

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 969 567**

51 Int. Cl.:

C12Q 1/6886 (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 19171630 (7)**

97 Fecha y número de publicación de la concesión europea: **29.11.2023** **EP 3536807**

54 Título: **Determinación no invasiva de metiloma de tumor 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:
21.05.2024

73 Titular/es:

THE CHINESE UNIVERSITY OF HONG KONG
(100.0%)

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 2 969 567 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 de metiloma de tumor de plasma

5 La presente divulgación se refiere, en general, a la determinación de un patrón de metilación (metiloma) de ADN, y más particularmente, a analizar una muestra biológica (por ejemplo, plasma) que incluye una mezcla de ADN de diferentes genomas (por ejemplo, del feto y de la madre, o de células tumorales y normales) para determinar el patrón de metilación (metiloma) del genoma minoritario. También se describen usos del metiloma determinado.

10 El desarrollo embrionario y fetal es un proceso complejo e implica una serie de acontecimientos genéticos y epigenéticos altamente orquestados. El desarrollo del cáncer también es un proceso complejo que implica típicamente múltiples etapas genéticas y epigenéticas. Las anomalías en el control epigenético de procesos del desarrollo están implicadas en la infertilidad, el aborto espontáneo, anomalías del crecimiento intrauterino y consecuencias postnatales.

15 La metilación del ADN es uno de los mecanismos epigenéticos más frecuentemente estudiados. La metilación del ADN se produce principalmente en el contexto de la adición de un grupo de metilo al carbono 5' de restos de citosina entre dinucleótidos de CpG. La metilación de citosina añade una capa de control a la transcripción génica y la función del ADN. Por ejemplo, la hipermetilación de promotores génicos enriquecidos con dinucleótidos de CpG, denominados islas de CpG, se asocia típicamente con represión de la función génica.

20 A pesar del papel importante de los mecanismos epigenéticos en la mediación de procesos del desarrollo, los tejidos embrionarios y fetales humanos no son fácilmente accesibles para su análisis (los tumores pueden de forma similar no ser accesibles). Los estudios de los cambios dinámicos de dichos procesos epigenéticos en salud y enfermedad durante el periodo prenatal en seres humanos son prácticamente imposibles. Los tejidos extraembrionarios, particularmente la placenta, que pueden obtenerse como parte de procedimientos de diagnóstico prenatales o después del nacimiento, han proporcionado una de las principales vías para dichas investigaciones. Sin embargo, dichos tejidos requieren procedimientos invasivos.

El perfil de metilación de ADN de la placenta humana ha intrigado a los investigadores durante décadas. La placenta humana muestra una multitud de características fisiológicas peculiares que implican metilación de ADN. A nivel global, los tejidos placentarios están hipometilados en comparación con la mayoría de los tejidos somáticos. Al nivel génico, el estado de metilación de locus genómicos seleccionados es una identificación específica de tejidos placentarios. Los perfiles de metilación tanto globales como específicos de locus muestran cambios dependientes de la edad gestacional. Los genes con impronta, concretamente genes para los que la expresión es dependiente del origen parental de los alelos, cumplen funciones clave en la placenta. La placenta se ha descrito como pseudomaligna y se ha observado hipermetilación de varios genes supresores de tumores.

Los estudios del perfil de metilación de ADN de tejidos placentarios han proporcionado información acerca de la patofisiología de enfermedades asociadas con el embarazo o relacionadas con el desarrollo tales como preeclampsia y restricción del crecimiento intrauterino. Los trastornos en la impronta genómica se asocian con trastornos del desarrollo, tales como síndrome de Prader-Willi y síndrome de Angelman. Se han observado perfiles alterados de impronta genómica y metilación de ADN global en tejidos placentarios y fetales en embarazos resultantes de técnicas de reproducción asistida (H Hiura *et al.*, 2012 Hum Reprod; 27: 2541-2548). Varios factores ambientales tales como tabaquismo materno (KE Haworth *et al.*, 2013 Epigenomics; 5: 37-49), factores dietéticos maternos (X Jiang y otros, 2012 FASEB J; 26: 3563-3574) y estado metabólico materno tal como diabetes (N Hajj *et al.*, Diabetes. doi: 10.2337/db12-0289) se han asociado con aberraciones epigenéticas de la descendencia.

A pesar de décadas de intentos, no ha habido ningún medio práctico disponible para estudiar el metiloma fetal o tumoral y para supervisar los cambios dinámicos durante el embarazo o durante los procesos de enfermedad, tales como tumores malignos. Por lo tanto, es deseable proporcionar métodos para analizar todos o parte de un metiloma fetal y un metiloma tumoral de forma no invasiva.

En el documento WO 2011/038507 se desvela un método para analizar perfiles de metilación de ADN de ADN sin células en fluidos corporales. En el documento WO 2012/031329 se enseña un ensayo para detectar perfiles epigenéticos asociados con el cáncer, incluyendo su forma o tipo, estado y estatus de enfermedad residual mínima.

De acuerdo con la presente invención, se proporciona un método para analizar una muestra de plasma de un organismo humano, comprendiendo la muestra de plasma ADN sin células que se origina de células normales y posiblemente de células asociadas con cáncer, comprendiendo el método: analizar una pluralidad de moléculas de ADN sin células de la muestra de plasma, en donde analizar una molécula de ADN sin células incluye: determinar una localización de la molécula de ADN sin células en un genoma del organismo humano mapeando, mediante lecturas de secuencia, la molécula de ADN sin células en una parte de un genoma humano; y determinar si la molécula de ADN sin células está metilada en uno o más sitios CpG de una pluralidad de sitios CpG organizados en una pluralidad de islas CpG, incluyendo cada isla CpG más de un sitio CpG; determinar un número respectivo de moléculas de ADN sin células en cada una de la pluralidad de islas CpG que están metiladas; calcular un primer nivel de metilación basado en los números respectivos de moléculas de ADN sin células que están metiladas en la pluralidad de islas CpG; comparar el primer nivel de metilación con un primer valor de corte indicativo de la presencia de cáncer; y

determinar una primera clasificación de la presencia de cáncer basada en la comparación.

En el contexto del cáncer, la medición de los cambios metilómicos en el plasma puede permitir detectar el cáncer (por ejemplo, con fines de exploración), para el seguimiento (por ejemplo, para detectar la respuesta tras el tratamiento contra el cáncer; y para detectar la recidiva del cáncer) y para el pronóstico (por ejemplo, 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).

Puede obtenerse un mejor entendimiento de la naturaleza y de las ventajas de las realizaciones de la presente invención con referencia a la siguiente descripción detallada y a los dibujos adjuntos.

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

La FIGURA 1B muestra la densidad de metilación en ventanas de 1 Mb de muestras secuenciadas de acuerdo con realizaciones de la presente divulgación.

Las FIGURAS 2A-2C muestran representaciones de los valores beta contra los índices de metilación: (A) Células sanguíneas maternas, (B) Muestra de vellosidad coriónica, (C) Tejido placentario a término.

Las FIGURAS 3A y 3B muestran diagramas de barras del porcentaje de sitios de CpG metilados en plasma y células sanguíneas recogidas de un hombre adulto y de una mujer adulta no embarazada: (A) Autosomas, (B) Cromosoma X.

Las FIGURAS 4A y 4B muestran representaciones de densidades de metilación de locus correspondientes en ADN de células sanguíneas y ADN de plasma: (A) Mujer adulta no embarazada, (B) Hombre adulto.

Las FIGURAS 5A y 5B muestran gráficos de barras del porcentaje de sitios de CpG metilados entre muestras recogidas del embarazo: (A) Autosomas, (B) Cromosoma X.

La FIGURA 6 muestra un gráfico de barras del nivel de metilación de diferentes clases repetidas del genoma humano para sangre materna, placenta y plasma materno.

La FIGURA 7A muestra una representación de Circos 700 para muestras del primer trimestre. La FIGURA 7B muestra una representación de Circos 750 para muestras del tercer trimestre.

Las FIGURAS 8A-8D muestran representaciones de comparaciones de las densidades de metilación de ADN del tejido genómico frente a ADN del plasma materno para sitios de CpG que rodean a los polimorfismos de un único nucleótido informativos.

La FIGURA 9 es un diagrama de flujo que ilustra un método 900 para determinar un primer perfil de metilación de una muestra biológica de un organismo de acuerdo con realizaciones de la presente divulgación.

La FIGURA 10 es un diagrama de flujo que ilustra un método 1000 de determinación de un primer perfil de metilación de una muestra biológica de un organismo de acuerdo con realizaciones de la presente divulgación.

Las FIGURAS 11A y 11B muestran gráficos del rendimiento del algoritmo de predicción usando datos de plasma materno y concentración de ADN fetal fraccional de acuerdo con realizaciones de la presente divulgación.

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

La FIGURA 13 es un diagrama de flujo de un método 1300 para detectar una anomalía cromosómica fetal de una muestra biológica de un sujeto mujer embarazada con al menos un feto.

La FIGURA 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 divulgación.

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

La FIGURA 16 es una tabla 1600 que muestra los números de locus que se ha predicho que están hipermetilados o hipometilados basándose en el análisis directo de los datos de secuenciación por bisulfito de plasma materno.

La FIGURA 17A es una representación 1700 que muestra la distribución de tamaños del plasma materno, plasma de control de mujer no embarazada, ADN placentario y de sangre periférica. La FIGURA 17B es una representación 1750 de la distribución de tamaño y el perfil de metilación del plasma materno, plasma de control de mujer adulta, tejido placentario y sangre de control de mujer adulta.

Las FIGURAS 18A y 18B son representaciones de densidades de metilación y tamaño de moléculas de ADN de plasma de acuerdo con realizaciones de la presente divulgación.

La FIGURA 19A muestra una representación 1900 de densidades de metilación y los tamaños de lecturas secuenciadas para una mujer adulta no embarazada. La FIGURA 19B es una representación 1950 que muestra la distribución de tamaño y el perfil de metilación de moléculas de ADN específicas fetales y específicas maternas en plasma materno.

La FIGURA 20 es un diagrama de flujo de un método 2000 para estimar un nivel de metilación de ADN en una muestra biológica de un organismo de acuerdo con realizaciones de la presente divulgación.

La FIGURA 21A es una tabla 2100 que muestra las densidades de metilación del plasma preoperatorio y las muestras tisulares de un paciente con carcinoma hepatocelular (CHC). La FIGURA 21B es una tabla 2150 que

muestra el número de lecturas de secuencias y la profundidad de secuenciación conseguida por muestra.

La FIGURA 22 es una tabla 220 que muestra las densidades de metilación en los autosomas, que varían de 71,2 % a 72,5 %, en las muestras de plasma de los controles sanos.

Las FIGURAS 23A y 23B muestran densidad de metilación de la capa leucoplaquetaria, tejido tumoral, tejido hepático no tumoral, el plasma preoperatorio y plasma postoperatorio del paciente con CHC.

La FIGURA 24A es una representación 2400 que muestra las densidades de metilación del plasma preoperatorio del paciente con CHC. La FIGURA 24B es una representación 2450 que muestra las densidades de metilación del plasma postoperatorio del paciente con CHC.

Las FIGURAS 25A y 25B muestran puntuaciones z de las densidades de metilación de ADN de plasma para las muestras de plasma preoperatorias (representación 2500) y postoperatorias (representación 2550) del paciente con CHC usando los datos del metiloma de plasma de los cuatro sujetos de control sanos como referencia para el cromosoma 1.

La FIGURA 26A es una tabla 2600 que muestra datos para puntuaciones z para plasma preoperatorio y postoperatorio. La FIGURA 26B es una representación de Circos 2620 que muestra la puntuación z de las densidades de metilación de ADN de plasma para las muestras de plasma preoperatorias y postoperatorias del paciente con CHC usando los cuatro sujetos de control sano como referencia para grupos de 1 Mb analizados de todos los autosomas. La FIGURA 26C es una tabla 2640 que muestra una distribución de las puntuaciones z de los grupos de 1 Mb para el genoma completo en las muestras de plasma tanto preoperatorias como postoperatorias del paciente con CHC. La FIGURA 26D es una tabla 2660 que muestra los niveles de metilación del tejido tumoral y muestra de plasma preoperatoria solapante con algunas de las muestras de plasma de control cuando se usan los contextos de CHH y CHG.

La FIGURA 27A-H muestran representaciones de Circos de densidad de metilación de 8 pacientes con cáncer de acuerdo con realizaciones de la presente divulgación. La FIGURA 27I es la tabla 2780 que muestra el número de lecturas de secuencias y la profundidad de secuenciación conseguida por muestra. La FIGURA 27J es una tabla 2790 que muestra una distribución de las puntuaciones z de los grupos de 1 Mb para el genoma completo en plasma de pacientes con diferentes tumores malignos. CP = adenocarcinoma de pulmón; CNF = carcinoma nasofaríngeo; CCR = carcinoma colorrectal; NE = carcinoma neuroendocrino; SML = sarcoma de músculo liso.

La FIGURA 28 es un diagrama de flujo del método 2800 para analizar una muestra biológica de un organismo para determinar una clasificación de un nivel de cáncer de acuerdo con realizaciones de la presente divulgación.

La FIGURA 29A es una representación 2900 que muestra la distribución de las densidades de metilación en sujetos de referencia suponiendo que esta distribución siga una distribución normal. La FIGURA 29B es una representación 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 de metilación medio está 2 desviaciones típicas por debajo del punto de corte.

La FIGURA 30 es una representación 3000 que muestra la distribución de densidades de metilación del ADN de plasma de sujetos sanos y pacientes con cáncer.

La FIGURA 31 es un gráfico 3100 que muestra la distribución de las diferencias en las densidades de metilación entre la media del ADN de plasma de sujetos sanos y el tejido tumoral del paciente con CHC.

La FIGURA 32A es una tabla 3200 que muestra el efecto de la reducción de la profundidad de secuenciación cuando la muestra de plasma contenía ADN tumoral 5 % o 2 %.

La FIGURA 32B es un gráfico 3250 que muestra las densidades de metilación de los elementos repetidos y regiones no repetidas en el plasma de los cuatro sujetos de control sanos, la capa leucoplaquetaria, el tejido de hígado normal, el tejido tumoral, el plasma preoperatorio y las muestras de plasma postoperatorias del paciente con CHC.

La FIGURA 33 muestra un diagrama de bloques de un sistema informático ejemplar 3300 que puede utilizarse con el sistema y métodos de acuerdo con realizaciones de la presente divulgación.

La FIGURA 34A muestra una distribución de tamaños de ADN de plasma en el paciente con lupus eritematoso sistémico (LES) SLE04. Las FIGURAS 34B y 34C muestran análisis de metilación para ADN de plasma de un paciente con LES SLE04 (FIGURA 34B) y un paciente con CHC TBR36 (FIGURA 34C).

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

La FIGURA 36 es un diagrama de flujo de un método 3600 para analizar una muestra biológica de un organismo usando una pluralidad de regiones cromosómicas de acuerdo con realizaciones de la presente divulgación.

La FIGURA 37A muestra análisis de ANC para tejidos tumorales, ADN de plasma no tratado con bisulfito (BS) y ADN de plasma tratado con bisulfito (de dentro a fuera) para el paciente TBR36. La FIGURA 37B es una representación de dispersión que muestra la relación entre las puntuaciones z para la detección de ANC usando plasma tratado con bisulfito y no tratado con bisulfito de los grupos de 1 Mb para el paciente TBR36.

La FIGURA 38A muestra análisis de ANC para tejidos tumorales, ADN de plasma no tratado con bisulfito (BS) y ADN de plasma tratado con bisulfito (de dentro a fuera) para el paciente TBR34. La FIGURA 38B es una representación de dispersión que muestra la relación entre las puntuaciones z para la detección de ANC usando plasma tratado con bisulfito y no tratado con bisulfito de los grupos de 1 Mb para el paciente TBR34.

La FIGURA 39A es una representación de Circos que muestra el análisis de ANC (anillo interno) y metilación (anillo externo) para el plasma tratado con bisulfito para un paciente con CHC TBR240. La FIGURA 39B es una representación de Circos que muestra el análisis de ANC (anillo interno) y metilación (anillo externo) para el plasma tratado con bisulfito para un paciente con CHC TBR164.

La FIGURA 40A muestra el análisis de ANC para el paciente TBR36 para la muestra pretratamiento y la muestra

postratamiento.

La FIGURA 40B muestra el análisis de metilación para el paciente TBR36 para la muestra pretratamiento y la muestra postratamiento. La FIGURA 41A muestra el análisis de ANC para el paciente TBR34 para la muestra pretratamiento y la muestra postratamiento. La FIGURA 41B muestra el análisis de metilación para el paciente

TBR34 para la muestra pretratamiento y la muestra postratamiento. La FIGURA 42 muestra un diagrama de rendimiento de diagnóstico de análisis de hipometilación en todo el genoma con diferente número de lecturas secuenciadas.

La FIGURA 43 es un diagrama que muestra curvas de ROC para la detección de cáncer basándose en análisis de hipometilación de todo el genoma con diferentes tamaños de grupos (50 kb, 100 kb, 200 kb y 1 Mb).

La FIGURA 44A muestra un rendimiento de diagnóstico para probabilidad acumulada (PA) y porcentaje de grupos con aberraciones. La FIGURA 44B muestra rendimientos de diagnóstico para el análisis de plasma con respecto a hipometilación global, hipermetilación de islas de CpG y ANC.

La FIGURA 45 muestra una tabla con resultados para hipometilación global, hipermetilación de islas de CpG y ANC en pacientes con carcinoma hepatocelular.

La FIGURA 46 muestra una tabla con resultados para hipometilación global, hipermetilación de islas de CpG y ANC en pacientes que padecen cánceres distintos de carcinoma hepatocelular.

La FIGURA 47 muestra un análisis en serie de metilación en plasma para el caso TBR34.

La FIGURA 48A muestra representaciones de Circos que demuestran la ANC (anillo interno) y los cambios de metilación (anillo externo) en el ADN de plasma tratado con bisulfito para el paciente con CHC TBR36. La FIGURA 48B es una representación de puntuaciones z de metilación para regiones con ganancia y pérdida cromosómica, y regiones sin cambio del número de copias para el paciente con CHC TBR36.

La FIGURA 49A muestra representaciones de Circos que demuestran la ANC (anillo interno) y los cambios de metilación (anillo externo) en el ADN de plasma tratado con bisulfito para el paciente con CHC TBR34. La FIGURA 49B es una representación de puntuaciones z de metilación para regiones con ganancia y pérdida cromosómica, y regiones sin cambio del número de copias para el paciente con CHC TBR34.

Las FIGURAS 50A y 50B muestran resultados de hipometilación de plasma y análisis de ANC para los pacientes con LES SLE04 y SLE10.

Las FIGURAS 51A y 51B muestran análisis de Z_{met} para regiones con y sin ANC para el plasma de dos pacientes con CHC (TBR34 y TBR36). Las FIGURAS 51C y 51D muestran análisis de Z_{met} para regiones con y sin ANC para el plasma de dos pacientes con LES (SLE04 y SLE10).

La FIGURA 52A muestra análisis de agrupamiento jerárquico para muestras de plasma de pacientes con CHC, pacientes con cáncer distinto de CHC y sujetos de control sanos usando características de grupo A para ANC, metilación global y metilación de islas de CpG. La FIGURA 52B muestra agrupamiento jerárquico usando características del grupo B para ANC, metilación global y metilación de islas de CpG.

La FIGURA 53A muestra análisis de agrupamiento jerárquico para muestras de plasma de pacientes con CHC, pacientes con cáncer distinto de CHC y sujetos de control sanos usando los elementos de metilación de islas de CpG del grupo A. La FIGURA 53B muestra análisis de agrupamiento jerárquico para muestras de plasma de pacientes con CHC, pacientes con cáncer distinto de CHC y sujetos de control sanos usando las densidades de metilación globales del grupo A.

La FIGURA 54A muestra un análisis de agrupamiento jerárquico para muestras de plasma de pacientes con CHC, pacientes con cáncer distinto de CHC y sujetos de control sanos usando los ANC globales del grupo A. La FIGURA 54B muestra un análisis de agrupamiento jerárquico para muestras de plasma de pacientes con CHC, pacientes con cáncer distinto de CHC y sujetos de control sanos usando las densidades de metilación de islas de CpG del grupo B.

La FIGURA 55A muestra un análisis de agrupamiento jerárquico para muestras de plasma de pacientes con CHC, pacientes con cáncer distinto de CHC y sujetos de control sanos usando las densidades de metilación globales del grupo B. La FIGURA 55B muestra un análisis de agrupamiento jerárquico para muestras de plasma de pacientes con CHC, pacientes con cáncer distinto de CHC y sujetos de control sanos usando las densidades de metilación globales del grupo B.

La FIGURA 56 muestra la densidad de metilación media de grupos de 1 Mb (puntos rojos) entre 32 sujetos sanos.

Un "metiloma" proporciona una medida de una cantidad de metilación de ADN en una pluralidad de sitios o locus en un genoma. El metiloma puede corresponder a todo el genoma como una parte sustancial del genoma o partes relativamente pequeñas del genoma. Un "metiloma fetal" corresponde al metiloma de un feto de una mujer embarazada. El metiloma fetal puede determinarse usando una diversidad de tejidos fetales o fuentes de ADN fetal, incluyendo tejidos placentarios y ADN fetal sin células en plasma materno. Un "metiloma tumoral" corresponde al metiloma de un tumor de un organismo (por ejemplo, un ser humano). El metiloma tumoral puede determinarse usando tejido tumoral o ADN tumoral sin células en plasma materno. El metiloma fetal y el metiloma tumoral son ejemplos de un metiloma de interés. Otros ejemplos de metilomas de interés son los metilomas de órganos (por ejemplo metilomas de células cerebrales, huesos, los pulmones, el corazón, los músculos y los riñones, etc.) que pueden aportar ADN en un fluido corporal (por ejemplo, plasma, suero, sudor, saliva, orina, secreciones genitales, semen, líquido de heces, líquido diarreico, líquido cefalorraquídeo, secreciones del tracto gastrointestinal, secreciones pancreáticas, secreciones intestinales, esputo, lágrimas, líquidos de aspiración de mama y tiroides, etc.). Los órganos pueden ser

Un "metiloma de plasma" es el metiloma determinado a partir del plasma o suero de un animal (por ejemplo, un ser

humano). El metiloma de plasma es un ejemplo de un metiloma sin células ya que el plasma y el suero incluyen ADN sin células. El metiloma de plasma también es un ejemplo de un metiloma mixto ya que es una mezcla de metiloma fetal/materno o metiloma de tumor/paciente. El “metiloma placentario” puede determinarse a partir de una muestra de vellosidad coriónica (CVS) o una muestra tisular placentaria (por ejemplo, obtenida después del parto). El “metiloma celular” corresponde al metiloma determinado a partir de células (por ejemplo, células sanguíneas) del paciente. El metiloma de las células sanguíneas se denomina metiloma de células sanguíneas (o metiloma sanguíneo).

Un “sitio” corresponde a un único sitio, que puede ser una única posición de base o un grupo de posiciones de bases correlacionadas, por ejemplo, un sitio de CpG. Un “locus” puede corresponder a una región que incluye múltiples sitios. Un locus puede incluir solamente un sitio, lo que haría al locus equivalente a un sitio en ese contexto.

El “índice de metilación” para cada sitio genómico (por ejemplo, un sitio de CpG) se refiere a la proporción de lecturas de secuencias que muestran metilación en el sitio frente al número total de lecturas que abarcan ese sitio. La “densidad de metilación” de una región es el número de lecturas en sitios dentro de la región que muestran metilación dividida por el número total de lecturas que abarcan los sitios en la región. Los sitios pueden tener características específicas, por ejemplo, que son sitios de CpG. Por lo tanto, la “densidad de metilación de CpG” de una región es el número de lecturas que muestran metilación de CpG dividido por el número total de lecturas que abarcan sitios de CpG en la región (por ejemplo, un sitio de CpG particular, sitios de CpG dentro de una isla de CpG, o una región mayor). Por ejemplo, la densidad de metilación para cada grupo de 100 kb en el genoma humano puede determinarse a partir del número total de citosinas no convertidas después del tratamiento con bisulfito (que corresponde a citosina metilada) en sitios de CpG como una proporción de todos los sitios de CpG abarcados por lecturas de secuencias mapeadas a la región de 100 kb. Este análisis también puede realizarse para otros tamaños de grupos, por ejemplo 50 kb o 1 Mb, etc. Una región podría ser el genoma completo o un cromosoma o parte de un cromosoma (por ejemplo, una rama cromosómica). El índice de metilación de un sitio de CpG es el mismo que la densidad de metilación para una región cuando la región incluye solamente ese sitio de CpG. La “proporción de citosinas metiladas” se refiere al número de sitios de citosina, “C”, que se ha mostrado que están metilados (por ejemplo, no convertidos después de la conversión por bisulfito) frente al 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”.

Un “perfil de metilación” (también denominado estado de metilación) incluye información relacionada con la metilación de ADN para una región. La información relacionada con metilación de ADN puede incluir, pero sin limitación, un índice de metilación de un sitio de CpG, una densidad de metilación de sitios de CpG en una región, una distribución de sitios de CpG sobre una región contigua, un patrón o nivel de metilación para cada sitio de CpG individual dentro de una región que contiene más de un sitio de CpG y metilación distinta de CpG. Un perfil de metilación de una parte sustancial del genoma puede considerarse equivalente al metiloma. La “metilación de ADN” en genomas de mamífero se refiere típicamente a la adición de un grupo metilo al carbono 5' de restos de citosina (es decir 5-metilcitosinas) entre dinucleótidos de CpG. Puede producirse metilación de ADN en citosinas en otros contextos, por ejemplo, CHG y CHH, donde H es adenina, citosina o timina. La metilación de citosina también puede estar en forma de 5-hidroximetilcitosina. También se ha indicado la metilación distinta de citosina, tal como N6-metiladenina.

Un “tejido” corresponde a cualquier célula. Diferentes tipos de tejido pueden corresponder a diferentes tipos de células (por ejemplo, hígado, pulmón o sangre), pero también pueden corresponder a tejido de diferentes órganos (madre frente a feto) o a células sanas frente a células tumorales. Una “muestra biológica” se refiere a cualquier muestra que se tome de un sujeto (por ejemplo, un ser humano, tal como una mujer embarazada, una persona con cáncer o una persona que se sospecha que tiene cáncer, un receptor de trasplante de órgano o un sujeto que se sospecha que tiene un proceso de enfermedad que implica un órgano (por ejemplo, el corazón en infarto de miocardio, o el cerebro en ictus) y contiene una o más moléculas de ácido nucleico de interés. La muestra biológica puede ser un líquido corporal, tal como sangre, plasma, suero, orina, líquido vaginal, líquidos de lavado uterino o vaginal, líquido pleural, líquido ascítico, líquido cefalorraquídeo, saliva, sudor, lágrimas, esputo, líquido de lavado broncoalveolar, etc. También pueden usarse muestras de heces.

La expresión “nivel de cáncer” puede referirse a si existe cáncer, un estadio de un cáncer, un tamaño de tumor, si hay metástasis, la carga tumoral total del cuerpo y/u otra medida de una 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 condiciones premalignas o precancerosas (estados) asociados con mutaciones o un número de mutaciones. El nivel de cáncer puede usarse de diversas maneras. Por ejemplo, la exploración puede comprobar si está presente cáncer en alguien de quien no se ha sabido previamente que tuviera cáncer. La evaluación puede investigar a alguien a quien se ha diagnosticado que tiene cáncer para supervisar el progreso del cáncer 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 un tiempo específico, o la probabilidad de que el cáncer se metastatice. La detección puede significar “explorar” o puede significar comprobar si alguien, con elementos que sugieren cáncer (por ejemplo, síntomas u otros ensayos positivos), tiene cáncer.

Los mecanismos epigenéticos desempeñan un papel importante en el desarrollo embrionario y fetal. Sin embargo, los tejidos embrionarios y fetales humanos (incluyendo tejidos placentarios) no están fácilmente accesibles (Patente de

Estados Unidos 6.927.028). Ciertas realizaciones han abordado este problema analizando una muestra que tiene moléculas de ADN fetal sin células en circulación materna. El metiloma fetal puede deducirse de diversas formas. Por ejemplo, el metiloma de plasma materno puede compararse con un metiloma celular (de células sanguíneas de la madre) y se muestra que la diferencia está correlacionada con el metiloma fetal. Como otro ejemplo, pueden usarse los alelos específicos fetales para determinar la metilación del metiloma fetal en locus específicos. Adicionalmente, el tamaño de un fragmento puede usarse como un indicador de un porcentaje de metilación, ya que se muestra una correlación entre el tamaño y el porcentaje de metilación.

En una realización, se usa secuenciación por bisulfito de todo el genoma para analizar el perfil de metilación (parte de o todo un metiloma) de ADN de plasma materno a 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 usaron diferencias polimórficas, pero puede usarse un diferencial entre el metiloma de plasma y el metiloma de células sanguíneas.

En otra realización, aprovechando variaciones de un único nucleótido y/o aberraciones del número de copias entre un genoma tumoral y un genoma no tumoral, y datos de secuenciación de plasma (u otra muestra), puede realizarse el perfil de metilación de un tumor en la muestra de un paciente que se sospecha o se sabe que tiene cáncer. Una diferencia en un nivel de metilación en una muestra de plasma de un individuo de ensayo en comparación con el nivel de metilación en plasma de un control sano o un grupo de controles sanos puede permitir que se identifique que el individuo de ensayo alberga cáncer. Adicionalmente, el perfil de metilación puede actuar como una identificación que revela el tipo de cáncer, por ejemplo, de qué órgano, que la persona ha desarrollado y si se ha producido metástasis.

Debido a la naturaleza no invasiva de este enfoque, los inventores pudieron evaluar en serie los metilomas de plasma fetal y materno 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 a partir del plasma materno durante el embarazo se asemejó al metiloma placentario. Los genes con impronta y regiones metiladas diferencialmente se identificaron a partir de los datos de plasma materno.

Se ha desarrollado por lo tanto un enfoque para estudiar el metiloma fetal de forma no invasiva, en serie y de forma exhaustiva, ofreciendo de este modo la posibilidad de identificar biomarcadores o ensayo directo de patologías relacionadas con el embarazo. También pueden usarse realizaciones para estudiar el metiloma tumoral de forma no invasiva, en serie y de forma exhaustiva, para explorar o detectar si un sujeto padece cáncer, para supervisar enfermedades malignas en un paciente de cáncer y para pronóstico. Pueden aplicarse realizaciones a cualquier tipo de cáncer, incluyendo, pero sin limitación, cáncer de pulmón, cáncer de mama, cáncer colorrectal, cáncer de próstata, cáncer nasofaríngeo, cáncer gástrico, cáncer testicular, cáncer de piel (por ejemplo melanoma), cáncer que afecta al sistema nervioso, cáncer de hueso, cáncer ovárico, cáncer de hígado (por ejemplo carcinoma hepatocelular), tumores malignos hematológicos, cáncer pancreático, carcinoma endometrial, cáncer de riñón, cáncer del cuello uterino, cáncer de vejiga, etc.

Se analiza en primer lugar una descripción de cómo determinar un metiloma o un perfil de metilación, y después se describen metilomas diferentes (tales como metilomas fetales, un metiloma tumoral, metilomas de la madre o un paciente, y un metiloma mixto, por ejemplo, de plasma). La determinación de un perfil de metilación fetal se describe después usando marcadores específicos fetales o comparando un perfil de metilación mixto con un perfil de metilación celular. Se determinan marcadores de metilación fetales comparando perfiles de metilación. Se analiza una relación entre el tamaño y la metilación. También se proporcionan usos de perfiles de metilación para detectar cáncer.

I. DETERMINACIÓN DE UN METILOMA

Se ha usado una multitud de enfoques para investigar el metiloma placentario, pero cada enfoque tiene sus limitaciones. Por ejemplo, el bisulfito sódico, un producto químico que modifica restos de citosina no metilados a uracilo y deja citosina metilada sin cambios, convierte las diferencias en metilación de citosina en una diferencia de secuencia genética para análisis adicional. El método de patrón oro de estudio de metilación de citosina se basa en el tratamiento de ADN tisular con bisulfito sódico seguido de secuenciación directa de clones individuales de moléculas de ADN convertidas con bisulfito. Después del análisis de múltiples clones de moléculas de ADN, puede obtenerse el patrón de metilación de citosina y perfil cuantitativo por cada sitio de CpG. Sin embargo, la secuenciación por bisulfito clonada es un procedimiento de bajo rendimiento y laborioso que no puede aplicarse fácilmente a una escala de todo el genoma.

Las enzimas de restricción sensibles a metilación que digieren típicamente ADN no metilado proporcionan un enfoque de bajo coste para estudiar la metilación del ADN. Sin embargo, los datos generados a partir de dichos estudios están limitados a locus con los motivos de reconocimiento de enzimas y los resultados no son cuantitativos. Puede usarse inmunoprecipitación de ADN unido con anticuerpos anticitosina metilada para investigar grandes segmentos del genoma, pero tiende a desplazarse hacia locus con metilación densa debido a mayor fuerza de unión del anticuerpo con dichas regiones. Los enfoques basados en micromatrices dependen del diseño a priori de las sondas de consulta y eficacias de hibridación entre las sondas y el ADN diana.

Para consultar un metiloma de forma exhaustiva, algunas realizaciones usan secuenciación paralela masiva (SPM) para proporcionar información de todo el genoma y evaluación cuantitativa del nivel de metilación por nucleótido y por alelo. Recientemente, se ha hecho viable la conversión de bisulfito seguida de SPM de todo el genoma (R Lister *et al.*, 2008 Cell; 133: 523-536).

Entre el número pequeño 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 han aplicado secuenciación de bisulfito de todo el genoma para la investigación de metilomas humanos, dos estudios se han centrado en células madre embrionarias y fibroblastos fetales (R Lister *et al.* Nature; 462: 315-322; L Laurent *et al.*, 2010 Genome Res; 20: 320-331). Ambos estudios analizaron ADN derivado de línea celular.

A. Secuenciación de bisulfito de todo el genoma

Ciertas realizaciones pueden superar los retos anteriormente mencionados y permitir la consulta de un metiloma fetal de forma exhaustiva, no invasiva y en serie. En una realización, se usó secuenciación por bisulfito de todo el genoma para analizar moléculas de ADN fetal sin células que se encuentran en la circulación de mujeres embarazadas. A pesar de la baja abundancia y la naturaleza fragmentada de las moléculas de ADN de plasma, los inventores fueron capaces de ensamblar un metiloma fetal de alta resolución a partir de plasma materno y observar en serie los cambios con la progresión del embarazo. Dado el intenso interés en los ensayos prenatales no invasivos (EPNI), las realizaciones pueden proporcionar una nueva herramienta potente para el descubrimiento de biomarcadores fetales o actuar como una plataforma directa para conseguir EPNI de enfermedades asociadas con el embarazo. Se proporcionan ahora datos de la secuenciación por bisulfito de todo el genoma de diversas muestras, de los que puede obtenerse el metiloma fetal. En una realización, esta tecnología puede aplicarse para perfil de metilación en embarazos complicados con preeclampsia, o retardo del crecimiento intrauterino, o parto prematuro. Para dicho embarazos complicados, esta tecnología puede usarse en serie debido a su naturaleza no invasiva, para permitir la supervisión y/o el pronóstico y/o la respuesta al tratamiento.

La FIGURA 1A muestra una tabla 100 de resultados de secuenciación para sangre materna, placenta y plasma materno de acuerdo con realizaciones de la presente divulgación. En una realización, se realizó secuenciación de genoma completo en bibliotecas de ADN convertidas por bisulfito, preparadas usando adaptadores de bibliotecas de ADN metilado (Illumina) (R Lister *et al.*, 2008 Cell; 133: 523-536), de células sanguíneas de la muestra de sangre recogida en el primer trimestre, la CVS, el tejido placentario recogido a término, las muestras de plasma materno recogidas durante el primer y el tercer trimestre y el periodo posparto. También se analizaron muestras de ADN de células sanguíneas y plasma obtenidas de un hombre adulto y una mujer adulta no embarazada. Se generaron en este estudio un total de 9.500 millones de pares de lecturas de secuencias en bruto. La cobertura de secuenciación de cada muestra se muestra en la tabla 100.

Las lecturas de secuencias que eran mapeables de forma única en el genoma de referencia humano alcanzaron coberturas genómicas haploides promedio de 50 veces, 34 veces y 28 veces, respectivamente, para las muestras de plasma materno del primer trimestre, tercer trimestre y posparto. La cobertura de los sitios de CpG en el genoma varió del 81 % al 92 % para las muestras obtenidas del embarazo. Las lecturas de secuencias que abarcaron sitios de CpG sumaron coberturas haploides promedio de 33 veces por cadena, 23 veces por cadena y 19 veces por cadena, respectivamente, para las muestras de plasma materno del primer trimestre, tercer trimestre y posparto. Las eficacias de conversión por bisulfito para todas las muestras fueron >99,9 % (tabla 100).

En la tabla 100, la tasa ambigua (marcada "a") se refiere a la proporción de lecturas mapeadas en ambas cadenas de Watson y Crick del genoma humano de referencia. La tasa de conversión lambda se refiere a la proporción de citosinas no metiladas en el control de ADN lambda interno que se convierten a los restos de "timina" por modificación con bisulfito. H es igual en general a A, C o T. "a" se refiere a lecturas que podían mapearse en un locus genómico específico, pero no pueden asignarse a la cadena de Watson o Crick. "b" se refiere a lecturas emparejadas con coordenadas de inicio y final idénticas. Para "c", se añadió ADN lambda a cada muestra antes de la conversión por bisulfito. La tasa de conversión lambda se refiere a la proporción de nucleótidos citosina que permanecen como citosina después de la conversión por bisulfito y se usa como un indicio 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 una secuencia de citosina después de conversión por bisulfito.

Durante la modificación por bisulfito, se convierten citosinas desmetiladas en uracilos y posteriormente timinas después de amplificaciones por PCR mientras que las citosinas metiladas permanecen intactas (M Frommer *et al.*, 1992 Proc Natl Acad Sci. USA; 89: 1827-31). Después de la secuenciación y el alineamiento, el estado de metilación de un sitio de CpG individual podría de este modo inferirse a partir del recuento de lecturas de secuencias metiladas "M" (metiladas) y el recuento de lecturas de secuencias desmetiladas "U" (desmetiladas) en el resto de citosina en el contexto de CpG. Usando los datos de secuenciación por bisulfito, se construyeron los metilomas completos de sangre materna, placenta y plasma materno. La densidad de CpG metilado media (también denominada densidad de metilación MD) de locus específicos en el plasma materno puede calcularse usando la ecuación:

$$MD = \frac{M}{M + U}$$

donde M es el recuento de lecturas metiladas y U es el recuento de lecturas no metiladas en los sitios de CpG dentro del locus genético. Si hay más de un sitio de CpG dentro de un locus, entonces M y U corresponden a los recuentos entre los sitios.

B. Diversas técnicas

Como se ha descrito anteriormente, el perfil de metilación puede realizarse usando secuenciación paralela masiva (SPM) de ADN de plasma convertido por bisulfito. La SPM del ADN de plasma convertido por bisulfito puede realizarse de una manera aleatoria o al azar. La profundidad de la secuenciación puede variarse de acuerdo con el tamaño de la región de interés.

En otra realización, la región o las regiones de interés en el ADN de plasma convertido por bisulfito puede capturarse en primer lugar usando un proceso basado en hibridación en fase de solución o en fase sólida, seguido de SPM. La secuenciación en paralelo masiva puede realizarse usando una plataforma de secuenciación por síntesis tal como la Illumina, una plataforma de secuenciación por ligamiento tal como la plataforma SOLiD de Life Technologies, un sistema de secuenciación basado en semiconductores tal como las plataformas Ion Torrent o Ion Proton de Life Technologies, o sistema de secuenciación de una única molécula tal como el sistema Helicos o el sistema de Pacific Biosciences o un sistema de secuenciación basado en nanoporos. La secuenciación basada en nanoporos incluye nanoporos que se construyen usando, por ejemplo, bicapas lipídicas y nanoporos de proteínas y nanoporos de estado sólido (tales como los que se basan en grafeno). Como se seleccionó las plataformas de secuenciación de una única molécula permitirían dilucidar el estado de metilación de moléculas de ADN (incluyendo N6-metiladenina, 5-metilcitosina y 5-hidroximetilcitosina) directamente sin conversión por bisulfito (BA 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 de ADN de muestra no convertido por bisulfito (por ejemplo, ADN de plasma).

Además de secuenciación, pueden usarse otras técnicas. En una realización, puede realizarse el perfil de metilación mediante PCR específica de metilación o digestión con enzimas de restricción sensibles a 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 única molécula o digital (B Vogelstein *et al.*, 1999 Proc Natl Acad Sci USA; 96: 9236-9241). En más realizaciones adicionales, la PCR puede ser una PCR en tiempo real. En otras realizaciones, la PCR puede ser PCR múltiple.

II. ANÁLISIS DE METILOMAS

Algunas realizaciones pueden determinar el perfil de metilación de ADN de plasma usando secuenciación por bisulfito de todo el genoma. El perfil de metilación de un feto puede determinarse por secuenciación de muestras de ADN de plasma materno, como se describe posteriormente. Por lo tanto, se accedió a las moléculas de ADN fetal (y metiloma fetal) de forma no invasiva durante el embarazo, y se supervisaron los cambios en serie a medida que progresó el embarazo. Debido a la exhaustividad de los datos de secuenciación, se pudieron estudiar los metilomas de plasma materno a una escala de todo el genoma a una resolución de un solo nucleótido.

Ya que se conocían las coordenadas genómicas de las lecturas secuenciadas, estos datos permitieron estudiar los niveles de metilación generales del metiloma o cualquier región de interés en el genoma y realizar comparaciones entre diferentes elementos genéticos. Además, múltiples lecturas de secuencias abarcaron cada sitio o locus de CpG. Se proporciona ahora una descripción de algunas de las medidas usadas para medir el metiloma.

A. Metilación de moléculas de ADN de plasma

Las moléculas de ADN están presentes en plasma humano a concentraciones bajas y en forma fragmentada, típicamente en longitudes que se asemejan a unidades mononucleosómicas (YMD Lo *et al.* 2010 Sci Transl Med; 2: 61ra91; e YW Zheng *et al.* 2012 Clin Chem; 58: 549-558). A pesar de estas limitaciones, un proceso de secuenciación por bisulfito de todo el genoma fue capaz de analizar la metilación de las moléculas de ADN de plasma. En otras realizaciones más, ya que las plataformas de secuenciación de moléculas individuales seleccionadas permitirían dilucidar el estado de metilación de moléculas de ADN directamente sin conversión por bisulfito (BA 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 usar el ADN de plasma no convertido por bisulfito para determinar los niveles de metilación de ADN de plasma o determinar el metiloma de plasma. Dichas plataformas pueden detectar N6-metiladenina, 5-metilcitosina y 5-hidroximetilcitosina, que pueden proporcionar resultados mejorados (por ejemplo, sensibilidad o especificidad mejoradas) en relación con las diferentes funciones biológicas de las diferentes formas de metilación. Dichos resultados mejorados pueden ser útiles cuando se aplican realizaciones para la detección o supervisión de trastornos específicos, por ejemplo, preeclampsia o un tipo de cáncer particular.

La secuenciación por bisulfito también puede diferenciar entre diferentes formas de metilación. En una realización, se

pueden incluir etapas adicionales que puedan diferenciar 5-metilcitosina de 5-hidroximetilcitosina. Uno de dichos enfoques es la secuenciación por bisulfito oxidativa (oxBS-seq), que puede dilucidar la localización de 5-metilcitosina y 5-hidroximetilcitosina a una resolución de una sola base (MJ Booth *et al.* 2012 Science; 336: 934-937; MJ Booth *et al.* 2013 Nature Protocols; 8: 1841-1851). En secuenciación por bisulfito, tanto 5-metilcitosina como 5-hidroximetilcitosina se leen como citosinas y por lo tanto no pueden diferenciarse. Por otro lado, en oxBS-seq, la oxidación específica de 5-hidroximetilcitosina a 5-formilcitosina por tratamiento con perrutenato de potasio (KRuO₄), seguido de la conversión de la 5-formilcitosina de nueva formación a uracilo usando conversión por bisulfito permitiría diferenciar 5-hidroximetilcitosina de 5-metilcitosina. Por lo tanto, puede obtenerse una lectura de 5-metilcitosina a partir de un único ciclo de oxBS-seq y los niveles de 5-hidroximetilcitosina se deducen por comparación con los resultados de secuenciación por bisulfito. En otra realización, puede distinguirse 5-metilcitosina de 5-hidroximetilcitosina usando secuenciación por bisulfito asistida por Tet (TAB-seq) (M Yu *et al.*, 2012 Nat Protoc; 7: 2159-2170). TAB-seq puede identificar 5-hidroximetilcitosina a una resolución de una sola base, así como determinar su abundancia en cada sitio de modificación. Este método implica protección mediada por β -glucosiltransferasa de 5-hidroximetilcitosina (glucosilación) y oxidación mediada por Tet1 de ratón (mTet1) recombinante de 5-metilcitosina a 5-carboxilcitosina. Después del tratamiento por bisulfito posterior y la amplificación por PCR, tanto citosina como 5-carboxilcitosina (derivada de 5-metilcitosina) se convierten a timina (T), mientras que 5-hidroximetilcitosina se leerá como C.

La FIGURA 1B muestra la densidad de metilación en ventanas de 1 Mb de muestras secuenciadas según realizaciones de la presente divulgación. La representación 150 es una representación de Circos que representa la densidad de metilación en el plasma materno y ADN genómico en ventanas de 1 Mb por todo el genoma. Desde fuera hacia dentro: los ideogramas de cromosomas pueden orientarse pter-qter en el sentido de las agujas del reloj (los centrómeros se muestran en rojo), sangre materna (rojo), placenta (amarillo), plasma materno (verde), lecturas compartidas en plasma materno (azul) y lecturas específicas fetales en plasma materno (morado). Los niveles de metilación de CpG generales (es decir, niveles de densidad) de células sanguíneas maternas, placenta y plasma materno pueden encontrarse en la tabla 100. El nivel de metilación de células sanguíneas maternas es en general mayor que el de la placenta en todo el genoma completo.

B. Comparación de secuenciación por bisulfito con otras técnicas

Se estudió el metiloma placentario usando secuenciación por bisulfito en paralelo masiva. Además, se estudió el metiloma placentario usando una plataforma de matriz de oligonucleótidos que abarcaba aproximadamente 480.000 sitios de CpG en el genoma humano (Illumina) (M Kulis *et al.* 2012 Nat Genet; 44: 1236-1242; y C Clark *et al.* 2012 PLoS One; 7: e50233). En una realización usando genotipado basado en beadchip y análisis de metilación, se realizó genotipado usando la matriz de genotipado Illumina HumanOmni2.5-8 de acuerdo con el protocolo del fabricante. Los genotipos se deciden usando el algoritmo GenCall del software Genome Studio (Illumina). Las tasas de decisión fueron de más del 99 %. Para el análisis de metilación basado en micromatrices, se trató ADN genómico (500–800 ng) con bisulfito sódico usando 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 μ l de ADN genómico convertido por bisulfito a 50 ng/ μ l de acuerdo con el protocolo de Ensayo de Metilación Infinium HD. La beadchip hibridada se exploró en un instrumento iScan de Illumina. Se analizaron los datos de metilación de ADN por el software de Módulo de Metilación (v1.9.0) de GenomeStudio (v2011.1), con normalización con respecto a controles internos y resta de fondos. El índice de metilación para sitio de CpG individual estuvo representado por un valor beta (β), que se calculó usando la relación de intensidades de fluorescencia entre alelos metilados y desmetilados:

$$\beta = \frac{\text{Intensidad de alelo metilado}}{\text{Intensidad de alelo desmetilado} + \text{Intensidad de alelo metilado} + 100}$$

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

Las FIGURAS. 2A-2C muestran representaciones de los valores beta determinados por la matriz de beadchip de 450K de Metilación Humana Infinium Illumina frente a los índices de metilación determinados por secuenciación por bisulfito de todo el genoma de sitios de CpG correspondientes que se consultaron para ambas plataformas: (A) Células sanguíneas maternas, (B) Muestra de la vellosoidad coriónica, (C) Tejido placentario a término. Los datos de ambas plataformas fueron altamente concordantes y los coeficientes de correlación de Pearson fueron de 0,972, 0,939 y 0,954, y los valores de R² fueron de 0,945, 0,882 y 0,910 para las células sanguíneas maternas, CVS y tejido placentario a término, respectivamente.

Se compararon además los datos de secuenciación de los inventores con los presentados por Chu *et al.*, que investigaron los perfiles de metilación de 12 pares de CVS y muestras de ADN de células sanguíneas maternas usando

una matriz de oligonucleótidos que abarcó aproximadamente 27.000 sitios de CpG (T Chu *et al.* 2011 PLoS One; 6: e14723). Los datos de correlación entre los resultados de secuenciación de la CVS y ADN de células sanguíneas maternas y cada uno de los 12 pares de muestras en el estudio previo proporcionó un coeficiente de Pearson promedio (0,967) y R^2 (0,935) para sangre materna y un coeficiente de Pearson promedio (0,943) y R^2 (0,888) para la CVS.

Entre los sitios de CpG representados en ambas matrices, los datos de los inventores se correlacionaron en alta medida con los datos publicados. Las tasas de metilación distinta de CpG fueron de <1 % para las células sanguíneas maternas, CVS y tejidos placentarios (tabla 100). Estos resultados fueron coherentes con la creencia actual de que cantidades sustanciales de metilación distinta de CpG se restringieron principalmente a células pluripotenciales (R Lister *et al.*, 2009 Nature; 462: 315-322; L Laurent *et al.*, 2010 Genome Res; 20: 320-331).

C. Comparación de metilomas de plasma y sangre para personas no embarazadas

Las FIGURAS 3A y 3B muestran diagramas de barras del porcentaje de sitios de CpG metilados en plasma y células sanguíneas recogidos de un hombre adulto y una mujer adulta no embarazada: (A) Autosomas, (B) Cromosoma X. Los diagramas muestran una similitud entre los metilomas de plasma y sangre de un hombre y una mujer no embarazada. Las proporciones generales de sitios de CpG que estaban metilados en las muestras de plasma del hombre y de la mujer no embarazada fueron casi iguales al ADN de células sanguíneas correspondientes (tabla 100 y FIGURAS 2A y 2B).

A continuación, se estudió la correlación de los perfiles de metilación de las muestras de plasma y células sanguíneas de una manera específica de locus. Se determinó la densidad de metilación de cada grupo de 100 kb en el genoma humano determinando el número total de citosinas no convertidas en sitios de CpG como una proporción de todos los sitios de CpG abarcados por lecturas de secuencias mapeadas en la región de 100 kb. Las densidades de metilación fueron altamente concordantes entre la muestra de plasma y ADN de células sanguíneas correspondiente de las muestras del hombre, así como de la mujer.

Las FIGURAS 4A y 4B muestran representaciones de densidades de metilación de locus correspondientes en ADN de células sanguíneas y ADN de plasma: (A) Mujer adulta no embarazada, (B) Hombre adulto. El coeficiente de correlación de Pearson y el valor de R^2 para las muestras de mujer no embarazada fueron respectivamente 0,963 y 0,927, y para las muestras del hombre fueron respectivamente 0,953 y 0,908. Estos datos son coherentes con hallazgos previos basados en la evaluación de genotipos de moléculas de ADN de plasma de receptores de trasplante de células madre hematopoyéticas alogénico que mostraron 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 entre metilomas

A continuación, se estudiaron los niveles de metilación de ADN de plasma materno, células sanguíneas maternas y tejido placentario para determinar los niveles de metilación. Los niveles se determinaron para regiones repetidas, regiones no repetidas, y generales.

Las FIGURAS 5A y 5B muestran diagramas de barras del porcentaje de sitios de CpG metilados entre muestras recogidas del embarazo: (A) Autosomas, (B) Cromosoma X. Las proporciones generales de CpG metilados fueron de 67,0 % y 68,2 % para las muestras de plasma materno del primer y tercer trimestre, respectivamente. A diferencia de los resultados obtenidos de las personas no embarazadas, estas proporciones fueron menores que las de la muestra de células sanguíneas maternas del primer trimestre, pero mayores que las de la CVS y muestras de tejido placentario a término (tabla 100). Debe observarse que el porcentaje de CpG metilados para la muestra de plasma materno posparto fue del 73,1 % lo que fue similar a los datos de células sanguíneas (tabla 100). Estas tendencias se observaron en CpG distribuidos por todos los autosomas, así como en el cromosoma X y abarcaron tanto las regiones no repetidas como múltiples clases de elementos repetidos del genoma humano.

Se descubrió que los elementos tanto repetidos como no repetidos en la placenta estaban hipometilados en relación con células sanguíneas maternas. Los resultados fueron concordantes con los hallazgos en la bibliografía de que la placenta está hipometilada en relación con otros tejidos, incluyendo células sanguíneas periféricas.

Entre el 71 % y el 72 % de los sitios de CpG secuenciados estaban metilados en el ADN de células sanguíneas de la mujer embarazada, la mujer no embarazada y el hombre adulto (tabla 100 de la FIGURA 1). Estos datos son comparables con el informe del 68,4 % de sitios de CpG de células mononucleares sanguíneas presentado por Y Li *et al.* 2010 PLoS Biol; 8: e1000533. De forma coherente con los informes previos sobre la naturaleza hipometilada de tejidos placentarios, el 55 % y el 59 % de los sitios de CpG estaban metilados en la CVS y el tejido placentario a término, respectivamente (tabla 100).

La FIGURA 6 muestra un diagrama de barras del nivel de metilación de diferentes clases repetidas del genoma humano para sangre materna, placenta y plasma materno. Las clases repetidas son como se definen por el buscador de genoma de UCSC. Los datos mostrados son para las muestras del primer trimestre. A diferencia de datos anteriores que sugieren que la naturaleza hipometilada de tejidos placentarios se observó principalmente en ciertas clases repetidas en el genoma (B Novakovic *et al.*, 2012 Placenta; 33: 959-970), aquí se muestra que la placenta estaba de

hecho hipometilada en la mayoría de las clases de elementos genómicos en referencia a células sanguíneas.

E. Similitud de metilomas

- 5 Las realizaciones pueden determinar los metilomas de tejidos placentarios, células sanguíneas y plasma usando la misma plataforma. Por lo tanto, fueron posibles comparaciones directas de los metilomas de esos tipos de muestras biológicas. El alto nivel de semejanza entre metilomas de las células sanguíneas y plasma para el hombre y la mujer no embarazada, así como entre las células sanguíneas maternas y la muestra de plasma materno posparto afirmaron adicionalmente que las células hematopoyéticas eran las principales fuentes de ADN en plasma humano (YW Zheng
10 *et al.* 2012 Clin Chem; 58: 549-558).

Las semejanzas son evidentes tanto con respecto a la proporción general de CpG metilados en el genoma, así como de la alta correlación de densidades de metilación entre locus correspondientes en el ADN de células sanguíneas y ADN del plasma. No obstante, las proporciones generales de CpG metilados en las muestras de plasma materno del
15 primer trimestre y tercer trimestre se redujeron en comparación con los datos de células sanguíneas maternas o la muestra de plasma materno posparto. Los niveles de metilación reducidos durante el embarazo se debieron a la naturaleza hipometilada de las moléculas de ADN fetal presentes en el plasma materno.

La inversión del perfil de metilación en la muestra de plasma materno posparto hasta hacerse más similar al de las
20 células sanguíneas maternas sugiere que las moléculas de ADN fetal se habían retirado de la circulación materna. El cálculo de las concentraciones de ADN fetal basado en marcadores de SNP del feto mostraron de hecho que la concentración cambió de 33,9 % antes del parto a solamente 4,5 % en la muestra posparto.

F. Otras aplicaciones

25 Las realizaciones han ensamblado con éxito metilomas de ADN mediante el análisis de SPM de ADN de plasma. La capacidad de determinar el metiloma placentario o fetal a partir del plasma materno proporciona un método no invasivo para determinar, detectar y supervisar los perfiles de metilación aberrante asociados con afecciones asociadas al embarazo tales como preeclampsia, restricción del crecimiento intrauterino, parto prematuro y otros. Por ejemplo, la
30 detección de una identificación de metilación aberrante específica de enfermedad permite la exploración, el diagnóstico y la supervisión de dichas afecciones asociadas a embarazo. La medición del nivel de metilación de plasma materno permite la exploración, el diagnóstico y la supervisión de dichas afecciones asociadas a embarazo. Además de las aplicaciones directas en la investigación de afecciones asociadas a embarazo, el enfoque podría aplicarse a otras áreas de la medicina en las que es interesante el análisis de ADN de plasma. Por ejemplo, los metilomas de cánceres
35 pudieron determinarse a partir de ADN de plasma de pacientes con cáncer. El análisis metilómico de cáncer a partir del plasma, como se describe en el presente documento, es potencialmente una tecnología sinérgica con el análisis genómico del cáncer a partir del plasma (KCA Chan *et al.* 2013 Clin Chem; 59: 211-224 y RJ Leary y col., 2012 Sci Transl Med; 4: 162ra154).

40 Por ejemplo, la determinación de un nivel de metilación de una muestra de plasma podría usarse para explorar con respecto al cáncer. Cuando el nivel de metilación de la muestra de plasma muestra niveles aberrantes en comparación con controles sanos, puede sospecharse cáncer. Después puede realizarse confirmación y evaluación adicional del tipo de cáncer u origen tisular del cáncer determinando el perfil de plasma de metilación en diferentes locus genómicos o mediante análisis genómico del plasma para detectar aberraciones en el número de copias asociadas con tumor,
45 translocaciones cromosómicas y variantes de un único nucleótido. De hecho, en una realización de la presente divulgación, el perfil metilómico y genómico del cáncer de plasma puede llevarse a cabo de forma simultánea. Como alternativa, podrían usarse investigaciones radiológicas y de captura de imágenes (por ejemplo, tomografía computarizada, imágenes por resonancia magnética, tomografía de emisión de positrones) o endoscopia (por ejemplo, endoscopia gastrointestinal superior o colonoscopia) para investigar adicionalmente individuos que se sospecha que
50 tienen cáncer basándose en el análisis del nivel de metilación en plasma.

Para exploración o detección de cáncer, la determinación de un nivel de metilación de una muestra de plasma (u otra biológica) puede usarse junto con otras modalidades para exploración o detección de cáncer tal como medición del
55 antígeno específico de próstata (por ejemplo, para cáncer de próstata), antígeno carcinoembrionario (por ejemplo para carcinoma colorrectal, carcinoma gástrico, carcinoma pancreático, carcinoma de pulmón, carcinoma de mama, carcinoma de tiroides medular), fetoproteína alfa (por ejemplo para cáncer de hígado o tumores de células germinales), CA125 (por ejemplo para cáncer ovárico y de mama) y CA19-9 (por ejemplo para carcinoma pancreático).

Adicionalmente, otros tejidos pueden secuenciarse para obtener un metiloma celular. Por ejemplo, el tejido hepático
60 puede analizarse para determinar un patrón de metilación específico para el hígado, que puede usarse para identificar patologías de hígado. Otros tejidos que pueden analizarse también incluyen células cerebrales, huesos, los pulmones, el corazón, los músculos y los riñones, etc. Los perfiles de metilación de diversos tejidos pueden cambiar de vez en cuando, por ejemplo, como resultado del desarrollo, el envejecimiento, procesos de enfermedad (por ejemplo, inflamación o cirrosis o procesos autoinmunitarios (tal como en lupus eritematoso sistémico)) o tratamiento (por
65 ejemplo, tratamiento con agentes desmetilantes tales como 5-azacitidina y 5-azadesoxicitidina). La naturaleza dinámica de la metilación de ADN hace a dicho análisis potencialmente muy valioso para supervisar los procesos

fisiológicos y patológicos. Por ejemplo, si se detecta un cambio en el metiloma de plasma de un individuo en comparación con un valor de línea basal obtenido cuando estaban sanos, se podrían entonces detectar procesos de enfermedad en órganos que aportan ADN de plasma.

Además, los metilomas de órganos trasplantados podrían determinarse a partir de ADN de plasma de receptores de trasplante de órganos. El análisis metilómico de trasplante de plasma, como se describe en esta divulgación, es potencialmente una tecnología sinérgica con el análisis genómico del trasplante de plasma (YW Zheng *et al.*, 2012; YMD Lo *et al.*, 1998 Lancet; 351: 1329-1330; y TM Snyder *et al.* 2011 Proc Natl Acad Sci USA; 108: 6229-6234). Como el ADN de plasma se considera en general un marcador de muerte celular, un aumento en el nivel en plasma de ADN liberado de un órgano trasplantado podría usarse como un marcador para muerte celular aumentada de ese órgano, tal como un episodio de rechazo u otros procesos patológicos que implican a ese órgano (por ejemplo, infección o absceso). En el caso de que se instituya con éxito terapia antirrechazo, se espera que se reduzca el nivel de plasma de ADN liberado por el órgano trasplantado.

III. DETERMINACIÓN DE METILOMA FETAL O TUMORAL USANDO SNP

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

En una realización, se usan diferencias de polimorfismos de un único nucleótido (SNP) entre la madre y el feto para identificar las moléculas de ADN fetal en plasma materno. Un objetivo fue identificar locus de SNP en los que la madre es homocigótica, pero el feto es heterocigoto; el alelo específico fetal puede usarse para determinar qué fragmentos de ADN son del feto. Se analizó ADN genómico de las células sanguíneas maternas usando una matriz de genotipado de SNP, la Illumina HumanOmni2.5-8. Por otro lado, para locus de SNP en los que la madre es heterocigota y el feto es homocigoto, entonces el alelo de SNP que es específico para la madre puede usarse para determinar qué fragmentos de ADN de plasma son 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 lecturas específicas fetales y metiloma placentario

Se identificaron locus que tienen dos alelos diferentes, en los que la cantidad de un alelo (B) era significativamente menor que la del otro alelo (A), a partir de resultados de secuenciación de una muestra biológica. Las lecturas que abarcaban los alelos B se consideraron específicas fetales (lecturas específicas fetales). Se determina que la madre es homocigota para A y el feto heterocigoto para A/B, y por lo tanto las lecturas que abarcan el alelo A fueron compartidas por la madre y el feto (lecturas compartidas).

En un caso de embarazo analizado que se usó para ilustrar varios de los conceptos de esta divulgación, se descubrió que la madre embarazada era homocigota en 1.945.516 locus en los autosomas. Se inspeccionaron las lecturas de secuenciación de ADN de plasma materno que abarcaban estos SNP. Se detectaron lecturas que portaban un alelo no materno en 107.750 locus y estos se consideraron locus informativos. En cada SNP informativo, el alelo que no era de la madre se denominó alelo específico fetal mientras que el otro se denominó alelo compartido.

Puede determinarse una concentración de ADN fetal/tumoral fraccional (también denominada porcentaje de ADN fetal) en el plasma materno. En una realización, la concentración de ADN fetal fraccional en el plasma materno, f , se determina por la ecuación:

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

donde p es el número de lecturas secuenciadas con el alelo específico fetal 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 descubrió que las proporciones de ADN fetal en las muestras de plasma materno del primer trimestre, tercer trimestre y posparto eran del 14,4 %, 33,9 % y 4,5 %, respectivamente. Las proporciones de ADN fetal también se calcularon usando los números de lecturas que se alineaban con el cromosoma Y. Basándose en los datos del cromosoma Y, los resultados fueron de 14,2 %, 34,9 % y 3,7 %, respectivamente, en las muestras de plasma materno del primer trimestre, tercer trimestre y posparto.

Analizando por separado estas lecturas de secuencias específicas fetales o compartidas, las realizaciones demuestran que las moléculas de ADN fetal en circulación estaban mucho más hipometiladas que las moléculas de ADN de fondo. Las comparaciones de las densidades de metilación de locus correspondientes en las lecturas de plasma materno específicas fetales y los datos de tejido placentario tanto para el primer como para el tercer trimestre revelaron altos niveles de correlación. Estos datos proporcionaron pruebas del nivel de genoma de que la placenta es la fuente

predominante de moléculas de ADN derivadas del feto en plasma materno y representaron un avance importante en comparación con pruebas previas basándose en información derivada de locus seleccionados.

Se determinó la densidad de metilación de cada región de 1 Mb en el genoma usando las lecturas específicas fetales o compartidas que abarcaban sitios de CpG adyacentes a los SNP informativos. Los metilomas específicos fetales y no específicos fetales ensamblados a partir de lecturas de secuencias de plasma materno pueden presentarse, por ejemplo, en representaciones de Circos (M Krzywinski *et al*, 2009 Genome Res; 19: 1639-1645). Las densidades de metilación por cada grupo de 1 Mb también se determinaron para las células sanguíneas maternas y muestras de tejido placentario.

La FIGURA 7A muestra una representación de Circos 700 para muestras del primer trimestre. La FIGURA 7B muestra una representación de Circos 750 para muestras del tercer trimestre. Las representaciones 700 y 750 muestran densidad de metilación por cada grupo de 1 Mb. Se orientan ideogramas de cromosomas (anillo más externo) pter- qter en el sentido de las agujas del reloj (los centrómeros se muestran en rojo). La segunda pista más externa muestra el número de sitios de CpG en las regiones de 1 Mb correspondientes. La escala de las barras rojas mostradas es de hasta 20.000 sitios por cada grupo de 1 Mb. Las densidades de metilación de las regiones de 1 Mb correspondientes se muestran en las otras pistas basándose en el esquema de color mostrado en el centro.

Para las muestras del primer trimestre (FIGURA 7A), de dentro a fuera, las pistas son: muestra de vellosidad coriónica, lecturas específicas fetales en plasma materno, lecturas específicas maternas en plasma materno, lecturas fetales y no fetales combinadas en plasma materno y células sanguíneas maternas. Para las muestras del tercer trimestre (FIGURA 7B), las pistas son: tejido placentario a término, lecturas específicas fetales en plasma materno, lecturas específicas maternas en plasma materno, lecturas fetales y no fetales combinadas en plasma materno, plasma materno posparto y células sanguíneas maternas (de la muestra de sangre del primer trimestre). Puede apreciarse que, para las muestras de plasma tanto del primer como del tercer trimestre, los metilomas fetales estaban más hipometilados que los de los metilomas no específicos fetales.

El perfil de metilación general de los metilomas fetales se asemejó más estrechamente al de las muestras del tejido placentario o CVS. Por el contrario, el perfil de metilación de ADN de las lecturas compartidas en el plasma, que era predominantemente ADN materno, se asemejó más estrechamente al de las células sanguíneas maternas. Se realizó después una comparación locus a locus sistemática de las densidades de metilación de las lecturas de ADN del plasma materno y los tejidos maternos o fetales. Se determinaron las densidades de metilación de sitios de CpG que estaban presentes en la misma lectura de secuencia que los SNP informativos y estuvieron abarcados por al menos 5 lecturas de secuencias de ADN de plasma materno.

Las FIGURAS 8A-8D muestran representaciones de comparaciones de las densidades de metilación de ADN de tejido genómico frente a ADN de plasma materno para sitios de CpG que rodean los polimorfismos de un único nucleótido informativos. La FIGURA 8A muestra densidades de metilación para lecturas específicas fetales en la muestra de plasma materno del primer trimestre en relación con densidades de metilación para lecturas en una muestra de CVS. Como puede verse, los valores específicos fetales se corresponden bien con los valores de CVS.

La FIGURA 8B muestra densidades de metilación para lecturas específicas fetales en la muestra de plasma materno del tercer trimestre en relación con 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 fetales.

La FIGURA 8C muestra densidades de metilación para lecturas compartidas en la muestra de plasma materno del primer trimestre en relación con densidades de metilación para lecturas en células sanguíneas maternas. Dado que la mayoría de las lecturas compartidas son de la madre, los dos conjuntos de valores se corresponden bien. La FIGURA 8D muestra densidades de metilación para lecturas compartidas en la muestra de plasma materno del tercer trimestre en relación con densidades de metilación para lecturas en células sanguíneas maternas.

Para las lecturas específicas fetales en plasma materno, el coeficiente de correlación de Spearman entre el plasma materno del primer trimestre y la CVS fue de 0,705 ($P < 2,2 \times 10^{-16}$); y entre el plasma materno del tercer trimestre y el tejido placentario a término fue de 0,796 ($P < 2,2 \times 10^{-16}$) (FIGURAS 8A y 8B). Se realizó una comparación similar para las lecturas compartidas en plasma materno con los datos de células sanguíneas maternas. 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 fue de 0,638 ($P < 2,2 \times 10^{-16}$) para la muestra de plasma del tercer trimestre (FIGURAS 8C y 8D).

B. Metiloma fetal

En una realización, para ensamblar el metiloma fetal a partir del plasma materno, se clasificó con respecto a lecturas de secuencias que abarcaban al menos un sitio de SNP fetal informativo y contenían al menos un sitio de CpG dentro de la misma lectura. Se incluyeron lecturas que mostraban los alelos específicos fetales en el ensamblaje del metiloma fetal. Se incluyeron lecturas que mostraban el alelo compartido, es decir alelo no específico fetal, en el ensamblaje del metiloma no específico fetal que estaba comprendido predominantemente por moléculas de ADN derivadas de la

madre.

Las lecturas específicas fetales abarcaban 218.010 sitios de CpG en los autosomas para las muestras de plasma materno del primer trimestre. Las cifras correspondientes para las muestras de plasma materno del tercer trimestre y posparto fueron de 263.611 y 74.020, respectivamente. En promedio, las lecturas compartidas abarcaron los sitios de CpG un promedio de 33,3, 21,7 y 26,3 veces, respectivamente. Las lecturas específicas fetales abarcaron esos sitios de CpG 3,0, 4,4 y 1,8 veces, respectivamente, para las muestras de plasma materno del primer trimestre, tercer trimestre y posparto.

El ADN fetal representa una población menor en el plasma materno y por lo tanto la cobertura de esos sitios de CpG por lecturas específicas fetales fue proporcional al porcentaje de ADN fetal de la muestra. Para la muestra de plasma materno del primer trimestre, el porcentaje general de CpG metilado 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 metilado 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 fetales en el plasma materno estaban más hipometiladas que las lecturas compartidas en el plasma materno.

C. Método

Las técnicas descritas anteriormente también pueden usarse para determinar un perfil de metilación tumoral. Se describen ahora métodos para determinar los perfiles de metilación fetal y tumoral.

La FIGURA 9 es un diagrama de flujo que ilustra un método 900 para determinar un primer perfil de metilación de una muestra biológica de un organismo de acuerdo con realizaciones de la presente divulgación. El Método 900 puede construir un mapa epigenético del feto a partir del perfil de metilación del plasma materno. La muestra biológica incluye ADN sin células que comprende una mezcla de ADN sin células que se origina de un primer tejido y de un segundo tejido. Como ejemplos, el primer tejido puede ser de un feto, un tumor o un órgano trasplantado.

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 determinar 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.

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

En el bloque 920, se identifica una pluralidad de primeros locus en los que un primer genoma del primer tejido es heterocigoto para un primer alelo respectivo y un segundo alelo respectivo y un segundo genoma del segundo tejido es homocigoto para el primer alelo respectivo. Por ejemplo, las lecturas específicas fetales pueden identificarse en la pluralidad de primeros locus. Por otro lado, las lecturas específicas de tumor pueden identificarse en la pluralidad de primeros locus. Las lecturas específicas de tejido pueden identificarse a partir de lecturas de secuenciación en las que el porcentaje de lecturas de secuencias del segundo alelo queda dentro de un intervalo particular, por ejemplo, de aproximadamente 3 %-25 %, indicando de este modo una población minoritaria de fragmento de ADN de un genoma heterocigoto en el locus y una población mayoritaria de un genoma homocigoto en el locus.

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

El locus puede definirse por un número específico de sitios, un conjunto específico de sitios, o un tamaño particular para una región alrededor de una variación que comprende el alelo específico de tejido. Un locus puede tener solamente un sitio. Los sitios pueden tener propiedades específicas, por ejemplo, que sean sitios de CpG. La determinación de un número de lecturas que están desmetiladas es equivalente, y está abarcado 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 basándose en los números de moléculas de ADN metiladas en los uno o más sitios del locus y correspondientes al segundo alelo respectivo del locus. Por ejemplo, puede determinarse una densidad de metilación para sitios de CpG correspondientes a un locus.

En el bloque 950, el primer perfil de metilación del primer tejido se crea a partir de las densidades de metilación para

los primeros locus. El primer perfil de metilación puede corresponder a sitios particulares, por ejemplo, sitios de CpG. El perfil de metilación puede ser para todos los locus que tienen un alelo específico fetal, o solamente algunos de esos locus.

5 IV. USO DE LA DIFERENCIA DE METILOMAS DE PLASMA Y SANGRE

Anteriormente se ha mostrado que las lecturas específicas fetales del plasma se correlacionan con el metiloma placentario. Como el componente materno del metiloma de plasma materno es aportado principalmente por las células sanguíneas, la diferencia entre el metiloma de plasma y el metiloma de sangre puede usarse para determinar el metiloma placentario para todos los locus, y no solamente localizaciones de alelos específicos fetales. Una diferencia entre el metiloma de plasma y el metiloma de sangre también puede usarse para determinar un metiloma de un tumor.

A. Método

La FIGURA 10 es un diagrama de flujo que ilustra un método 1000 para determinar un primer perfil de metilación de una muestra biológica de un organismo de acuerdo con realizaciones de la presente divulgación. La muestra biológica (por ejemplo, plasma) incluye ADN sin células que comprende una mezcla de ADN sin células que se origina de un primer tejido y de un segundo tejido. El primer perfil de metilación corresponde a un perfil de metilación del primer tejido (por ejemplo, tejido fetal o tejido tumoral). El método 1200 puede proporcionar una deducción de regiones metiladas diferencialmente de plasma materno.

En el bloque 1010, se recibe una muestra biológica. La muestra biológica podría recibirse simplemente en una máquina (por ejemplo, una máquina de secuenciación). La muestra biológica puede estar en una forma tomada del organismo o puede estar en una forma procesada, por ejemplo, la muestra puede ser plasma que se extrae de una muestra sanguínea.

En el bloque 1020, se obtiene un segundo perfil de metilación correspondiente a ADN del segundo tejido. El segundo perfil de metilación podría leerse de la memoria, ya que puede haberse determinado previamente. El segundo perfil de metilación puede determinarse a partir del segundo tejido, por ejemplo, una muestra diferente que contiene sola o predominantemente células del segundo tejido. El segundo perfil de metilación puede corresponder a un perfil de metilación celular y obtenerse de ADN celular. Como otro ejemplo, el segundo perfil puede determinarse a partir de una muestra de plasma recogida antes del embarazo, o antes del desarrollo de cáncer debido a que el metiloma de plasma de una persona no embarazada sin cáncer es muy similar al metiloma de células sanguíneas.

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 particular 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, en la que se usan sitios de CpG asociados con el locus 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 los dos valores son complementarios.

En una realización, el segundo perfil de metilación se obtiene realizando secuenciación sensible a la metilación de ADN celular a partir de una muestra del organismo. Un ejemplo de secuenciación sensible a la metilación incluye tratar ADN con bisulfito sódico y después realizar secuenciación de ADN. En otro ejemplo, la secuenciación sensible a la metilación puede realizarse sin usar bisulfito sódico, usando una plataforma de secuenciación de una única molécula que permitiría dilucidar el estado de metilación de las moléculas de ADN (incluyendo N6-metiladenina, 5-metilcitosina y 5-hidroximetilcitosina) directamente sin conversión por bisulfito (AB 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 citosina metilada (por ejemplo usando un anticuerpo contra metilcitosina o usando una proteína o un péptido de unión a ADN metilado (LG Acevedo *et al.*, 2011 Epigenomics; 3: 93-101) seguido de secuenciación; o mediante el uso de enzimas de restricción sensibles a metilación seguido de secuenciación. En otra realización, se usan técnicas sin secuenciación, tales como matrices, PCR digital y espectrometría de masas.

En otra realización, la segunda densidad de metilación del segundo tejido podría obtenerse previamente 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 tiene densidades de metilación de referencia. Las densidades de metilación de referencia pueden determinarse a partir de múltiples muestras, en las que un nivel medio (u otro valor estadístico) de las diferentes densidades de metilación en un locus pueden usarse como la densidad de metilación de referencia en el locus.

En el bloque 1030, se determina un perfil de metilación sin células a partir del ADN sin células de la mezcla. El perfil de metilación sin células proporciona una densidad de metilación en cada una de la pluralidad de locus. El perfil de metilación sin células puede determinarse recibiendo lecturas de secuencias de una secuenciación del ADN sin células, en el que la información de metilación se obtiene de las lecturas de secuencias. El perfil de metilación sin células puede determinarse de la misma manera que el metiloma celular.

En el bloque 1040, se determina un porcentaje del ADN sin células 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 diversas formas, por ejemplo, usando un alelo específico fetal o un alelo específico tumoral. También puede usarse el número de copia para determinar el porcentaje, por ejemplo, como se describe en la solicitud de patente de Estados Unidos 13/801.748 titulada "Mutational Analysis Of Plasma DNA For Cancer Detection" presentada el 13 de marzo de 2013.

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 usados para determinar el perfil de metilación sin células y el segundo perfil de metilación. Por lo tanto, la pluralidad de locus puede corresponderse. Es posible que puedan usarse más locus para determinar el perfil de metilación sin células y el segundo perfil de metilación.

En algunas realizaciones, pueden identificarse locus que estaban hipermetilados o hipometilados en el segundo perfil de metilación, por ejemplo, usando células sanguíneas maternas. Para identificar los locus que estaban hipermetilados en las células sanguíneas maternas, se puede explorar desde un extremo de un cromosoma con respecto a un sitio de CpG con un índice de metilación $\geq X\%$ (por ejemplo, 80 %). Se puede después buscar el siguiente sitio de CpG dentro de la región cadena abajo (por ejemplo, a una distancia de 200 pb cadena abajo). Si el sitio de CpG inmediatamente cadena abajo también tenía un índice de metilación $\geq X\%$ (u otra cantidad específica), pueden agruparse el primer y el segundo sitio de CpG. El agrupamiento puede continuar hasta que no haya ningún otro sitio de CpG dentro de la siguiente región cadena abajo; o el sitio de CpG inmediatamente cadena abajo tenga un índice de metilación $< X\%$. La región de los sitios de CpG agrupados pueden indicarse como hipermetilada en células sanguíneas maternas si la región contuviera al menos cinco sitios de CpG hipermetilados inmediatamente adyacentes. Puede realizarse un análisis similar para buscar locus que estaban hipometilados en células sanguíneas maternas para sitios de CpG con índices de metilación $\leq 20\%$. Las densidades de metilación para el segundo perfil de metilación pueden calcularse para los locus seleccionados y usarse para deducir el primer perfil de metilación (por ejemplo, densidad de metilación del tejido placentario) de los locus correspondientes, por ejemplo, a partir de datos de secuenciación por bisulfito de plasma materno.

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 sin células para cada uno de la pluralidad de locus. La diferencia se cambia de escala por el porcentaje.

En una realización, la primera densidad de metilación de un locus en el primer tejido (por ejemplo, placentario) (D) se dedujo usando la ecuación:

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

donde *mbc* indica la densidad de metilación del segundo perfil de metilación en un locus (por ejemplo, un locus seleccionado como se determina en los datos de secuenciación por bisulfito de células sanguíneas maternas); *mp* indica la densidad de metilación del locus correspondiente en los datos de secuenciación por bisulfito de plasma materno; *f* representó el porcentaje de ADN sin células del primer tejido (por ejemplo, concentración de ADN fetal fraccional) y CN representa el número de copias en el locus (por ejemplo, un valor mayor para amplificaciones o un número menor para supresiones en relación con la normalidad). Si no hay ninguna amplificación o delección en el primer tejido entonces CN puede ser uno. Para trisomía (o una duplicación de la región en un tumor o un feto), CN sería 1,5 (ya que el aumento es de 2 copias a 3 copias) y la monosomía sería 0,5. Puede aumentarse la amplificación mayor por incrementos de 0,5. En este ejemplo, D puede corresponder al parámetro diferencial.

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 explicar diferencias fijas entre los parámetros diferenciales y el perfil de metilación real del primer tejido. Por ejemplo, los valores pueden diferir por una constante fija o por una pendiente. La transformación puede ser lineal o no lineal.

En una realización, se descubrió que la distribución de los valores deducidos, *D*, era menor que el nivel de metilación real del tejido placentario. Por ejemplo, los valores deducidos pueden transformarse de forma lineal usando datos de islas de CpG, que fueron segmentos genómicos que tuvieron una sobrerrepresentación de sitios de CpG. Las posiciones genómicas de islas de CpG usadas en este estudio se obtuvieron de la base de datos del Buscador de Genoma de UCSC (NCBI versión 36/hg18) (PA Fujita *et al.*, 2011 Nucleic Acids Res; 39: D876-882). Por ejemplo, puede definirse una isla de CpG como un segmento genómico con contenido de GC $\geq 50\%$, longitud genómica > 200 pb y la relación de número de CpG observado/esperado $> 0,6$ (M Gardiner-Garden *et al* 1987 J Mol Biol; 196: 261-282).

En una implementación, para derivar la ecuación de transformación lineal, pueden incluirse islas de CpG con al menos 4 sitios de CpG y una profundidad de lectura promedio ≥ 5 por sitio de CpG en las muestras secuenciadas. Después de determinar las relaciones lineales entre las densidades de metilación de islas de CpG en la CVS o placenta a

término y los valores deducidos, D , se usaron las siguientes ecuaciones para determinar los valores predichos:

$$\text{Valores predichos del primer trimestre} = D \times 1,6 + 0,2$$

$$\text{Valores predichos del tercer trimestre} = D \times 1,2 + 0,05$$

B. Ejemplo fetal

Como se ha mencionado anteriormente, el método 1000 puede usarse para deducir un paisaje de metilación de la placenta del plasma materno. El ADN en circulación en plasma se origina predominantemente de células hematopoyéticas. Aún hay una proporción desconocida de ADN sin células al que contribuyen otros órganos internos. Además, el ADN sin células derivado de placenta representa aproximadamente el 5-40 % del ADN total en plasma materno, con una media de aproximadamente 15 %. Por lo tanto, se puede suponer que el nivel de metilación en plasma materno es equivalente a una metilación de fondo existente más una contribución placentaria durante el embarazo, como se ha descrito anteriormente.

El nivel de metilación de plasma materno, MP , puede determinarse usando la siguiente ecuación:

$$MP = BKG \times (1 - f) + PLN \times f$$

donde BKG es el nivel de migración de ADN de fondo en plasma obtenido de células sanguíneas y órganos internos, PLN es el nivel de metilación de la placenta y f es la concentración de ADN fetal fraccional en plasma materno.

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

$$PLN = \frac{MP - BKG \times (1 - f)}{f} \quad (2)$$

Las ecuaciones (1) y (2) son equivalentes cuando CN es igual a uno, D es igual a PLN y BKG es igual a mbc . En otra realización, la concentración de ADN fetal fraccional puede suponerse o establecerse en un valor específico, por ejemplo, como parte de una suposición de que esté presente un f mínimo.

Se ha considerado que el nivel de metilación de sangre materna representa la metilación de fondo del plasma materno. Además de los locus que estuvieron hipermetilados o hipometilados en células sanguíneas maternas, se exploró adicionalmente el enfoque de deducción centrándose en regiones definidas con relevancia clínica, por ejemplo, islas de CpG en el genoma humano.

La densidad de metilación media de un total de 27.458 islas de CpG (NCBI Versión36/hg18) en los autosomas y crX se obtuvo de los datos de secuenciación del plasma materno y la placenta. Solamente se seleccionaron los que tenían ≥ 10 sitios de CpG abarcados y una profundidad de lectura promedio de ≥ 5 por cada sitio de CpG abarcado en todas las muestras analizadas, incluyendo la placenta, la sangre materna y el plasma materno. Como resultado, 26.698 islas de CpG (97,2 %) permanecieron como válidas y su nivel de metilación se dedujo usando los datos de metilación de plasma y la concentración de ADN fetal fraccional según la ecuación anterior.

Se ha observado que la distribución de valores de PLN deducidos fue menor que el nivel de metilación real de islas de CpG en el tejido placentario. Por lo tanto, en una realización, los valores de PLN deducidos, o simplemente valores deducidos (D), se usaron como una unidad arbitraria para estimar el nivel de metilación de islas de CpG en la placenta. Después de una transformación, los valores deducidos de forma lineal y su distribución se hicieron más similares al conjunto de datos real. Los valores deducidos transformados se denominaron valores predictivos de metilación (VPM) y posteriormente se usaron para predecir el nivel de metilación de locus genéticos en la placenta.

En este ejemplo, las islas de CpG se clasificaron en 3 categorías basándose en sus densidades de metilación en la placenta: baja ($\geq 0,4$), intermedia ($>0,4 < 0,8$) y alta ($\geq 0,8$). Usando la ecuación de deducción, se calculó el VPM del mismo conjunto de islas de CpG y después se usaron los valores para clasificarlos en 3 categorías con los mismos puntos de corte. Comparando los conjuntos de datos reales y los deducidos, se descubrió que el 75,1 % de las islas de CpG seleccionadas podrían coincidir correctamente con las mismas categorías en los datos de tejidos de acuerdo con su VPM. Aproximadamente el 22 % de las islas de CpG se asignaron a grupos con una diferencia de 1 nivel (alto frente a intermedio o intermedio frente a bajo) y menos del 3 % estarían clasificadas de forma completamente errónea (alto frente a bajo) (FIGURA 12A). Se dedujo que el rendimiento de clasificación general también se determinó: 86,1 %, 31,4 % y 68,8 % de islas de CpG con densidades de metilación $\leq 0,4$, $>0,4 < 0,8$ y $\geq 0,8$ en la placenta era "Bajo", "Intermedio" y "Alto" correctamente (FIGURA 12B).

Las FIGURAS 11A y 11B muestran gráficos del rendimiento del algoritmo de predicción usando datos de plasma materno y la concentración de ADN fetal fraccional según las realizaciones de la presente divulgación. La FIGURA 11A es un gráfico 1100 que muestra la precisión de la clasificación de islas de CpG usando la clasificación con corrección de VPM (la categoría deducida coincide exactamente con el conjunto de datos real); diferencia de 1 nivel (la categoría deducida es 1 nivel diferente del conjunto de datos real); y clasificación errónea (la categoría deducida

es opuesta al conjunto de datos real). La FIGURA 11B es un gráfico 1150 que muestra la proporción de islas de CpG clasificadas en cada categoría deducida.

Siempre que la metilación de fondo materno sea baja en las regiones genómicas respectivas, la presencia de ADN derivado de placenta hipermetilado en la circulación aumentaría el nivel de metilación en plasma general hasta un grado dependiente de la concentración de ADN fetal fraccional. Podría observarse un cambio notable cuando el ADN fetal liberado esté completamente metilado. Por el contrario, cuando la metilación de fondo materno es alta, el grado de cambio en el nivel de metilación de plasma sería más significativo si se libera ADN fetal hipometilado. Por lo tanto, el esquema de deducción puede ser más práctico cuando el nivel de metilación se deduce para locus genéticos que se sabe que son distintos entre el fondo materno y la placenta, especialmente para los marcadores hipermetilados e hipometilados en la placenta.

La FIGURA 12A es una tabla 1200 que muestra detalles de 15 locus genómicos seleccionados para predicción de metilación según realizaciones de la presente divulgación. Para confirmar técnicas, se seleccionaron 15 locus genómicos metilados diferencialmente que se habían estudiado previamente. Los niveles de metilación de regiones seleccionadas se dedujeron y se compararon con los 15 locus genéticos metilados diferencialmente previamente estudiados (RWK Chiu *et al.* 2007 Am J Pathol; 170: 941-950; S.S.C. Chim *et al.* 2008 Clin Chem; 54: 500-511; SSC Chim *et al.* 2005 Proc Natl Acad Sci USA; 102: 14753-14758; DWY Tsui *et al.* 2010 PLoS One; 5: e15069).

La FIGURA 12B es un gráfico 1250 que muestra las categorías deducidas de los 15 locus genómicos seleccionados y sus niveles de metilación correspondientes en la placenta. Las categorías de metilación deducidas son: baja, $\leq 0,4$; Intermedia, $>0,4$ - $<0,8$; Alta, $\geq 0,8$. La Tabla 1200 y el gráfico 1300 muestran que sus niveles de metilación en la placenta podrían deducirse correctamente con varias excepciones: RASSF1A, CGI009, CGI137 y VAPA. De estos 4 marcadores, solamente CGI009 mostró una discrepancia notable con el conjunto de datos real. Los otros se clasificaron erróneamente de forma solamente marginal.

En la tabla 1200, "1" se refiere a los valores deducidos (D) que se calculan por la ecuación:

$$D = \frac{MP - BKG \times (1 - f)}{f}$$

donde f es la concentración de ADN fetal fraccional. La etiqueta "2" se refiere a los valores predictivos de metilación (VPM) que se refieren a los valores deducidos linealmente transformados usando la ecuación: $VPM = D \times 1,6 + 0,25$. La etiqueta "3" se refiere al punto de corte de clasificación para los valores deducidos: bajo, $\leq 0,4$; Inter(medio), $>0,4$ - $<0,8$; Alto, $\geq 0,8$. La etiqueta "4" se refiere al punto de corte de clasificación para el conjunto de datos placentario real: bajo, $\leq 0,4$; Inter(medio), $>0,4$ - $<0,8$; Alto, $\geq 0,8$. La etiqueta "5" indica que el estado placentario se refiere al estado de metilación de la placenta en relación con el de células sanguíneas maternas.

C. Cálculo de concentraciones fraccionales de ADN fetal

En una realización, el porcentaje de ADN fetal del primer tejido puede usar un cromosoma Y para un feto masculino. La proporción de secuencias de cromosoma Y (% de crY) en una muestra de plasma materno fue una combinación de las lecturas de cromosoma Y derivadas del feto masculino y el número de lecturas maternas (femeninas) que se alinearon incorrectamente con el cromosoma Y (RWK Chiu *et al.* 2011 BMJ; 342: c7401). Por lo tanto, la relación entre % de crY y la concentración de ADN fetal fraccional (f) en la muestra puede proporcionarse por:

$$\%crY = \%crY_{\text{masculino}} \times f + \%crY_{\text{femenino}} \times (1 - f)$$

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

El %crY puede determinarse a partir de lecturas que se alinearon con el cromosoma Y sin desapareamientos para una muestra de un embarazo femenino con un feto masculino, por ejemplo, donde las lecturas son de muestras convertidas por bisulfito. El valor de $\%crY_{\text{masculino}}$ puede obtenerse para la secuenciación por bisulfito de dos muestras de plasma de hombre adulto. El valor de $\%crY_{\text{femenino}}$ puede obtenerse a partir de la secuenciación por bisulfito de dos muestras de plasma de mujeres adultas no embarazadas.

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

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

El genoma placentario está más hipometilado que el genoma materno. Como se ha analizado anteriormente la metilación del plasma de una mujer embarazada 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 tejidos fetales al plasma materno. Por ejemplo,

en una mujer embarazada que porta un feto trisómico (por ejemplo, que padece trisomía del 21 o trisomía del 18 o trisomía del 13), el feto aportaría una cantidad adicional del ADN del cromosoma trisómico al plasma materno en comparación con los cromosomas disómicos. En esta situación, la densidad de metilación en plasma para el cromosoma trisómico (o cualquier región cromosómica que tenga una amplificación) sería menor que las de los cromosomas disómicos. El grado de diferencia puede predecirse por cálculo matemático teniendo en cuenta la concentración de ADN fetal fraccional en la muestra de plasma. Cuanto mayor sea la concentración de ADN fetal fraccional en la muestra de plasma mayor será la diferencia de la densidad de metilación entre los cromosomas trisómicos y disómicos. Para regiones que tengan una delección, la densidad de metilación sería mayor.

Un ejemplo de una delección es el síndrome de Turner, cuando un feto femenino tendría solamente una copia del cromosoma X. En esa situación, para una mujer embarazada que porta un feto que padece síndrome de Turner, la densidad de metilación del cromosoma X en su ADN en plasma sería mayor que la situación de la misma mujer embarazada que porte un feto femenino que tenga el número normal de cromosoma X. En una realización de esta estrategia, se podría analizar en primer lugar plasma materno con respecto a la presencia o ausencia de secuencias del cromosoma Y (por ejemplo, usando SPM o una técnica basada en PCR). Si están presentes secuencias del cromosoma Y, entonces el feto puede clasificarse como masculino y el siguiente análisis no sería necesario. Por otro lado, si están ausentes secuencias del cromosoma Y en el plasma materno, entonces el feto puede clasificarse como femenino. En esta situación, se puede analizar después la densidad de metilación del cromosoma X en el plasma materno. Una mayor densidad de metilación del cromosoma X que la normal indicaría que el feto tiene un alto riesgo de tener síndrome de Turner. Este enfoque también puede aplicarse para las otras aneuploidías cromosómicas sexuales. Por ejemplo, para un feto aquejado de XYY, la densidad de metilación para el cromosoma Y en plasma materno sería menor que la de un feto XY normal que tenga un nivel similar de ADN fetal en plasma materno. Como otro ejemplo, para un feto que padece síndrome de Klinefelter (XXY), están presentes secuencias del cromosoma Y en plasma materno, pero la densidad de metilación del cromosoma X en plasma materno será menor que la de un feto XY normal que tenga un nivel similar de ADN fetal en plasma materno.

A partir del análisis anterior, puede calcularse la densidad de metilación en plasma para un cromosoma disómico ($MP_{No\ aneu}$) como: $MP_{No\ aneu} = BKG \times (1 - f) + PLN \times f$, donde BKG es el nivel de metilación de ADN de fondo en plasma derivado de células sanguíneas y órganos internos, PLN es el nivel de metilación de la placenta y f es la concentración de ADN fetal fraccional en plasma materno.

La densidad de metilación en plasma para un cromosoma trisómico (MP_{Aneu}) puede calcularse como: $MP_{Aneu} = BKG \times (1 - f) + PLN \times f \times 1,5$, donde el 1,5 corresponde al número de copias CN y la adición de uno o más cromosomas es un aumento del 50 %. La diferencia entre un cromosoma trisómico y uno disómico (MP_{Diff}) sería

$$MP_{Diff} = PLN \times f \times 0,5.$$

En una realización, una comparación de la densidad de metilación del cromosoma (o región cromosómica) potencialmente aneuploide con uno o más cromosomas adicionales supuestamente no aneuploides o la densidad de metilación general del genoma puede usarse para normalizar eficazmente la concentración de ADN fetal en la muestra de plasma. La comparación puede ser mediante un cálculo de un parámetro (por ejemplo, que implica una relación o una diferencia) 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 (por ejemplo, determinado como un parámetro de las dos densidades de metilación).

Si la densidad de metilación del cromosoma potencialmente aneuploide no se normaliza con respecto a la densidad de metilación de uno o más cromosomas adicionales, u otros parámetros que reflejan la concentración fraccional de ADN fetal, la concentración fraccional sería un factor importante que afecta a la densidad de metilación en el plasma. Por ejemplo, la densidad de metilación en plasma del cromosoma 21 de una mujer embarazada que porta un feto con trisomía del 21 con una concentración de ADN fetal fraccional del 10 % sería igual que la de una mujer embarazada que porta un feto euploide y la concentración de ADN fetal fraccional es del 15 %, mientras que una densidad de metilación normalizada mostraría una diferencia.

En otra realización, la densidad de metilación del cromosoma potencialmente aneuploide puede normalizarse con respecto a la concentración de ADN fetal fraccional. Por ejemplo, la siguiente ecuación puede aplicarse para normalizar la densidad de metilación: $MP_{Normalizada} = MP_{no\ normalizada} + (BKG - PLN) \times f$, donde $MP_{Normalizada}$ es la densidad de metilación normalizada con la concentración de ADN fetal fraccional en el plasma, $MP_{no\ normalizada}$ es la densidad de metilación medida, BKG es la densidad de metilación de fondo de células sanguíneas o tejidos maternos, PLN es la densidad de metilación en los tejidos placentarios, y f es la concentración de ADN fetal fraccional. Las densidades de metilación de BKG y PLN podrían basarse en valores de referencia previamente establecidos de células sanguíneas maternas y tejidos placentarios obtenidos de embarazos sanos. Pueden usarse métodos genéticos y epigenéticos diferentes para la determinación de la concentración de ADN fetal fraccional en la muestra de plasma, por ejemplo, mediante la medición del porcentaje de lecturas de secuencias del cromosoma Y usando secuenciación paralela masiva o PCR en ADN no convertido por bisulfito.

En una implementación, la densidad de metilación normalizada para un cromosoma potencialmente aneuploide puede

compararse con un grupo de referencia que consiste en una mujer embarazada que porta fetos euploides. Puede determinarse la media y DT de la densidad de metilación normalizada del grupo de referencia. Después puede expresarse la densidad de metilación normalizada del caso ensayado como una puntuación z que indica el número de DT de la media del grupo de referencia por:

5

$$\text{puntuación } z = \frac{MP_{\text{Normalizada}} - \text{Media}}{DT}$$

donde $MP_{\text{Normalizada}}$ es la densidad de metilación normalizada para el caso ensayado, 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. Un punto de corte, por ejemplo, puntuación z <-3, puede usarse para clasificar si un cromosoma está hipometilado de forma significativa y, por lo tanto, para determinar el estado de aneuploidía de la muestra.

10

En otra realización, la MP_{Dif} puede usarse como la densidad de metilación normalizada. En dicha realización, puede deducirse PLN, por ejemplo, usando el método 1000. En algunas implementaciones, una densidad de metilación de referencia (que puede normalizarse usando f) puede determinarse a partir de un nivel de metilación de una región no aneuploide. Por ejemplo, la media podría determinarse a partir de una o más regiones cromosómicas de la misma muestra. El punto de corte podría cambiarse de escala por f, o ajustarse solo a un nivel suficiente siempre que exista una concentración mínima.

15

En consecuencia, una comparación de un nivel de metilación con respecto a una región con un punto de corte puede conseguirse de diversas formas. La comparación puede implicar una normalización (por ejemplo, como se ha descrito anteriormente), que puede realizarse de forma equivalente en el nivel de metilación o el valor de punto de corte, dependiendo de cómo se definan los valores. Por lo tanto, si el nivel de metilación determinado de una región es estadísticamente diferente de un nivel de referencia (determinado a partir de la misma muestra u otras muestras) puede determinarse de diversas maneras.

20

25

El análisis anterior puede aplicarse a análisis de regiones cromosómicas, que pueden incluir un cromosoma completo o partes del cromosoma, incluyendo subregiones contiguas o separadas de un cromosoma. En una realización, el cromosoma potencialmente aneuploide puede dividirse en varios grupos. Los grupos pueden ser del mismo tamaño o de tamaños diferentes. La densidad de metilación de cada grupo puede normalizarse con respecto a la concentración fraccional de la muestra o la densidad de metilación de uno o más cromosomas supuestamente no aneuploides o la densidad de metilación general del genoma. La densidad de metilación normalizada de cada grupo puede después compararse con un grupo de referencia para determinar si está hipometilada de forma significativa. Después puede determinarse el porcentaje de grupos que están significativamente hipometilados. Un punto de corte, por ejemplo, más de 5 %, 10 %, 15 %, 20 % o 30 % de los grupos que están significativamente hipometilados, puede usarse para clasificar el estado de aneuploidía del caso.

30

35

Cuando se ensaya con respecto a una amplificación o una delección, se puede comparar la densidad de metilación con una densidad de metilación de referencia, que puede ser específica para una región particular que se ensaye. Cada región puede tener una densidad de metilación de referencia diferente ya que la metilación puede variar entre regiones, particularmente dependiendo del tamaño de las regiones (por ejemplo, regiones más pequeñas mostrarán más variación).

40

Como se ha mencionado anteriormente, pueden usarse una o más mujeres embarazadas que portan cada una un feto euploide 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 la PLN (por ejemplo, mediante medición directa o como se deduce por el método 1000). En otras realizaciones, puede usarse una relación entre dos densidades de metilación, por ejemplo, de un cromosoma potencialmente aneuploide y puede usarse un cromosoma no aneuploide para el análisis en lugar de su diferencia. Este enfoque de análisis de metilación puede combinarse con enfoque de recuento de lecturas de secuencias (RWK Chiu *et al.* 2008 Proc Natl Acad Sci USA; 105: 20458-20463) y enfoques que implican el análisis del tamaño de ADN de plasma (Patente de Estados Unidos 2011/0276277) para determinar o confirmar una aneuploidía. El enfoque de recuento de lecturas de secuencias que se usa en combinación con análisis de metilación puede realizarse usando secuenciación aleatoria (RWK Chiu *et al.* 2008 Proc Natl Acad Sci USA; 105: 20458-20463; DW Bianchi *et al.* 2012 Obstet Gynecol 119: 890-901) o secuenciación dirigida (AB Sparks *et al.* 2012 Am J Obstet Gynecol 206: 319.e1-9; B Zimmermann *et al.* 2012 Prenat Diagn 32: 1233-1241; GJ Liao *et al.* 2012 PLoS One; 7:e38154).

45

50

55

El uso de BKG puede explicar variaciones en el fondo entre muestras. Por ejemplo, una mujer podría tener diferentes niveles de metilación de BKG que otra mujer, pero puede usarse una diferencia entre la BKG y la PLN entre muestras en dichas situaciones. El punto de corte para diferentes regiones cromosómicas puede ser diferente, por ejemplo, cuando una densidad de metilación de una región del genoma difiere con respecto a otra región del genoma.

60

Este enfoque puede generalizarse para detectar cualquier aberración cromosómica, incluyendo delección y amplificación, en el genoma fetal. Además, la resolución de este análisis puede ajustarse al nivel deseado, por ejemplo,

el genoma puede dividirse en grupos de 10 Mb, 5 Mb, 2 Mb, 1 Mb, 500 kb, 100 kb. Por lo tanto, esta tecnología también puede usarse para detectar duplicación subcromosómica o delección subcromosómica. Esta tecnología permitiría por lo tanto obtener de forma no invasiva un cariotipo molecular fetal prenatal. Cuando se usa de esta manera, esta tecnología puede usarse en combinación con los métodos de ensayo prenatales no invasivos que se basan en el recuento de moléculas (A Srinivasan *et al.* 2013 Am J Hum Genet; 92: 167-176; SCY Yu *et al.* 2013 PLoS One 8: e60968). En otras realizaciones, no es necesario que el tamaño de los grupos sea idéntico. Por ejemplo, el tamaño de los grupos puede ajustarse de modo que cada grupo contenga un número idéntico de dinucleótidos de CpG. En este caso, el tamaño físico de los grupos sería diferente.

La ecuación puede reescribirse para aplicar a diferentes tipos de aberraciones cromosómicas como $MP_{Dif} = (BKG - PLN) \times f \times 0,5 \times CN$. Aquí CN representa el número de cambios de número de copias en la región afectada. CN es 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 (por ejemplo, para detectar síndrome de Turner fetal en el que un feto femenino ha perdido uno de los cromosomas X, lo que conduce a un cariotipo XO). No es necesario cambiar esta ecuación cuando se cambia el tamaño de los grupos. Sin embargo, la sensibilidad y especificidad pueden reducirse cuando se usa un tamaño de grupos menor debido a que estaría presente un menor número de dinucleótidos de CpG (u otras combinaciones de nucleótidos que muestren metilación diferencial entre ADN fetal y ADN materno) en grupos más pequeños, lo que conduce a una variación estocástica aumentada en la medición de densidades de metilación. En una realización, el número de lecturas requerido puede determinarse analizando el coeficiente de variación de la densidad de metilación y el nivel deseado de sensibilidad.

Para demostrar la viabilidad de este enfoque, se han analizado las muestras de plasma de 9 mujeres embarazadas. En cinco mujeres embarazadas, cada una portaba un feto euploide y las otras cuatro portaban cada una un feto con trisomía del 21 (T21). Tres de los cinco embarazos euploides se seleccionaron de forma aleatoria para formar un grupo de referencia. Los dos casos de embarazo euploide restantes (Eu1 y Eu2) y los cuatro casos de T21 (T21-1, T21-2, T21-3 y T21-4) se analizaron usando este enfoque para ensayar con respecto a un estado de T21 potencial. El ADN del plasma se convirtió por bisulfito y se secuenció usando la plataforma HiSeq2000 de Illumina. En una realización, se calculó la densidad de metilación de cromosomas individuales. Después 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 usó la media y DT del grupo de referencia para el cálculo de la puntuación z de los seis casos de ensayo.

Tabla 1: usando un punto de corte de <-3 para puntuación z para clasificar una muestra como T21, la clasificación de todos los casos euploides y T21 fue correcta.

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

En otra realización, el genoma se dividió en grupos de 1 Mb y se determinó la densidad de metilación para cada grupo de 1 Mb. La densidad de metilación de todos los grupos en el cromosoma potencialmente aneuploide puede normalizarse con la mediante de la densidad de metilación de todos los grupos localizados en los cromosomas supuestamente no aneuploides. En una implementación, para cada grupo, puede calcularse la diferencia en la densidad de metilación de la mediana de los grupos no aneuploides. La puntuación z puede calcularse para estos valores usando la media y los valores de DT del grupo de referencia. Puede determinarse el porcentaje de grupos que muestran hipometilación (Tabla 2) y compararse con un porcentaje de punto de corte.

Tabla 2: usando el 5 % como un punto de corte para los grupos con significativamente más hipometilación en el cromosoma 21, todos los casos se clasificaron correctamente para el estado de T21

	Eu1	Eu2	T21-1	T21-2	T21-3	T21-4
Porcentaje de grupos en cr 21 que tienen una puntuación z de $MP_{Dif} <-3$	0 %	0 %	33,3 %	58,3 %	19,4 %	52,8 %

Este enfoque basado en metilación de ADN para detectar aberraciones cromosómicas o subcromosómicas fetales puede usarse junto con los basados en el recuento de moléculas tal como por secuenciación (RWK Chiu *et al.* 2008 Proc Natl Acad Sci USA; 105: 20458-20463) o PCR digital (YMD Lo *et al.* 2007 Proc Natl Acad Sci USA; 104: 13116-13121), o por la medición de las moléculas de ADN (Publicación de Patente de Estados Unidos 2011/0276277). Dicha combinación (por ejemplo, metilación de ADN más recuento molecular o metilación de ADN más medición del tamaño o metilación de ADN más recuento molecular más medición del tamaño) tendría un efecto sinérgico que sería ventajoso en una situación clínica, por ejemplo, mejorando la sensibilidad y/o especificidad. Por ejemplo, el número de moléculas de ADN que sería necesario analizar, por ejemplo, por secuenciación, puede reducirse sin influir de forma adversa a la precisión de diagnóstico. Esta característica permitiría realizar dichos ensayos de forma más económica. Como otro ejemplo, para un número dado de moléculas de ADN analizadas, un enfoque combinado permitiría detectar

aberraciones cromosómicas o subcromosómicas fetales a una concentración fraccional menor de ADN fetal.

La FIGURA 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 sin células que comprende una mezcla de ADN sin células que se origina desde un primer tejido y desde un segundo tejido. El primer tejido puede ser de un feto o un tumor y el segundo tejido puede ser de una mujer embarazada o un paciente.

En el bloque 1310, se analiza una pluralidad de moléculas de ADN de la muestra biológica. El análisis de una molécula de ADN puede incluir determinar una localización de la molécula de ADN en un genoma del organismo y determinar si la molécula de ADN está metilada en uno o más sitios. El análisis puede realizarse recibiendo lecturas de secuencias de una secuenciación sensible a la metilación, y por lo tanto el análisis puede realizarse solamente en datos previamente obtenidos del ADN. En otras realizaciones, el análisis puede incluir la secuenciación real u otras etapas activas para obtener los datos.

La determinación de una localización puede incluir mapear las moléculas de ADN (por ejemplo, mediante lecturas de secuencias) en partes respectivas del genoma humano, por ejemplo, en regiones específicas. En una implementación, si una lectura no se mapea en una región de interés, entonces la lectura puede ignorarse.

En el bloque 1320, un número respectivo de moléculas de ADN que están metiladas en el sitio se determina para cada uno de una pluralidad de sitios. En una realización, los sitios son sitios de CpG, y pueden ser solamente ciertos sitios de CpG, como se selecciona usando uno o más criterios mencionados en el presente documento. El número de ADN que están metilados es equivalente a la determinación del número que no están metilados una vez que se ha realizado normalización usando un número total de moléculas de ADN analizadas en un sitio particular, por ejemplo, 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 números respectivos de moléculas de ADN metiladas en sitios dentro de la primera región cromosómica. La primera región cromosómica puede ser de cualquier tamaño, por ejemplo, tamaños mencionados anteriormente. El nivel de metilación puede representar un número total de moléculas de ADN alineadas con la primera región cromosómica, por ejemplo, como parte de un procedimiento de normalización.

La primera región cromosómica puede ser de cualquier tamaño (por ejemplo, un cromosoma completo) y puede estar compuesta de subregiones inconexas, es decir, subregiones que están separadas entre sí. Pueden determinarse los niveles de metilación de cada subregión y combinarse, por ejemplo, como un promedio o una mediana, para determinar un nivel de metilación para la primera región cromosómica.

En el bloque 1340, el primer nivel de metilación se compara con un valor de punto de corte. El valor de punto de corte puede ser un nivel de metilación de referencia o estar relacionado con un nivel de metilación de referencia (por ejemplo, una distancia específica de un nivel normal). El valor de punto de corte puede determinarse a partir de otras mujeres embarazadas que portan fetos sin una anomalía cromosómica para la primera región cromosómica, de muestras de individuos sin cáncer, o de locus del organismo que se sabe que no están asociados con una aneuploidía (es decir, regiones que son disómicas).

En una realización, puede definirse que el valor de punto de corte tiene una diferencia de un nivel de metilación de referencia de $(BKG - PLN) \times f \times 0,5 \times CN$, donde BKG es el fondo de la mujer (o un promedio o una mediana de otros sujetos), f es la concentración fraccional de ADN sin células que se origina en el primer tejido y CN es un número de copias que se ensaya. CN es un ejemplo de un factor de escala correspondiente a un tipo de anomalía (deleción o duplicación). Puede usarse un punto de corte para un CN de 1 para ensayar todas las amplificaciones inicialmente, y después pueden usarse puntos de corte adicionales para determinar el grado de amplificación. El valor de punto de corte puede basarse en una concentración fraccional de ADN sin células que se origina del primer tejido para determinar el nivel esperado de metilación para un locus, por ejemplo, si no está presente ninguna aberración del número de copias.

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 aumento del riesgo del feto de tener una anomalía cromosómica. En diversas realizaciones, la anomalía cromosómica puede ser trisomía del 21, trisomía del 18, trisomía del 13, síndrome de Turner o síndrome de Klinefelter. Otros ejemplos son una deleción subcromosómica, duplicación subcromosómica o síndrome de DiGeorge.

V. DETERMINACIÓN DE MARCADORES

Como se ha observado anteriormente, ciertas partes del genoma fetal están metiladas de forma diferente al genoma materno. Estas diferencias pueden ser comunes entre embarazos. Las regiones de diferente metilación pueden usarse para identificar fragmentos de ADN que son del feto.

A. Método para determinar DMR de tejido placentario y tejido materno

La placenta tiene identificaciones de metilación específicas de tejido. Se han desarrollado marcadores de metilación del ADN específicos fetales para detección en plasma materno y para aplicaciones de diagnóstico prenatal no invasivas basándose en locus que están metilados diferencialmente entre tejidos placentarios y células sanguíneas maternas (SSC Chim *et al.* 2008 Clin Chem; 54: 500-511; EA Papageorgiou *et al* 2009 Am J Pathol; 174: 1609-1618; y T Chu *et al.* 2011 PLoS One; 6: e14723). Se proporcionan realizaciones para aprovechar dichas regiones metiladas diferencialmente (DMR) en todo el genoma.

La FIGURA 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 (por ejemplo, determinado a partir de células sanguíneas) de acuerdo con realizaciones de la presente divulgación. El método 1400 también puede usarse para determinar marcadores para un tumor comparando un perfil de metilación tumoral con un perfil de metilación correspondiente a tejido sano.

En el bloque 1410, se obtiene un metiloma placentario y un metiloma de sangre. El metiloma placentario puede determinarse a partir de una muestra placentaria, por ejemplo, CVS o una placenta a término. Debería entenderse que el metiloma posiblemente incluya densidades de metilación de solamente parte de un genoma.

En el bloque 1420, se identifica una región que incluye un número específico de sitios (por ejemplo, 5 sitios de 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 contenía al menos cinco sitios de CpG cualificados. Un sitio de CpG puede considerarse cualificado si el sitio está abarcado por al menos cinco lecturas de secuencias.

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

En el bloque 1440, los índices de metilación se compararon entre las células sanguíneas maternas y la muestra placentaria para determinar si los conjuntos de índices eran diferentes entre sí. Por ejemplo, los índices de metilación se compararon entre las células sanguíneas maternas y la CVS o la placenta a término usando, por ejemplo, el ensayo de Mann-Whitney. Un valor de P, por ejemplo, $\leq 0,01$, se consideró estadísticamente significativamente diferente, aunque pueden usarse otros valores, en los que un número menor reduciría las regiones de falso positivo.

En una realización, si el número de sitios de CpG cualificados era menor de cinco o el ensayo de Mann-Whitney no era significativo, la región de 500 pb se desplazó cadena abajo para 100 pb. La región continuó desplazándose cadena abajo hasta que el ensayo de Mann-Whitney se hizo significativo para una región de 500 pb. Después se consideraría la siguiente región de 500 pb. Si se descubriera que la siguiente región muestra significación estadística por ensayo de Mann-Whitney, se añadiría a la región actual siempre que la región contigua combinada no fuera mayor de 1.000 pb.

En el bloque 1450, pueden mezclarse regiones adyacentes que sean estadísticamente significativamente diferentes (por ejemplo, por el ensayo de Mann-Whitney). Obsérvese que la diferencia es entre los índices de metilación para las dos muestras. En una realización, si las regiones adyacentes están a una distancia específica (por ejemplo, 1.000 pb) entre sí y si mostraron un perfil de metilación similar entonces podrían mezclarse. En una implementación, la similitud del perfil de metilación entre regiones adyacentes puede definirse usando cualquiera de los siguientes: (1) mostrar la misma tendencia en el tejido placentario en referencia a las células sanguíneas maternas, por ejemplo ambas regiones estaban más metiladas en los tejidos placentarios que las células sanguíneas; (2) con diferencias en las densidades de metilación de menos del 10 % para las regiones adyacentes en el tejido placentario; y (3) con diferencias en las densidades de metilación de menos del 10 % para las regiones adyacentes en las células sanguíneas maternas.

En el bloque 1460, se calcularon las densidades de metilación del metiloma sanguíneo de ADN de células sanguíneas maternas y muestras placentarias (por ejemplo, CVS 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 DMR potenciales donde la densidad de metilación placentaria total y una densidad de metilación sanguínea total para todos los sitios en la región son estadísticamente significativamente diferentes. En una realización, todos los sitios de CpG cualificados dentro de una región mezclada se someten a un ensayo de χ^2 . El ensayo de χ^2 evaluó si el número de citosinas metiladas como una proporción de las citosinas metiladas y desmetiladas entre todos los sitios de CpG cualificados dentro de la región mezclada eran estadísticamente significativamente diferentes entre las células sanguíneas maternas y el tejido placentario. En una implementación, para el ensayo de χ^2 un valor de $P \leq 0,01$ puede considerarse estadísticamente significativamente diferente. Los segmentos mezclados que mostraron significación por el ensayo de χ^2 se consideraron DMR potenciales.

En el bloque 1480, se identificaron locus en los que las densidades de metilación del ADN de células sanguíneas maternas estaban por encima de un punto de corte alto o por debajo de un punto de corte bajo. En una realización, se identificaron locus en los que las densidades de metilación del ADN de células sanguíneas maternas eran ≤ 20 % o

>80 %. En otras realizaciones, pueden usarse líquidos corporales distintos de sangre materna, incluyendo, pero sin limitación, saliva, líquido de lavado uterino o del cuello uterino del tracto genital femenino, lágrimas, sudor, saliva y orina.

5 Una clave para el desarrollo exitoso de marcadores de metilación de ADN que son específicos fetales en plasma materno puede ser que el estado de metilación de las células sanguíneas maternas sea altamente metilado o tan desmetilado como sea posible. Esto puede reducir (por ejemplo, minimizar) la probabilidad de tener moléculas de ADN materno que interfieran con el análisis de las moléculas de ADN fetal derivadas de placenta que muestran un perfil de metilación opuesto. Por lo tanto, en una realización, las DMR candidatas se seleccionaron por filtración adicional. Los
10 locus hipometilados candidatos eran los que mostraron densidad de metilación ≤ 20 % en las células sanguíneas maternas y con densidades de metilación al menos 20 % mayores en los tejidos placentarios. Los locus hipermetilados candidatos fueron los que mostraron densidades de metilación ≥ 80 % en las células sanguíneas maternas y con densidades de metilación al menos 20 % menores en los tejidos placentarios. Pueden usarse otros porcentajes.

15 En el bloque 1490, se identificaron después DMR entre el subconjunto de locus en los que las densidades de metilación placentarias son significativamente diferentes de las densidades de metilación sanguíneas comparando la diferencia con un umbral. En una realización, el umbral es del 20 %, de modo que las densidades de metilación difirieron en al menos 20 % de las densidades de metilación de las células sanguíneas maternas. En consecuencia, puede calcularse una diferencia entre las densidades de metilación placentarias y las densidades de metilación sanguíneas en cada
20 locus identificado. La diferencia puede ser una simple resta. En otras realizaciones, pueden usarse factores de cambio de escala y otras funciones para determinar la diferencia (por ejemplo, la diferencia puede ser el resultado de una función aplicada a la simple resta).

En una implementación, usando este método, se identificaron 11.729 locus hipermetilados y 239.747 hipometilados a
25 partir de la muestra placentaria del primer trimestre. Los 100 locus más hipermetilados se enumeran en la tabla S2A del apéndice. Los 100 locus más hipometilados se enumeran en la tabla S2B del apéndice. Las tablas S2A y S2B enumeran el cromosoma, la localización de partida y final, el tamaño de la región, la densidad de metilación en la sangre materna, la densidad de metilación en la muestra placentaria, los valores de P (que son todos muy bajos) y la diferencia de metilación. Las localizaciones corresponden al genoma de referencia hg18, que puede encontrarse en
30 hgdownload.soe.ucsc.edu/goldenPath/hg18/chromosomes.

Se identificaron 11.920 locus hipermetilados y 204.768 hipometilados de la muestra placentaria del tercer trimestre. Los 100 locus más hipermetilados para el 3^{er} trimestre se enumeran en la tabla S2C y los 100 locus más hipometilados se enumeran en la tabla S2D. Se usaron treinta y tres locus que se había indicado previamente que estaban metilados
35 diferencialmente entre células sanguíneas maternas y tejidos placentarios de primer trimestre para validar la lista de los inventores de candidatos del primer trimestre. Se habían identificado 79 % de los 33 locus como DMR usando el algoritmo de los inventores.

La FIGURA 15A es una tabla 1500 que muestra una realización de algoritmo de identificación de DMR usando datos
40 del primer trimestre con referencia a 33 marcadores del primer trimestre previamente indicados. En la tabla, "a" indica que los locus 1 a 15 se habían descrito previamente en (RWK Chiu *et al.* 2007 Am J Pathol; 170: 941-950 y SSC Chim *et al.* 2008 Clin Chem; 54: 500-511); los locus 16 a 23 se habían descrito previamente en (KC Yuen, tesis 2007, La Universidad China de Hong Kong, Hong Kong); y los locus 24 a 33 se habían descrito previamente en (EA Papageorgiou *et al.* 2009 Am J Pathol; 174: 1609-1618). "b" indica que estos datos se obtuvieron de las publicaciones anteriores. "c" indica que se observaron densidades de metilación de células sanguíneas maternas y muestra de vellosidad coriónica y sus diferencias a partir de los datos de secuenciación generados en el estudio presente, pero basándose en las coordenadas genómicas proporcionadas por los estudios originales. "d" indica que los datos de los
45 locus se identificaron usando realizaciones del método 1400 en los datos de secuenciación por bisulfito sin tomar referencia de las publicaciones citadas anteriormente por Chiu *et al.* (2007), Chim *et al.* (2008), Yuen (2007) y Papageorgiou *et al.* (2009). La extensión de los locus incluyó las regiones genómicas previamente indicadas, pero en general abarcó regiones mayores. "e" indica que una DMR candidata se clasificó como verdadero positivo (VP) o falso negativo (FN) basándose en el requisito de observar una diferencia $>0,20$ entre las densidades de metilación de las coordenadas de genoma correspondientes de las DMR en células sanguíneas maternas y muestra de vellosidad coriónica.

55 La FIGURA 15B es una tabla 1550 que muestra una realización de algoritmo de identificación de DMR usando datos del tercer trimestre y en comparación con la muestra de placenta obtenida en el parto. "a" indica que se usó la misma lista de 33 locus que se describe en la FIGURA 17A. "b" indica que como los 33 locus se identificaron previamente a partir de muestras de embarazo tempranas, no podrían ser aplicables a los datos del tercer trimestre. Por lo tanto, se
60 revisaron los datos de secuenciación por bisulfito generados en el presente estudio en el tejido placentario a término basándose en las coordenadas genómicas proporcionadas por los estudios originales. Se usó una diferencia de $>0,20$ en las densidades de metilación entre las células sanguíneas maternas y el tejido placentario a término para determinar si los locus eran de hecho verdaderas DMR en el tercer trimestre. "c" indica que los datos de los locus se identificaron usando el método 1400 en los datos de secuenciación por bisulfito sin tomar referencia de publicaciones previamente
65 citadas en Chiu *et al.* (2007), Chim *et al.* (2008), Yuen (2007) y Papageorgiou *et al.* (2009). La extensión de los locus incluyó las regiones genómicas previamente indicadas, pero en general abarcó regiones mayores. "d" indica que DMR

candidatas que contenían locus que se clasificaron como metiladas 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 de genoma correspondientes de las DMR en células sanguíneas maternas y tejido placentario a término. Para locus que no se clasificaron como diferencialmente metilados en el tercer trimestre, su ausencia en la lista de DMR o la presencia de una DMR que contenía los locus pero que mostraba diferencia de metilación de $<0,20$ se consideró como DMR verdaderas negativas (VN).

B. DMR de datos de secuenciación de plasma materno

Debería poderse identificar DMR de tejido placentario directamente a partir de los datos de secuenciación por bisulfito de ADN de plasma materno siempre que se conociera también la concentración de ADN fetal fraccional de la muestra. Es posible debido a que la placenta es la fuente predominante de ADN fetal en el plasma materno (SSC Chim *et al.* 2005 Proc Natl Acad Sci USA 102, 14753-14758) y los inventores han mostrado en este estudio que el estado de metilación de ADN específico fetal en plasma materno se correlacionaba con el metiloma placentario.

Por lo tanto, pueden implantarse aspectos del método 1400 usando un metiloma de plasma para determinar un metiloma placentario deducido en lugar de usar una muestra placentaria. Por lo tanto, el método 1000 y el método 1400 pueden combinarse para determinar las DMR. El método 1000 puede usarse para determinar los valores predichos para el perfil de metilación placentario y usarlos en el método 1400. Para este análisis, el ejemplo también se centra en locus que estaban $\leq 20\%$ o $\geq 80\%$ metilados en las células sanguíneas maternas.

En una implementación, para deducir locus que estaban hipermetilados en los tejidos placentarios con respecto a células sanguíneas maternas, se clasificó con respecto a locus que mostraban $\leq 20\%$ de metilación en células sanguíneas maternas, y $\geq 60\%$ de metilación según el valor predicho con una diferencia de al menos 50 % entre la densidad de metilación de células sanguíneas y el valor predicho. Para deducir locus que están hipometilados en los tejidos placentarios con respecto a células sanguíneas maternas, se clasificó con respecto a locus que mostraban $>80\%$ de metilación en células sanguíneas maternas y $\leq 40\%$ de metilación de acuerdo con el valor predicho con una diferencia de al menos 50 % entre la densidad de metilación de células sanguíneas y el valor predicho.

La FIGURA 16 es una tabla 1600 que muestra los números de locus que se ha predicho que están hipermetilados o hipometilados basándose en el análisis directo de los datos de secuenciación por bisulfito de plasma materno. "N/A" significa no aplicable. "a" indica que la búsqueda de locus hipermetilados partió de la lista de locus que mostraban densidades de metilación $<20\%$ en las células sanguíneas maternas. "b" indica que la búsqueda de locus hipometilados partió de la lista de locus que mostraban densidades de metilación $>80\%$ en las células sanguíneas maternas. "c" indica que se usaron datos de secuenciación por bisulfito de la muestra de vellosidad coriónica para verificar los datos de plasma materno del primer trimestre, y se usó el tejido placentario a término para verificar los datos de plasma materno del tercer trimestre.

Como se muestra en la tabla 1600, una mayoría de los locus deducidos de forma no invasiva mostraron el patrón de metilación esperado en los tejidos y solaparon con las DMR obtenidas de los datos de tejido y presentadas en la sección anterior. El apéndice enumera DMR identificadas del plasma. La Tabla S3A enumera los 100 locus superiores que se ha deducido que están hipermetilados a partir de los datos de secuenciación por bisulfito de plasma materno del primer trimestre. La Tabla S3B enumera los 100 locus superiores que se ha deducido que están hipometilados de los datos de secuenciación por bisulfito de plasma materno del primer trimestre. La Tabla S3C enumera los 100 locus superiores que se ha deducido que están hipermetilados de los datos de secuenciación por bisulfito de plasma materno del tercer trimestre. La Tabla S3D enumera los 100 locus superiores que se ha deducido que están hipometilados de los datos de secuenciación por bisulfito del plasma materno del tercer trimestre.

C. Variación gestacional en metilomas placentarios y fetales

La proporción general de CpG metilados en la CVS fue del 55 % mientras que fue del 59 % para la placenta a término (tabla 100 de la FIGURA 1). Pudieron identificarse más DMR hipometiladas a partir de CVS que la placenta a término mientras que el número de DMR hipermetiladas fue similar para los dos tejidos. Por lo tanto, fue evidente que la CVS estaba más hipometilada que la placenta a término. Esta tendencia gestacional también resultó evidente en los datos de plasma materno. La proporción de CpG metilados entre las lecturas específicas fetales fue del 47,0 % en el plasma materno del primer trimestre, pero fue del 53,3 % en el plasma materno del tercer trimestre. Los números de locus hipermetilados validados fueron similares en las muestras de plasma materno del primer (1.457 locus) y tercer trimestre (1.279 locus) pero hubo sustancialmente más locus hipometilados en las muestras del primer (21.812 locus) que del tercer trimestre (12.677 locus) (tabla 1600 de la FIGURA 16).

D. Uso de marcadores

Los marcadores metilados diferencialmente, o DMR, son útiles en varios aspectos. La presencia de dichos marcadores en plasma materno indica y confirma la presencia de ADN fetal o placentario. Esta confirmación puede usarse como un control de calidad para ensayos prenatales no invasivos. Las DMR pueden actuar como marcadores de ADN fetal genéricos en plasma materno y tienen ventajas sobre marcadores que se basan en diferencias genotípicas entre la

madre y el feto, tales como marcadores basados en polimorfismo o los basados en el cromosoma Y. Las DMR son marcadores fetales genéricos que son útiles para todos los embarazos. Los marcadores basados en polimorfismo solamente son aplicables al subconjunto de embarazos cuando el feto ha heredado el marcador de su padre y cuando la madre no posee este marcador en su genoma. Además, se podría medir la concentración de ADN fetal en una muestra de plasma materno cuantificando las moléculas de ADN que se originan de esas DMR. Conociendo el perfil de DMR esperado para embarazos normales, podrían detectarse complicaciones asociadas con el embarazo, particularmente las que implican cambios del tejido placentario, observando una desviación en el perfil de DMR de plasma materno o el perfil de metilación con respecto al esperado para embarazos normales. Las complicaciones asociadas con el embarazo que implican cambios de tejido placentario incluyen, pero sin limitación, aneuploidías cromosómicas fetales. Los ejemplos incluyen la trisomía del 21, preeclampsia, retardo del crecimiento intrauterino y parto prematuro.

E. Kits que usan marcadores

Las realizaciones pueden proporcionar composiciones y kits para practicar los métodos descritos en el presente documento y otros métodos aplicables. Los kits pueden usarse para llevar a cabo ensayos para analizar el ADN fetal, por ejemplo, ADN fetal sin células en plasma materno. En una realización, un kit puede incluir al menos un oligonucleótido útil para hibridación específica con uno o más locus identificados en el presente documento. Un kit también puede incluir al menos un oligonucleótido útil para hibridación específica con uno o más locus de referencia. En una realización, se miden marcadores hipermetilados placentarios. El locus de ensayo puede ser el ADN metilado en plasma materno y el locus de referencia puede ser el ADN metilado en plasma materno. Un kit similar puede componerse para analizar ADN tumoral en plasma.

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

En diversas realizaciones, se proporciona un kit para analizar ADN fetal en una muestra biológica que contiene una mezcla de ADN fetal y ADN de una mujer embarazada con 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 lo tanto, puede usarse cualquier número de oligonucleótidos de entre todas las tablas o solamente de una tabla. Los oligonucleótidos pueden actuar como cebadores y pueden organizarse como pares de cebadores, en los que un par corresponde a una región particular de las tablas.

VI. RELACIÓN DEL TAMAÑO Y LA DENSIDAD DE METILACIÓN

Se sabe que existen moléculas de ADN en plasma en circulación en forma de moléculas cortas, siendo la mayoría de las moléculas de aproximadamente 160 pb de longitud (YMD Lo *et al.* 2010 Sci Transl Med; 2: 61ra91, YW Zheng *et al.* 2012 Clin Chem; 58: 549-558). Resulta interesante que los datos de los inventores han revelado una relación entre el estado de metilación y el tamaño de las moléculas de ADN de plasma. Por lo tanto, la longitud de los fragmentos de ADN de plasma está ligada al nivel de metilación del ADN. Los perfiles de tamaño característicos de moléculas de ADN en plasma sugieren que la mayoría están asociados con mononucleosomas, posiblemente derivados de degradación enzimática durante la apoptosis.

El ADN en circulación está fragmentado en la naturaleza. En particular, el ADN fetal en circulación es más corto que el ADN derivado de la madre en muestras de plasma materno (KCA Chan *et al.* 2004 Clin Chem; 50: 88-92). Ya que el alineamiento de extremos emparejados permite el análisis del tamaño de ADN tratado por bisulfito, se podría evaluar directamente si existe alguna correlación entre el tamaño de las moléculas de ADN de plasma y sus niveles de metilación respectivos. Se ha explorado esto en el plasma materno, así como una muestra de plasma de control de mujer adulta no embarazada.

Se usó secuenciación de extremos emparejados (que incluye secuenciación de una molécula completa) para ambos extremos de cada molécula de ADN para analizar cada muestra en este estudio. Alineando el par de secuencias finales de cada molécula de ADN con el genoma humano de referencia y observando las coordenadas del genoma de los extremos de las lecturas secuenciadas, se pueden determinar las longitudes de las moléculas de ADN secuenciadas. Las moléculas de ADN de plasma se fragmentan de forma natural en moléculas pequeñas y las bibliotecas de secuenciación para ADN de plasma se preparan típicamente sin ninguna etapa de fragmentación. Por lo tanto, las longitudes deducidas por la secuenciación representaron los tamaños de las moléculas de ADN de plasma originales.

En un estudio previo, se han determinado los perfiles de tamaño de las moléculas de ADN fetal y materno en plasma materno (YMD Lo *et al.* 2010 Sci Transl Med; 2: 61ra91). Se ha mostrado que las moléculas de ADN de plasma tenían

tamaños que se asemejaban a mononucleosomas y las moléculas de ADN fetal eran más cortas que las maternas. En este estudio, se ha determinado la relación del estado de metilación de moléculas de ADN de plasma con sus tamaños.

5 A. Resultados

La FIGURA 17A es una representación 1700 que muestra la distribución de tamaños de ADN de plasma materno, plasma de control de mujer no embarazada, placentario y de sangre periférica. Para la muestra materna y el plasma de control de mujer no embarazada, las dos muestras de plasma tratadas por bisulfito presentaron la misma
10 distribución de tamaños característica que se ha indicado previamente (YMD 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 moléculas de ADN más cortas de 143 pb.

La FIG. 17B es una representación 1750 de la distribución de tamaños y perfil de metilación de 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 contienen al menos un sitio de CpG, se calculó su densidad de metilación media. Después se representó la relación entre los tamaños de las moléculas de ADN y sus densidades de metilación. Específicamente, la densidad de metilación media se determinó para cada longitud de fragmento que variaba de 50 pb hasta 180 pb para lecturas secuenciadas que abarcaban al menos 1 sitio de CpG. Resulta interesante que la densidad de metilación
15 aumentó con el tamaño del ADN de plasma y alcanzó un máximo a aproximadamente 166-167 pb. Este patrón, sin embargo, no se observó en las muestras de ADN de placenta y sangre de control que se fragmentaron usando un sistema de ultrasonidos.

La FIGURA 18 muestra representaciones de densidades de metilación y tamaño de moléculas de ADN de plasma. La FIGURA 18A es una representación 1800 para el plasma materno del primer trimestre. La FIGURA 18B es una representación 1850 para el plasma materno del tercer trimestre. Los datos para todas las lecturas secuenciadas que abarcaban al menos un sitio de CpG se representan por la curva azul 1805. Los datos para lecturas que también contenían un alelo de SNP específico fetal se representan por la curva roja 1810. Los datos para lecturas que también contenían un alelo de SNP específico materno se representan por la curva verde 1815.
25

Se consideró que las lecturas que contenían un alelo de SNP específico fetal habían derivado de moléculas de ADN fetal. Se consideró que las lecturas que contenían un alelo de SNP específico materno habían derivado de moléculas de ADN materno. En general, las moléculas de ADN con densidades de metilación altas tenían un tamaño mayor. Esta tendencia estaba presente en moléculas de ADN tanto fetales como maternas tanto en el primer como en el tercer trimestre. Los tamaños generales de las moléculas de ADN fetal fueron más cortos que los maternos como se ha indicado previamente.
30

La FIGURA 19A muestra una representación 1900 de densidades de metilación y los tamaños de lecturas secuenciadas para una mujer no embarazada adulta. La muestra de ADN de plasma de la mujer no embarazada adulta también mostró la misma relación entre los tamaños y el estado de metilación de las moléculas de ADN. Por otro lado, las muestras de ADN genómico se fragmentaron por una etapa de ultrasonidos antes del análisis de SPM. Como se muestra en la representación 1900, los datos de las muestras de células sanguíneas y tejido placentario no revelaron la misma tendencia. Ya que la fragmentación de las células es artificial, se esperaba que no hubiera relación de tamaño y densidad. Ya que las moléculas de ADN fragmentadas de forma natural en el plasma sí muestran una dependencia del tamaño, se puede suponer que las densidades de metilación menores hacen más probable que las moléculas se rompan en fragmentos más pequeños.
35

La FIGURA 19B es una representación 1950 que muestra la distribución de tamaños y el perfil de metilación de moléculas de ADN específicas fetales y específicas maternas en plasma materno. Las moléculas de ADN de plasma específicas fetales y específicas maternas también mostraron la misma correlación entre el tamaño de los fragmentos y el nivel de metilación. La longitud de los fragmentos de ADN sin células tanto derivado de placenta como en circulación materno aumentó con el nivel de metilación. Además, la distribución de su estado de metilación no solapó entre sí, lo que sugiere que el fenómeno existe independientemente de la longitud del fragmento original de las fuentes de moléculas de ADN en circulación.
40

55 B. Método

En consecuencia, puede usarse una distribución de tamaños para estimar un porcentaje de metilación total de una muestra de plasma. Esta medición de metilación puede después seguirse durante el embarazo, durante la supervisión del cáncer o durante el tratamiento por medición en serie de las distribuciones de tamaños del ADN de plasma de acuerdo con la relación mostrada en las FIGURAS 18A y 18B. La medición de metilación también puede usarse para buscar la liberación aumentada o reducida de ADN de un órgano o un tejido de interés. Por ejemplo, se pueden buscar específicamente identificaciones de metilación de ADN específicas para un órgano específico (por ejemplo, el hígado) y medir las concentraciones de estas identificaciones en plasma. Ya que se libera ADN al plasma cuando las células mueren, un aumento en los niveles podría significar un aumento en la muerte celular o daño en ese órgano o tejido particular. Una reducción en el nivel de un órgano particular puede significar que el tratamiento para contrarrestar el
60

daño o procesos patológicos en ese órgano está bajo control.

La FIGURA 20 es un diagrama de flujo de un método 2000 para estimar un nivel de metilación de ADN en una muestra biológica de un organismo de acuerdo con realizaciones de la presente divulgación. El nivel de metilación puede estimarse para una región particular de un genoma o el genoma completo. Si se desea una región específica, entonces pueden usarse fragmentos de ADN solamente de esa región específica.

En el bloque 2010, se miden las cantidades de fragmentos de ADN correspondientes a diversos tamaños. Para cada tamaño de una pluralidad de tamaños, se puede medir una cantidad de una pluralidad de fragmentos de ADN de la muestra biológica correspondiente al tamaño. Por ejemplo, se puede medir el número de fragmentos de ADN que tienen una longitud de 140 bases. Las cantidades pueden guardarse como un histograma. En una realización, se mide un tamaño de cada uno de la pluralidad de ácidos nucleicos de la muestra biológica, lo que puede realizarse individualmente (por ejemplo, mediante secuenciación de moléculas individuales de una molécula completa o solamente los extremos de la molécula) o en grupos (por ejemplo, mediante electroforesis). Los tamaños pueden corresponder a un intervalo. Por lo tanto, una cantidad puede ser para fragmentos de ADN que tienen un tamaño dentro de un intervalo particular. Cuando se realiza secuenciación de extremos emparejados, los fragmentos de ADN (como se determinan por las lecturas de secuencias emparejadas) que se mapean (se alinea) en una región particular pueden usarse para determinar el nivel de metilación de la región.

En el bloque 2020, se calcula un primer valor de un primer parámetro basándose en las cantidades de fragmentos de ADN en múltiples tamaños. En un aspecto, el primer parámetro proporciona una medida estadística de un perfil de tamaños (por ejemplo, un histograma) de fragmentos de ADN en la muestra biológica. El parámetro puede denominarse parámetro de tamaños ya que se determina a partir de los tamaños de la pluralidad de fragmentos de ADN.

El primer parámetro puede ser de diversas formas. Un parámetro es el porcentaje de fragmento de ADN de un tamaño particular o un 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 en un tamaño particular dividido por el número total de fragmentos, que puede obtenerse a partir de un histograma (cualquier estructura de datos que proporciona recuentos absolutos o relativos de fragmentos a tamaños particulares). Como otro ejemplo, un parámetro podría ser un número de fragmentos en un tamaño particular o dentro de un intervalo particular dividido por un número de fragmentos de otro tamaño o intervalo. La división puede actuar como una normalización para representar un número diferente de fragmentos de ADN que se analizan para diferentes muestras. Puede realizarse una normalización analizando un mismo número de fragmentos de ADN para cada muestra, lo que proporciona de forma eficaz un mismo resultado que la división por un número total de fragmentos analizados. Pueden encontrarse ejemplos adicionales de parámetros y acerca de análisis de tamaños en la Solicitud de Patente de Estados Unidos 13/789.553.

En el bloque 2030, el primer valor de tamaño se compara con un valor de tamaño de referencia. El valor de tamaño de referencia puede calcularse a partir de fragmentos de ADN de una muestra de referencia. Para determinar los valores de tamaños 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 lo tanto, cuando el primer valor de tamaño se compara con el valor de tamaño de referencia, puede determinarse un nivel de metilación.

En el bloque 2040, el nivel de metilación se estima basándose en la comparación. En una realización, se puede determinar si el primer valor del primer parámetro está por encima o por debajo del valor de tamaño de referencia, y de este modo determinar si el nivel de metilación de la presente muestra está por encima o por debajo del nivel de metilación del valor de tamaño de referencia. En otra realización, la comparación se consigue introduciendo el primer valor en una función de calibración. La función de calibración puede comparar eficazmente el primer valor con valores de calibración (un conjunto de valores de tamaño de referencia) identificando el punto en una curva correspondiente al primer valor. El nivel de metilación estimado se proporciona después como el valor de salida de la función de calibración.

En consecuencia, se puede calibrar un parámetro de tamaño con un nivel de metilación. Por ejemplo, puede medirse un nivel de metilación y asociarse con un parámetro de tamaño particular para esa muestra. Después puntos de datos de diversas muestras pueden ajustarse a una función de calibración. En una implementación, pueden usarse diferentes funciones de calibración para diferentes subconjuntos de ADN. Por lo tanto, puede haber alguna forma de calibración basada en el conocimiento previo acerca de la relación entre la metilación y el tamaño para un subconjunto particular de ADN. Por ejemplo, la calibración para ADN fetal y materno podría ser diferente.

Como se ha mostrado anteriormente, la placenta está más hipometilada en comparación con la sangre materna, y por lo tanto el ADN fetal es más pequeño debido a la menor metilación. En consecuencia, puede usarse un tamaño promedio de los fragmentos de una muestra (u otro valor estadístico) para estimar la densidad de metilación. Como los tamaños de los fragmentos pueden medirse usando secuenciación de extremos emparejados, en lugar de la secuenciación sensible a la metilación potencialmente técnicamente más compleja, este enfoque sería potencialmente rentable si se usara clínicamente. Este enfoque puede usarse para supervisar los cambios de metilación asociados con el progreso de embarazo o con trastornos asociados con el embarazo tales como preeclampsia, parto prematuro

y trastornos fetales (tales como los provocados por anomalías cromosómicas o genéticas o retardo del crecimiento intrauterino).

En otra realización, este enfoque puede usarse para detectar y supervisar cáncer. Por ejemplo, con el tratamiento exitoso del cáncer, el perfil de metilación en plasma u otro fluido corporal como se mide usando este enfoque basado en el tamaño cambiaría hacia el de individuos sanos sin cáncer. Por el contrario, en el caso de que el cáncer esté progresando, entonces el perfil de metilación en plasma u otro fluido corporal sería distinto del de individuos sanos sin cáncer.

En resumen, las moléculas hipometiladas fueron más cortas que las hipermetiladas en plasma. Se observó la misma tendencia en las moléculas de ADN tanto fetales como maternas. Ya que se sabe que la metilación de ADN influye en el empaquetamiento de nucleosomas, los datos de los inventores sugieren que quizás las moléculas de ADN hipometiladas estén empaquetadas de forma menos densa con histonas y por lo tanto sean más susceptibles a la degradación enzimática. Por otro lado, los datos presentados en las FIGURAS 18A y 18B también han mostrado que a pesar de que el ADN fetal está mucho más hipometilado que las lecturas maternas, la distribución por tamaños del ADN fetal y materno no lo separa entre sí completamente. En la FIGURA 19B, se puede ver que incluso para la misma categoría de tamaño, el nivel de metilación de lecturas específicas fetales y maternas difieren entre sí. Esta observación sugiere que el estado hipometilado del ADN fetal no es el único factor que explicaba su bajo tamaño relativo en referencia al ADN materno.

VII. ESTADO DE IMPRONTA DE LOCUS GÉNICOS

Pueden detectarse moléculas de ADN derivadas de fetos que comparten el mismo genotipo, pero con diferentes identificaciones epigenéticas a la madre en el plasma materno (LLM Poon *et al.* 2002 Clin Chem; 48: 35-41). Para demostrar que el enfoque de secuenciación es sensible para captar moléculas de ADN derivadas de fetos en plasma materno, se aplicó la misma estrategia para detectar los alelos fetales con impronta en muestra de plasma materno. Se identificaron dos regiones con impronta genómica: H19 (cr11: 1.977.419-1.977.821, NCBI versión 36/hg18) y MEST (cr7: 129.917.976-129.920.347, NCBI versión 36/hg18). Ambas contienen SNP informativos para diferenciación entre las secuencias maternas y fetales. Para H19, un gen expresado por vía materna, la madre era homocigota (A/A) y el feto era heterocigoto (A/C) para el SNP rs2071094 (cr11: 1.977.740) en la región. Uno de los alelos maternos A estaba completamente metilado y el otro estaba desmetilado. En la placenta, sin embargo, el alelo A estaba desmetilado mientras que el alelo C heredado por vía paterna estaba completamente metilado. Se detectaron dos lecturas metiladas con el genotipo C, correspondientes a los alelos paternos con impronta derivados de la placenta, en plasma materno.

MEST, también conocido como PEG1, es un gen expresado por vía paterna. Tanto la madre como el feto eran heterocigotos (A/G) para el SNP rs2301335 (cr7: 129.920.062) dentro del locus con impronta. El alelo G estaba metilado mientras que el alelo A estaba desmetilado en sangre materna. El patrón de metilación se invirtió en la placenta estando el alelo materno A metilado y el alelo paterno G desmetilado. Tres alelos G desmetilados, que derivaban por vía paterna, eran detectables en plasma materno. Por el contrario, VAV1, un locus génico sin impronta en el cromosoma 19 (cr19: 6.723.621-6.724.121), no presentó ningún patrón de metilación alélico en el tejido, así como en las muestras de ADN de plasma.

Por lo tanto, el estado de metilación puede usarse para determinar que los fragmentos de ADN son del feto. Por ejemplo, no puede usarse solamente la detección del alelo A en el plasma materno como un marcador fetal cuando la madre es heterocigota GA. Pero si se distingue el estado de metilación de las moléculas A en plasma, las moléculas A metiladas son específicas fetales mientras que las moléculas A desmetiladas son específicas maternas, o viceversa.

A continuación, los inventores se centraron en locus que se había indicado que demostraban impronta genómica en tejidos placentarios. Basándose en la lista de locus presentados por Woodfine *et al.* (2011 Epigenetics Chromatin; 4: 1), se clasificaron adicionalmente los que contenían SNP dentro de la región de control de la impronta. Cuatro locus cumplieron los criterios y fueron H19, KCNQ10T1, MEST y NESP.

Con respecto a las lecturas de la muestra de células sanguíneas maternas para H19 y KCNQ10T1, las lecturas maternas fueron homocigotas para el SNP y hubo proporciones aproximadamente iguales de lecturas metiladas y desmetiladas. La CVS y la muestra de tejido placentario a término revelaron que el feto era heterocigoto para ambos locus y cada alelo se metiló o desmetiló de forma exclusiva, es decir que mostraban metilación monoalélica. En las muestras de plasma materno, las moléculas de ADN fetal heredadas por vía paterna se detectaron para ambos locus. Para H19, las moléculas heredadas por vía paterna se representaron por las lecturas secuenciadas que contenían el alelo específico fetal y se metilaron. Para KCNQ10T1, las moléculas heredadas por vía paterna se representaron por las lecturas secuenciadas que contenían el alelo específico fetal y estaban desmetiladas.

Por otro lado, la madre era heterocigota tanto para MEST como para NESP. Para MEST, tanto la madre como el feto eran heterocigotos GA para el SNP. Sin embargo, como resulta evidente a partir de los datos para la cadena de Watson para las células sanguíneas maternas y el tejido placentario, el estado de metilación para los CpG adyacentes al SNP era opuesto en la madre y el feto. El alelo A estaba desmetilado en el ADN de la madre, pero estaba metilado

en el ADN del feto. Para MEST, el alelo materno estaba metilado. Por lo tanto, se podría apuntar a que el feto había heredado el alelo A de su madre (metilado en la CVS) y la madre había heredado el alelo A de su padre (desmetilado en las células sanguíneas maternas). Resulta interesante que, en las muestras de plasma materno, los cuatro grupos de moléculas podrían distinguirse fácilmente, incluyendo cada uno de los dos alelos de la madre y cada uno de los dos alelos del feto. Por lo tanto, combinando la información genotípica con el estado de metilación en los locus con impronta, se podrían distinguir fácilmente las moléculas de ADN fetal heredadas por vía materna de las moléculas de ADN maternas de fondo (LLM Poon *et al.* 2002 Clin Chem; 48: 35-41).

Este enfoque podría usarse para detectar disomía uniparental. Por ejemplo, si se sabe que el padre de este feto es homocigoto para el alelo G, la ausencia de detección del alelo G desmetilado en plasma materno significa la falta de contribución del alelo paterno. Además, en dicha circunstancia, cuando tanto el alelo G metilado como el alelo A metilado se detectaron en el plasma de este embarazo, se sugeriría que el feto tiene heterodisomía de la madre, es decir, que hereda dos alelos diferentes de la madre sin ninguna herencia del padre. Como alternativa, si tanto el alelo A metilado (alelo fetal heredado de la madre) como el alelo A desmetilado (alelo materno heredado del abuelo materno) se detectaron en plasma materno sin el alelo G desmetilado (alelo paterno que debería haberse heredado por el feto), se sugeriría que el feto tiene isodisomía de la madre, es decir que hereda dos alelos idénticos de la madre sin ninguna herencia del padre.

Para NESP, la madre era una heterocigota GA en el SNP mientras que el feto era homocigoto para el alelo G. El alelo paterno estaba metilado para NESP. En las muestras de plasma materno, los alelos G fetales heredados por vía paterna que estaban metilados podrían distinguirse fácilmente de los alelos G maternos de fondo que estaban desmetilados.

VIII. CÁNCER/DONANTES

Algunas realizaciones pueden usarse para la detección, exploración, supervisión (por ejemplo, con respecto a recaída, remisión o respuesta (por ejemplo, presencia o ausencia) al tratamiento), estadificación, clasificación (por ejemplo, para ayudar a elegir la modalidad de tratamiento más apropiada) y pronóstico del cáncer usando análisis de metilación de ADN de plasma/suero en circulación.

Se sabe que el ADN de cáncer demuestra metilación de ADN aberrante (JG Herman *et al.* 2003 N Engl J Med; 349: 2042-2054). Por ejemplo, los promotores de islas de CpG de genes, por ejemplo, genes supresores de tumores, están hipermetilados mientras que los sitios de CpG en el cuerpo del gen están hipometilados en comparación con células no cancerosas. Siempre que el perfil de metilación de las células cancerosas pueda reflejarse por el perfil de metilación de las moléculas de ADN de plasma derivadas de tumores usando métodos descritos en el presente documento, se espera que el perfil de metilación general en plasma sea diferente entre 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 con respecto a 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 tejidos cancerosos (Gama-Sosa MA *et al.* 1983 Nucleic Acids Res; 11: 6883-6894), se observaría reducción en las densidades de metilación en el metiloma de plasma o segmentos del genoma en plasma de pacientes con cáncer.

Los cambios cualitativos en el perfil de metilación también deberían reflejarse entre los datos del metiloma de plasma. Por ejemplo, las moléculas de ADN de plasma que se originan de genes que están hipermetilados solamente en células cancerosas mostrarían hipermetilación en plasma de un paciente con cáncer en comparación con moléculas de ADN de plasma que se originan de los mismos genes, pero en una muestra de un control sano. Debido a que se produce metilación aberrante en la mayoría de los cánceres, los métodos descritos en el presente documento podrían aplicarse a la detección de todas las formas de tumores malignos con metilación aberrante, por ejemplo, tumores malignos en, pero sin limitación, el pulmón, la mama, el colon y el recto, la próstata, la nasofaringe, el estómago, los testículos, la piel, el sistema nervioso, el hueso, el ovario, el hígado, tejidos hematológicos, el páncreas, el útero, el riñón, la vejiga, tejidos linfoides, etc. Los tumores malignos pueden ser de una diversidad de subtipos histológicos, por ejemplo, carcinomas, adenocarcinomas, sarcomas, fibroadenocarcinoma, neuroendocrinos e indiferenciados, etc.

Por otro lado, se espera que las moléculas de ADN derivadas de tumores puedan distinguirse de las moléculas de ADN no derivadas de tumores de fondo debido a que el perfil de tamaños en general cortos del ADN derivado de tumores se acentúa para moléculas de ADN que se originan de locus con hipometilación aberrante asociada a tumores lo que tendría un efecto adicional en el tamaño de la molécula de ADN. Además, las moléculas de ADN de plasma derivadas de tumores pueden distinguirse de las moléculas de ADN de plasma no derivadas de tumores de fondo usando múltiples elementos característicos que se asocian con ADN tumoral, incluyendo, pero sin limitación, variantes de un único nucleótido, ganancias y pérdidas del número de copias, translocaciones, inversiones, hipermetilación o hipometilación aberrante y perfiles de tamaños. Como todos estos cambios podrían producirse de forma independiente, el uso combinado de estos elementos puede proporcionar una ventaja aditiva para la detección sensible y específica de ADN de cáncer en plasma.

A. Tamaño y cáncer

El tamaño de moléculas de ADN derivadas de tumores en plasma también se asemeja a los tamaños de unidades mononucleosómicas y son más cortas que las moléculas de ADN no derivadas de tumores de fondo, que coexisten en el plasma de enfermos de cáncer. Se ha mostrado que los parámetros de tamaño están correlacionados con el

Ya que el ADN tanto derivado del feto como derivado de la madre en plasma ha mostrado 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

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

En este ejemplo, se analizó el plasma y las muestras tisulares de un paciente con carcinoma hepatocelular (CHC). Se recogieron muestras de sangre del paciente con CHC antes de la resección quirúrgica del tumor y una semana después. Se recogieron el plasma y la capa leucoplaquetaria después de centrifugación de las muestras sanguíneas. Se recogieron el tumor resecado y el tejido de hígado no tumoral adyacente. Las muestras de ADN extraídas del plasma y muestras tisulares se analizaron usando secuenciación paralela masiva con y sin tratamiento por bisulfito previo. El ADN de plasma de cuatro individuos sanos sin cáncer se analizó también como controles. El tratamiento por bisulfito de una muestra de ADN convertiría los restos de citosina no metilados en uracilo. En la reacción en cadena de la polimerasa cadena abajo y la secuenciación, estos restos de uracilo se comportarían como timidina. Por otro lado, el tratamiento por bisulfito no convertiría los restos de citosina metilados en uracilo. Después de secuenciar de forma paralela masiva, las lecturas de secuenciación se analizaron mediante el Methy-Pipe (P Jiang, *et al.* Methy-Pipe: An integrated bioinformatics data analysis pipeline for whole genome methylome analysis, artículo presentado en los Talleres de la Conferencia Internacional sobre Bioinformática y Biomedicina de IEEE, Hong Kong, del 18 al 21 de diciembre de 2010), para determinar el estado de metilación de los restos de citosina en todas las posiciones de dinucleótidos CG, es decir sitios de CpG.

La FIGURA 21A es una tabla 2100 que muestra las densidades de metilación del plasma preoperatorio y las muestras tisulares de un paciente con CHC. La densidad de metilación de CpG para las regiones de interés (por ejemplo, sitios de CpG, regiones promotoras o repetidas etc.) se refiere a la proporción de lecturas que muestran metilación de CpG sobre el número total de lecturas que abarcan dinucleótidos de CpG genómicos. Las densidades de metilación de la capa leucoplaquetaria y el tejido hepático no tumoral son similares. La densidad de metilación general del tejido tumoral, basándose en datos de todos los autosomas, fue 25 % menor que la de la capa leucoplaquetaria y el tejido hepático no tumoral. La hipometilación fue uniforme entre cada cromosoma individual. La densidad de metilación del plasma estuvo entre los valores de los tejidos no malignos y los tejidos de cáncer. Esta observación es coherente con el hecho de que los tejidos tanto de cáncer como no de cáncer contribuirían al ADN en circulación de un paciente con cáncer. Se ha mostrado que el sistema hematopoyético es la fuente principal del ADN en circulación en individuos sin una afección maligna activa (YYN Lui, *et al.*, 2002 Clin Chem; 48: 421-7). También se han analizado por lo tanto muestras de plasma obtenidas de cuatro controles sanos. El número de lecturas de secuencias y la profundidad de secuenciación conseguida por muestra se muestran en la tabla 2150 de la FIGURA 21B.

La FIGURA 22 es una tabla 220 que muestra que las densidades de metilación en los autosomas variaron de 71,2 % a 72,5 % en las muestras de plasma de los controles sanos. Estos datos mostraron el nivel esperado de metilación de 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 a la circulación (KCA Chan *et al.* 2013 Clin Chem; 59: 211-224); RJ Leary *et al.* 2012 Sci Transl Med; 4: 162ra154). Debido a la naturaleza hipometilada del tumor de CHC, la presencia de ADN tanto derivado de tumor como no derivado de tumor en el plasma preoperatorio del paciente dio como resultado una reducción de la densidad de metilación en comparación con los niveles en plasma de controles sanos. De hecho, la densidad de metilación de la muestra de plasma preoperatoria estuvo 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 de plasma de 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 de tumor en la circulación. Una densidad de metilación menor del tejido tumoral y una mayor concentración fraccional de ADN derivado de tumor en la circulación conduciría a una menor densidad de metilación del ADN de plasma en un paciente con cáncer. Se ha indicado que la mayoría de los tumores muestran hipometilación global (JG Herman *et al.* 2003 N Engl J Med; 349: 2042-2054; MA Gama-Sosa *et al.* 1983 Nucleic Acids Res; 11: 6883-6894). Por lo tanto, las observaciones actuales vistas en las muestras de CHC también serían aplicables a otros tipos de tumores.

En una realización, la densidad de metilación del ADN de plasma puede usarse para determinar la concentración fraccional de ADN derivado de tumor en una muestra de plasma/suero cuando se conoce el nivel de metilación del tejido tumoral. El nivel de metilación, por ejemplo, densidad de metilación, del tejido tumoral puede obtenerse si la muestra tumoral está disponible o si está disponible una biopsia del tumor. En otra realización, la información con respecto al nivel de metilación del tejido tumoral puede obtenerse de una investigación del nivel de metilación en un grupo de tumores de un tipo similar y esta información (por ejemplo, un nivel medio o una mediana del nivel) se aplica al paciente para analizar usando la tecnología descrita en esta divulgación. El nivel de metilación del tejido tumoral

puede determinarse por el análisis del tejido tumoral del paciente o inferirse a partir del análisis de los tejidos tumorales de otros pacientes con el mismo tipo de cáncer o uno similar. La metilación de tejidos tumorales puede determinarse usando una serie de plataformas sensibles a metilación, incluyendo, pero sin limitación secuenciación paralela masiva, secuenciación molecular individual, micromatrices (por ejemplo, matrices de oligonucleótidos) o espectrometría de masas (tales como el análisis de Epityper, Sequenom, Inc.). En algunas realizaciones, dichos análisis pueden estar precedidos por procedimientos que son sensibles al estado de metilación de moléculas de ADN, incluyendo, pero sin limitación, inmunoprecipitación de citosina y digestión con enzimas de restricción sensibles a metilación. Cuando se conoce el nivel de metilación de un tumor, la concentración fraccional de ADN tumoral en el plasma de pacientes con cáncer podría calcularse después del análisis del metiloma en plasma.

La relación entre el nivel de metilación de plasma, P , con la concentración de ADN tumoral fraccional, f , y el nivel de metilación de tejido tumoral, TUM , puede describirse como: $P = BKG \times (1-f) + TUM \times f$, donde BKG es el nivel de metilación de ADN de fondo en plasma derivado de células sanguíneas y otros órganos internos. Por ejemplo, se ha mostrado que la densidad de metilación general de todos los autosomas era del 42,9 % en el tejido de biopsia tumoral obtenido de este paciente con CHC, es decir el valor de TUM para este caso. La densidad de metilación media de las muestras de plasma de los cuatro controles sanos fue del 71,6 %, es decir el valor de BKG de este caso. La densidad de metilación de plasma para el plasma preoperatorio fue del 59,7 %. Usando estos valores, se estima que f es 41,5 %.

En otra realización, el nivel de metilación del tejido tumoral puede estimarse de forma no invasiva basándose en los datos de metiloma de plasma cuando se conoce la concentración fraccional del ADN derivado de tumor en la muestra de plasma. La concentración fraccional del ADN derivado de tumor en la muestra de plasma puede determinarse por otros análisis genéticos, por ejemplo, el análisis de todo el genoma de pérdida alélica (GAAL) y el análisis de mutaciones de un único nucleótido como se ha descrito previamente (Solicitud de Patente de Estados Unidos 13/308.473; KCA Chan *et al.* 2013 Clin Chem; 59: 211-24). El cálculo se basa en la misma relación descrita anteriormente excepto que en esta realización se conoce el valor de f y el valor de TUM es el desconocido. La deducción puede realizarse para el genoma completo o para partes del genoma, de forma similar a los datos observados para el contexto de determinar el nivel de metilación de tejido placentario a partir de datos del plasma materno.

En otra realización, se puede usar la variación entre grupos o perfil en las densidades de metilación para diferenciar sujetos con cáncer y los que no tienen cáncer. La resolución del análisis de metilación puede aumentarse adicionalmente dividiendo el genoma en grupos de un tamaño particular, por ejemplo, 1 Mb. En dicha realización, la densidad de metilación de cada grupo de 1 Mb se calculó para las muestras recogidas, por ejemplo, la capa leucoplaquetaria, el tejido de CHC reseado, 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 que los tamaños de los grupos se mantengan constantes. En una implementación, el número de sitios de CpG se mantiene constante dentro de cada grupo mientras que el tamaño del grupo en sí mismo puede variar.

Las FIGURAS 23A y 23B muestran densidad de metilación de la capa leucoplaquetaria, el tejido tumoral, el tejido hepático no tumoral, el plasma preoperatorio y plasma postoperatorio del paciente con CHC. La FIGURA 23A es una representación 2300 de resultados para el cromosoma 1. La FIGURA 23B es una representación 2350 de resultados para el cromosoma 2.

Para la mayoría de las ventanas de 1 Mb, las densidades de metilación para la capa leucoplaquetaria y el tejido hepático no tumoral adyacente al tumor fueron similares mientras que las de los tejidos tumorales eran menores. Las densidades de metilación del plasma preoperatorio quedan entre las del tumor y los tejidos no malignos. Las densidades de metilación de las regiones genómicas consultadas en los tejidos tumorales podrían deducirse usando los datos de metilación del plasma preoperatorio y la concentración de ADN tumoral fraccional. El método es el mismo que se ha descrito anteriormente usando los valores de densidad de metilación de todos los autosomas. La deducción de la metilación tumoral descrita también puede realizarse usando estos datos de metilación de mayor resolución del ADN de plasma. También pueden usarse otros tamaños de grupos, tales como 300 kb, 500 kb, 2 Mb, 3 Mb, 5 Mb o más de 5 Mb. En una realización, no es necesario mantener constantes los tamaños de los grupos. En una implementación, el número de sitios de CpG se mantiene constante dentro de cada grupo mientras que el tamaño del grupo en sí mismo puede variar.

C. Comparación de la densidad de metilación en plasma entre el paciente con cáncer e individuos sanos

Como se muestra en 2100, las densidades de metilación del ADN de plasma preoperatorio fueron menores que las de tejidos no malignos en el paciente con cáncer. Esto probablemente se deba a la presencia de ADN del tejido tumoral que estaba hipometilado. Esta densidad de metilación de ADN de plasma inferior puede usarse potencialmente como un biomarcador para la detección y supervisión de cáncer. Para supervisión de cáncer, si un cáncer está progresando, entonces habrá una cantidad aumentada de ADN derivado de cáncer en plasma con el tiempo. En este ejemplo, una cantidad aumentada de ADN derivado de cáncer en circulación en plasma conducirá a una reducción adicional en la densidad de metilación de ADN de plasma en todo el genoma.

Por el contrario, si un cáncer responde al tratamiento, entonces la cantidad de ADN derivado de cáncer en plasma se

reducirá con el tiempo. En este ejemplo, una reducción de la cantidad de ADN derivado de cáncer en plasma conducirá a un aumento en la densidad de metilación de ADN de plasma. Por ejemplo, si un paciente con cáncer de pulmón con mutación del receptor del factor de crecimiento epidérmico se ha tratado con una terapia dirigida, por ejemplo, inhibición de tirosina quinasa, entonces un aumento en la densidad de metilación de ADN de plasma significaría una respuesta. Posteriormente, la aparición de un clon tumoral resistente a la inhibición de tirosina quinasa se asociaría con una reducción en la densidad de metilación de ADN de plasma lo que indicaría una recaída.

Pueden realizarse mediciones de densidad de metilación en plasma en serie y puede calcularse la tasa de cambio de dichas mediciones y usarse para predecir o correlacionar con la progresión clínica o remisión o pronóstico. Para locus genómicos seleccionados que están hipermetilados en tejidos de cáncer, pero hipometilados en tejidos normales, por ejemplo, las regiones promotoras de varios genes supresores de tumores, la relación entre la progresión de cáncer y la respuesta favorable al tratamiento se opondrá a los patrones descritos anteriormente.

Para demostrar la viabilidad de este enfoque, se compararon las densidades de metilación de ADN de muestras de plasma recogidas del paciente con cáncer antes y después de la retirada quirúrgica del tumor con ADN de plasma obtenido de cuatro sujetos de control sanos.

La Tabla 2200 muestra las densidades de metilación de 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 los de los cuatro sujetos de control sanos. Para todos los cromosomas, las densidades de metilación de la muestra de ADN de plasma preoperatoria fueron menores que las de la muestra postoperatoria y las muestras de plasma de los cuatro sujetos sanos. La diferencia en las densidades de metilación de ADN de plasma entre las muestras preoperatorias y postoperatorias proporcionó pruebas de apoyo de que las densidades de metilación menores en la muestra de plasma preoperatoria se debían a la presencia de ADN del tumor de CHC.

La inversión de las densidades de metilación de ADN en la muestra de plasma postoperatoria hasta niveles similares a las muestras de plasma de los controles sanos ha sugerido que mucho del ADN derivado de tumor ha desaparecido debido a la retirada quirúrgica de la fuente, es decir el tumor. Estos datos sugieren que la densidad de metilación del plasma preoperatorio como se determina usando datos disponibles de regiones genómicas grandes, tales como todos los autosomas o cromosomas individuales, fue de un nivel de metilación menor que el de los controles sanos para permitir la identificación, es decir el diagnóstico o la exploración, de que el caso de ensayo tenga cáncer.

Los datos del plasma preoperatorio también mostraron un nivel de metilación mucho menor que el del plasma postoperatorio lo que indica que el nivel de metilación del plasma también podría usarse para supervisar la carga tumoral, por lo tanto, para pronosticar y supervisar el progreso del cáncer en el paciente. Los valores de referencia pueden determinarse a partir del plasma de controles sanos o personas que están en riesgo de cáncer pero que actualmente no tienen cáncer. Las personas en riesgo de CHC incluyen las que tienen hepatitis B crónica o infección por hepatitis C, las que tienen hemocromatosis y las que tienen cirrosis hepática.

Pueden usarse valores de densidad de metilación de plasma más allá de, por ejemplo, menores que, un punto de corte definido basado en los valores de referencia para evaluar si el plasma de una persona no embarazada tiene ADN tumoral o no. Para detectar la presencia de ADN tumoral en circulación hipometilado, el punto de corte puede definirse tan bajo como el 5º o 1º percentiles 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 de densidad de metilación medios de los controles, o basarse en la determinación de un múltiplo de la mediana (MoM). Para ADN tumoral hipermetilado, el punto de corte puede definirse más alto que el 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 de densidad de metilación medios de los controles, o basarse en la determinación de un múltiplo de la mediana (MoM). En una realización, la población de control coincide en edad con el sujeto de ensayo. No es necesario que la coincidencia de edad sea exacta y puede realizarse en bandas de edad (por ejemplo, de 30 a 40 años, para un sujeto de ensayo de 35 años).

A continuación, se compararon las densidades de metilación de grupos de 1 Mb entre las muestras de plasma del paciente con cáncer y los cuatro sujetos de control. Para fines de ilustración, se muestran los resultados del cromosoma 1.

La FIGURA 24A es una representación 2400 que muestra las densidades de metilación del plasma preoperatorio del paciente con CHC. La FIGURA 24B es una representación 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 fueron menores que las de los sujetos de control para la mayoría de los grupos. Se observaron patrones similares para otros cromosomas. 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 para la mayoría de los grupos. Se observaron patrones similares para otros cromosomas.

Para evaluar si un sujeto ensayado tiene cáncer, el resultado del sujeto ensayado se compararía con los valores de un grupo de referencia. En una realización, el grupo de referencia puede comprender varios sujetos sanos. En otra realización, el grupo de referencia puede comprender sujetos con afecciones no malignas, por ejemplo, infección por hepatitis B crónica o cirrosis. La diferencia en las densidades de metilación entre el sujeto ensayado y el grupo de referencia puede después cuantificarse.

En una realización, puede obtenerse un intervalo de referencia a partir de los valores del grupo de control. Después pueden usarse desviaciones en el resultado del sujeto de ensayo de 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 de tumor en el plasma y la diferencia en el nivel de metilación entre tejidos malignos y no malignos. Una concentración fraccional mayor de ADN derivado de tumor en plasma conduciría a mayores diferencias de densidad de metilación entre la muestra de plasma de ensayo y los controles. Un mayor grado de diferencia en el nivel de metilación de los tejidos malignos y no malignos también se asocia con diferencias de densidad de metilación mayores entre la muestra de plasma de ensayo y los controles. En otra realización más, se eligen grupos de referencia diferentes para sujetos de ensayo de diferentes 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 grupo de 1 Mb. Después para grupos 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, esta diferencia se dividió después por la DT del grupo correspondiente para determinar la puntuación z . En otras palabras, la puntuación z representa la diferencia en las densidades de metilación entre las muestras de plasma de ensayo y de control expresadas como un número de DT desde la media de los sujetos de control. Una puntuación $z > 3$ de un grupo indica que el ADN de plasma del paciente con CHC está más hipermetilado que los sujetos de control en más de 3 DT en ese grupo mientras que una puntuación z de ≤ -3 en un grupo indica que el ADN de plasma del paciente con CHC está más hipometilado que los sujetos de control en más de 3 DT en ese grupo.

Las FIGURAS 25A y 25B muestran puntuaciones z de las densidades de metilación de ADN de plasma para las muestras de plasma preoperatorias (representación 2500) y postoperatorias (representación 2550) del paciente con CHC usando los datos del metiloma de plasma de los cuatro sujetos de control sanos como referencia para el cromosoma 1. Cada punto representa el resultado de un grupo de 1 Mb. Los puntos negros representan los grupos con puntuación z entre -3 y 3 . Los puntos rojos representan grupos con puntuación $z \leq -3$.

La FIGURA 26A es una tabla 2600 que muestra datos para puntuaciones z para plasma preoperatorio y postoperatorio. La mayoría de los grupos en el cromosoma 1 (80,9 %) en la muestra de plasma preoperatoria tuvo una puntuación z de ≤ -3 lo que indica que el ADN de plasma preoperatorio del paciente con CHC estaba significativamente más hipometilado que el de los sujetos de control. Por el contrario, el número de puntos rojos se redujo sustancialmente en la muestra de plasma postoperatoria (8,3 % de los grupos en el cromosoma 1) lo que sugiere que la mayoría del ADN tumoral se había retirado de la circulación debido a resección quirúrgica de la fuente de ADN tumoral en circulación.

La FIGURA 26B es una representación de Circos 2620 que muestra la puntuación z de las densidades de metilación de ADN de plasma para las muestras de plasma preoperatorias y postoperatorias del paciente con CHC usando los cuatro sujetos de control sanos como referencia para grupos de 1 Mb analizados de todos los autosomas. El anillo más externo muestra los ideogramas de los autosomas humanos. El anillo medio muestra los datos para la muestra de plasma preoperatoria. El anillo más interno muestra los datos para la muestra de plasma postoperatoria. Cada punto representa el resultado de un grupo de 1 Mb. Los puntos negros representan los grupos con puntuaciones z entre -3 y 3 . Los puntos rojos representan grupos con puntuaciones $z \leq -3$. Los puntos verdes representan grupos con puntuaciones $z > 3$.

La FIGURA 26C es una tabla 2640 que muestra una distribución de las puntuaciones z de los grupos de 1 Mb para el genoma completo en las muestras de plasma tanto preoperatorias como postoperatorias del paciente con CHC. Los resultados indican que el ADN de plasma preoperatorio del paciente con CHC estaba más hipometilado que el de los controles para la mayoría de las regiones (85,2 % de los grupos de 1 Mb) en el genoma completo. Por el contrario, la mayoría de las regiones (93,5 % de los grupos de 1 Mb) en la muestra de plasma postoperatoria no mostraron hipermetilación o hipometilación significativas en comparación con los controles. Estos datos indican que mucho del ADN tumoral, principalmente hipometilado en la naturaleza para este CHC, ya no estaba presente en la muestra de plasma postoperatoria.

En una realización, el número, porcentaje o proporción de grupos con puntuaciones $z \leq -3$ pueden usarse para indicar si está presente un cáncer. Por ejemplo, como se muestra en la tabla 2640, 2330 de los 2734 grupos analizados (85,2 %) mostraron puntuaciones $z \leq -3$ en el plasma preoperatorio mientras que solamente 171 de los 2734 grupos analizados (6,3 %) mostraron puntuaciones $z \leq -3$ 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.

Los valores de puntos de corte del número de grupos pueden determinarse usando métodos estadísticos. Por ejemplo,

se esperaría que aproximadamente 0,15 % de los grupos tuvieran una puntuación z de <-3 basándose en una distribución normal. Por lo tanto, el punto de corte del número de grupos puede ser 0,15 % del número total de grupos que se analicen. En otras palabras, si una muestra de plasma de una mujer no embarazada muestra más del 0,15 % de los grupos con puntuaciones z <-3 , hay una fuente de ADN hipometilado en el plasma, concretamente cáncer. Por ejemplo, 0,15 % de los 2734 grupos de 1 Mb que se han analizado en este ejemplo es aproximadamente 4 grupos. Usando este valor como punto de corte, las muestras de plasma tanto preoperatorias como postoperatorias contenían ADN derivado de tumor hipometilado, aunque la cantidad es mucho mayor en la muestra de plasma preoperatoria que en la muestra de plasma postoperatoria. Para los cuatro sujetos de control sanos, ninguno de los grupos mostró hipermetilación o hipometilación significativa. Pueden usarse otros valores de punto de corte (por ejemplo 1,1 %) y pueden variar dependiendo del requisito del ensayo que se use. Como otros ejemplos, el porcentaje de punto de corte puede variar basándose en la distribución estadística, así como la sensibilidad deseada y una especificidad aceptable.

En otra realización, el número de punto de corte puede determinarse por análisis de curva de característica operadora receptora (ROC) analizando varios enfermos de cáncer e individuos sin cáncer. Para validar adicionalmente la especificidad de este enfoque, se analizó una muestra de plasma de un paciente que buscaba consulta médica para una afección no maligna (C06). 1,1 % de los grupos tuvieron una puntuación z de <-3 . En una realización, pueden usarse diferentes umbrales para clasificar diferentes niveles de estado de enfermedad. Puede usarse un umbral de menor porcentaje para diferenciar el estado sano de las afecciones benignas y un umbral de mayor porcentaje para diferenciar afecciones benignas de tumores malignos.

El rendimiento de diagnóstico para análisis de hipometilación de plasma usando secuenciación paralela masiva parece ser superior al obtenido usando amplificación basada en reacción en cadena de la polimerasa (PCR) de clases específicas de elementos repetitivos, por ejemplo, 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 dominante en el genoma tumoral, tiene algún grado de heterogeneidad de una región genómica a la siguiente.

De hecho, se ha observado que las densidades de metilación de plasma medias de los sujetos de referencia variaron a lo largo del genoma (FIGURA 56). Cada punto rojo en la FIGURA 56 muestra la densidad de metilación media de un grupo de 1 Mb entre 32 sujetos sanos. La representación muestra todos los grupos de 1 Mb analizados en todo el genoma. El número dentro de cada caja representa el número de cromosoma. Se observó que las densidades de metilación medias variaban entre grupos.

Un ensayo sencillo basado en PCR no podría abarcar dicha heterogeneidad de región a región en su algoritmo de diagnóstico. Dicha heterogeneidad ampliaría el intervalo de densidades de metilación observadas entre los individuos sanos. Sería necesaria entonces una mayor magnitud de reducción en la densidad de metilación para que se considerara que una muestra mostraba hipometilación. Esto daría como resultado una reducción de la sensibilidad del ensayo.

Por el contrario, un enfoque basado en secuenciación paralela masiva divide el genoma en grupos de 1 Mb (o grupos de otros tamaños) y mide las densidades de metilación para dichos grupos individualmente. Este enfoque reduce la influencia de las variaciones en las densidades de metilación de línea basal entre diferentes regiones genómicas ya que cada región se compara entre una muestra de ensayo y los controles. De hecho, dentro del mismo grupo, la variación entre individuos a través de los 32 controles sanos fue relativamente pequeña. El 95 % de los grupos tuvieron un coeficiente de variación (CV) a través de los 32 controles sanos de $\leq 1,8$ %. Aun así, para potenciar adicionalmente la sensibilidad para la detección de hipometilación asociada al cáncer, la comparación puede realizarse a través de múltiples regiones genómicas. La sensibilidad se potenciaría ensayando múltiples regiones genómicas debido a que protegería contra el efecto de la variación biológica cuando resulte que la muestra de cáncer no demuestra hipometilación para una región particular cuando se ensaya solamente una región.

El enfoque de comparar las densidades de metilación de regiones genómicas equivalentes entre controles y muestras de ensayo (por ejemplo, ensayando cada región genómica por separado, y después posiblemente combinando dichos resultados) y realizar esta comparación para múltiples regiones genómicas tiene una relación de señal con respecto a ruido mayor para la detección de hipometilación asociada con cáncer. Este enfoque de secuenciación paralela masiva se muestra como ilustración. Se predeciría que otras metodologías que podrían determinar las densidades de metilación de múltiples regiones genómicas y permitir la comparación de densidades de metilación de regiones correspondientes entre controles y muestras de ensayo conseguirían un efecto similar. Por ejemplo, podrían diseñarse sondas de hibridación o sondas de inversión molecular que podrían dirigirse a las moléculas de ADN de plasma que se originan de regiones genómicas específicas, así como determinar un nivel de metilación de la región para conseguir el efecto deseado.

En otra realización más, la suma de las puntuaciones z para todos los grupos puede usarse para determinar si el cáncer está presente o usarse para la supervisión de los cambios en serie del nivel de metilación de ADN de plasma. Debido a la naturaleza hipometilada general del ADN tumoral, la suma de puntuaciones z sería menor en plasma recogido de un individuo con cáncer que de controles sanos. La suma de puntuaciones z para la muestra de plasma pre y postoperatoria del paciente con CHC serían -49843,8 y -3132,13, respectivamente.

En otras realizaciones, pueden usarse otros métodos para explorar el nivel de metilación de ADN de plasma. Por ejemplo, la proporción de restos de citosina metilados sobre el contenido total de restos de citosina puede determinarse usando espectrometría de masas (ML Chen *et al.*, 2013 Clin Chem; 59: 824-832) o secuenciación paralela masiva. Sin embargo, como la mayoría de los restos de citosina no están en el contexto de dinucleótidos de CpG, la proporción de citosina metilada entre restos totales de citosina totales sería relativamente pequeña en comparación con los niveles de metilación estimados en el contexto de dinucleótidos de CpG. Se determinó el nivel de metilación de las muestras tisulares y de plasma obtenidas del paciente con CHC, así como 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 usando los datos de secuenciación paralela masiva de todo el genoma. H se refiere a restos de adenina, timina o citosina.

La FIGURA 26D es una tabla 2660 que muestra los niveles de metilación del tejido tumoral y muestra de plasma preoperatoria que solapa con algunas de las muestras de plasma de control cuando se usan los contextos de CHH y CHG. Los niveles de metilación del tejido tumoral y la muestra de plasma preoperatoria fueron uniformemente menores en comparación con la capa leucoplaquetaria, el tejido hepático no tumoral, la muestra de plasma postoperatoria y las muestras de plasma de control sano en ambos entre las citosinas de CpG y no especificadas. Sin embargo, los datos basados en los CpG metilados, es decir, 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 de plasma puede determinarse por métodos que usan anticuerpos contra citosina metilada, por ejemplo, inmunoprecipitación con ADN metilado (MeDIP). Sin embargo, se espera que la precisión de estos métodos sea inferior a los métodos basados en secuenciación debido a la variabilidad en la unión de anticuerpos. En otra realización más, puede determinarse el nivel de 5-hidroximetilcitosina en ADN de plasma. A este respecto, se ha descubierto que una reducción en el nivel de 5-hidroximetilcitosina es un elemento epigenético de ciertos cánceres, por ejemplo, melanoma (CG Lian, *et al.*, 2012 Cell; 150: 1135-1146).

Además de CHC, también se ha investigado si este enfoque pudiera aplicarse a otros tipos de cánceres. Se han analizado las muestras de plasma de 2 pacientes con adenocarcinoma del pulmón (CL1 y CL2), 2 pacientes con carcinoma nasofaríngeo (NPC1 y NPC2), 2 pacientes con cáncer colorrectal (CRC1 y CRC2), 1 paciente con tumor neuroendocrino metastásico (NE1) y 1 paciente con sarcoma de músculo liso metastásico (SMS1). El ADN de plasma de estos sujetos se convirtió por bisulfito y se secuenció usando la plataforma Illumina HiSeq2000 para 50 pb en un extremo. Los cuatro sujetos de control sanos mencionados anteriormente se usaron como grupo de referencia para el análisis de estos 8 pacientes. Se usaron 50 pb de las lecturas de secuencias en un extremo. El genoma completo se dividió en grupos de 1 Mb. La media y DT de la densidad de metilación se calcularon para cada grupo usando los datos del grupo de referencia. Después se expresaron los resultados de los 8 enfermos de cáncer como puntuaciones z, que representan el número de DT de la media del grupo de referencia. Un valor positivo indica que la densidad de metilación del caso de ensayo es menor que la media del grupo de referencia, y viceversa. El número de lecturas de secuencias y la profundidad de secuenciación conseguida por muestra se muestran en la tabla 2780 de la FIGURA 27I.

Las FIGURAS 27A-H muestran representaciones de Círculos de densidad de metilación de 8 enfermos de cáncer de acuerdo con realizaciones de la presente divulgación. Cada punto representa el resultado de un grupo de 1 Mb. Los puntos negros representan los grupos con puntuaciones z entre -3 y 3. Los puntos rojos representan grupos con puntuaciones z <-3. Los puntos verdes representan grupos con puntuaciones z >3. El intervalo entre dos líneas consecutivas representa una diferencia de puntuación z de 20.

Se observó hipometilación significativa en múltiples regiones a través de los genomas para pacientes con la mayoría de los tipos de cánceres, incluyendo cáncer de pulmón, carcinoma nasofaríngeo, cáncer colorrectal y tumor neuroendocrino metastásico. Resulta interesante que, además de la hipometilación, se observó hipermetilación significativa en múltiples regiones a través del genoma en el caso de 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 cánceres en los 7 pacientes restantes es el ectodermo. Por lo tanto, es posible que el patrón de metilación de ADN del sarcoma pueda ser diferente de el del carcinoma.

Como puede verse a partir de este caso, el patrón de metilación de ADN de plasma puede ser útil para diferenciar diferentes tipos de cáncer, lo que en este ejemplo es una diferenciación de carcinoma y sarcoma. Estos datos también sugieren que el enfoque podría usarse para detectar hipermetilación aberrante asociada con el tumor maligno. Para estos 8 casos, solamente estuvieron disponibles muestras de plasma y no se hubo analizado ningún tejido tumoral. Esto mostró que incluso sin el perfil de metilación anterior o los niveles de metilación del tejido tumoral, puede detectarse fácilmente ADN derivado de tumor en plasma usando los métodos descritos.

La FIGURA 27J es una tabla 2790 que es una tabla que muestra una distribución de las puntuaciones z de los grupos de 1 Mb para el genoma completo en plasma de pacientes con diferentes tumores malignos. Los porcentajes de grupos con puntuación z <-3, -3 a 3 y >3 se muestran para cada caso. Más del 5 % de los grupos tuvieron una puntuación z de <-3 para todos los casos. Por lo tanto, si se usa un punto de corte de 5 % de los grupos que están hipometilados de forma significativa para clasificar una muestra que es positiva para el cáncer, entonces todos estos casos se

clasificarían como positivos para cáncer. Los resultados de los inventores muestran que la hipometilación probablemente sea un fenómeno general para diferentes tipos de cánceres y el análisis de metiloma de plasma sería útil para detectar diferentes tipos de cánceres.

5 D. Método

La FIGURA 28 es un diagrama de flujo del método 2800 para analizar una muestra biológica de un organismo para determinar una clasificación de un nivel de cáncer de acuerdo con realizaciones de la presente divulgación. La muestra biológica incluye ADN que se origina de células normales y puede incluir potencialmente ADN de células asociadas con cáncer. Al menos parte del ADN puede estar sin células en la muestra biológica.

En el bloque 2810, se analiza una pluralidad de moléculas de ADN de la muestra biológica. El análisis de una molécula de ADN puede incluir determinar una ubicación de la molécula de ADN en un genoma del organismo y determinar si la molécula de ADN está metilada en uno o más sitios. El análisis puede realizarse recibiendo lecturas de secuencias de una secuenciación sensible a metilación, y por lo tanto el análisis puede realizarse solamente en datos previamente obtenidos del ADN. En otras realizaciones, el análisis puede incluir la secuenciación real u otras etapas activas para obtener los datos.

En el bloque 2820, 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 de CpG, y pueden ser solamente ciertos sitios de CpG, como se selecciona usando uno o más criterios mencionados en el presente documento. El número de moléculas de ADN que están metiladas es equivalente a la determinación del número que están desmetiladas una vez que se ha realizado normalización usando un número total de moléculas de ADN analizadas en un sitio particular, por ejemplo, un número total de lecturas de secuencias. Por ejemplo, un aumento en la densidad de metilación de CpG de una región es equivalente a una reducción en la densidad de CpG desmetilados 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 basándose en el número de moléculas de ADN correspondientes a la pluralidad de sitios. Los sitios pueden corresponder a una pluralidad de locus o solamente un locus.

En el bloque 2840, el primer nivel de metilación se compara con un primer valor de punto de corte. El primer valor de punto de corte puede ser un nivel de metilación de referencia o estar relacionado con un nivel de metilación de referencia (por ejemplo, una distancia específica de un nivel normal). El nivel de metilación de referencia puede determinarse a partir de muestras 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 punto de corte puede establecerse a partir de un nivel de metilación de referencia determinado a partir de una muestra biológica previa del organismo obtenido antes de ensayarse la muestra biológica.

En una realización, el primer valor de punto de corte es una distancia específica (por ejemplo, un número específico de desviaciones 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 comparando después la diferencia con un umbral correspondiente al primer valor de punto de corte (por ejemplo, para determinar si el nivel de metilación es estadísticamente diferente al nivel de metilación de referencia).

En el bloque 2850, se determina una clasificación de un nivel de cáncer basándose en la comparación. Los 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 punto de corte puede determinarse a partir de una muestra previamente obtenida del sujeto (por ejemplo, un nivel de metilación de referencia puede determinarse a partir de la muestra previa).

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 usando criterios mencionados en el presente documento, por ejemplo, de ciertas longitudes o cierto número de sitios. Pueden identificarse uno o más sitios (por ejemplo, sitios de CpG) dentro de cada una de las regiones. Puede calcularse un nivel de metilación regional para cada región. El primer nivel de metilación es para una primera región. Cada uno de los niveles de metilación regional se compara con un valor de punto de corte regional respectivo, que puede ser el mismo o variar entre regiones. El valor de punto de corte regional para la primera región es el primer valor de punto de corte. Los valores de punto de corte regional respectivos pueden ser una cantidad específica (por ejemplo, 0,5) de un nivel de metilación de referencia, contando de este modo solamente regiones que tengan una diferencia significativa con respecto a una referencia, lo que puede determinarse a partir de sujetos sin cáncer.

Puede determinarse un primer número de regiones cuyo nivel de metilación regional supera el valor de punto de corte regional respectivo, y compararse con un valor umbral para determinar la clasificación. En una implementación, el

valor umbral es un porcentaje. La comparación del primer número con un valor umbral puede incluir dividir el primer número de regiones por un segundo número de regiones (por ejemplo, todas las regiones) antes de comparar con el valor umbral, por ejemplo, como parte de un proceso de normalización.

- 5 Como se ha descrito anteriormente, puede usarse una concentración fraccional de ADN tumoral en la muestra biológica para calcular el primer valor de punto de corte. Puede estimarse simplemente que la concentración fraccional es mayor que un valor mínimo, mientras que puede marcarse una muestra con una concentración fraccional menor que el valor mínimo, por ejemplo, como no adecuada para su análisis. El valor mínimo puede determinarse basándose en una diferencia esperada en los niveles de metilación para un tumor en relación con un nivel de metilación de referencia. Por ejemplo, si una diferencia es de 0,5 (por ejemplo, como se usa como un valor de punto de corte), entonces se requeriría que una cierta concentración tumoral fuera suficientemente alta para ver esta diferencia.

- 15 Pueden aplicarse técnicas específicas del método 1300 para el método 2800. En el método 1300, las variaciones del número de copias pueden determinarse para un tumor (por ejemplo, cuando la primera región cromosómica de un tumor pueda ensayarse con respecto a tener un cambio del número de copias relativo a una segunda región cromosómica del tumor). Por lo tanto, el método 1300 puede suponer que existe un tumor. En el método 2800, puede ensayarse una muestra con respecto a si hay un indicio de la existencia de ningún tumor en absoluto, independientemente de cualquier característica del número de copias. Algunas técnicas de los dos métodos pueden ser similares. Sin embargo, los valores de punto de corte y parámetros de metilación (por ejemplo, niveles de metilación normalizados) para el método 2800 pueden detectar una diferencia estadística a partir de un nivel de metilación de referencia para ADN no de cáncer en oposición a una diferencia de un nivel de metilación de referencia para una mezcla de ADN de cáncer y ADN no de cáncer teniendo algunas regiones posiblemente variaciones del número de copias. Por lo tanto, los valores de referencia para el método 2800 pueden determinarse a partir de muestras sin cáncer, tales como de organismos sin cáncer o de tejido no canceroso del mismo paciente (por ejemplo, plasma tomado previamente o de muestras adquiridas de forma contemporánea que se sabe que no tienen cáncer, lo que puede determinarse a partir de ADN celular).

E. Predicción de la concentración fraccional mínima de ADN tumoral para detectar usando análisis de metilación de ADN de plasma

- 30 Un modo de medir la sensibilidad del enfoque para detectar cáncer usando el nivel de metilación de ADN de plasma está relacionado con la concentración de ADN derivado de tumor fraccional mínima que se requiere para revelar un cambio en el nivel de metilación de ADN de plasma en comparación con los de controles. La sensibilidad del ensayo también depende del alcance de la diferencia en la metilación de ADN entre el tejido tumoral y los niveles de metilación de ADN de plasma de línea basal en controles sanos o ADN de células sanguíneas. Las células sanguíneas son la fuente predominante de ADN en el plasma de individuos sanos. Cuanto mayor sea la diferencia, más fácilmente se podrá diferenciar los enfermos de cáncer de los individuos sin cáncer y esto se reflejaría como un límite de detección menor de derivado tumoral en plasma y una mayor sensibilidad clínica en la detección de los enfermos de cáncer. Además, las variaciones en la metilación de ADN de plasma en los sujetos sanos o en sujetos con diferentes edades (G Hannum *et al.*, 2013 Mol Cell; 49: 359-367) también afectarían a la sensibilidad de detección de los cambios de metilación asociados con la presencia de un cáncer. Una menor variación en la metilación de ADN de plasma en los sujetos sanos facilitaría la detección del cambio causado por la presencia de una cantidad pequeña de ADN derivado de cáncer.

- 45 La FIGURA 29A es una representación 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 solamente proporciona un valor de densidad de metilación, por ejemplo, la densidad de metilación de todos los autosomas o de un cromosoma particular. Ilustra cómo la especificidad del análisis se vería afectada. En una realización, un punto de corte de 3 DT por debajo de la media de la densidad de metilación de ADN de los sujetos de referencia se usa para determinar si una muestra de ensayo está significativamente más hipometilada que muestras de los sujetos de referencia. Cuando se usa este punto de corte, se espera que aproximadamente 0,15 % de los sujetos sin cáncer tengan resultados falsos positivos de clasificación como enfermos de cáncer dando como resultado una especificidad del 99,85 %.

- 55 La FIGURA 29B es una representación 2950 que muestra las distribuciones de densidades de metilación en sujetos de referencia y enfermos de cáncer. El valor de punto de corte está 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 enfermos de cáncer está 2 DT por debajo del punto de corte (es decir 5 DT por debajo de la media de los sujetos de referencia), se esperaría que el 97,5 % de los sujetos con cáncer tuvieran una densidad de metilación por debajo del valor de punto de corte. En otras palabras, 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 del genoma completo, de todos los autosomas o de un cromosoma particular. La diferencia entre las densidades de metilación medias de las dos poblaciones se ve afectada por dos factores, concretamente el grado de diferencia en el nivel de metilación entre tejidos cancerosos y no cancerosos y la concentración fraccional de ADN derivado de tumor en la muestra de plasma. 60 Cuanto mayor sean los valores de estos dos parámetros, mayor será la diferencia en el valor de las densidades de metilación de estas dos poblaciones. Además, cuanto menor sea la DT de las distribuciones de densidades de 65

metilación de las dos poblaciones, menor será el solapamiento de las distribuciones de las densidades de metilación de las dos poblaciones.

Aquí se usa un ejemplo hipotético para ilustrar este concepto. Supongamos que la densidad de metilación del tejido tumoral es de aproximadamente 0,45 y que la del ADN de plasma de los sujetos sanos es de aproximadamente 0,7. Estos valores supuestos son similares a los obtenidos del paciente con CHC de los inventores donde la densidad de metilación general 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 de ADN de plasma para el genoma completo es del 1 %, el valor de punto de corte sería de $0,7 \times (100 \% - 3 \times 1 \%) = 0,679$. Para conseguir una sensibilidad del 97,5 %, es necesario que la densidad de metilación media del ADN de plasma para los enfermos de cáncer sea de aproximadamente $0,679 - 0,7 \times (2 \times 1 \%) = 0,665$. Si f representa la concentración fraccional de ADN derivado de tumor en la muestra de plasma, entonces f puede calcularse como $(0,7 - 0,665) \times f = 0,7 - 0,665$. Por lo tanto, f es aproximadamente 14 %. A partir de este cálculo, se ha estimado que la concentración fraccional mínima que puede detectarse en el plasma es del 14 % para conseguir una sensibilidad de diagnóstico del 97,5 % si la densidad de metilación total del genoma completo se usa como el parámetro de diagnóstico.

A continuación, se realizó este análisis sobre los datos obtenidos del paciente con CHC. Para esta ilustración, solamente se preparó una medición de densidad de metilación basada en el valor estimado a partir de todos los autosomas para cada muestra. La densidad de metilación media fue del 71,6 % entre 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, sería necesario que el valor de punto de corte para la densidad de metilación de plasma fuera de $71,6 \% - 3 \times 0,631 \% = 69,7 \%$ para alcanzar una puntuación $z < -3$ y una especificidad de 99,85 %. Para conseguir una sensibilidad del 97,5 %, sería necesario que la densidad de metilación de plasma media de los pacientes con cáncer estuviera 2 DT por debajo del punto de corte, es decir, 68,4 %. Ya que la densidad de metilación del tejido tumoral fue del 42,9 % y usando la fórmula: $P = BKG \times (1-f) + TUM \times f$, sería necesario que f fuera al menos 11,1 %.

En otra realización, las densidades de metilación de diferentes regiones genómicas pueden analizarse por separado, por ejemplo, como se muestra en las FIGURAS 25A o 26B. En otras palabras, se realizaron múltiples mediciones del nivel de metilación para cada muestra. Como se muestra posteriormente, podría detectarse hipometilación significativa a una concentración de ADN tumoral fraccional mucho menor en plasma y por lo tanto se potenciaría el rendimiento de diagnóstico del análisis de metilación de ADN de plasma para detección de cáncer. Puede contarse el número de regiones genómicas que muestran una desviación significativa en las densidades de metilación con respecto a la población de referencia. Después puede compararse el número de regiones genómicas con un valor de punto de corte para determinar si hay una hipometilación significativa general de ADN de plasma a través de la población de regiones genómicas investigadas, por ejemplo, los grupos de 1 Mb del genoma completo. El valor de punto de corte puede establecerse por el análisis de un grupo de sujetos de referencia sin un cáncer o derivarse matemáticamente, por ejemplo, de acuerdo con la función de distribución normal.

La FIGURA 30 es una representación 3000 que muestra la distribución de densidades de metilación del ADN de plasma de sujetos sanos y enfermos de cáncer. La densidad de metilación de cada grupo de 1 Mb se compara con los valores correspondientes del grupo de referencia. Se determinó el porcentaje de grupos que mostraban hipometilación significativa (3 DT por debajo de la media del grupo de referencia). Se usó un punto de corte de 10 % que está significativamente hipometilado para determinar si está presente ADN derivado de tumor en la muestra de plasma. Pueden usarse también otros valores de punto de corte tales como 5 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 60 %, 70 %, 80 % o 90 % según la sensibilidad y especificidad deseada del ensayo.

Por ejemplo, para clasificar que una muestra contiene ADN derivado de tumor, se puede usar 10 % de los grupos de 1 Mb que muestran hipometilación significativa (puntuación $z < -3$) como un punto de corte. Si hay más del 10 % de los grupos que están significativamente más hipometilados que el grupo de referencia, entonces la muestra se clasifica como positiva para el ensayo de cáncer. Para cada grupo de 1 Mb, se usa 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 grupos de 1 Mb, si la densidad de metilación de ADN de plasma media de los enfermos de cáncer está 1,72 DT por debajo de las densidades de metilación de ADN de plasma medias de los sujetos de referencia, entonces hay una probabilidad del 10 % de que el valor de densidad de metilación de cualquier grupo particular de un paciente con cáncer sea menor que el punto de corte (es decir puntuación $z < -3$) y proporciona un resultado positivo. Por lo tanto, si se observan todos los grupos de 1 Mb para el genoma completo, entonces se esperaría que aproximadamente el 10 % de los grupos mostraran resultados positivos de tener densidades de metilación significativamente menores (es decir puntuaciones $z < -3$). Suponiendo que la densidad de metilación general del ADN de plasma de los sujetos sanos es de aproximadamente 0,7 y el coeficiente de variación (CV) de la medición de la densidad de metilación de ADN de plasma para cada grupo de 1 Mb es del 1 %, sería necesario que la densidad de metilación media del ADN de plasma de los enfermos de cáncer fuera de $0,7 \times (100 \% - 1,72 \times 1 \%) = 0,68796$. f puede ser la concentración fraccional de ADN derivado de tumor en plasma para conseguir esta densidad de metilación de ADN de plasma media. Suponiendo que la densidad de metilación del tejido tumoral es 0,45, entonces f puede calcularse usando la ecuación

$$(\bar{M}_{P_{ref}} - M_{tumor}) \times f = \bar{M}_{P_{ref}} - \bar{M}_{P_{enfermos}}$$

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

Usando esta ecuación, $(0,7-0,45) \times f = 0,7 - 0,68796$, por lo tanto, la concentración fraccional mínima que puede detectarse usando este enfoque se deduciría como 4,8 %. La sensibilidad puede potenciarse adicionalmente reduciendo el porcentaje de punto de corte de grupos que están significativamente más hipometilados, por ejemplo, del 10 % al 5 %.

Como se ha mostrado en el ejemplo anterior, la sensibilidad de este método se determina por el grado de diferencia en el nivel de metilación entre tejidos cancerosos y no cancerosos, por ejemplo, células sanguíneas. En una realización, solamente se seleccionan las regiones cromosómicas que muestran una gran diferencia en las densidades de metilación entre el ADN de plasma de los sujetos sin cáncer y el tejido tumoral. En una realización, solamente se seleccionan regiones con una diferencia en la densidad de metilación de $>0,5$. En otras realizaciones se puede usar 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 su lugar, las regiones genómicas se definen, por ejemplo, basándose en una profundidad de lectura fija o un número fijo de sitios de CpG. Los niveles de metilación en una multitud de estas regiones genómicas se evalúan para cada muestra.

La FIGURA 31 es un gráfico 3100 que muestra la distribución de las diferencias en las densidades de metilación entre la media del ADN de plasma 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 de plasma de los sujetos sanos y un valor negativo significa que la densidad de metilación es mayor en el tejido tumoral.

En una realización, pueden seleccionarse los grupos con la mayor diferencia entre la densidad de metilación de los tejidos cancerosos y no cancerosos, por ejemplo, los que tengan una diferencia de $>0,5$, independientemente de si el tumor está hipometilado o hipermetilado para estos grupos. El límite de detección de concentración fraccional de ADN derivado de tumor en plasma puede reducirse centrándose en estos grupos debido a las mayores diferencias entre las distribuciones de los niveles de metilación de ADN de plasma entre sujetos con cáncer y sin cáncer dada la misma concentración fraccional de ADN derivado de tumor en el plasma. Por ejemplo, si solamente se usan grupos con diferencias $>0,5$ y se adopta un punto de corte de 10 % de los grupos que están significativamente más hipometilados para determinar si un individuo sometido a ensayo tiene cáncer, la concentración fraccional mínima (f) de ADN derivado de tumor detectado puede calcularse usando la siguiente ecuación: $(\bar{M}_{P_{ref}} - M_{tumor}) \times f = \bar{M}_{P_{ref}} - \bar{M}_{P_{enfermos}}$, donde $\bar{M}_{P_{ref}}$ representa la densidad de metilación media de ADN de plasma en los individuos de referencia; M_{tumor} representa la densidad de metilación del tejido tumoral en el paciente con cáncer; y $\bar{M}_{P_{enfermos}}$ representa la densidad de metilación media de ADN de plasma en los enfermos de cáncer.

Cuando 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, entonces, se tiene $0,5 \times f = 0,7 - 0,68796$ y $f = 2,4$ %. Por lo tanto, centrándose en los grupos con una mayor diferencia en la densidad de metilación entre tejidos cancerosos y no cancerosos, el límite inferior de ADN derivado de tumor fraccional puede reducirse del 4,8 % al 2,4 %. La información con respecto a qué grupos mostrarían mayores grados de diferencias de metilación entre tejidos cancerosos y no cancerosos, por ejemplo, células sanguíneas, podría determinarse a partir de tejidos tumorales del mismo órgano o el mismo tipo histológico obtenido de otros individuos.

En otra realización, puede obtenerse un parámetro de la densidad de metilación del ADN de plasma de todos los grupos y teniendo en cuenta la diferencia en las densidades de metilación entre tejidos cancerosos y no cancerosos. Puede darse un mayor peso a grupos con mayor diferencia. En una realización, la diferencia en la densidad de metilación entre tejido canceroso y no canceroso de cada grupo puede usarse directamente como el peso del grupo particular en el cálculo del parámetro final.

En otra realización más, diferentes tipos de cáncer pueden tener diferentes patrones de metilación en el tejido tumoral. Puede obtenerse un perfil de peso específico de cáncer del grado de metilación del tipo específico de cáncer.

En otra realización más, la relación entre grupos de la densidad de metilación puede determinarse en sujetos con y sin cáncer. En la Figura 8, se puede observar que, en un número pequeño de grupos, los tejidos tumorales estaban más metilados que el ADN de plasma de los sujetos de referencia. Por lo tanto, pueden seleccionarse los grupos con los valores de diferencia más extremos, por ejemplo, diferencia $>0,5$ y diferencia <0 . La relación de la densidad de metilación de estos grupos puede usarse después para indicar si el individuo sometido a ensayo tiene cáncer. En otras realizaciones, la diferencia y el cociente de la densidad de metilación de diferentes grupos pueden usarse como parámetros para indicar la relación entre grupos.

Se evaluó adicionalmente la sensibilidad de detección del enfoque para detectar o evaluar el tumor usando las densidades de metilación de múltiples regiones genómicas como se ilustra por los datos obtenidos del paciente con CHC. En primer lugar, se mezclaron lecturas del plasma preoperatorio con las obtenidas de las muestras de plasma de los controles sanos para simular muestras de plasma que contenían una concentración de ADN tumoral fraccional que variaba de 20 % a 0,5 %. Después se puntuó el porcentaje de grupos de 1 Mb (de 2.734 grupos en el genoma completo) con densidades de metilación equivalentes a puntuaciones $z < -3$. Cuando la concentración de ADN tumoral fraccional en plasma fue del 20 %, 80,0 % de los grupos mostraron hipometilación significativa. Los datos correspondientes para concentración de ADN tumoral fraccional en plasma de 10 %, 5 %, 2 %, 1 % y 0,5 % fueron de 67,6 %, 49,7 %, 18,9 %, 3,8 % y 0,77 % de los grupos que mostraron hipometilación, respectivamente. Ya que el límite teórico del número de grupos que muestran puntuaciones $z < -3$ en las muestras de control es de 0,15 %, los datos de los inventores muestran que hubo aún más grupos (0,77 %) más allá del límite de punto de corte teórico incluso cuando la concentración fraccional tumoral fue de solamente 0,5 %.

La FIGURA 32A es una tabla 3200 que muestra el efecto de reducir la profundidad de secuenciación cuando la muestra de plasma contenía 5 % o 2 % de ADN tumoral. Aún pudo detectarse una alta proporción de grupos ($>0,15$ %) que mostraban hipometilación significativa cuando la profundidad de secuenciación media fue de solamente 0,022 veces el genoma haploide.

La FIGURA 32B es un gráfico 3250 que muestra las densidades de metilación de los elementos repetidos y 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 de plasma postoperatorio del paciente con CHC. Se puede observar que los elementos repetidos estaban más metilados (mayor densidad de metilación) que las regiones no repetidas en tejidos tanto cancerosos como no cancerosos. Sin embargo, la diferencia en la metilación entre elementos repetidos y regiones no repetidas fue mayor en los tejidos no cancerosos y el ADN de plasma de los sujetos sanos cuando se compararon con los tejidos tumorales.

Como resultado, el ADN de plasma del paciente con cáncer tuvo una mayor reducción en la densidad de metilación en los elementos repetidos que en las regiones no repetidas. La diferencia en la densidad de metilación de ADN de plasma 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 en las muestras de plasma preoperatorias y postoperatorias también mostraron que el intervalo dinámico en el cambio en la densidad de metilación fue mayor en las regiones repetidas que en las no repetidas. En una realización, la densidad de metilación de ADN de plasma de los elementos repetidos puede usarse para determinar si un paciente está aquejado de cáncer o para supervisión de la progresión de enfermedad.

Como se ha analizado anteriormente, la variación en las densidades de metilación en el plasma de los sujetos de referencia también afectaría a la precisión de la diferenciación de enfermos de cáncer de individuos sin cáncer. Cuanto más ajustada sea la distribución de densidades de metilación (es decir menor desviación típica), con mayor precisión se podrán diferenciar los sujetos con cáncer de los sujetos sin cáncer. En otra realización, el coeficiente de variación (CV) de las densidades de metilación de los grupos de 1 Mb puede usarse como un criterio para seleccionar los grupos con baja variabilidad de densidades de metilación de ADN de plasma en el grupo de referencia. Por ejemplo, solamente se seleccionan grupos con $CV < 1$ %. Otros valores, por ejemplo 0,5 %, 0,75 %, 1,25 % y 1,5 % también pueden usarse como criterios para seleccionar los grupos con baja variabilidad en la densidad de metilación. En otra realización más, los criterios de selección pueden incluir tanto el CV del grupo como la diferencia en la densidad de metilación entre tejidos cancerosos y no cancerosos.

La densidad de metilación también puede usarse para estimar la concentración fraccional de ADN derivado de tumor en una muestra de plasma cuando se conoce la densidad de metilación del tejido tumoral. Esta información puede obtenerse por el análisis del tumor del paciente o de la investigación de los tumores de varios pacientes que tienen el mismo tipo de cáncer. Como se ha analizado anteriormente, la densidad de metilación de plasma (P) puede expresarse usando la siguiente ecuación: $P = BKG \times (1 - f) + TUM \times f$ donde BKG es la densidad de metilación de fondo de las células sanguíneas y otros órganos, TUM es la densidad de metilación en el tejido tumoral, y f es la concentración fraccional de ADN derivado de tumor en la muestra de plasma. Esto puede reescribirse como:

$$f = \frac{BKG - P}{BKG - TUM}.$$

Los valores de BKG pueden determinarse analizando la muestra de plasma del paciente en un punto temporal en el que no está presente el cáncer o de la investigación de un grupo de referencia de individuos sin cáncer. Por lo tanto, después de medir la densidad de metilación del plasma, puede determinarse f.

F. Combinación con otros métodos

Pueden usarse enfoques de análisis de metilación descritos en el presente documento en combinación con otros

métodos que se basan en los cambios genéticos de ADN derivado de tumor en plasma. Los ejemplos de dichos métodos incluyen el análisis con respecto a aberraciones cromosómicas asociadas con cáncer (KCA Chan *et al.*, 2013 Clin Chem; 59: 211-224; RJ Leary *et al.*, 2012 Sci Transl Med; 4: 162ra154) y variaciones de un único nucleótido asociadas con cáncer en plasma (KCA Chan *et al.*, 2013 Clin Chem; 59: 211-224). Existen ventajas para el enfoque de análisis de metilación frente a los enfoques genéticos.

Como se muestra en la FIGURA 21A, la hipometilación del ADN tumoral es un fenómeno global que implica regiones distribuidas por casi todo el genoma completo. Por lo tanto, los fragmentos de ADN de todas las regiones cromosómicas serían informativos con respecto a la contribución potencial del ADN hipometilado derivado de tumor al ADN del plasma/suero en el paciente. Por el contrario, las aberraciones cromosómicas (bien amplificación o bien delección de una región cromosómica) solamente están presentes en algunas regiones cromosómicas y los fragmentos de ADN de las regiones sin una aberración cromosómica en el tejido tumoral no serían informativos en el análisis (KCA Chan *et al.*, 2013 Clin Chem; 59: 211-224). De forma similar solamente se observan algunos miles de alteraciones de nucleótidos individuales en cada genoma de cáncer (KCA Chan *et al.*, 2013 Clin Chem; 59: 211-224). Los fragmentos de ADN que no solapan con estos cambios de nucleótidos individuales no serían informativos en la determinación de si está presente ADN derivado de tumor en el plasma. Por lo tanto, este enfoque de análisis de metilación es potencialmente más rentable que los enfoques genéticos para detectar cambios asociados con cáncer en la circulación.

En una realización, la rentabilidad del análisis de metilación de ADN de plasma puede potenciarse adicionalmente enriqueciendo con respecto a fragmentos de ADN de las regiones más informativas, por ejemplo, regiones con mayor diferencia de metilación diferencial entre tejidos cancerosos y no cancerosos. Los ejemplos de los métodos para enriquecer estas regiones incluyen el uso de sondas de hibridación (por ejemplo, sistema Nimblegen SeqCap y sistema de enriquecimiento de diana Agilent SureSelect), amplificación por PCR e hibridación en fase sólida.

G. Análisis específico de tejidos/donantes

Las células derivadas de tumor invaden y metastatizan en órganos adyacentes o distantes. Los tejidos invadidos o focos metastásicos aportan ADN al plasma como resultado de la muerte celular. Analizando el perfil de metilación de ADN en el plasma de enfermos de cáncer y detectando la presencia de identificaciones de metilación específicas de tejido, se podrían detectar los tipos de tejidos que están implicados en el proceso de 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 los sitios primarios y metastásicos. La supervisión de las concentraciones relativas de las identificaciones de metilación de los órganos implicados en plasma también permitiría evaluar la carga tumoral de esos órganos y determinar si el proceso de cáncer en ese órgano está deteriorándose o mejorando o se ha curado. Por ejemplo, si un gen X está específicamente metilado en el hígado, entonces se esperaría que la implicación metastásica del hígado por un cáncer (por ejemplo, cáncer colorrectal) aumentara la concentración de secuencias metiladas del gen X en el plasma. Habría también otra secuencia o grupos de secuencias con características de metilación similares al gen X. Se podría después combinar los resultados de dichas secuencias. Son aplicables consideraciones similares a otros tejidos, por ejemplo, el cerebro, huesos, pulmones y riñones, etc.

Por otro lado, se sabe que el ADN de diferentes órganos muestra diferentes identificaciones de metilación específicas de tejido (BW Futscher *et al.*, 2002 Nat Genet; 31: 175-179; SSC Chim *et al.*, 2008 Clin Chem; 54: 500-511). Por lo tanto, puede usarse el perfil de metilación en plasma para dilucidar la contribución de tejidos de diversos órganos al plasma. La dilucidación de dicha contribución puede usarse para evaluar el daño orgánico, ya que se cree que se libera ADN de plasma cuando mueren las células. Por ejemplo, la patología hepática tal como hepatitis (por ejemplo, por virus, procesos autoinmunitarios, etc.) o hepatotoxicidad (por ejemplo, sobredosis de fármacos (tal como paracetamol) o toxinas (tal como alcohol) provocada por fármacos está asociada con el daño a células hepáticas y se esperaría que estuviera asociada con un aumento del nivel de ADN derivado del hígado en el plasma. Por ejemplo, si un gen X está específicamente metilado en el hígado, entonces se esperaría que la patología hepática aumentara la concentración de secuencias metiladas del gen X en el plasma. Por el contrario, si un gen Y está específicamente hipometilado en el hígado, entonces se esperaría que la patología hepática redujera la concentración de secuencias metiladas del gen Y en el plasma. En otra realización más, el gen X o Y puede reemplazarse por cualquier secuencia genómica que pueda no ser un gen y que muestre metilación diferencial en diferentes tejidos dentro del cuerpo.

Las técnicas descritas en el presente documento también podrían aplicarse a la evaluación de ADN derivado de donante en el plasma de receptores de trasplante de órganos (YMD Lo *et al.*, 1998 Lancet; 351: 1329-1330). Se han usado diferencias polimórficas entre el donante y el receptor para distinguir el ADN derivado de donante del ADN derivado de receptor en el plasma (YW Zheng *et al.*, 2012 Clin Chem; 58: 549-558). Los inventores proponen que las identificaciones de metilación específicas de tejido del órgano trasplantado también podrían usarse como un método para detectar el ADN del donante en el plasma del receptor.

Supervisando la concentración del ADN del donante, se podría evaluar de forma no invasiva el estado del órgano trasplantado. Por ejemplo, el rechazo de trasplante se asocia con una mayor tasa de muerte celular y por lo tanto la concentración del ADN del donante en el plasma del receptor (o suero), como se refleja por la identificación de metilación del órgano trasplantado, aumentaría en comparación con el momento en que el paciente está en una

condición estable o cuando se compara con otros receptores de trasplante estables o controles sanos sin trasplante. De forma similar a lo que se ha descrito para el cáncer, el ADN derivado del donante podría identificarse en el plasma de receptores de trasplante detectando todos o algunos de los elementos característicos, incluyendo diferencias polimórficas, ADN de tamaño más corto para los órganos sólidos trasplantados (YW Zheng *et al.* al. 2012 Clin Chem; 58: 549-558) y perfil de metilación específico de tejido.

H. Normalización de la metilación basándose en el tamaño

Como se ha descrito anteriormente y en Lun *et al.* (FMM Lun *et al.* Clin. Chem. 2013; doi:10.1373/clinchem.2013.212274), la densidad de metilación (por ejemplo, de ADN de plasma) se correlaciona con el tamaño de los fragmentos de ADN. La distribución de las densidades de metilación para fragmentos de ADN de plasma más cortos fue significativamente menor que la de fragmentos más largos. Los inventores proponen que algunas afecciones no cancerosas (por ejemplo, lupus eritematoso sistémico (LES)) con patrones de fragmentación anómalos de ADN de plasma pueden mostrar una hipometilación aparente de ADN de plasma debido a la presencia de fragmentos de ADN de plasma cortos más abundantes, que están menos metilados. En otras palabras, la distribución de tamaños de ADN de plasma puede ser un factor de confusión para la densidad de metilación para el ADN de plasma.

La FIGURA 34A muestra una distribución de tamaños de ADN de plasma en el paciente con LES SLE04. Las distribuciones de tamaños de nueve sujetos de control sanos se muestran como líneas grises discontinuas y las de SLE04 se muestran como una línea negra continua. Los fragmentos de ADN de plasma cortos fueron más abundantes en SLE04 que en los nueve sujetos de control sanos. Como los fragmentos de ADN más cortos están en general menos metilados, este patrón de distribución de tamaños puede confundir el análisis de metilación de ADN de plasma y conducir a una hipometilación más evidente.

En algunas realizaciones, un nivel de metilación medido puede normalizarse para reducir el efecto de confusión de la distribución de tamaños en análisis de metilación de ADN de plasma. Por ejemplo, puede medirse un tamaño de moléculas de ADN en la pluralidad de sitios. En diversas implementaciones, la medición puede proporcionar un tamaño específico (por ejemplo, longitud) a una molécula de ADN o determinar sencillamente que el tamaño queda dentro de un intervalo específico, lo que también puede corresponder a un tamaño. El nivel de metilación normalizado puede después compararse con un valor de punto de corte. Existen varias formas de realizar la normalización para reducir el efecto de confusión de la distribución de tamaños en análisis de metilación de ADN de plasma.

En una realización, puede realizarse fraccionamiento por tamaños del ADN (por ejemplo, ADN de plasma). El fraccionamiento por tamaños puede asegurar que se usen fragmentos de ADN de un tamaño similar para determinar el nivel de metilación de una manera coherente con el valor de punto de corte. Como parte del fraccionamiento por tamaños, pueden seleccionarse fragmentos de ADN que tengan un primer tamaño (por ejemplo, un primer intervalo de longitudes), donde el primer valor de punto de corte corresponde al primer tamaño. La normalización puede conseguirse calculando el nivel de metilación usando solamente los fragmentos de ADN seleccionados.

Puede conseguirse el fraccionamiento por tamaños de diversas maneras, por ejemplo, por separación física de moléculas de ADN de diferentes tamaños (por ejemplo, mediante electroforesis o tecnologías basadas en microfluídica, o tecnologías basadas en centrifugación) o mediante análisis por ordenador. Para análisis por ordenador, en una realización, se puede realizar secuenciación paralela masiva de extremos emparejados de las moléculas de ADN de plasma. Se puede deducir después el tamaño de las moléculas secuenciadas por comparación con la localización de cada uno de dos extremos de una molécula de ADN de plasma con un genoma humano de referencia. Después, se puede realizar análisis posterior por la selección de moléculas de ADN secuenciadas que coinciden con uno o más criterios de selección por tamaño (por ejemplo, estando los criterios del tamaño dentro de un intervalo específico). Por lo tanto, en una realización, la densidad de metilación puede analizarse con respecto a fragmentos con un tamaño similar (por ejemplo, dentro de un intervalo específico). El valor de punto de corte (por ejemplo, en el bloque 2840 del método 2800) puede determinarse basándose en fragmentos dentro del mismo intervalo de tamaños. Por ejemplo, los niveles de metilación pueden determinarse a partir de muestras que se sabe que tienen cáncer o no tienen cáncer, y los valores de punto de corte pueden determinarse a partir de estos niveles de metilación.

En otra realización, puede determinarse una relación funcional entre la densidad de metilación y el tamaño del ADN en circulación. La relación funcional puede definirse por puntos de datos o coeficientes de una función. La relación funcional puede proporcionar valores de cambio de escala correspondientes a tamaños respectivos (por ejemplo, tamaños más cortos pueden tener aumentos correspondientes a la metilación). En diversas implementaciones, el valor de cambio de escala puede estar entre 0 y 1 o ser mayor de 1.

La normalización puede realizarse basándose en un tamaño promedio. Por ejemplo, puede calcularse un tamaño promedio correspondiente a moléculas de ADN usadas para calcular el primer nivel de metilación, y el primer nivel de metilación puede multiplicarse por el valor de cambio de escala correspondiente (es decir, correspondiente al tamaño promedio). Como otro ejemplo, la densidad de metilación de cada molécula de ADN puede normalizarse de acuerdo con el tamaño de la molécula de ADN y la relación entre el tamaño de ADN y la metilación.

En otra implementación, la normalización puede realizarse por molécula. Por ejemplo, puede obtenerse un tamaño respectivo de una molécula de ADN en un sitio particular (por ejemplo, como se ha descrito anteriormente), y puede identificarse un valor de cambio de escala correspondiente al tamaño respectivo a partir de la relación funcional. Para un cálculo no normalizado, cada molécula se contaría igualmente en la determinación de un índice de metilación en el sitio. Para el cálculo normalizado, la contribución de una molécula al índice de metilación puede ponderarse por el factor de cambio de escala que corresponde al tamaño de la molécula.

Las FIGURAS 34B y 34C muestran análisis de metilación para ADN de plasma de un paciente con LES SLE04 (FIGURA 34B) y un paciente con CHC TBR36 (FIGURA 34C). Los círculos externos muestran los resultados de Z_{met} para ADN de plasma sin fraccionamiento por tamaños por ordenador. Los círculos internos muestran los resultados de Z_{met} para ADN de plasma de 130 pb o mayor. Para el paciente con LES SLE04, el 84 % de los grupos mostraron hipometilación sin fraccionamiento por tamaños por ordenador. El porcentaje de los grupos que muestran hipometilación se redujo al 15 % cuando solamente se analizaron fragmentos de 130 pb o mayores. Para el paciente con CHC TBR36, el 98,5 % y el 98,6 % de los grupos mostraron hipometilación para el ADN de plasma con y sin fraccionamiento por tamaños por ordenador, respectivamente. Estos resultados sugieren que el fraccionamiento por tamaños por ordenador puede reducir eficazmente los resultados de hipometilación falsos positivos relacionados con la fragmentación aumentada de ADN de plasma, por ejemplo, en pacientes con LES o en otras afecciones inflamatorias.

En una realización, los resultados para los análisis con y sin fraccionamiento por tamaños pueden compararse para indicar si hay algún efecto de confusión del tamaño en los resultados de metilación. Por lo tanto, además de o en lugar de la normalización, puede usarse el cálculo de un nivel de metilación en un tamaño particular para determinar si hay probabilidad de un falso positivo cuando el porcentaje de grupos por encima de un valor de punto de corte difiera con y sin fraccionamiento por tamaños, o si solamente difiere un nivel de metilación particular. Por ejemplo, la presencia de una diferencia significativa entre los resultados para muestras con y sin fraccionamiento por tamaños puede usarse para indicar la posibilidad de un resultado falso positivo debido a un patrón de fraccionamiento anómalo. El umbral para determinar si la diferencia es significativa puede establecerse mediante el análisis de una cohorte de enfermos de cáncer y una cohorte de sujetos de control sin cáncer.

I. Análisis para hipermetilación de islas de CpG en todo el genoma en plasma

Además de la hipometilación general, también se observa habitualmente hipermetilación de islas de CpG en cánceres (SB Baylin *et al.* 2011 Nat Rev Cancer; 11: 726-734; PA 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, se describe el uso de análisis de todo el genoma con respecto a hipermetilación de islas de CpG para la detección y la supervisión de cánceres.

La FIGURA 35 es un diagrama de flujo de un método 3500 que determina una clasificación de un nivel de cáncer basándose en la hipermetilación de islas de CpG de acuerdo con realizaciones de la presente divulgación. La pluralidad de sitios del método 2800 puede incluir sitios de CpG, en los que los sitios de CpG se organizan en una pluralidad de islas de CpG, incluyendo cada isla de CpG uno o más sitios de CpG. Los niveles de metilación para cada isla de CpG pueden usarse para determinar la clasificación del nivel de cáncer.

En el bloque 3510, se identifican islas de CpG para analizar. En este análisis, como ejemplo, se determinó en primer lugar un conjunto de islas de CpG para analizar, que se caracterizan con densidades de metilación relativamente bajas en el plasma de los sujetos de referencia sanos. En un aspecto, la variación de las densidades de metilación en el grupo de referencia puede ser relativamente pequeña para permitir la detección de hipermetilación asociada con cáncer más fácilmente. En una realización, las islas de CpG tienen una densidad de metilación media menor que un primer porcentaje en un grupo de referencia, y un coeficiente de variación para la densidad de metilación en el grupo de referencia es menor que un segundo porcentaje.

Como ejemplo, para fines de ilustración, se usan los siguientes criterios para identificación de las islas de CpG útiles:

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

Estos parámetros pueden ajustarse para una aplicación específica. A partir del conjunto de datos de los inventores, 454 islas de CpG en el genoma cumplieron estos criterios.

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

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

hipermetilada.

En una realización, un primer valor de punto de corte puede corresponder a una media de densidades de metilación para el grupo de referencia más un porcentaje específico. Otro valor de punto de corte puede corresponder a la media de densidades de metilación para el grupo de referencia más un número específico de desviaciones típicas. En una implementación, se calculó una puntuación z (Z_{met}) y se comparó con valores de punto de corte. Como ejemplo, una isla de CpG en un sujeto de ensayo (por ejemplo, un sujeto que se explora con respecto a cáncer) se consideró significativamente hipermetilada si cumplía los siguientes criterios:

- i. Su densidad de metilación fue mayor que la media del grupo de referencia en 2 %, y
- ii. $Z_{met} > 3$

Estos parámetros también pueden ajustarse para una aplicación específica.

- En el bloque 3540, las densidades de metilación (por ejemplo, como puntuaciones z) de las islas de CpG hipermetiladas se usan para determinar una puntuación acumulada. Por ejemplo, después de la identificación de todas las islas de CpG hipermetiladas de forma significativa, puede calcularse una puntuación que implica una suma de puntuaciones z o funciones de puntuaciones z de todas las islas de CpG hipermetiladas. Un ejemplo de una puntuación es una puntuación de probabilidad acumulada (PA), como se describe en otra sección. La puntuación de probabilidad acumulada usa Z_{met} para determinar la probabilidad de tener dicha observación al azar según una distribución de probabilidades (por ejemplo, distribución de probabilidades de t de Student con 3 grados de libertad).

- 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 de las islas de CpG identificadas es suficientemente grande, entonces puede identificarse que el organismo tiene cáncer. En una realización, el umbral acumulado corresponde a una mayor puntuación acumulada del grupo de referencia.

IX. METILACIÓN Y ANC

- Como se ha mencionado anteriormente, los enfoques de análisis de metilación descritos en el presente documento pueden usarse en combinación con otros métodos que se basan en los cambios genéticos de ADN derivado de tumor en plasma. Los ejemplos de dichos métodos incluyen el análisis con respecto a aberraciones cromosómicas asociadas a cáncer (KCA Chan *et al.* 2013 Clin Chem; 59: 211-224; RJ Leary *et al.* 2012 Sci Transl Med; 4: 162ra154). Se describen aspectos de aberraciones del número de copias (ANC) en la Solicitud de Patente de Estados Unidos n.º 13/308.473.

A. ANC

- Pueden detectarse aberraciones del número de copias contando fragmentos de ADN que se alinean con una parte particular del genoma, normalizando el recuento y comparando el recuento con un valor de punto de corte. En diversas realizaciones, la normalización puede realizarse por un recuento de fragmentos de ADN alineados con otro haplotipo de la misma parte del genoma (dosificación de haplotipo relativa (DOHR)) o por un recuento de fragmentos de ADN alineados con otra parte del genoma.

- El método de DOHR se basa en el uso de locus heterocigotos. También pueden usarse realizaciones descritas en esta sección para locus que son homocigotos comparando dos regiones y no dos haplotipos de la misma región, y por lo tanto no son específicos de haplotipo. En un método de dosificación de región cromosómica relativa, el número de fragmentos de una región cromosómica (por ejemplo, como se determina contando las lecturas de secuencias alineadas con esa región) se compara con un valor esperado (que puede ser de una región de cromosoma de referencia o de la misma región en otra muestra que se sabe que está sana). De esta manera, se contaría un fragmento para una región cromosómica independientemente del haplotipo de procedencia del marcador secuenciado. Por lo tanto, aún podrían usarse lecturas de secuencias que no contengan ningún locus heterocigoto. Para realizar la comparación, una realización puede normalizar el recuento de marcadores antes de la comparación. Cada región se define por al menos dos locus (que están separados entre sí), y fragmentos en estos locus pueden usarse para obtener un valor colectivo alrededor de la región.

- Un valor normalizado para las lecturas secuenciadas (marcadores) para una región particular puede calcularse dividiendo el número de lecturas secuenciadas que se alinean con esa región por el número total de lecturas secuenciadas alineables con el genoma completo. El recuento de marcadores normalizado permite comparar resultados de una muestra con los resultados de otra muestra. Por ejemplo, el valor normalizado puede ser la proporción (por ejemplo, porcentaje o fracción) de lecturas secuenciadas que se espera que sean de la región particular, 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 por el número de recuentos de una región de referencia (en el caso anterior, la región de referencia es simplemente el genoma completo). Este recuento de marcadores normalizado puede después compararse frente a un valor umbral, que puede determinarse a partir de una o más muestras de referencia que no muestran cáncer.

El recuento de marcadores normalizado del caso ensayado se compararía después con el recuento de marcadores normalizado de uno o más sujetos de referencia, por ejemplo, los que no tienen cáncer. En una realización, la comparación se realiza calculando la puntuación z del caso para la región cromosómica particular. La puntuación z puede calcularse usando la siguiente ecuación: puntuación $z = (\text{recuento de marcadores normalizado del caso} - \text{media})/\text{DT}$, donde "media" es el recuento de marcadores normalizado medio que se alinea con la región cromosómica particular para las muestras de referencia; y DT es la desviación típica del número de recuentos de marcadores normalizados que se alinea con la región particular para las muestras de referencia. Por lo tanto, la puntuación z es el número de desviaciones típicas por las que el recuento de marcadores normalizado de una región cromosómica para el caso ensayado se aleja del recuento de marcadores normalizado medio para la misma región cromosómica de los sujetos de referencia.

En la situación en la que el organismo sometido a ensayo tiene cáncer, las regiones cromosómicas que se amplifican en los tejidos tumorales estarían sobrerrepresentadas en el ADN del plasma. Esto daría como resultado un valor positivo de la puntuación z . Por otro lado, las regiones cromosómicas que están suprimidas en los tejidos tumorales estarían infrarrepresentadas en el ADN de plasma. Esto daría como resultado un valor negativo de la puntuación z . La magnitud de la puntuación z se determina por varios factores.

Un factor es la concentración fraccional de ADN derivado de tumor en la muestra biológica (por ejemplo, plasma). Cuanto mayor sea la concentración fraccional de ADN derivado de tumor en la muestra (por ejemplo, plasma), mayor será la diferencia entre el recuento de marcadores normalizado del caso ensayado y los casos de referencia. Por lo tanto, se produciría una mayor magnitud de la puntuación z .

Otro factor es la variación del recuento de marcadores normalizado en el o los casos de referencia. Con el mismo grado de la sobrerrepresentación de la región cromosómica en la muestra biológica (por ejemplo, plasma) del caso ensayado, una variación más pequeña (es decir una desviación típica más pequeña) del recuento de marcadores normalizado en el grupo de referencia daría como resultado una puntuación z mayor. De forma similar, con el mismo grado de infrarrepresentación de la región cromosómica en la muestra biológica (por ejemplo, plasma) del caso ensayado, una desviación típica más pequeña del recuento de marcadores normalizado en el grupo de referencia daría como resultado una puntuación z más negativa.

Otro factor es la magnitud de aberración cromosómica en los tejidos tumorales. La magnitud de aberración cromosómica se refiere a los cambios del número de copias para la región cromosómica particular (bien ganancia o bien pérdida). Cuanto mayor sean los cambios de números de copias en los tejidos tumorales, mayor será el grado de sobre o infrarrepresentación de la región cromosómica particular en el ADN de plasma. Por ejemplo, la pérdida de ambas copias del cromosoma daría como resultado mayor infrarrepresentación de la región cromosómica en el ADN de plasma que la pérdida de una de las dos copias del cromosoma y, por lo tanto, daría como resultado una puntuación z más negativa. Típicamente, hay múltiples aberraciones cromosómicas en los cánceres. Las aberraciones cromosómicas en cada cáncer pueden variar adicionalmente por su naturaleza (es decir amplificación o delección), su grado (ganancia o pérdida de una única copia o múltiples copias) y su alcance (tamaño de la aberración con respecto a longitud cromosómica).

La precisión de la medición del recuento de marcadores normalizado se ve afectada por el número de moléculas analizadas. Se espera que sea necesario analizar 15.000, 60.000 y 240.000 moléculas para detectar aberraciones cromosómicas con un cambio de una copia (bien ganancia o bien pérdida) cuando la concentración fraccional sea de aproximadamente 12,5 %, 6,3 % y 3,2 % respectivamente. Se describen detalles adicionales del recuento de marcadores para detección de cáncer para diferentes regiones cromosómicas en la Publicación de Patente de Estados Unidos n.º 2009/0029377 titulada "Diagnosing Fetal Chromosomal Aneuploidy Using Massively Parallel Genomic Sequencing" de Lo *et al.*

Las realizaciones también pueden usar análisis de tamaños, en lugar del método de recuento de marcadores. También puede usarse análisis de tamaños, en lugar de un recuento de marcadores normalizado. El análisis de tamaños puede usar diversos parámetros, como se ha mencionado en el presente documento, y en la Solicitud de Patente de Estados Unidos n.º 12/940.992. Por ejemplo, pueden usarse los valores Q o F anteriores. Dichos valores de tamaños no necesitan una normalización por recuentos de otras regiones ya que esos valores no cambian de escala con el número de lecturas. Pueden usarse técnicas de los métodos específicos de haplotipo, tales como el método de DOHR descrito anteriormente y en más detalle en la Solicitud de Patente de Estados Unidos n.º 13/308.473, para los métodos no específicos también. Por ejemplo, pueden usarse técnicas que impliquen la profundidad y el refinamiento de una región. En algunas realizaciones, puede tenerse en cuenta un desplazamiento a GC para una región particular cuando se comparen dos regiones. Ya que el método de DOHR usa la misma región, no es necesaria dicha corrección.

Aunque ciertos cánceres pueden presentarse típicamente aberraciones en regiones cromosómicas particulares, dichos cánceres no presentan siempre exclusivamente aberraciones en dichas regiones. Por ejemplo, regiones cromosómicas adicionales podrían mostrar aberraciones, y la localización de dichas regiones adicionales podría ser desconocida. Además, cuando se exploran pacientes para identificar estadios tempranos de cáncer, se puede querer identificar una amplia serie de cánceres, que podrían mostrar aberraciones presentes por todo el genoma. Para

abordar estas situaciones, las realizaciones pueden analizar una pluralidad de regiones de una manera sistemática para determinar qué regiones muestran aberraciones. Puede usarse el número de aberraciones y su localización (por ejemplo, si son contiguas), por ejemplo, para confirmar aberraciones, determinar un estadio del cáncer, proporcionar un diagnóstico de cáncer (por ejemplo, si el número es mayor que un valor umbral), y proporcionar un pronóstico basado en el número y la localización de diversas regiones que muestran una aberración.

En consecuencia, las realizaciones pueden identificar si un organismo tiene cáncer basándose en el número de regiones que muestran una aberración. Por lo tanto, se puede ensayar una pluralidad de regiones (por ejemplo, 3.000) para identificar un número de regiones que muestren una aberración. Las regiones pueden abarcar el genoma completo o solamente partes del genoma, por ejemplo, región no repetida.

La FIGURA 36 es un diagrama de flujo de un método 3600 para analizar una muestra biológica de un organismo usando una pluralidad de regiones cromosómicas de acuerdo con realizaciones de la presente divulgación. La muestra biológica incluye moléculas de ácido nucleico (también denominadas fragmentos).

En el bloque 3610, se identifican una pluralidad de regiones (por ejemplo, regiones no solapantes) del genoma del organismo. Cada región cromosómica incluye una pluralidad de locus. Una región puede ser de 1 Mb de tamaño, o algún otro tamaño igual. Para la situación de una región que sea de 1 Mb de tamaño, el genoma completo puede entonces incluir aproximadamente 3.000 regiones, cada una de un tamaño y una localización predeterminados. Dichas regiones predeterminadas pueden variar para acomodar una longitud de un cromosoma particular o un número específico de regiones para usar, y cualquier otro criterio mencionado en el presente documento. Si algunas regiones tienen diferentes longitudes, dichas longitudes pueden usarse para normalizar los resultados, por ejemplo, como se describe en el presente documento. Las regiones pueden seleccionarse específicamente basándose en ciertos criterios del organismo específico y/o basándose en el conocimiento del cáncer que se ensaya. Las regiones también pueden seleccionarse de forma arbitraria.

En el bloque 3620, una localización de la molécula de ácido nucleico en un genoma de referencia del organismo se identifica para cada una de una pluralidad de moléculas de ácido nucleico. La localización puede determinarse de cualquiera de las maneras mencionadas en el presente documento, por ejemplo, secuenciando los fragmentos para obtener marcadores secuenciados y alinear los marcadores secuenciados con el genoma de referencia. También puede determinarse un 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 localizaciones identificadas. El grupo respectivo puede incluir al menos una molécula de ácido nucleico localizada en cada uno de la pluralidad de locus de la región cromosómica. En una realización, el grupo puede ser fragmentos que se alinean con un haplotipo particular de una región cromosómica, por ejemplo, como en el método de DOHR 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 mencionados en el presente documento. Por ejemplo, el valor puede ser el número de fragmentos en el grupo o un valor estadístico de una distribución por tamaños de los fragmentos en el grupo. El valor respectivo también puede ser un valor normalizado, por ejemplo, un recuento de marcadores de la región dividido por el número total de recuentos de marcadores para la muestra o el número de recuentos de marcadores para una región de referencia. El valor respectivo también puede ser una diferencia o relación de otro valor (por ejemplo, en DOHR), proporcionando de ese modo la propiedad de una diferencia para 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 muestra una delección o una amplificación. Este valor de referencia puede ser cualquier valor umbral o de referencia descrito en el presente documento. Por ejemplo, el valor de referencia podría ser un valor umbral determinado para muestras normales. Para DOHR, el valor respectivo podría ser la diferencia o relación de recuentos de marcadores 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 recuento de marcadores o el valor de los tamaños para otro haplotipo o región, y la comparación puede incluir tomar una diferencia o una relación (o una función de estas) y después determinar si la diferencia o la relación es mayor que un valor umbral.

El valor de referencia puede variar basándose en los resultados de otras regiones. Por ejemplo, si las regiones adyacentes también muestran una desviación (aunque pequeña en comparación con un umbral, por ejemplo, una puntuación z de 3), entonces puede usarse un umbral inferior. Por ejemplo, si tres regiones consecutivas están todas por encima de un primer umbral, entonces el cáncer puede ser más probable. Por lo tanto, este primer umbral puede ser menor que otro umbral que es necesario para identificar el cáncer de regiones no consecutivas. La existencia de tres regiones (o más de tres) que tengan una desviación incluso pequeña puede tener una probabilidad suficientemente baja de un efecto al azar para que se pueda conservar la sensibilidad y especificidad.

En el bloque 3660, se determina una cantidad de regiones genómicas clasificadas como presentadoras de una delección o amplificación. Las regiones cromosómicas que se cuentan pueden tener restricciones. Por ejemplo, pueden contarse solamente regiones que sean contiguas con al menos otra región (o pueden requerirse regiones contiguas que sean de un cierto tamaño, por ejemplo, 4 o más regiones). Para realizaciones en las que las regiones no son iguales, el número también puede representar las longitudes respectivas (por ejemplo, el número podría ser una longitud total de las regiones aberrantes).

En el bloque 3670, la cantidad se compara con una cantidad de valor umbral 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 cáncer. En una realización, se cuentan todas las regiones aberrantes y se usa un único valor umbral independientemente de donde aparezcan las regiones. En otra realización, un valor umbral puede variar basándose en las localizaciones y el tamaño de las regiones que se cuentan. Por ejemplo, la cantidad de regiones en un cromosoma o una rama de un cromosoma particular puede compararse con un umbral para ese cromosoma (o rama) particular. Pueden usarse múltiples umbrales. Por ejemplo, la cantidad de regiones aberrantes en un cromosoma (o rama) particular puede ser mayor que un primer valor umbral, y la cantidad total de regiones aberrantes en el genoma debe ser mayor que un segundo valor umbral. El valor umbral puede ser un porcentaje de las regiones que se ha determinado que muestran una delección o una amplificación.

Este valor umbral para la cantidad de regiones también puede depender de lo fuerte que sea el desequilibrio para las regiones contadas. Por ejemplo, la cantidad de regiones que se usan como el umbral para determinar una clasificación de cáncer puede depender de la especificidad y sensibilidad (umbral aberrante) usadas para detectar una aberración en cada región. Por ejemplo, si el umbral aberrante es bajo (por ejemplo, puntuación z de 2), entonces la cantidad umbral puede seleccionarse para que sea alta (por ejemplo 150). Sin embargo, si el umbral aberrante es alto (por ejemplo, una puntuación z de 3), entonces la cantidad umbral puede ser menor (por ejemplo, 50). La cantidad de regiones que muestran una aberración también puede ser un valor ponderado, por ejemplo, una región que muestra un alto desequilibrio puede ponderarse más alta que una región que solamente muestra un desequilibrio pequeño (es decir, hay más clasificaciones aparte de solamente positivo o negativo para la aberración). Como ejemplo, puede usarse una suma de puntuaciones z , usando de ese modo los valores ponderados.

En consecuencia, la cantidad (que puede incluir el número y/o el tamaño) de regiones cromosómicas que muestran sobre o infrarrepresentación significativa de un recuento de marcadores normalizado (u otro valor respectivo para la propiedad del grupo) puede usarse para reflejar la gravedad de la enfermedad. La cantidad de regiones cromosómicas con un recuento de marcadores normalizado aberrante puede determinarse por dos factores, concretamente el número (o tamaño) de aberraciones cromosómicas en los tejidos tumorales y la concentración fraccional de ADN derivado de tumor en la muestra biológica (por ejemplo plasma). Cánceres más avanzados tienden a mostrar más (y mayores) aberraciones cromosómicas. Por lo tanto, más aberraciones cromosómicas asociadas con cáncer serían potencialmente detectables en la muestra (por ejemplo, plasma). En pacientes con cáncer más avanzado, la mayor carga tumoral conduciría a una mayor concentración fraccional de ADN derivado de tumor en el plasma. Como resultado, las aberraciones cromosómicas asociadas a tumor serían más fácilmente detectables 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 punto de corte para la puntuación z sigue siendo >2 y <-2 . Sin embargo, una región cromosómica se clasificaría como potencialmente aberrante solamente cuando dos segmentos consecutivos muestren el mismo tipo de aberraciones, por ejemplo, ambos segmentos tengan una puntuación z de >2 . En otras realizaciones, las puntuaciones z de segmentos adyacentes pueden sumarse entre sí usando un valor de punto de corte más alto. Por ejemplo, las puntuaciones z de tres segmentos consecutivos pueden sumarse y puede usarse un valor de punto de corte de 5. Este concepto puede extenderse a más de tres segmentos consecutivos.

La combinación de cantidad y umbrales aberrantes también puede depender del fin del análisis, y cualquier conocimiento previo del organismo (o falta del mismo). Por ejemplo, si se explora una población sana normal con respecto a cáncer, entonces se usaría típicamente alta especificidad, potencialmente tanto en la cantidad de regiones (es decir alto umbral para el número de regiones) como en un umbral aberrante para cuando se identifica una región como poseedora de una aberración. Sin embargo, en un paciente con mayor riesgo (por ejemplo, un paciente que se queje de un bulto o tenga un historial familiar, fumador, portador de papilomavirus humano crónico (VPH), portador del virus de la hepatitis o portador de otro virus), entonces los umbrales podrían ser menores para tener más sensibilidad (menos falsos negativos).

En una realización, si se usa una resolución de 1 Mb y un límite de detección menor de 6,3 % de ADN derivado de tumor para detectar una aberración cromosómica, sería necesario que el número de moléculas en cada segmento de 1 Mb fuera de 60.000. Esto se traduciría en aproximadamente 180 millones (60.000 lecturas/Mb x 3.000 Mb) de lecturas alineables para el genoma completo.

Un tamaño de segmento menor proporcionaría una mayor resolución para detectar aberraciones cromosómicas menores. Sin embargo, esto aumentaría la cantidad requerida del número de moléculas para analizar en total. Un

tamaño de segmento mayor reduciría el número de moléculas requerido para el análisis a costa de resolución. Por lo tanto, solamente pueden detectarse aberraciones mayores. En una implementación, podrían usarse regiones mayores, los segmentos que muestran una aberración podrían subdividirse y estas subregiones analizarse para obtener una mejor resolución (por ejemplo, como se ha descrito anteriormente). Si se tiene una estimación de un tamaño de la delección o amplificación para detectar (o concentración mínima para detectar), puede determinarse el número de moléculas para analizar.

B. ANC basada en la secuenciación de ADN de plasma tratado por bisulfito

Puede observarse con frecuencia hipometilación en todo el genoma y ANC en tejidos tumorales. Aquí, se demuestra que la información de ANC y cambios de metilación asociados con cáncer puede obtenerse simultáneamente de la secuenciación por bisulfito de ADN de plasma. Como los dos tipos de análisis pueden llevarse a cabo en el mismo conjunto de datos, prácticamente no hay ningún coste adicional para el análisis de ANC. Otras realizaciones pueden usar diferentes procedimientos para obtener la información de metilación y la información genética. En otras realizaciones, se puede realizar un análisis similar para hipermetilación asociada con cáncer junto con el análisis de ANC.

La FIGURA 37A muestra análisis de ANC para tejidos tumorales, ADN de plasma no tratado por bisulfito (BS) y ADN de plasma tratado por bisulfito (de dentro a fuera) para el paciente TBR36. La FIGURA 37A muestra análisis de ANC para tejidos tumorales, ADN de plasma no tratado por bisulfito (BS) y ADN de plasma tratado por bisulfito (de dentro a fuera) para el paciente TBR36. El anillo más externo muestra el ideograma del cromosoma. 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 análisis de plasma, se muestran las puntuaciones z. Una diferencia de 5 está presente entre dos líneas concéntricas. Para análisis de tejido tumoral, se muestra el número de copias. Una diferencia de una copia está presente entre dos líneas concéntricas. La FIGURA 38A muestra análisis de ANC para tejidos tumorales, ADN de plasma no tratado por bisulfito (BS) y ADN de plasma tratado por bisulfito (de dentro a fuera) para el paciente TBR34. Los patrones de ANC detectados en las muestras de plasma tratadas por bisulfito y no tratadas por bisulfito fueron concordantes.

Los patrones de ANC detectados en los tejidos tumorales, plasma no tratado por bisulfito y plasma tratado por bisulfito fueron concordantes. Para evaluar adicionalmente la concordancia entre los resultados del plasma tratado por bisulfito y no tratado por bisulfito, se construye una representación de dispersión. La FIGURA 37B es una representación de dispersión que muestra la relación entre las puntuaciones z para la detección de ANC usando plasma tratado por bisulfito y no tratado por bisulfito de los grupos 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 FIGURA 38B es una representación de dispersión que muestra la relación entre las puntuaciones z para la detección de ANC usando plasma tratado por bisulfito y no tratado por bisulfito de los grupos 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).

C. Análisis sinérgico de ANC asociado a cáncer y cambios de metilación

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

En una realización, la presencia de ANC o cambio de metilación asociado a cáncer puede usarse para indicar la presencia potencial de un cáncer. En dicha realización, la sensibilidad de detección de cáncer puede aumentarse cuando están presentes ANC o cambios de metilación en el plasma de un sujeto ensayado. En otra realización, la presencia de ambos cambios puede usarse para indicar la presencia de un cáncer. En dicha realización, la especificidad del ensayo puede mejorarse debido a que uno de los dos tipos de cambios puede detectarse potencialmente en algunos sujetos sin cáncer. Por lo tanto, la tercera clasificación puede ser positiva para cáncer solamente cuando tanto la primera clasificación como la segunda clasificación indican cáncer.

Se reclutaron 26 enfermos de CHC y 22 sujetos sanos. Se recogió una muestra de sangre de cada sujeto y el ADN de plasma se secuenció después del tratamiento por bisulfito. Para los enfermos de CHC, las muestras de sangre se recogieron en el momento del diagnóstico. La presencia de cantidades significativas de ANC se definió, por ejemplo, por tener $>5\%$ de los grupos que muestran una puntuación z de <-3 o >3 . Se definió la presencia de cantidades significativas de hipometilación asociada a cáncer como el hecho de que $>3\%$ de los grupos muestran una puntuación z <-3 . Como ejemplos, la cantidad de regiones (grupos) puede expresarse como un recuento sin procesar de grupos, un porcentaje y una longitud de los grupos.

La Tabla 3 muestra la detección de cantidades significativas de ANC y cambios de metilación en el plasma de 26 enfermos de CHC usando secuenciación paralela masiva en ADN de plasma tratado por bisulfito.

Tabla 3

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

Las tasas de detección del cambio de metilación asociado a cáncer y ANC fueron del 69 % y 50 %, respectivamente. La tasa de detección (es decir sensibilidad de diagnóstico) mejoró hasta el 73 % si la presencia de uno de los criterios se usó para indicar la potencial presencia de un cáncer.

Se muestran los resultados de dos pacientes que muestran la presencia de ANC (FIGURA 39A) o cambios de metilación (FIGURA 39B). La FIGURA 39A es una representación de Circos que muestra el ANC (anillo interno) y análisis de metilación (anillo externo) para el plasma tratado por bisulfito para un paciente con CHC TBR240. Para el análisis de ANC, los puntos verdes, rojos y grises representan regiones con ganancia cromosómica, 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 hipermetilación, hipometilación y metilación normal, respectivamente. En este paciente, se detectó ANC asociada a cáncer en el plasma mientras que el análisis de metilación no reveló cantidades significativas de hipometilación asociada a cáncer. La FIGURA 39B es una representación de Circos que muestra el análisis de ANC (anillo interno) y metilación (anillo externo) para el plasma tratado por bisulfito para un paciente con CHC TBR164. En este paciente, se detectó hipometilación asociada a cáncer en el plasma. Sin embargo, no se observaron cantidades significativas de ANC. Los resultados de dos pacientes que muestran la presencia tanto de ANC como de cambios de metilación se muestran en las FIGURAS 48A (TBR36) y 49A (TBR34).

La Tabla 4 muestra la detección de cantidades significativas de ANC y cambios de metilación en el plasma de 22 sujetos de control usando secuenciación paralela masiva en ADN de plasma tratado por bisulfito. Se usó un enfoque de bootstrap (es decir dejando uno fuera) para la evaluación de cada uno de los sujetos de control. Por lo tanto, cuando se evaluó a un sujeto particular, los otros 21 sujetos se usaron para el cálculo de la media y DT del grupo de control.

Tabla 4

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

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

En una realización, las muestras positivas para ANC y/o hipometilación se consideran positivas para cáncer, y las muestras cuando ambas son indetectables se consideran negativas. El uso de la lógica "o" proporciona mayor sensibilidad. En otra realización, solamente se consideran positivas para cáncer muestras que sean positivas tanto para ANC como para hipometilación, proporcionando de este modo mayor especificidad. En otra realización más, se pueden usar tres niveles de clasificación. Los sujetos se clasifican en i. ambos normales; ii. uno anómalo; iii. ambos anómalos.

Pueden usarse diferentes estrategias de seguimiento para estas tres clasificaciones. Por ejemplo, los sujetos para (iii) pueden someterse al protocolo de seguimiento más intensivo, por ejemplo que implica captura de imágenes de cuerpo completo; los sujetos para (ii) pueden someterse a un protocolo de seguimiento menos intensivo, por ejemplo secuenciación de ADN de plasma repetida después de un intervalo de tiempo corto relativo de varias semanas; y los sujetos para (i) pueden someterse al protocolo de seguimiento menos intensivo tal como volver a ensayar después de varios años. En otras realizaciones, la metilación y las mediciones de ANC pueden usarse junto con otros parámetros clínicos (por ejemplo, resultados de captura de imágenes o bioquímica de suero) para refinar adicionalmente la clasificación.

D. Valor pronóstico del análisis de ADN de plasma después de tratamiento con intención de curar

La presencia de ANC asociada a cáncer y/o cambios de metilación en plasma indicaría la presencia de ADN derivado de tumor en la circulación del paciente con cáncer. Se esperaría una reducción o eliminación de estos cambios asociados a cáncer después del tratamiento (por ejemplo, cirugía). Por otro lado, la persistencia de estos cambios en el plasma después del tratamiento podría indicar la eliminación incompleta de todas las células tumorales del cuerpo y puede ser un pronosticador útil para reaparición de enfermedad.

Se recogieron muestras de sangre de los dos enfermos de CHC TBR34 y TBR36 una semana después de la resección quirúrgica con intención de curar de los tumores. Se realizaron análisis de ANC y metilación en las muestras de plasma postratamiento tratadas por bisulfito.

La FIGURA 40A muestra análisis de ANC en ADN de plasma tratado por bisulfito recogido antes (anillo interno) y después (anillo externo) de la resección quirúrgica del tumor para el 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 del número de copias, pérdida del número de copias y sin cambio del número de copias, respectivamente. La mayoría de la ANC observada antes del tratamiento desapareció después de resección tumoral. La proporción de grupos que muestran una puntuación z de <-3 o >3 se redujo del 25 % al 6,6 %.

La FIGURA 40B muestra análisis de metilación en ADN de plasma tratado por bisulfito recogido antes (anillo interno) y después (anillo externo) de la resección quirúrgica del tumor para el paciente con CHC TBR36. Los puntos verdes, rojos y grises representan regiones con hipermetilación, hipometilación y metilación normal, respectivamente. Hubo una reducción notable en la proporción de grupos que mostraban hipometilación significativa del 90 % al 7,9 % y el grado de hipometilación también mostró una reducción notable. Este paciente tuvo una remisión clínica completa a los 22 meses después de la resección tumoral.

La FIGURA 41A muestra análisis de ANC en ADN de plasma tratado por bisulfito recogido antes (anillo interno) y después (anillo externo) de resección quirúrgica de tumor para el paciente con CHC TBR34. Aunque hay una reducción tanto en el número de grupos que muestran ANC como en la magnitud de ANC en los grupos afectados después de la resección quirúrgica del tumor, podría observarse ANC residual en la muestra de plasma postoperatoria. El círculo rojo destaca la región en la que las ANC residuales eran más evidentes. La proporción de grupos que muestran una puntuación z de <-3 o >3 se redujo del 57 % al 12 %.

La FIGURA 41B muestra análisis de metilación en ADN de plasma tratado por bisulfito recogido antes (anillo interno) y después (anillo externo) de resección quirúrgica de tumor para el paciente con CHC TBR34. La magnitud de la hipometilación se redujo después de la resección tumoral habiéndose reducido la puntuación z media para los grupos hipometilados de -7,9 a -4,0. Sin embargo, la proporción de grupos que tenían 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 después del tratamiento. Clínicamente, se detectaron múltiples focos de nódulos tumorales en el hígado no resecado restante a los 3 meses después de la resección tumoral. Se observaron metástasis de pulmón desde el cuarto mes después de la cirugía. El paciente murió por reaparición local y enfermedad metastásica 8 meses después de la operación.

Las observaciones en estos dos pacientes (TBR34 y TBR36) sugieren que la presencia de cambios asociados a cáncer residual de ANC e hipometilación puede usarse para supervisar y pronosticar a enfermos de cáncer después de tratamientos con intención de curar. Los datos también mostraron que el grado de cambio en la cantidad de ANC en plasma detectada puede usarse de forma sinérgica con evaluación del grado de cambio en el alcance de la metilación de ADN de plasma para pronóstico y supervisión de la eficacia del tratamiento.

En consecuencia, en algunas realizaciones, se obtiene una muestra biológica antes del tratamiento y se obtiene una segunda muestra biológica después del tratamiento (por ejemplo, cirugía). Se obtienen primeros valores para la primera muestra, tales como las puntuaciones z de regiones (por ejemplo, niveles de metilación regional y valores normalizados para ANC) y el número de regiones que mostraban hipometilación y ANC (por ejemplo, amplificación o delección). Pueden obtenerse segundos valores para la segunda muestra. En otra realización, pueden obtenerse una tercera muestra, o incluso muestras adicionales, después del tratamiento. El número de regiones que muestran hipometilación y ANC (por ejemplo, amplificación o delección) puede obtenerse de la tercera muestra o incluso muestras adicionales.

Como se ha descrito anteriormente para las FIGURAS 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 FIGURAS 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 el primer número con el segundo número puede usarse para determinar un pronóstico del tratamiento. En diversas realizaciones, solamente una de las comparaciones puede ser determinante de pronóstico o pueden usarse ambas comparaciones. En realizaciones en las que se obtienen la tercera o incluso muestras adicionales, una o más de estas muestras pueden usarse para determinar un pronóstico del tratamiento, por sí solas, o junto con la segunda muestra.

En una implementación, se predice que el pronóstico será peor cuando una primera diferencia entre la primera cantidad y la segunda cantidad esté por debajo de un primer umbral de diferencia. En otra implementación, se predice que el pronóstico será peor cuando una segunda diferencia entre el primer número y el segundo número esté por debajo de un segundo umbral de diferencia. El umbral podría ser igual o diferente. En una realización, el primer umbral de diferencia y el segundo umbral de diferencia son cero. Por lo tanto, para el ejemplo anterior, la diferencia entre los

valores para metilación indicaría un pronóstico peor 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 umbrales respectivos. La clasificación del pronóstico puede depender de cuánto más arriba o más abajo del umbral estén las diferencias. Podrían usarse múltiples umbrales para proporcionar diversas clasificaciones. Diferencias mayores pueden predecir mejores resultados y diferencias menores (e incluso valores negativos) pueden predecir resultados peores.

En algunas realizaciones, los puntos de tiempo en los que se toman las diversas muestras también se indican. Con dichos parámetros temporales, se podría determinar la cinética o la tasa de cambio de la cantidad. En una realización, una reducción rápida en la hipometilación asociada a tumor en plasma y/o una reducción rápida en la ANC asociada a tumor en plasma serán predictivas de buen pronóstico. Por el contrario, una hipometilación asociada a tumor estática o un aumento rápido en la hipometilación asociada a tumor en plasma y/o una ANC asociada a tumor estática o un aumento rápido en la ANC asociada a tumor serán predictivos de mal pronóstico. Las mediciones de metilación y ANC pueden usarse junto con otros parámetros clínicos (por ejemplo, resultados de captura de imágenes o bioquímica de suero o marcadores de proteínas) para predicción del resultado clínico.

Las realizaciones pueden usar otras muestras además de plasma. Por ejemplo, pueden medirse aberraciones de metilación asociadas a tumor (por ejemplo, hipometilación) y/o ANC asociadas a tumor a partir de células tumorales que circulan en la sangre de enfermos de cáncer, de ADN sin células o de células tumorales en la orina, heces, saliva, esputo, líquido biliar, líquido pancreático, hisopos del cuello uterino, secreciones del tracto reproductor (por ejemplo, del vaginal), líquido ascítico, líquido pleural, semen, sudor y lágrimas.

En diversas realizaciones, pueden detectarse aberraciones de metilación asociadas a tumor (por ejemplo, hipometilación) y/o ANC asociadas a tumor de la sangre o el plasma de pacientes con cáncer de mama, cáncer de pulmón, cáncer colorrectal, cáncer pancreático, cáncer ovárico, carcinoma nasofaríngeo, cáncer de cuello uterino, melanoma, tumores cerebrales, etc. De hecho, como las alteraciones de metilación y genéticas tales como ANC son fenómenos universales en el cáncer, los enfoques descritos pueden usarse para todos los tipos de cáncer. Las mediciones de metilación y ANC pueden usarse junto con otros parámetros clínicos (por ejemplo, resultados de captura de imágenes) para la predicción del resultado clínico. También pueden usarse realizaciones para la exploración y supervisión de pacientes con lesiones preneoplásicas, por ejemplo, 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 ha determinado que muestran una delección o una amplificación y pueden producir una segunda cantidad posterior de regiones que se ha determinado que tienen un nivel de metilación regional que supera el valor de punto de corte regional 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 determinar una primera diferencia entre la primera cantidad y la primera cantidad posterior, y la primera diferencia puede compararse con uno o más primeros umbrales de diferencia para determinar un pronóstico. La comparación para determinar el pronóstico del organismo también puede incluir determinar una segunda diferencia entre la segunda cantidad y la segunda cantidad posterior, y la segunda diferencia puede compararse con uno o más segundos umbrales de diferencia. Los umbrales pueden ser cero u otro número.

Puede predecirse que el pronóstico sea peor cuando la primera diferencia esté por debajo de un primer umbral de diferencia que cuando la primera diferencia esté por encima del primer umbral de diferencia. Puede predecirse que el pronóstico sea peor cuando la segunda diferencia esté por debajo de un segundo umbral de diferencia que cuando la segunda diferencia esté por encima del segundo umbral de diferencia. Los ejemplos de tratamientos incluyen inmunoterapia, cirugía, radioterapia, quimioterapia, terapia basada en anticuerpos, terapia génica, terapia epigenética o terapia dirigida.

E. Rendimiento

Se describe ahora el rendimiento de diagnóstico para diferentes números de lecturas de secuencias y de tamaño de grupos para análisis de ANC y de metilación.

1. Número de lecturas de secuencias

De acuerdo con una realización, se analizó el ADN de plasma de 32 sujetos de control sanos, 26 pacientes que padecen carcinoma hepatocelular y 20 pacientes que padecen otros tipos de cánceres, incluyendo carcinoma nasofaríngeo, cáncer de mama, cáncer de pulmón, cáncer neuroendocrino y sarcoma del músculo liso. Veintidós de los 32 sujetos sanos se seleccionaron aleatoriamente como el grupo de referencia. Se usaron la media y la desviación típica (DT) de estos 22 individuos de referencia para determinar el intervalo normal de densidad de metilación y representación genómica. Se usó ADN extraído de la muestra de plasma de cada individuo para construcción de

bibliotecas de secuenciación usando el kit de secuenciación de extremos emparejados Illumina. Las bibliotecas de secuenciación se sometieron después a tratamiento por bisulfito que convirtió restos de citosina desmetilados en uracilo. La biblioteca de secuenciación convertida por bisulfito para cada muestra de plasma se secuenció usando un carril de un secuenciador HiSeq2000 Illumina.

Después de determinar las bases, se retiraron secuencias adaptadoras y bases de baja calidad (es decir puntuación de calidad <5) de los extremos de los fragmentos. Las lecturas recortadas en formato FASTQ se procesaron después por un proceso de análisis de datos de metilación denominado Methy-Pipe (P Jiang *et al.* 2010, Conferencia Internacional sobre Bioinformática y Biomedicina de IEEE, doi:10.1109/BIB-MW.2010.5703866). Para alinear las lecturas de secuenciación convertidas por bisulfito, se realizó en primer lugar conversión por ordenador de todos los restos de citosina a timinas en las cadenas de Watson y Crick por separado usando el genoma humano de referencia (NCBI versión 36/hg19). Después, se realizó conversión por ordenador de cada citosina a timina en todas las lecturas procesadas y se mantuvo la información posicional de cada resto convertido. Se usó 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 desapareamientos permitidos para cada lectura alineada. Solamente se usaron lecturas mapeables en una única localización genómica para análisis cadena abajo. Se retiraron lecturas ambiguas mapeadas en las cadenas tanto de Watson como de Crick y lecturas duplicadas (clonales). Se usaron restos de citosina en el contexto del dinucleótido de CpG para análisis de metilación cadena abajo. Después del alineamiento, las citosinas originalmente presentes en las lecturas secuenciadas se recuperaron basándose en la información posicional mantenida durante la conversión por ordenador. Las citosinas recuperadas entre los dinucleótidos de CpG se puntuaron como metiladas. Las timinas entre los dinucleótidos de CpG se puntuaron como desmetiladas.

Para análisis de metilación, el genoma se dividió en grupos del mismo tamaño. El tamaño de los grupos ensayados incluye 50 kb, 100 kb, 200 kb y 1 Mb. La densidad de metilación para cada grupo se calculó como el número de citosinas metiladas en el contexto del dinucleótido de CpG dividido por el número total de citosinas en posiciones de CpG. En otras realizaciones, el tamaño del grupo puede no ser igual a lo largo del genoma. En una realización, cada grupo entre dichos grupos de tamaños desiguales se compara entre múltiples sujetos.

Para determinar si la densidad de metilación de plasma de un caso ensayado era normal, se comparó la densidad de metilación con los resultados del grupo de referencia. Se seleccionaron aleatoriamente veintidós de los 32 sujetos sanos como el grupo de referencia para el cálculo de la puntuación z de metilación (Z_{met}).

$$Z_{met} = \frac{DM_{ensayo} - \overline{DM}_{ref}}{DM_{DT}}$$

donde DM_{ensayo} fue la densidad de metilación del caso ensayado para un grupo de 1 Mb; \overline{DM}_{ref} fue la densidad de metilación media del grupo de referencia para el grupo correspondiente; y DM_{DT} fue la DT de la densidad de metilación del grupo de referencia para el grupo correspondiente.

Para análisis de ANC, se determinó el número de lecturas secuenciadas que se mapeaban en cada grupo de 1 Mb (KCA Chan *et al.* 2013 Clin Chem 59: 211-24). Se determinó la densidad de lectura secuenciada para cada grupo después de corrección con respecto a desviación de GC usando regresión de suavizado de representación de dispersión ponderado localmente como se ha descrito previamente (EZ Chen *et al.* 2011 PLoS One 6: e21791). Para análisis de plasma, la densidad de lecturas secuenciadas del caso ensayado se comparó con el grupo de referencia para calcular la puntuación z de ANC (Z_{ANC}).

$$Z_{ANC} = \frac{DL_{ensayo} - \overline{DL}_{ref}}{DL_{DT}}$$

donde DL_{ensayo} fue la densidad de lecturas secuenciadas del caso ensayado para un grupo de 1 Mb particular; \overline{DL}_{ref} fue la densidad de lecturas secuenciadas media del grupo de referencia para el grupo correspondiente; y DL_{DT} fue la DT de la densidad de lecturas secuenciadas del grupo de referencia para el grupo correspondiente. Se definió que un grupo mostraba ANC si la Z_{ANC} del grupo 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 de diagnóstico, se seleccionaron aleatoriamente 10 millones de lecturas alineadas de cada caso. Se usó el mismo conjunto de individuos de referencia para establecer el intervalo de referencia de cada grupo de 1 Mb para el conjunto de datos con lecturas secuenciadas reducidas. El porcentaje de grupos que muestran hipometilación significativa, es decir, Z_{met} <-3 y el porcentaje de grupos con ANC, es decir, Z_{ANC} <-3 o >3, se determinaron para cada caso. Se usaron curvas de características operadoras receptoras (ROC) para ilustrar el rendimiento de diagnóstico de análisis de hipometilación y ANC de todo el genoma 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, los 32 sujetos sanos se usaron para el análisis.

La FIGURA 42 muestra un diagrama de rendimiento de diagnóstico de análisis de hipometilación de todo el genoma con diferente número de lecturas secuenciadas. Para análisis de 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 análisis de ANC, el rendimiento de diagnóstico se deterioró con una reducción significativa en las áreas bajo la curva cuando el número de lecturas secuenciadas se redujo del uso de los datos de un carril a 10 millones ($P < 0,001$).

2. Efecto del uso de diferente tamaño de grupos

Además de dividir el genoma en grupos de 1 Mb, también se exploró si pueden usarse tamaños de grupos más pequeños. Teóricamente, el uso de grupos más pequeños puede reducir potencialmente la variabilidad en la densidad de metilación dentro de un grupo. Esto se debe a que la densidad de metilación entre diferentes regiones genómicas puede variar ampliamente. Cuando un grupo es mayor, la probabilidad de incluir regiones con diferentes densidades de metilación aumentaría y, por lo tanto, conduciría a un aumento general de la variabilidad en la densidad de metilación de los grupos.

Aunque el uso de un tamaño de grupo menor puede reducir potencialmente la variabilidad de la densidad de metilación relacionada con la diferencia interregional, esto por otro lado reduciría el número de lecturas secuenciadas mapeadas en un grupo en particular. La reducción de lecturas que se mapean en grupos individuales aumentaría la variabilidad debido a la variación del número de muestras. El tamaño de grupo óptimo que puede dar lugar a menor variabilidad general en la densidad de metilación puede determinarse experimentalmente para los requisitos de una aplicación de diagnóstico particular, por ejemplo, el número total de lecturas secuenciadas por muestra y el tipo de secuenciador de ADN usado.

La FIGURA 43 es un diagrama que muestra curvas de ROC para la detección de cáncer basándose en análisis de hipometilación de todo el genoma con diferentes tamaños de grupos (50 kb, 100 kb, 200 kb y 1 Mb). Los valores de P mostrados son para comparación de área bajo la curva con un tamaño de grupo de 1 Mb. Puede verse una tendencia a la mejora cuando el tamaño del grupo se redujo de 1 Mb a 200 kb.

F. Puntuación de probabilidad acumulada

La cantidad de regiones para metilación y ANC puede tener diversos valores. Los ejemplos anteriores han descrito varias regiones que superan un valor de punto de corte o un porcentaje de dichas regiones que han mostrado hipometilación o ANC significativas como un parámetro para clasificar si una muestra estaba asociada con cáncer. Dichos enfoques no tienen en cuenta la magnitud de la aberración para grupos individuales. Por ejemplo, un grupo con una Z_{met} de -3,5 sería igual que un grupo con una Z_{met} de -30 ya que ambos se clasificarían como poseedores de hipometilación significativa. Sin embargo, el grado de cambios de hipometilación en el plasma, es decir, la magnitud del valor de Z_{met} , se ve afectado por la cantidad de ADN asociado a cáncer en la muestra y, por lo tanto, puede complementar la información del porcentaje de grupos que muestran aberraciones para reflejar la carga tumoral. Una mayor concentración fraccional de ADN tumoral en la muestra de plasma conduciría a una menor densidad de metilación y esto se traduciría a un valor de Z_{met} menor.

1. Puntuación de probabilidad acumulada como un parámetro de diagnóstico

Para hacer uso de la información de la magnitud de las aberraciones, se desarrolló un enfoque denominado puntuación de probabilidad acumulada (PA). Basándose en la función de probabilidad de distribución normal, cada valor de Z_{met} se tradujo a una probabilidad de tener dicha observación al azar.

La puntuación de PA se calculó como:

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

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

Una puntuación de PA mayor indica una menor probabilidad de tener dicha densidad de metilación desviada en una población normal al azar. Por lo tanto, una puntuación de PA alta indicaría una mayor probabilidad de tener ADN hipometilado de forma anómala en la muestra, por ejemplo, la presencia de ADN asociado a cáncer.

En comparación con el porcentaje de grupos que muestran aberración, la medición de puntuación de PA tiene un mayor intervalo dinámico. Aunque las cargas tumorales entre diferentes pacientes pueden variar ampliamente, el mayor intervalo de valores de PA sería útil para reflejar las cargas tumorales de pacientes con cargas tumorales

relativamente altas y relativamente bajas. Además, el uso de puntuaciones de PA puede ser potencialmente más sensible para detectar los cambios en la concentración de ADN asociado a tumor en plasma. Esto es ventajoso para la supervisión de respuesta al tratamiento y pronóstico. Por lo tanto, una reducción en las puntuaciones de PA durante el tratamiento es indicativa de una buena respuesta al tratamiento. Una falta de reducción o incluso aumento en las puntuaciones de PA durante el tratamiento indicaría escasa respuesta o falta de respuesta. Para pronóstico, una alta puntuación de PA es indicativa de alta carga tumoral y sugiere un mal pronóstico (por ejemplo, mayor probabilidad de muerte o progresión tumoral).

La FIGURA 44A muestra un rendimiento de diagnóstico para probabilidad acumulada (PA) y porcentaje de grupos con aberraciones. No hubo ninguna diferencia significativa entre las áreas bajo la curva para los dos tipos de algoritmo de diagnóstico ($P = 0,791$).

La FIGURA 44B muestra rendimientos de diagnóstico para el análisis de plasma con respecto a hipometilación global, hipermetilación de islas de CpG y ANC. Con un carril de secuenciación por muestra (200 kb de tamaño de grupo para análisis de hipometilación y 1 Mb de tamaño de grupo para ANC, e islas de CpG definidas de acuerdo con la base de datos alojada por la Universidad de California, Santa Cruz (UCSC)), las áreas bajo la curva para los tres tipos de análisis estuvieron por encima de 0,90.

En los análisis posteriores, se usó la mayor puntuación de PA en los sujetos de control como el punto de corte para cada uno de los tres tipos de análisis. La selección de estos puntos de corte proporcionó una especificidad de diagnóstico del 100 %. Las sensibilidades de diagnóstico para análisis de hipometilación general, hipermetilación de islas de CpG y ANC fueron del 78 %, 89 % y 52 %, respectivamente. En 43 de los 46 enfermos de cáncer, se detectó al menos uno de los tres tipos de aberraciones, dando lugar por lo tanto a una sensibilidad del 93,4 % y una especificidad del 100 %. Los resultados de los inventores indican que los tres tipos de análisis pueden usarse de forma sinérgica para la detección de cáncer.

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

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

2. Aplicación de puntuación de PA para supervisión de cáncer

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

La FIGURA 47 muestra un análisis en serie para metilación de plasma para el caso TBR34. El anillo más interno muestra la densidad de metilación de la capa leucoplaquetaria (negra) y tejidos tumorales (púrpura). Para las muestras de plasma, la Z_{met} se muestra para cada grupo de 1 Mb. La diferencia entre dos líneas representa una diferencia de Z_{met} de 5. Los puntos rojos y grises representan grupos con hipometilación y sin cambios en la densidad de metilación en comparación con el grupo de referencia. A partir del segundo anillo interno hacia fuera están las muestras de plasma tomadas antes del tratamiento, a los 3 días y a los 2 meses después de la resección tumoral, respectivamente. Antes del tratamiento, pudo observarse un mayor grado de hipometilación en el plasma y más del 18,5 % de los grupos tuvieron una Z_{met} de <-10 . A los 3 días después de la resección tumoral, se pudo observar que el grado de metilación se redujo en el plasma con ninguno de los grupos con Z_{met} de <-10 .

Tabla 5

N.º de caso	Punto temporal	Análisis de metilación		
		Porcentaje de grupos que muestran hipometilación significativa	Puntuación de probabilidad acumulada (PA)	Puntuación z sumatoria
TBR34	Antes de OT	62,6 %	37.573	14.285
	3 días después de OT	80,5 %	17.777	9.195
	2 meses después de OT	40,1 %	15.087	5.201

La Tabla 5 muestra que, aunque la magnitud de los cambios de hipometilación se redujo a los 3 días después de la resección quirúrgica del tumor, el porcentaje de grupos que mostraban aberración mostró un aumento paradójico. Por otro lado, la puntuación de PA reveló con más precisión la reducción en el grado de hipometilación en plasma y puede reflejar mejor los cambios en la carga tumoral.

5 A los 2 meses después de OT, aún hubo un porcentaje significativo de grupos que mostraban cambios de hipometilación. La puntuación de PA también permaneció estática a aproximadamente 15.000. Se diagnosticó posteriormente que este paciente tenía depósitos tumorales multifocales (previamente desconocidos en el momento de la cirugía) en el hígado no resecado restante a los 3 meses y se observó que tenía múltiples metástasis pulmonares a los 4 meses después de la operación. El paciente murió por enfermedad metastásica a los 8 meses después de la operación. Estos resultados sugirieron que la puntuación de PA podía ser más potente que el porcentaje de grupos con aberración para reflejar la carga tumoral.

15 En general, la PA puede ser útil para aplicaciones que requieran medir la cantidad de ADN tumoral en plasma. Los ejemplos de dichas aplicaciones incluyen: pronóstico y supervisión de enfermos de cáncer (por ejemplo, para observar respuesta a tratamiento, o para observar progresión tumoral).

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

X. Influencia de ANC en la metilación

25 El uso de ANC y metilación para determinar clasificaciones respectivas con respecto a un nivel de cáncer, donde las clasificaciones se combinan para proporcionar una tercera clasificación, se ha descrito anteriormente. Además de dicha combinación, ANC puede usarse para cambiar los valores de punto de corte para el análisis de metilación e identificar falsos positivos comparando los niveles de metilación para grupos de regiones que tengan diferentes características de ANC. Por ejemplo, el nivel de metilación para superabundancia (por ejemplo, $Z_{ANC} > 3$) puede compararse con el nivel de metilación para abundancia normal (por ejemplo, $-3 < Z_{ANC} < 3$). En primer lugar, se describe la influencia de ANC en los niveles de metilación.

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

35 Como los tejidos tumorales generalmente muestran una hipometilación general, la presencia de ADN derivado de tumor en el plasma de enfermos de cáncer conduciría a la reducción en la densidad de metilación en comparación con sujetos sin cáncer. El grado de hipometilación en el plasma de enfermos de cáncer es teóricamente proporcional a la concentración fraccional de ADN derivado de tumor en la muestra de plasma.

40 Para regiones que muestran una ganancia cromosómica en los tejidos tumorales, se liberaría una dosificación adicional de ADN tumoral de los segmentos de ADN amplificados al plasma. Esta contribución aumentada de ADN tumoral al plasma teóricamente conduciría a un mayor grado de hipometilación en el ADN de plasma para la región afectada. Un factor adicional es que se esperaría que regiones genómicas que mostraban amplificación confirieran ventaja de crecimiento a las células tumorales, y por lo tanto se esperaría que se expresaran. Dichas regiones están en general hipometiladas.

50 Por el contrario, para regiones que muestran pérdida cromosómica en el tejido tumoral, la contribución reducida de ADN tumoral al plasma conduciría a un menor grado de hipometilación en comparación con regiones sin cambio del número de copias. Un factor adicional es que las regiones genómicas que se suprimen en células tumorales podrían contener genes supresores tumorales y podría ser ventajoso para células tumorales que dichas regiones estuvieran silenciadas. Por lo tanto, se espera que dichas regiones tengan una mayor probabilidad de estar hipermetiladas.

55 Aquí, se usan los resultados de dos enfermos de CHC (TBR34 y TBR36) para ilustrar este efecto. Las FIGURAS 48A (TBR36) y 49A (TBR34) tienen círculos que destacan regiones con ganancias o pérdidas cromosómicas y el análisis de metilación correspondiente. Las FIGURAS 48B y 49B muestran representaciones de puntuaciones z de metilación para pérdidas, normales y ganancias para los pacientes TBR36 y TBR34, respectivamente.

60 La FIGURA 48A muestra representaciones de Circos que demuestran los cambios de ANC (anillo interno) y metilación (anillo externo) en el ADN de plasma tratado por 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 regiones sin cambios del número de copias. Las regiones que mostraban pérdidas cromosómicas estaban menos hipometiladas que regiones sin cambios del número de copias. La FIGURA 48B es una representación de puntuaciones z de metilación para regiones con ganancias y pérdidas cromosómicas, y regiones sin cambio del número de copias para el paciente con CHC TBR36. En comparación con regiones sin cambios de copias, las regiones con ganancias cromosómicas tuvieron más puntuaciones z negativas (más hipometilación) y regiones con pérdidas cromosómicas tuvieron menos puntuaciones z negativas (menos hipometiladas).

La FIGURA 49A muestra representaciones de Circos que demuestran los cambios de ANC (anillo interno) y metilación (anillo externo) en el ADN de plasma tratado por bisulfito para el paciente con CHC TBR34. La FIGURA 49B es una representación de las puntuaciones z de metilación para regiones con ganancias y pérdidas cromosómicas, y regiones sin cambio del número de copias para el paciente con CHC TBR34. La diferencia en las densidades de metilación entre regiones con ganancias y pérdidas cromosómicas fue mayor en el paciente TBR36 que en el paciente TBR34 debido a que la concentración fraccional de ADN derivado del tumor en el paciente anterior era mayor.

En este ejemplo, las regiones usadas para determinar ANC son las mismas que las regiones usadas para determinar la metilación. En una realización, los valores de punto de corte regionales respectivos dependen de si la región respectiva muestra una delección o una amplificación. En una implementación, un valor de punto de corte regional respectivo (por ejemplo, el punto de corte de puntuación z usado para determinar la hipometilación) tiene una mayor magnitud cuando la región respectiva muestra una amplificación que cuando no se muestra amplificación (por ejemplo, la magnitud podría ser mayor de 3, y puede usarse un punto de corte de menos de -3). Por lo tanto, para ensayar la hipometilación, un valor de punto de corte regional respectivo puede tener un mayor valor negativo cuando la región respectiva muestra una amplificación que cuando no se muestra amplificación. Se espera que dicha implementación mejore la especificidad del ensayo para detectar cáncer.

En otra implementación, un valor de punto de corte regional respectivo tiene una menor magnitud (por ejemplo, menos de 3) que cuando la región respectiva muestra una delección que cuando no se muestra delección. Por lo tanto, para ensayar la hipometilación, un valor de punto de corte regional respectivo puede tener un valor menos negativo que cuando la región respectiva muestra una delección que cuando no se muestra delección. Se espera que dicha implementación mejore la sensibilidad del ensayo para detectar cáncer. El ajuste de los valores de punto de corte en las implementaciones anteriores puede cambiarse dependiendo de la sensibilidad y especificidad deseadas para un escenario de diagnóstico particular. En otras realizaciones, las mediciones de metilación y ANC pueden usarse junto con otros parámetros clínicos (por ejemplo, resultados de captura de imágenes o bioquímica de suero) para predicción del cáncer.

B. Uso de ANC para seleccionar regiones

Como se ha descrito anteriormente, se ha mostrado que la densidad de metilación de plasma se alteraría en regiones que tengan aberraciones del número de copias en los tejidos tumorales. En regiones con ganancia del número de copias en el tejido tumoral, un aumento de la contribución de ADN tumoral hipometilado al plasma conduciría a un mayor grado de hipometilación de ADN de plasma en comparación con regiones sin una aberración del número de copias. Por el contrario, en regiones con pérdida de número de copias en el tejido tumoral, la contribución reducida de ADN derivado de cáncer hipometilado al plasma conduciría a un menor grado de hipometilación de ADN de plasma. Esta relación entre la densidad de metilación de ADN de plasma y la representación relativa puede usarse potencialmente para diferenciar resultados de hipometilación asociados con la presencia de ADN asociado a cáncer y otras causas no cancerosas (por ejemplo, LES) de hipometilación en ADN de plasma.

Para ilustrar este enfoque, se analizaron las muestras de plasma de dos enfermos de carcinoma hepatocelular (CHC) y dos pacientes con LES sin cáncer. Estos dos enfermos de LES (SLE04 y SLE10) mostraron la presencia aparente de hipometilación y ANC en plasma. Para el paciente SLE04, 84 % de los grupos mostraron hipometilación y 11,2 % de los grupos mostraron ANC. Para el paciente SLE10, 10,3 % de los grupos mostraron hipometilación y 5,7 % de los grupos mostraron ANC.

Las FIGURAS 50A y 50B muestran resultados de análisis de hipometilación y ANC de plasma para enfermos de LES SLE04 y SLE10. El círculo externo muestra las puntuaciones z de metilación (Z_{met}) a una resolución de 1 Mb. Los grupos con metilación $Z_{\text{met}} < -3$ estaban en rojo y los de $Z_{\text{met}} > -3$ estaban en gris. El círculo interno muestra las puntuaciones z de ANC (Z_{ANC}). Los puntos verdes, rojos y grises representan grupos con $Z_{\text{ANC}} > 3$, < 3 y entre -3 y 3, respectivamente. En estos dos enfermos de LES, se observaron cambios de hipometilación y ANC en plasma.

Para determinar si los cambios en metilación y ANC eran coherentes con la presencia de ADN derivado de cáncer en plasma, se comparó la Z_{met} para regiones con $Z_{\text{ANC}} > 3$, < -3 y entre -3 y 3. Para cambios de metilación y ANC a los que contribuyó ADN derivado de cáncer en plasma, se esperaba que regiones con $Z_{\text{ANC}} < -3$ estuvieran menos hipometiladas y tuvieran menos Z_{met} negativa. Por el contrario, se esperaba que regiones con $Z_{\text{ANC}} > 3$ estuvieran más hipometiladas y tuvieran más Z_{met} negativa. Para fines de ilustración, se aplicó un ensayo de suma de rangos unilaterales para comparar la Z_{met} para regiones con ANC (es decir regiones con $Z_{\text{ANC}} < -3$ o > 3) con regiones sin ANC (es decir regiones con Z_{ANC} entre -3 y 3). En otras realizaciones, pueden usarse otros ensayos estadísticos, por ejemplo, pero sin limitación, ensayo de t de Student, ensayo de análisis de varianza (ANOVA) y ensayo de Kruskal-Wallis.

Las FIGURAS 51A y 51B muestran análisis de Z_{met} para regiones con y sin ANC para el plasma de dos enfermos de CHC (TBR34 y TBR36). Las regiones con $Z_{\text{ANC}} < -3$ y > 3 representan regiones con infra y sobrerrepresentación en plasma, respectivamente. Tanto en TBR34 como en TBR36, las regiones que estaban infrarrepresentadas en plasma (es decir regiones con $Z_{\text{ANC}} < -3$) tuvieron Z_{met} significativamente mayor (Valor de $P < 10^{-5}$, ensayo de sumas de rangos

unilaterales) que 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. Para regiones con sobrerrepresentación en plasma (es decir regiones con $Z_{ANC} > 3$), tuvieron Z_{met} significativamente menor que regiones con representación normal en plasma (valor de $P < 10^{-5}$, ensayo de sumas de rangos unilateral). Todos estos cambios fueron coherentes con la presencia de ADN tumoral hipometilado en las muestras de plasma.

Las FIGURAS 51C y 51D muestran análisis de Z_{met} para regiones con y sin ANC para el plasma de dos enfermos de LES (SLE04 y SLE10). Las regiones con $Z_{ANC} < -3$ y > 3 representan regiones con infra y sobrerrepresentación en plasma, respectivamente. Para SLE04, las regiones que estaban infrarrepresentadas en plasma (es decir regiones con $Z_{ANC} < -3$) no tuvieron Z_{met} significativamente mayor (Valor de $P = 0,99$, ensayo de sumas de rangos unilateral) que regiones con representación normal en plasma (es decir regiones con Z_{ANC} entre -3 y 3) y regiones con sobrerrepresentación en plasma (es decir regiones con $Z_{ANC} > 3$) no tuvieron Z_{met} significativamente menor que regiones con representación normal en plasma (Valor de $P = 0,68$, ensayo de suma de rangos unilateral). Estos resultados fueron diferentes de los cambios esperados debido a la presencia de ADN hipometilado derivado de tumor en plasma. De forma similar, para SLE10, las regiones con $Z_{ANC} < -3$ no tuvieron Z_{met} significativamente mayor que regiones con Z_{ANC} entre -3 y 3 (Valor de $P = 0,99$, ensayo de suma de rangos unilateral).

Una razón para no tener el patrón asociado a cáncer típico entre Z_{met} y Z_{ANC} en los enfermos de LES es que, en los enfermos de LES, la ANC no está presente en un tipo celular específico que también muestra hipometilación. En su lugar, la presencia aparente observada de ANC e hipometilación se debe a la distribución de tamaños alterada de ADN en circulación en enfermos de LES. La distribución de tamaños alterada podría alterar potencialmente las densidades de lectura secuenciadas para diferentes regiones genómicas lo que conduce a ANC aparentes ya que las referencias se obtuvieron de sujetos sanos. Como se ha descrito en las secciones anteriores, existe una correlación entre el tamaño de un fragmento de ADN en circulación y su densidad de metilación. Por lo tanto, la distribución de tamaños alterada también puede conducir a una metilación aberrante.

Aunque las regiones con $Z_{ANC} > 3$ tuvieron niveles de metilación ligeramente menores que regiones con Z_{ANC} entre -3 y 3, el valor de P para la comparación fue mucho mayor que los observados en dos enfermos de cáncer. En una realización, el valor de P puede usarse como un parámetro para determinar la probabilidad de que un caso ensayado tenga cáncer. En otra realización, la diferencia en Z_{met} entre regiones con representación normal y aberrante puede usarse como un parámetro para indicar la probabilidad de la presencia de cáncer. En una realización, puede usarse un grupo de enfermos de cáncer para establecer la correlación entre Z_{met} y Z_{ANC} y para determinar los umbrales para diferentes parámetros para indicar que los cambios son coherentes con la presencia de ADN hipometilado derivado de cáncer en la muestra de plasma ensayada.

En consecuencia, en una realización, puede realizarse un análisis de ANC para determinar un primer conjunto de regiones que muestran todas: una delección, una amplificación o representación normal. Por ejemplo, el primer conjunto de regiones puede mostrar, todas ellas, una delección, o puede mostrar, todas ellas una amplificación o mostrar, todas ellas, una representación normal (por ejemplo, tienen una primera cantidad normal de regiones, tal como una Z_{met} normal). Puede determinarse un nivel de metilación para este primer conjunto de regiones (por ejemplo, el primer nivel de metilación del método 2800 puede corresponder al primer conjunto de regiones).

El análisis de ANC puede determinar un segundo conjunto de regiones que muestran, todas ellas, una segunda de: una delección, una amplificación o representación normal. El segundo conjunto de regiones se mostraría de forma diferente al primer conjunto. Por ejemplo, si el primer conjunto de regiones fuera normal, entonces el segundo conjunto de regiones puede mostrar una delección o una amplificación. Puede calcularse un segundo nivel de metilación basándose en los números respectivos de moléculas de ADN metiladas en sitios en el segundo conjunto de regiones.

Después puede calcularse un parámetro entre el primer nivel de metilación y la segunda metilación. Por ejemplo, puede calcularse una diferencia o relación y compararse con un valor de punto de corte. La diferencia o relación también puede someterse a una distribución de probabilidad (por ejemplo, como parte de un ensayo estadístico) para determinar la probabilidad de obtener el valor, y esta probabilidad puede compararse con un valor de punto de corte para determinar un nivel de cáncer basándose en niveles de metilación. Dicho punto de corte puede seleccionarse para diferenciar muestras que tengan cáncer y las que no tengan cáncer (por ejemplo, LES).

En una realización, puede determinarse un nivel de metilación para el primer conjunto de región o una mezcla de regiones (es decir, mezcla de regiones que muestran amplificación, delección y normales). Este nivel de metilación puede después compararse con un primer punto de corte como parte de un primer estadio de análisis. Si se supera el punto de corte, lo que indica de este modo una posibilidad de cáncer, entonces el análisis anterior puede realizarse para determinar si la indicación fue un falso positivo. La clasificación final para el nivel de cáncer puede por lo tanto incluir la comparación del parámetro para los dos niveles de metilación con un segundo punto de corte.

El primer nivel de metilación puede ser un valor estadístico (por ejemplo, promedio o mediana) de los niveles de metilación regionales calculados para cada región del primer conjunto de regiones. El segundo nivel de metilación también puede ser un valor estadístico de niveles de metilación regionales calculados para cada región del segundo conjunto de regiones. Como ejemplos, los valores estadísticos pueden determinarse usando de suma de rangos

unilateral, ensayo de t de Student, ensayo de análisis de varianza (ANOVA) o ensayo de Kruskal-Wallis.

XI. CLASIFICACIÓN DEL TIPO DE CÁNCER

- 5 Además de determinar si un organismo tiene cáncer o no, las realizaciones pueden identificar un tipo de cáncer asociado con la muestra. La identificación del tipo de cáncer puede usar patrones de hipometilación global, hipermetilación de islas de CpG y/o ANC. Los patrones pueden implicar el agrupamiento de pacientes con un diagnóstico conocido usando niveles de metilación regional medidos, valores de ANC respectivos para regiones y nivel de metilación para las islas de CpG. Los resultados posteriores muestran que organismos con un tipo similar de cáncer
- 10 tienen valores similares para las regiones e islas de CpG, así como los pacientes sin cáncer que tienen valores similares. En el agrupamiento, cada uno de los valores para una región o isla puede ser una dimensión separada en el proceso de agrupamiento.

- 15 Se ha sabido que el mismo tipo de cánceres compartirían cambios genéticos y epigenéticos similares (E Gebhart *et al.*, 2004 Cytogenet Genome Res; 104: 352-358; PA Jones *et al.*, 2007 Cell; 128: 683-692). A continuación, se describe cómo los patrones de ANC y cambio de metilación detectados en el plasma son útiles para inferir el origen o tipo del cáncer. Las muestras de ADN de plasma de los enfermos de CHC, no enfermos de CHC y sujetos de control sanos se clasificaron usando, por ejemplo, análisis de agrupamiento jerárquico. El análisis se realizó usando, por ejemplo, la función heatmap.2 en el paquete de script R (cran.r-project.org/web/packages/gplots/gplots.pdf).

- 20 Para ilustrar el potencial de este enfoque, se usaron dos conjuntos de criterios (grupo A y grupo B) como ejemplos para identificar elementos útiles para la clasificación de las muestras de plasma (véase Tabla 6). En otras realizaciones, pueden usarse otros criterios para identificar los elementos. Los elementos usados incluyeron ANC global a una resolución de 1 Mb, densidad de metilación global a una resolución de 1 Mb y metilación de islas de CpG.

25

Metilación global a una resolución de 1 Mb		
	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	1.911

Elementos de 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 $z > 2,5$ o $< -2,5$
n.º de elementos identificados	355	759

Metilación de islas de 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 2 % en las islas de CpG particulares	>1 caso de cáncer con una densidad de metilación que difiere de la media de la referencia en 2 % en las islas de CpG particulares
n.º de elementos identificados	110	191

Tabla 6

- En los dos primeros ejemplos, se usaron todos los elementos de ANC, metilación global a una resolución de 1 Mb y metilación de islas de CpG para la clasificación. En otras realizaciones, pueden usarse otros criterios, por ejemplo,
- 30 pero sin limitación, la precisión de medición del elemento en el plasma del grupo de referencia.

- La FIGURA 52A muestra análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sano usando los 1.130 elementos del grupo A incluyendo 355 ANC, 584 elementos de metilación global a una resolución de 1 Mb y el estado de metilación de 110 islas de CpG.
- 35 La barra de color de la parte superior representa los grupos de muestras: verde, azul y rojo representan los sujetos sanos, enfermos de CHC y enfermos de cáncer distinto de CHC, respectivamente. En general, los tres grupos de sujetos tienden a agruparse entre sí. El eje vertical representa los elementos de clasificación. Se agruparon entre sí elementos con patrones similares entre diferentes sujetos. Estos resultados sugieren que los patrones de cambios de metilación de islas de CpG, cambios de metilación de todo el genoma a una resolución de 1 Mb y ANC en plasma
- 40 pueden usarse potencialmente para determinar el origen del cáncer en pacientes con primarios desconocidos.

La FIGURA 52B muestra análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sanos usando los 2.780 elementos del grupo B incluyendo 759 ANC, 1.911 de metilación global a una resolución de 1 Mb y el estado de metilación de 191 islas de CpG. La barra de color de la parte superior representa los grupos de muestras: verde, azul y rojo representan los sujetos sanos, enfermos de CHC y enfermos de cáncer distinto de CHC, respectivamente. En general, los tres grupos de sujetos tendieron a agruparse entre sí. El eje vertical representa los elementos de clasificación. Se agruparon entre sí elementos con patrones similares entre diferentes sujetos. Estos resultados sugieren que los patrones de diferentes conjuntos de cambios de metilación de islas de CpG, cambios de metilación de todo el genoma a una resolución de 1 Mb y ANC en plasma pueden usarse para determinar el origen del cáncer en pacientes con primarios desconocidos. La selección de los elementos de clasificación puede ajustarse para aplicaciones específicas. Además, puede darse peso a la predicción del tipo de cáncer según las probabilidades anteriores de los sujetos para diferentes tipos de cánceres. 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 lo tanto, puede calcularse una probabilidad ponderada del tipo de cáncer usando, por ejemplo, pero sin limitación, regresión logística, múltiple o de agrupamiento.

En otras realizaciones, puede usarse un único tipo de elementos para el análisis de clasificación. Por ejemplo, en los ejemplos siguientes, solamente se usaron la metilación global a una resolución de 1 Mb, la hipermetilación de islas de CpG o las ANC a una resolución de 1 Mb para el análisis de agrupamiento jerárquico. La potencia de diferenciación puede ser diferente cuando se usan diferentes elementos. El refinamiento adicional de los elementos de clasificación puede mejorar potencialmente las precisiones de clasificación.

La FIGURA 53A muestra análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sanos usando los elementos de metilación de islas de CpG del grupo A. En general, los enfermos de cáncer se agruparon entre sí y los sujetos sin cáncer estuvieron en otro grupo. Sin embargo, los enfermos de CHC y sin CHC estuvieron menos separados en comparación con el uso de los tres tipos de elementos.

La FIGURA 53B muestra el análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sanos usando las densidades de metilación globales del grupo A a una resolución de 1 Mb como elementos de clasificación. Se observó agrupamiento preferente de enfermos de CHC y sin CHC.

La FIGURA 54A muestra un análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sanos usando las ANC globales del grupo A a una resolución de 1 Mb como elementos de clasificación. Se vio agrupamiento preferente de enfermos de CHC y sin CHC.

La FIGURA 54B muestra un análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sanos usando las densidades de metilación de islas de CpG del grupo B como elementos de clasificación. Pudo observarse agrupamiento preferente de enfermos de CHC y de cáncer distinto de CHC.

La FIGURA 55A muestra un análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sanos usando las densidades de metilación globales del grupo B a una resolución de 1 Mb como elementos de clasificación. Se pudo observar agrupamiento preferente de enfermos de CHC y cáncer distinto de CHC.

La FIGURA 55B muestra un análisis de agrupamiento jerárquico para muestras de plasma de enfermos de CHC, enfermos de cáncer distinto de CHC y sujetos de control sanos usando las ANC globales del grupo B a una resolución de 1 Mb como elementos de clasificación. Se pudo observar agrupamiento preferente de enfermos de CHC y de cáncer distinto de CHC.

Estos resultados de agrupamiento jerárquico para muestras de plasma sugieren que la combinación de diferentes elementos puede usarse potencialmente para la identificación de los tipos de cáncer primarios. El refinamiento adicional de los criterios de selección puede mejorar adicionalmente de forma potencial la precisión de la clasificación.

En consecuencia, en una realización, cuando una clasificación de metilación indica que existe cáncer para el organismo, puede identificarse un tipo de cáncer asociado con el organismo comparando un nivel de metilación (por ejemplo, primera metilación del método 2800 o cualquier nivel de metilación regional) 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. Se identifican al menos dos de los otros organismos como poseedores de diferentes tipos de cáncer. Por ejemplo, los valores correspondientes pueden organizarse en grupos, donde se asocian dos grupos con diferentes cánceres.

Además, cuando se usan ANC y metilación juntas para obtener una tercera clasificación del nivel de cáncer, pueden compararse los elementos de ANC y metilación con valores correspondientes de otros organismos. Por ejemplo, la primera cantidad de regiones (por ejemplo, de la FIGURA 36) que muestran una delección o amplificación puede compararse con valores correspondientes determinados a partir de los otros organismos para identificar el tipo de cáncer asociado con el organismo.

En algunas realizaciones, los elementos de metilación son los niveles de metilación regionales de una pluralidad de regiones del genoma. Pueden usarse regiones que se ha determinado que tienen un nivel de metilación regional que supera el valor de punto de corte regional respectivo, por ejemplo, pueden compararse niveles de metilación regionales del organismo con niveles de metilación regionales de otros organismos para las mismas regiones del genoma. La comparación puede permitir diferenciar tipos de cáncer, o solamente proporcionar un filtro adicional para confirmar el cáncer (por ejemplo, para identificar falsos positivos). Por lo tanto, se puede determinar si el organismo tiene el primer tipo de cáncer, ausencia de cáncer o el segundo tipo de cáncer basándose en la comparación.

Los otros organismos (junto con el que se ensaya) pueden agruparse usando los niveles de metilación regionales. Por lo tanto, puede usarse una comparación de los niveles de metilación regionales para determinar a qué grupo pertenece el organismo. El agrupamiento también puede usar valores normalizados de ANC para regiones que se ha determinado que muestra una delección o una amplificación, como se ha descrito anteriormente. Además, el agrupamiento puede usar las densidades de metilación respectivas de islas de CpG hipermetiladas.

Para ilustrar el principio de este método, se ha mostrado un ejemplo del uso de regresión logística para la clasificación de dos muestras desconocidas. El fin de esta clasificación fue determinar si estas dos muestras eran CHC o cánceres distintos de CHC. Se compiló un conjunto de entrenamiento de muestras que incluían 23 muestras de plasma recogidas de enfermos de CHC y 18 muestras de enfermos que padecen cáncer distinto de CHC. Por lo tanto, hubo un total de 41 casos en el conjunto de entrenamiento. En este ejemplo, se seleccionaron 13 elementos, incluyendo cinco elementos en la metilación de islas de CpG (X1-X5), seis elementos en la metilación de regiones de 1 Mb (X6-X11) y 2 elementos en la ANC de regiones de 1 Mb (X12-X13). Los elementos de metilación de CpG se seleccionaron basándose en el criterio de que al menos 15 casos en el conjunto de entrenamiento tengan una puntuación z de >3 o <-3 . Los elementos de metilación de 1 Mb se seleccionaron basándose en el criterio de que al menos 39 casos en el conjunto de entrenamiento tengan una puntuación z de >3 o <-3 . Los elementos de ANC se seleccionaron basándose en el criterio de que al menos 20 casos tengan una puntuación z de >3 o <-3 . Se realizó regresión logística en las muestras de este conjunto de entrenamiento para determinar el coeficiente de regresión para cada uno de los elementos (X1-X13). Los elementos con coeficientes de regresión de las magnitudes mayores (independientemente de si está en un sentido positivo o negativo) ofrecen mejor diferenciación entre muestras de CHC y distintas de CHC. Las puntuaciones z de cada caso para los elementos respectivos se usaron como los valores de aporte de las variables independientes. Después se analizaron dos muestras de plasma, una de un paciente con CHC (TBR36) y una de un paciente que padece cáncer de pulmón (TBR177) con respecto a los 13 elementos.

En este análisis de clasificación de tipo de cáncer, se supuso que estas dos muestras se recogían de pacientes que padecían cánceres de origen desconocido. Para cada muestra, las puntuaciones z para el elemento respectivo se pusieron en la ecuación de regresión logística para determinar el algoritmo natural de la relación de probabilidad ($\ln(\text{relación de probabilidad})$) donde la relación de probabilidad representaba la relación de probabilidades de tener CHC y no tener CHC (CHC/no CHC).

La Tabla 7 muestra los coeficientes de regresión para los 13 elementos de la ecuación de regresión logística. También se muestran las puntuaciones z para los elementos respectivos de los dos casos ensayados (TBR36 y TBR177). El $\ln(\text{relación de probabilidad})$ de CHC para TBR36 y TBR177 fue de 37,03 y -4,37, respectivamente. A partir de estas relaciones de probabilidad, se calculó la probabilidad de recoger muestras de plasma de enfermos de CHC como $>99,9\%$ y 1% , respectivamente. Brevemente, TBR36 tuvo una mayor probabilidad de ser una muestra de un paciente con CHC, mientras que TBR177 tuvo una menor probabilidad de ser una muestra de un paciente con CHC.

Tabla 7

Elemento	Coeficiente de regresión	Puntuación z del elemento respectivo	
		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
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

(continuación)

Elemento	Coeficiente de regresión	Puntuación z del elemento respectivo	
		TBR36	TBR177
X12	-1,1826	4,8	3,3
X13	-0,3805	-11,7	-1,4
	ln(relación de probabilidad)	37,03	-4,37463

En otras realizaciones, puede usarse regresión de agrupamiento jerárquico, análisis de árboles de clasificación y otros modelos de regresión para determinar el origen primario probable del cáncer.

5

XII. MATERIALES Y MÉTODOS

A. Preparación de bibliotecas de ADN tratado por bisulfito y secuenciación

10 Se fragmentó ADN genómico (5 µg) al que se añadió ADN lambda desmetilado 0,5 % (p/p) (Promega) por un sistema Covaris S220 (Covaris) hasta aproximadamente 200 pb de longitud. Se prepararon bibliotecas de ADN usando el Kit de Preparación de Muestras de Secuenciación de Extremos Emparejados (Illumina) según las instrucciones del fabricante, excepto que se ligaron adaptadores metilados (Illumina) a los fragmentos de ADN. Después de dos ciclos de purificación usando perlas magnéticas AMPure XP (Beckman Coulter), los productos de ligamiento se dividieron

15 en 2 partes, una de las cuales se sometió a 2 ciclos de modificación por bisulfito con un kit de Bisulfito EpiTect (Qiagen). Las citosinas desmetiladas en sitios de CpG en los insertos se convirtieron en uracilos mientras que las citosinas metiladas permanecieron sin cambios. Las moléculas de ADN ligadas a adaptadores, bien tratadas o bien no tratadas con bisulfito sódico, se enriquecieron por 10 ciclos de PCR usando la siguiente receta: 2,5 U de ADN polimerasa de inicio en caliente PfuTurboCx (Agilent Technologies), tampón de reacción PfuTurboCx 1X, 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 termociclación fue: 95 °C durante 2 min, 98 °C durante 30 s, después 10 ciclos de 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 PCR se purificaron usando perlas magnéticas AMPure XP.

20

25 Se añadió ADN de plasma extraído de 3,2-4 ml de muestras de plasma materno con ADN lambda fragmentado (25 pg por ml de plasma) y se sometió a construcción de bibliotecas como se ha descrito anteriormente (RWK Chiu *et al.* 2011 BMJ; 342: c7401). Después de ligar con los adaptadores metilados, los productos de ligamiento se dividieron en 2 mitades y una parte se sometió a 2 ciclos de modificación por bisulfito. Los productos de ligamiento tratados o no tratados por bisulfito se enriquecieron después por 10 ciclos de PCR como se ha descrito anteriormente.

30

Las bibliotecas de ADN tratado o no tratado por bisulfito se secuenciaron para 75 pb en un formato de extremos emparejados en instrumentos HiSeq2000 (Illumina). Se generaron grupos de ADN con un Kit de Generación de Grupos de Extremos Emparejados v3 en un instrumento cBot (Illumina). Se realizó análisis de imágenes en tiempo real y determinación de bases usando el Software de Control HiSeq (HCS) v1.4 y Software de Análisis en Tiempo Real (RTA) v1.13 (Illumina), por los que la matriz automática y los cálculos de fase se basaron en el control de PhiX v3 añadido secuenciado con las bibliotecas de ADN.

35

B. Alineamiento de secuencias e identificación de citosinas metiladas

40 Después de la determinación de las bases, se retiraron las secuencias adaptadoras y bases de baja calidad (es decir puntuación de calidad <20) en los extremos de los fragmentos. Las lecturas recortadas en formato FASTQ se procesaron después por un proceso de análisis de datos de metilación denominado Methy-Pipe (P Jiang, *et al.* Methy-Pipe: An integrated bioinformatics data analysis pipeline for whole genome methylome analysis, artículo presentado en los Talleres de la Conferencia Internacional sobre Bioinformática y Biomedicina de IEEE, Hong Kong, del 18 al 21 de diciembre de 2010). Para alinear las lecturas de secuenciación convertidas por bisulfito, se realizó en primer lugar conversión por ordenador de todos los restos de citosina en timinas, en las cadenas de Watson y Crick por separado, usando el genoma humano de referencia (NCBI versión 36/hg18). Después se realizó conversión por ordenador de cada citosina en timina en todas las lecturas procesadas y se mantuvo la información posicional de cada resto convertido. Se usó 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 desapareamientos permitidos para cada extremo alineado. Solamente se seleccionaron lecturas mapeables en una única localización genómica. Las lecturas ambiguas que se mapearon en las cadenas tanto de Watson como de Crick y lecturas duplicadas (clonales) que tenían en las mismas posiciones genómicas de inicio y fin se retiraron. Se conservaron lecturas secuenciadas con tamaño de inserto ≤600 pb para los análisis de metilación y tamaño.

45

50

55

Los restos de citosina en el contexto del dinucleótido de CpG fueron las dianas principales para los estudios de metilación de ADN cadena abajo. Después del alineamiento, las citosinas presentes originalmente en las lecturas secuenciadas se recuperaron basándose en la información posicional mantenida durante la conversión por ordenador. Las citosinas recuperadas entre los dinucleótidos de CpG se puntuaron como metiladas. Las timinas entre los dinucleótidos de CpG se puntuaron como desmetiladas. El ADN lambda desmetilado incluido durante la preparación

60

de bibliotecas actuó como un control interno para estimar la eficacia de la modificación por bisulfito sódico. Todas las citosinas en el ADN lambda deberían haberse convertido en timinas si la eficacia de conversión por bisulfito fuera del 100 %.

5 XIII. SUMARIO

Con el uso de las realizaciones descritas en el presente documento, se podría explorar, detectar, supervisar o pronosticar el cáncer de forma no invasiva usando por ejemplo el plasma de un sujeto. También se podrían llevar a cabo exploración prenatal, diagnóstico, investigación o supervisión de un feto deduciendo el perfil de metilación de ADN fetal del plasma materno. Para ilustrar la potencia del enfoque, se ha mostrado que información que convencionalmente se ha obtenido mediante el estudio de tejidos placentarios podría evaluarse directamente a partir de plasma materno. Por ejemplo, el estado de impronta de locus génicos, identificación de locus con metilación diferencial entre el ADN fetal y materno y la variación gestacional en el perfil de metilación de locus génicos se consiguieron mediante el análisis directo de ADN de plasma materno. La principal ventaja del enfoque de los inventores es que el metiloma fetal podría evaluarse exhaustivamente durante el embarazo sin alteración del embarazo o la necesidad de tomar muestras de tejidos fetales de forma invasiva. Dada la asociación conocida entre el estado de metilación de ADN alterado y las muchas afecciones asociadas al embarazo, el enfoque descrito en este estudio puede actuar como una herramienta importante para investigar la patofisiología y la identificación de biomarcadores de esas afecciones. Centrándose en los locus con impronta, se ha mostrado que los perfiles de metilación fetales tanto transmitidos por vía paterna como transmitidos por vía materna podrían evaluarse a partir del plasma materno. Este enfoque puede ser potencialmente útil para la investigación de enfermedades con impronta. Las realizaciones también pueden aplicarse directamente para la evaluación prenatal de enfermedades asociadas al embarazo o fetales.

Se ha demostrado que la secuenciación por bisulfito de todo el genoma puede aplicarse para investigar el perfil de metilación de ADN de tejidos placentarios. Hay aproximadamente 28 M sitios de CpG en el genoma humano (C Clark *et al.*, 2012 PLoS One; 7: e50233). Los datos de secuenciación por bisulfito de la CVS y muestra de tejido placentario a término abarcaron más del 80 % de los CpG. Esto representa una cobertura sustancialmente más amplia que la que puede conseguirse usando otras plataformas de alto rendimiento. Por ejemplo, la matriz de beadchip de 27K de Metilación Humana Illumina Infinium que se usó en un estudio previo en tejidos placentarios (T Chu *et al.*, 2011 PLoS One; 6: e14723) solamente abarcó 0,1 % de los CpG en el genoma. La matriz de beadchip de 450K de Metilación Humana Illumina Infinium que estuvo disponible más recientemente solamente abarcó 1,7 % de los CpG (C Clark *et al.*, 2012 PLoS One; 7: e50233). Debido a que el enfoque de SPM está libre de restricciones relacionadas con el diseño de sondas, la eficacia de hibridación o la fuerza de captura del anticuerpo, podrían evaluarse CpG dentro de o más allá de las islas de CpG y en la mayoría de los contextos de secuencia.

35 XIV. SISTEMA INFORMÁTICO

Cualquiera de los sistemas informáticos mencionados en el presente documento puede utilizar cualquier número adecuado de subsistemas. Se muestran ejemplos de dichos subsistemas en la FIGURA 33 en el aparato informático 3300. En algunas realizaciones, un sistema informático incluye un único aparato informático, en el que los subsistemas pueden ser los componentes del aparato informático. En otras realizaciones, un sistema informático puede incluir múltiples aparatos informáticos, siendo cada uno un subsistema, con componentes internos.

Los subsistemas mostrados en la FIGURA 33 están interconectados mediante un bus de sistema 3375. Se muestran subsistemas adicionales tales como una impresora 3374, teclado 3378, dispositivo o dispositivos de almacenamiento 3379, monitor 3376, que está acoplado a un adaptador de presentación 3382, y otros. Pueden conectarse periféricos y dispositivos de entrada/salida (I/O), que se acoplan al controlador de I/O 3371, al sistema informático por cualquiera de varios medios conocidos en la técnica, tales como puerto en serie 3377. Por ejemplo, el puerto en serie 3377 o la interfaz externa 3381 (por ejemplo, Ethernet, Wi-Fi, etc.) pueden usarse para conectar el sistema informático 3300 a una red de área amplia tal como Internet, un dispositivo de entrada de tipo ratón o un escáner. La interconexión mediante bus de sistema 3375 permite que el procesador central 3373 se comunique con cada subsistema y controle la ejecución de instrucciones de la memoria del sistema 3372 o el dispositivo o los dispositivos de almacenamiento 3379 (por ejemplo, un disco fijo), así como el intercambio de información entre subsistemas. La memoria del sistema 3372 y/o el dispositivo o los dispositivos de almacenamiento 3379 pueden incorporar un medio leible por ordenador. Cualquiera de los valores mencionados en el presente documento puede extraerse en un componente a otro componente y puede presentarse al usuario.

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

Debería entenderse que cualquiera de las realizaciones de la presente divulgación puede implementarse en forma de lógica de control usando hardware (por ejemplo, un circuito integrado específico de aplicación o matriz de puertas programable *in situ*) y/o usando software informático con un procesador programable en general de una manera

modular o integrada. Como se usa en el presente documento, un procesador incluye un procesador multinúcleo en la misma microplaca integrada, o múltiples unidades de procesamiento en una única placa de circuito o en red. Basándose en la divulgación y enseñanzas proporcionadas en el presente documento, un experto habitual en la materia conocerá y apreciará otros modos y/o métodos para implementar realizaciones de la presente divulgación usando hardware y una combinación de hardware y software.

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

Dichos programas también pueden codificarse y transmitirse usando señales transportadoras adaptadas para transmisión mediante redes alámbricas, ópticas y/o inalámbricas que se adaptan a una diversidad de protocolos, incluyendo Internet. Como tal, un medio leíble por ordenador de acuerdo con una realización de la presente divulgación puede crearse usando una señal de datos codificada con dichos programas. Los medios leíbles por ordenador codificados con el código de programa pueden empaquetarse con un dispositivo compatible o proporcionarse por separado de otros dispositivos (por ejemplo, mediante descarga por Internet). Cualquiera de dichos medios leíbles por ordenador puede residir en o dentro de un único producto de programa informático (por ejemplo, un disco duro, un CD o un sistema informático completo) y puede estar presente en o dentro de diferentes productos de programas informáticos dentro de un sistema o una red. Un sistema informático puede incluir un monitor, una impresora u otro sistema de representación adecuado para proporcionar cualquiera de los resultados mencionados en el presente documento a un usuario.

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 lo 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 respectiva o un grupo respectivo de etapas. Aunque se presentan como etapas numeradas, las etapas de métodos del presente documento pueden realizarse en el mismo momento o en un orden diferente. Adicionalmente, partes de estas etapas pueden usarse con partes de otras etapas de otros métodos. Además, toda o partes de una etapa 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 realizaciones particulares pueden combinarse de cualquier manera adecuada sin alejarse del alcance de realizaciones de la divulgación. Sin embargo, otras realizaciones de la divulgación pueden dirigirse a realizaciones específicas relacionadas con cada aspecto individual o combinaciones específicas de estos aspectos individuales.

La descripción anterior de realizaciones ejemplares de la divulgación se ha presentado para los fines de ilustración y descripción. No se pretende que sean exhaustivas o que limiten la invención a la forma precisa descrita, y muchas modificaciones y variaciones son posibles a la luz de la enseñanza anterior. Las realizaciones se han elegido y descrito para explicar mejor los principios de la divulgación y sus aplicaciones prácticas para permitir de este modo que otros expertos en la materia utilicen mejor la divulgación en diversas realizaciones y con diversas modificaciones según sean adecuadas para el uso particular contemplado.

Se pretende que una cita de "un" o "el" signifique "uno o más" a no ser que se indique específicamente lo contrario.

Tabla S2A. Lista de las 100 regiones más hipermetiladas identificadas de la muestra de velloidad coriónica del primer trimestre y células sanguíneas maternas.

Cromosoma	Inicio	Final	Tamaño (pb)	Células sanguíneas maternas	CVS	Valores de P	Diferencia de Metilación
cr13	113063600	113064100	500	0,009	0,9	3,67E-15	0,891
cr6	36279700	36280200	500	0,0068	0,8957	2,39E-22	0,8889
cr16	66876000	66876500	500	0,0327	0,9211	3,82E-21	0,8884
cr10	163500	164000	500	0,0195	0,9034	3,60E-35	0,8839
cr9	3518300	3518800	500	0,0263	0,9045	1,32E-26	0,8782
cr12	31877100	31877600	500	0,007	0,8784	3,08E-22	0,8714
cr22	37477400	37478400	1000	0,0152	0,8848	0,00E+00	0,8696
cr4	148940500	148941000	500	0,0055	0,8717	4,40E-29	0,8662
cr5	131836300	131836800	500	0,075	0,9403	1,54E-10	0,8653
cr17	26661700	26663600	1900	0,0187	0,875	2,95E-38	0,8563
cr2	105758600	105759600	1000	0,0305	0,8828	1,19E-53	0,8523
cr22	39188800	39189800	1000	0	0,8514	2,05E-46	0,8514
cr3	153443900	153444900	1000	0,0436	0,8945	5,43E-34	0,8509
cr6	25149600	25150600	1000	0,0135	0,8632	0,00E+00	0,8497
cr5	98296800	98297300	500	0,0432	0,8925	4,97E-23	0,8493
cr7	150679900	150680400	500	0,0496	0,8944	6,50E-17	0,8448
cr7	107563100	107563600	500	0,0495	0,8895	9,58E-26	0,84
cr7	37348300	37349300	1000	0,0012	0,8409	0,00E+00	0,8397
cr14	58837800	58838300	500	0,0097	0,848	3,35E-16	0,8383
cr6	119238100	119238600	500	0,0899	0,928	2,38E-19	0,8381
cr15	93669900	93670400	500	0,0753	0,913	2,19E-10	0,8377
cr17	26669200	26670200	1000	0,0221	0,859	1,44E-29	0,8369
cr2	88108100	88108600	500	0,075	0,9109	3,55E-17	0,8359
cr13	98363800	98364300	500	0,11	0,9457	1,28E-11	0,8357
cr16	66948000	66948500	500	0,0331	0,8685	0,00E+00	0,8354
cr6	42098000	42098500	500	0,0484	0,8835	3,73E-16	0,8351
cr3	129876000	129876500	500	0,0565	0,8897	8,81E-17	0,8332
cr3	142700300	142700800	500	0,0063	0,8393	2,59E-22	0,833
cr8	145883800	145884300	500	0,0392	0,872	0,00E+00	0,8328
cr10	8320700	8321200	500	0,0566	0,8871	9,40E-09	0,8305
cr3	120438100	120438600	500	0,102	0,9292	7,09E-16	0,8272
cr3	173792600	173793100	500	0,0182	0,8453	2,84E-39	0,8271
cr17	40320700	40321200	500	0,0539	0,8788	6,50E-30	0,8249
cr15	72076200	72076700	500	0,0299	0,8525	6,48E-10	0,8226
cr16	29663900	29665400	1500	0,0081	0,8305	0,00E+00	0,8224
cr11	66961100	66962100	1000	0,0489	0,8712	0,00E+00	0,8223
cr9	27083100	27084100	1000	0,097	0,9177	2,37E-55	0,8207
cr9	111249600	111250100	500	0,0613	0,8795	1,99E-20	0,8182
cr14	101412400	101412900	500	0	0,8167	8,26E-32	0,8167
cr1	242549200	242549700	500	0	0,8155	3,50E-21	0,8155
cr8	38642800	38643300	500	0,0191	0,8346	3,22E-41	0,8155
cr4	85893600	85894100	500	0,0394	0,8533	1,45E-15	0,8139
cr5	142368600	142369100	500	0,0385	0,8523	1,18E-18	0,8138
cr8	130969500	130970000	500	0,069	0,8824	2,42E-24	0,8134
cr2	196783900	196784400	500	0	0,8123	3,63E-40	0,8123
cr16	49258100	49258600	500	0,0733	0,8851	4,29E-18	0,8118
cr1	232601200	232601700	500	0,0594	0,8707	1,73E-13	0,8113
cr1	109039500	109040000	500	0,0366	0,8471	1,07E-11	0,8105
cr17	59491300	59491800	500	0,0662	0,8758	2,15E-17	0,8096
cr21	42194100	42194600	500	0,11	0,9182	1,61E-12	0,8082
cr9	116174500	116175500	1000	0,0062	0,8132	1,98E-60	0,807
cr15	73429200	73429700	500	0	0,8066	9,81E-33	0,8066
cr6	157462800	157463300	500	0,0758	0,8819	7,94E-16	0,8061
cr3	16858500	16859500	1000	0,0021	0,8068	4,76E-68	0,8047
cr9	96662800	96663300	500	0,0614	0,8651	6,79E-28	0,8037
cr9	88143000	88143500	500	0,1538	0,9559	7,43E-09	0,8021
cr19	16090000	16091000	1000	0,0899	0,8904	1,60E-53	0,8005
cr15	29436300	29437300	1000	0,0553	0,8556	1,18E-80	0,8003
cr11	77816100	77816600	500	0,1069	0,9068	2,31E-17	0,7999

Cromosoma	Inicio	Final	(continuación)		CVS	Valores de P	Diferencia de Metilación
			Tamaño (pb)	Células sanguíneas maternas			
cr10	30346800	30347300	500	0,1212	0,9211	7,48E-07	0,7999
cr1	89510300	89511300	1000	0,0203	0,8191	3,53E-77	0,7988
cr3	125986100	125986600	500	0,1686	0,9674	5,24E-22	0,7988
cr19	60162800	60163300	500	0,0127	0,8113	9,99E-19	0,7986
cr16	73655900	73656900	1000	0,082	0,8806	4,48E-41	0,7986
cr16	30104300	30105800	1500	0,0298	0,8282	0,00E+00	0,7984
cr10	118642400	118642900	500	0,0588	0,8571	8,63E-11	0,7983
cr16	4495000	4496000	1000	0,0632	0,8615	2,27E-44	0,7983
cr1	2048300	2048800	500	0,0309	0,8289	1,19E-80	0,798
cr2	136481800	136482800	1000	0,0554	0,8533	8,50E-48	0,7979
cr10	29959200	29959700	500	0,1429	0,94	2,60E-08	0,7971
cr6	139642400	139642900	500	0,0618	0,8585	2,16E-29	0,7967
cr14	69825300	69825800	500	0,0654	0,8615	6,85E-14	0,7961
cr8	49739700	49740200	500	0,0324	0,828	2,88E-30	0,7956
cr17	42205700	42206200	500	0,057	0,852	2,11E-30	0,795
cr4	77445300	77445800	500	0,0442	0,8377	1,79E-35	0,7935
cr17	53762700	53766300	3600	0,0003	0,7926	0,00E+00	0,7923
cr17	44269900	44270400	500	0,026	0,8182	3,49E-21	0,7922
cr6	42462700	42463200	500	0,0761	0,8678	4,74E-22	0,7917
cr2	23396200	23396700	500	0,0333	0,8235	1,25E-14	0,7902
cr9	100921100	100921600	500	0,0244	0,814	3,32E-21	0,7896
cr7	74016100	74016600	500	0,1442	0,9333	6,74E-10	0,7891
cr6	157879000	157879500	500	0,133	0,9219	6,36E-17	0,7889
cr3	3189400	3190400	1000	0,0693	0,8571	1,38E-24	0,7878
cr16	29581500	29584500	3000	0,0081	0,7956	0,00E+00	0,7875
cr17	42201800	42202800	1000	0,0884	0,8751	0,00E+00	0,7867
cr11	94257000	94257500	500	0,1122	0,8986	4,29E-10	0,7864
cr10	14741600	14742100	500	0,0139	0,8	1,73E-20	0,7861
cr21	33826900	33827400	500	0,0879	0,8739	2,81E-11	0,786
cr4	130057200	130057700	500	0,0893	0,875	1,76E-13	0,7857
cr21	35343400	35343900	500	0	0,7853	7,43E-18	0,7853
cr12	105372800	105373300	500	0,0923	0,8767	8,67E-22	0,7844
cr5	10799800	10800300	500	0,1429	0,9263	8,21E-17	0,7834
cr5	16753100	16753600	500	0,041	0,8241	1,40E-15	0,7831
cr3	135746000	135746500	500	0,1429	0,9259	2,86E-09	0,783
cr6	53708300	53708800	500	0,0412	0,8235	2,74E-31	0,7823
cr2	128122900	128123400	500	0,0634	0,8455	4,82E-21	0,7821
cr5	150574200	150574700	500	0,0876	0,8696	1,56E-20	0,782
cr16	84326000	84327000	1000	0,1071	0,8891	3,58E-61	0,782
cr1	26744500	26745500	1000	0,0336	0,8152	0,00E+00	0,7816
cr2	234882000	234882500	500	0,0392	0,819	7,63E-14	0,7798

Tabla S2B. Lista de las 100 regiones más hipometiladas identificadas de muestra de vello de corión del primer trimestre y células sanguíneas maternas

Cromosoma	Inicio	Final	Tamaño (pb)	Células sanguíneas maternas	CVS	Valores de P	Diferencia de Metilación
cr18	12217500	12218500	1000	0,9873	0	3,05E-25	0,9873
cr17	22885400	22885900	500	0,9714	0,0161	8,92E-12	0,9553
cr3	184827100	184827600	500	0,9875	0,033	4,79E-16	0,9545
cr5	148968300	148968800	500	0,98	0,0426	6,70E-09	0,9374
cr10	104794500	104795000	500	0,973	0,0385	9,33E-10	0,9345
cr4	84977900	84978400	500	0,9643	0,0417	2,98E-08	0,9226
cr3	180395300	180395800	500	0,9877	0,0667	6,72E-08	0,921
cr2	138908300	138908800	500	0,939	0,0208	1,10E-16	0,9182
cr6	139873100	139873600	500	0,9667	0,0526	1,29E-07	0,914
cr8	59604700	59605200	500	0,9468	0,033	2,88E-14	0,9138
cr6	167622300	167622800	500	0,9452	0,0316	3,86E-14	0,9136
cr3	175701300	175701800	500	0,9846	0,0735	7,43E-10	0,9111
cr13	59246400	59246900	500	0,9402	0,0313	2,31E-11	0,9089
cr12	71263600	71264100	500	0,9296	0,0213	1,08E-08	0,9083
cr5	39459400	39459900	500	0,9219	0,014	5,01E-22	0,9079

Cromosoma	Inicio	Final	Tamaño (pb)	(continuación) Células sanguíneas maternas	CVS	Valores de P	Diferencia de Metilación
cr17	24904700	24905200	500	0,9161	0,0092	5,04E-35	0,9069
cr12	31889900	31890400	500	0,9524	0,0465	6,78E-13	0,9059
cr3	152897800	152898300	500	0,9402	0,0345	1,70E-17	0,9057
cr1	40378700	40379200	500	0,9565	0,0526	3,31E-09	0,9039
cr12	43979300	43979800	500	0,952	0,05	6,68E-13	0,902
cr18	1395900	1397400	1500	0,9308	0,0293	0,00E+00	0,9015
cr1	223482900	223483400	500	0,9579	0,0575	3,36E-24	0,9004
cr9	130357000	130357500	500	0,9282	0,0286	9,19E-13	0,8996
cr3	72878300	72878800	500	0,9612	0,0625	8,20E-14	0,8987
cr7	84347200	84348700	1500	0,9401	0,0418	0,00E+00	0,8983
cr15	37317500	37318000	500	0,9358	0,0385	7,58E-14	0,8973
cr8	42528600	42529100	500	0,9302	0,0337	1,73E-14	0,8965
cr6	134914000	134914500	500	0,9037	0,0076	4,84E-21	0,8961
cr13	56207100	56208100	1000	0,9184	0,0245	0,00E+00	0,894
cr2	209074000	209074500	500	0,9309	0,037	6,13E-27	0,8938
cr12	74021100	74022100	1000	0,9513	0,058	0,00E+00	0,8933
cr4	118939300	118939800	500	0,9192	0,0276	6,58E-27	0,8916
cr5	12626600	12628600	2000	0,9266	0,0355	0,00E+00	0,8911
cr5	105517300	105518300	1000	0,927	0,0359	0,00E+00	0,891
cr12	70056300	70057300	1000	0,9488	0,0609	0,00E+00	0,888
cr6	153238200	153239200	1000	0,9123	0,0244	0,00E+00	0,8879
cr17	60374800	60375300	500	0,9655	0,0777	3,64E-14	0,8878
cr14	68272700	68273200	500	0,9389	0,0523	1,23E-22	0,8866
cr19	54533800	54534800	1000	0,9117	0,0262	0,00E+00	0,8855
cr12	15392200	15393200	1000	0,9307	0,0457	0,00E+00	0,885
cr1	212517400	212517900	500	0,9266	0,0417	9,81E-12	0,8849
cr10	49344400	49345400	1000	0,9422	0,0579	0,00E+00	0,8844
cr3	47410400	47410900	500	0,9213	0,0381	5,59E-16	0,8832
cr3	879500	880000	500	0,9455	0,0625	8,06E-06	0,883
cr2	31572400	31573400	1000	0,9176	0,0357	0,00E+00	0,8819
cr1	89131200	89131700	500	0,9314	0,0498	5,15E-70	0,8816
cr8	94832000	94832500	500	0,9156	0,0351	2,16E-65	0,8805
cr7	14008300	14009800	1500	0,9349	0,0545	0,00E+00	0,8804
cr12	12971300	12972300	1000	0,9361	0,0559	0,00E+00	0,8802
cr5	43114700	43115200	500	0,9638	0,0842	1,79E-13	0,8796
cr11	107872400	107872900	500	0,9472	0,0677	2,31E-32	0,8794
cr8	49757600	49758100	500	0,9048	0,0269	3,15E-52	0,8779
cr13	33106400	33106900	500	0,9384	0,0606	9,54E-15	0,8778
cr3	190658800	190659300	500	0,9388	0,0617	2,71E-22	0,877
cr1	181508000	181508500	500	0,9259	0,0495	3,78E-15	0,8764
cr1	180436900	180437400	500	0,9412	0,0652	2,36E-13	0,876
cr6	122642800	122643800	1000	0,9218	0,0458	0,00E+00	0,8759
cr5	166429300	166429800	500	0,9551	0,08	5,26E-05	0,8751
cr12	14972900	14973400	500	0,9483	0,0733	2,10E-18	0,8749
cr5	123933900	123934400	500	0,943	0,0683	1,12E-39	0,8746
cr2	15969400	15970400	1000	0,8939	0,0196	9,43E-46	0,8743
cr3	167635200	167636200	1000	0,9363	0,0625	9,03E-41	0,8738
cr5	159442700	159443200	500	0,9174	0,044	6,27E-14	0,8734
cr4	48027200	48027700	500	0,9839	0,1111	8,89E-06	0,8728
cr6	140071500	140072000	500	0,9234	0,0506	4,39E-33	0,8728
cr10	22356300	22356800	500	0,9548	0,0822	1,04E-18	0,8726
cr8	61007300	61007800	500	0,9197	0,0476	1,24E-15	0,8721
cr11	95463500	95464000	500	0,9348	0,0629	1,16E-20	0,8718
cr2	216399800	216400300	500	0,938	0,0667	5,98E-06	0,8713
cr18	57359700	57360200	500	0,9293	0,0584	1,77E-19	0,871
cr3	102734400	102734900	500	0,8917	0,0207	7,94E-22	0,871
cr1	173605700	173606200	500	0,96	0,0891	5,46E-13	0,8709
cr2	86993700	86995700	2000	0,8965	0,0261	0,00E+00	0,8704
cr3	162621100	162621600	500	0,9226	0,0526	7,89E-38	0,8699
cr12	10144800	10145300	500	0,929	0,0598	3,45E-17	0,8691
cr3	113855100	113855600	500	0,9667	0,0982	3,97E-14	0,8685
cr2	156958200	156959200	1000	0,9252	0,0571	8,89E-50	0,8681
cr2	55775000	55776000	1000	0,9159	0,0483	0,00E+00	0,8676

(continuación)							
Cromosoma	Inicio	Final	Tamaño (pb)	Células sanguíneas maternas	CVS	Valores de P	Diferencia de Metilación
cr6	124898400	124898900	500	0,8987	0,0313	1,91E-15	0,8675
cr5	42003700	42004700	1000	0,9262	0,0588	0,00E+00	0,8674
cr3	24162200	24162700	500	0,883	0,0161	1,75E-27	0,8668
cr6	35394000	35395000	1000	0,9204	0,0539	0,00E+00	0,8665
cr17	8451800	8453300	1500	0,9376	0,0714	0,00E+00	0,8662
cr14	53487700	53488700	1000	0,9013	0,0353	0,00E+00	0,866
cr7	98572800	98573300	500	0,9651	0,0995	8,37E-26	0,8656
cr6	52298700	52299200	500	0,9427	0,0772	1,31E-28	0,8655
cr6	159047900	159048400	500	0,908	0,0426	3,34E-08	0,8655
cr14	22152600	22153100	500	0,9085	0,0435	4,43E-17	0,865
cr12	103285000	103285500	500	0,9321	0,0674	0,00E+00	0,8647
cr7	43302200	43302700	500	0,968	0,1037	6,40E-16	0,8643
cr14	22247400	22247900	500	0,9804	0,1163	3,83E-06	0,8641
cr2	66780900	66781400	500	0,9355	0,0714	8,37E-09	0,8641
cr12	97393000	97393500	500	0,9045	0,0408	3,46E-21	0,8637
cr5	162797900	162798900	1000	0,9271	0,0635	1,75E-57	0,8636
cr2	83598400	83599400	1000	0,9354	0,0719	0,00E+00	0,8635
cr11	111358800	111359300	500	0,9156	0,0523	4,15E-24	0,8632
cr11	104891100	104892600	1500	0,9164	0,0533	0,00E+00	0,863
cr1	184583600	184584100	500	0,9647	0,1026	3,02E-13	0,8621
cr5	132350500	132351500	1000	0,9042	0,0426	1,86E-36	0,8616
cr5	53268300	53268800	500	0,972	0,1111	4,76E-16	0,8609

Tabla S2C. Lista de las 100 regiones más hipermetiladas identificadas de tejido placentario de tercer trimestre y células sanguíneas maternas

Cromosoma	Inicio	Final	Tamaño (pb)	Células sanguíneas maternas	Placenta a término	Valores de P	Diferencia de Metilación
cr4	78129700	78130200	500	0,0488	0,9747	3,97E-33	0,926
cr5	131467400	131467900	500	0,0213	0,9275	7,10E-27	0,9063
cr17	26661700	26663600	1900	0,0187	0,9226	1,79E-41	0,9039
cr4	148940500	148941000	500	0,0055	0,9079	1,82E-29	0,9024
cr9	100921100	100921600	500	0,0244	0,9242	1,38E-25	0,8998
cr6	137114200	137114700	500	0	0,8934	8,87E-14	0,8934
cr3	173792600	173793100	500	0,0182	0,9091	1,70E-42	0,8908
cr5	98296800	98297300	500	0,0432	0,9333	2,58E-23	0,8901
cr12	44898000	44898500	500	0	0,8889	4,47E-11	0,8889
cr3	197328900	197329400	500	0,0169	0,9048	5,55E-10	0,8878
cr8	49739700	49740200	500	0,0324	0,9194	5,71E-34	0,887
cr12	122279300	122279800	500	0,0135	0,8969	3,46E-21	0,8834
cr17	43092200	43092700	500	0	0,8824	4,34E-10	0,8824
cr7	107563100	107563600	500	0,0495	0,931	1,05E-28	0,8815
cr11	72543200	72543700	500	0,0377	0,9167	2,94E-09	0,8789
cr14	58837800	58838300	500	0,0097	0,886	9,16E-18	0,8763
cr3	153443900	153444900	1000	0,0436	0,9197	6,24E-39	0,876
cr3	16953200	16953700	500	0,0896	0,9655	6,78E-09	0,876
cr17	42205700	42206200	500	0,057	0,933	1,13E-31	0,8759
cr6	53217600	53218100	500	0,0818	0,9571	1,54E-19	0,8754
cr3	112749000	112749500	500	0,0403	0,9154	4,11E-22	0,8752
cr8	22453700	22454200	500	0,003	0,8765	1,64E-50	0,8735
cr1	162860900	162861400	500	0,023	0,8932	8,37E-14	0,8702
cr6	36279700	36280200	500	0,0068	0,8762	1,14E-21	0,8694
cr5	80962500	80963000	500	0	0,8679	2,08E-15	0,8679
cr16	11312500	11313000	500	0	0,8679	2,14E-10	0,8679
cr16	29663900	29665400	1500	0,0081	0,8759	0,00E+00	0,8678
cr3	120438100	120438600	500	0,102	0,9639	2,98E-15	0,8618
cr8	134157000	134157500	500	0,0625	0,9219	6,10E-20	0,8594
cr6	42620900	42621400	500	0	0,8571	5,68E-08	0,8571
cr5	131836300	131836800	500	0,075	0,9315	1,26E-10	0,8565
cr14	60290000	60290500	500	0	0,8544	5,63E-14	0,8544
cr6	42850300	42851300	1000	0,0676	0,9211	2,38E-24	0,8534

Cromosoma	Inicio	Final	Tamaño (pb)	(continuación) Células sanguíneas maternas	Placenta a término	Valores de P	Diferencia de Metilación
cr8	28974100	28974600	500	0,0394	0,8927	2,03E-51	0,8533
cr22	22368500	22369000	500	0,0248	0,8778	1,18E-70	0,8529
cr14	69825300	69825800	500	0,0654	0,9174	2,73E-14	0,852
cr3	142700300	142700800	500	0,0063	0,8582	2,56E-23	0,8519
cr17	59491300	59491800	500	0,0662	0,9175	1,81E-16	0,8513
cr15	30881700	30882200	500	0,0493	0,8995	2,38E-26	0,8502
cr15	91496300	91496800	500	0	0,85	3,13E-17	0,85
cr17	18745300	18745800	500	0,0294	0,8775	3,47E-51	0,848
cr15	29436500	29437000	500	0,0336	0,8811	2,62E-66	0,8476
cr2	217795300	217795800	500	0	0,8472	1,78E-22	0,8472
cr11	16328100	16328600	500	0,0278	0,875	3,43E-11	0,8472
cr13	113063500	113064000	500	0,0102	0,8571	1,82E-15	0,8469
cr5	40472400	40472900	500	0,0197	0,8664	7,54E-35	0,8467
cr1	242549200	242549700	500	0	0,8462	8,53E-23	0,8462
cr11	58099100	58099600	500	0,0162	0,8612	4,45E-35	0,845
cr9	16020400	16020900	500	0,0132	0,8555	8,05E-23	0,8423
cr8	37550700	37551200	500	0,0093	0,8512	1,11E-16	0,8419
cr5	75722400	75722900	500	0,1215	0,9627	5,97E-23	0,8411
cr19	60454700	60455200	500	0,0316	0,8722	2,44E-62	0,8405
cr4	99587100	99587600	500	0,0128	0,8526	1,49E-12	0,8398
cr6	25149600	25150600	1000	0,0135	0,8514	0,00E+00	0,8379
cr1	32065200	32065700	500	0	0,8371	1,09E-44	0,8371
cr7	5337200	5337700	500	0,0727	0,9098	2,18E-14	0,8371
cr17	44269900	44270400	500	0,026	0,8621	3,94E-22	0,8361
cr1	36180800	36181300	500	0,0714	0,9067	1,23E-09	0,8352
cr18	10472700	10473700	1000	0,0713	0,9064	9,35E-70	0,8351
cr5	350000	350500	500	0,0297	0,8643	1,49E-16	0,8346
cr2	136481800	136482800	1000	0,0554	0,8887	1,87E-52	0,8332
cr4	89241100	89241600	500	0,1091	0,9423	1,05E-12	0,8332
cr17	40320700	40321200	500	0,0539	0,8859	6,94E-31	0,832
cr7	133897200	133897700	500	0,0769	0,9077	1,64E-24	0,8308
cr8	98060600	98061100	500	0,0741	0,9048	3,11E-07	0,8307
cr8	134141500	134142000	500	0	0,829	2,77E-58	0,829
cr14	80250600	80251100	500	0,0839	0,9122	2,05E-18	0,8283
cr2	100730900	100731400	500	0,0787	0,9067	4,85E-11	0,828
cr2	88108100	88108600	500	0,075	0,901	2,10E-16	0,826
cr19	16338500	16339500	1000	0,0011	0,8259	0,00E+00	0,8247
cr5	141791900	141792900	1000	0,0225	0,8467	0,00E+00	0,8243
cr11	116227400	116227900	500	0	0,8242	1,01E-17	0,8242
cr22	48705500	48706000	500	0,0649	0,8891	4,00E-76	0,8242
cr9	3518300	3518800	500	0,0263	0,8493	6,66E-25	0,823
cr11	16791000	16791500	500	0,1095	0,9322	1,05E-22	0,8228
cr3	135746000	135746500	500	0,1429	0,9651	7,63E-10	0,8223
cr1	19323400	19323900	500	0,0411	0,8624	2,36E-20	0,8213
cr9	96662800	96663300	500	0,0614	0,8826	1,07E-28	0,8212
cr7	37348300	37349300	1000	0,0012	0,821	0,00E+00	0,8198
cr2	234882000	234882500	500	0,0392	0,8591	5,36E-15	0,8198
cr6	44694000	44694500	500	0,1024	0,9222	5,68E-19	0,8198
cr17	18320500	18321000	500	0	0,8197	2,78E-39	0,8197
cr22	28992000	28994000	2000	0,0012	0,8195	0,00E+00	0,8183
cr17	53762700	53766300	3600	0,0003	0,8179	0,00E+00	0,8176
cr1	114215500	114216000	500	0	0,8169	2,24E-20	0,8169
cr6	13381700	13382700	1000	0,0037	0,8206	1,03E-40	0,8169
cr5	17045400	17045900	500	0,0235	0,84	3,24E-13	0,8165
cr12	110924300	110924800	500	0,0855	0,9016	1,01E-18	0,816
cr1	200499800	200500300	500	0,011	0,8269	9,73E-24	0,8159
cr4	8311000	8311500	500	0,053	0,8687	7,82E-18	0,8157
cr8	6535300	6535800	500	0,0667	0,8824	5,25E-09	0,8157
cr6	42462700	42463200	500	0,0761	0,8919	5,78E-23	0,8157
cr1	91969900	91970400	500	0,0172	0,8325	5,23E-18	0,8152
cr2	105758600	105759600	1000	0,0305	0,8455	5,15E-52	0,815
cr21	37538500	37539000	500	0,1595	0,9745	8,81E-16	0,815

Cromosoma	Inicio	Final	Tamaño (pb)	(continuación) Células sanguíneas maternas	Placenta a término	Valores de P	Diferencia de Metilación
cr9	92953000	92953500	500	0,0189	0,8333	1,88E-15	0,8145
cr16	30104400	30105900	1500	0,0505	0,8636	0,00E+00	0,8131
cr1	234184400	234185400	1000	0,0346	0,8477	9,66E-31	0,813
cr8	19116400	19116900	500	0	0,8125	9,33E-11	0,8125
cr4	141194300	141195300	1000	0,0865	0,899	5,46E-30	0,8125

Tabla S2D. Lista de las 100 regiones más hipometiladas identificadas de tejido placentario de tercer trimestre y células sanguíneas maternas.

Cromosoma	Inicio	Final	Tamaño (pb)	Células sanguíneas maternas	Placenta a término	valores de P	Diferencia de Metilación
cr9	40380300	40380800	500	0,9667	0	1,13E-06	0,9667
cr1	31769200	31769700	500	0,9548	0,0256	5,57E-25	0,9291
cr18	12217600	12218100	500	0,9873	0,0602	1,63E-19	0,9271
cr20	19704400	19704900	500	0,9426	0,018	4,34E-18	0,9246
cr15	37317500	37318000	500	0,9358	0,0132	1,90E-25	0,9226
crX	83368400	83368900	500	0,913	0	3,15E-07	0,913
cr11	27549100	27549600	500	0,9224	0,0123	3,92E-24	0,9101
cr18	58141500	58142000	500	0,9737	0,0645	1,07E-09	0,9092
cr1	159897000	159897500	500	0,9067	0	2,53E-16	0,9067
cr7	84347200	84348700	1500	0,9401	0,0407	0,00E+00	0,8994
cr2	216916100	216916600	500	0,9695	0,0714	2,13E-16	0,8981
cr7	144200000	144200500	500	0,9294	0,0317	1,24E-10	0,8977
cr1	241331600	241332100	500	0,9198	0,0227	0,00E+00	0,8971
cr7	123190000	123191000	1000	0,9341	0,0384	0,00E+00	0,8957
cr5	12626600	12628600	2000	0,9266	0,0321	0,00E+00	0,8944
cr12	12971300	12972300	1000	0,9361	0,0438	0,00E+00	0,8923
cr22	20936500	20937000	500	0,9528	0,0606	1,87E-06	0,8922
cr13	31321900	31322400	500	0,9231	0,0313	1,43E-06	0,8918
cr22	21701500	21702000	500	0,9579	0,0667	1,30E-09	0,8912
cr10	104794400	104794900	500	1	0,1111	6,10E-09	0,8889
cr7	21835800	21836300	500	0,9156	0,0267	3,85E-13	0,8889
cr1	16134800	16135300	500	0,95	0,0635	4,79E-10	0,8865
cr3	47410400	47410900	500	0,9213	0,0357	6,63E-17	0,8855
cr10	49344400	49345400	1000	0,9422	0,0571	0,00E+00	0,8851
cr2	209073900	209074400	500	0,9196	0,0353	1,63E-22	0,8843
cr1	89131200	89131700	500	0,9314	0,0472	1,05E-75	0,8842
cr3	167118500	167119500	1000	0,9365	0,0527	0,00E+00	0,8838
cr18	1395900	1397400	1500	0,9308	0,0472	0,00E+00	0,8836
cr2	59670300	59670800	500	0,9433	0,0599	5,09E-23	0,8834
cr14	28368900	28369400	500	0,9446	0,0619	5,03E-64	0,8827
cr3	126028800	126029300	500	0,9379	0,0556	1,83E-20	0,8823
cr9	69378900	69379900	1000	0,8816	0	6,02E-51	0,8816
cr5	105517300	105518300	1000	0,927	0,0461	0,00E+00	0,8808
cr2	31572400	31573400	1000	0,9176	0,037	0,00E+00	0,8806
cr5	42003700	42004700	1000	0,9262	0,0462	0,00E+00	0,88
cr14	94718300	94718800	500	0,9548	0,0764	6,67E-19	0,8784
cr19	56417800	56418300	500	0,9615	0,0833	1,71E-06	0,8782
cr2	70183000	70183500	500	0,9694	0,0914	9,49E-39	0,878
cr4	118939300	118939800	500	0,9192	0,0412	2,20E-34	0,878
cr13	59246400	59246900	500	0,9402	0,0633	5,40E-12	0,8769
cr12	74021100	74022100	1000	0,9513	0,0752	0,00E+00	0,8761
cr2	173432500	173433000	500	0,9529	0,0778	5,39E-12	0,8752
cr16	24004400	24004900	500	0,9239	0,0488	3,25E-23	0,8751
cr13	27596300	27597300	1000	0,9538	0,0795	0,00E+00	0,8743
cr15	88904300	88904800	500	0,9212	0,0481	7,69E-27	0,8731
cr18	12720200	12721200	1000	0,9346	0,0618	0,00E+00	0,8728
cr15	60975900	60976900	1000	0,9311	0,0587	0,00E+00	0,8724
cr21	39630100	39631100	1000	0,9423	0,07	4,68E-43	0,8723
cr5	123933900	123934400	500	0,943	0,0707	2,60E-38	0,8722
cr8	77382600	77383600	1000	0,9117	0,0395	0,00E+00	0,8722

Cromosoma	Inicio	Final	Tamaño (pb)	(continuación) Células sanguíneas maternas	Placenta a término	valores de P	Diferencia de Metilación
cr21	32238800	32239300	500	0,9396	0,0677	1,28E-18	0,8719
cr5	175019600	175020100	500	0,9542	0,0828	4,34E-20	0,8714
cr8	134437400	134438400	1000	0,9083	0,037	4,79E-29	0,8713
cr5	69668800	69669300	500	0,9194	0,0492	2,88E-09	0,8702
cr1	60877900	60878900	1000	0,9378	0,068	0,00E+00	0,8698
cr16	80650400	80650900	500	0,9309	0,0611	2,49E-32	0,8698
cr18	59388800	59389300	500	0,9706	0,1008	4,84E-15	0,8697
cr2	15969400	15970400	1000	0,8939	0,0244	3,07E-52	0,8695
cr13	56207100	56208100	1000	0,9184	0,0505	0,00E+00	0,868
cr3	180395300	180395800	500	0,9877	0,12	1,73E-09	0,8677
cr6	153238200	153239200	1000	0,9123	0,0452	0,00E+00	0,8671
cr18	61635100	61635600	500	0,9268	0,06	2,77E-13	0,8668
cr3	177562200	177563200	1000	0,9121	0,0455	0,00E+00	0,8666
cr4	160368300	160370200	1900	0,9272	0,0606	0,00E+00	0,8665
cr6	144626900	144627400	500	0,9114	0,046	2,10E-12	0,8654
cr16	59885500	59886500	1000	0,9407	0,0757	1,12E-62	0,865
cr1	55667100	55667600	500	0,9095	0,0446	8,62E-39	0,8649
cr2	83598300	83599300	1000	0,9366	0,0718	0,00E+00	0,8648
cr4	105135200	105136200	1000	0,913	0,0486	0,00E+00	0,8644
cr14	32048400	32048900	500	0,9142	0,0499	5,43E-53	0,8643
cr1	223482700	223483700	1000	0,9636	0,0997	2,69E-34	0,864
cr14	47487700	47488200	500	0,915	0,0514	5,45E-33	0,8636
cr3	104515000	104515500	500	1	0,1373	1,08E-06	0,8627
cr7	14008300	14009800	1500	0,9349	0,0725	0,00E+00	0,8624
cr1	243134000	243135500	1500	0,9208	0,0588	0,00E+00	0,8619
cr10	14156400	14156900	500	0,9105	0,0489	0,00E+00	0,8616
cr2	118616200	118617200	1000	0,9178	0,0565	0,00E+00	0,8613
cr17	8455500	8456000	500	0,8941	0,0331	1,94E-18	0,8611
cr12	15392200	15393200	1000	0,9307	0,0697	0,00E+00	0,861
cr8	81275900	81276900	1000	0,9291	0,0684	0,00E+00	0,8606
cr1	234269300	234269800	500	0,9471	0,087	2,25E-25	0,8602
cr1	181970300	181970800	500	0,9167	0,0566	2,15E-08	0,8601
cr2	55775000	55776000	1000	0,9159	0,0559	0,00E+00	0,8599
cr3	88338000	88339000	1000	0,8909	0,0311	0,00E+00	0,8598
cr5	140078700	140079200	500	0,8852	0,0253	0,00E+00	0,8598
cr21	16720900	16721400	500	0,9317	0,0721	1,38E-15	0,8596
cr11	104891100	104892600	1500	0,9164	0,0569	0,00E+00	0,8595
cr1	184204700	184205200	500	0,9194	0,0603	8,16E-16	0,859
cr6	160732500	160733000	500	0,9191	0,0606	1,31E-10	0,8585
cr8	37134300	37134800	500	0,9151	0,0567	1,21E-26	0,8584
cr18	5869800	5870300	500	0,913	0,0548	1,21E-09	0,8582
cr1	98448100	98448600	500	0,9574	0,1	2,08E-05	0,8574
cr3	152897800	152898300	500	0,9402	0,0828	3,28E-18	0,8574
cr1	110304000	110304500	500	0,9783	0,121	2,60E-18	0,8572
cr2	86993600	86995600	2000	0,8965	0,0395	0,00E+00	0,857
cr19	15428100	15430600	2500	0,9424	0,0862	0,00E+00	0,8563
cr13	75176800	75177800	1000	0,9258	0,0697	3,09E-47	0,8561
cr13	24126700	24127200	500	0,9498	0,0938	1,57E-17	0,856
cr16	28238500	28240500	2000	0,9427	0,0868	0,00E+00	0,8559
cr2	158079500	158080500	1000	0,9199	0,0642	0,00E+00	0,8557

Tabla S3A. Lista de los 100 locus superiores que se ha deducido que están hipermetilados a partir de los datos de secuenciación por bisulfito de plasma materno de primer trimestre.

Cromosoma	Inicio	Final	Células sanguíneas maternas	CVS	Diferencia de Metilación
cr22	39189067	39189863	0	0,8444	0,8444
cr17	53763065	53764027	0	0,7922	0,7922
cr7	41887694	41888212	0	0,7614	0,7614
cr2	1,14E+08	1,14E+08	0	0,751	0,751
cr12	25096242	25097206	0	0,7098	0,7098
cr1	66574104	66574793	0	0,7025	0,7025

ES 2 969 567 T3

Cromosoma	Inicio	Final	(continuación)	CVS	Diferencia de Metilación
			Células sanguíneas maternas		
cr6	11489985	11490755	0	0,7004	0,7004
cr6	1,07E+08	1,07E+08	0	0,6978	0,6978
cr10	30858286	30858871	0	0,6693	0,6693
cr17	21131574	21132167	0	0,6496	0,6496
cr18	13454740	13455292	0	0,5468	0,5468
cr16	11298755	11299326	0	0,5373	0,5373
cr2	1,75E+08	1,75E+08	0	0,5196	0,5196
cr19	44060511	44061036	0	0,5128	0,5128
cr6	1,08E+08	1,08E+08	0	0,5	0,5
cr3	71261611	71262501	0	0,4587	0,4587
cr9	36247847	36248885	0	0,447	0,447
cr19	17819240	17820082	0	0,4279	0,4279
cr17	53769900	53770731	0	0,4102	0,4102
cr1	1,12E+08	1,12E+08	0,0002	0,6167	0,6166
cr7	1,34E+08	1,34E+08	0,0003	0,4351	0,4348
cr3	11658550	11659929	0,0004	0,4299	0,4295
cr17	53764417	53765963	0,0005	0,7967	0,7961
cr10	11246762	11249052	0,0005	0,4002	0,3997
cr22	28992647	28993434	0,0006	0,8092	0,8087
cr15	62460278	62461007	0,0006	0,4334	0,4328
cr1	31002038	31003474	0,0007	0,5926	0,5919
cr19	3129246	3132159	0,0008	0,7725	0,7717
cr12	1,21E+08	1,21E+08	0,0008	0,7303	0,7295
cr19	12304446	12305741	0,0009	0,6986	0,6978
cr3	67788734	67789395	0,001	0,9131	0,9121
cr9	1,32E+08	1,32E+08	0,001	0,7047	0,7037
cr19	6723370	6724479	0,001	0,689	0,688
cr3	1,84E+08	1,84E+08	0,001	0,4384	0,4374
cr2	53848089	53849214	0,001	0,4368	0,4358
cr17	59450886	59452113	0,0012	0,469	0,4678
cr5	1,72E+08	1,72E+08	0,0014	0,578	0,5766
cr21	35342527	35343373	0,0014	0,5392	0,5378
cr21	45164804	45165437	0,0015	0,4251	0,4236
crX	3742417	3744601	0,0016	0,4486	0,447
cr21	45158293	45159003	0,0017	0,7799	0,7782
cr7	39839340	39839876	0,0017	0,4074	0,4057
cr2	1,75E+08	1,75E+08	0,0018	0,4816	0,4797
cr12	1,24E+08	1,24E+08	0,0019	0,6306	0,6287
cr3	50352688	50353823	0,002	0,624	0,622
cr9	97264382	97265523	0,0021	0,5008	0,4987
cr7	64178628	64179354	0,0021	0,4088	0,4066
cr9	94767202	94767802	0,0023	0,7568	0,7544
cr5	42986308	42988304	0,0023	0,4882	0,4859
cr17	63854127	63854693	0,0024	0,8266	0,8242
cr12	1,22E+08	1,22E+08	0,0024	0,4869	0,4844
cr17	16260170	16260909	0,0026	0,6404	0,6378
cr4	39874787	39875456	0,0027	0,7233	0,7206
cr12	6441080	6441608	0,0027	0,6228	0,6201
cr19	45015653	45016886	0,0028	0,5444	0,5416
cr6	30757752	30758823	0,0028	0,4783	0,4755
cr6	41636176	41637112	0,0028	0,4254	0,4226
cr12	6315199	6315765	0,0029	0,4613	0,4584
cr14	76576283	76577070	0,0029	0,4365	0,4336
cr16	48857790	48858300	0,0031	0,5625	0,5594
cr5	1,7E+08	1,7E+08	0,0031	0,4752	0,4721
cr13	26897813	26898557	0,0032	0,4354	0,4322
cr14	52753948	52754571	0,0032	0,4221	0,4189
cr1	1,66E+08	1,66E+08	0,0033	0,5579	0,5545
cr12	56157424	56158348	0,0033	0,47	0,4667
cr22	16079971	16080532	0,0034	0,6226	0,6193
cr7	1946410	1946975	0,0036	0,6826	0,6789
cr11	258799	259749	0,0036	0,5072	0,5037
cr6	13381944	13382477	0,0037	0,5945	0,5908

Cromosoma	Inicio	Final	(continuación)	CVS	Diferencia de Metilación
			Células sanguíneas maternas		
cr7	1,27E+08	1,27E+08	0,0037	0,5096	0,5058
cr13	23743886	23744467	0,0037	0,4534	0,4497
cr2	1,21E+08	1,21E+08	0,0038	0,7175	0,7137
cr21	25855853	25857105	0,0039	0,4661	0,4622
cr2	43211724	43212565	0,0039	0,4345	0,4306
cr12	1,08E+08	1,08E+08	0,0041	0,6024	0,5983
cr15	92928924	92929575	0,0041	0,4074	0,4033
cr19	10731043	10731636	0,0042	0,5868	0,5826
cr6	1,45E+08	1,45E+08	0,0043	0,5783	0,574
cr1	52875323	52875907	0,0044	0,4145	0,4101
cr14	75058186	75058956	0,0045	0,602	0,5975
cr12	1,21E+08	1,21E+08	0,0045	0,4821	0,4776
cr17	76873737	76874417	0,0046	0,6012	0,5966
cr2	2,38E+08	2,38E+08	0,0049	0,7654	0,7604
cr2	1,98E+08	1,98E+08	0,0049	0,7228	0,7179
cr6	1,47E+08	1,47E+08	0,0049	0,4967	0,4918
cr9	1,36E+08	1,36E+08	0,0049	0,4584	0,4535
cr1	67545402	67546771	0,005	0,4971	0,4921
cr6	1,58E+08	1,58E+08	0,0052	0,6145	0,6093
cr3	1,7E+08	1,7E+08	0,0052	0,5845	0,5794
cr1	2,34E+08	2,34E+08	0,0053	0,7033	0,6979
cr10	80715722	80716751	0,0053	0,6515	0,6462
cr4	48602901	48603736	0,0053	0,6315	0,6262
cr19	13957965	13958580	0,0053	0,599	0,5937
cr1	90081114	90082367	0,0053	0,4574	0,4521
cr2	1,06E+08	1,06E+08	0,0054	0,8858	0,8804
cr16	29664213	29665369	0,0054	0,8339	0,8285
cr1	1,59E+08	1,59E+08	0,0054	0,7663	0,7608
cr13	97926489	97927025	0,0054	0,6229	0,6175
cr1	41604452	41605277	0,0054	0,6011	0,5956
cr9	1,28E+08	1,28E+08	0,0054	0,5871	0,5818

Tabla S3B. Lista de los 100 locus superiores que se ha deducido que están hipometilados a partir de los datos de secuenciación por bisulfito de plasma materno de primer trimestre.

Cromosoma	Inicio	Final	(continuación)	CVS	Diferencia de Metilación
			Células sanguíneas maternas		
cr1	235771917	235772426	0,9868	0,549	0,4377
cr1	97357972	97358622	0,9835	0,4805	0,503
cr1	4490516	4491074	0,9826	0,4793	0,5032
cr4	181124168	181124671	0,9825	0,4725	0,5099
cr16	71908694	71909213	0,982	0,5581	0,4239
cr3	182727915	182728477	0,981	0,3577	0,6233
cr5	115339535	115340038	0,9802	0,5455	0,4347
cr3	195855575	195856122	0,9801	0,3793	0,6008
cr6	155437621	155438161	0,9799	0,5991	0,3808
cr9	20468093	20468904	0,9798	0,4271	0,5527
cr10	90702298	90702987	0,9787	0,3324	0,6463
cr1	170581654	170582162	0,9785	0,4817	0,4968
cr3	108816849	108817794	0,9783	0,4793	0,4989
cr20	36912749	36913319	0,9783	0,5	0,4783
cr13	72517281	72517839	0,9782	0,4855	0,4927
cr12	103553001	103553677	0,9774	0,492	0,4854
cr22	27638905	27639408	0,9766	0,5385	0,4382
cr7	17290850	17291462	0,9763	0,59	0,3863
cr6	17227866	17228510	0,976	0,4058	0,5703
cr15	56998547	56999107	0,9754	0,3766	0,5988
cr7	70965945	70966842	0,9753	0,5893	0,386
cr3	32159338	32160065	0,9752	0,5379	0,4372
cr16	17043258	17043854	0,9752	0,5521	0,4231
cr16	22776223	22776850	0,9752	0,5735	0,4017
cr5	169344029	169344869	0,9751	0,4211	0,5541
cr11	34324955	34325722	0,975	0,5561	0,4189

ES 2 969 567 T3

Cromosoma	Inicio	Final	(continuación)	CVS	Diferencia de Metilación
			Células sanguíneas maternas		
cr8	58554745	58555376	0,9747	0,5784	0,3964
cr1	153933389	153934121	0,9746	0,463	0,5116
cr14	88003983	88004485	0,9745	0,5379	0,4366
cr3	151738501	151739120	0,9741	0,4901	0,484
cr14	105618699	105619606	0,974	0,3457	0,6283
cr16	24060085	24060702	0,9738	0,3991	0,5747
cr8	68941792	68942711	0,9738	0,5449	0,429
cr12	53208707	53209304	0,9737	0,4847	0,489
cr7	76892564	76893249	0,9736	0,5664	0,4072
cr3	69464294	69464971	0,9736	0,5893	0,3843
cr19	61401137	61401745	0,9732	0,4933	0,4799
cr11	124569867	124570490	0,9732	0,5136	0,4595
cr18	42618440	42619096	0,9732	0,5942	0,379
cr5	169398896	169399637	0,9731	0,498	0,4751
cr5	169328124	169328983	0,9731	0,572	0,401
cr20	34679880	34680448	0,9731	0,5922	0,3809
cr16	9042198	9042702	0,973	0,4286	0,5444
cr10	90205044	90205701	0,973	0,4407	0,5323
cr13	33236454	33236997	0,973	0,5906	0,3824
cr16	73284579	73285087	0,9729	0,5602	0,4127
cr8	29100691	29101428	0,9728	0,505	0,4678
cr2	202383851	202384447	0,9727	0,5461	0,4267
cr3	179501620	179502300	0,9722	0,5766	0,3956
cr6	107674976	107675906	0,9719	0,4434	0,5285
cr6	107880632	107881161	0,9718	0,5623	0,4095
cr12	56350283	56350933	0,9718	0,5909	0,3809
cr19	40636458	40637339	0,9717	0,4941	0,4776
cr2	223472599	223473287	0,9714	0,1824	0,7891
cr22	20709067	20709787	0,9714	0,5149	0,4565
cr19	46095583	46096190	0,9713	0,5385	0,4328
cr6	90258338	90259318	0,9712	0,3415	0,6297
cr2	54598347	54598933	0,9712	0,5894	0,3819
cr3	114810453	114811493	0,9711	0,5166	0,4545
cr19	15851125	15851654	0,9711	0,5236	0,4476
cr8	42889138	42890084	0,9711	0,5652	0,4059
cr18	52354390	52355064	0,971	0,598	0,373
cr15	38206236	38207010	0,9709	0,4186	0,5523
cr7	99700554	99701110	0,9708	0,305	0,6658
cr12	19487336	19487855	0,9708	0,4105	0,5603
cr7	87996908	87997437	0,9708	0,5462	0,4246
cr6	63628653	63629378	0,9707	0,529	0,4417
cr15	38209108	38209618	0,9706	0,5882	0,3824
cr19	6623769	6624450	0,9704	0,5179	0,4526
cr2	10794513	10795242	0,9704	0,5976	0,3728
cr2	118472785	118474454	0,9704	0,5992	0,3712
cr5	57820209	57820801	0,9701	0,5815	0,3886
cr10	100183380	100184702	0,9701	0,5826	0,3875
cr2	8151989	8152646	0,97	0,4701	0,4999
cr10	3938374	3938914	0,9699	0,1741	0,7958
cr9	123724524	123725439	0,9697	0,57	0,3997
cr14	89085469	89086097	0,9696	0,3278	0,6418
cr16	14129437	14130133	0,9695	0,5304	0,4392
cr5	60746367	60747191	0,9695	0,5571	0,4124
cr1	92002953	92003729	0,9694	0,52	0,4494
cr6	31264677	31265413	0,9693	0,5135	0,4558
cr7	99317013	99318281	0,9692	0,5117	0,4574
cr8	8808867	8809422	0,9692	0,5691	0,4002
cr19	20052165	20052720	0,969	0,2792	0,6898
cr8	129139026	129139573	0,969	0,3458	0,6232
cr11	122314929	122315458	0,969	0,4232	0,5458
cr13	98377663	98378165	0,9688	0,3319	0,6369
cr9	107606194	107606872	0,9688	0,449	0,5198
cr8	56096904	56097736	0,9688	0,5267	0,4422

Cromosoma	Inicio	Final	(continuación)	CVS	Diferencia de Metilación
			Células sanguíneas maternas		
cr7	128093836	128094339	0,9688	0,5929	0,3758
cr2	103109370	103109916	0,9686	0,3333	0,6352
cr3	101803534	101804063	0,9686	0,5027	0,4659
cr10	69505720	69506278	0,9684	0,2515	0,7169
cr13	26608225	26608754	0,9683	0,3614	0,6069
cr1	90993315	90993828	0,9683	0,5519	0,4164
cr6	11361243	11361801	0,9681	0,2578	0,7103
cr21	36529300	36529981	0,968	0,1944	0,7736
cr21	37813953	37814521	0,9679	0,2175	0,7505
cr2	15226273	15227211	0,9679	0,5134	0,4545
cr19	4102809	4103443	0,9679	0,5646	0,4034

Tabla S3C. Lista de los 100 locus superiores que se ha deducido que están hipermetilados a partir de los datos de secuenciación por bisulfito de plasma materno del tercer trimestre.

Cromosoma	Inicio	Final	(continuación)	Placenta a término	Diferencia de Metilación
			Células sanguíneas maternas		
cr17	53763065	53764027	0,0000	0,8680	0,8680
cr22	39189067	39189863	0,0000	0,8233	0,8233
cr10	30858286	30858871	0,0000	0,7713	0,7713
cr7	41887694	41888212	0,0000	0,7578	0,7578
cr2	1,14E+08	1,14E+08	0,0000	0,7500	0,7500
cr12	25096242	25097206	0,0000	0,7332	0,7332
cr6	1,07E+08	1,07E+08	0,0000	0,7229	0,7229
cr1	66574104	66574793	0,0000	0,7136	0,7136
cr16	11298755	11299326	0,0000	0,7005	0,7005
cr6	11489985	11490755	0,0000	0,6935	0,6935
cr18	13454740	13455292	0,0000	0,6594	0,6594
cr6	1,08E+08	1,08E+08	0,0000	0,6231	0,6231
crX	3627885	3628549	0,0000	0,6133	0,6133
cr12	7979754	7980413	0,0000	0,6118	0,6118
cr3	71261611	71262501	0,0000	0,5938	0,5938
cr17	53769900	53770731	0,0000	0,5586	0,5586
cr11	1,18E+08	1,18E+08	0,0000	0,5558	0,5558
cr19	44060511	44061036	0,0000	0,5464	0,5464
cr2	2,38E+08	2,38E+08	0,0000	0,5330	0,5330
cr1	1,91E+08	1,91E+08	0,0000	0,5294	0,5294
cr1	1,44E+08	1,44E+08	0,0000	0,4857	0,4857
cr2	1,75E+08	1,75E+08	0,0000	0,4785	0,4785
cr4	15366889	15367646	0,0000	0,4729	0,4729
cr2	19537237	19537737	0,0000	0,4599	0,4599
cr1	1,15E+08	1,15E+08	0,0000	0,4351	0,4351
cr1	1,54E+08	1,54E+08	0,0000	0,4299	0,4299
cr14	51383387	51384149	0,0000	0,4186	0,4186
cr1	1,12E+08	1,12E+08	0,0002	0,5350	0,5348
cr3	11658550	11659929	0,0004	0,5579	0,5575
cr17	53764417	53765963	0,0005	0,7894	0,7889
cr22	28992647	28993434	0,0006	0,8053	0,8047
cr6	27214981	27215823	0,0006	0,4593	0,4587
cr1	31002038	31003474	0,0007	0,6309	0,6302
cr12	1,12E+08	1,12E+08	0,0008	0,7360	0,7352
cr19	3129246	3132159	0,0008	0,7257	0,7249
cr19	12304446	12305741	0,0009	0,6397	0,6388
cr6	28723918	28724965	0,0009	0,4344	0,4335
cr19	6723370	6724479	0,0010	0,7280	0,7270
cr2	53848089	53849214	0,0010	0,4060	0,4050
cr9	1,32E+08	1,32E+08	0,0010	0,7558	0,7548
cr3	67788734	67789395	0,0010	0,9219	0,9209
cr19	18276368	18277132	0,0011	0,5136	0,5125
cr17	59450886	59452113	0,0012	0,5196	0,5184
cr17	74243852	74244670	0,0012	0,4117	0,4105
cr3	1,85E+08	1,85E+08	0,0014	0,4961	0,4948

Cromosoma	Inicio	Final	(continuación)	Placenta a término	Diferencia de Metilación
			Células sanguíneas maternas		
cr21	35342527	35343373	0,0014	0,5126	0,5112
cr5	1,72E+08	1,72E+08	0,0014	0,6531	0,6516
cr21	45164804	45165437	0,0015	0,4364	0,4349
crX	3742417	3744601	0,0016	0,7517	0,7501
cr21	45158293	45159003	0,0017	0,8180	0,8163
cr2	1,75E+08	1,75E+08	0,0018	0,6214	0,6196
cr12	1,24E+08	1,24E+08	0,0019	0,5906	0,5887
cr3	50352688	50353823	0,0020	0,6082	0,6062
cr9	94767202	94767802	0,0023	0,8327	0,8304
cr17	63854127	63854693	0,0024	0,7886	0,7862
cr12	1,22E+08	1,22E+08	0,0024	0,5021	0,4997
cr17	16260170	16260909	0,0026	0,5780	0,5754
cr12	6441080	6441608	0,0027	0,7471	0,7444
cr4	39874787	39875456	0,0027	0,7962	0,7935
cr18	50536032	50536649	0,0027	0,4920	0,4893
cr6	30757752	30758823	0,0028	0,4029	0,4001
cr19	2571763	2572292	0,0031	0,4200	0,4169
cr5	1,7E+08	1,7E+08	0,0031	0,4218	0,4187
cr13	26897813	26898557	0,0032	0,6485	0,6453
cr12	56157424	56158348	0,0033	0,5541	0,5508
cr1	1,66E+08	1,66E+08	0,0033	0,5147	0,5113
cr22	16079971	16080532	0,0034	0,6265	0,6231
cr6	16820551	16821134	0,0035	0,4800	0,4765
cr11	258799	259749	0,0036	0,5475	0,5439
cr7	1946410	1946975	0,0036	0,8251	0,8215
cr6	13381944	13382477	0,0037	0,8221	0,8183
cr7	1,27E+08	1,27E+08	0,0037	0,4767	0,4730
cr2	1,21E+08	1,21E+08	0,0038	0,6734	0,6697
cr2	43211724	43212565	0,0039	0,4256	0,4217
cr15	92928924	92929575	0,0041	0,5605	0,5564
cr12	1,08E+08	1,08E+08	0,0041	0,7313	0,7271
cr19	10731043	10731636	0,0042	0,5668	0,5626
cr6	1,45E+08	1,45E+08	0,0043	0,5910	0,5867
cr1	52875323	52875907	0,0044	0,6115	0,6071
cr12	1,21E+08	1,21E+08	0,0045	0,5884	0,5839
cr14	75058186	75058956	0,0045	0,6534	0,6489
cr17	76873737	76874417	0,0046	0,5658	0,5612
cr6	1,47E+08	1,47E+08	0,0049	0,5826	0,5777
cr2	1,98E+08	1,98E+08	0,0049	0,7944	0,7895
cr2	2,38E+08	2,38E+08	0,0049	0,7328	0,7278
cr8	1,42E+08	1,42E+08	0,0050	0,7728	0,7679
cr3	1,7E+08	1,7E+08	0,0052	0,7227	0,7176
cr6	1,58E+08	1,58E+08	0,0052	0,6389	0,6337
cr2	2,38E+08	2,38E+08	0,0052	0,4238	0,4185
cr1	56948854	56949496	0,0053	0,4484	0,4431
cr4	48602901	48603736	0,0053	0,5920	0,5867
cr5	1,31E+08	1,31E+08	0,0053	0,4858	0,4805
cr10	80715722	80716751	0,0053	0,5249	0,5196
cr19	13957965	13958580	0,0053	0,4379	0,4326
cr1	2,34E+08	2,34E+08	0,0053	0,8440	0,8387
cr13	97926489	97927025	0,0054	0,7233	0,7179
cr9	1,28E+08	1,28E+08	0,0054	0,7312	0,7258
cr2	1,06E+08	1,06E+08	0,0054	0,8513	0,8459
cr2	96556705	96557637	0,0054	0,4095	0,4041
cr16	29664213	29665369	0,0054	0,8837	0,8783

Tabla S3D. Lista de los 100 locus superiores que se ha deducido que están hipometilados a partir de los datos maternos.

Cromosoma	Inicio	Final	Células sanguíneas maternas	Placenta a término	Diferencia de Metilación
cr10	7548948	7549483	0,9866	0,5685	0,4181
cr1	4490516	4491074	0,9826	0,5015	0,4810
cr4	1,81E+08	1,81E+08	0,9825	0,5981	0,3843
cr3	1,83E+08	1,83E+08	0,9810	0,2925	0,6886
cr3	1,96E+08	1,96E+08	0,9801	0,4643	0,5158
cr6	1,55E+08	1,55E+08	0,9799	0,4610	0,5189
cr1	1,71E+08	1,71E+08	0,9785	0,5122	0,4662
cr20	36912749	36913319	0,9783	0,4513	0,5269
cr22	38583100	38583616	0,9783	0,5428	0,4355
cr1	19391314	19392207	0,9778	0,5273	0,4505
cr5	1,74E+08	1,74E+08	0,9770	0,5852	0,3918
cr19	13678906	13679531	0,9760	0,5812	0,3949
cr14	83650790	83651395	0,9760	0,5378	0,4382
cr15	56998547	56999107	0,9754	0,4691	0,5063
cr16	22776223	22776850	0,9752	0,5114	0,4638
cr5	1,69E+08	1,69E+08	0,9751	0,4809	0,4943
cr8	58554745	58555376	0,9747	0,5977	0,3770
cr14	1,06E+08	1,06E+08	0,9740	0,2069	0,7671
cr8	68941792	68942711	0,9738	0,5872	0,3866
cr16	24060085	24060702	0,9738	0,3470	0,6268
cr12	53208707	53209304	0,9737	0,5278	0,4459
cr5	1,69E+08	1,69E+08	0,9731	0,5057	0,4673
cr16	9042198	9042702	0,9730	0,1860	0,7869
cr10	90205044	90205701	0,9730	0,5922	0,3808
cr3	1,89E+08	1,89E+08	0,9720	0,4949	0,4771
cr6	1,08E+08	1,08E+08	0,9719	0,5825	0,3894
cr2	2,23E+08	2,23E+08	0,9714	0,3333	0,6381
cr19	46095583	46096190	0,9713	0,5065	0,4648
cr8	1,41E+08	1,41E+08	0,9713	0,5753	0,3959
cr6	90258338	90259318	0,9712	0,4357	0,5355
cr13	51403556	51404069	0,9710	0,3980	0,5731
cr18	66875048	66875726	0,9710	0,5259	0,4451
cr7	99700554	99701110	0,9708	0,3757	0,5951
cr7	87996908	87997437	0,9708	0,5720	0,3988
cr19	6623769	6624450	0,9704	0,4774	0,4930
cr1	97639047	97639749	0,9701	0,4148	0,5553
cr16	23892096	23892772	0,9701	0,5000	0,4701
cr10	3938374	3938914	0,9699	0,1148	0,8551
cr14	89085469	89086097	0,9696	0,2964	0,6732
cr8	1,29E+08	1,29E+08	0,9690	0,3565	0,6126
cr13	98377663	98378165	0,9688	0,3123	0,6566
cr8	56096904	56097736	0,9688	0,4562	0,5126
cr2	1,03E+08	1,03E+08	0,9686	0,3459	0,6227
cr13	26608225	26608754	0,9683	0,4562	0,5121
cr2	22738157	22738760	0,9682	0,5122	0,4560
cr6	11361243	11361801	0,9681	0,2646	0,7035
cr21	36529300	36529981	0,9680	0,1829	0,7852
cr21	37813953	37814521	0,9679	0,3061	0,6619

ES 2 969 567 T3

(continuación)

Cromosoma	Inicio	Final	Células sanguíneas maternas	Placenta a término	Diferencia de Metilación
cr2	2,43E+08	2,43E+08	0,9679	0,5750	0,3929
cr4	12413543	12414103	0,9679	0,5944	0,3735
cr3	1,27E+08	1,27E+08	0,9677	0,4030	0,5648
cr7	33509047	33509556	0,9676	0,4627	0,5048
cr14	59284846	59285553	0,9674	0,5254	0,4420
cr17	42623453	42624024	0,9673	0,4318	0,5355
cr19	6778363	6779377	0,9671	0,4416	0,5255
cr4	41798250	41798788	0,9670	0,5000	0,4670
cr5	88054080	88054588	0,9669	0,2238	0,7431
cr16	24109379	24110289	0,9669	0,5062	0,4607
cr10	13847159	13847895	0,9667	0,3188	0,6479
cr10	1,27E+08	1,27E+08	0,9667	0,5423	0,4244
cr12	1,12E+08	1,12E+08	0,9663	0,3722	0,5941
cr10	17220886	17221845	0,9662	0,4455	0,5207
cr8	5947355	5947862	0,9662	0,5171	0,4491
cr3	73740840	73741439	0,9659	0,3657	0,6002
cr14	57945953	57946875	0,9658	0,5357	0,4301
cr14	50905777	50906333	0,9658	0,3008	0,6650
cr15	90275374	90276000	0,9657	0,5409	0,4248
cr22	24717299	24718197	0,9657	0,5160	0,4497
cr7	36530128	36530987	0,9656	0,5194	0,4462
cr2	1,31E+08	1,31E+08	0,9655	0,4384	0,5271
cr4	42116988	42117788	0,9654	0,5195	0,4459
cr12	1,16E+08	1,16E+08	0,9653	0,5594	0,4059
cr2	7491785	7492736	0,9652	0,4556	0,5097
cr19	6599638	6600187	0,9652	0,5488	0,4163
cr6	25326803	25327398	0,9651	0,3974	0,5677
cr4	1,7E+08	1,7E+08	0,9651	0,4933	0,4718
cr7	99875338	99876155	0,9650	0,2696	0,6953
cr14	97144328	97145208	0,9649	0,5377	0,4272
cr3	11718596	11719163	0,9649	0,5521	0,4128
cr14	1E+08	1E+08	0,9649	0,3794	0,5855
cr7	1,5E+08	1,5E+08	0,9648	0,3327	0,6322
cr12	56357827	56358328	0,9648	0,4217	0,5430
cr10	8275750	8276276	0,9647	0,3100	0,6547
cr11	16999685	17000209	0,9647	0,2765	0,6882
cr22	34419356	34419861	0,9646	0,4245	0,5401
cr18	72453151	72453725	0,9646	0,4700	0,4946
cr5	49919879	49920699	0,9645	0,3169	0,6476
cr1	24580891	24581805	0,9643	0,3565	0,6078
cr22	18233774	18234492	0,9641	0,5205	0,4436
cr14	45356178	45356903	0,9640	0,3934	0,5706
cr3	53007193	53008661	0,9638	0,4902	0,4737
cr4	55027912	55028539	0,9637	0,5254	0,4384
cr5	1,37E+08	1,37E+08	0,9637	0,5290	0,4347
cr1	2,23E+08	2,23E+08	0,9636	0,0997	0,8640
cr7	1,35E+08	1,35E+08	0,9636	0,2959	0,6677
cr5	80350438	80351169	0,9636	0,4969	0,4667
cr12	31889600	31890343	0,9636	0,1745	0,7891
cr12	8365395	8366096	0,9636	0,5721	0,3914

(continuación)					
Cromosoma	Inicio	Final	Células sanguíneas maternas	Placenta a término	Diferencia de Metilación
cr19	15424819	15425355	0,9635	0,2836	0,6799
cr10	10985469	10986409	0,9635	0,4877	0,4759

REIVINDICACIONES

1. Un método para analizar una muestra de plasma de un organismo humano, comprendiendo la muestra de plasma ADN sin células que se origina de células normales y posiblemente de células asociadas con cáncer, comprendiendo el método:
 - analizar una pluralidad de moléculas de ADN sin células de la muestra de plasma, en donde analizar una molécula de ADN sin células incluye:
 - determinar una localización de la molécula de ADN sin células en un genoma del organismo humano mapeando, mediante lecturas de secuencia, la molécula de ADN sin células en una parte de un genoma humano; y
 - determinar si la molécula de ADN sin células está metilada en uno o más sitios CpG de una pluralidad de sitios CpG organizados en una pluralidad de islas CpG, incluyendo cada isla CpG más de un sitio CpG;
 - determinar un número respectivo de moléculas de ADN sin células en cada una de la pluralidad de islas CpG que están metiladas;
 - calcular un primer nivel de metilación basado en los números respectivos de moléculas de ADN sin células que están metiladas en la pluralidad de islas CpG;
 - comparar el primer nivel de metilación con un primer valor de corte indicativo de la presencia de cáncer; y
 - determinar una primera clasificación de la presencia de cáncer basada en la comparación.
2. El método de la reivindicación 1, en donde determinar si la molécula de ADN sin células está metilada en uno o más sitios CpG comprende realizar una secuenciación con reconocimiento de metilación.
3. El método de la reivindicación 2, en donde realizar la secuenciación con reconocimiento de metilación comprende realizar una secuenciación paralela masiva con reconocimiento de metilación.
4. El método de la reivindicación 2 o 3, en donde realizar la secuenciación con reconocimiento de metilación genera entre 39 y 142 millones de lecturas.
5. El método de una cualquiera de las reivindicaciones anteriores, en donde realizar la secuenciación con reconocimiento de metilación incluye:
 - tratar las moléculas de ADN sin células con bisulfito sódico; y
 - realizar la secuenciación de las moléculas de ADN sin células tratadas.
6. El método de la reivindicación 5, en donde tratar las moléculas de ADN sin células con bisulfito sódico forma parte de la secuenciación por bisulfito oxidativa o conversión con bisulfito asistida por Tet para la detección de 5-hidroximetilcitosina.
7. El método de una cualquiera de las reivindicaciones anteriores, en donde realizar la secuenciación con reconocimiento de metilación comprende la secuenciación de al menos 60.000 moléculas de ADN sin células.
8. El método de la reivindicación 1, en donde el primer valor de corte es una distancia específica de un nivel de metilación de referencia establecido a partir de una muestra biológica obtenida de un organismo sano.
9. El método de la reivindicación 8, en donde la distancia específica es un número específico de desviaciones típicas a partir del nivel de metilación de referencia.
10. El método de la reivindicación 1, en donde el primer valor de corte se establece a partir de un nivel de metilación de referencia determinado a partir de una muestra biológica previa del organismo humano obtenida antes de analizar la muestra de plasma.
11. El método de la reivindicación 1, en donde comparar el primer nivel de metilación con el primer valor de corte incluye:
 - determinar una diferencia entre el primer nivel de metilación y un nivel de metilación de referencia; y
 - comparar la diferencia con un umbral correspondiente al primer valor de corte.
12. El método de una cualquiera de las reivindicaciones 1 a 11, en donde la pluralidad de sitios se encuentra en una pluralidad de cromosomas.
13. El método de una cualquiera de las reivindicaciones 1 a 11, en donde se determina que la pluralidad de sitios son islas CpG hipermetiladas en comparación con un grupo de referencia.

14. El método de una cualquiera de las reivindicaciones 1 a 13, en donde para un perfil de metilación se calcula el primer nivel de metilación.

5 15. El método de una cualquiera de las reivindicaciones 1 a 13, en donde el análisis de la pluralidad de moléculas de ADN sin células comprende analizar al menos 10 millones de lecturas de secuencia.

16. Un producto de programa informático que comprende una pluralidad de instrucciones con capacidad de ejecución por un sistema informático, que cuando se ejecutan de este modo controlan el sistema informático para realizar el método de una cualquiera de las reivindicaciones 1 a 15.

10

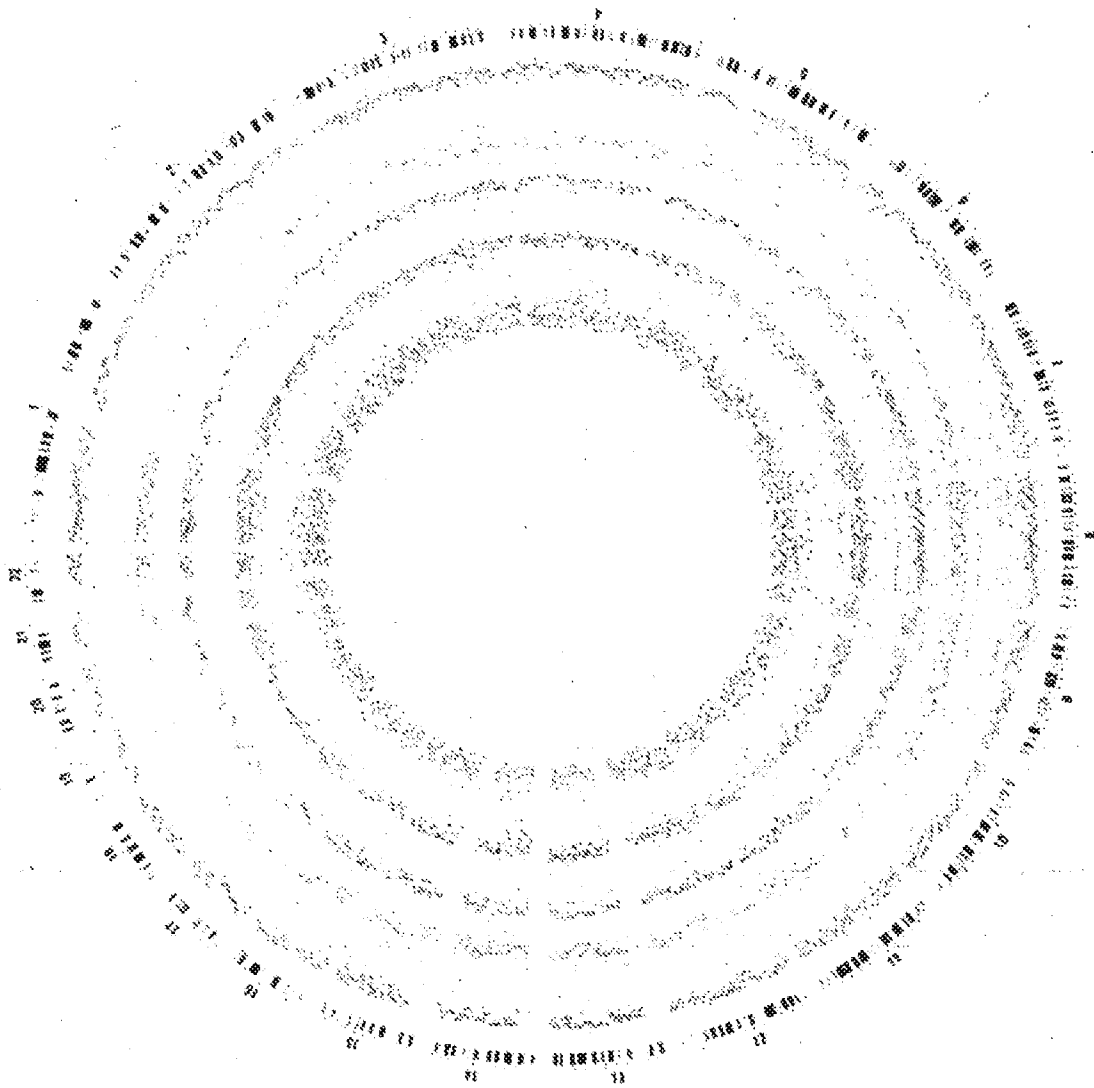
FIG. 1A

100

	Sangre materna (1er trimestre)		Cvs		Placenta a término		Plasma materno (1er trimestre)		Plasma materno (3er trimestre)		Plasma materno (pos-parto)	
	watson	crick	watson	crick	watson	crick	watson	crick	watson	crick	watson	crick
Recuentos de lecturas sin procesar	547522712	547522712	422873853	422873853	837637130	837637170	4367226854	4567226854	1289337674	1289337674	1277615263	1277615263
Capacidad de mapeo total (%)	35.88	44.35	43.83	43.83	46.70	44.40	29.82	29.48	35.81	35.39	37.96	37.49
Lecturas mapeables totales	246808492	243023086	186980568	165345026	287735873	283959215	1367896504	1346252254	481639866	456245835	443637803	477778018
Tasa ambigua (%) ^a	6.72	6.83	6.89	6.89	6.26	5.54	6.35	6.43	5.86	5.53	5.96	6.05
Tasa de duplicación (%) ^b	6.60	5.86	9.52	9.27	32.94	30.39	64.66	64.41	25.70	25.49	41.81	40.56
Profundidad de genoma completo después de la retirada de la tasa de conversión ambigua y lambda (%) ^c	11.29	11.45	8.36	8.31	9.43	9.25	25.49	25.29	16.24	16.85	14.28	14.00
Recuento de M de sec. C genómica ^d	268496805	265662565	186460378	186460378	226378356	223772855	648597052	635295175	455306420	433669856	397672185	381338159
Recuento de U de sec. C genómica	4484584136	4218629557	432782985	4261356735	4868446023	4793558224	1169385481	1164414410	8676357118	8538105685	6998711473	686335640
Cobertura de C genómica (%)	66.45	66.28	62.96	62.96	67.40	67.40	80.48	80.46	69.94	68.76	67.68	67.61
Profundidad de sec. de C genómica	8.11	8.07	7.57	7.51	8.89	8.78	21.12	21.01	15.45	15.39	12.48	12.43
Densidad de metilación de C genómica (%)	5.65	5.63	4.31	4.29	4.36	4.36	5.39	5.37	5.05	5.04	5.45	5.44
Recuento M de sec. de CpG	265303863	262504317	182423216	182423216	213375565	211326313	627868879	622753476	443226176	440725023	389810638	387475682
Recuento de U de sec. de CpG	164163857	163575011	145223817	145223817	147248361	146617702	316066675	308133584	206327402	205477045	142856625	142271154
Cobertura de CpG (%)	95.17	85.12	81.02	80.92	84.23	84.15	92.05	92.02	86.13	86.26	84.93	84.86
Profundidad de sec. de CpG	13.12	13.01	11.68	11.71	12.82	12.71	33.30	33.06	23.06	23.94	18.91	18.91
Densidad de metilación de CpG (%)	71.80	71.63	65.72	64.97	69.21	69.04	66.96	66.98	68.23	68.20	73.18	73.14
Recuento de M de sec. noCpG	3279043	3147347	2285715	2285715	12502145	12351436	13818674	12532920	12579169	12847842	7859875	7749841
Recuento de U de sec. noCpG	438239353	428505404	3379625287	3371152468	4560807486	4537934784	1135614810	1130854410	8307572669	8330832598	5749162946	5719112682
Cobertura noCpG (%)	0.13	0.13	0.12	0.13	0.13	0.13	0.16	0.16	0.14	0.14	0.13	0.13
Profundidad de sec. noCpG	1.52	1.52	1.41	1.41	1.82	1.81	4.02	4.00	2.96	2.96	2.38	2.38
Densidad de metilación noCpG (%)	0.08	0.07	0.16	0.16	0.27	0.27	5.11	5.11	0.15	0.15	0.12	0.12
Recuento de M de sec. de CHG	1825754	1800267	2160977	2071816	3223342	3176761	3766562	3691990	3625562	3692617	3347693	3314323
Recuento de U de sec. de CHG	1345439602	1341393086	1271982855	121669371	1367682765	1357568840	2416655003	2404225988	2519517702	2509439547	2036336055	2030358603
Cobertura de CHG (%)	71.90	71.82	73.26	73.13	77.59	77.38	88.16	88.12	79.68	79.59	77.84	77.75
Profundidad de sec. de CHG	10.89	10.85	9.85	9.85	11.04	11.03	27.65	27.56	20.40	20.32	16.48	16.42
Densidad de metilación de CHG (%)	0.08	0.07	0.17	0.17	0.24	0.24	0.11	0.11	0.14	0.14	0.12	0.11
Recuento de M de sec. de CHH	2249309	2146768	5163738	5181211	9278803	9124667	9052012	8941730	9353607	9251325	5512083	5426318
Recuento de U de sec. de CHH	2959584349	2943760945	2761843432	2754572597	3198719641	3180765944	7938542388	7903616496	5949615967	5821092051	4712844541	4691052079
Cobertura de CHH (%)	61.96	61.90	58.43	58.32	63.54	63.45	77.53	77.52	65.96	65.88	63.65	63.64
Profundidad de sec. de CHH	5.63	5.61	6.40	6.38	7.42	7.38	18.38	18.30	13.55	13.48	10.91	10.86
Densidad de metilación de CHH (%)	0.08	0.07	0.19	0.18	0.29	0.29	0.11	0.11	0.16	0.16	0.12	0.12

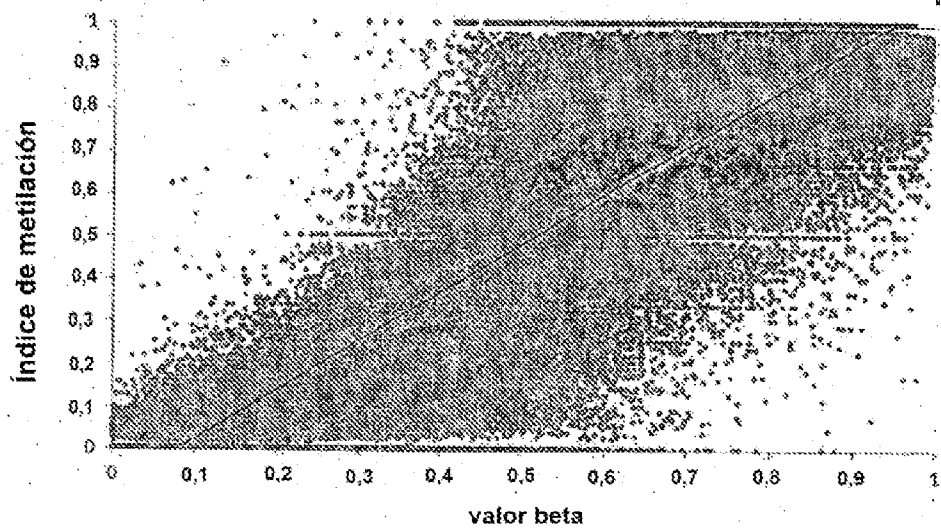
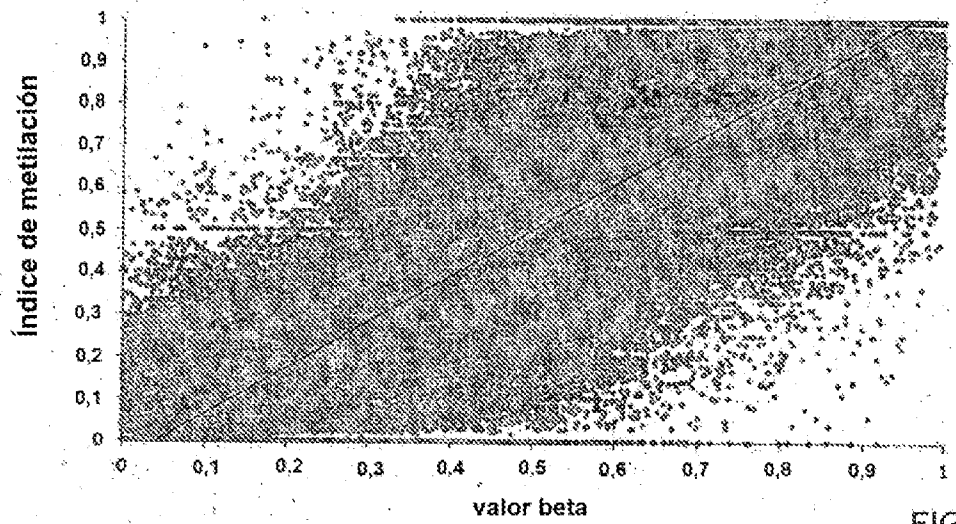
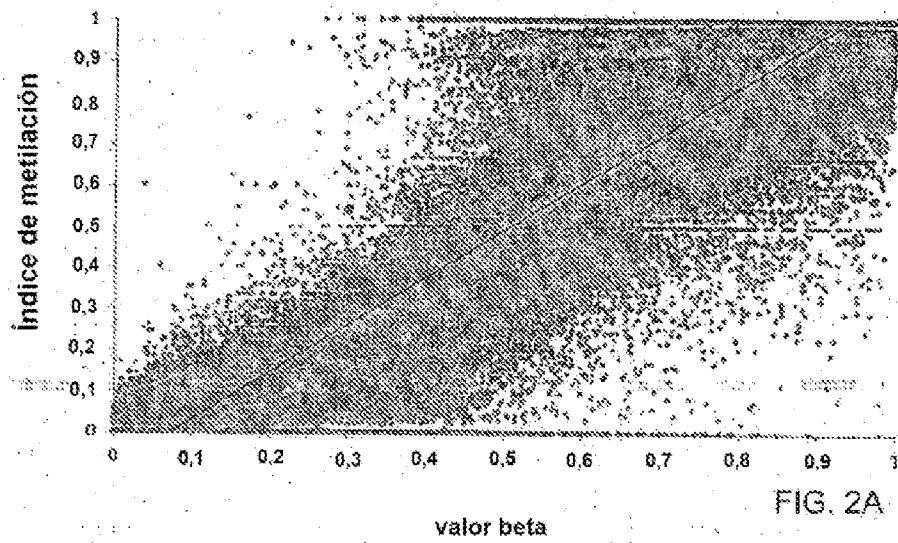
	Sangre de hombre adulto		Sangre de mujer adulta		Plasma de hombre adulto		Plasma de mujer adulta	
	Watson	Crick	Watson	Crick	Watson	Crick	Watson	Crick
Recuentos de lecturas sin procesar	195544792	196544792	197387755	197387755	179688600	179688600	141199439	141199439
Capacidad de mapeo total (%)	43,52	42,97	44,94	44,36	22,47	22,29	16,96	18,79
Lecturas mapeables totales	85526774	84462495	88699242	87542953	40300936	40048964	26795699	2637762
Tasa ambigua (%) ^a	6,05	6,13	6,29	6,37	6,09	6,14	6,29	6,35
Tasa de duplicación (%) ^b	1,99	1,73	1,68	1,53	7,89	7,64	11,71	11,47
Profundidad de genoma completo después de la retirada de la tasa de conversión ambigua y lambda (%) ^c	4,13	4,09	4,28	4,24	1,83	1,82	1,16	1,15
Recuento de M de sec. C genómica ^d	99,96	99,97	99,97	99,97	99,96	99,97	99,96	99,97
Recuento de U de sec. C genómica	127154246	125487793	124863462	122238508	53049489	54630306	37403326	37048460
Cobertura de C genómica (%)	2047971331	2026731827	2058153525	2039855644	938185429	933377753	598496391	594974253
Profundidad de sec. de C genómica	51,86	51,55	55,17	55,06	43,76	43,66	38,39	38,28
Densidad de metilación de C genómica (%)	3,72	3,68	3,74	3,70	1,70	1,69	1,09	1,08
Recuento M de sec. de CpG	5,85	5,83	5,72	5,70	5,54	5,53	5,88	5,85
Recuento de U de sec. de CpG	125161042	123579650	123020926	121374309	54155037	53756413	36788819	36453865
Cobertura de CpG (%)	49564325	49336003	47582363	47812716	21930864	21830555	13971713	13867179
Profundidad de sec. de CpG	73,37	73,25	76,24	76,12	63,93	63,80	57,87	57,77
Densidad de metilación de CpG (%)	6,20	6,16	6,05	6,01	2,70	2,68	1,80	1,79
Recuento de M de sec. noCpG	71,83	71,22	72,12	71,74	71,18	71,12	72,48	72,44
Recuento de U de sec. noCpG	1979410	1897750	1831150	1856842	894139	873671	614025	594289
Cobertura noCpG (%)	1950312392	1950307695	1976031255	1966232257	915375235	910910490	583173346	580050454
Profundidad de sec. noCpG	0,10	0,10	0,11	0,11	0,08	0,08	0,07	0,07
Densidad de metilación noCpG (%)	0,89	0,69	0,70	0,70	0,32	0,32	0,21	0,21
Recuento de M de sec. de CHG	0,10	0,10	0,10	0,09	0,10	0,10	0,11	0,10
Recuento de U de sec. de CHG	818974	598717	594356	573203	275945	272311	191120	186135
Cobertura de CHG (%)	512387336	510125996	507312325	503053927	283916067	282826050	160374603	179512404
Profundidad de sec. de CHG	63,24	63,11	66,75	66,62	54,69	54,56	48,57	48,45
Densidad de metilación de CHG (%)	4,96	4,94	4,91	4,90	2,30	2,29	1,46	1,45
Recuento de M de sec. de CHH	0,10	0,10	0,10	0,09	0,10	0,10	0,11	0,10
Recuento de U de sec. de CHH	1360436	1301033	1356524	1283439	614154	601360	422905	408154
Cobertura de CHH (%)	1347925056	1340180599	1368718940	1361136330	631459158	628084440	402796543	400438060
Profundidad de sec. de CHH	45,93	45,63	50,48	50,36	39,32	39,22	34,20	34,11
Densidad de metilación de CHH (%)	3,12	3,10	3,17	3,15	1,46	1,45	0,93	0,93
	0,10	0,10	0,10	0,09	0,10	0,10	0,10	0,10

FIG. 1A (Cont.)



150

FIG. 1B



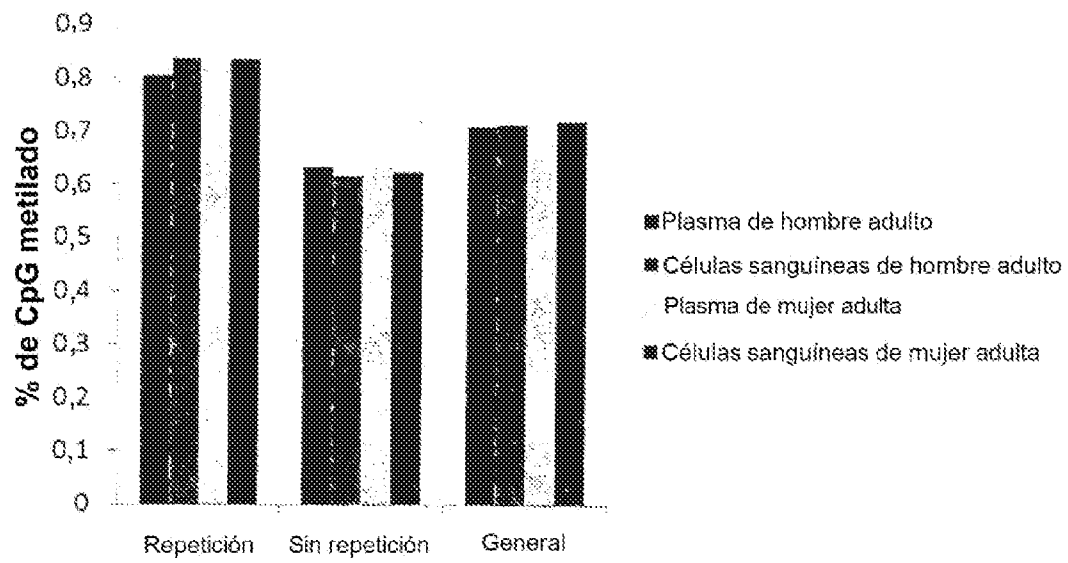


FIG. 3A

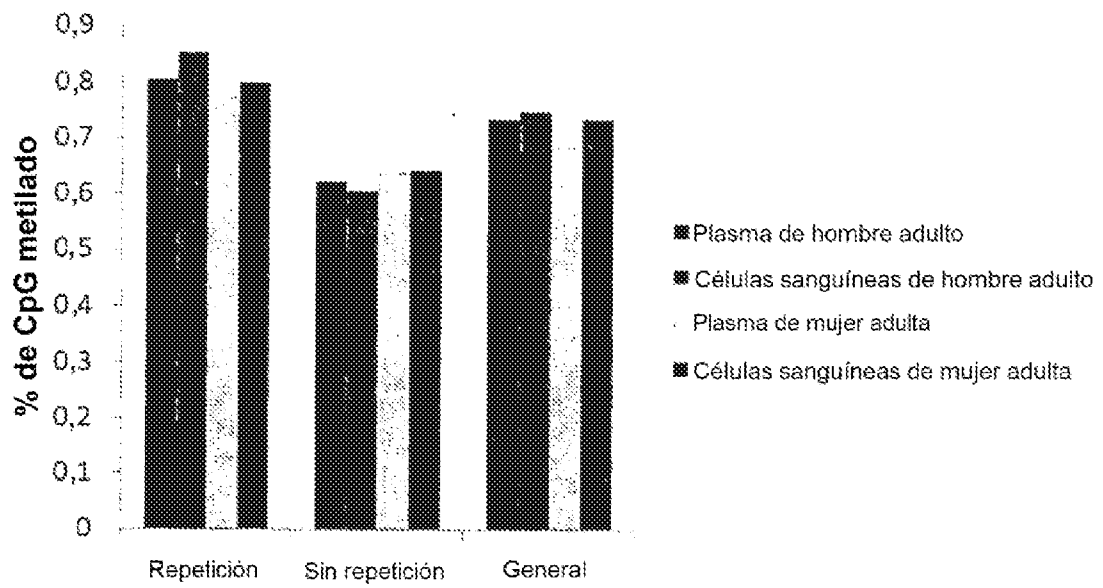


FIG. 3B

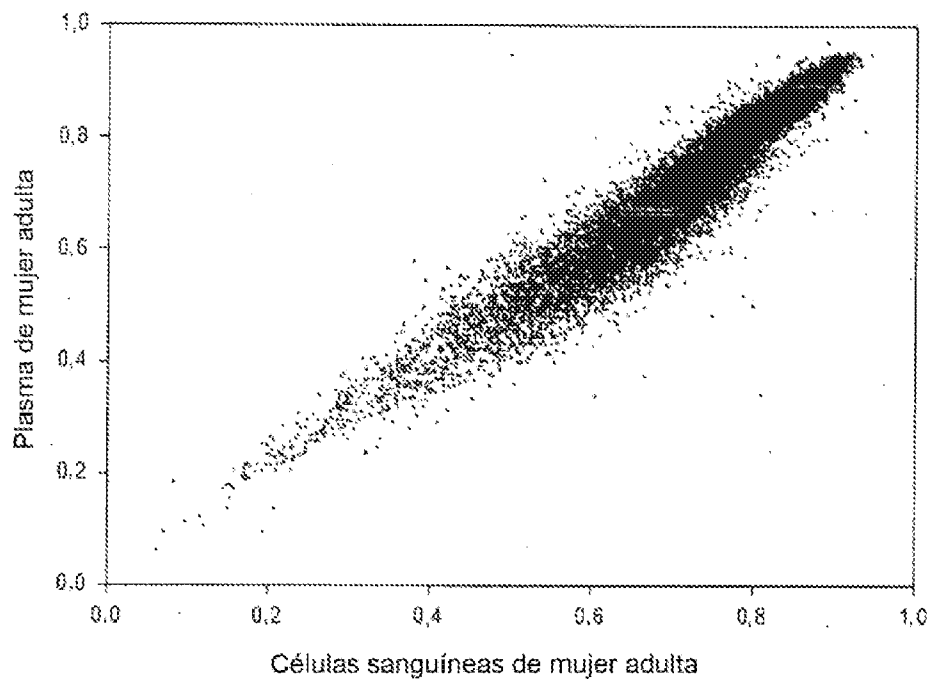


FIG. 4A

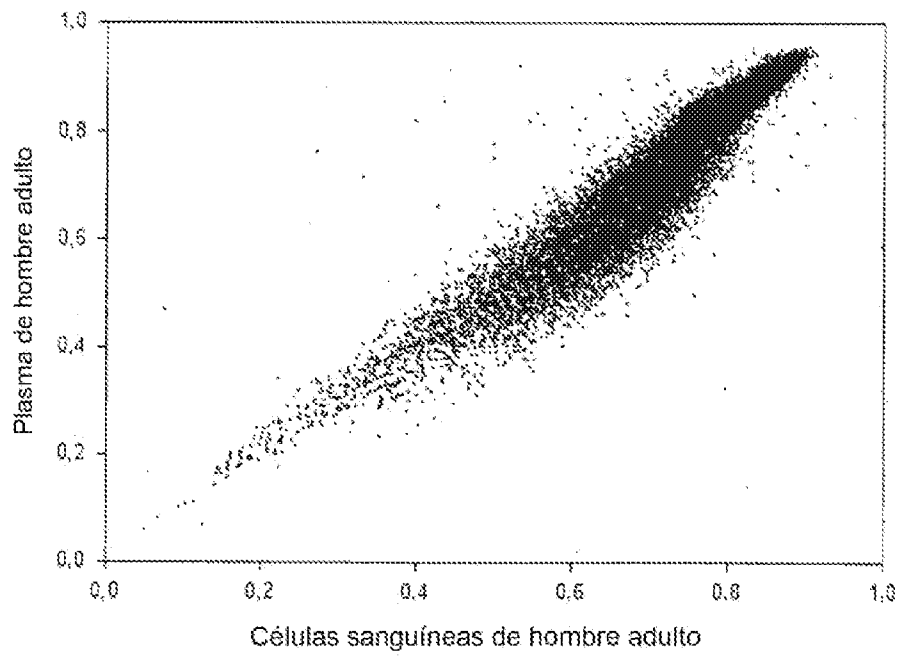


FIG. 4B

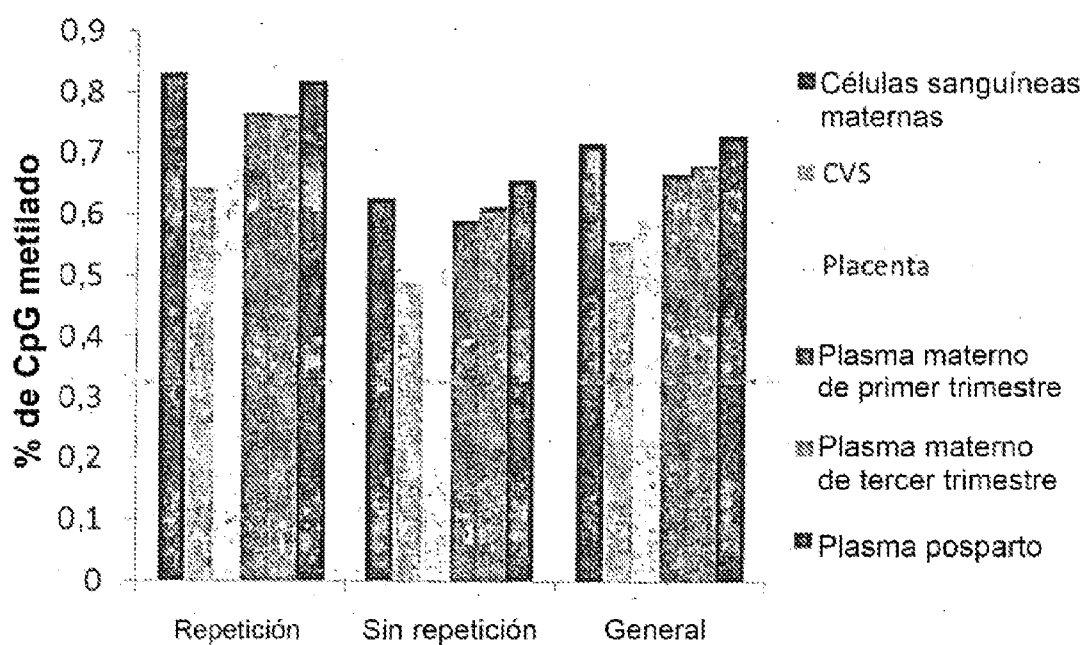


FIG. 5A

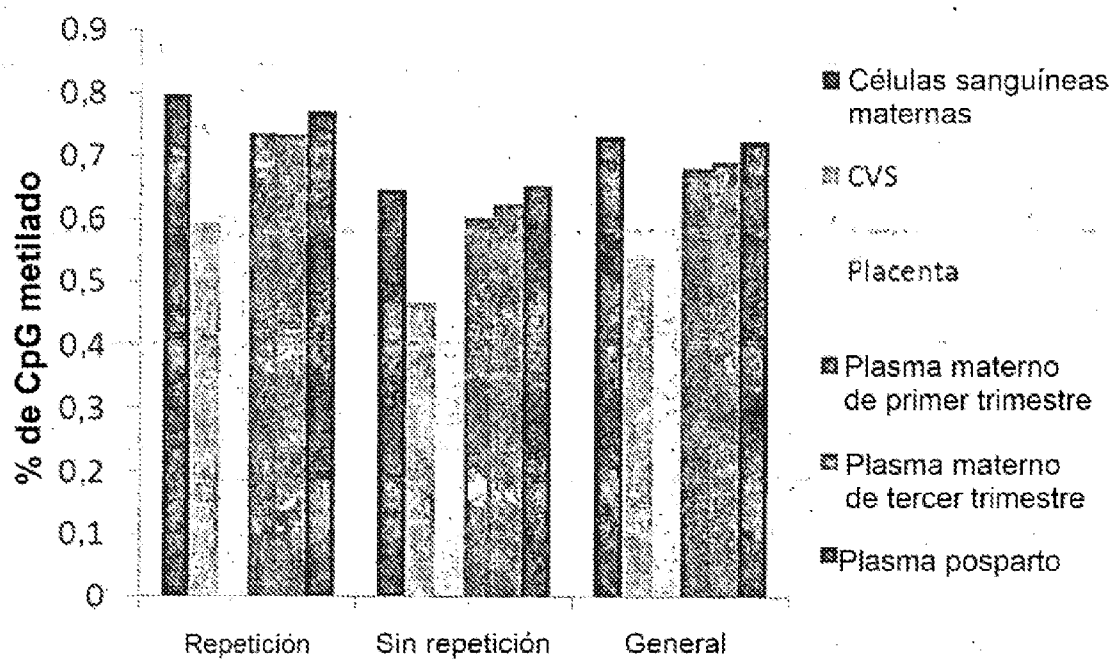


FIG. 5B

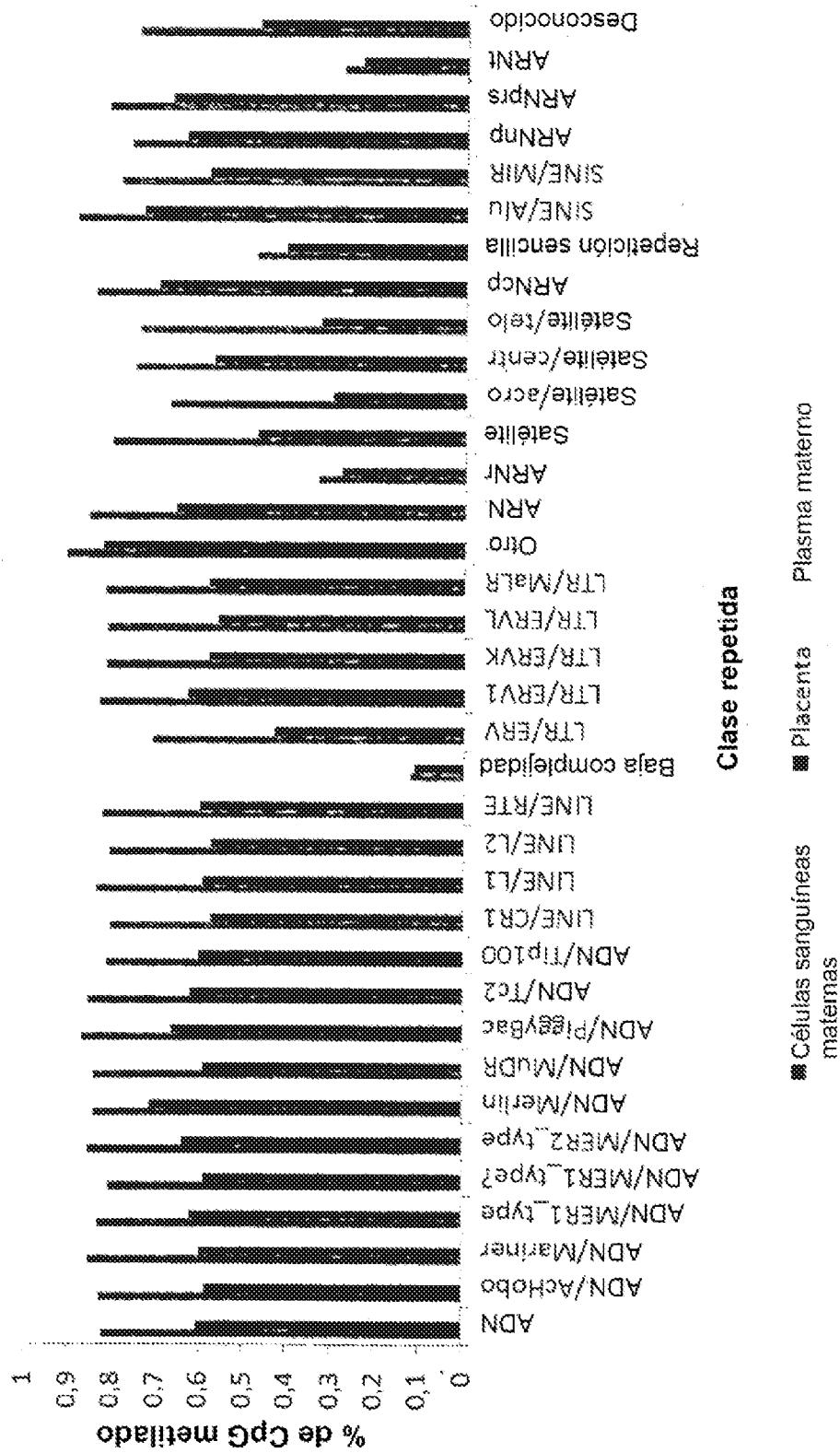


FIG. 6

De fuera a dentro:

Densidad de CpG
Célula sanguínea materna
Plasma materno (primer trimestre)
Lecturas de plasma compartido
Lecturas de plasma específico
fetal CVS

700

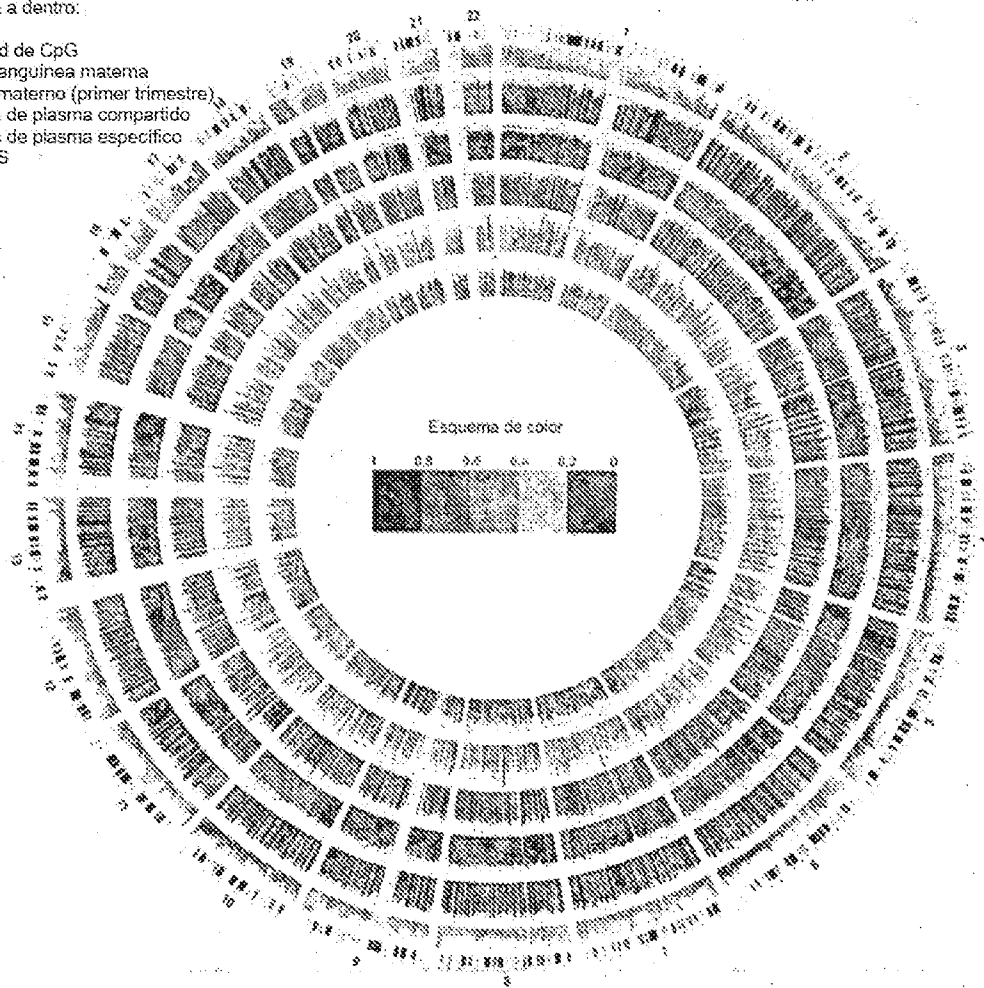


FIG. 7A

De fuera a dentro:

Densidad de CpG
Célula sanguínea materna
Plasma pos-parto
Plasma materno (tercer trimestre)
Lecturas de plasma compartido
Lecturas de plasma específico fetal
Tejido placentario

750

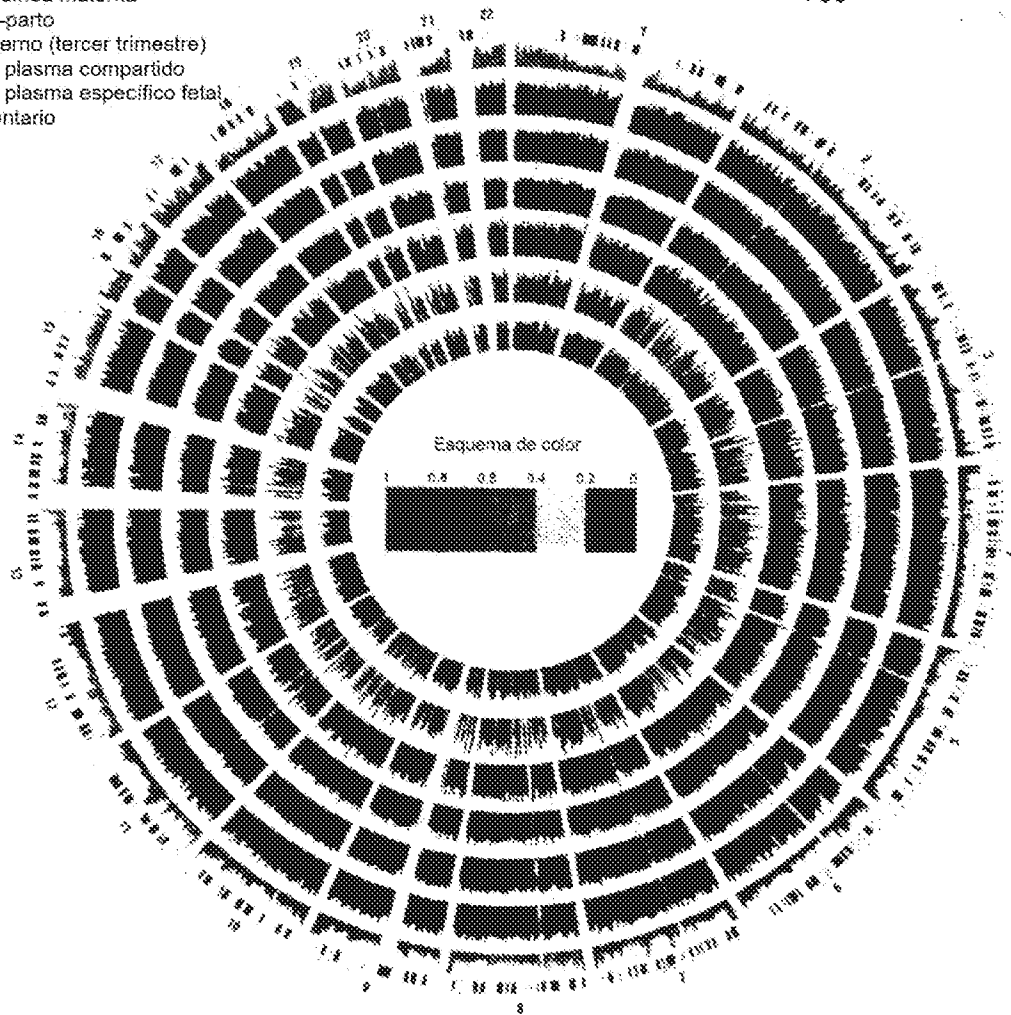


FIG. 7B

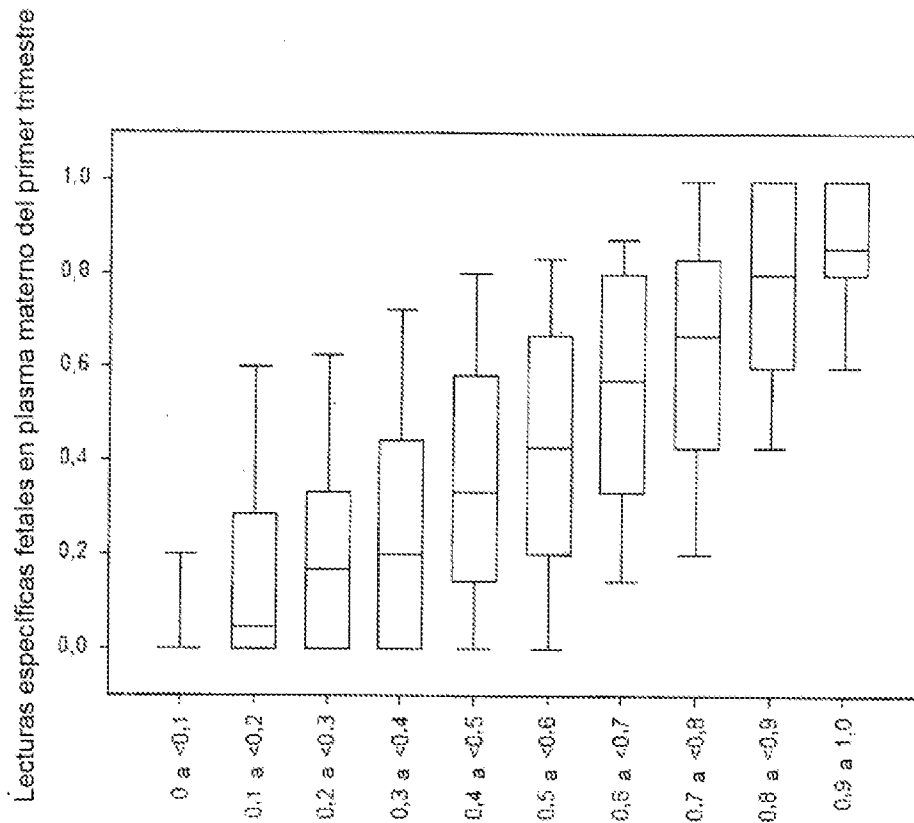


FIG. 8A

Lecturas de muestra de vellosidad coriónica

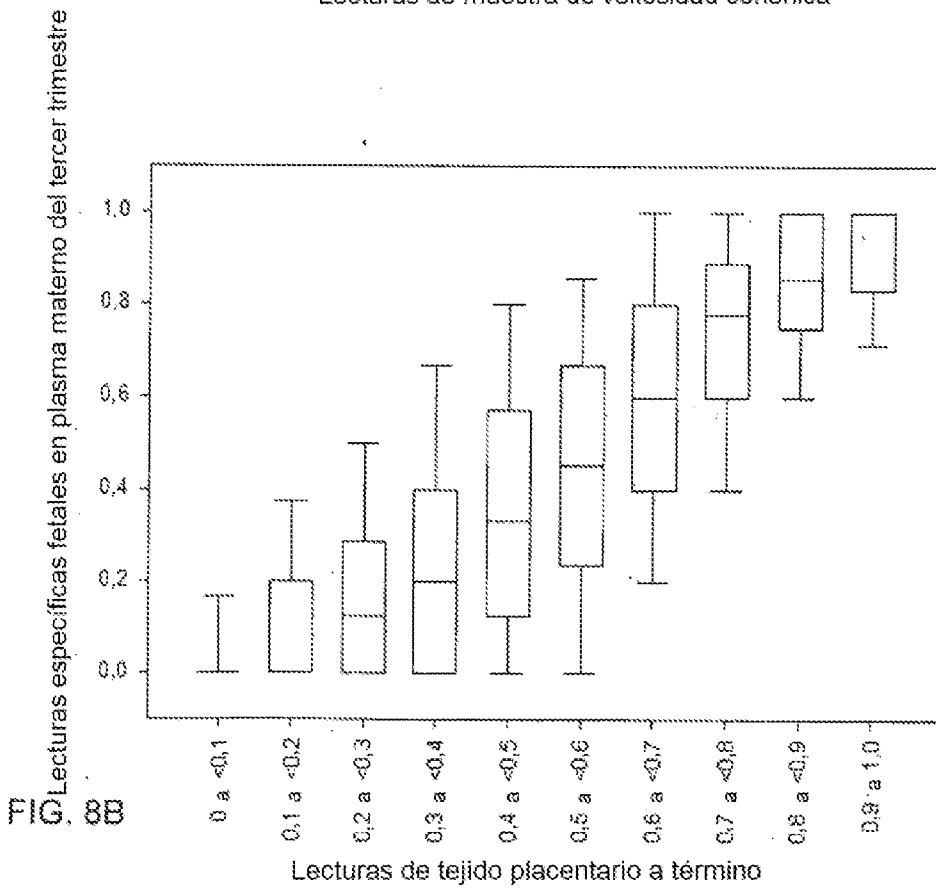
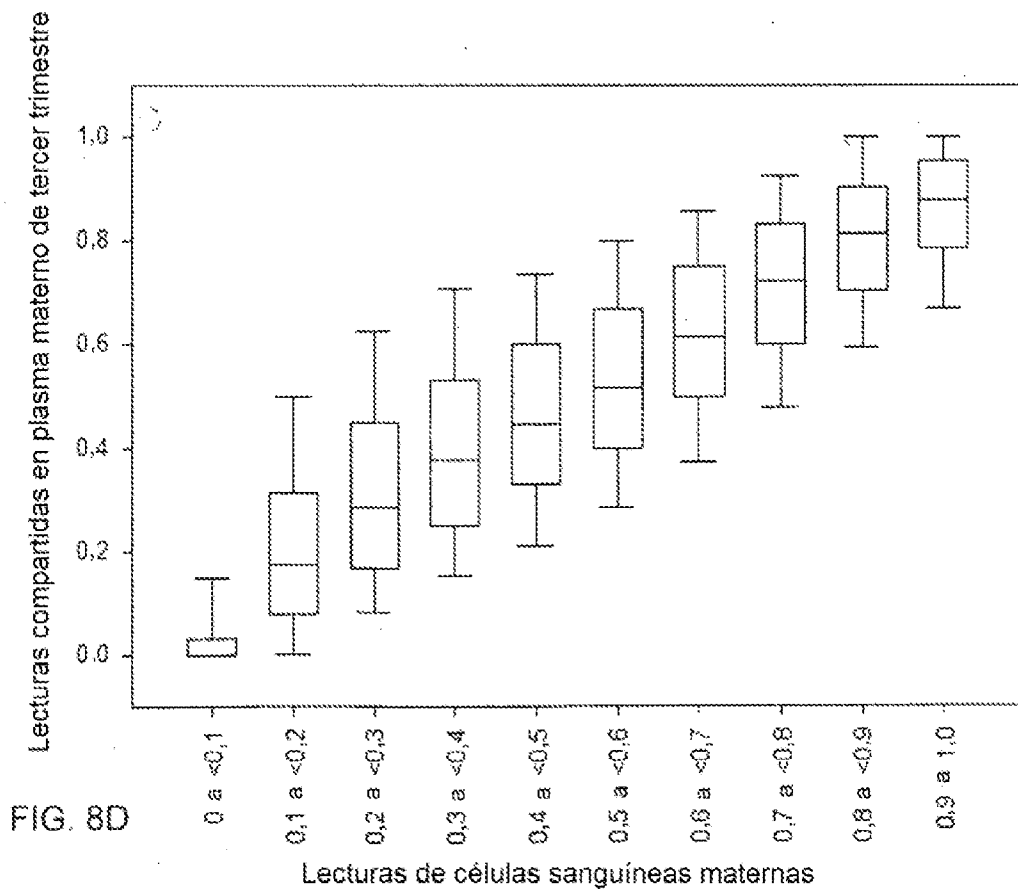
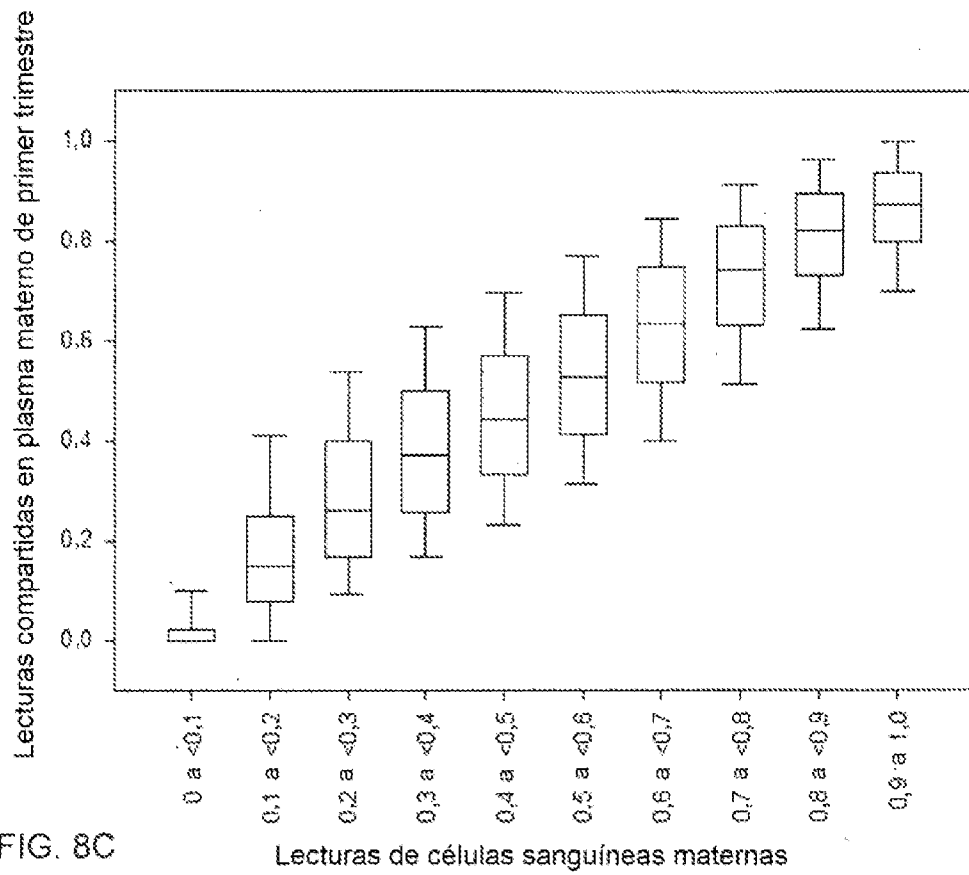


FIG. 8B

Lecturas de tejido placentario a término



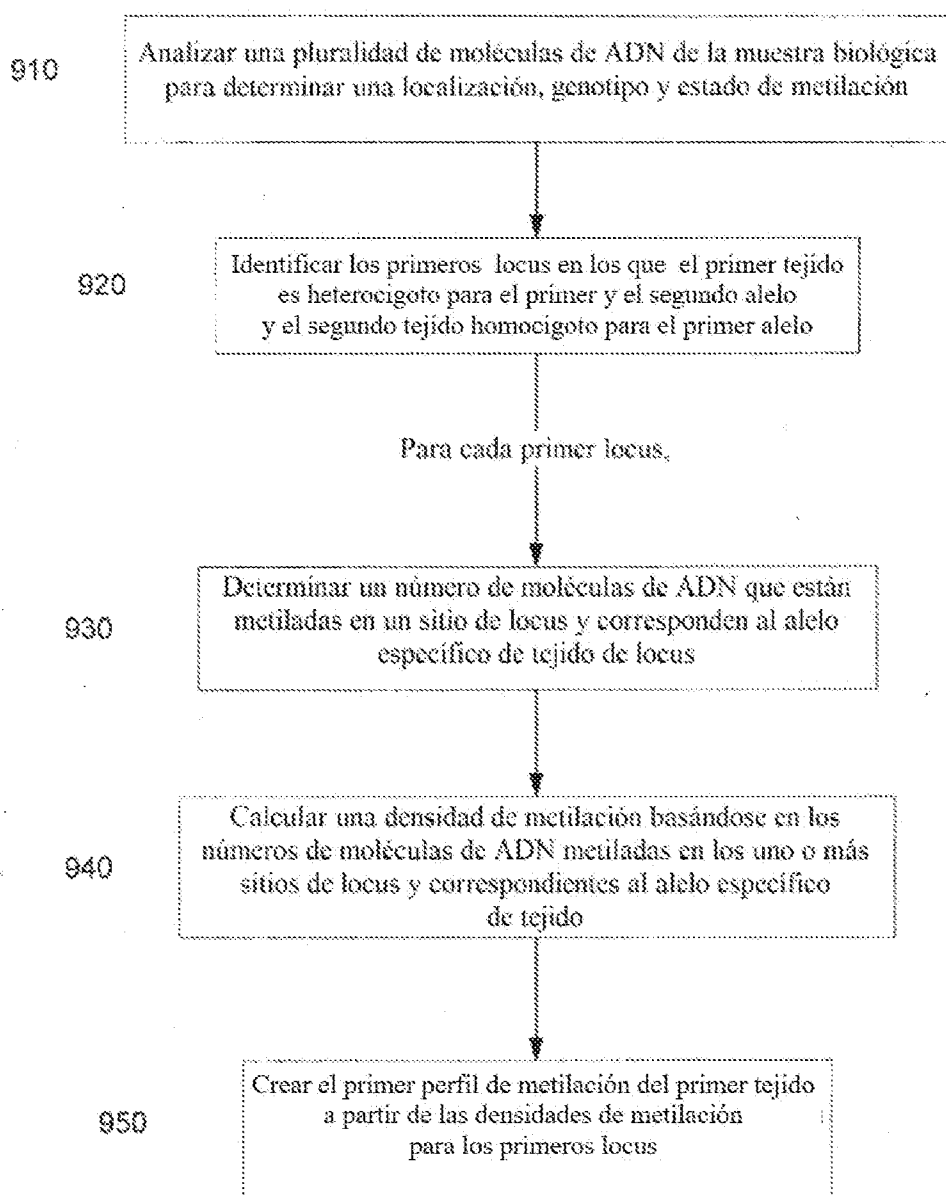


FIG. 9

900

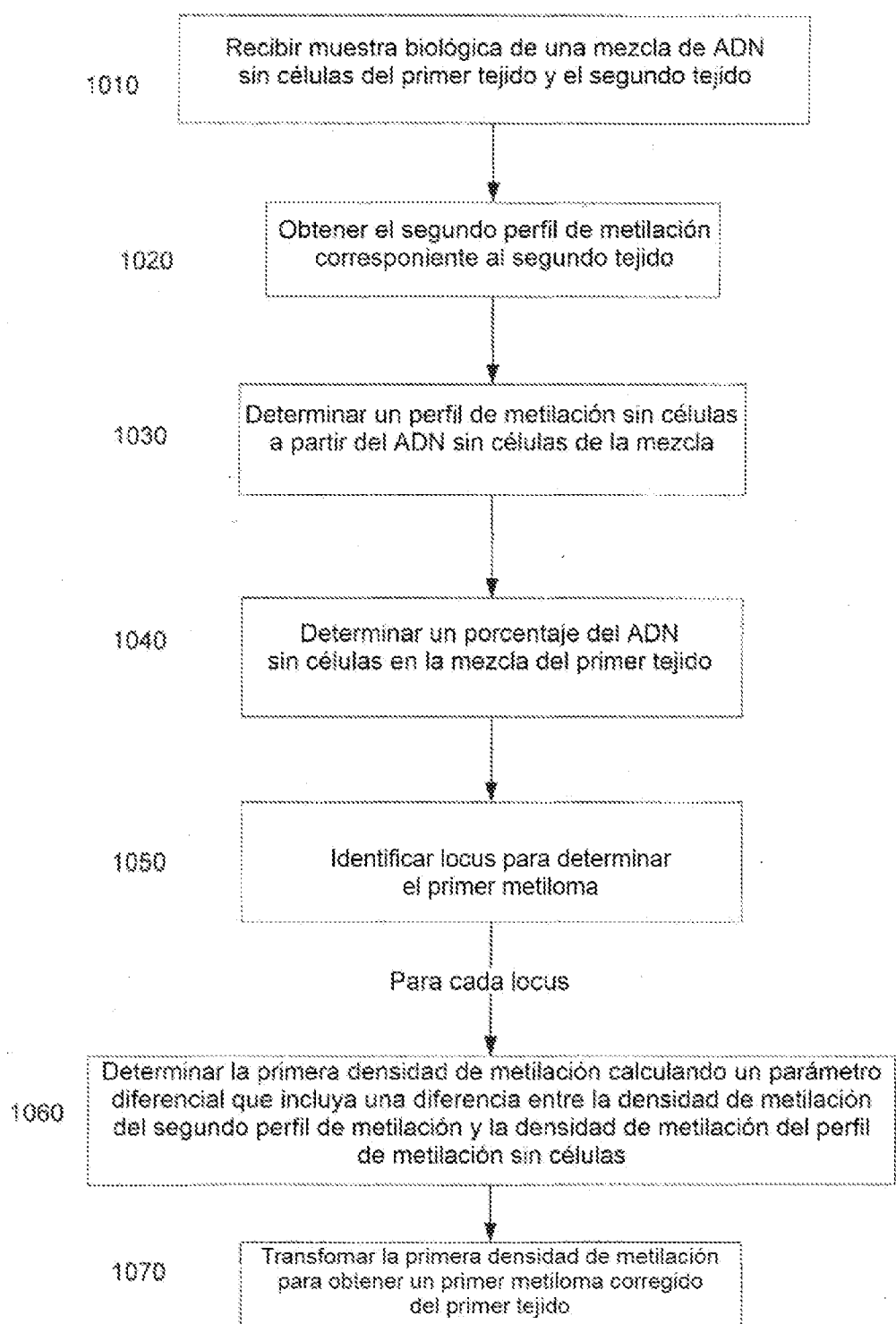
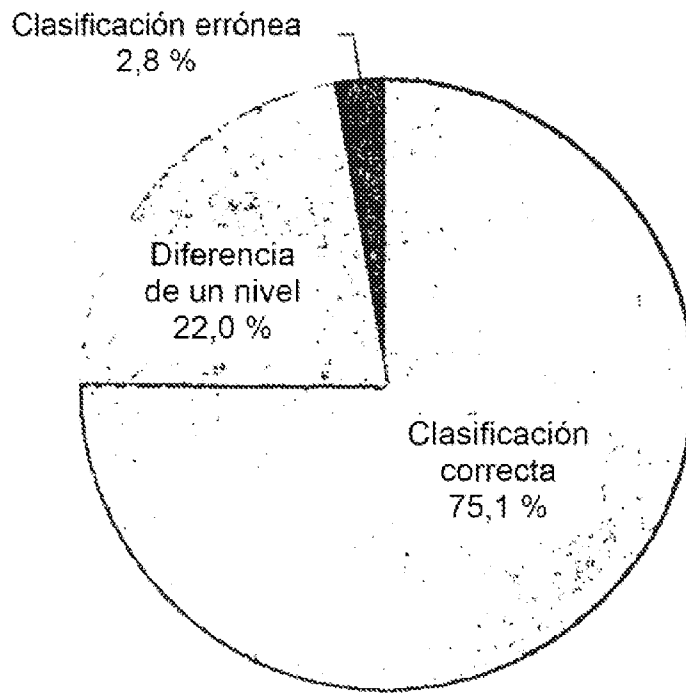


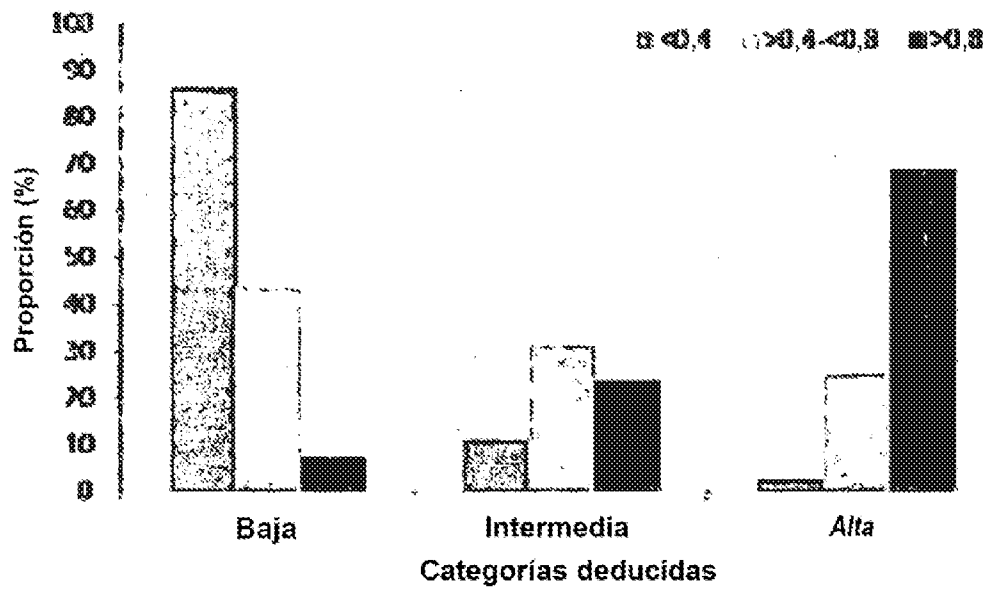
FIG. 10

1000



1100

FIG. 11A



1150

FIG. 11B

Densidad de metilación (datos de secuenciación)					Deducción		Estado de metilación en la bibliografía				
N.º	Locus génico	Localización genómica	Sangre materna (BKG)	Placenta	Plasma materno (PM)	Valores deductivos (D) ²	Puntuación predictiva de metilación (PPM) ²	Categorías de metilación deducidas ³	Nivel de metilación placentario ³	Estado placentario ⁶	Referencias
1	RASSF1A	chr3:50352938-50353234	0.030	0.637	0.067	0.506	1.060	Alto	Inter	Hipermetilado	{1}
2	CG009	chr21:25856489-25856756	0.001	0.432	0.050	0.374	0.849	Alto	inter	Hipermetilado	{2}
3	CG028	chr21:33321942-33322182	0.035	0.331	0.040	0.073	0.367	Bajo	Bajo	Hipermetilado	{2}
4	CG051	chr21:36995232-36995509	0.043	0.303	0.045	0.072	0.365	Bajo	Bajo	Hipermetilado	{2}
5	CG071	chr21:41140466-41140898	0.031	0.531	0.045	0.145	0.483	Inter	Inter	Hipermetilado	{2}
6	CG113	chr21:44953820-44953972	0.013	0.547	0.051	0.306	0.729	Inter	Inter	Hipermetilado	{2}
7	CG137	chr21:46249872-46250196	0.940	0.413	0.820	0.030	0.299	Bajo	Inter	Hipometilado	{2}
8	CG149	chr21:46911575-46912533	0.024	0.440	0.030	0.202	0.573	Inter	Inter	Hipermetilado	{2}
9	HUCS	chr21:37274945-37275031	0.038	0.576	0.059	0.203	0.575	Inter	inter	Hipermetilado	{2}
10	PDE9A	chr21:42979270-42979525	0.950	0.303	0.823	-0.018	0.222	Bajo	Bajo	Hipometilado	{2}
11	-	chr21:36589647-36589995	0.952	0.149	0.027	0.007	0.262	Bajo	Bajo	Hipometilado	No publicada
12	SERPIN85	chr18:59294739-59295038	0.936	0.590	0.956	0.329	0.777	Inter	Inter	Hipometilado	{3}
13	VAPA	chr18:10022563-10023186	0.056	0.752	0.082	0.382	0.861	Alto	inter	Hipometilado	{4}
14	CAB1E51	chr18:19035954-19036069	0.335	0.801	0.491	1.519	2.680	Alto	Alto	Hipermetilado	{4}
15	BAGALT6	chr18:27486040-27486932	0.113	0.827	0.331	1.764	3.072	Alto	Alto	Hipermetilado	{4}

1200

FIG. 12A

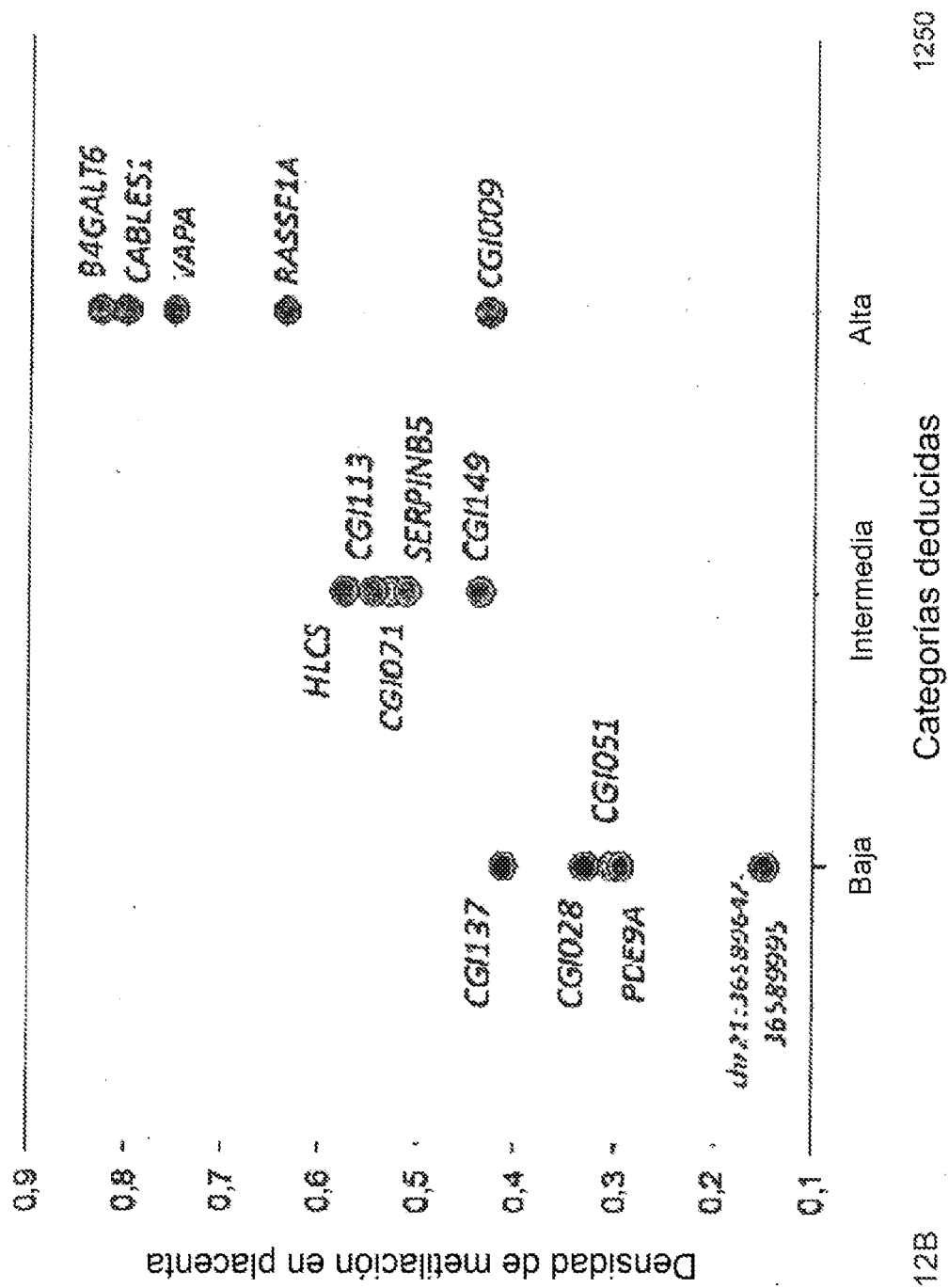


FIG. 12B

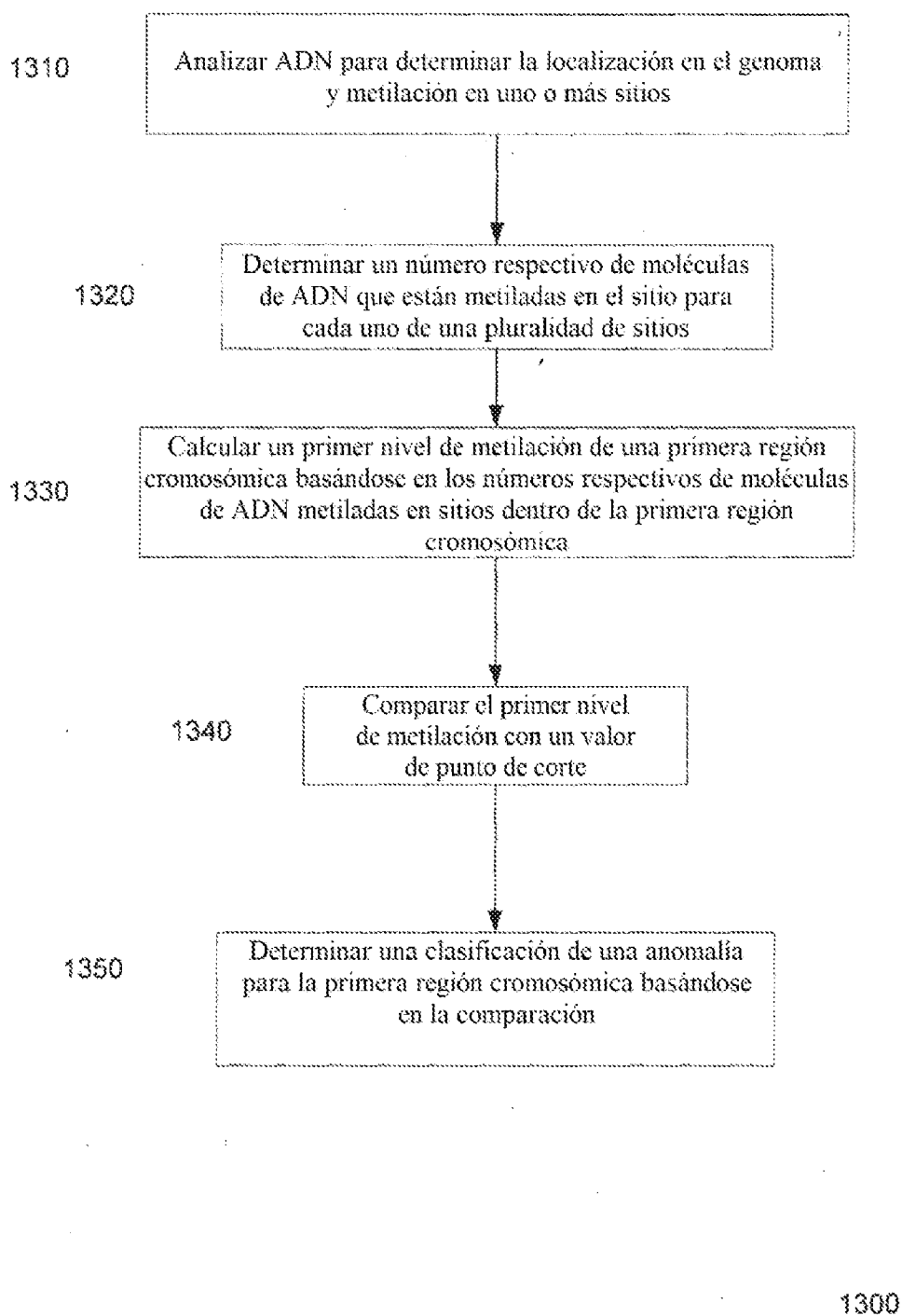


FIG. 13

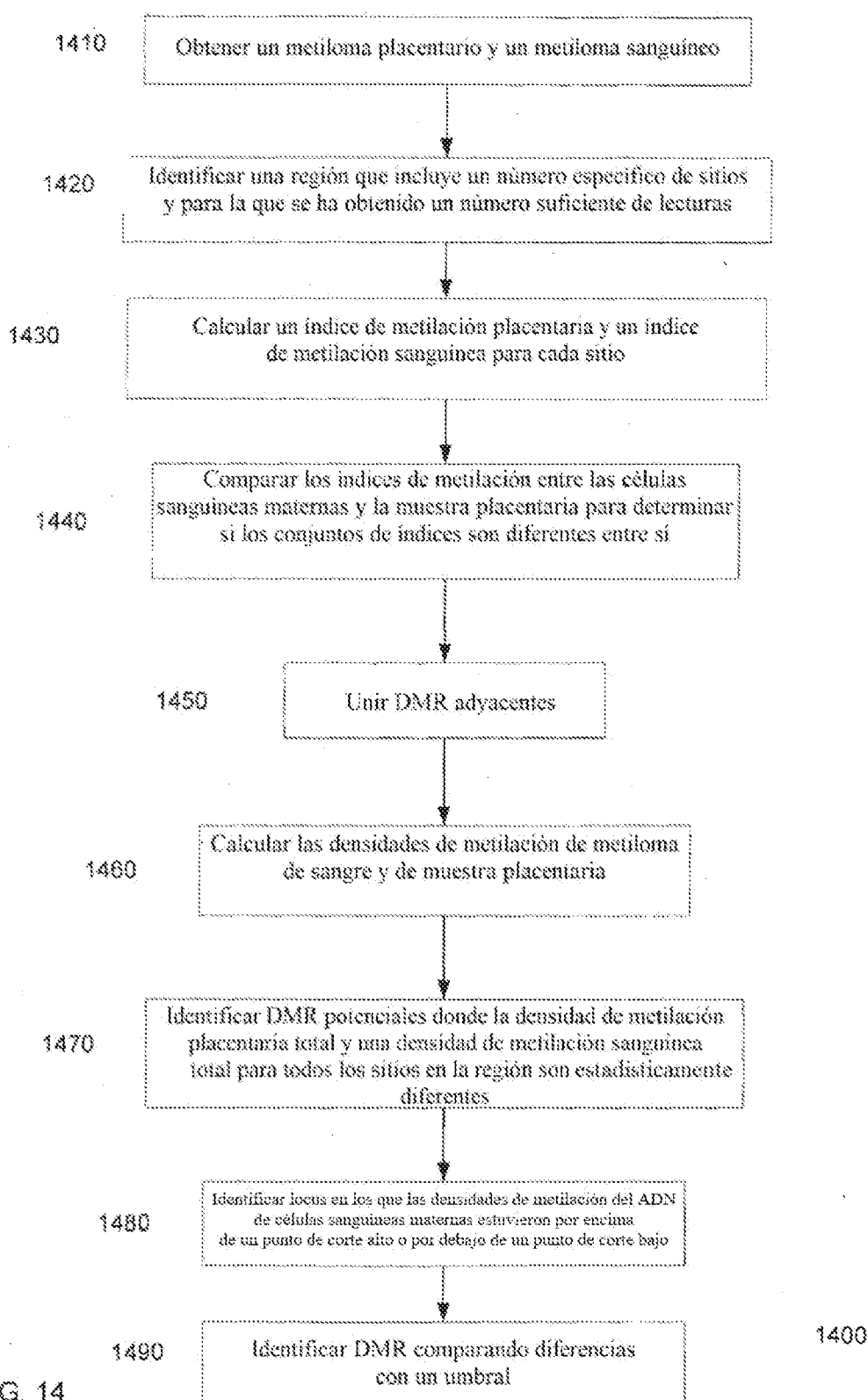


FIG. 14

N.º	Locus génico	Localización genómica ^a	Tamaño ^b	Estado polimórfico ^c	BS-Seq ^d			DMR ^e							
					Sequencia conservada	CVS de ref.	Cr	Sequencia	Final	Tamaño	Sequencia conservada	CVS de p.	valor de p.	cl. de ref.	Caso ^f
1	MASS1A	cr3:50352338-50352334	257	Hipometilado	0,00	60,73	60,73	cr3:50351900	50354350	3000	0,28	66,66	0,00E+00	60,42	TP
2	CGI694	cr2:25856489-25856488	368	Hipometilado	0,08	43,20	43,13	cr2:125855500	25857500	2000	4,87	48,70	0,00E+00	43,73	TP
3	CGI628	cr2:133321942-133321942	741	Hipometilado	3,61	33,33	29,56	cr2:133316400	33323900	7500	4,52	73,85	0,00E+00	19,23	FN
4	CGI691	cr2:138294232-38265302	278	Hipometilado	4,26	30,33	26,07	cr2:138292800	37003900	11100	3,12	21,61	0,00E+00	19,78	FN
5	CGI671	cr2:141140489-41140488	453	Hipometilado	3,29	52,06	49,97	cr2:141140000	41141000	1000	0,40	52,00	0,00E+00	49,80	TP
6	CGI113	cr2:144825820-44825820	153	Hipometilado	1,27	56,09	53,42	cr2:144848000	44854700	6200	7,18	53,23	0,00E+00	50,05	TP
7	CGI137	cr2:148249872-48220195	323	Hipometilado	84,80	41,34	52,85	cr2:148227800	48225460	22900	88,74	52,87	0,00E+00	37,27	TP
8	CGI148	cr2:148911375-48912533	959	Hipometilado	0,40	48,96	43,36	cr2:148911200	48912700	1500	6,31	49,17	0,00E+00	40,88	TP
9	HLCS	cr2:172727945-27275031	87	Hipometilado	0,77	57,61	53,85	cr2:17274200	37275200	1000	7,30	54,24	0,00E+00	46,81	TP
10	PDE5A	cr2:149787070-4978925	256	Hipometilado	85,01	30,33	64,88	cr2:14978100	40034800	25700	80,97	51,44	0,00E+00	38,53	TP
11		cr2:138585647-38585965	349	Hipometilado	96,17	14,89	85,23	cr2:138586200	38589200	1000	84,46	12,27	0,00E+00	75,82	TP
12	SERPINC8	cr18:58294739-58295058	360	Hipometilado	93,64	63,00	60,64	cr18:58294800	58296300	1500	96,47	40,70	3,32E-07	47,71	TP
13	VAPA	cr18:10222963-10223187	525	Hipometilado	3,83	75,17	71,54	cr18:10221800	10223300	1500	24,08	80,33	0,00E+00	68,25	TP
14	CABLES1	cr18:19031804-19032070	117	Hipometilado	33,47	80,09	68,62								FN
15	BAGAL76	cr18:27488533-27488533	604	Hipometilado	1,34	82,75	71,41	cr18:27488100	27489100	1000	11,34	82,75	0,00E+00	71,41	TP
16	CASPR	cr2:201803042-201831985	1544	Hipometilado	0,59	47,54	47,00	cr2:201803000	20183200	2980	3,77	46,11	4,20E-02	45,94	TP
17	BARB	cr3:254448403-254450200	261	Hipometilado	2,25	36,93	34,85	cr3:25443900	25445100	1500	3,44	22,97	0,00E+00	19,53	FN
18	APC	cr5:112150999-112151943	945	Hipometilado	0,71	48,78	45,68	cr5:112150300	112151800	1500	0,59	48,08	0,00E+00	44,60	TP
19	DAB2IP _{int}	cr9:121541221-121541485	240	Hipometilado	84,87	82,83	82,04	cr9:121540400	121541500	1100	84,21	80,88	5,28E-08	49,36	TP
20	DAB2IP _{ex}	cr9:123118261-123129077	817	Hipometilado	95,08	32,89	62,10	cr9:123127700	123128200	1500	90,72	62,61	1,47E-06	40,11	TP
21	THY1	cr11:118798852-118798852	431	Hipometilado	6,88	39,32	38,44	cr11:118797600	118798800	2000	2,25	27,28	0,00E+00	25,03	TP
22	PTPNS	cr12:8939211-8939211	3562	Hipometilado	1,52	70,86	74,29	cr12:8930000	8934000	4000	4,26	80,13	0,00E+00	75,87	TP
23	PVCARD _{ex}	cr16:31121905-31122148	332	Hipometilado	4,90	48,51	41,51								FN
24	cr18A	cr18:55090284-55090303	323	Hipometilado	6,14	31,79	25,85	cr18:55089700	55090400	4500	4,76	22,81	0,00E+00	18,13	FN
25	cr21A	cr2:138276723-38230304	382	Hipometilado	12,89	64,62	51,21	cr2:138275300	38230800	900	4,74	62,67	7,17E-02	53,83	TP
26	cr21B	cr2:144161027-44161371	345	Hipometilado	0,85	83,36	83,84	cr2:144160700	44161700	1000	11,66	77,32	0,00E+00	68,28	TP
27	cr21C	cr2:133320644-33320823	286	Hipometilado	5,49	37,92	32,48	cr2:133316400	33323600	7500	4,52	33,86	0,00E+00	19,33	FN
28	cr21D	cr2:14218823-42188623	461	Hipometilado	13,02	85,84	72,82	cr2:142188700	42189700	1000	7,76	81,87	3,82E-07	73,89	TP
29	cr21E	cr2:143355712-43355708	197	Hipometilado	0,00	57,85	57,85	cr2:143355000	43357500	2500	6,21	62,23	0,00E+00	57,02	TP
30	cr21F	cr2:142357215-42357341	127	Hipometilado	0,57	49,25	44,86	cr2:142357000	42357300	2500	0,21	52,23	5,00E+00	53,02	TP
31	cr21H	cr2:132288803-32288943	141	Hipometilado	0,49	78,84	78,35	cr2:132288300	32289300	1500	20,43	85,29	4,86E-08	84,78	TP
32	cr21I	cr2:144079403-44079503	301	Hipometilado	1,23	87,44	86,21	cr2:144079400	44080400	1000	15,19	88,06	4,95E-08	82,87	TP
33	SERPINC5_P	cr18:58294811-58294811	465	Hipometilado	91,98	39,71	52,25	cr18:58294800	58295300	1500	88,47	40,76	3,32E-06	47,71	TP

FIG. 15A

1500

R.N. ^a	Locus genico	BS-Seq			DMR ^b								Clase ^c
		Cálculo variaciones maternos	Placenta de met ^d	diff	Cr	Inicio	Final	Tamaño	Cálculo variaciones paternos	Placenta	valor de p	diff de met	
1	RASSF1A	0,03	60,84	60,84	cr3	59362100	60354100	2000	2,16	62,21	0,00E+00	60,85	TP
2	CGI009	0,08	20,50	20,43	cr21	25855600	25857100	1500	2,04	24,37	0,00E+00	22,33	TP
3	CGI028	3,51	17,38	13,85	-	-	-	-	-	-	-	-	TN
4	CGI051	4,38	13,93	9,57	-	-	-	-	-	-	-	-	TN
5	CGI071	3,39	24,85	21,78	cr21	41139900	41141400	1500	6,83	29,28	0,00E+00	22,45	TP
6	CGI113	1,27	40,24	38,97	cr21	44948500	44954700	6200	7,18	45,00	0,00E+00	37,82	TP
7	CGI137	94,00	37,10	56,90	cr21	46248400	46250400	2000	92,28	43,55	1,00E-76	48,43	TP
8	CGI148	0,40	22,19	21,79	cr21	46911200	46912700	1500	5,31	25,73	0,00E+00	30,42	TP
9	HUCS	3,77	58,49	54,72	cr21	37274200	37275200	1000	7,33	56,83	0,00E+00	46,50	TP
10	PDE5A	95,01	32,41	62,60	cr21	42978900	42990100	11200	90,89	44,38	0,00E+00	46,51	TP
11	-	95,17	19,92	75,24	cr21	36589200	36590200	1000	94,49	22,66	0,00E+00	71,83	TP
12	SERPINE5	93,64	28,06	65,57	cr18	59294500	59295500	1000	92,11	36,53	3,85E-31	55,58	TP
13	VAPA	3,63	73,66	70,25	cr18	10022100	10023100	1000	11,67	75,76	0,00E+00	84,09	TP
14	CABLES1	33,47	81,40	47,93	-	-	-	-	-	-	-	-	FN
15	BAGALT6	11,34	88,57	77,23	cr18	27485800	27486800	1000	28,67	90,00	4,3779E-59	81,03	TP
16	CASPB	0,52	56,27	55,75	cr2	201830200	201832200	2000	3,77	57,80	1,33E-52	53,83	TP
17	RARB	2,28	38,16	35,90	cr3	25444000	25445000	1000	2,36	24,42	8,00E-54	22,06	TP
18	APC	0,71	34,89	34,18	cr5	112100700	112101700	1000	0,38	32,97	0,00E+00	32,59	TP
19	CAS2IP_on	84,87	40,28	44,60	cr9	121540400	121541400	1000	84,79	30,63	1,7308E-14	54,15	TP
20	CAS2IP_ex	95,09	15,14	79,95	cr9	123128600	123129100	500	95,14	17,42	9,5247E-47	77,72	TP
21	THY1	0,88	22,28	21,40	cr11	118797800	118799300	1500	1,22	20,50	0,00E+00	19,28	FN
22	PTPN6	1,52	81,33	79,82	cr12	6930000	6934000	4000	4,26	81,58	0,00E+00	77,32	TP
23	PYCARD_ex	4,90	43,73	43,83	cr16	31121700	31122300	500	4,85	32,87	1,1158E-34	28,02	TP
24	chr18A	5,14	16,26	10,10	-	-	-	-	-	-	-	-	TN
25	chr21A	12,80	74,65	61,85	cr21	39278300	39280300	2000	9,96	76,07	0,00E+00	65,11	TP
26	cr21B	0,86	81,56	80,70	cr21	44160700	44161700	1000	11,06	78,9	0,00E+00	67,54	TP
27	cr21C	5,49	14,91	9,42	-	-	-	-	-	-	-	-	TN
28	cr21D	13,02	88,71	75,70	cr21	42158700	42189700	1000	7,78	84,42	3,5806E-77	76,54	TP
29	cr21E	0,03	73,21	73,21	cr21	42355300	42357800	2500	4,66	76,19	0,00E+00	74,53	TP
30	cr21E1	3,57	84,00	80,43	cr21	42355300	42367800	2500	4,66	76,19	0,00E+00	74,53	TP
31	cr21H	0,49	77,49	77,00	cr21	32268400	32268900	500	18,4	83,85	2,1972E-30	66,45	TP
32	cr21I	1,23	58,62	57,39	cr21	44079300	44080300	1000	12,16	72,55	2,3559E-50	60,39	TP
33	SERPINE5_P	91,95	28,61	63,36	cr18	59294500	59295500	1000	92,11	36,53	3,85E-31	55,58	TP

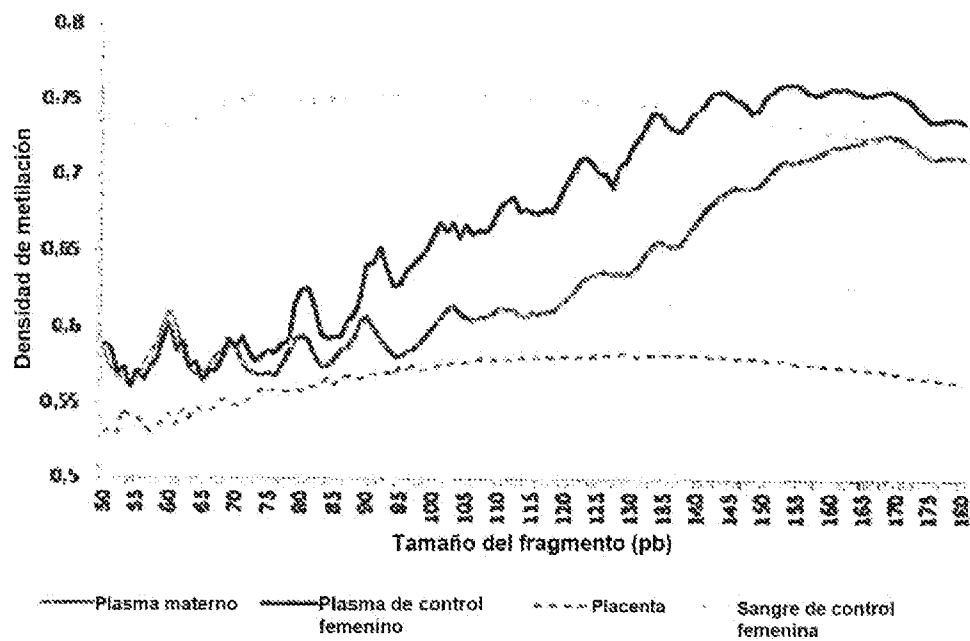
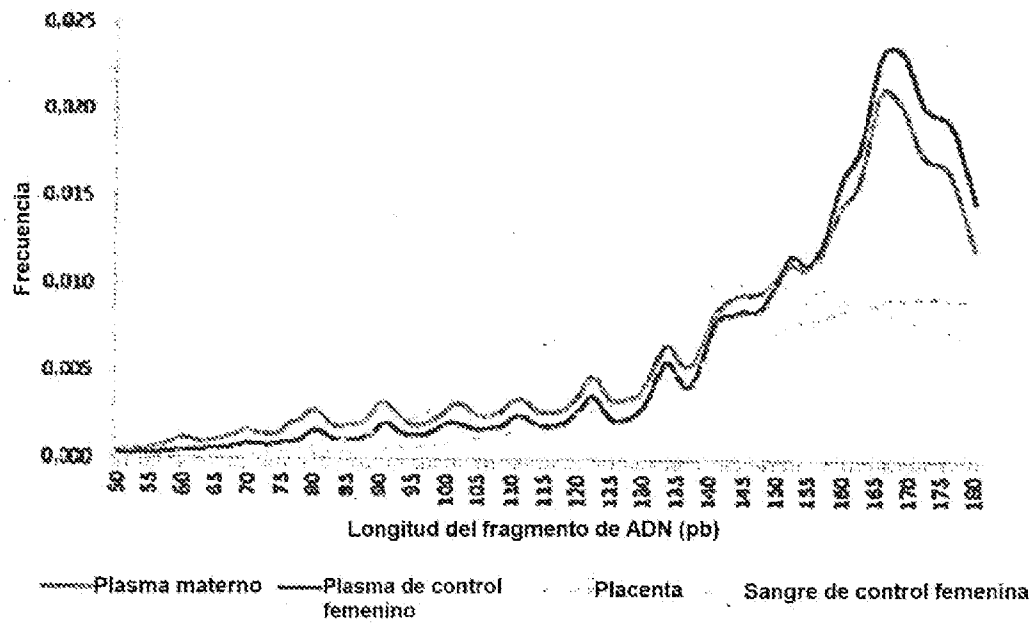
FIG. 15B

1550

	Plasma materno de primer trimestre		Plasma materno de tercer trimestre	
	Hipermetilado ^a	Hipometilado ^b	Hipermetilado ^a	Hipometilado ^b
N.º de locus predicho	3081	44455	1746	14930
N.º de locus con densidades de metilación > 40% en los datos de tejido placentario ^c	1.678	N/A	1.525	N/A
N.º de locus con densidades de metilación < 50% en los datos de tejido placentario ^c	N/A	23468	N/A	13475
N.º de locus solapantes con DMR obtenidas de los datos del tejido placentario ^c y células sanguíneas maternas	1457	21912	1279	12677

FIG. 16

1600



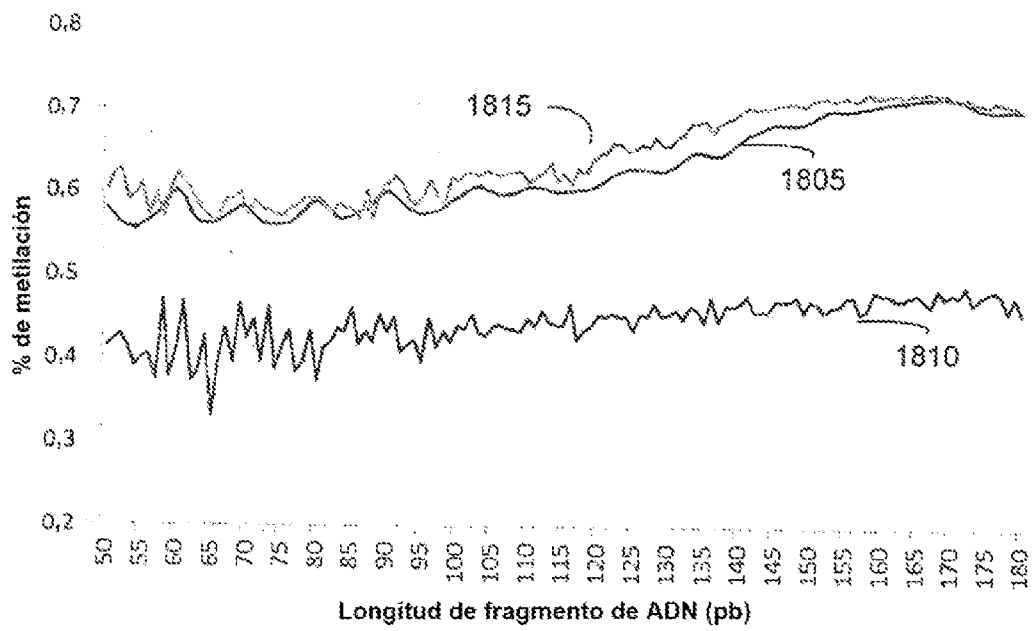


FIG. 18A

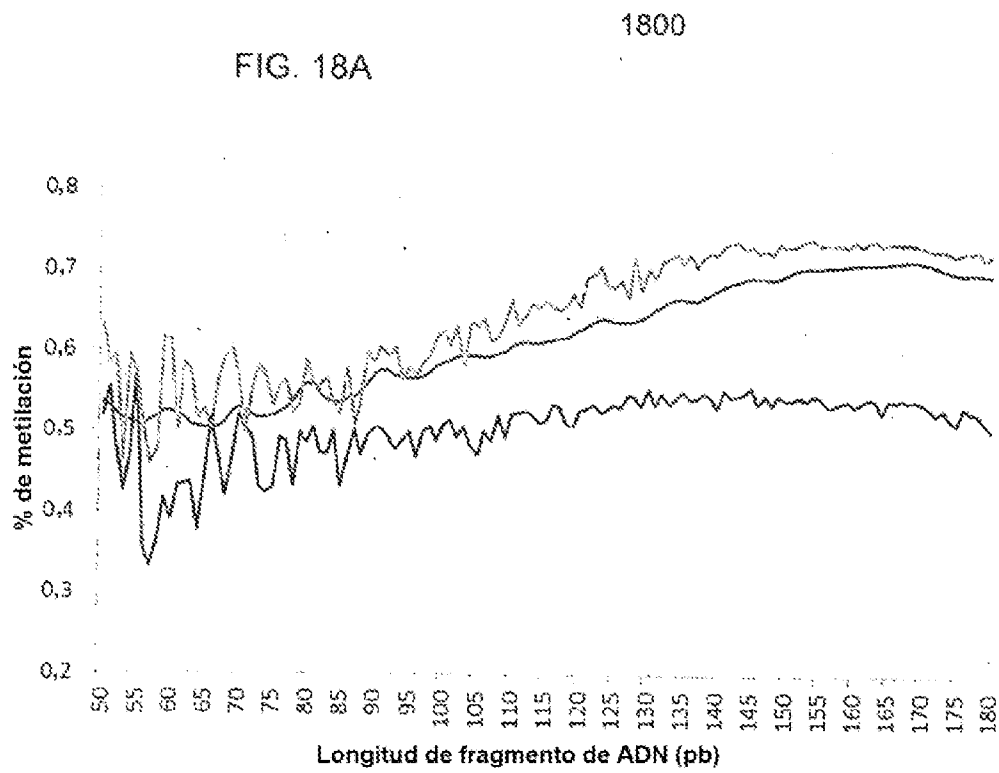


FIG. 18B

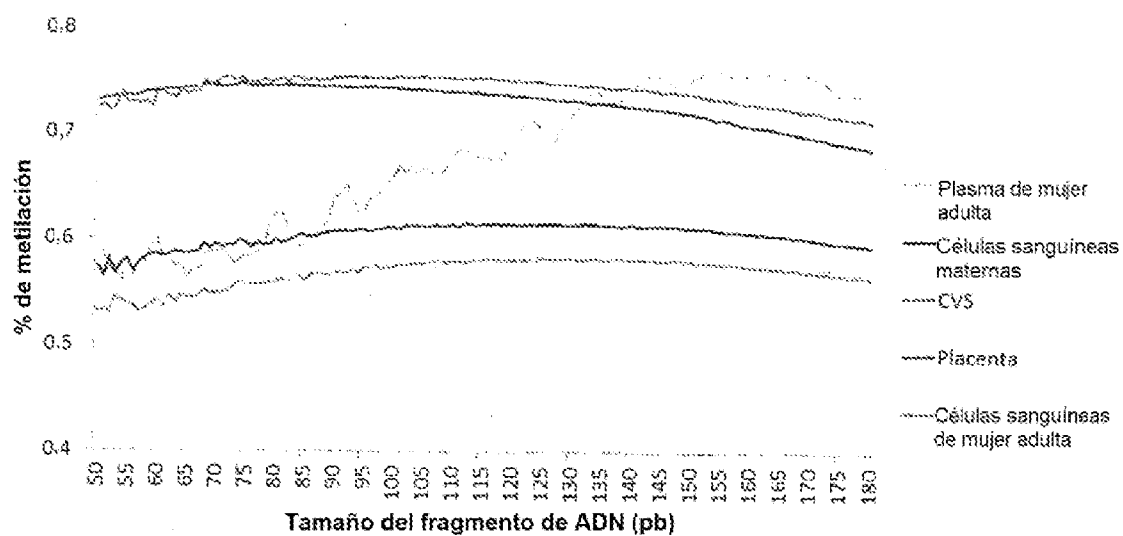
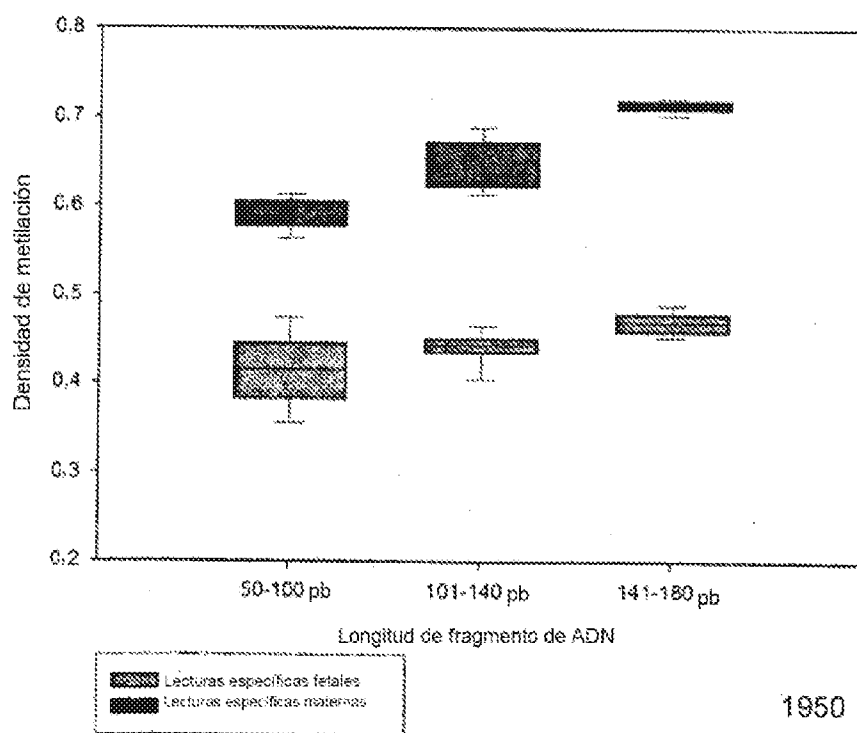


FIG. 19A

1900



1950

FIG. 19B

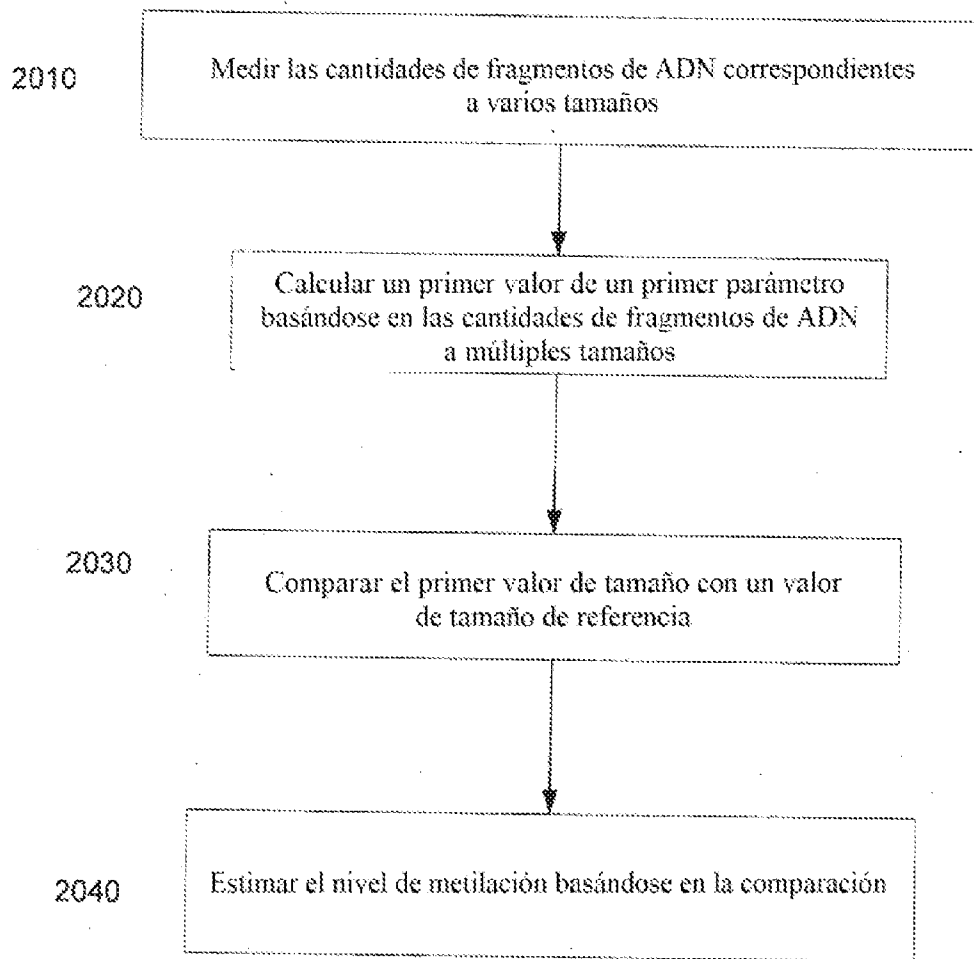


FIG. 20

2000

Cromosoma	Paciente de CHC			
	Capa leucoplaquetaria	Tejido hepático no tumoral	Tejido tumoral	Plasma pre-operatorio
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

	Recuentos de lecturas sin procesar	Recuento de lecturas mapeadas	Fragmentos no duplicados (solamente autosomas)	Proporción de lecturas utilizables %	Profundidad por cadena	Cobertura de CpG %	Profundidad en sitios de CpG con lecturas cubiertas por cadena
Sumario							
Control 1	159554537	56323206	41653660	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	83268888	33%	2,11	64,31%	3,65
Capa leucoplacuetaria	579329822	439663895	375813248	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 pre-operatorio	798183998	637735722	474751039	60%	12,72	87,18%	18,69
Plasma pos-operatorio	921833848	624851729	550104801	60%	14,73	89,94%	24,81

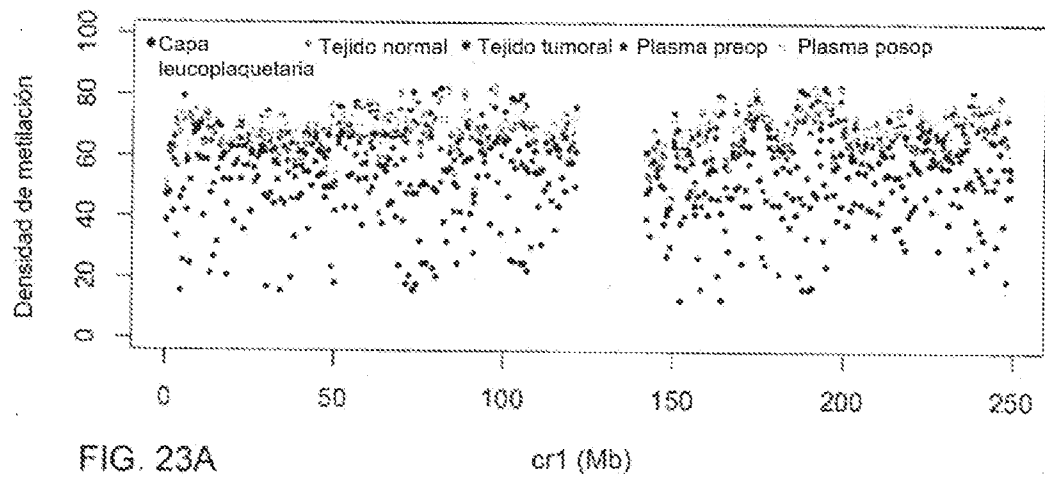
FIG. 21B

2150

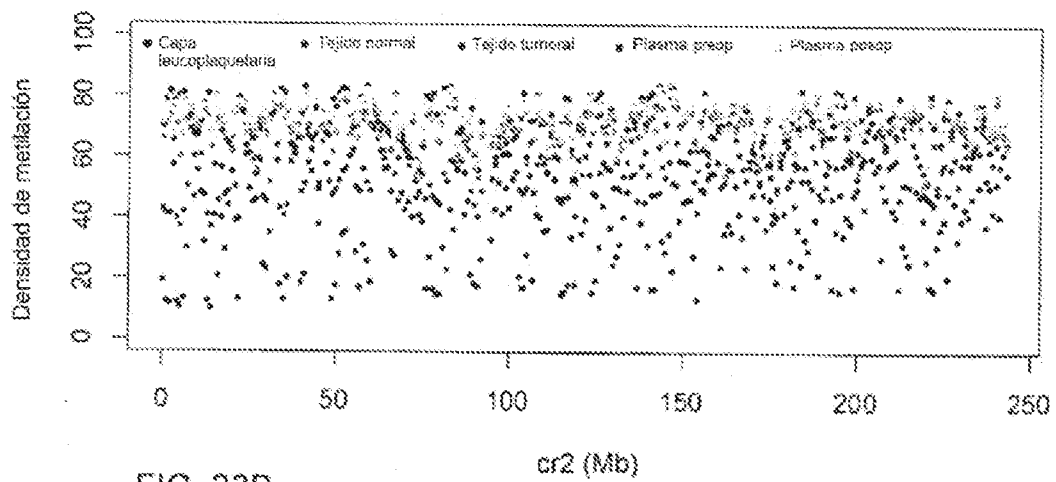
Cr	Densidad de metilación de ADN de plasma					
	Paciente de CHC		Control 1	Control 2	Control 3	Control 4
	Preoperatorio	Posoperatorio				
1	60,2%	69,6%	70,0%	70,3%	70,1%	71,3%
2	58,6%	71,9%	72,2%	72,4%	72,3%	73,5%
3	60,9%	72,3%	72,7%	73,1%	72,9%	73,9%
4	56,9%	71,8%	71,9%	72,2%	72,2%	73,4%
5	56,9%	70,9%	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%	66,8%	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



2300



2350

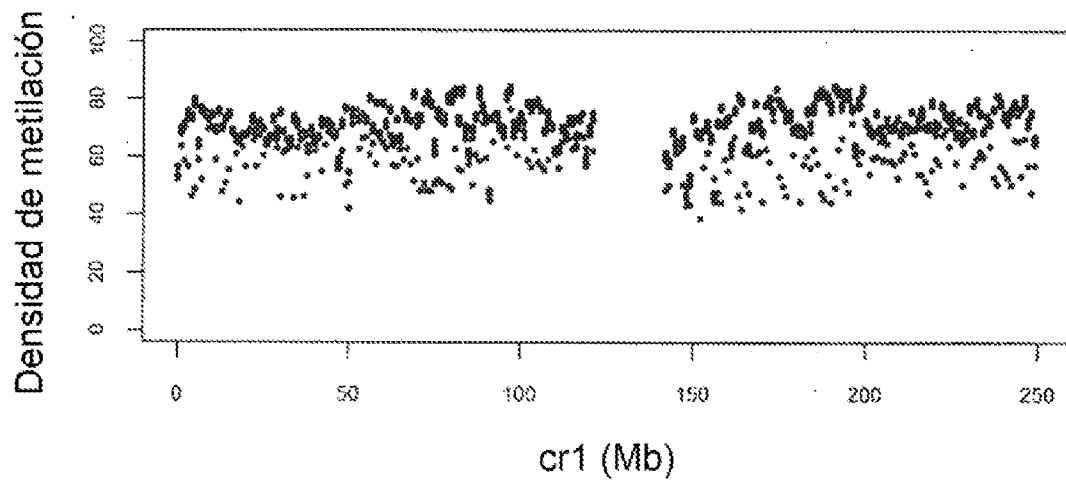


FIG. 24A

2400

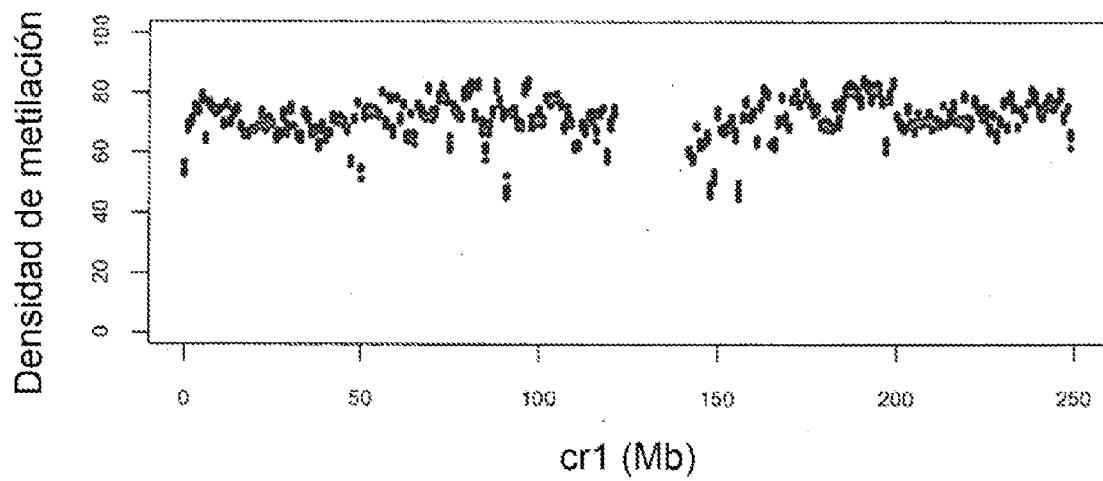


FIG. 24B

2450

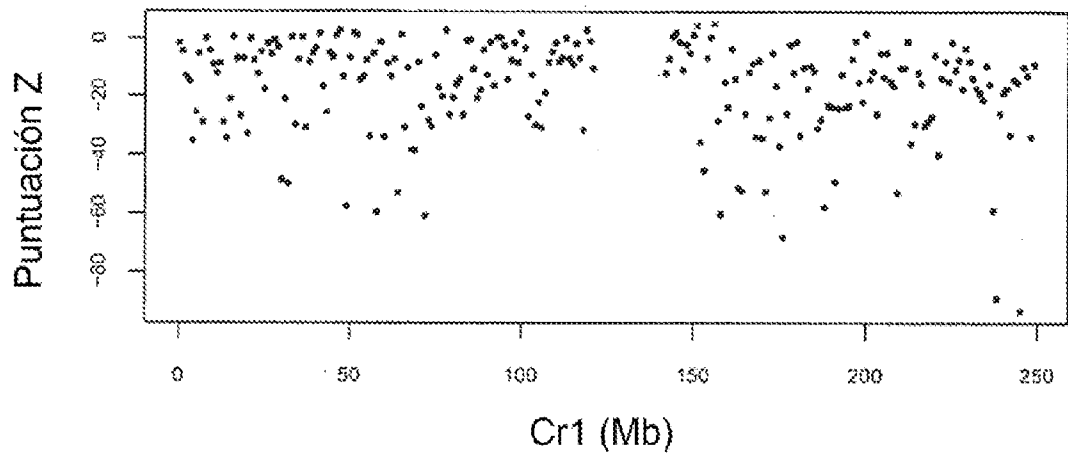


FIG. 25A

2500

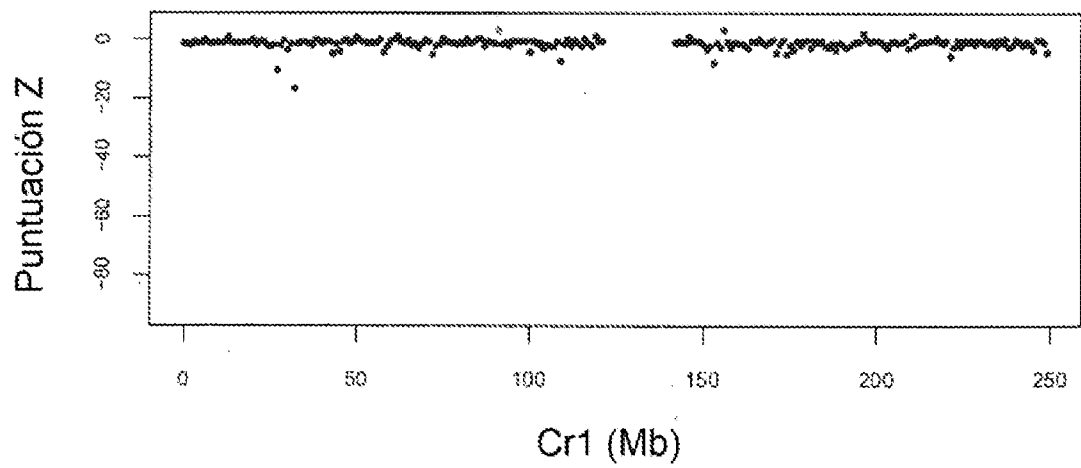


FIG. 25B

2550

Puntuación Z	Plasma preoperatorio (N.º de grupos)	Plasma posoperatorio (N.º de grupos)
<-3	186 (80,9%)	19 (8,3%)
-3 a 3	42 (18,3%)	209 (90,3%)
>3	2 (0,9%)	2 (0,9%)
Total	230 (100%)	230 (100%)

FIG. 26A

2600

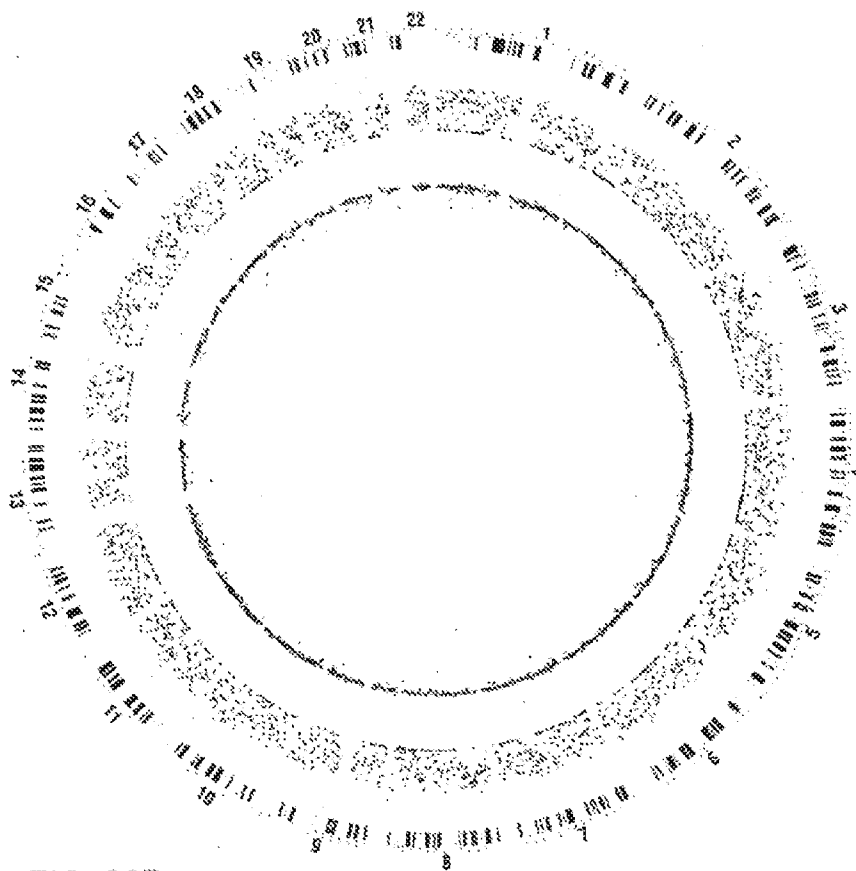


FIG. 26B

2620

puntuación z	Plasma preoperatorio (N.º de grupos)	Plasma postoperatorio (N.º de grupos)	C06	Control 1 (N.º de grupos)	Control 2 (N.º de grupos)	Control 3 (N.º de grupos)	Control 4 (N.º de grupos)
<-3	2330 (85,2%)	171 (6,3%)	29 (1,1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
-3 a 3	362 (13,3%)	2556 (93,5%)	2688 (98,3%)	2734 (100%)	2734 (100%)	2734 (100%)	2734 (100%)
>3	42 (1,5%)	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 metilado (densidad de metilación)
Capa leucoplaquetaria de paciente de 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 posop	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

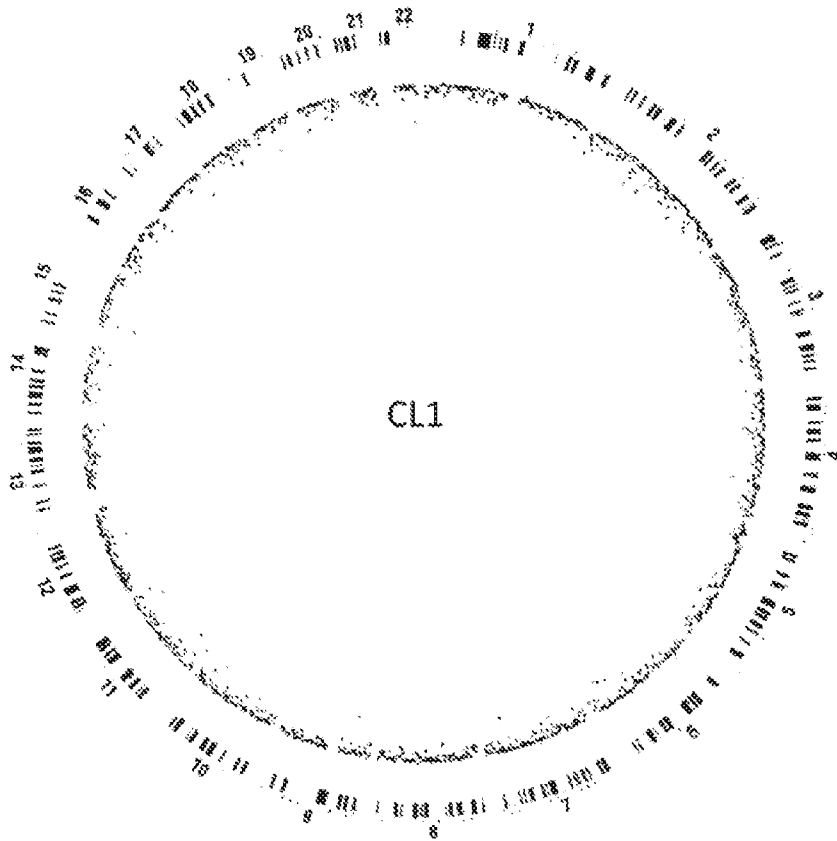


FIG. 27A

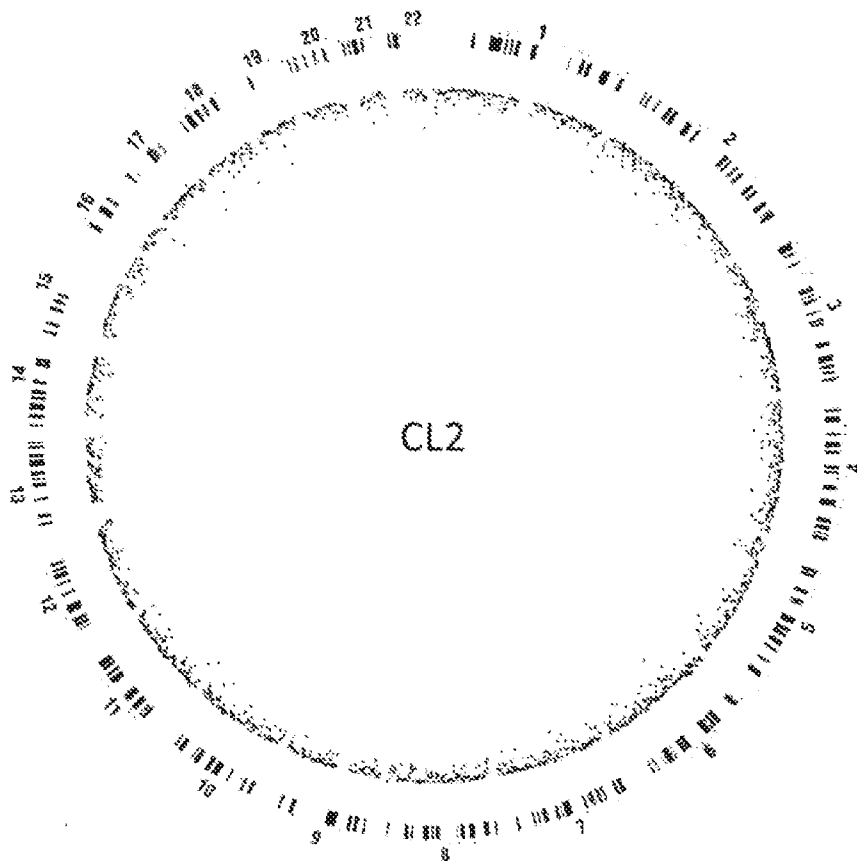


FIG. 27B

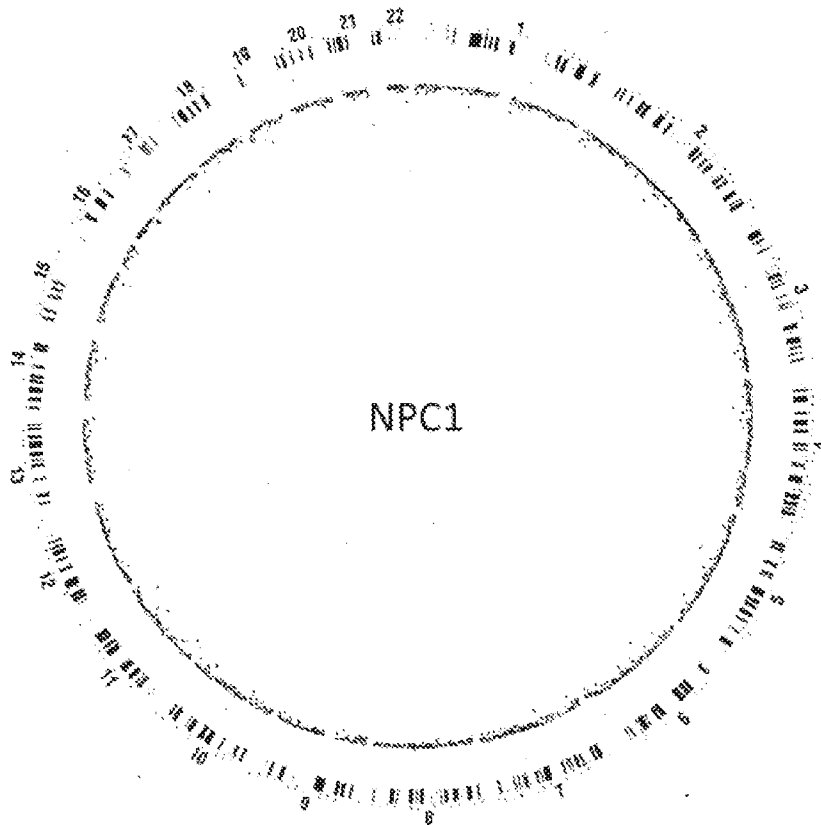


FIG. 27C

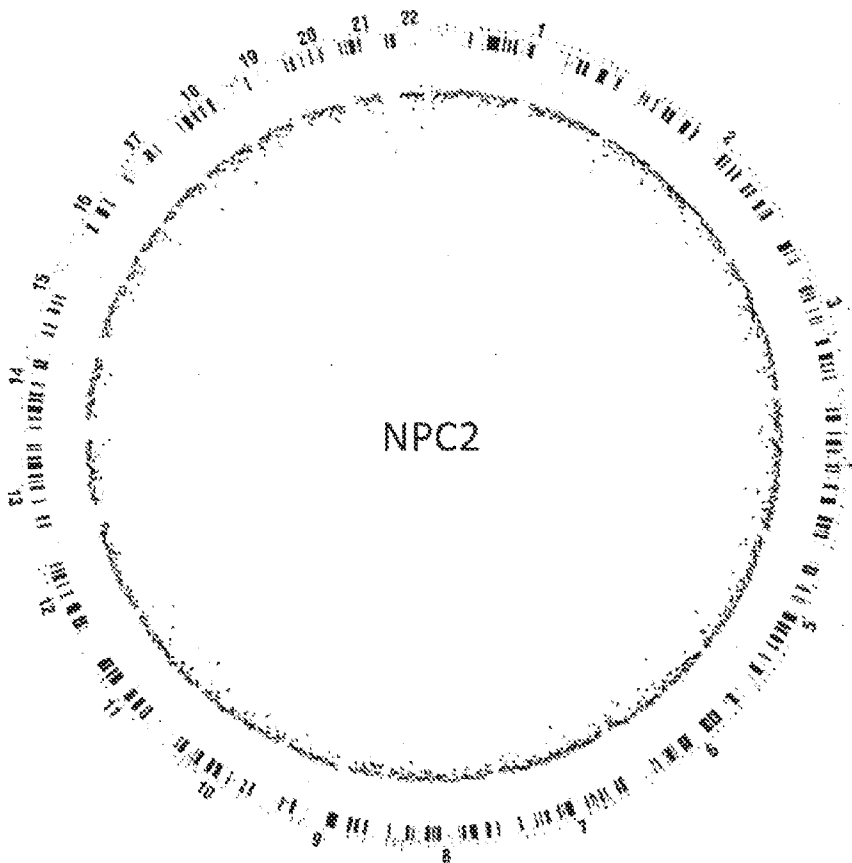


FIG. 27D

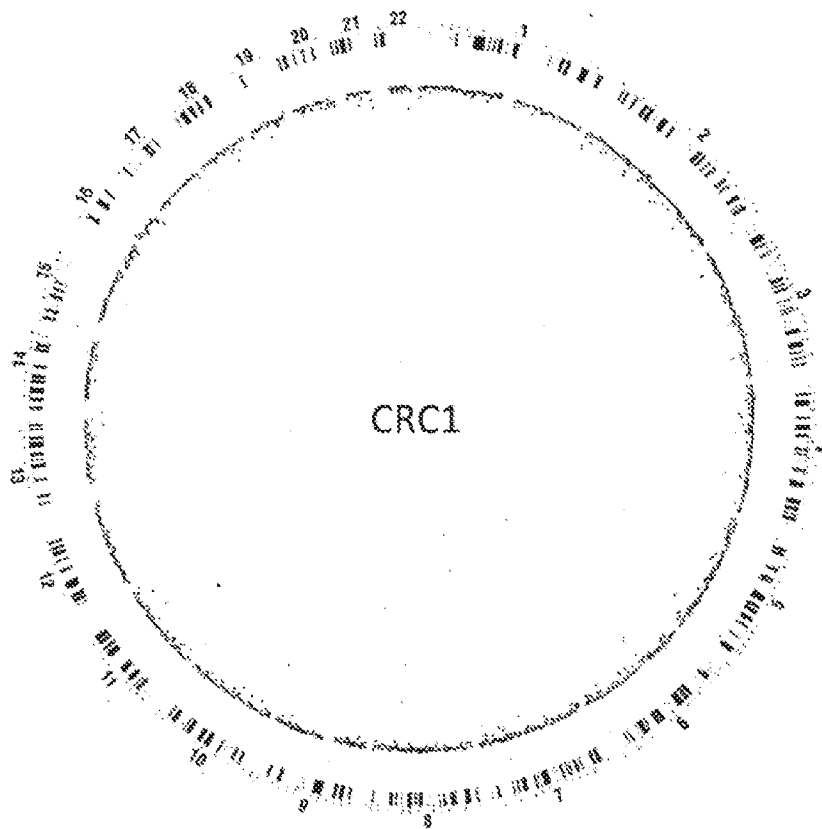


FIG. 27E

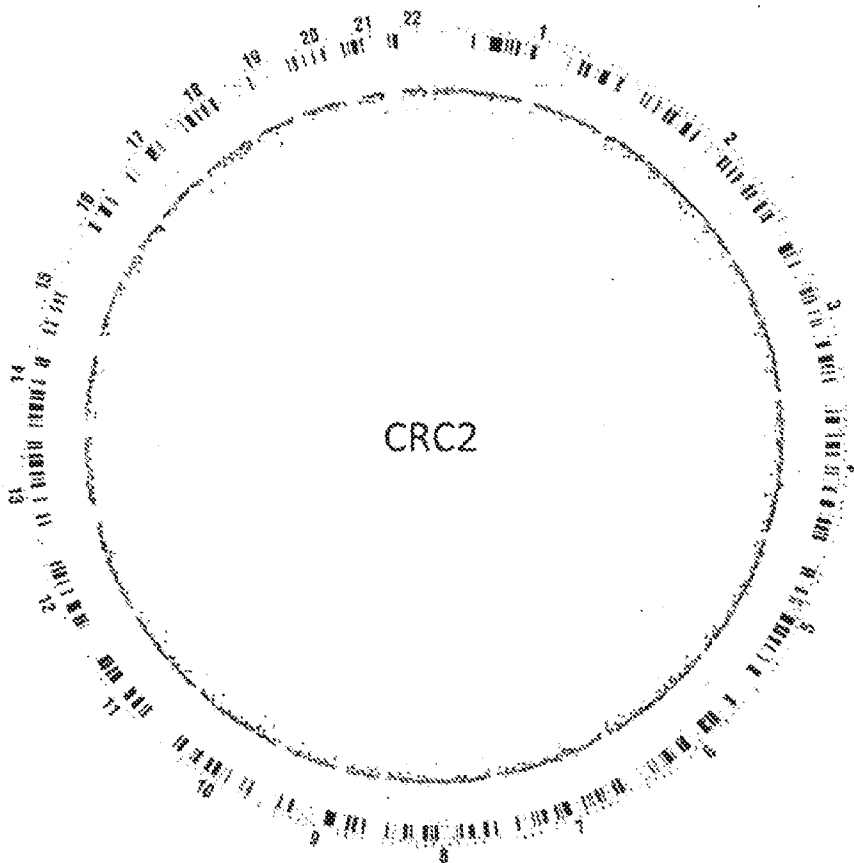


FIG. 27F

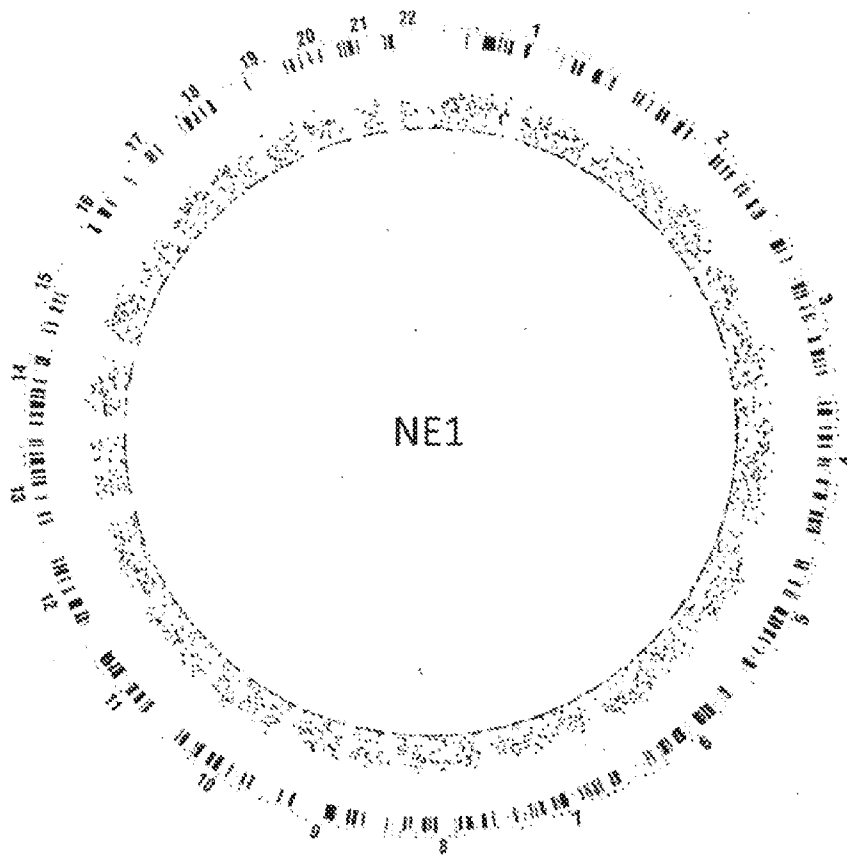


FIG. 27G

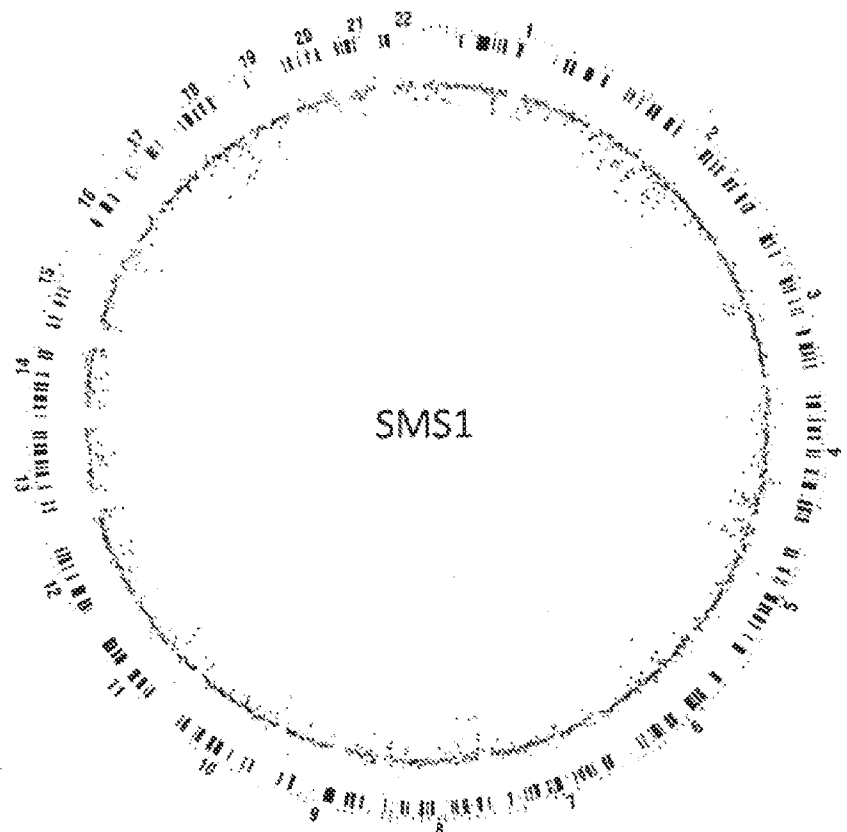


FIG. 27H

FIG. 27I

2780

	Recuento de lecturas sin procesar	Recuento de lecturas mapeadas	Fragmentos no duplicados	Proporción de lecturas utilizables %	Profundidad por cadena	Cobertura de CpG %	Profundidad en sitios de CpG con lecturas cubiertas por cadena
CL1	120807588	91750018	76890830	64%	0,59	33,32%	2,87
CL2	101344850	76473792	66246435	65%	0,59	30,26%	2,79
NPC1	101883643	75281330	57236188	56%	0,51	27,87%	2,72
NPC2	102745073	74211781	56841611	55%	0,51	28,74%	2,70
CRC1	87732645	66456695	56112543	64%	0,50	27,38%	2,63
CRC2	109955730	82594007	67253101	61%	0,60	29,91%	2,90
NE1	105514000	79923333	64534726	61%	0,58	27,86%	2,58
SMS1	94727779	71698718	64426529	68%	0,58	29,23%	2,79

Puntuación Z	CL1	CL2	NPC1	NPC2	CRC1	CRC2	NE1	SMS1
< -3	34,9%	33,9%	17,6%	54,3%	13,5%	9,8%	98,1%	11,2%
-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

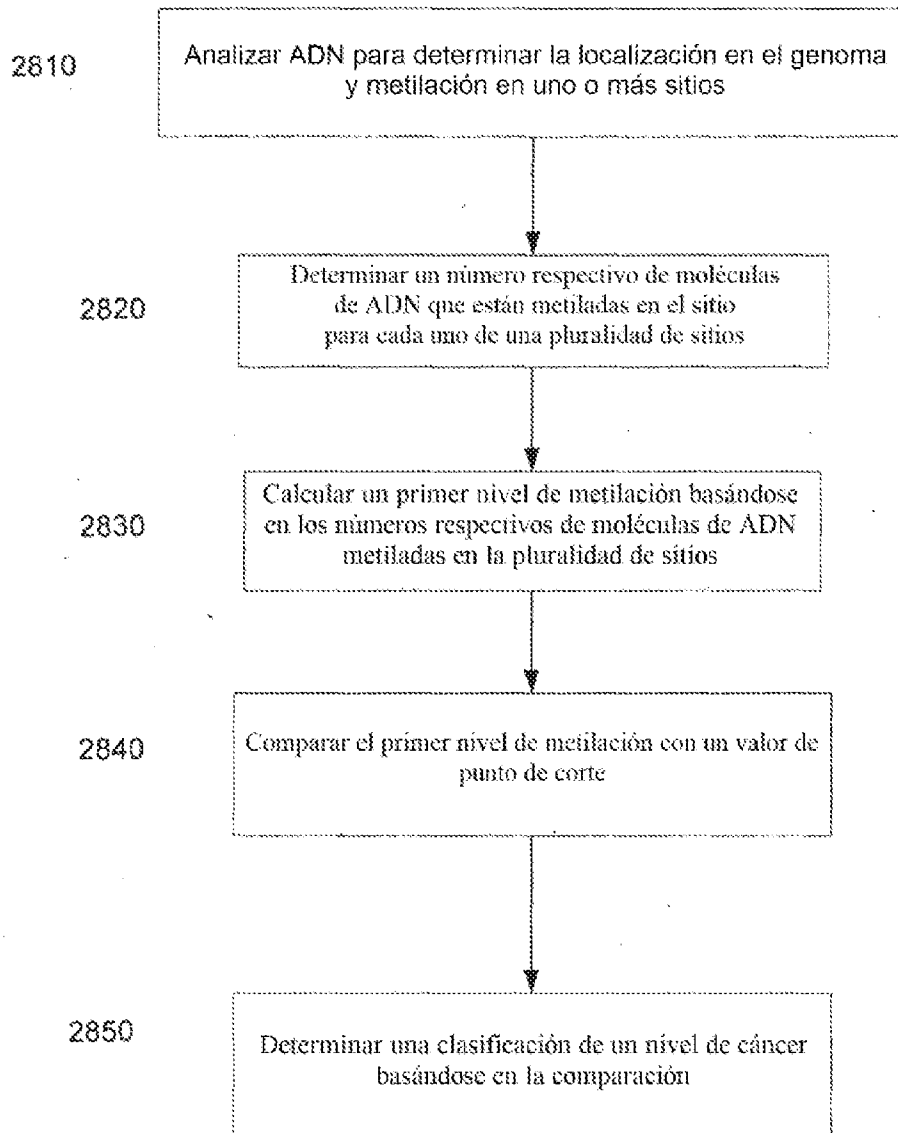
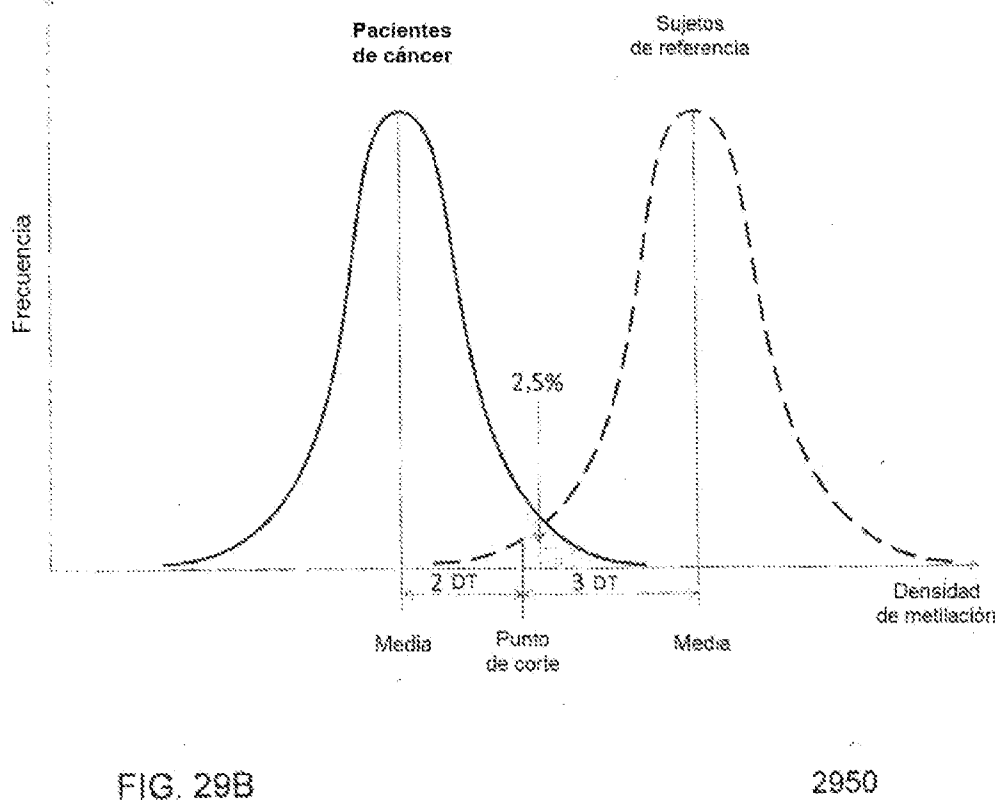
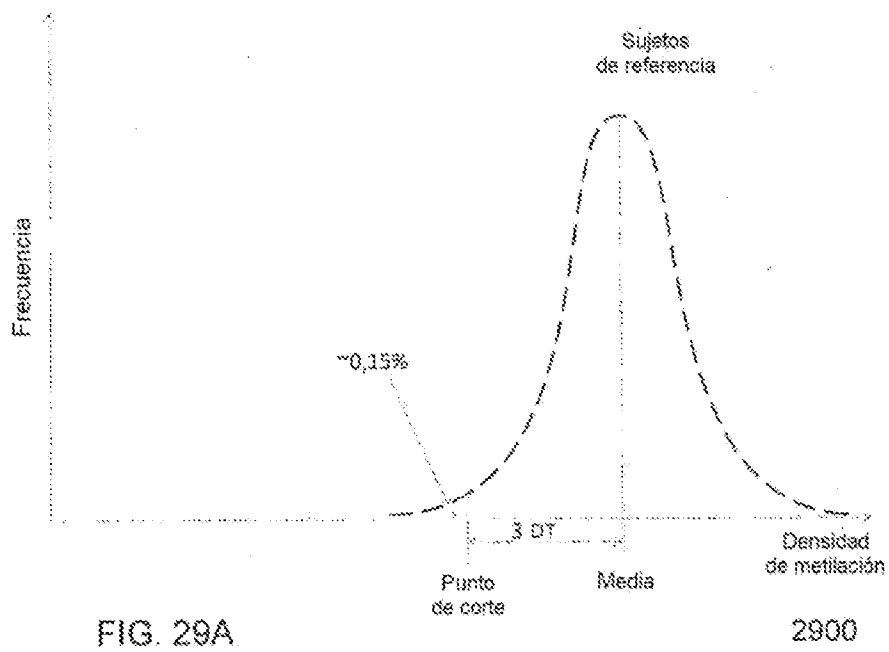


FIG. 28

2800



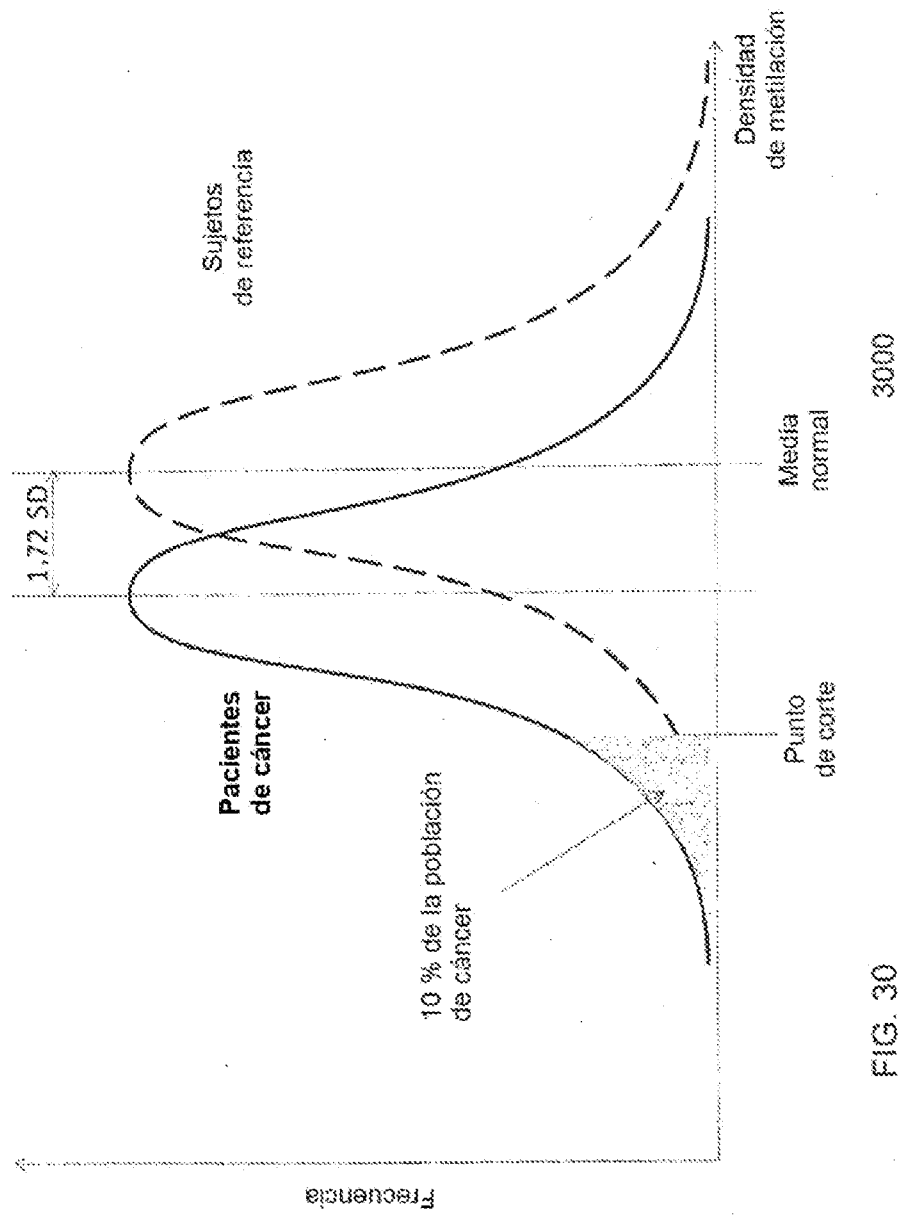
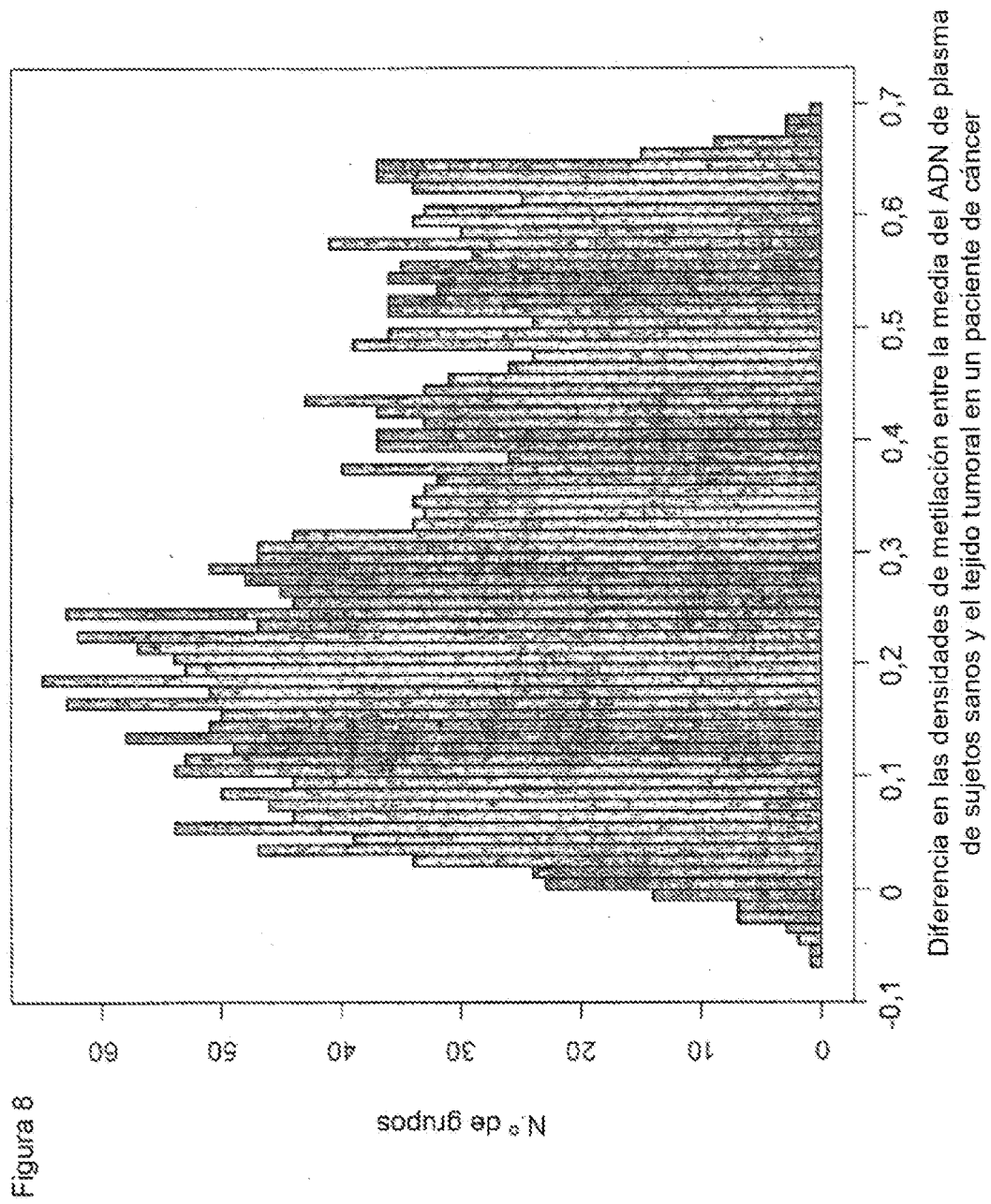


FIG. 30



Factor de dilución en profundidad a 5 %	Profundidad de secuenciación (factor de genoma haploide)	N.º de grupos hipometilados (puntuaciones z < -3) en plasma preoperatorio	% de grupos 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 en profundidad a 2 %	Profundidad de secuenciación (factor de genoma haploide)	N.º de grupos hipometilados (puntuaciones z < -3) en plasma preoperatorio	% de grupos 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	356	13,02
0,05	0,11	297	10,86
0,01	0,022	231	8,45

FIG. 32A

3200

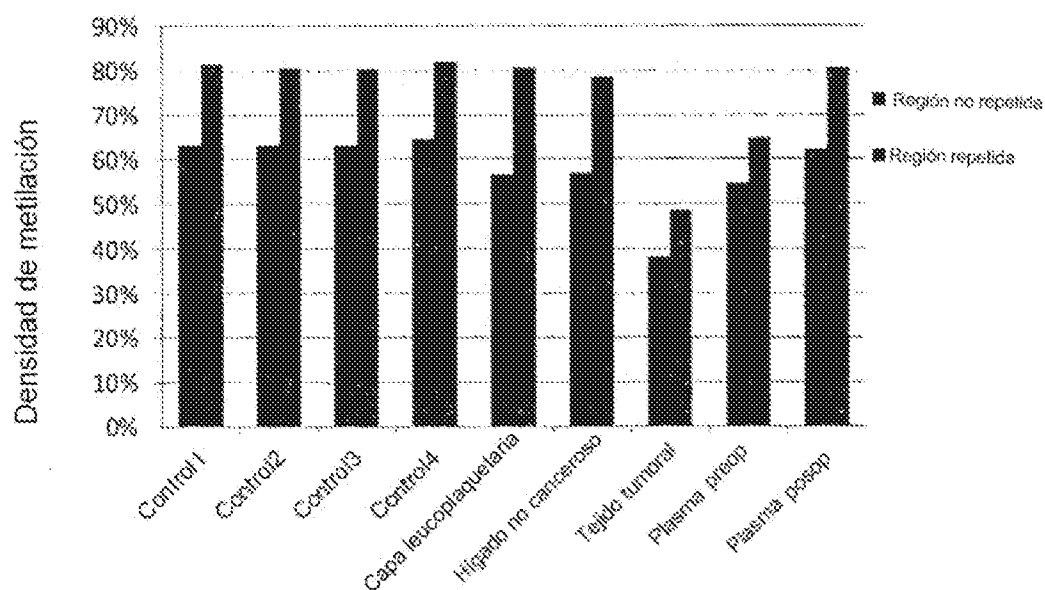


FIG. 32B

3250

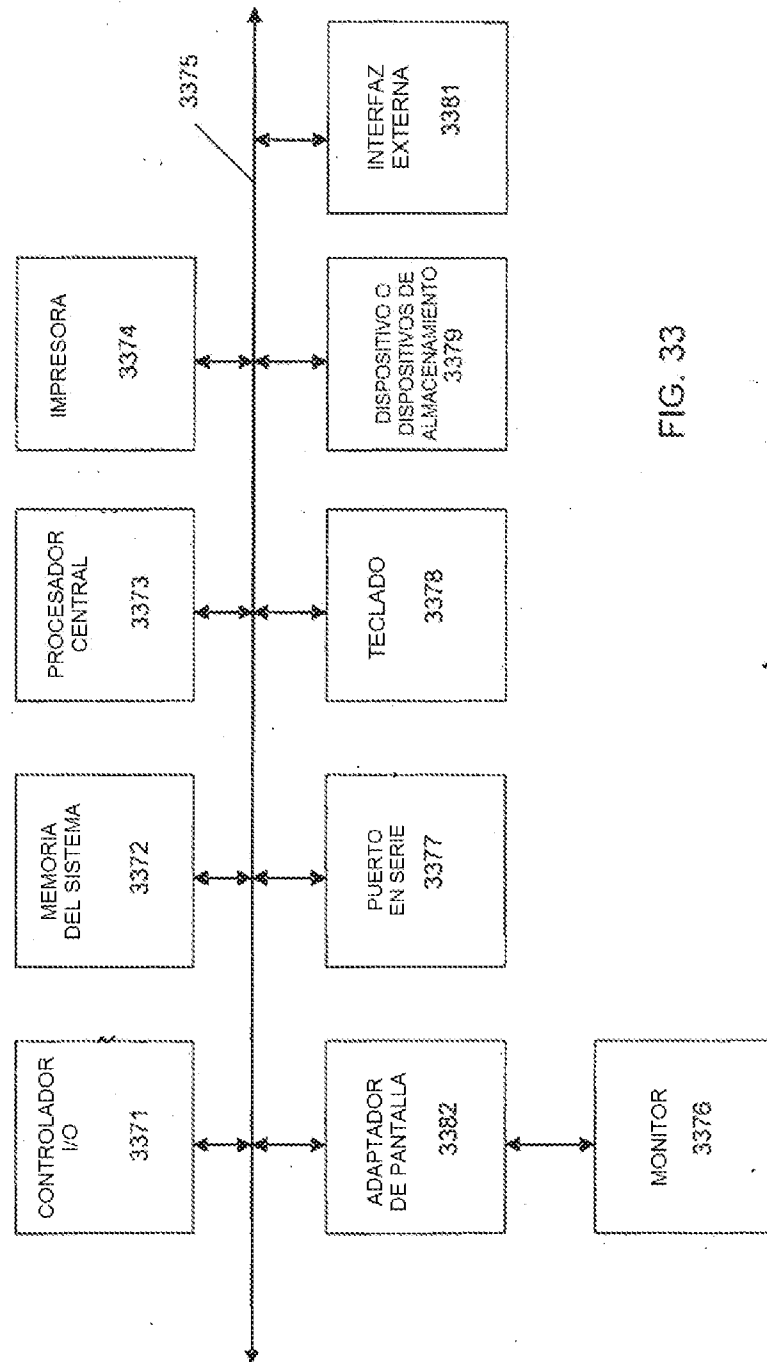


FIG. 33

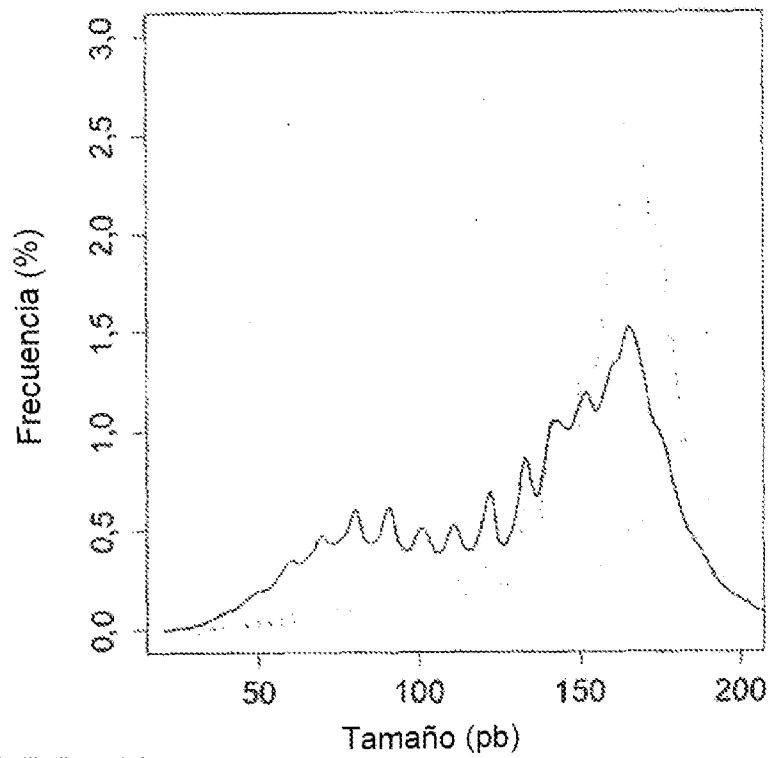


FIG. 34A

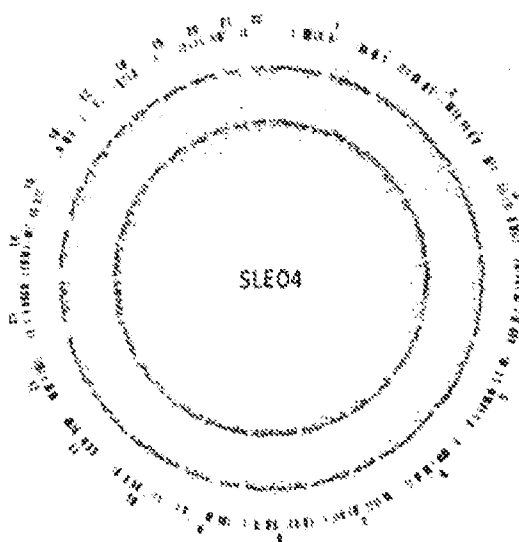


FIG. 34B

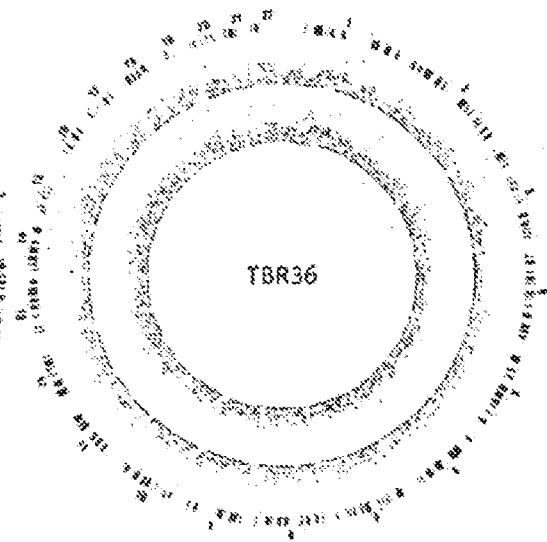


FIG. 34C

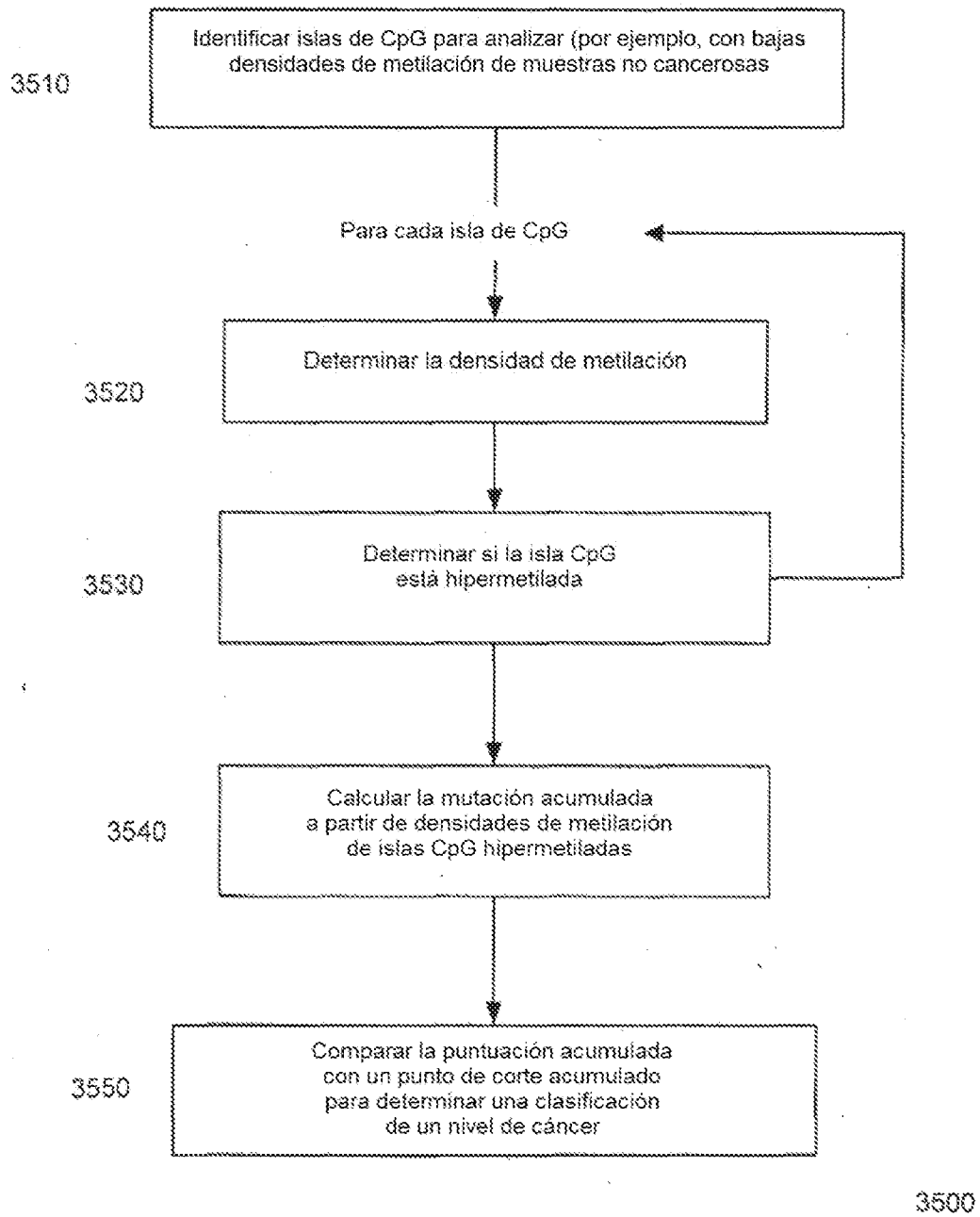


FIG. 35

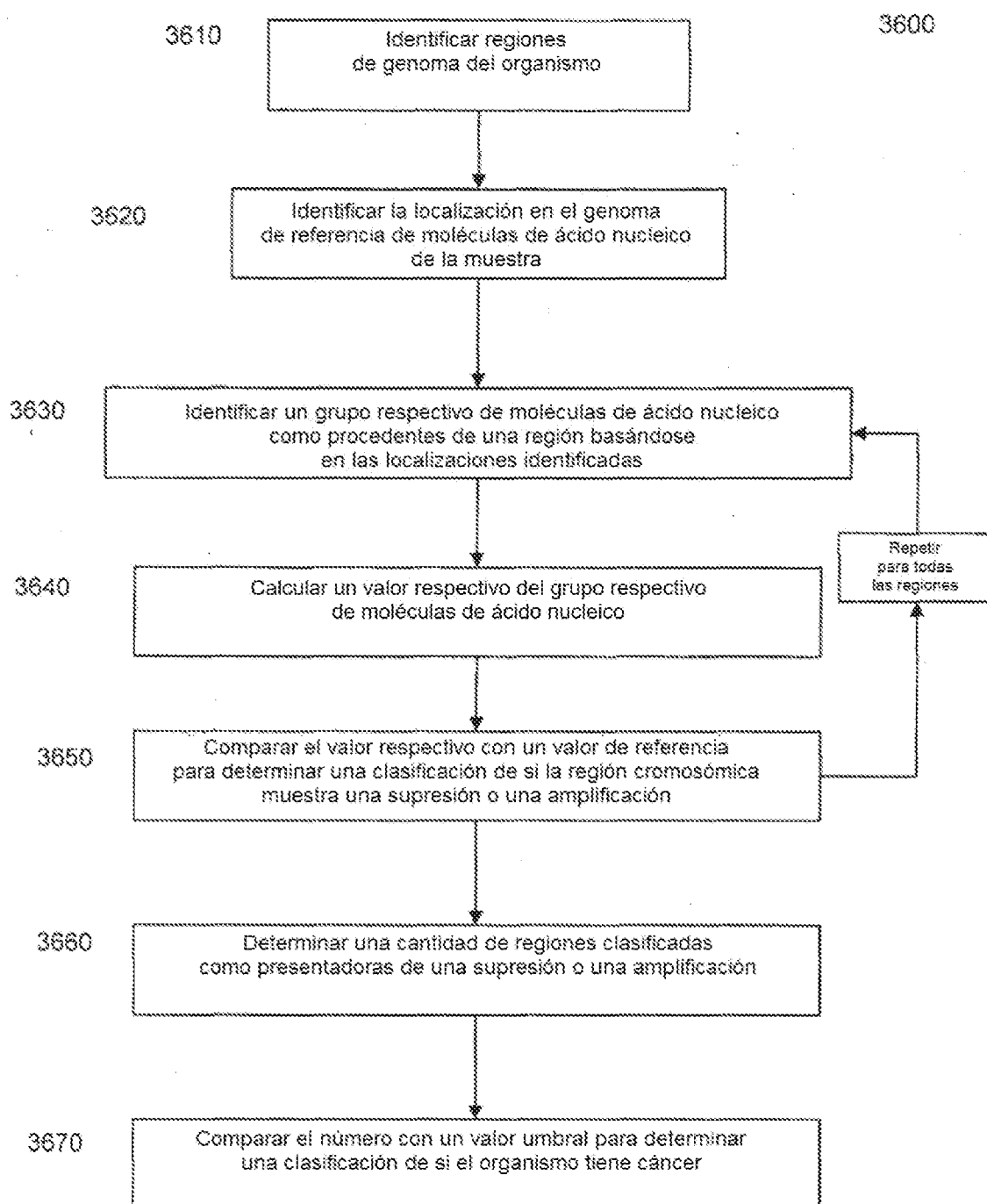


FIG. 36

Muestra: TBR36
Fracción de ADN tumoral: 40 %

De dentro a fuera:
1) ANC tumoral
2) ANC de plasma (sin BS)
3) ANC de plasma (BS)

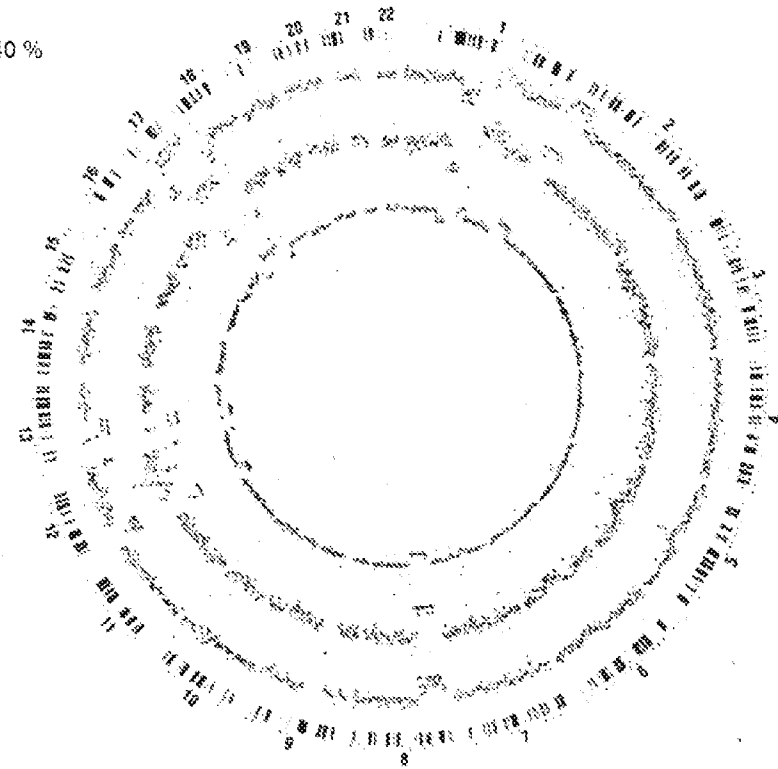


FIG. 37A

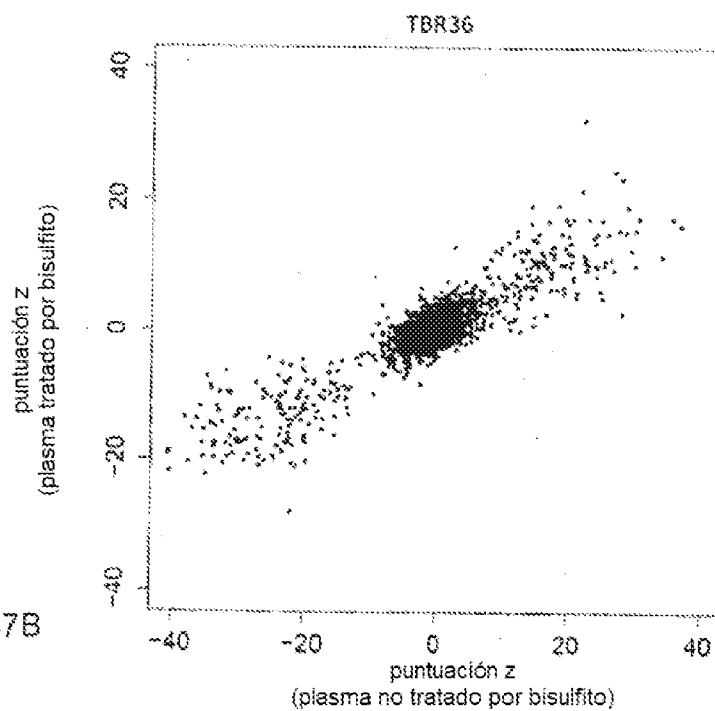


FIG. 37B

Muestra: TBR34
Fracción de ADN tumoral: 30 %

De dentro a fuera:
1) ANC tumoral
2) ANC de plasma (sin BS)
3) ANC de plasma (BS)

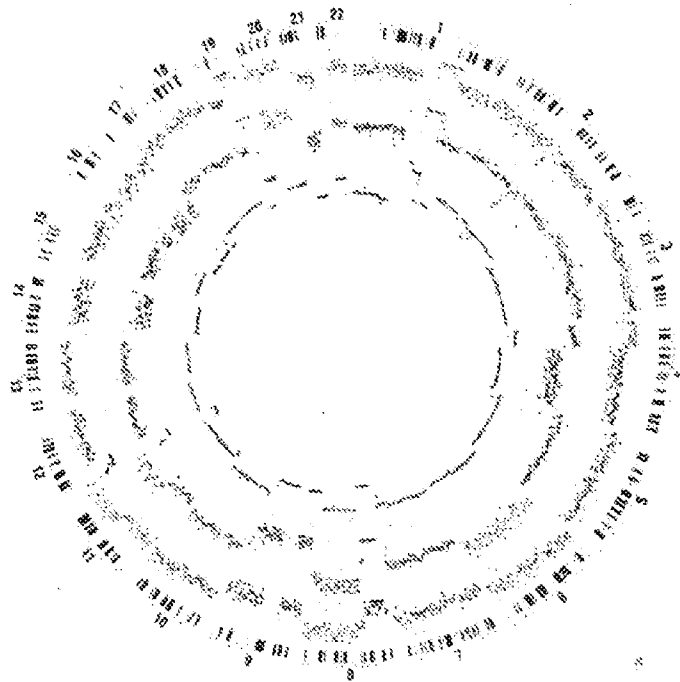


FIG. 38A

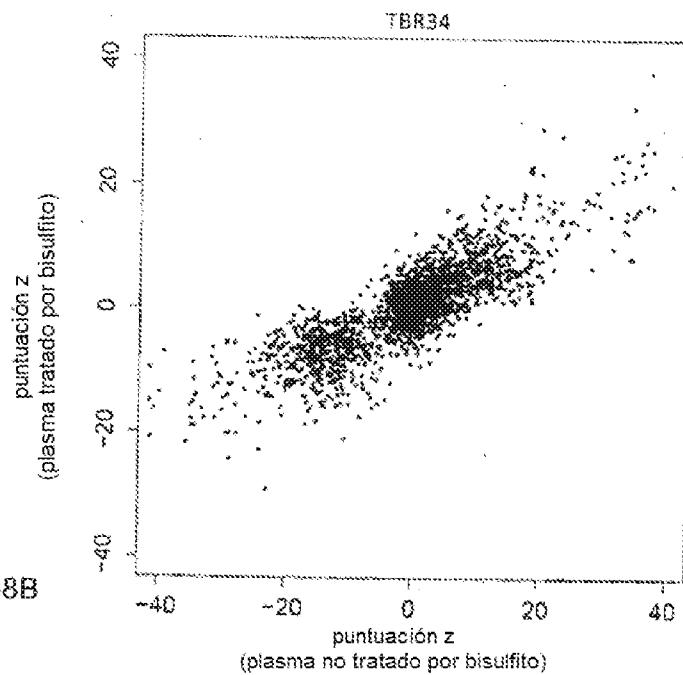


FIG. 38B

Muestra: TBR240
Tamaño tumoral: 2,8 cm x 2,3 cm

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



FIG. 39A

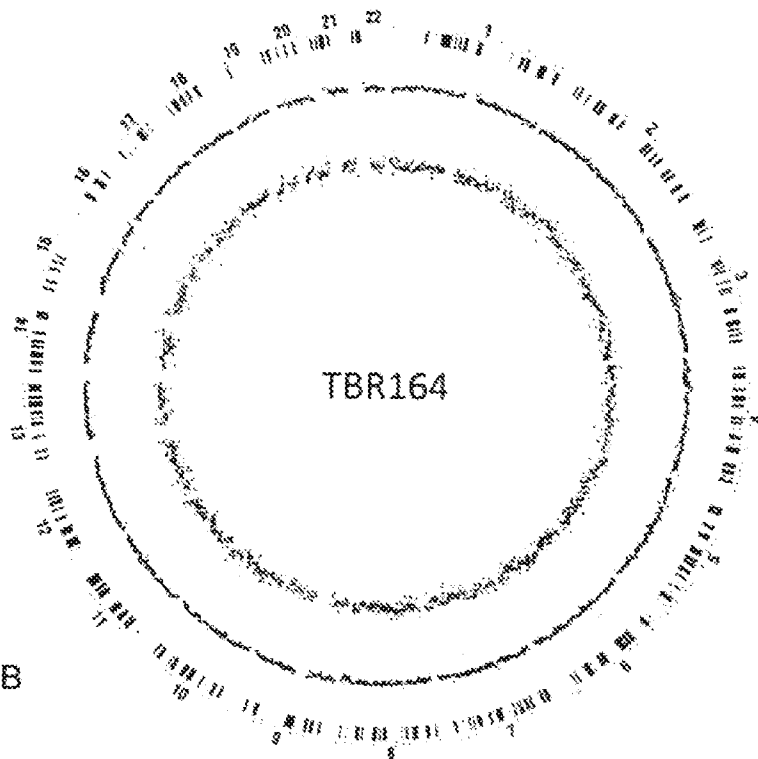


FIG. 39B

Anillo interno:
Plasma pretratamiento

Anillo externo:
Plasma postratamiento

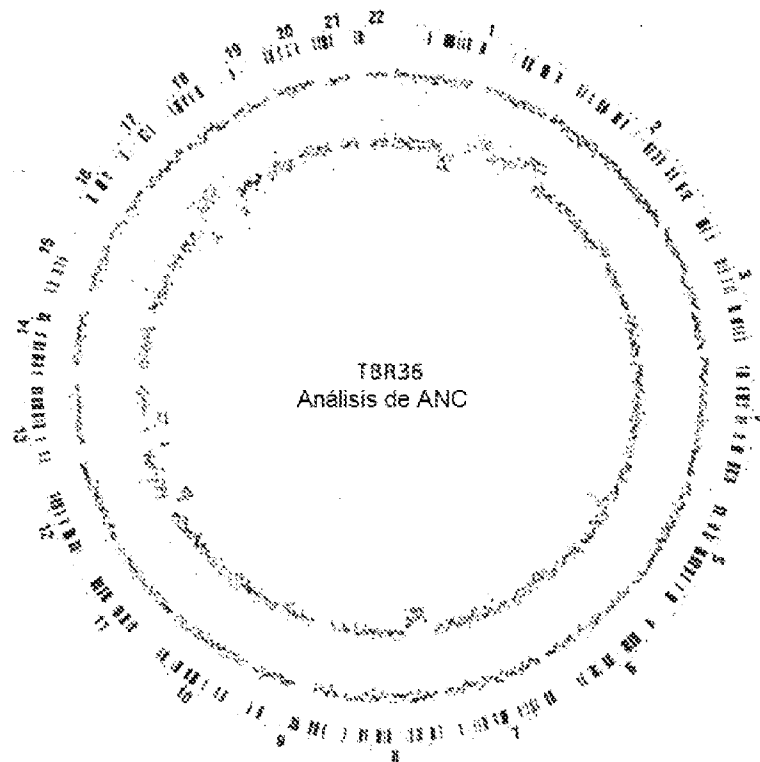


FIG. 40A

Anillo interno:
Plasma pretratamiento

Anillo externo:
Plasma postratamiento

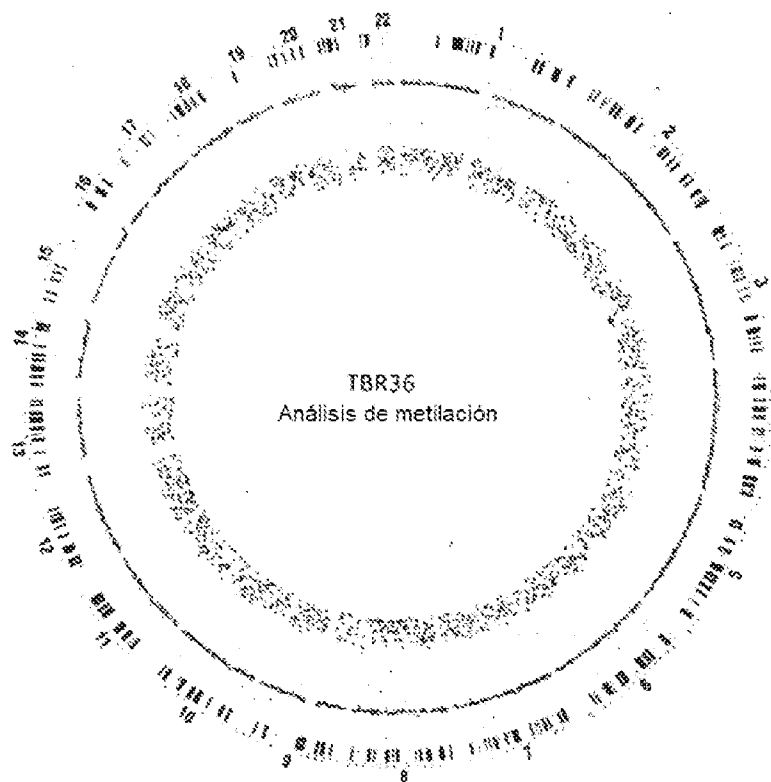


FIG. 40B

Muestra: TBR34
Fracción de ADN tumoral: 30 %
De fuera a dentro:
1) CNV posop
2) CNV preop

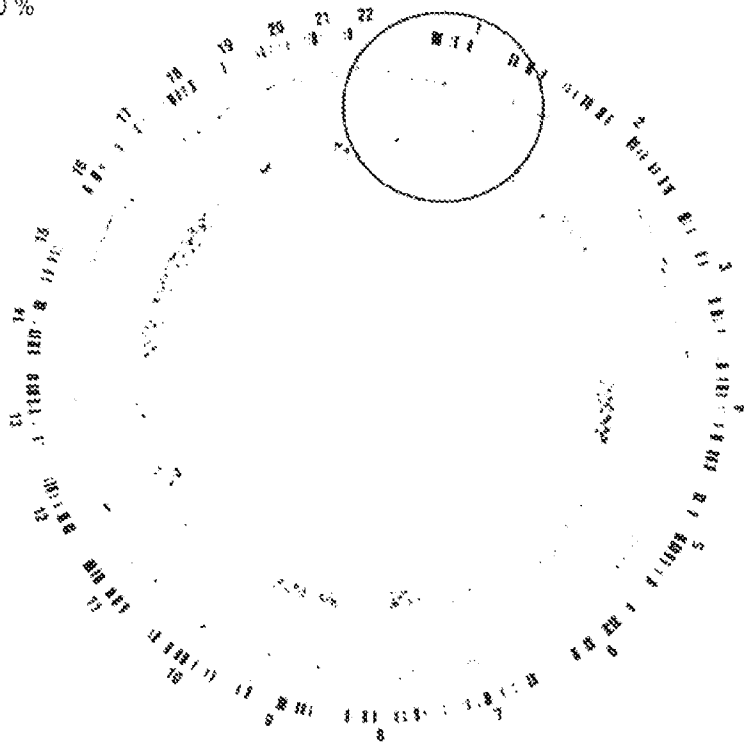


FIG. 41A

Anillo interno:
Plasma pretratamiento
Anillo externo:
Plasma postratamiento

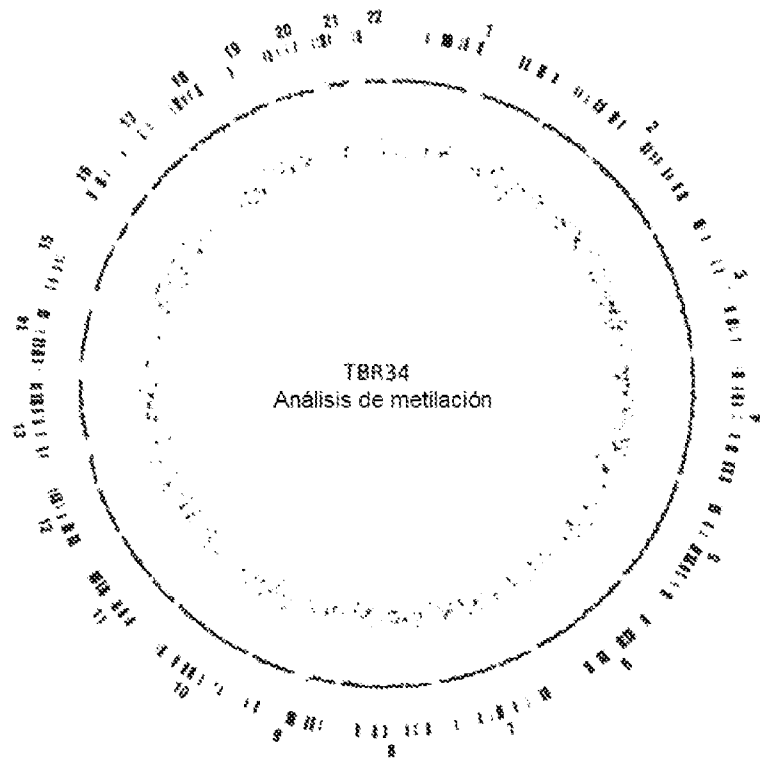


FIG. 41B

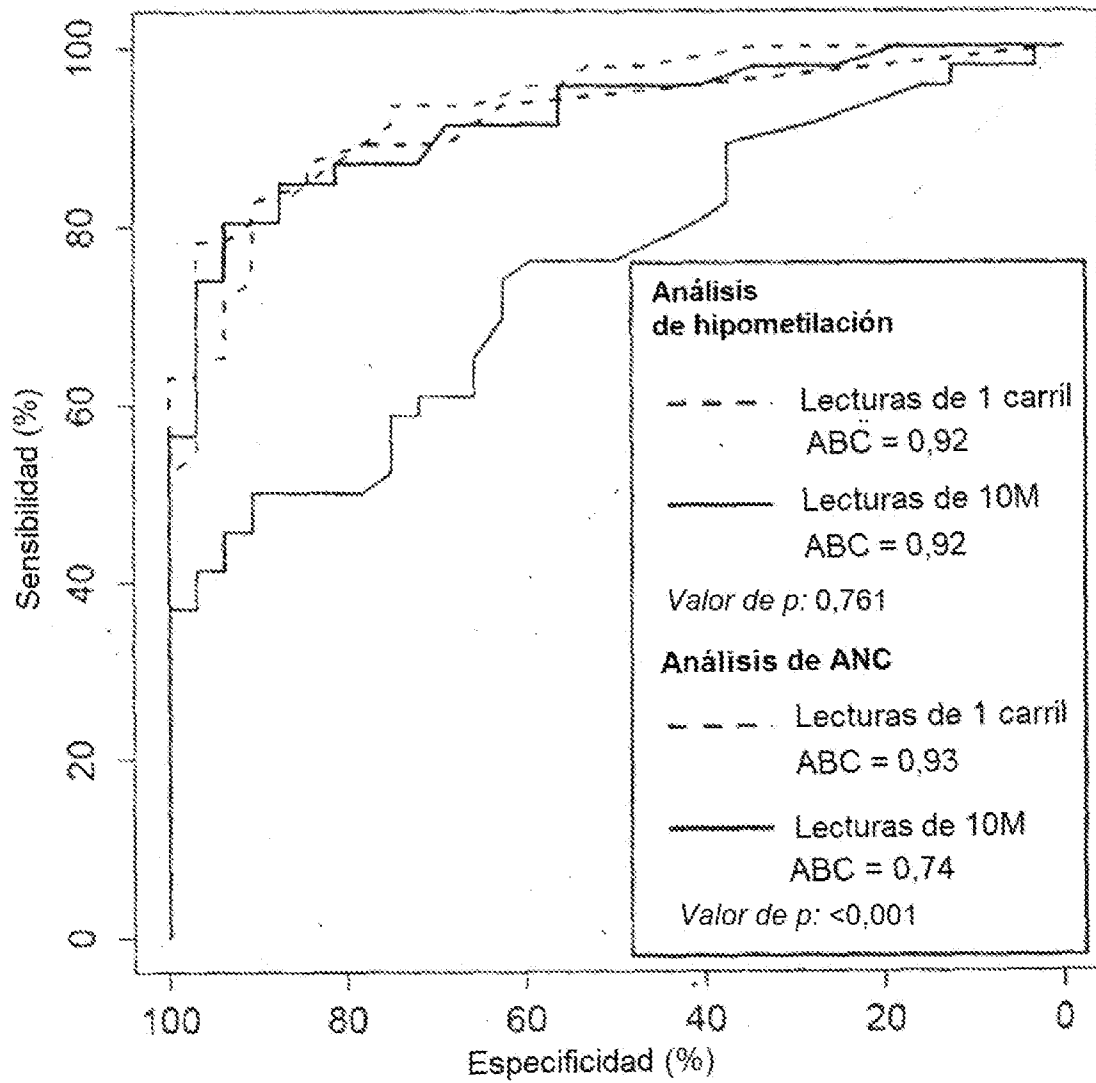


FIG. 42

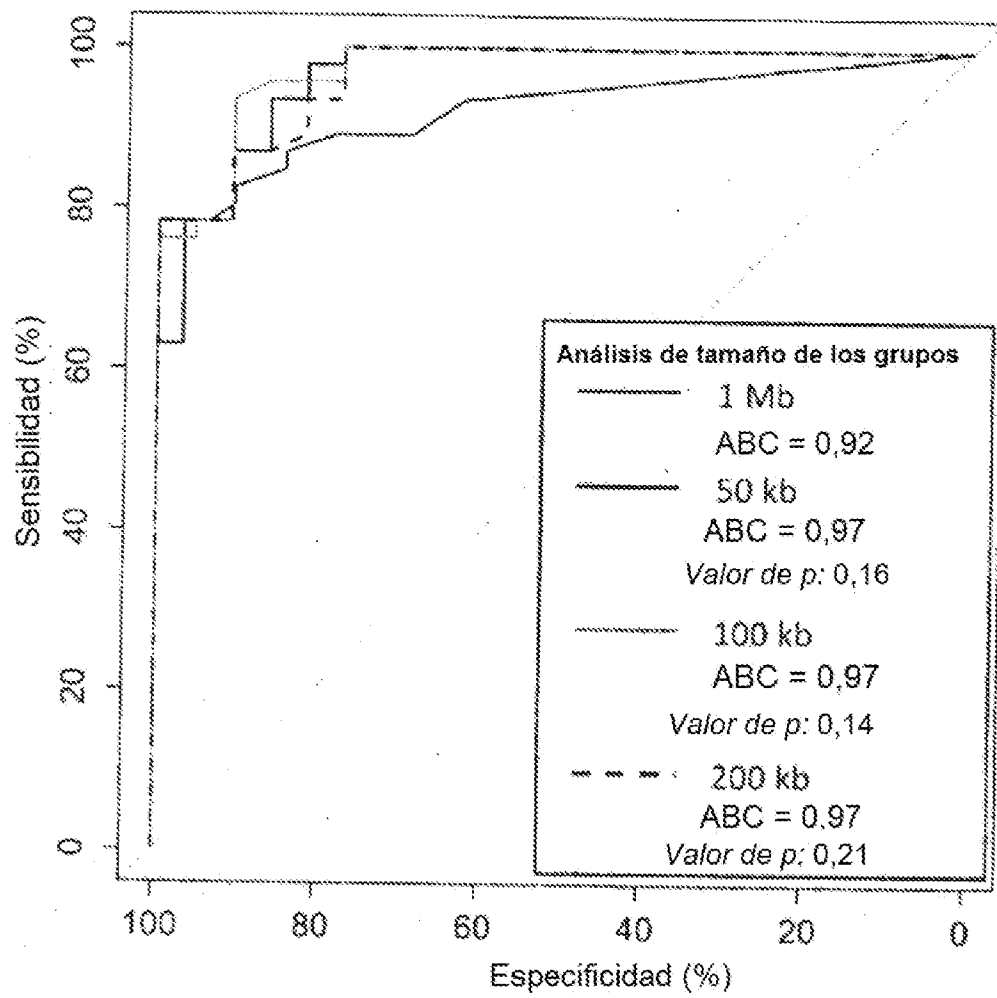
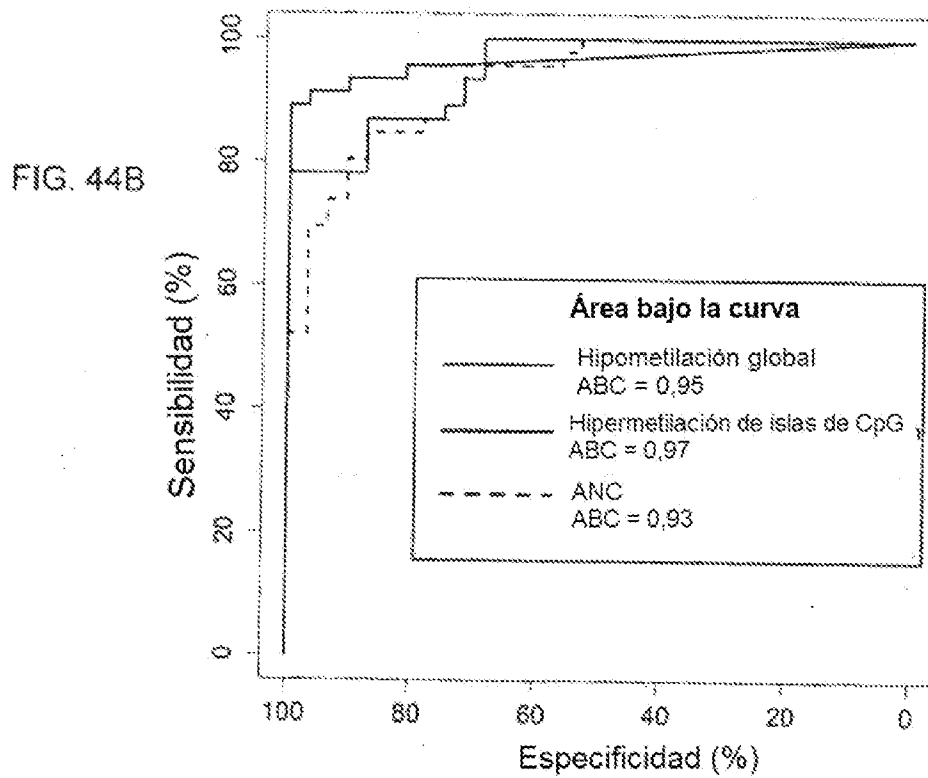
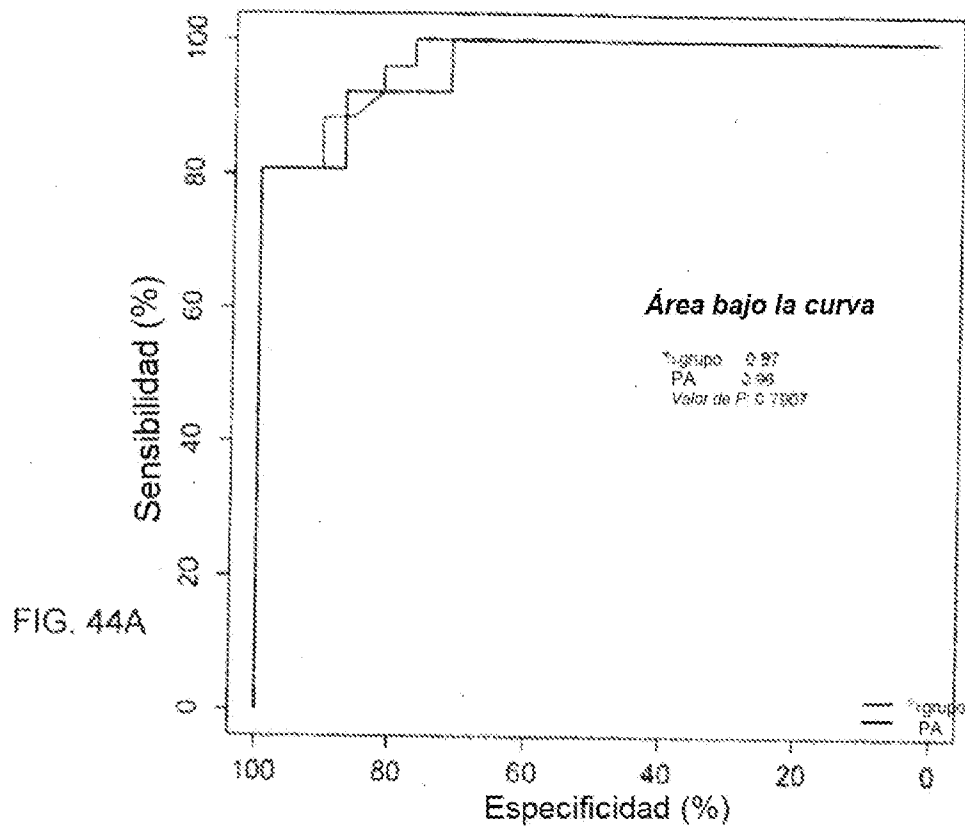


FIG. 43



N.º de caso	Estadio de BCLC	Dimensión mayor del tumor o los tumores (cm)	Puntuación de PA			Presencia de cualquier tipo de cambios
			Hipometilación general	Hipermetilación de islas CpG	ANC	
TBR36	A1	16	<u>88585</u>	<u>80,9</u>	<u>3777</u>	Positivo
TBR34	A1	13	<u>37573</u>	<u>221,7</u>	<u>8650</u>	Positivo
HOT198	A1	12	<u>93132</u>	<u>52,5</u>	<u>6803</u>	Positivo
HOT205	A1	6	<u>1146</u>	<u>1,7</u>	<u>175</u>	Positivo
HOT192	A1	5,5	<u>25914</u>	<u>119,5</u>	<u>1668</u>	Positivo
HOT227	A1	5,5	<u>36333</u>	<u>155,4</u>	<u>5008</u>	Positivo
HOT238	A1	5,5	<u>12187</u>	<u>105,3</u>	<u>1978</u>	Positivo
HOT197	B	5,3	<u>1714</u>	<u>9,5</u>	<u>28</u>	Positivo
HOT156	A1	5	<u>16529</u>	<u>115,9</u>	<u>1947</u>	Positivo
HOT236	A1	3,7	<u>7993</u>	<u>5,9</u>	<u>85</u>	Positivo
HOT170	A1	3,5	<u>22271</u>	<u>81,9</u>	<u>1542</u>	Positivo
HOT229	A1	3,1	<u>383</u>	<u>5,4</u>	<u>71</u>	Positivo
HOT233	A1	3	<u>189</u>	<u>7,5</u>	<u>322</u>	Positivo
HOT240	A1	2,8	<u>1277</u>	<u>14,0</u>	<u>972</u>	Positivo
HOT222	A1	2,6	<u>18832</u>	<u>10,5</u>	<u>172</u>	Positivo
HOT162	A1	2,5	<u>26112</u>	<u>41,0</u>	<u>3048</u>	Positivo
HOT172	A1	2,5	<u>12909</u>	<u>98,1</u>	<u>1305</u>	Positivo
HOT164	A1	2,3	<u>12604</u>	<u>4,8</u>	<u>210</u>	Positivo
HOT215	A1	2,3	<u>25824</u>	<u>54,8</u>	<u>714</u>	Positivo
HOT207	A1	2,1	<u>3377</u>	<u>12,0</u>	<u>101</u>	Positivo
HOT224	A1	2	<u>367</u>	<u>0,0</u>	<u>13</u>	Negativo
HOT159	A1	1,5	<u>1367</u>	<u>1,4</u>	<u>26</u>	Positivo
HOT204	A1	1,5	<u>102</u>	<u>4,5</u>	<u>198</u>	Positivo
HOT151	A1	1,5	<u>153</u>	<u>2,8</u>	<u>9</u>	Negativo
HOT208	A1	1,2	<u>5867</u>	<u>4,7</u>	<u>214</u>	Positivo
HOT167	A1	1	<u>21544</u>	<u>4,5</u>	<u>54</u>	Positivo

FIG. 45

Tipo de cáncer	N.º de caso	Estadificación/ dimensión máxima del tumor	Puntuación de PA			Presencia de cualquier tipo de cambios
			Hipometilación general	Hipermetilación de islas CpG	ANC	
Cáncer de mama	TBR150	Metastásico	<u>1481</u>	<u>146,7</u>	<u>5123</u>	Positivo
	TBR126	T2N1M0 4,5 cm	<u>10415</u>	<u>34,9</u>	<u>1951</u>	Positivo
	TBR127	T2N1M0 2,5 cm	<u>2598</u>	<u>4,3</u>	54	Positivo
	TBR128	T2N0M0 2 cm	157	<u>7,7</u>	<u>2258</u>	Positivo
	TBR111	T1N0M0 0,9 cm	293	0,0	128	Negativo
Sarcoma de músculo liso	TBR051	Metastásico	<u>4787</u>	<u>79,8</u>	<u>7986</u>	Positivo
Tumor neuroendocrino	TBR052	Metastásico	<u>112518</u>	<u>119,8</u>	<u>14681</u>	Positivo
Cáncer de pulmón	TBR164	Metastásico	<u>51818</u>	<u>47,3</u>	<u>4714</u>	Positivo
	TBR012	Metastásico	<u>4351</u>	<u>7,7</u>	<u>2213</u>	Positivo
	TBR014	Metastásico	<u>2356</u>	<u>12,2</u>	<u>825</u>	Positivo
	TBR177	T2N1M0 3,9 cm	<u>1072</u>	<u>10,6</u>	25	Positivo
Cáncer nasofaríngeo	TBR031	Metastásico	<u>1028</u>	<u>4,4</u>	<u>1860</u>	Positivo
	TBR125	Metastásico	<u>2004</u>	<u>162,5</u>	<u>6708</u>	Positivo
	TBR124	T4N1	270	<u>4,9</u>	37	Positivo
	TBR123	T1N2M0	<u>6871</u>	<u>11,8</u>	80	Positivo
	TBR062	T3N1M0	111	<u>30,9</u>	41	Positivo
	TBR108	T3N1M0	93	<u>7,7</u>	26	Positivo
	TBR107	T3N0M0	<u>2439</u>	<u>41,2</u>	205	Positivo
	TBR106	T1N1M0	<u>2281</u>	<u>5,9</u>	97	Positivo
	TBR099	T1N0M0	<u>6187</u>	<u>10,2</u>	26	Positivo

FIG. 46

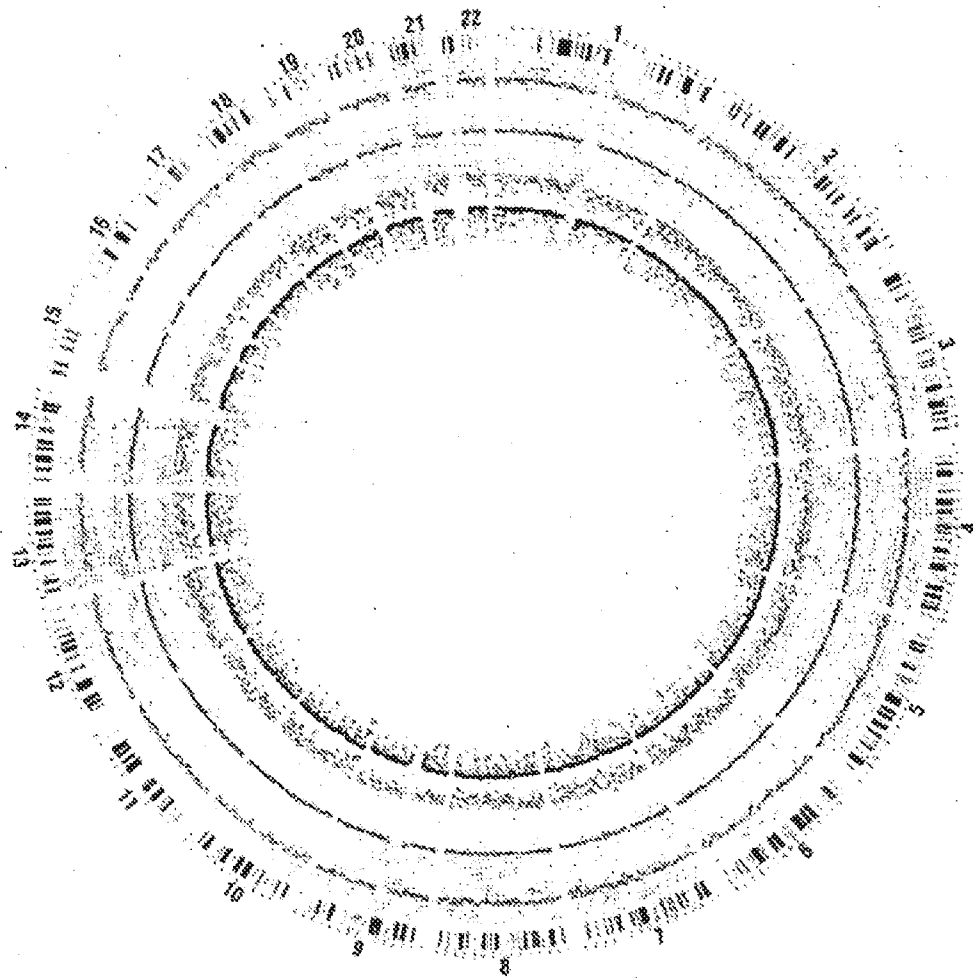


FIG. 47

Muestra: TBR36
Fracción de ADN tumoral: 40 %

De dentro a fuera
1) Puntuación z de ANC
2) Puntuación z de metilación

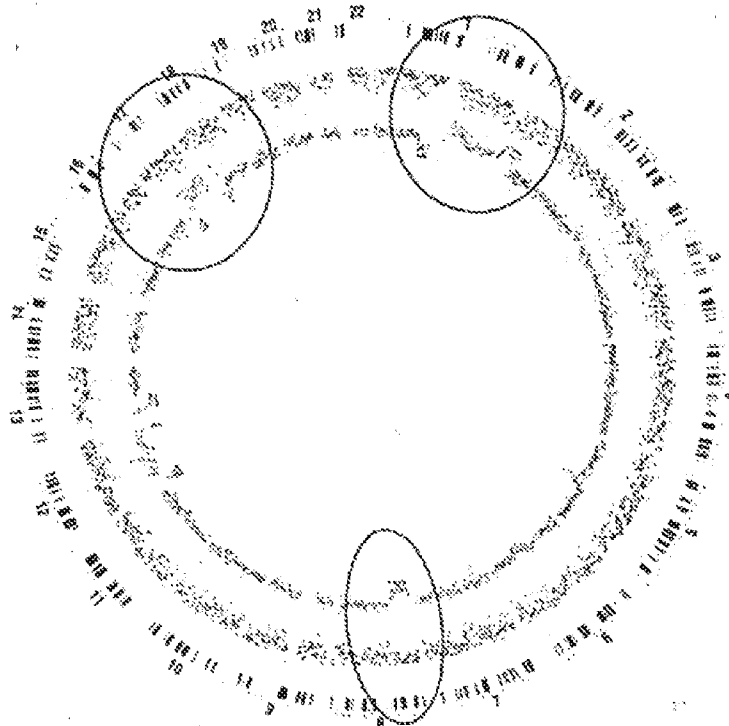


FIG. 48A

Muestra: TBR36
Fracción de ADN tumoral: 40 %

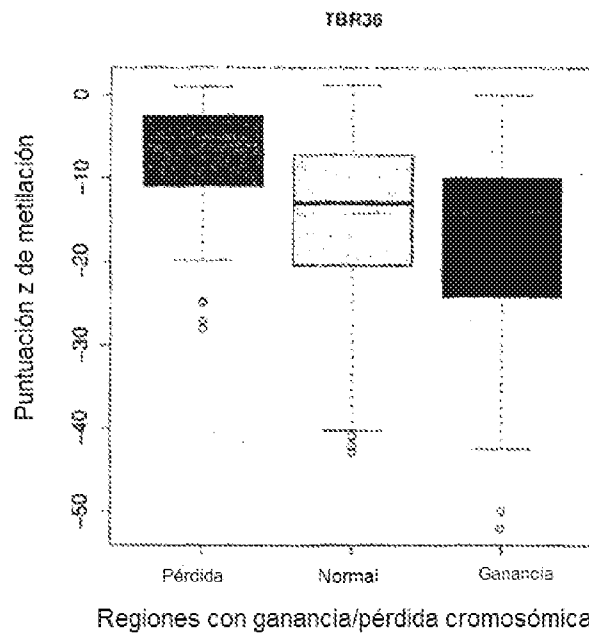


FIG. 48B

Muestra: TBR34
Fracción de ADN tumoral: 30 %

De dentro a fuera
1) Puntuación z de ANC
2) Puntuación z de metilación

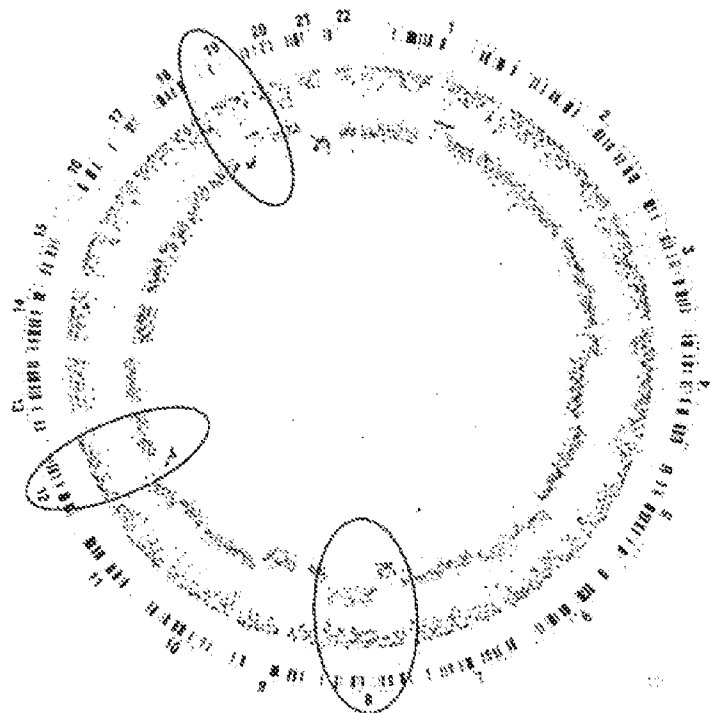


FIG. 49A

Muestra: TBR34
Fracción de ADN tumoral: 30 %

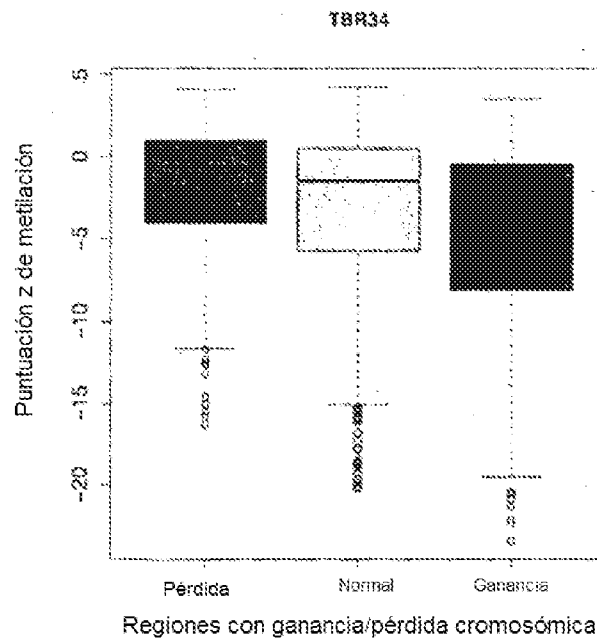


FIG. 49B

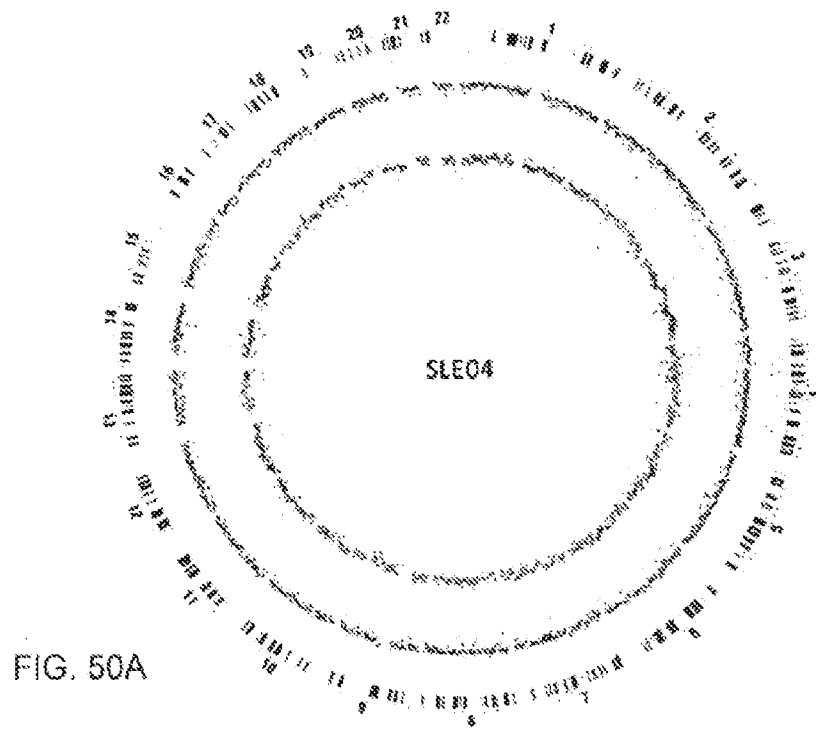


FIG. 50A

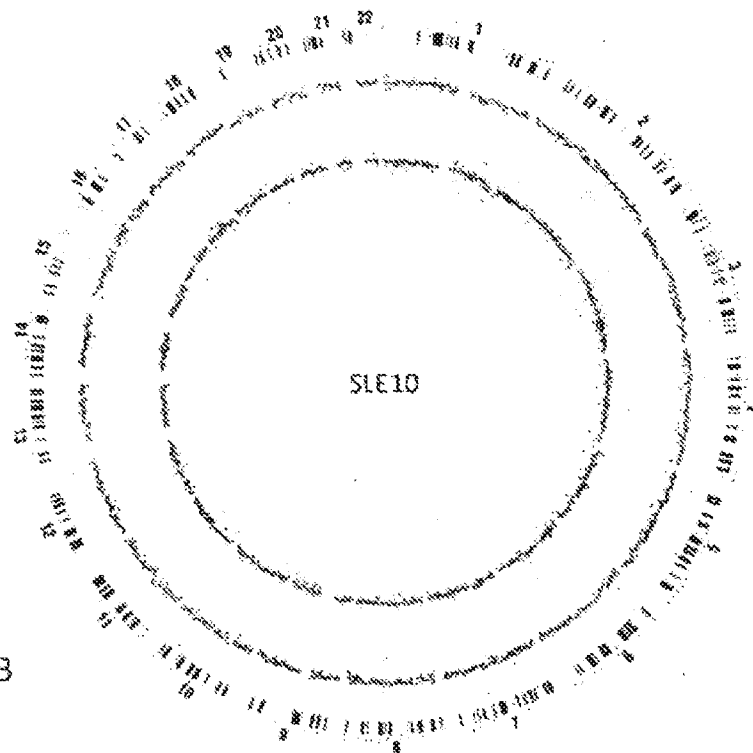


FIG. 50B

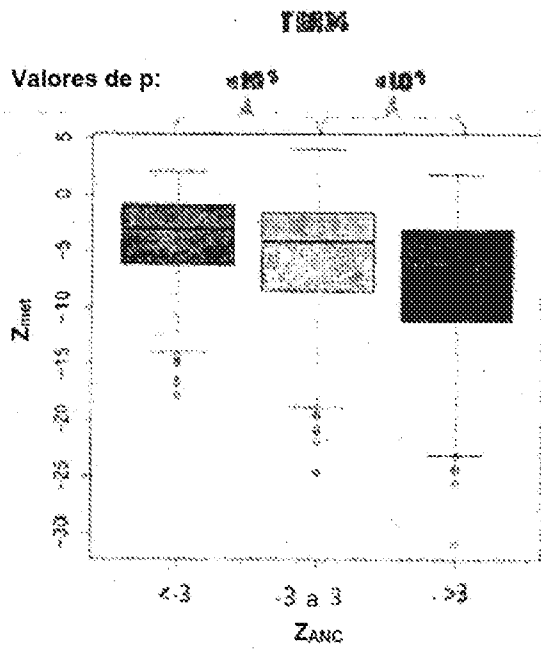


FIG. 51A

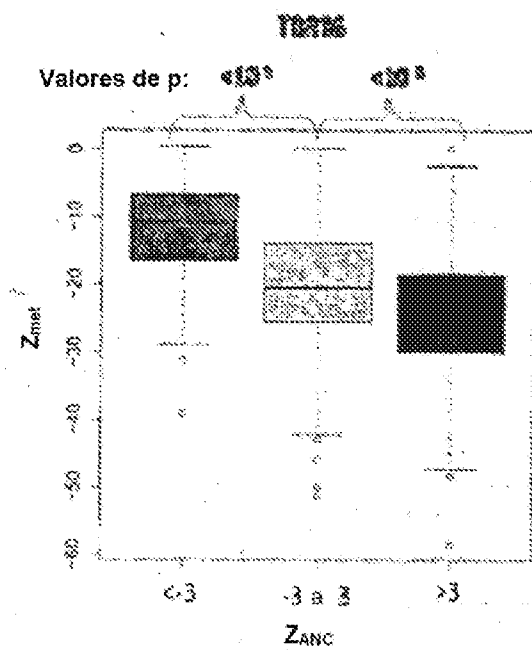


FIG. 51B

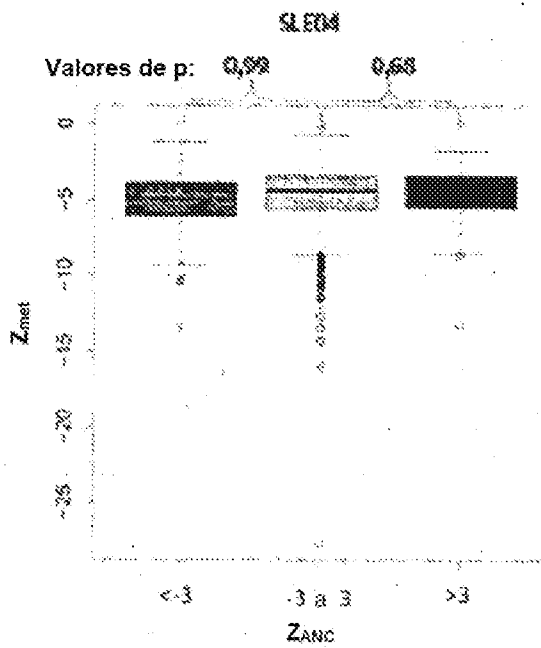


FIG. 51C

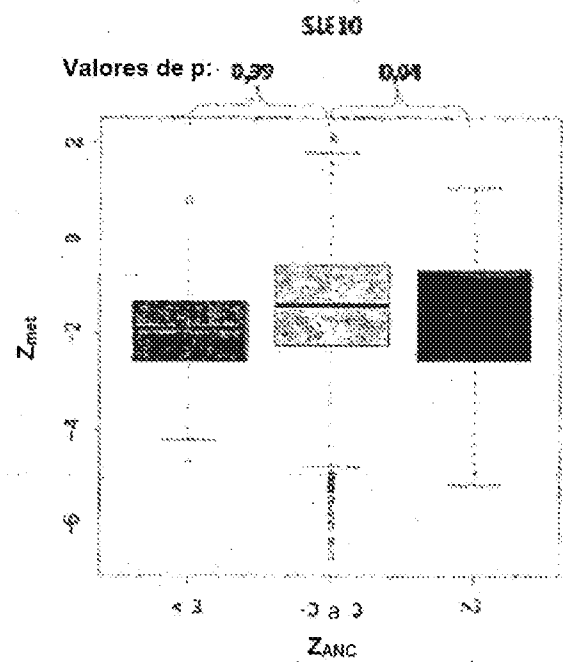


FIG. 51D

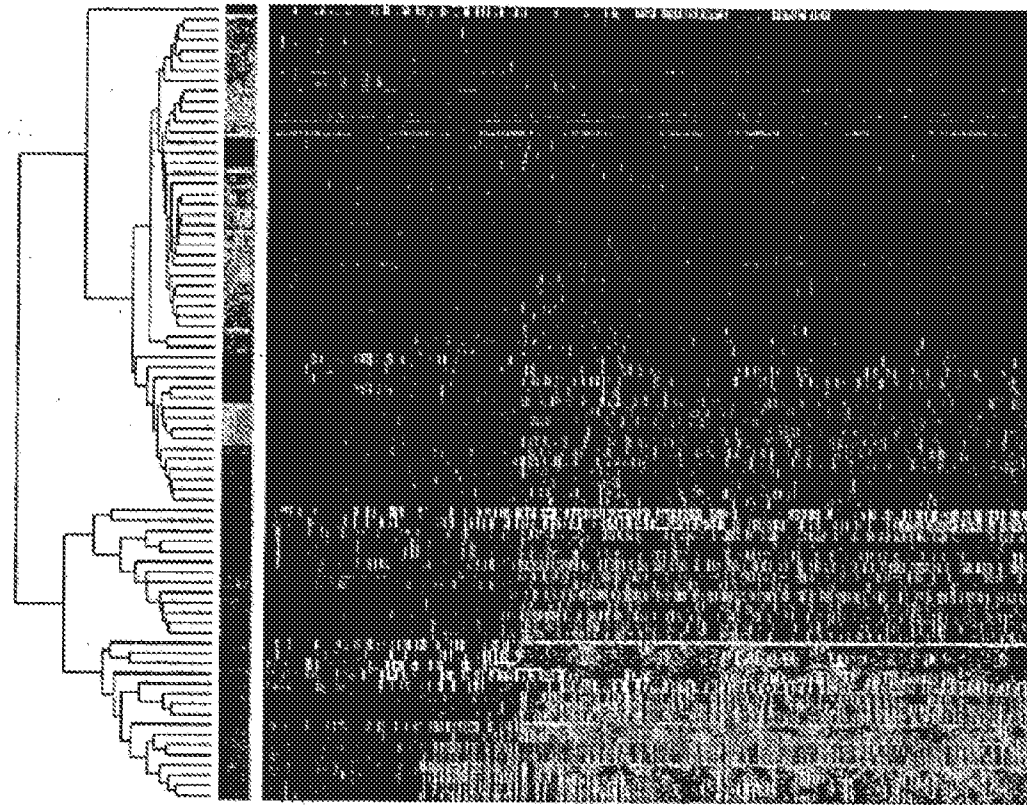


FIG. 52B

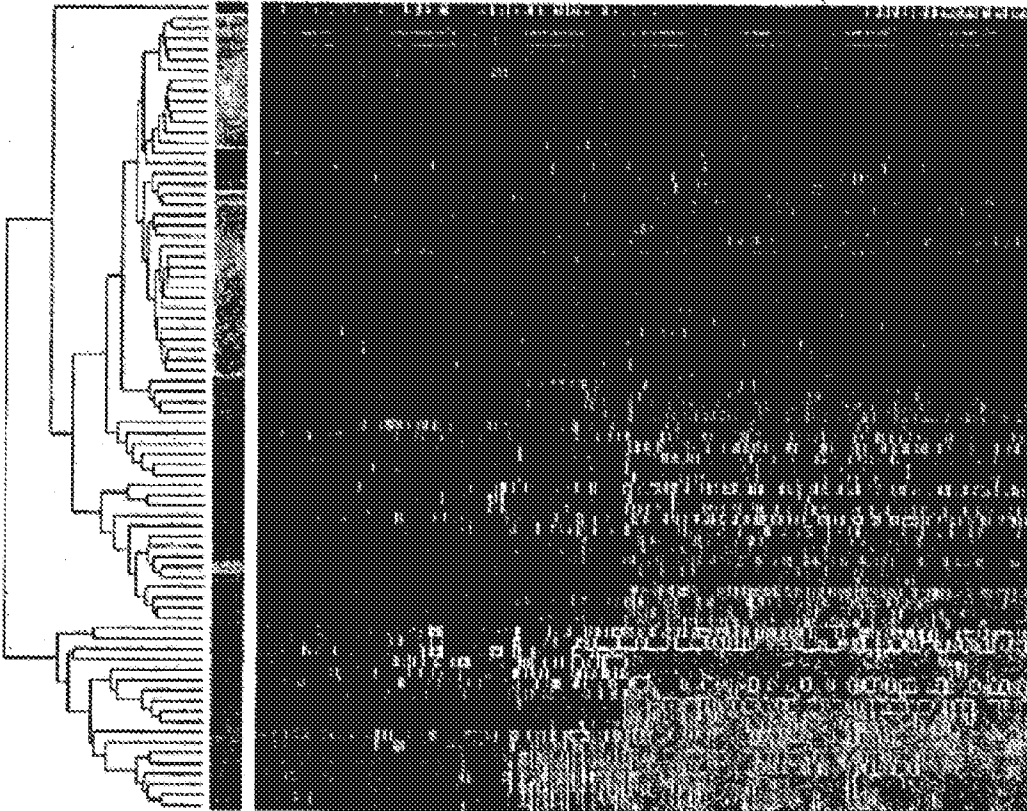


FIG. 52A

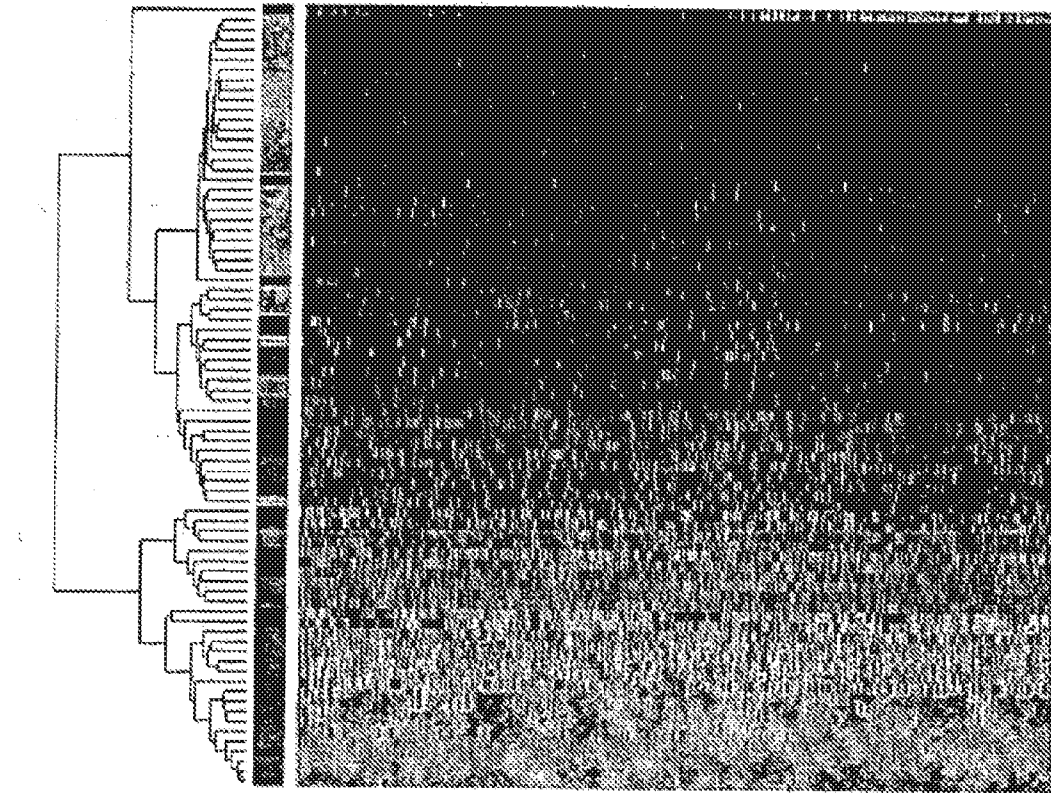


FIG. 53B

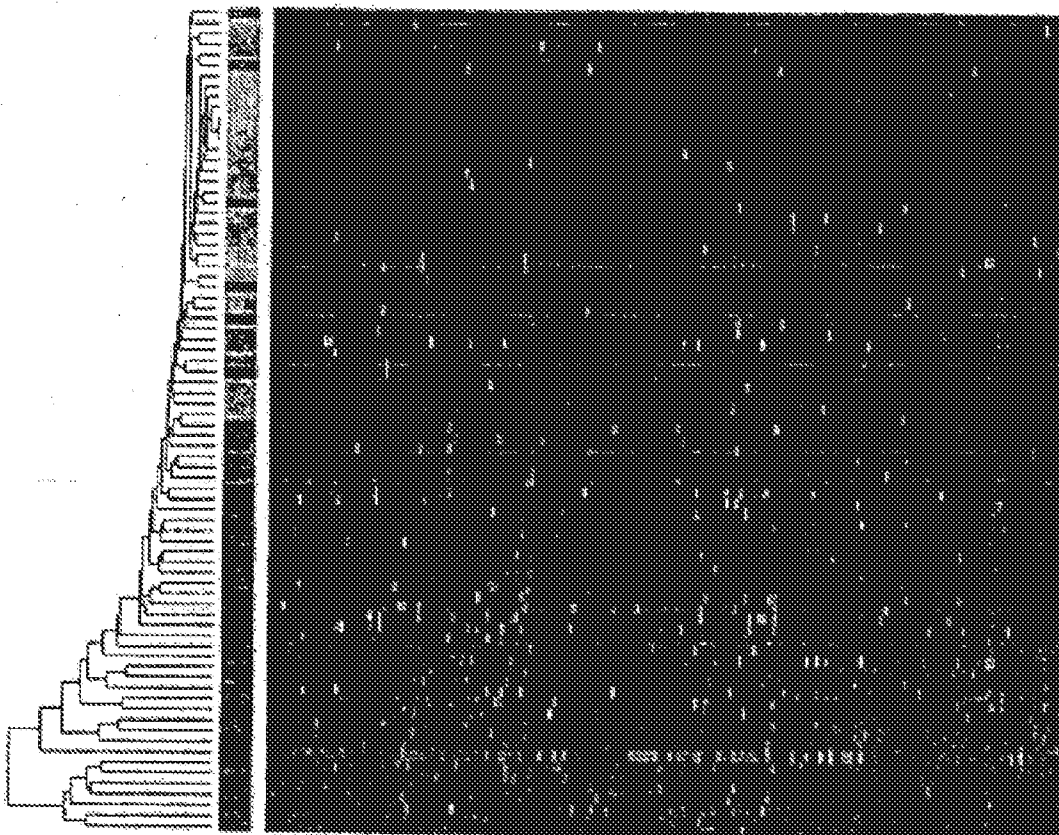


FIG. 53A

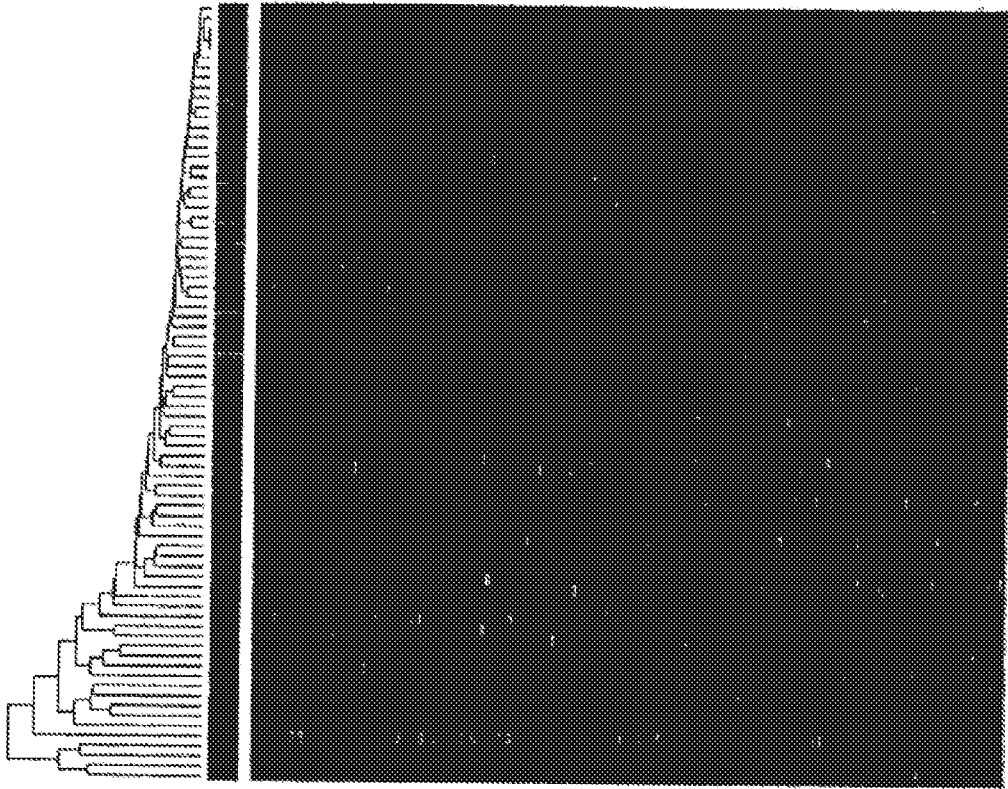


FIG. 54B

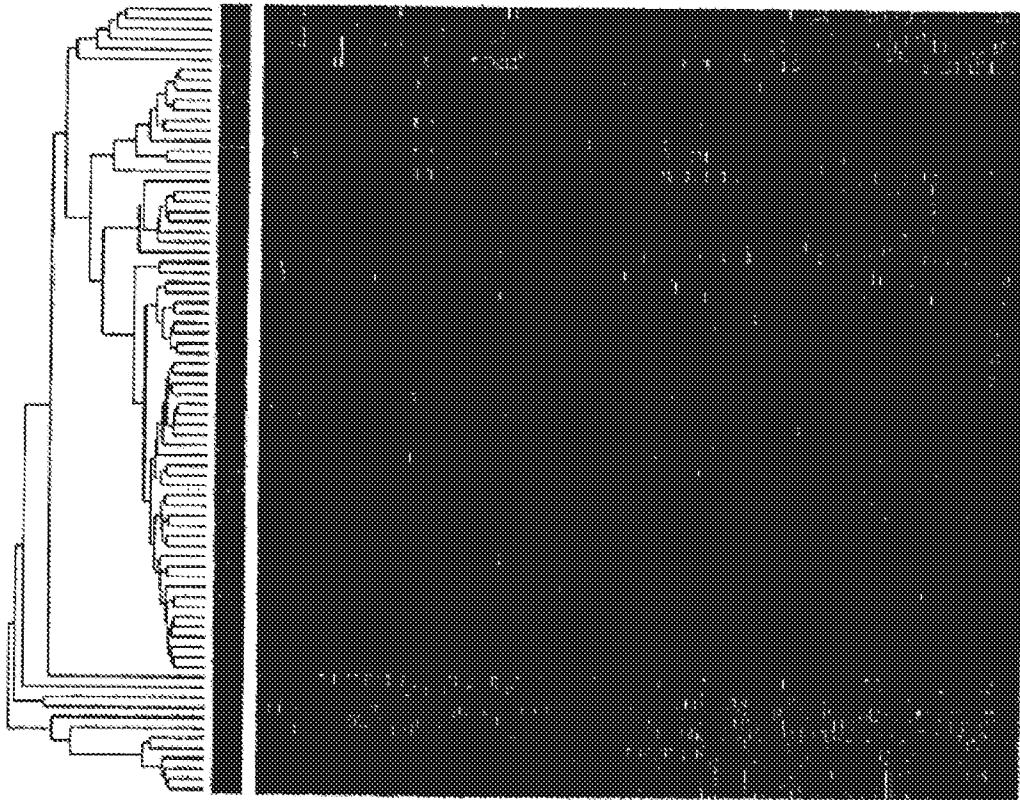


FIG. 54A

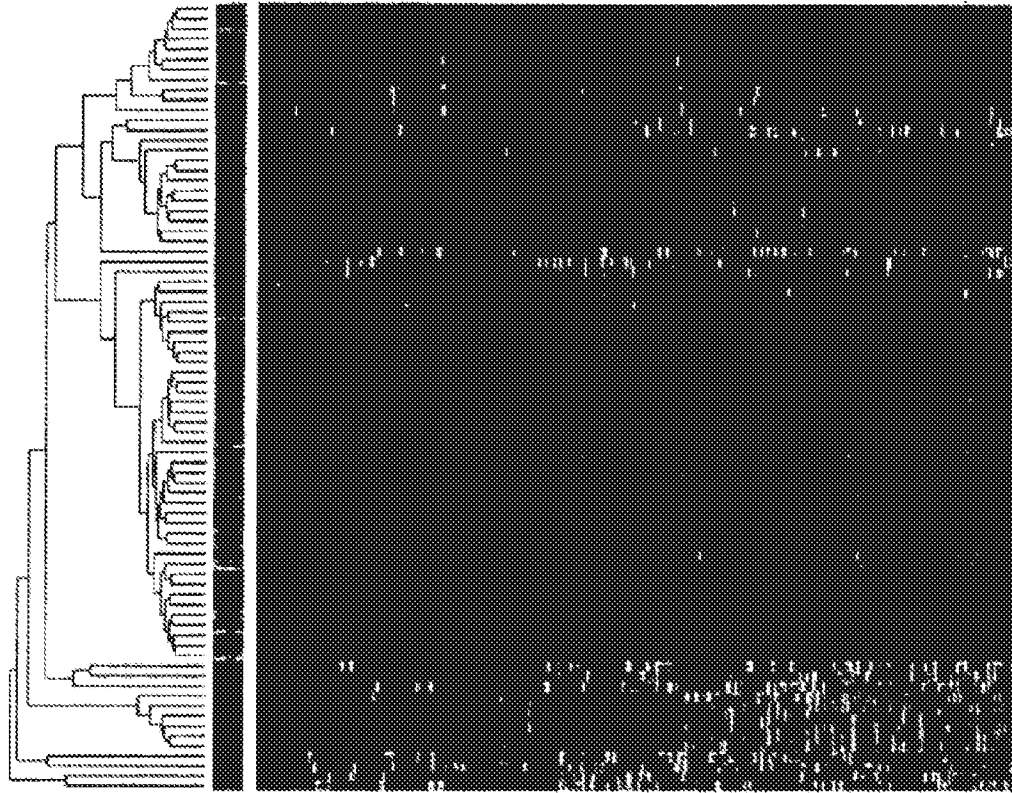


FIG. 55B

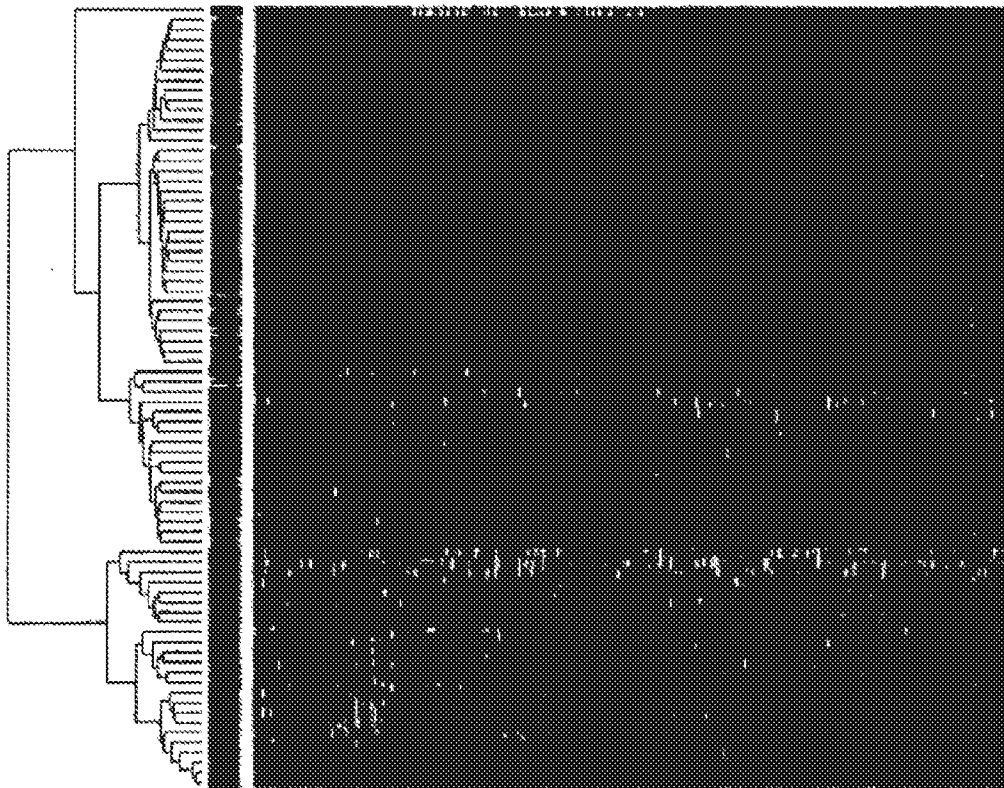


FIG. 55A

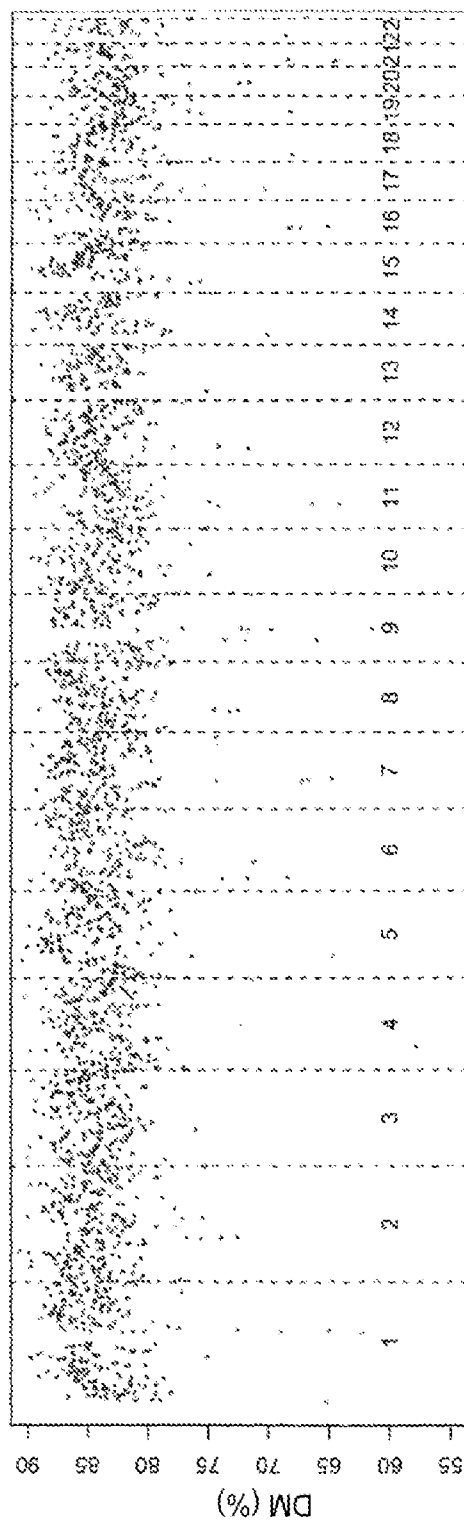


FIG. 56