2734 (100 %)	2734 (100 %)	2734 (100 %)	2734 (100 %)

FIG. 26C

2640

	Proporción de CHH metilado	Proporción de CHG metilado	Proporción de citosinas metiladas	Proporción de CpG metilados (densidad de metilación)
Capa leucocitaria de paciente con CHC	0,13	0,14	5,74	67,08
Tejido hepático no tumoral	0,16	0,15	5,45	66,85
Tejido tumoral	0,09	0,09	3,29	42,88
Plasma preop.	0,09	0,10	3,99	59,67
Plasma postop.	1,25	0,96	6,52	70,75
Plasma de control 1	0,12	0,12	6,38	71,17
Plasma de control 2	0,10	0,10	5,62	71,33
Plasma de control 3	0,09	0,09	5,57	71,20
Plasma de control 4	0,10	0,10	5,93	72,48

FIG. 26D

2660

ES 2 994 426 T3

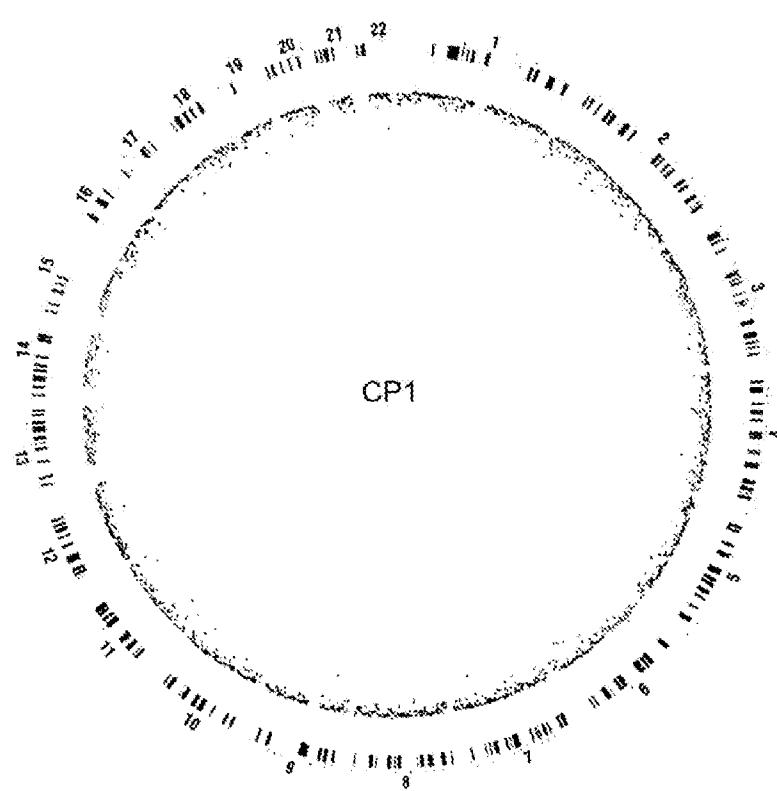


FIG. 27A

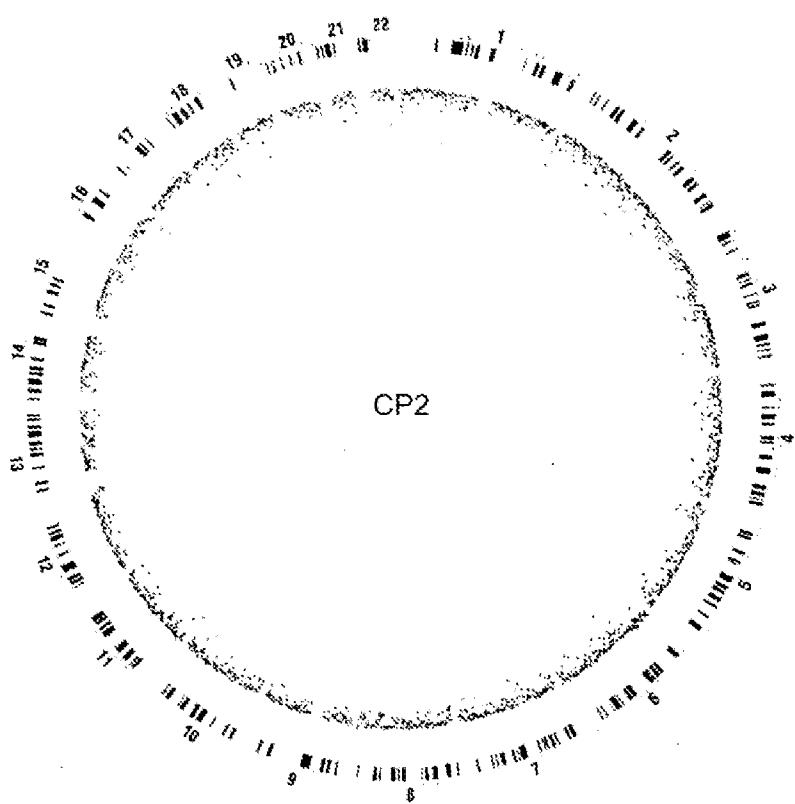


FIG. 27B

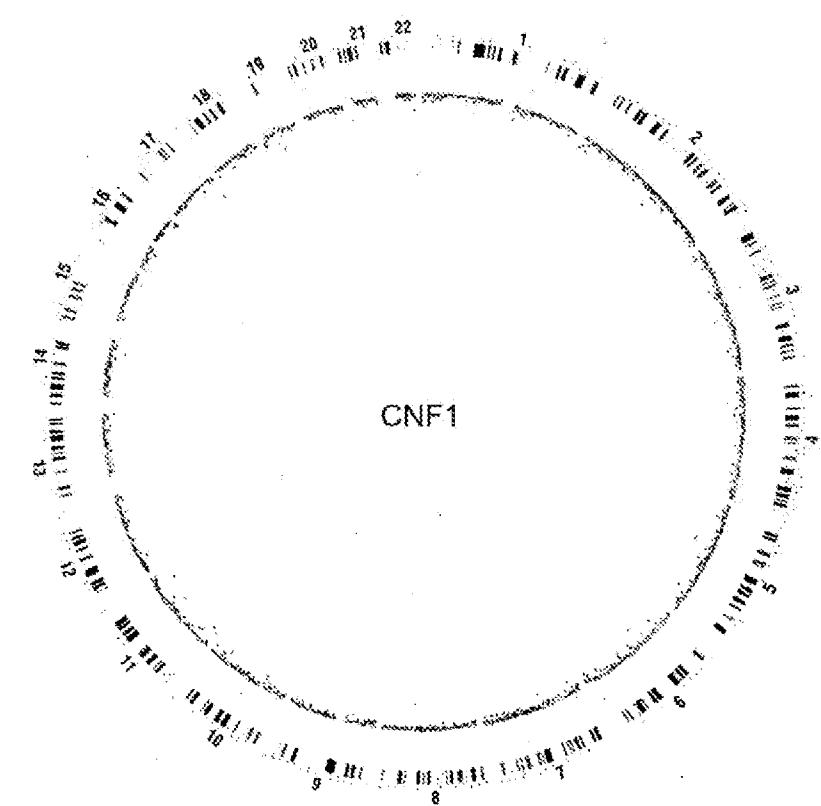


FIG. 27C

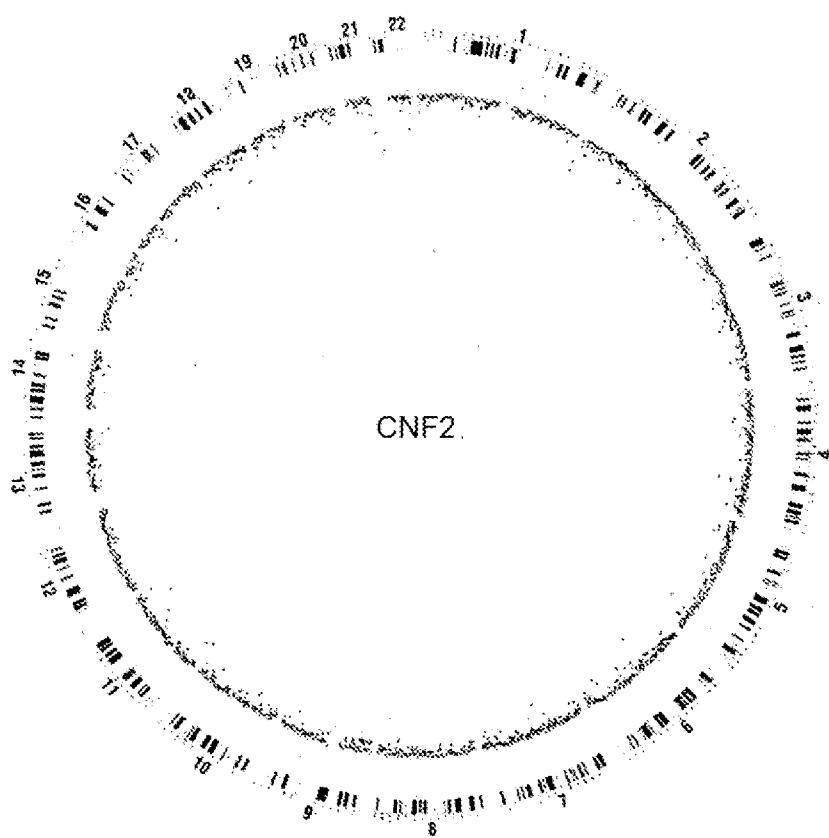


FIG. 27D

ES 2 994 426 T3

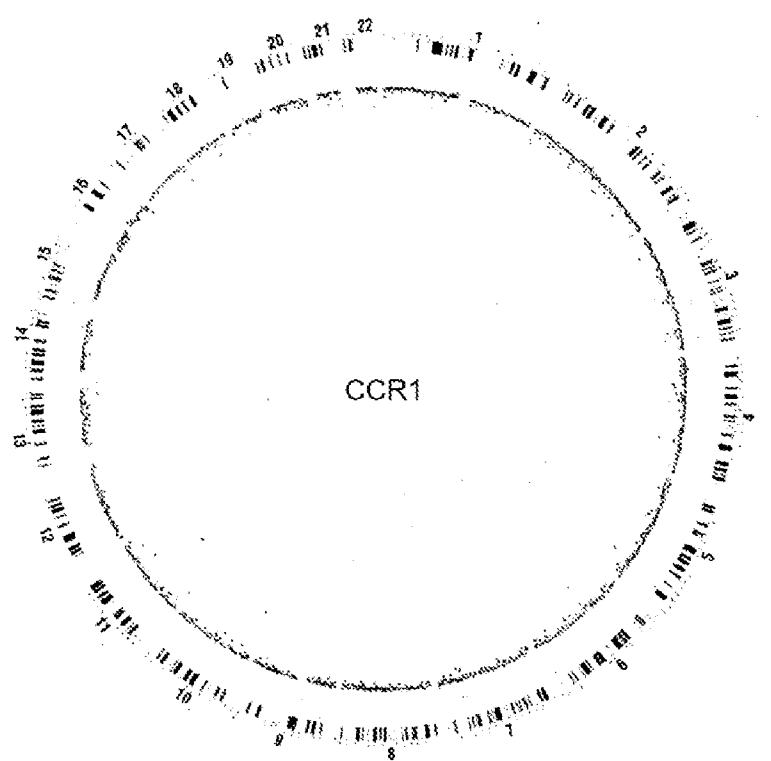


FIG. 27E

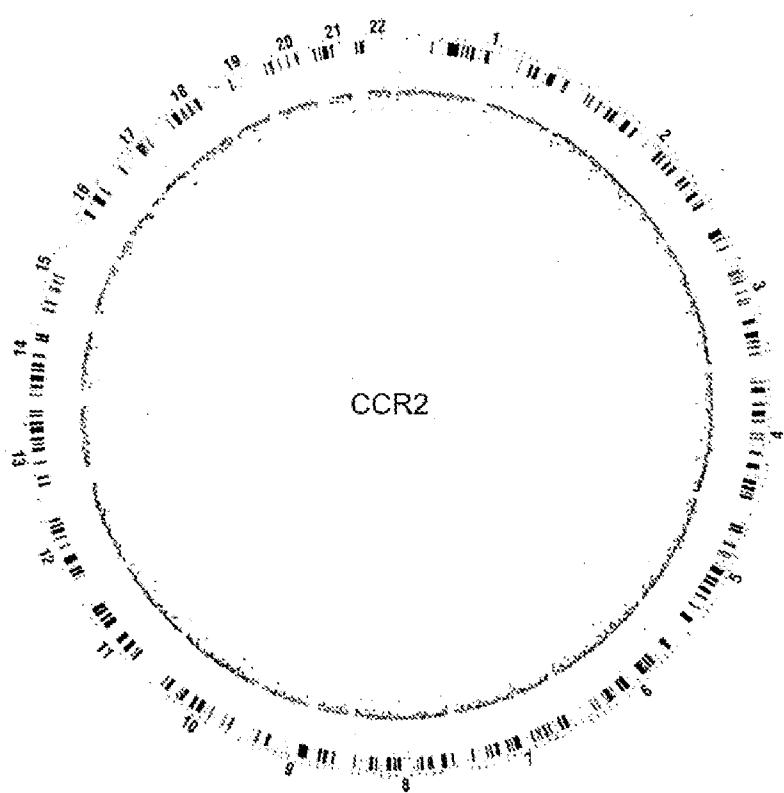


FIG. 27F

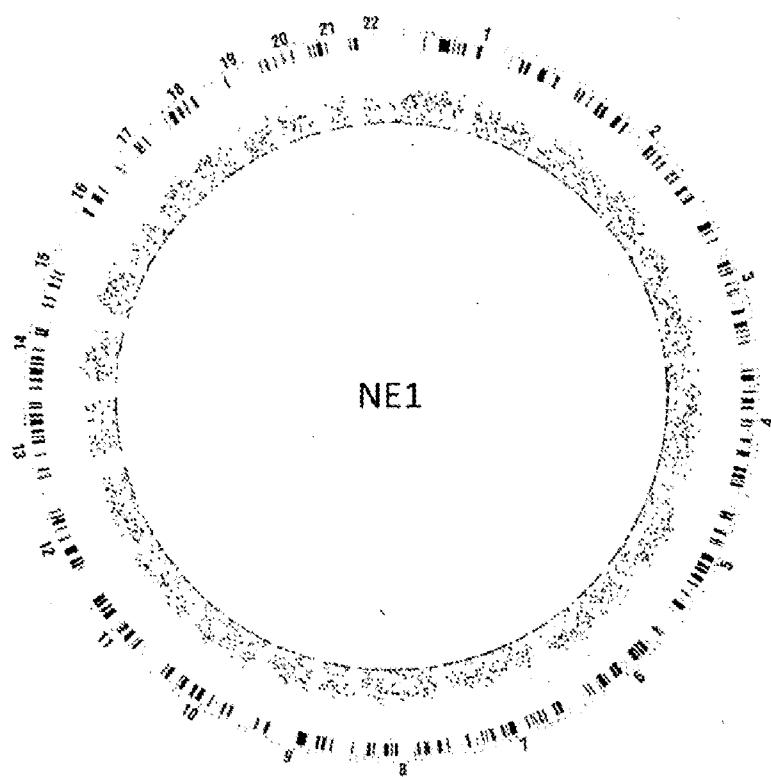


FIG. 27G

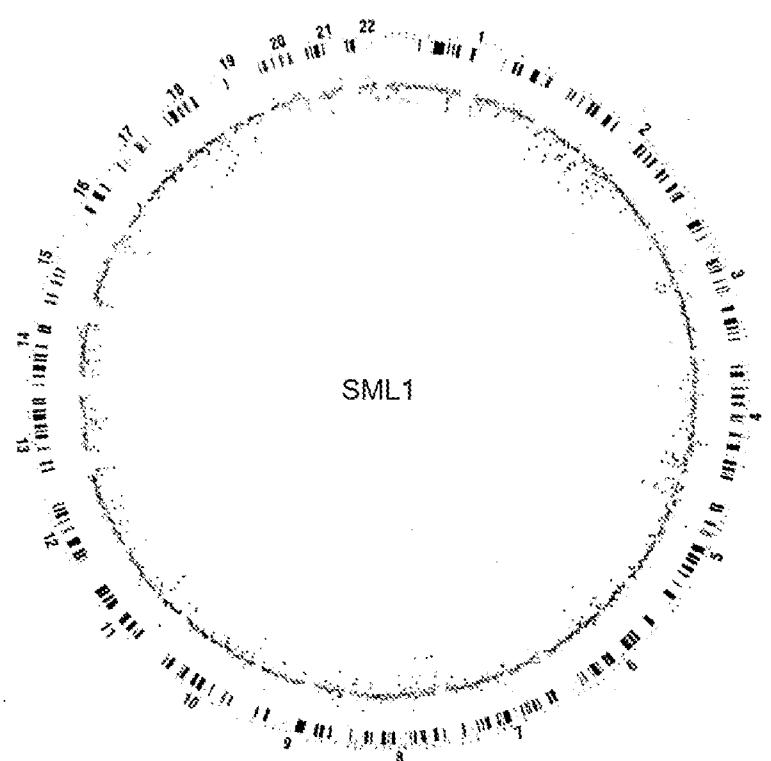


FIG. 27H

FIG. 27I

Número bruto de lecturas	Número de lecturas cartografiadas	Fragmentos no duplicados	Proporción de lecturas utilizables (%)	Profundidad por cadena	% de cobertura de CpG	Profundidad en los sitios CpG con las lecturas cubiertas por cadena
CP1 120807588	91750018	76890830	64 %	0,69	33,32 %	2,87
CP2 101344850	76473792	66245435	65 %	0,59	30,26 %	2,79
CNF1 101883643	75281330	57736188	56 %	0,51	27,87 %	2,72
CNF2 102745073	74211781	56841611	55 %	0,51	28,74 %	2,70
CCR1 87732645	66456695	56112543	64 %	0,50	27,38 %	2,63
CCR2 109955730	82594007	67253101	61 %	0,60	29,91 %	2,90
NE1 105514000	79923333	64534726	61 %	0,58	27,86 %	2,58
SML1 947277279	71658718	64426529	68 %	0,58	29,23 %	2,79

Puntuación z	CP1	CP2	CNF1	CNF2	CCR1	CCR2	NE1	SML1
< -3	34,9 %	33,8 %	17,6 %	54,3 %	13,5 %	9,8 %	98,1 %	11,2 %
de -3 a 3	65,0 %	65,9 %	82,4 %	45,6 %	86,5 %	90,2 %	1,7 %	70,8 %
>3	0,07 %	0,15 %	0,04 %	0 %	0,07 %	0 %	0,15 %	18,0 %

FIG. 27J

2790

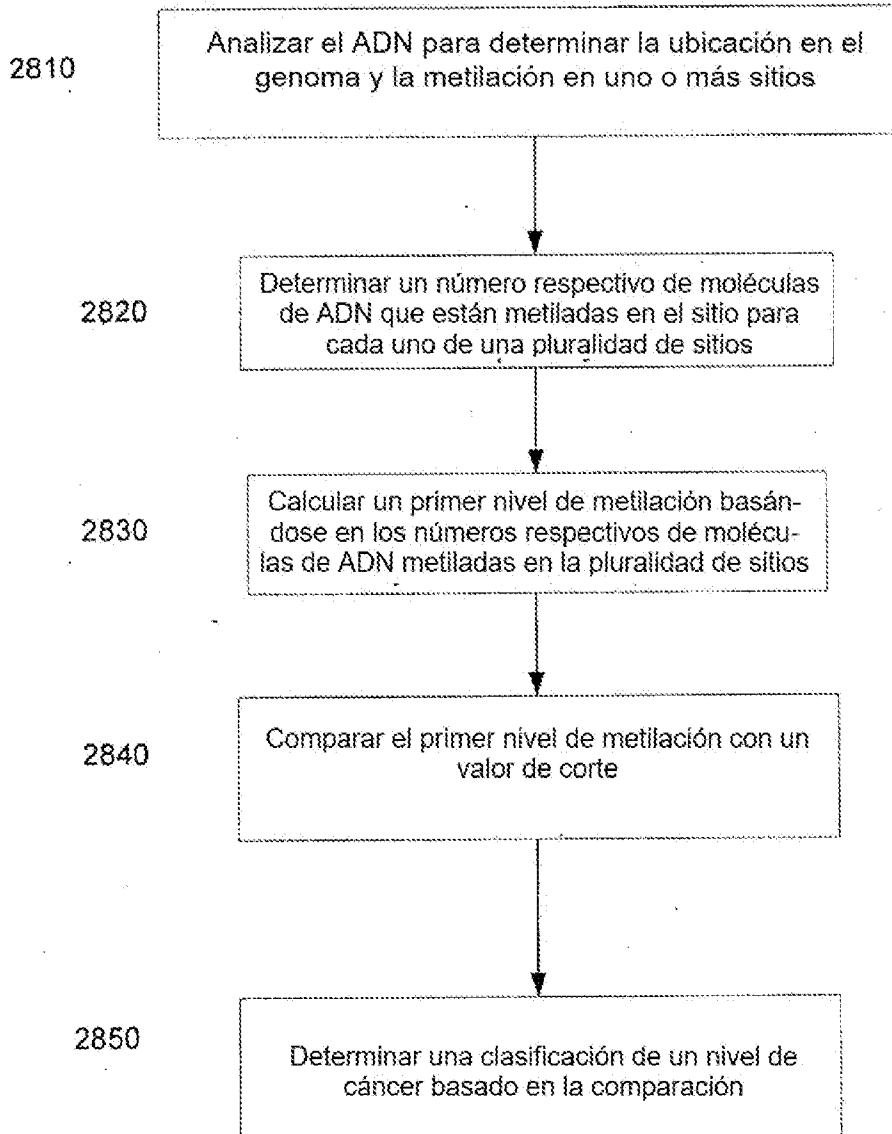


FIG. 28

2800

ES 2 994 426 T3

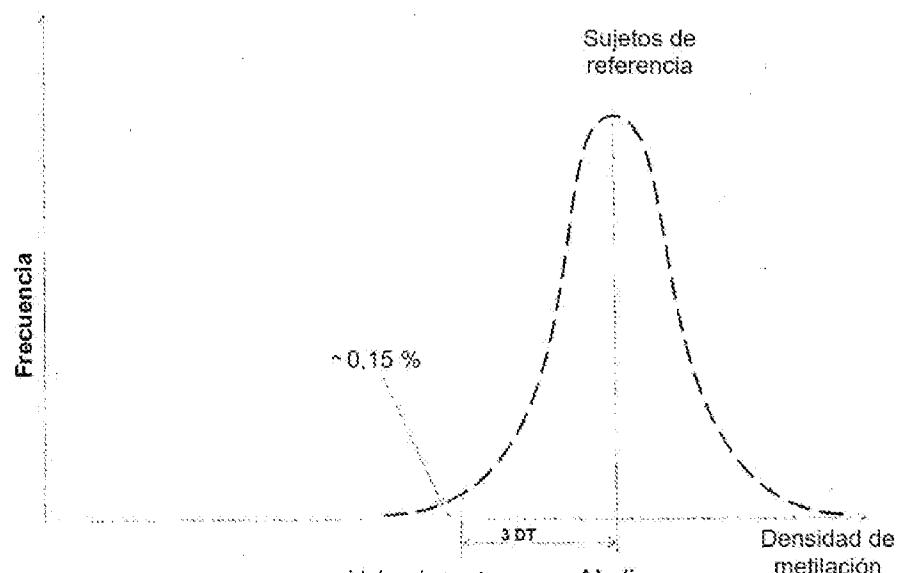


FIG. 29A

2900

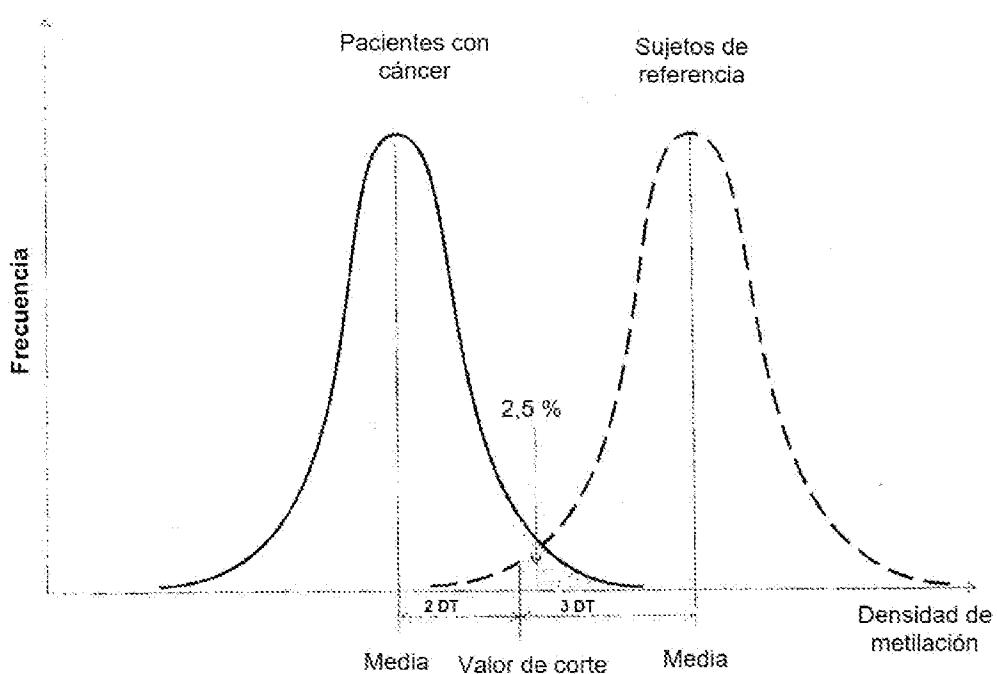


FIG. 29B

2950

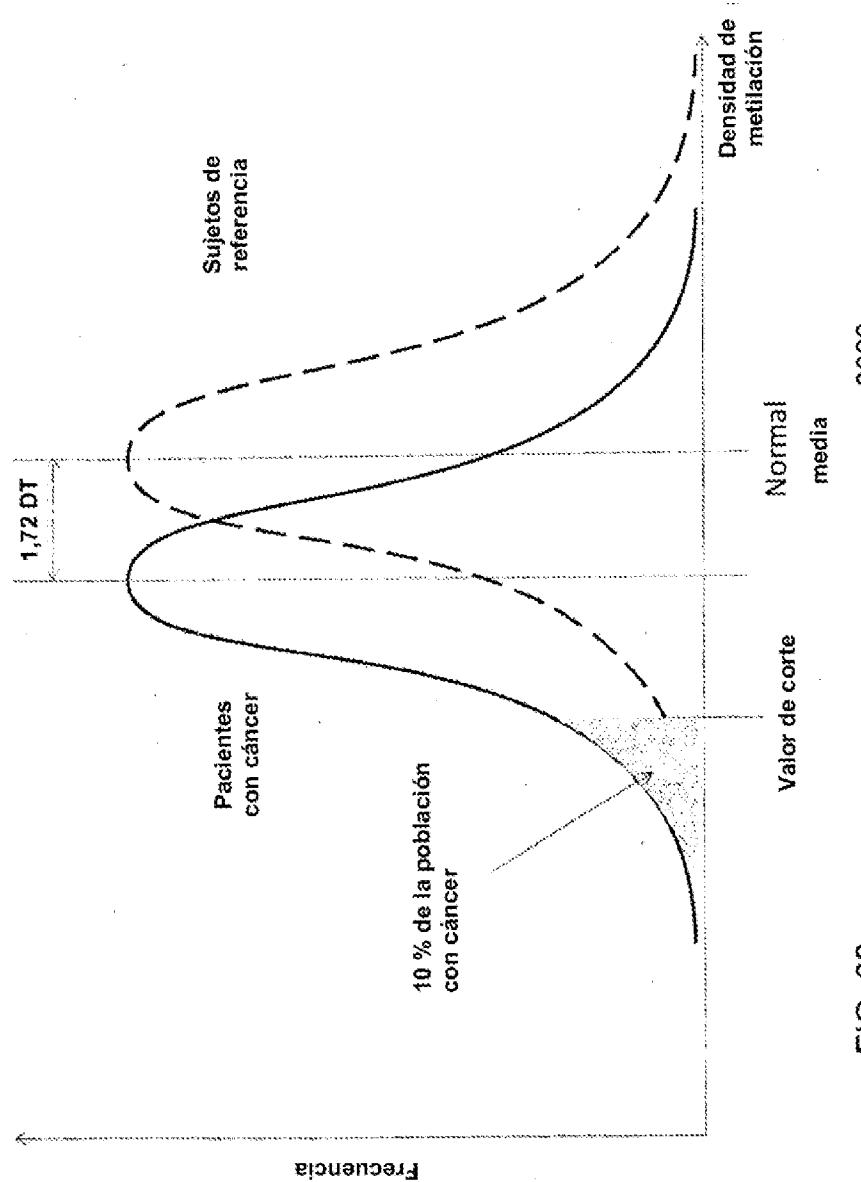
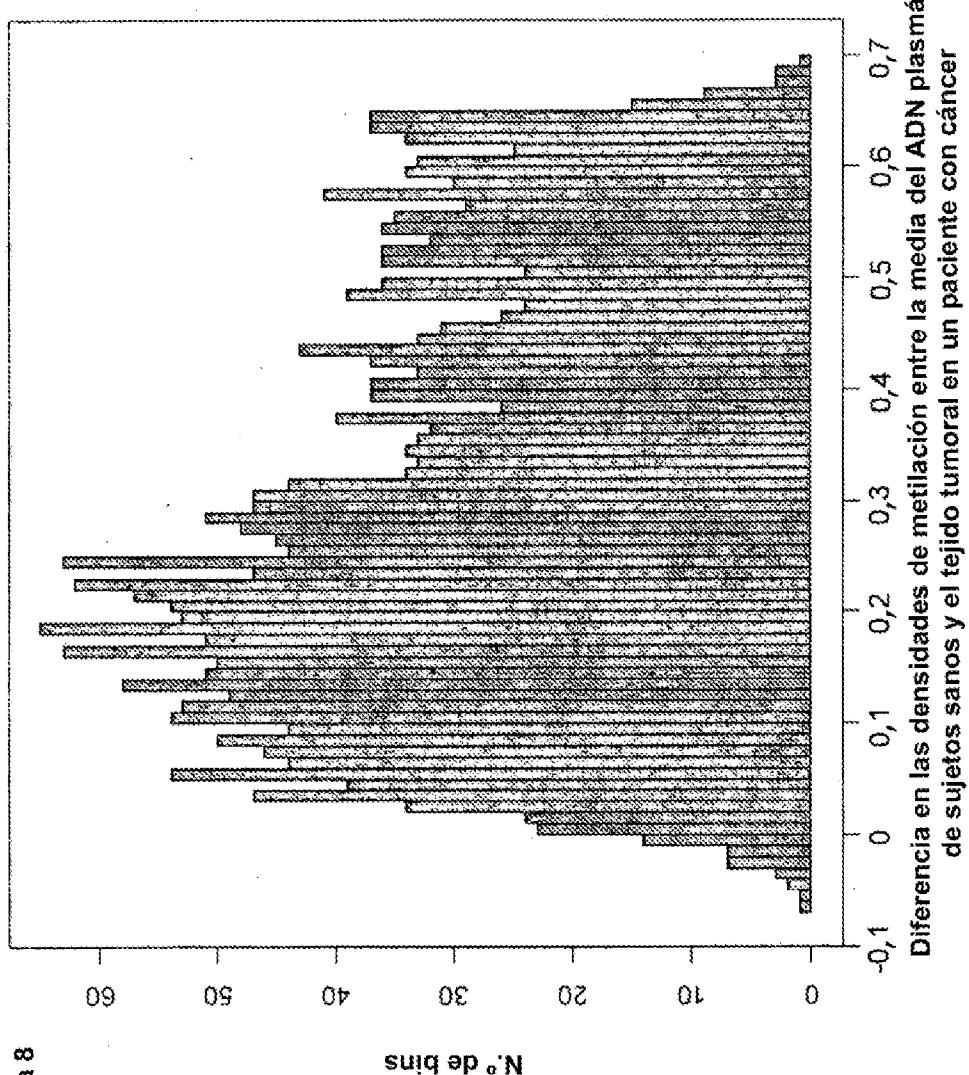


FIG. 30

Figura 8



Diferencia en las densidades de metilación entre la media del ADN plasmático de sujetos sanos y el tejido tumoral en un paciente con cáncer

FIG. 31
3100

Factor de dilución de la profundidad al 5 %	Profundidad de secuenciación (veces de genoma haploide)	N.º de bins hipometilados (puntuaciones z <-3) en plasma preoperatorio	% de hipometilados (puntuaciones z <-3) en plasma preoperatorio
0,8	1,76	1314	48,06
0,6	1,32	1323	48,39
0,4	0,88	1220	44,62
0,2	0,44	1094	40,01
0,1	0,22	900	32,92
0,05	0,11	645	23,59
0,01	0,022	207	7,57

Factor de dilución de la profundidad al 2 %	Profundidad de secuenciación (veces de genoma haploide)	N.º de bins hipometilados (puntuaciones z <-3) en plasma preoperatorio	% de hipometilados (puntuaciones z <-3) en plasma preoperatorio
0,8	1,76	491	17,96
0,6	1,32	493	18,03
0,4	0,88	460	16,83
0,2	0,44	439	16,08
0,1	0,22	358	13,02
0,05	0,11	297	10,86
0,01	0,022	231	8,45

FIG. 32A

3200

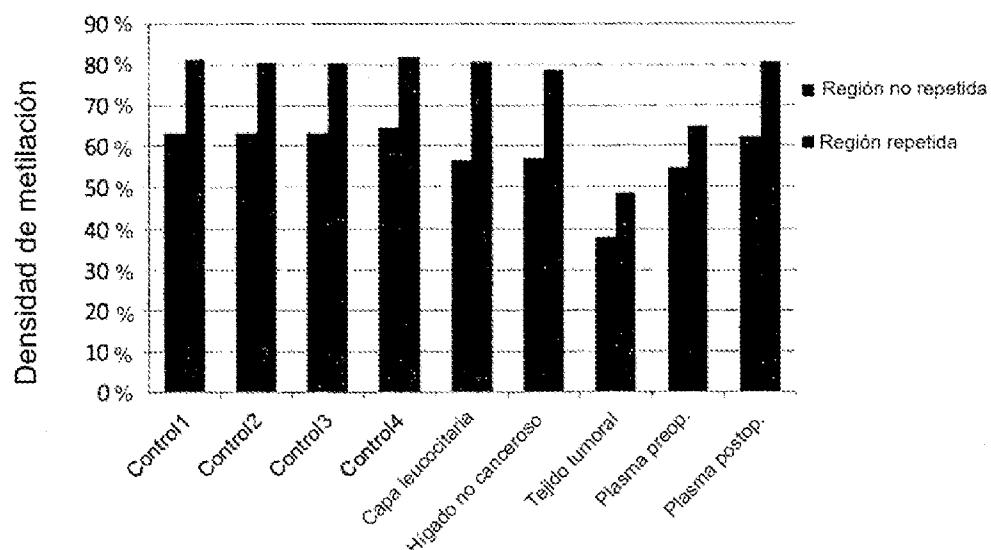


FIG. 32B

3250

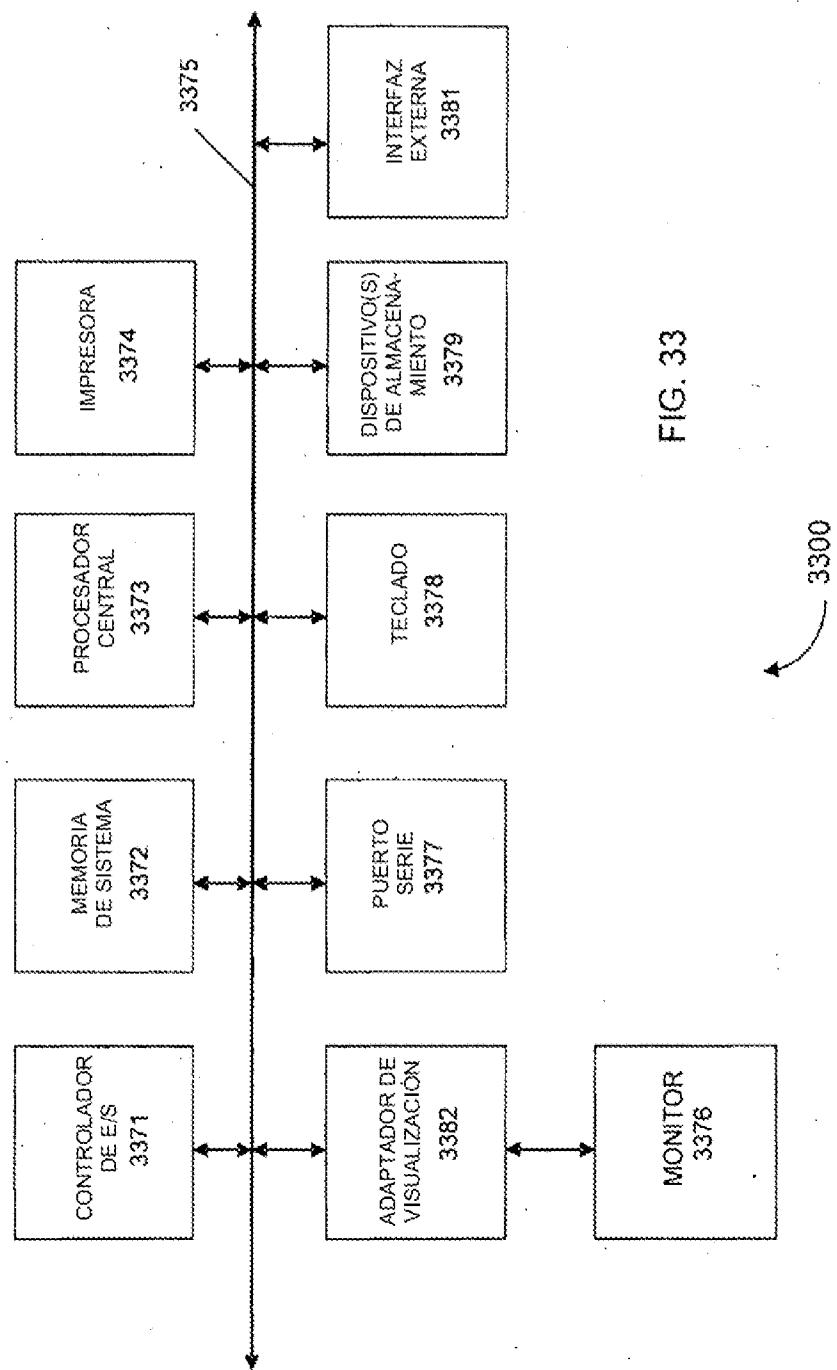


FIG. 33

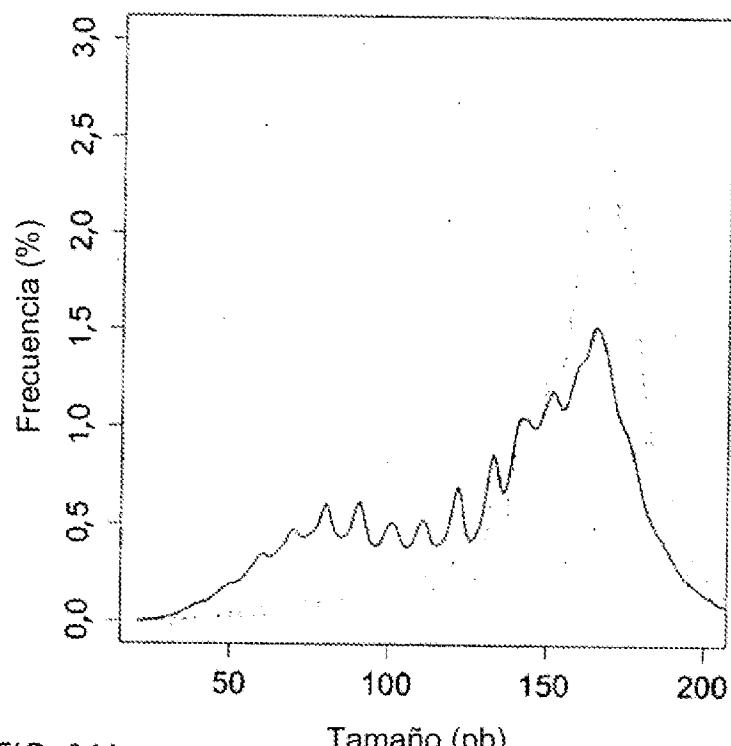


FIG. 34A

Tamaño (pb)

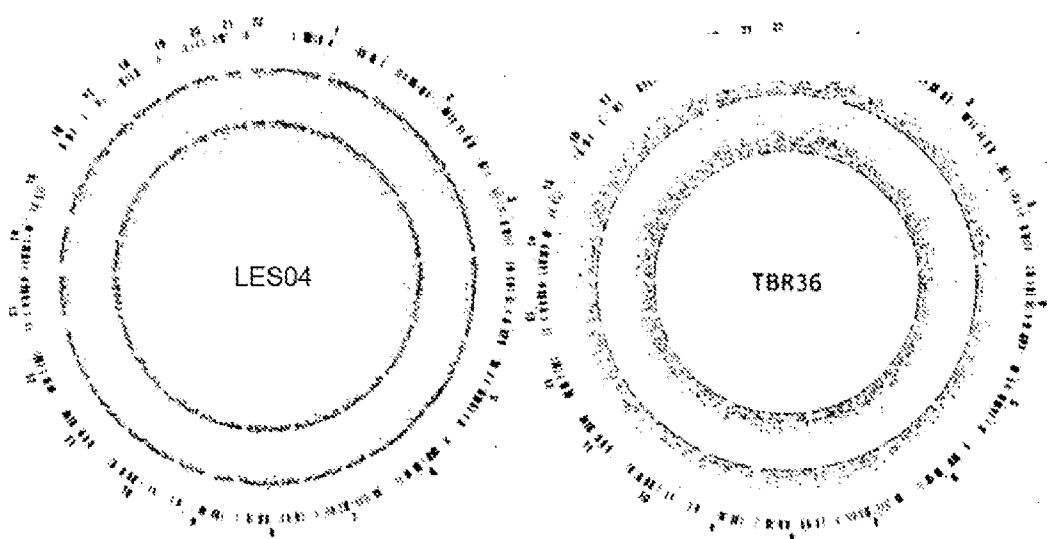


FIG. 34B

FIG. 34C

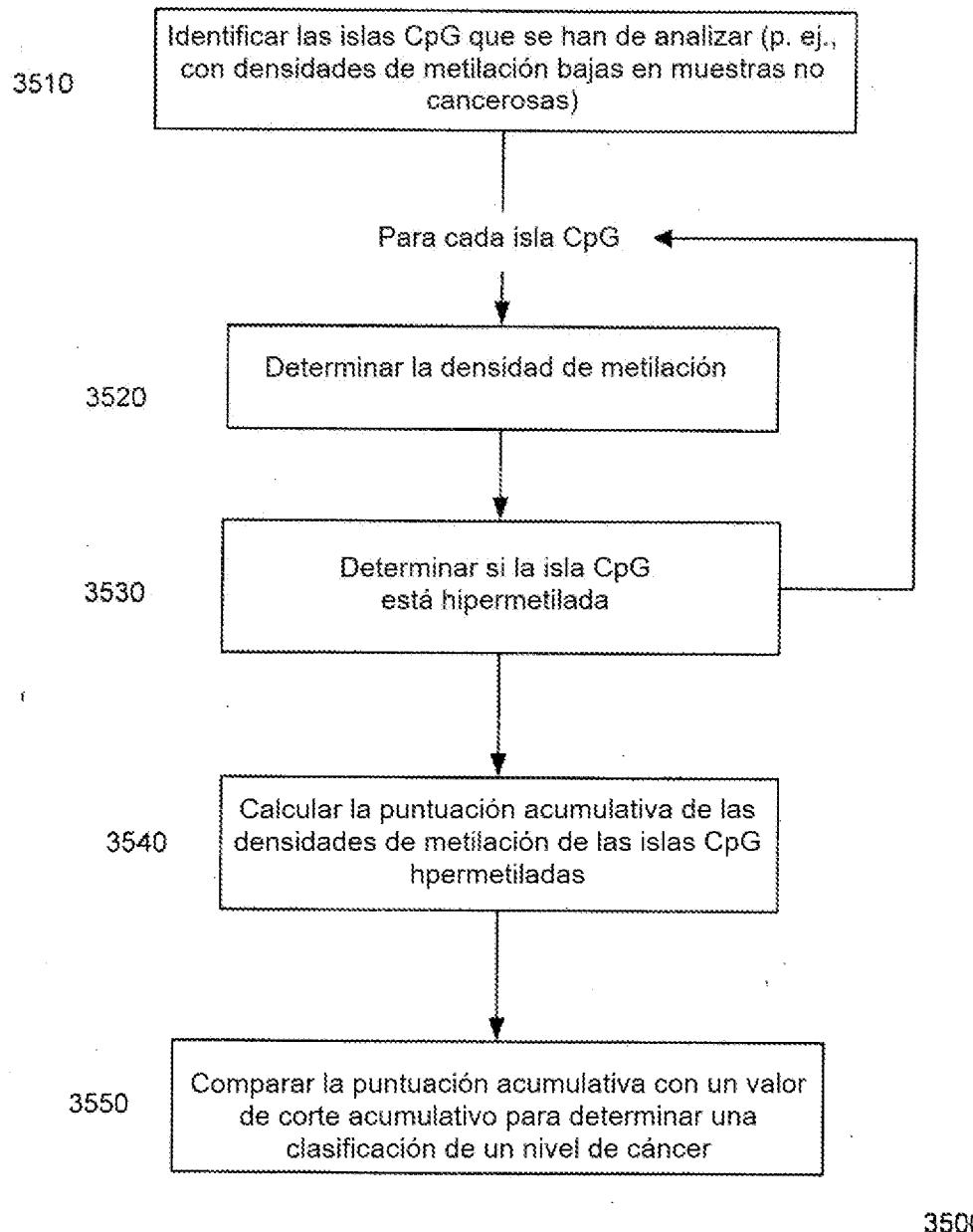


FIG. 35

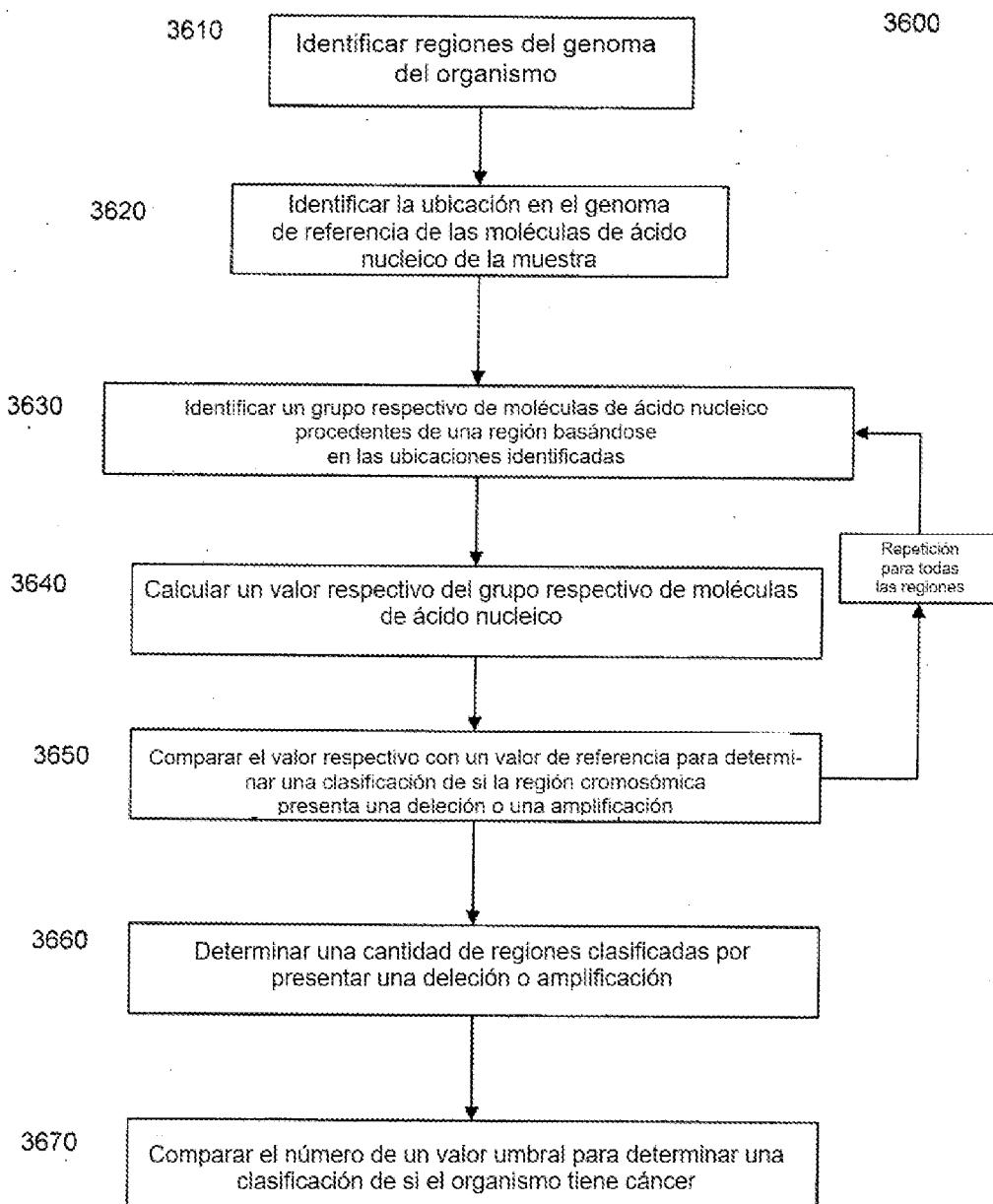


FIG. 36

ES 2 994 426 T3

Muestra: TBR36
Fracción de ADN tumoral: 40 %

De dentro hacia fuera:

- 1) ANC tumorales
- 2) ANC plasmáticas (sin BS)
- 3) ANC plasmáticas (BS)

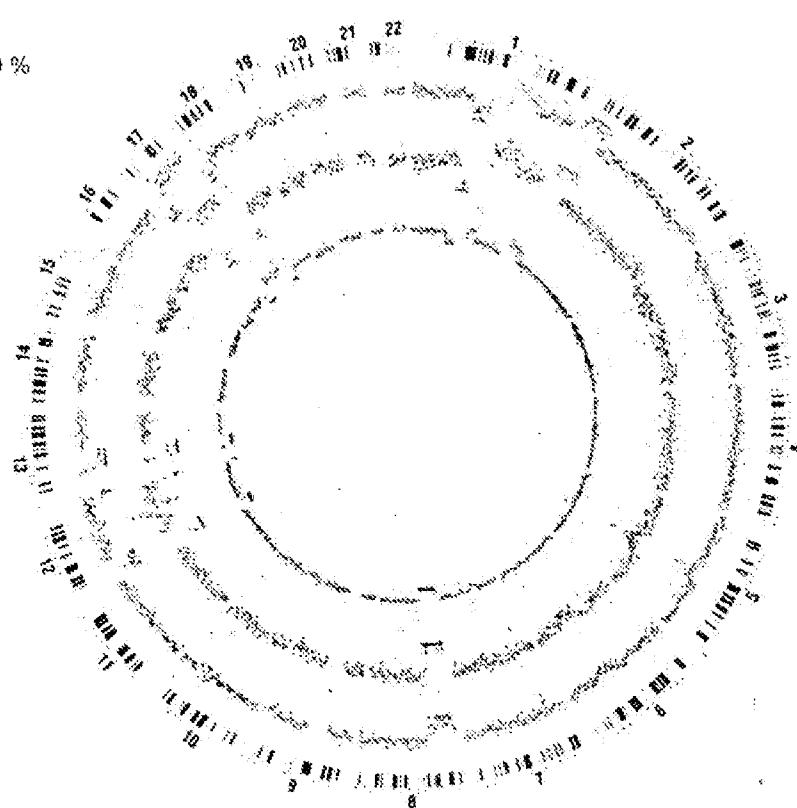


FIG. 37A

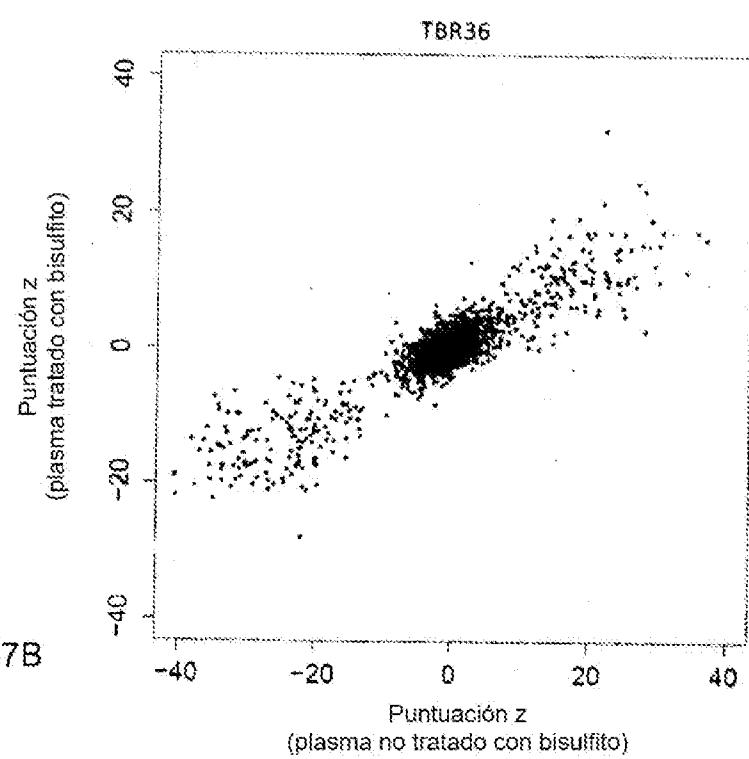


FIG. 37B

ES 2 994 426 T3

Muestra: TBR34

Fracción de ADN tumoral: 30 %

De dentro hacia fuera:

- 1) ANC tumorales
- 2) ANC plasmáticas (sin BS)
- 3) ANC plasmáticas (BS)

FIG. 38A

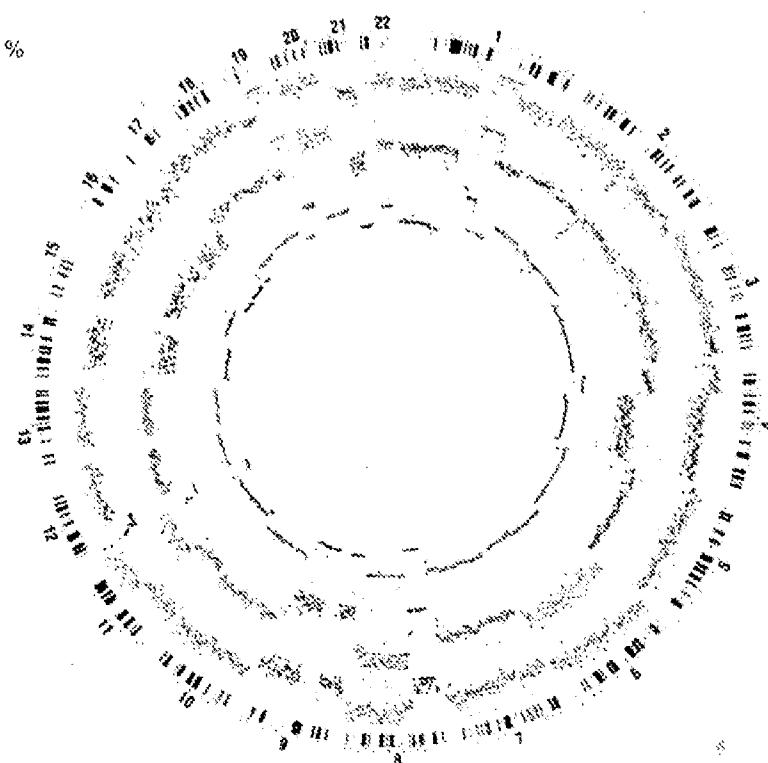
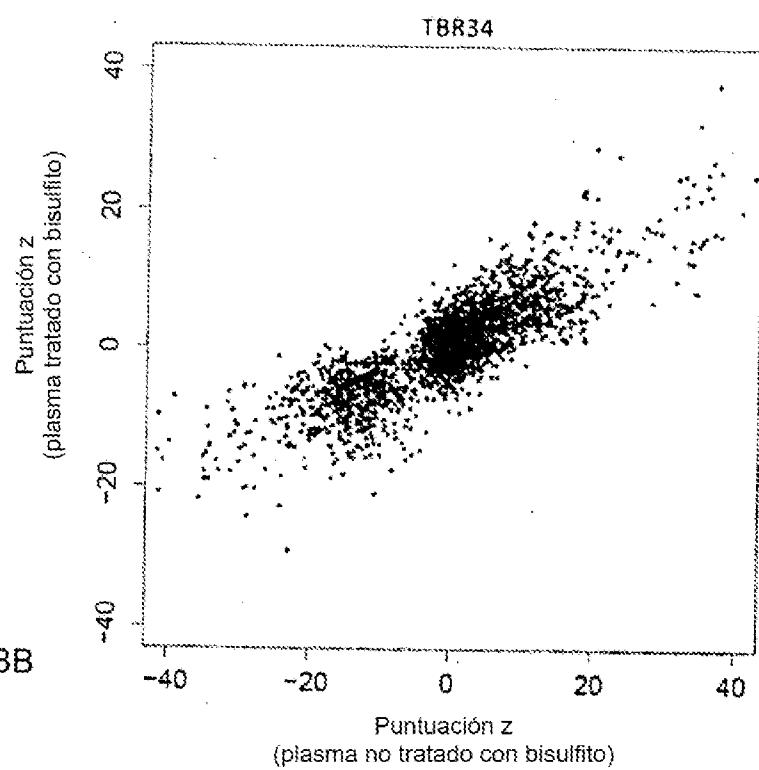


FIG. 38B



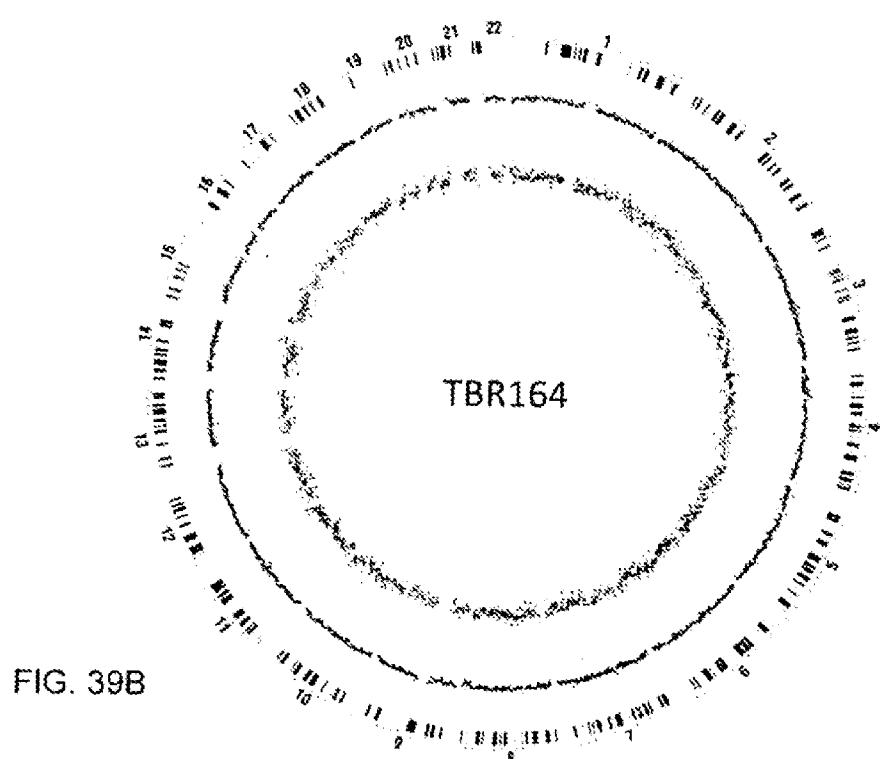
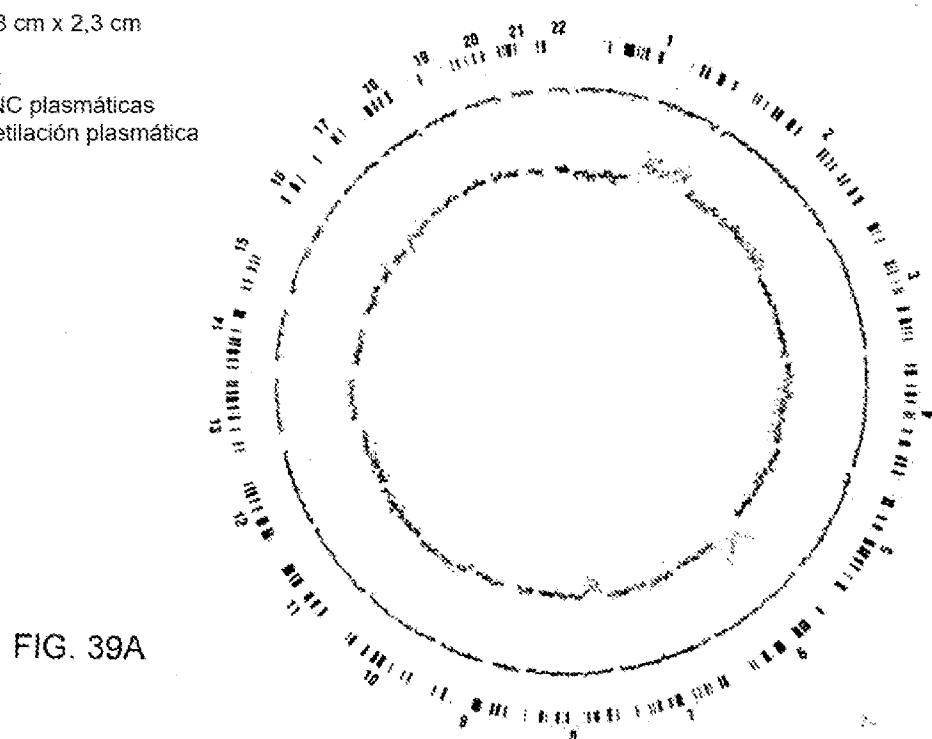
ES 2 994 426 T3

Muestra: TBR240

Tamaño del tumor: 2,8 cm x 2,3 cm

De dentro hacia fuera:

- 1) Puntuación z de ANC plasmáticas
- 2) Puntuación z de metilación plasmática



ES 2 994 426 T3

Anillo interior:
Plasma previo al tratamiento

Anillo exterior:
Plasma posterior al tratamiento

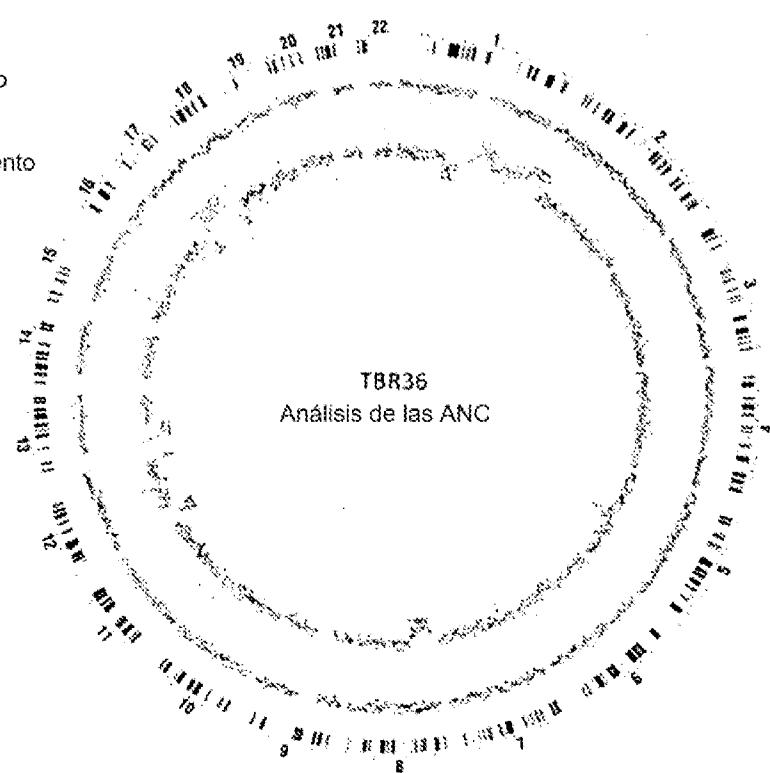


FIG. 40A

Anillo interior:
Plasma previo al tratamiento

Anillo exterior:
Plasma posterior al tratamiento

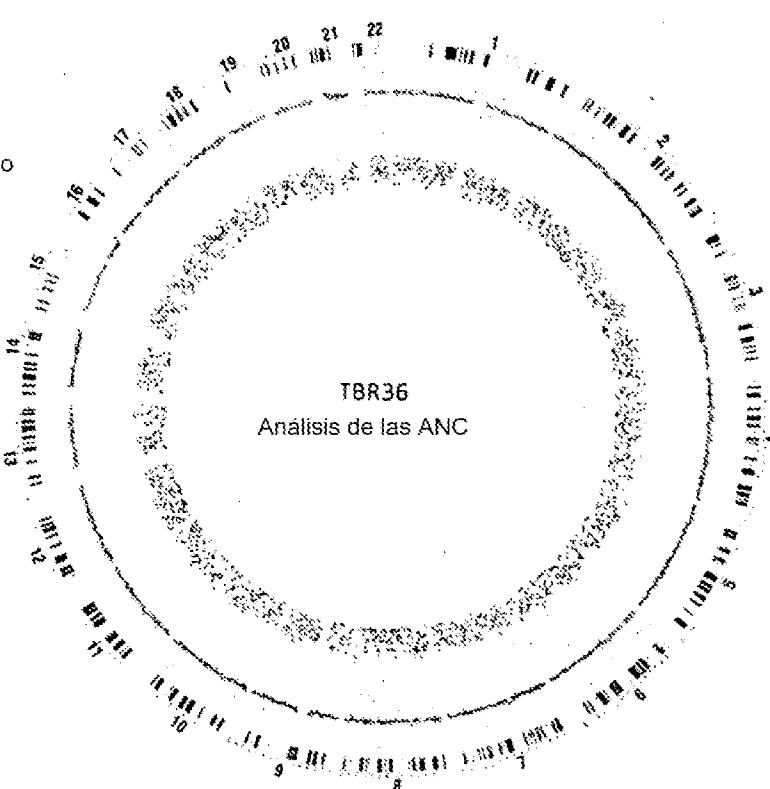


FIG. 40B

Muestra: TBR34
Fracción de ADN tumoral: 30 %
De fuera hacia dentro:
1) CNV postop.
2) CNV preop.

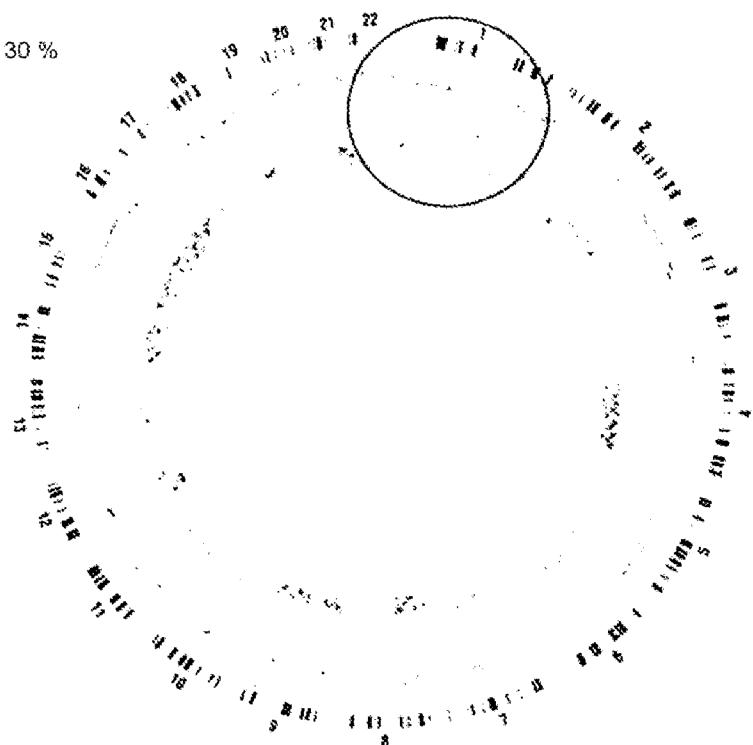


FIG. 41A

Anillo interior:
Plasma previo al tratamiento

Anillo exterior:
Plasma posterior al tratamiento

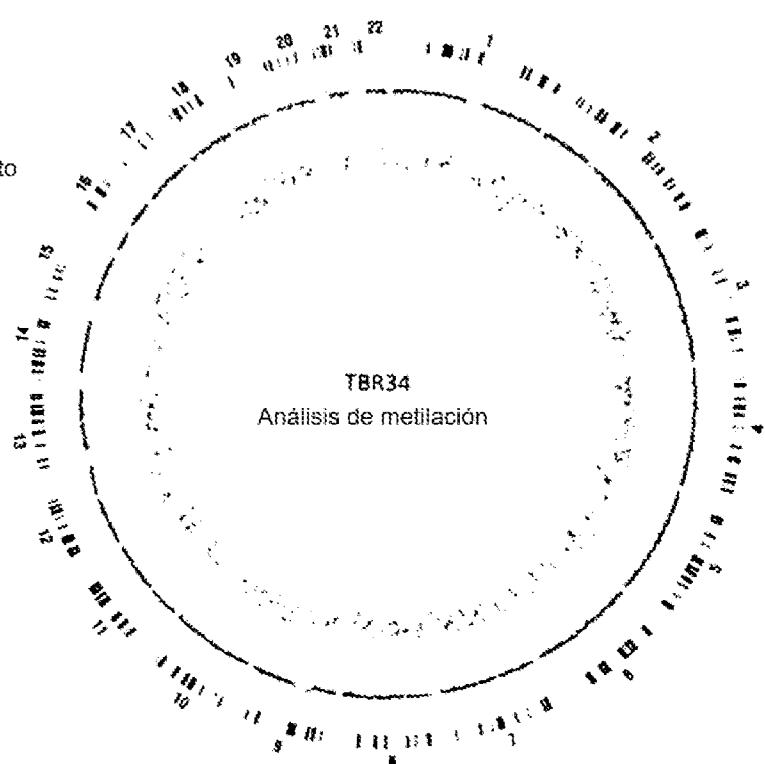


FIG. 41B

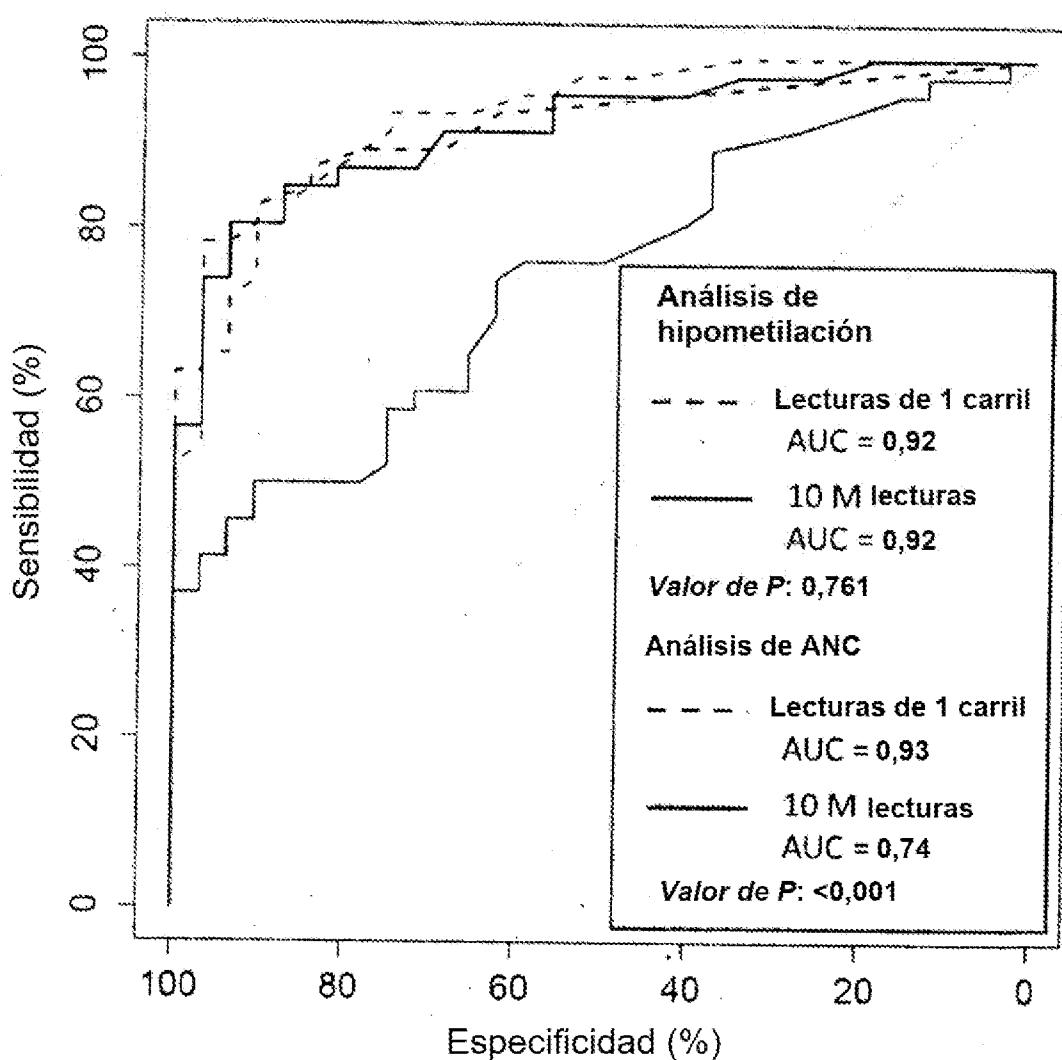


FIG. 42

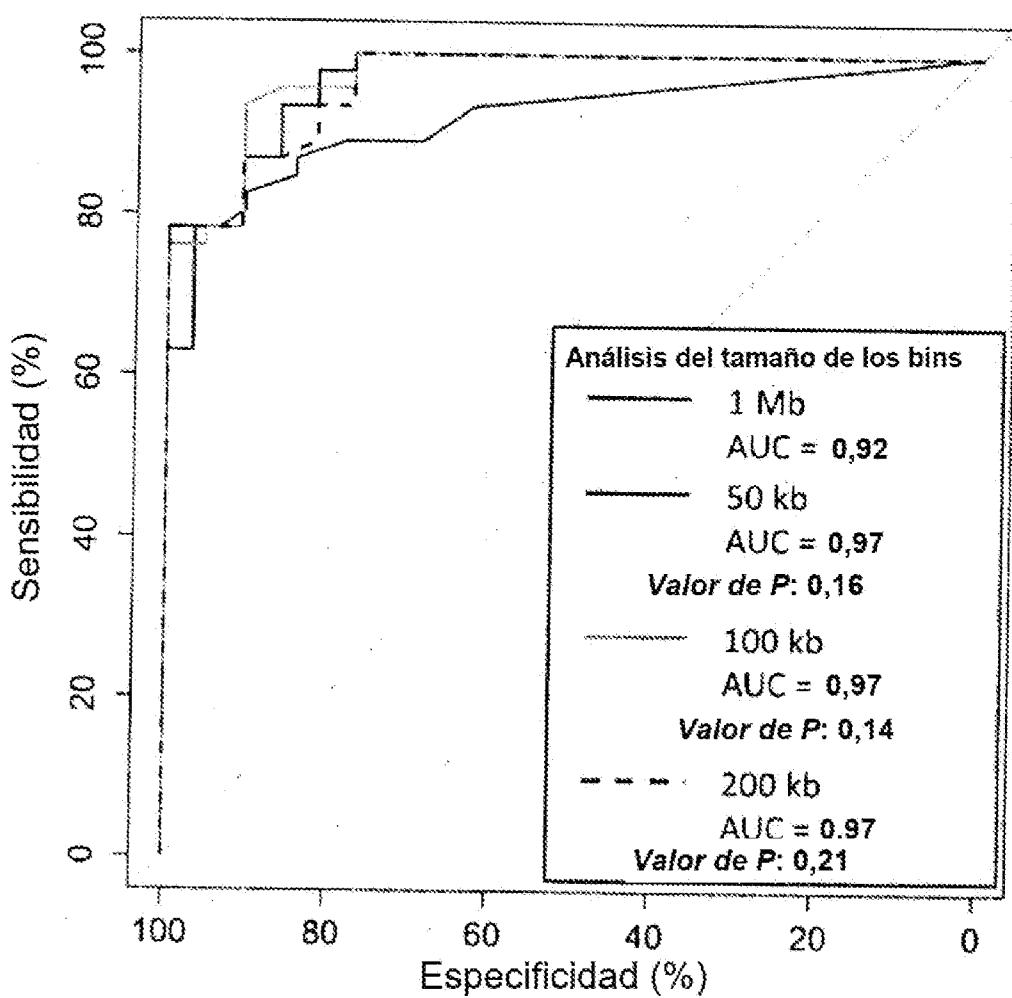


FIG. 43

FIG. 44A

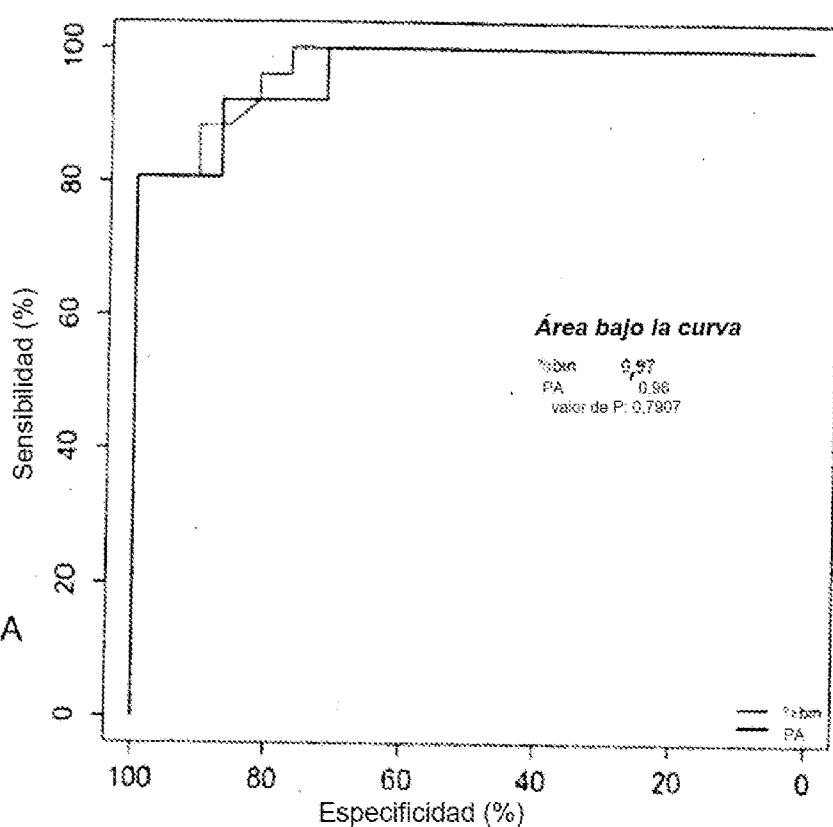
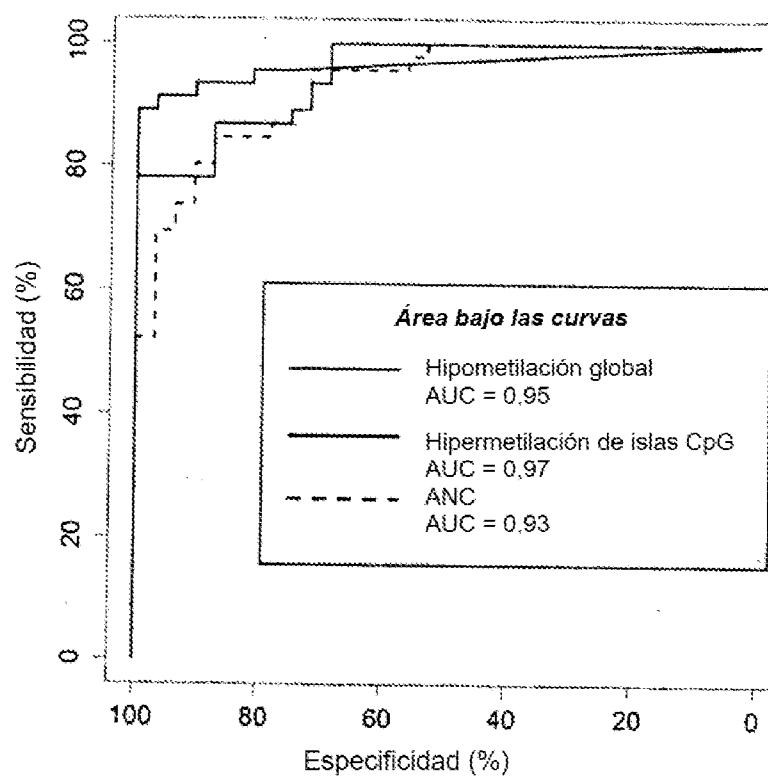


FIG. 44B



N.º de caso	Estadio del SCLC	Dimensión más larga del (de los) tumor(es) (cm)	Puntuación de PA			Presencia de cualquier tipo de cambio
			Hipometilación general	Hipermetilación de islas CpG	ANC	
TBR36	A1	16	<u>88585</u>	<u>80,9</u>	<u>3777</u>	Positiva
TBRG4	A1	13	<u>37573</u>	<u>221,7</u>	<u>8650</u>	Positiva
HOT198	A1	12	<u>93132</u>	<u>52,5</u>	<u>6803</u>	Positiva
HOT205	A1	6	<u>1146</u>	<u>1,7</u>	<u>175</u>	Positiva
HOT192	A1	5,5	<u>25914</u>	<u>119,5</u>	<u>1668</u>	Positiva
HOT227	A1	5,5	<u>36333</u>	<u>155,4</u>	<u>5008</u>	Positiva
HOT238	A1	5,5	<u>12187</u>	<u>105,3</u>	<u>1978</u>	Positiva
HOT197	B	5,3	<u>1714</u>	<u>9,5</u>	<u>28</u>	Positiva
HOT156	A1	5	<u>16529</u>	<u>115,9</u>	<u>1947</u>	Positiva
HOT236	A1	3,7	<u>7993</u>	<u>5,9</u>	<u>85</u>	Positiva
HOT170	A1	3,5	<u>22271</u>	<u>81,9</u>	<u>1542</u>	Positiva
HOT229	A1	3,1	<u>383</u>	<u>5,4</u>	<u>71</u>	Positiva
HOT233	A1	3	<u>189</u>	<u>7,5</u>	<u>322</u>	Positiva
HOT240	A1	2,8	<u>1277</u>	<u>14,0</u>	<u>972</u>	Positiva
HOT222	A1	2,6	<u>18832</u>	<u>10,5</u>	<u>172</u>	Positiva
HOT162	A1	2,5	<u>26112</u>	<u>41,0</u>	<u>3048</u>	Positiva
HOT172	A1	2,5	<u>12909</u>	<u>98,1</u>	<u>1305</u>	Positiva
HOT164	A1	2,3	<u>12604</u>	<u>4,8</u>	<u>210</u>	Positiva
HOT215	A1	2,3	<u>25824</u>	<u>54,8</u>	<u>714</u>	Positiva
HOT207	A1	2,1	<u>3377</u>	<u>12,0</u>	<u>101</u>	Positiva
HOT224	A1	2	<u>367</u>	<u>0,0</u>	<u>13</u>	Negativa
HOT159	A1	1,5	<u>1367</u>	<u>1,4</u>	<u>26</u>	Positiva
HOT204	A1	1,5	<u>102</u>	<u>4,5</u>	<u>198</u>	Positiva
HOT151	A1	1,5	<u>153</u>	<u>2,8</u>	<u>9</u>	Negativa
HOT208	A1	1,2	<u>5867</u>	<u>4,7</u>	<u>214</u>	Positiva
HOT167	A1	1	<u>21544</u>	<u>4,5</u>	<u>54</u>	Positiva

FIG. 45

Tipo de cáncer	N.º de caso	Estadificación/ Dimensión máx. del tumor	Puntuación de PA			Presencia de cualquier tipo de cambio
			Hipometilación general	Hipermetilación de islas CpG	ANC	
Cáncer de mama	TBR150	Metastásico	<u>1481</u>	<u>146,7</u>	<u>5123</u>	Positiva
	TBR126	T2N1M0 4,5 cm	<u>10415</u>	<u>34,9</u>	<u>1951</u>	Positiva
	TBR127	T2N1M0 2,5 cm	<u>2598</u>	<u>4,3</u>	54	Positiva
	TBR128	T2N0M0 2 cm	157	<u>7,7</u>	<u>2250</u>	Positiva
	TBR111	T1N0M0 0,9 cm	298	0,0	128	Negativa
Sarcoma del músculo liso	TBR051	Metastásico	<u>4787</u>	<u>79,8</u>	<u>7986</u>	Positiva
Tumor neuroendocrino	TBR032	Metastásico	<u>112518</u>	<u>119,8</u>	<u>14681</u>	Positiva
Cáncer de pulmón	TBR164	Metastásico	<u>51818</u>	<u>47,3</u>	<u>4714</u>	Positiva
	TBR012	Metastásico	<u>4351</u>	<u>7,7</u>	<u>2213</u>	Positiva
	TBR014	Metastásico	<u>2356</u>	<u>12,2</u>	<u>825</u>	Positiva
	TBR177	T2N1M0 3,9 cm	<u>1072</u>	<u>10,6</u>	25	Positiva
Cáncer nasofaringeo	TBR031	Metastásico	<u>1028</u>	<u>4,4</u>	<u>1860</u>	Positiva
	TBR125	Metastásico	<u>20949</u>	<u>162,5</u>	<u>6708</u>	Positiva
	TBR124	T4N1	270	<u>4,9</u>	37	Positiva
	TBR123	T1N2M0	<u>6871</u>	<u>11,8</u>	80	Positiva
	TBR062	T3N1M0	111	<u>30,9</u>	41	Positiva
	TBR108	T3N1M0	93	<u>7,7</u>	26	Positiva
	TBR107	T3N0M0	<u>2439</u>	<u>41,2</u>	205	Positiva
	TBR106	T1N1M0	<u>2281</u>	<u>5,9</u>	97	Positiva
	TBR099	T1N0M0	<u>6167</u>	<u>10,2</u>	26	Positiva

FIG. 46

ES 2 994 426 T3

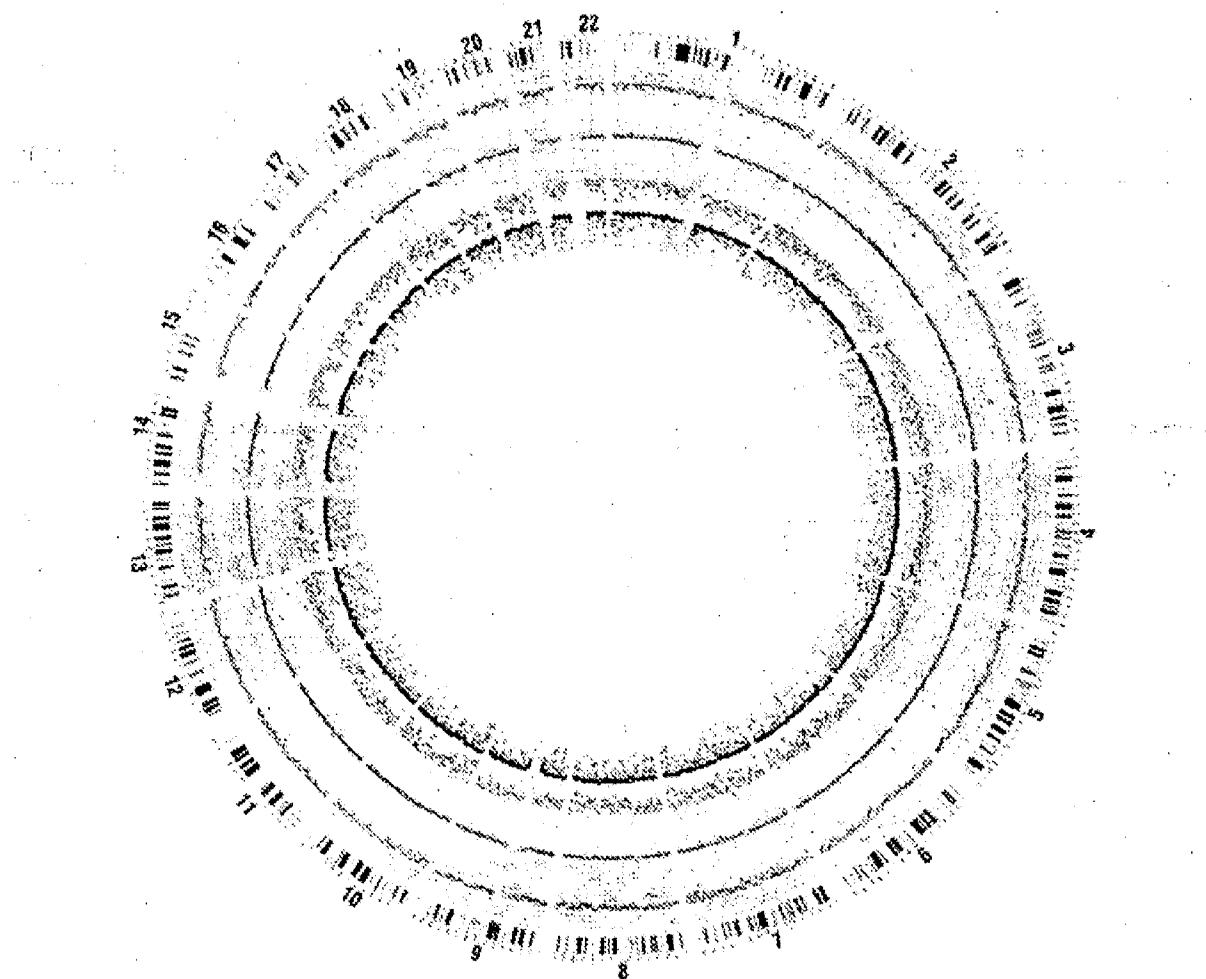


FIG. 47

ES 2 994 426 T3

Muestra: TBR36
Fracción de ADN tumoral: 40 %

De dentro hacia fuera:
1) Puntuación z de ANC
2) Puntuación z de metilación

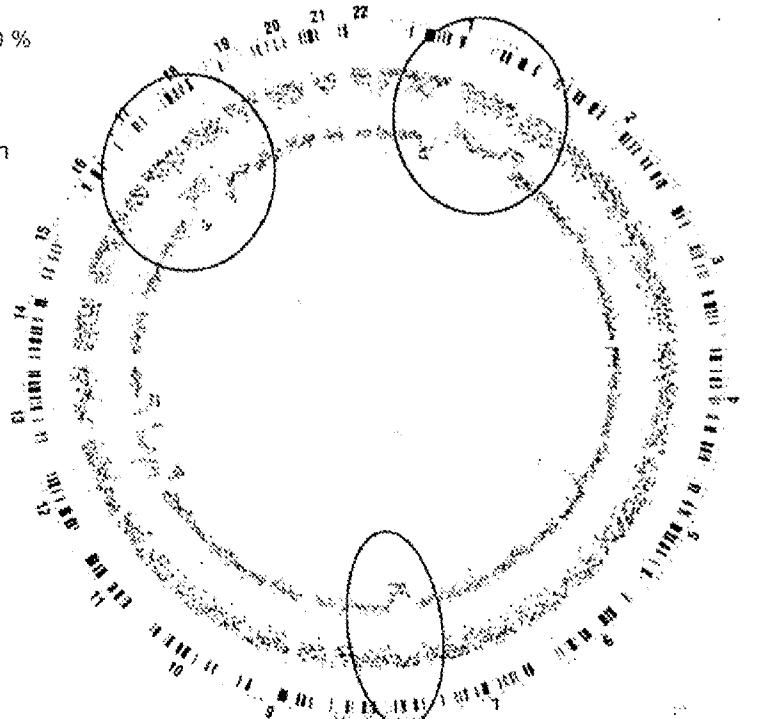


FIG. 48A

Muestra: TBR36
Fracción de ADN tumoral: 40 %

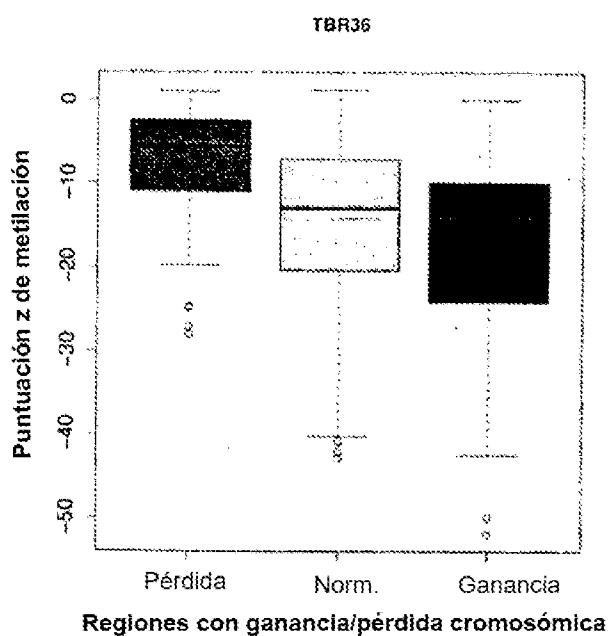


FIG. 48B

ES 2 994 426 T3

Muestra: TBR34
Fracción de ADN tumoral: 30 %

De dentro hacia fuera:
1) Puntuación z de ANC
2) Puntuación z de metilación

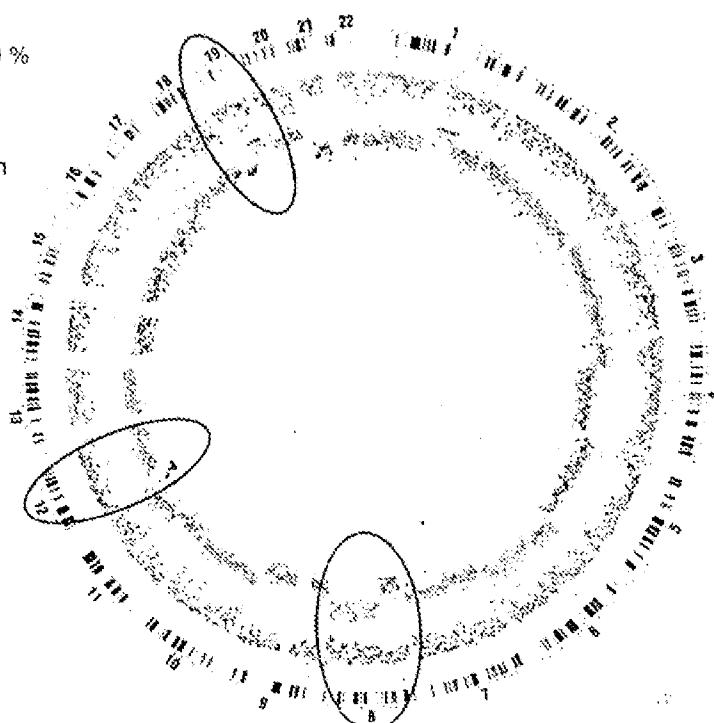


FIG. 49A

Muestra: TBR34
Fracción de ADN tumoral: 30 %

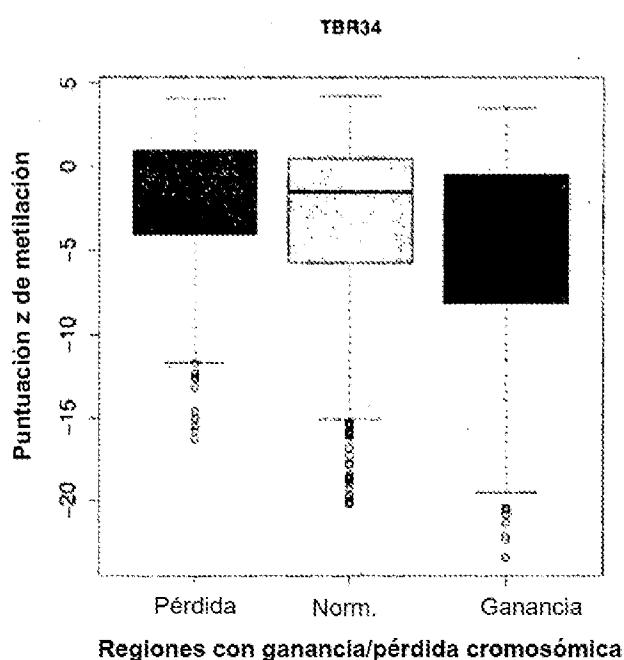


FIG. 49B

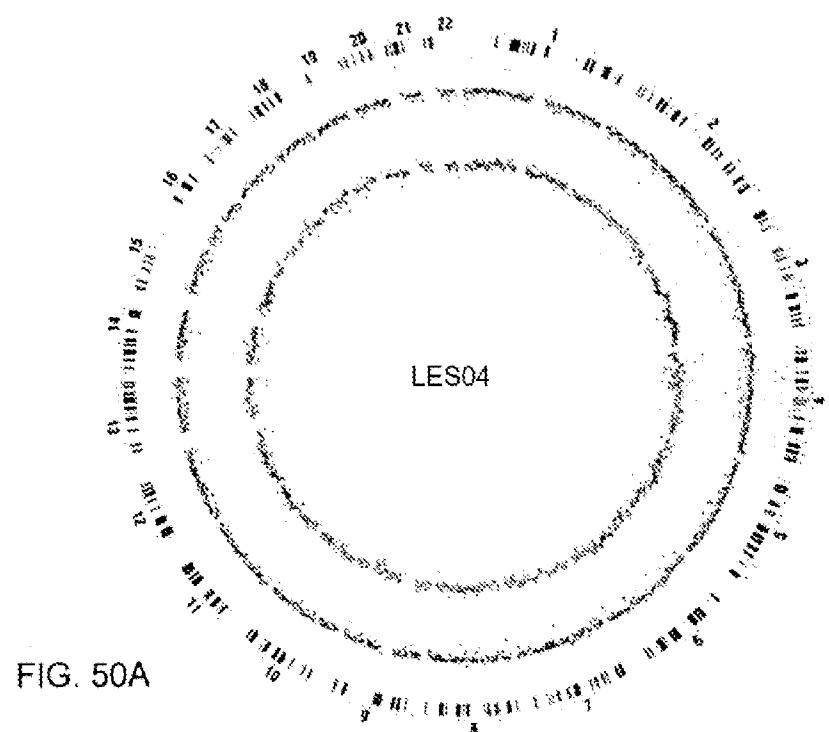


FIG. 50A

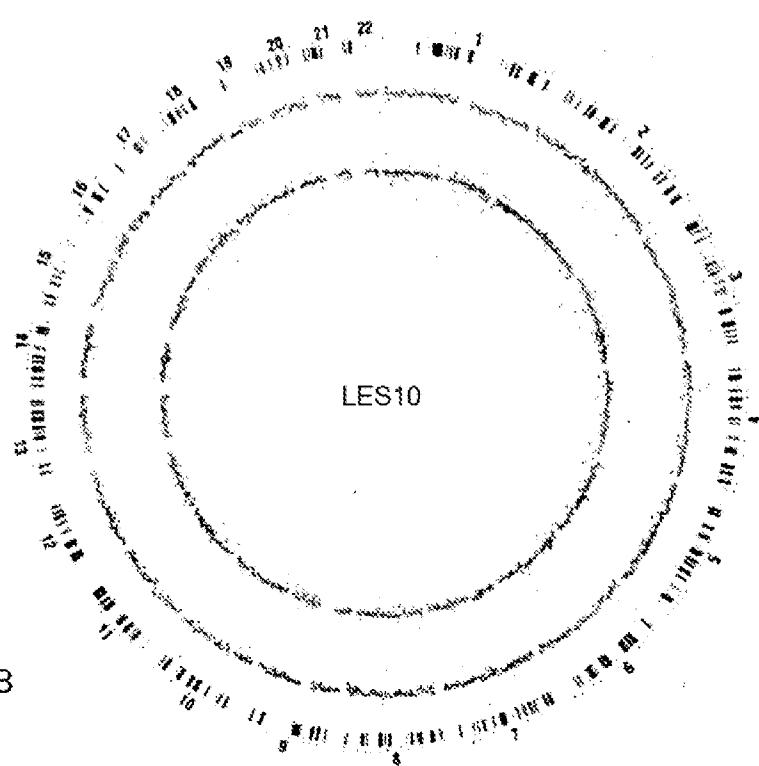


FIG. 50B

ES 2 994 426 T3

T003

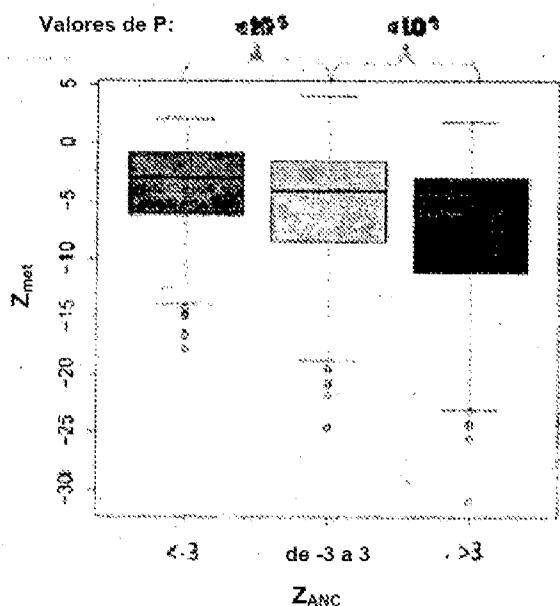


FIG. 51A

T003

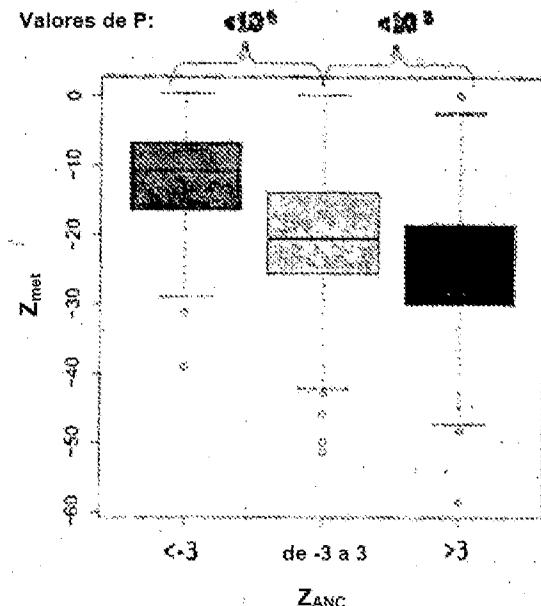


FIG. 51B

LES04

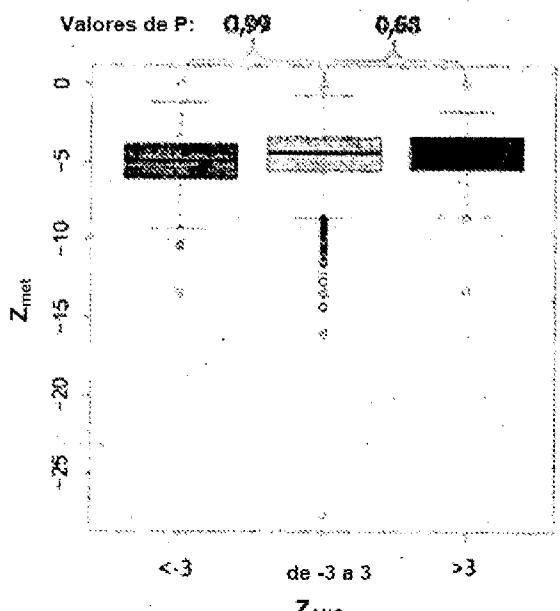


FIG. 51C

LES10

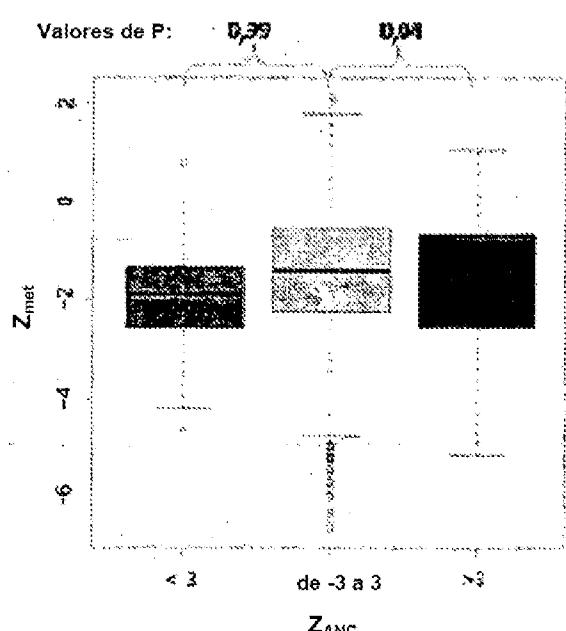


FIG. 51D

ES 2 994 426 T3

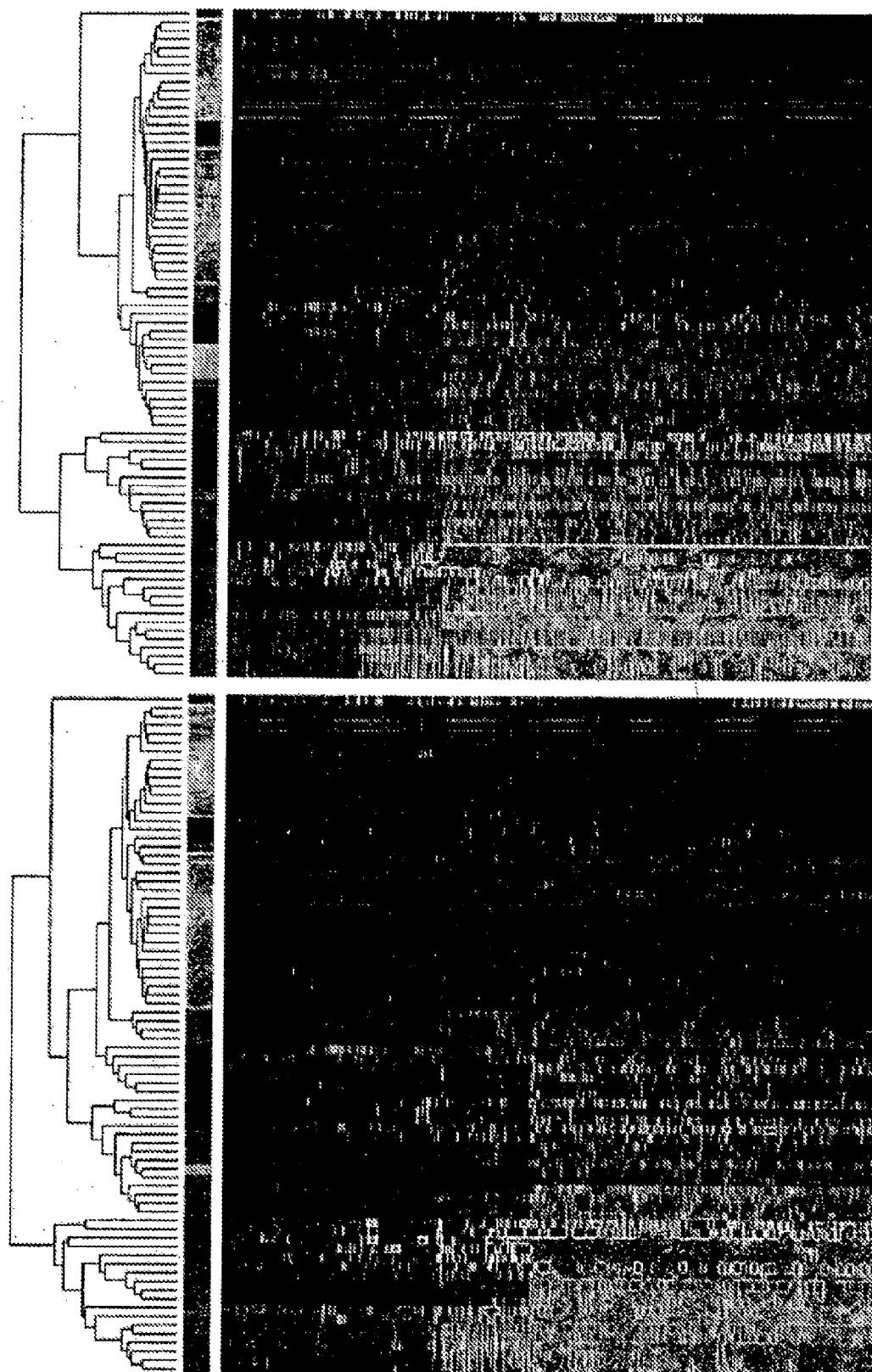


FIG. 52A

FIG. 52B

ES 2 994 426 T3

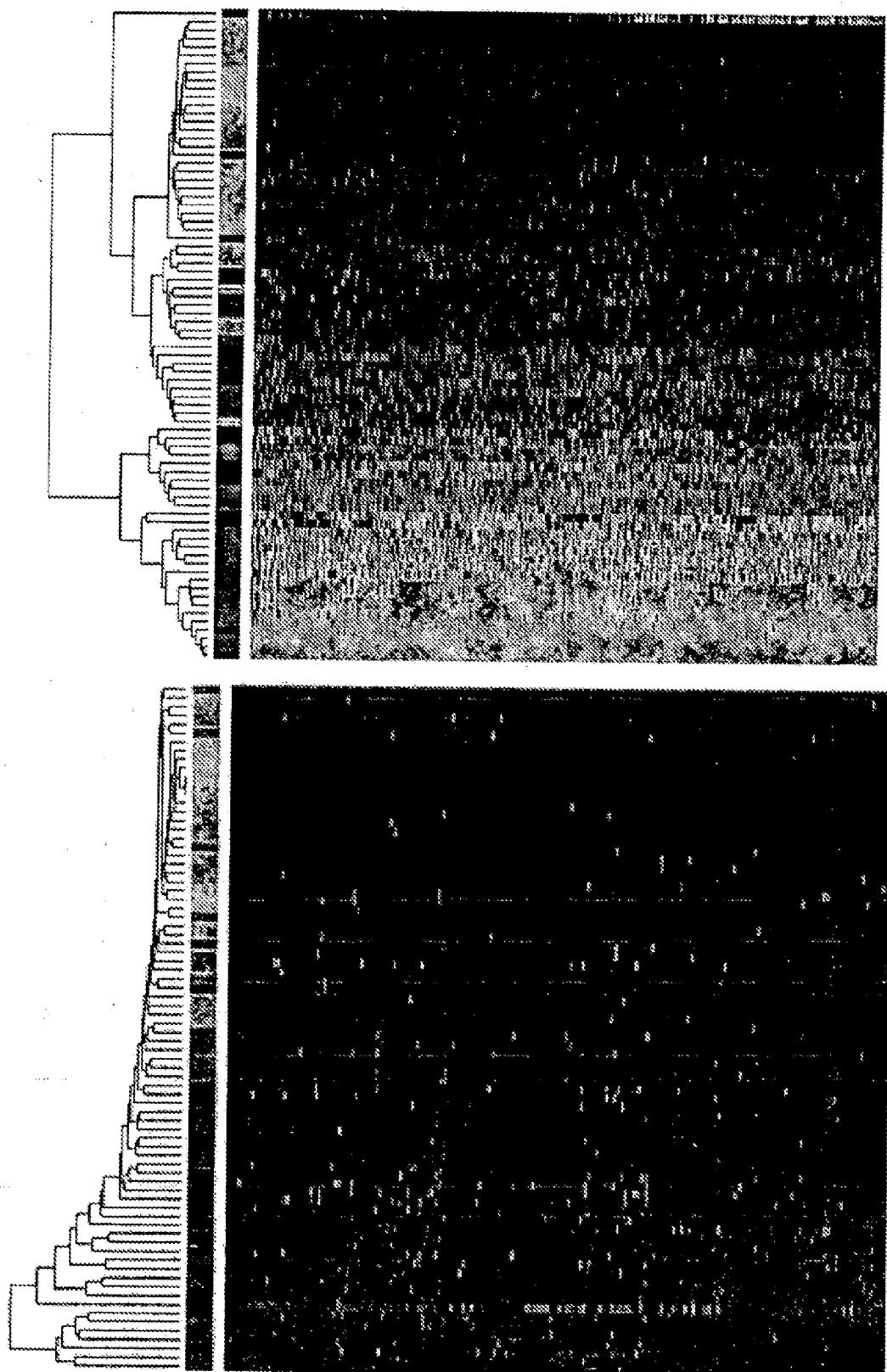


FIG. 53A

FIG. 53B

ES 2 994 426 T3

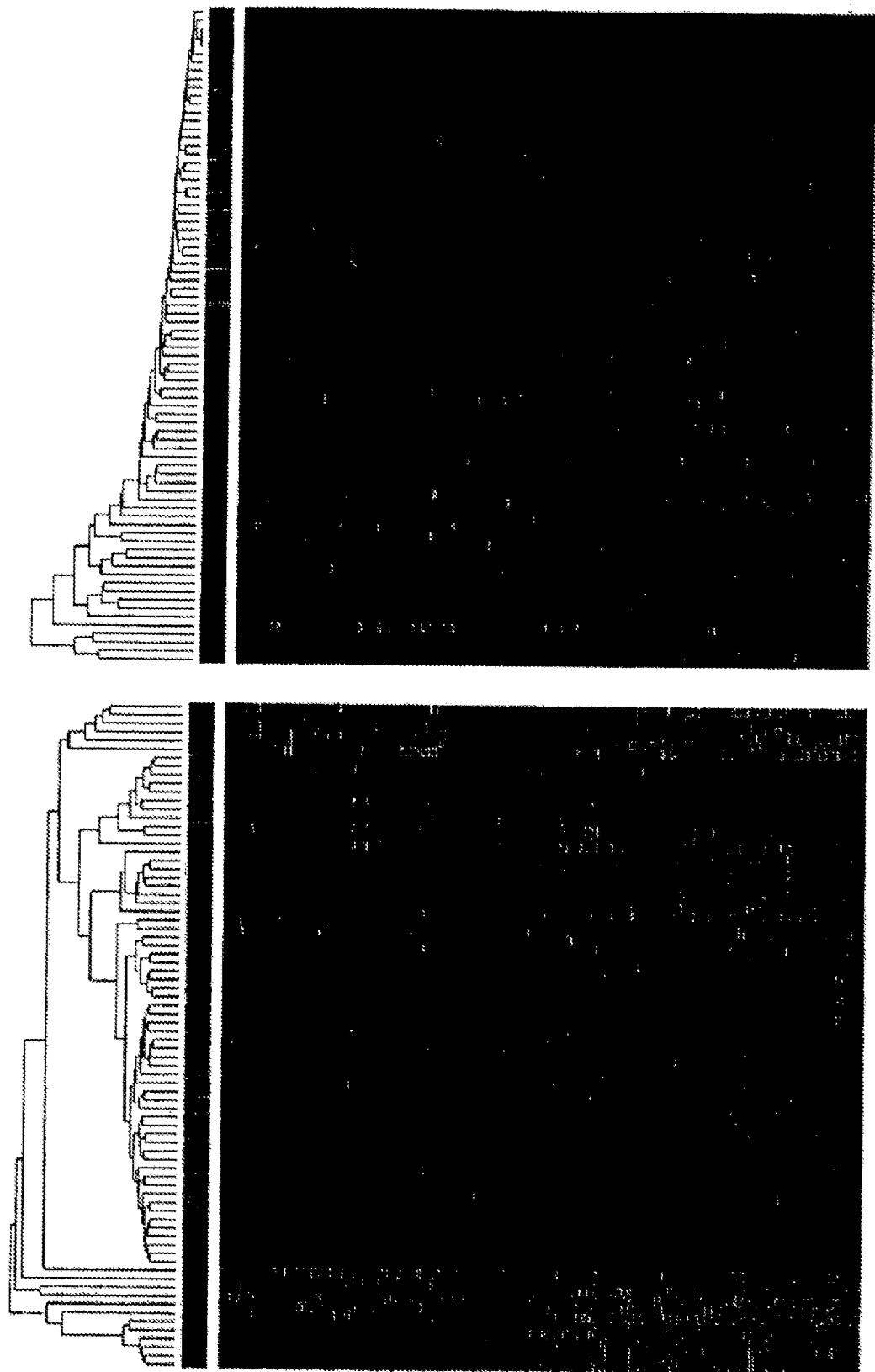


FIG. 54A
FIG. 54B

ES 2 994 426 T3

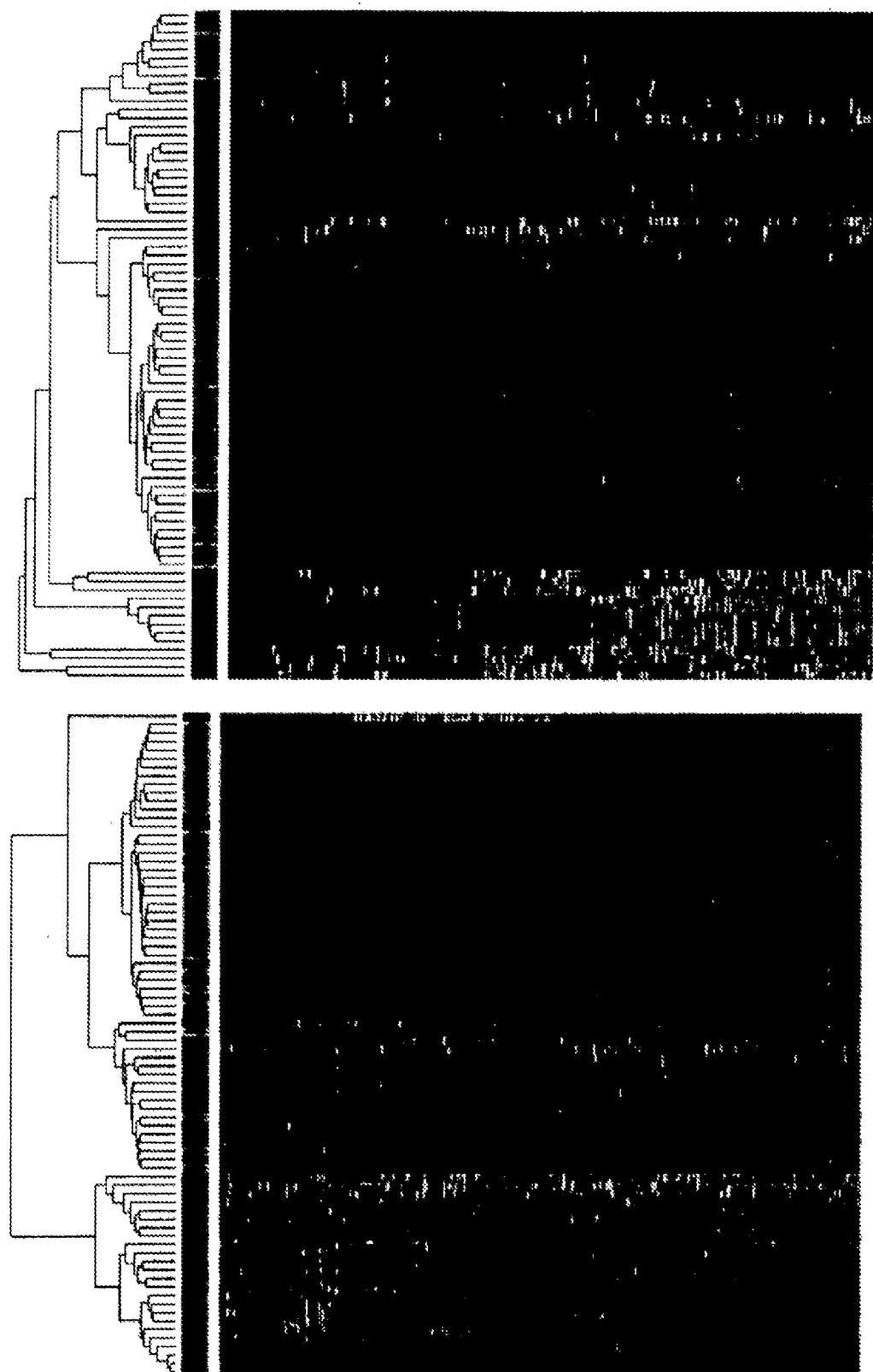


FIG. 55A

FIG. 55B

ES 2 994 426 T3

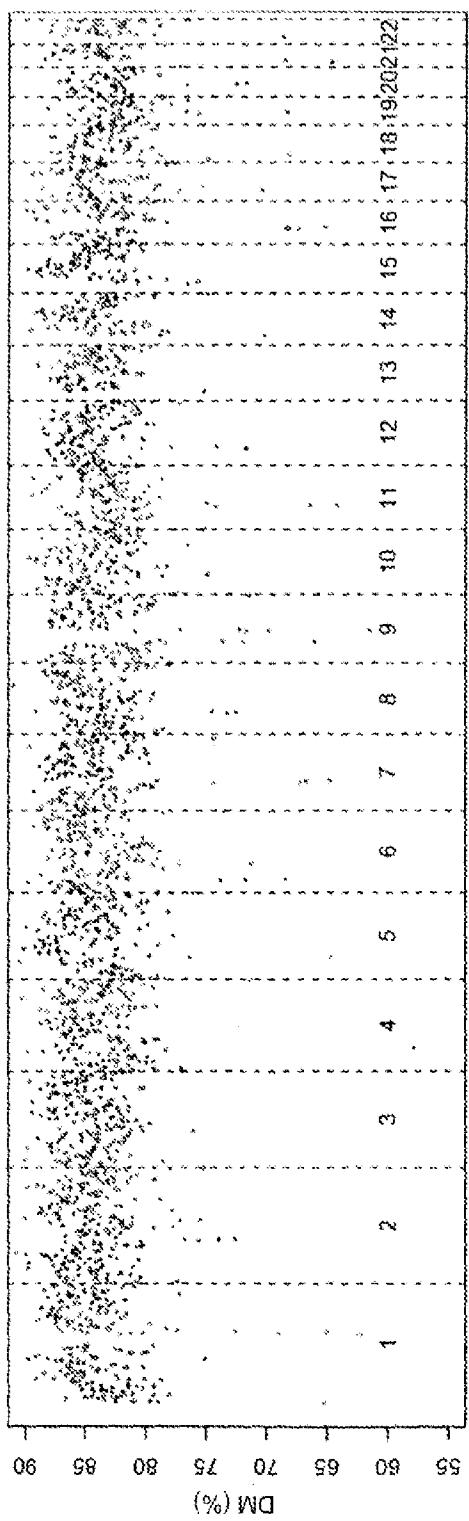


FIG. 56