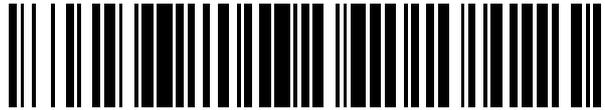


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 560 529**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.06.2008 E 08748039 (8)**

97 Fecha y número de publicación de la concesión europea: **30.09.2015 EP 2150626**

54 Título: **Métodos epigenéticos**

30 Prioridad:

07.06.2007 AU 2007903075 P

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

19.02.2016

73 Titular/es:

HAPLOMIC TECHNOLOGIES PTY LTD. (100.0%)

**24 Mary Street
St Kilda West, VIC 3182, AU**

72 Inventor/es:

SIMONS, MALCOLM JAMES

74 Agente/Representante:

SUGRAÑES MOLINÉ, Pedro

ES 2 560 529 T3

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

DESCRIPCIÓN

Métodos epigenéticos

5 **Campo de la invención**

La presente invención se refiere al campo de la genética, y más específicamente, a la epigenética. En particular, la invención se refiere a métodos para investigar las características epigenéticas del estado haploide de un organismo.

10 **Antecedentes de la invención**

La herencia epigenética es la transmisión de información de una célula u organismo multicelular a los descendientes, sin que esa información esté codificada en la secuencia de nucleótidos del gen. Un tipo de herencia epigenética es la metilación del ADN, la cual, se ha demostrado que está implicada en un número de enfermedades humanas. Por ejemplo, los cambios en los perfiles de metilación del ADN son frecuentes en trastornos tales como el cáncer, y los síndromes de Beckwith-Wiedemann, Prader-Willi y Angelman. También se sabe que la metilación del ADN está implicada en la modulación de la expresión génica en el curso del desarrollo humano. Se piensa que la metilación intensa de las regiones promotoras es importante para la regulación negativa de la transcripción, proporcionando así un "cambio" en la expresión génica.

La metilación del ADN es una modificación epigenética que se produce típicamente en los sitios CpG (es decir, donde una citosina está inmediatamente seguida de una guanina en la secuencia de ADN); la metilación da como resultado la conversión de la citosina a 5-metilcitosina. La formación de Me-CpG está catalizada por la enzima ADN metil-transferasa. Los sitios CpG son poco frecuentes en los genomas de los vertebrados, pero suelen encontrarse en mayor densidad cerca de los promotores de los genes de vertebrados, donde se les denomina colectivamente islas CpG. El estado de metilación de estos sitios CpG puede tener un impacto crucial sobre la actividad y/o la expresión génica.

El patrón de metilación se ha convertido recientemente en un tema importante para la investigación. Por ejemplo, tanto la hipometilación (pérdida de metilación) del ADN como la hipermetilación (aumento de la metilación) del ADN se han relacionado con estudios que examinan genes que están metilados de forma diferente en el tejido normal y en el tejido canceroso. Los genes relacionados con el cáncer que se han relacionado con una metilación alterada incluyen los implicados en la regulación del ciclo celular, la reparación del ADN, la señalización RAS y la invasión. Los estudios han encontrado que en el tejido normal, la metilación de un gen se localiza principalmente en la región codificante, que es pobre en CpG. Por el contrario, la región promotora del gen está sin metilar, a pesar de la alta densidad de islas CpG en la región.

El grado de metilación también puede ser importante en la regulación génica. En el cáncer, el silenciamiento génico mediado epigenéticamente se produce de forma gradual. Empieza con un leve descenso de la transcripción, que promueve un descenso de la protección de la isla CpG contra la propagación de la heterocromatina adyacente y la metilación dentro de la isla. Esta pérdida da como resultado incrementos graduales de los sitios CpG individuales, que varían entre las copias del mismo gen en diferentes células.

Otro tipo de herencia epigenética es la que surge del efecto de las proteínas asociadas al ADN. Se sabe, por ejemplo, que el estado de acetilación de las histonas puede afectar a la expresión del gen con el que están asociadas. Se piensa que la acetilación de ciertas histonas conduce al silenciamiento del gen, debido al empaquetamiento más denso del ADN en la estructura de la cromatina.

Aunque la técnica previa ha divulgado la asociación entre las modificaciones epigenéticas en el ADN y las proteínas asociadas, y el fenotipo, las interacciones son complejas y aún no se han esclarecido completamente. Es un aspecto de la presente invención superar o atenuar un problema de la técnica previa para proporcionar métodos para utilizar patrones de modificaciones epigenéticas en el ADN y patrones de las proteínas asociadas para asignar fenotipos a organismos, además de los que se pensaba anteriormente que eran posibles.

La discusión de documentos, actos, materiales, dispositivos, artículos y similares se incluye en esta especificación únicamente con el propósito de proporcionar un contexto para la presente invención. No se sugiere ni se representa que cualquiera o todas estas materias formasen parte de la base de la técnica previa o constituyeran conocimiento general común en el campo pertinente para la presente invención, ya que existían antes de la fecha de prioridad de cada reivindicación de la presente invención.

60 **Sumario de la invención**

En un primer aspecto, la presente invención proporciona un método para obtener información epigenética para una muestra biológica obtenida de un sujeto humano, animal o vegetal, incluyendo el método la etapa de:

analizar la muestra biológica, conteniendo la muestra una molécula de ADN sustancialmente aislada mediante:

la determinación de la presencia o ausencia de dos o más bases metiladas en la molécula de ADN, en la que la presencia o ausencia de las bases metiladas puede modular la expresión de la molécula de ADN *in vivo*; y la determinación de si cualquiera de las dos bases metiladas está presente en *cis* en la molécula de ADN; caracterizada por que la molécula de ADN sustancialmente aislada, se ha aislado usando un método físico; y en la que la molécula de ADN sustancialmente aislada consiste en un cromosoma único, una cromátida, o un fragmento de la misma.

La etapa de analizar también puede incluir un método *in situ* que puede de analizar selectivamente el ADN. El método *in situ* permite examinar el material poliploide para proporcionar información epigenética haploide definitiva, lo que hace innecesario separar físicamente las moléculas de ADN materno y paterno.

En otra forma del método la molécula de ADN está presente o se obtiene de una célula diploide. Aunque los gametos contienen información haploide, estas células pueden ser difíciles de obtener en la clínica y/o proporcionar información incorrecta sobre las modificaciones epigenéticas presentes en las células somáticas.

En otra forma de la invención en la que la etapa de análisis se realiza en el ADN, la modificación es metilación. El análisis puede ejecutarse utilizando cualquier metodología adecuada, sin embargo, típicamente, la etapa de análisis de la presencia o ausencia de metilación en los dos o más sitios comprende un método seleccionado del grupo que consiste en secuenciación de ADN usando tratamiento con bisulfito, exploración genómica por sitios de restricción, PCR arbitraria sensible a metilación, análisis de *Southern* usando una enzima de restricción sensible a metilación, PCR específica de metilación, digestión con enzimas de restricción de productos de PCR amplificados a partir de ADN convertido con bisulfito y combinaciones de los mismos.

La información epigenética proporcionada por el presente método puede proporcionar información fenotípica para el sujeto, tal como la presencia o ausencia de una enfermedad, afección o trastorno; o una predisposición a una enfermedad, afección o trastorno.

Descripción detallada de la invención

En un primer aspecto, la presente invención proporciona un método para obtener información epigenética para una muestra biológica obtenida de un sujeto humano, animal o vegetal, incluyendo el método la etapa de:

analizar la muestra biológica, conteniendo la muestra una molécula de ADN sustancialmente aislada mediante:

la determinación de la presencia o ausencia de dos o más bases metiladas en la molécula de ADN, en la que la presencia o ausencia de las bases metiladas puede modular la expresión de la molécula de ADN *in vivo*; y la determinación de si cualquiera de las dos bases metiladas está presente en *cis* en la molécula de ADN; caracterizada por que la molécula de ADN sustancialmente aislada, se ha aislado usando un método físico; y en la que la molécula de ADN sustancialmente aislada consiste en un cromosoma único, una cromátida, o un fragmento de la misma.

Según el mejor saber y entender del solicitante, en la invención divulgada en el presente documento es la primera vez que se ha apreciado la importancia de la fase cuando se estudia la herencia epigenética por medio de las modificaciones en el ADN y la proteína. Por primera vez, se obtiene información específica de la fase sobre fenómenos epigenéticos tales como la metilación del ADN y la acetilación de las proteínas asociadas al ADN. La presente invención proporciona, por tanto, la capacidad de discernir si una modificación epigenética en dos sitios está presente en el mismo cromosoma (es decir, una relación *cis*). Esta información es importante, entre otras cosas, para identificar efectos epigenéticos específicos de fase en individuos y poblaciones.

El campo de la epigenética es relativamente nuevo en la técnica de la genética, y se refiere al estudio de cambios en la función genómica que no se basa en la secuencia específica de nucleótidos dentro del ADN de un organismo. La epigenética incluye el estudio de los efectos que se heredan de una generación celular a la siguiente, tanto si estos se producen durante la morfogénesis embrionaria, la regeneración, la renovación normal de las células, los tumores, el cultivo celular, o la replicación de los organismos unicelulares. Los procesos epigenéticos específicos de interés incluyen paramutación, marcaje, silenciamiento, inactivación del cromosoma X, efecto de posición, reprogramación, transvección, impronta genética, efectos maternos, la progresión de la carcinogénesis, los efectos de numerosos teratógenos, y la regulación de las modificaciones de las histonas y la heterocromatina.

En otra forma del método, la presencia o ausencia de las modificaciones puede modular la expresión de la molécula de ADN *in vivo*. Tal como aprecia el experto, la expresión de los genes está controlada, al menos en parte, por modificaciones epigenéticas tales como la metilación del ADN. La metilación del ADN es una modificación epigenética del ADN que se ha propuesto que es universal en los eucariotas. En los seres humanos,

aproximadamente el 1 % de las bases del ADN experimentan metilación del ADN. En los tejidos somáticos del adulto, la metilación del ADN ocurre típicamente en un contexto de dinucleótido CpG; la no metilación de los sitios CpG es frecuente en las células embrionarias pluripotenciales. En las plantas, las citosinas se metilan de forma tanto simétrica (CpG o CpNpG) como asimétrica (CpNpNp), donde N puede ser cualquier nucleótido.

5 En una realización, el análisis incluye la determinación de la metilación de un dinucleótido CpG.

10 En los mamíferos están metilados entre el 60-70 % de todos los CpG. Los CpG no metilados se agrupan en grupos llamados "islas CpG", que están presentes en las regiones reguladoras en 5' de numerosos genes. En numerosos procesos patológicos tales como el cáncer, las islas CpG de los promotores de los genes adquieren una hipermetilación anómala, que da como resultado un silenciamiento de la transcripción heredable. El refuerzo del silenciamiento de la transcripción está mediado por proteínas que pueden unirse a CpG metilados. Estas proteínas, que se llaman proteínas de unión a CpG metilados, reclutan histona-desacetilasas y otras proteínas remodeladoras de la cromatina que pueden modificar las histonas, formando así una cromatina compacta inactiva denominada heterocromatina. Esta conexión entre la metilación del ADN y la proteína a través de la alteración en la estructura de la cromatina está relacionada con el desarrollo de diversos fenotipos. Por ejemplo, la pérdida de la proteína 2 de unión a metil-CpG (MeCP2) se ha implicado en el síndrome de Rett, y el dominio de unión a metil-CpG de la proteína 2 (MBD2, *Methyl-CpG binding domain protein 2*) media el silenciamiento de la transcripción de los genes hipermetilados en el cáncer.

20 Como se describe anteriormente, las proteínas asociadas con el ADN pueden estar implicadas en la modulación de la expresión génica. Por ejemplo, la estructura física del ADN, dado que existe compactado en la cromatina, puede afectar a la capacidad de las proteínas reguladoras de la transcripción (denominadas factores de transcripción) y de las ARN-polimerasas para encontrar el acceso a genes específicos y activar la transcripción a partir de ellos.

25 Cromatina es un término que designa la estructura en la que existe el ADN dentro de las células. La estructura de la cromatina está determinada y estabilizada a través de la interacción del ADN con las proteínas de unión al ADN. Hay 2 clases de proteínas de unión al ADN. Las histonas son la principal clase de proteínas de unión al ADN implicadas en el mantenimiento de la estructura compacta de la cromatina. Hay 5 proteínas histónicas diferentes, identificadas como H1, H2A, H2B, H3 y H4.

30 La otra clase de proteínas de unión al ADN es un grupo diverso de proteínas denominadas simplemente proteínas no histónicas. Esta clase de proteínas incluye los diversos factores de transcripción, polimerasas, receptores de hormonas y otras enzimas nucleares. En cualquier célula dada hay más de 1000 tipos diferentes de proteínas no histónicas unidas al ADN.

35 La unión al ADN de las histonas genera una estructura llamada el nucleosoma. El núcleo del nucleosoma contiene una estructura proteica octamérica que consiste en dos subunidades de cada una, H2A, H2B, H3 y H4. La histona H1 ocupa el ADN internucleosómico y está identificada como la histona engarzadora. El núcleo del nucleosoma contiene aproximadamente 150 pb de ADN. El ADN engarzador entre cada nucleosoma puede variar desde 20 hasta más de 200 pb. Estas estructuras del núcleo nucleosómico aparecerían como cuentas en una hebra si los ADN se estirasen en una estructura lineal.

40 Los propios núcleos del nucleosoma se enrollan en una forma de solenoide que se enrolla sobre sí misma para compactar aún más el ADN. Estas espirales finales se compactan más en la característica cromatina que se ve en una extensión cariotípica. La estructura de proteína-ADN de la cromatina se estabiliza mediante la unión a un armazón de proteína no histónica llamado matriz nuclear.

45 Las modificaciones postraduccionales (MPT) de las histonas se asocian con estados transcripcionales positivos y negativos. Un modelo típico para el papel de estas MPT es que, en respuesta a la señalización citoplasmática para los factores de transcripción, se establecen MPT de acción positiva a lo largo de los promotores y de las fases de lectura abierta mediante activadores unidos al ADN y ARN-polimerasas. Las marcas de acción negativa se efectúan a lo largo de los genes durante la represión, mediante el reclutamiento por represores unidos al ADN a lo largo de las regiones de heterocromatina del genoma. Ambos conjuntos de modificaciones alteran las superficies nucleosómicas que reclutan entonces complejos protéicos reguladores.

50 Las MPT de las histonas incluyen, pero no se limitan a acetilación de la histona 3 (H3), la histona 4 (H4), la histona 2A (H2A), la histona 2B (H2B); fosforilación de H3, H2A y H2B; metilación de arginina de H3 y H4; metilación de lisina de H3 y H4; ubiquitilación de H2A y H2B; sumoilación de lisina de H2A y H2B; e isomerización de prolina en H3; ribosilación de ADP, desiminación (conversión de arginina a citrulina). Las histonas variantes H2AX, H3.1, H3.3 y Hzt1 también se modifican mediante MPT.

55 Este repertorio de MPT de las histonas, conocido como el "código de las histonas", sirve como superficies de unión para la asociación de proteínas efectoras que contienen dominios de interacción específicos. Por ejemplo, los bromodomínios reconocen acetilación, los dominios similares al cromo de la familia Real (cromo, Tudor, MBT) y dominios PHD no relacionados reconocen metilación, y un dominio dentro de las proteínas 14-3-3 reconoce

fosforilación.

Las histona-acetil-transferasas (HAT) de interés incluyen la familia de las GNAT [GCN5 (acetil-transferasas relacionadas con el control general de la síntesis de aminoácidos, *general control of amino-acid synthesis 5*)-related acetyltransferase], la familia de las CBP/P300 (proteína de unión a CREB, *CREB-binding protein*), y la familia de las MYST (proteínas MOZ, YBF2/SAS3, SAS2, TIP60). Estas HAT y las histona-acetil-transferasas nucleares y de factores de transcripción relacionadas con hormonas también incluyen GCN5, PCAF (factor asociado a la proteína de unión a P300/CREB, *P300/CREB binding protein-associated factor*), GCN5L (control general de la síntesis de aminoácidos 2, similar al 5, *general control of amino-acid synthesis 5-like 2*), ELP3, P300 (proteína de unión a e1a p300), CBP (proteína de unión a CREB), TIP60, MOF/MYST1, MOZ (proteína dedo de cinc de leucemia monocítica, *monocytic leukaemia zinc finger protein*)/MYST3, MORF (factor relacionado con MOZ, *MOZ-related factor*)/MYST4, HB01 (histona-acetil-transferasa de unión a ORC, *histone acetyltransferase binding to ORC*)/MYST2, ATF2, TAF1 (factor 1 asociado a secuencias TATA, *TATA box-associated factor 1*), GTF3C4: Factor general de transcripción 3c (*general transcription factor 3c*, polipéptido 4), ACTR: activin receptor, SRC-1 (coactivador del receptor de esteroides 1, *steroid receptor coactivator 1*)/ NCOA1/2 (coactivador del receptor nuclear 1, *nuclear receptor coactivator 1*), ACTR (receptor de activina, *activin receptor*), SRC-1 (coactivador del receptor de esteroides 1), CDYL y HAT1 (histona-acetil-transferasa 1).

Las histona-desacetilasas (HDAC) de interés incluyen HDAC de clase I y clase II (tales como HDAC 1 hasta 8 y HDAC10), y las enzimas de clase III dependientes de NAD de la familia Sir (tales como SirT2).

Las histona-metil-transferasas (HMT) de interés incluyen HMT de arginina, tanto de tipo I como de tipo II (tales como PRMT1, PRMT4, PRMT5 y PRMT7) y HMT de lisina. Las HMT de lisina de interés incluyen las de LLM (leucemia de linaje mixto)-1 hasta 5, las de la familia Set1, que tienen homología con la proteína Set1 de levaduras (incluyendo SET1A y SET1B), SET2, SYMD2, Pr-SET 7/8, CLL8, NSD1, DOT1, SUV42H1/2, EZH2, RIZ1, SUV39H1/2, EuHMTasa, GLP, ESET, SETDB1, y G9a.

Las desmetilasas de arginina y lisina de interés incluyen las proteínas arginina-metil-transferasas (PMRT4, PRAT5, y CARM1), LSD1, JHDM1a, JHDM1b, JHDM2a, JHDM2b, JMJD2A/JHDM3A, JMJD2B, JMJD2C, y JMJD2D.

Enzimas que median otras MPT de histonas tales como desiminación, fosforilación, sumoilación y ubicuitilación, que son de interés, son PAD14, Bmi/Ring1A, RNF20/RNF40 y UbcH6, mono-ADP-ribosil-transferasas y poli-ADP-ribosil-polimerasas, prolina-isomerasas y cinasas tales como haspina, MSK1/2, CKII y Mst1. Son también de interés las proteínas de unión a ADN metilado, tales como proteínas con dominio de unión a metil-CpG (MBD), incluyendo MeCP2, MBD1-4, y MBD3L-1, y MBD3L-2.

Además, dado que el ARNi está implicado en la formación de heterocromatina en insectos, plantas y hongos, y el trabajo reciente sugiere que estos mecanismos están presentes en los vertebrados, la presente invención incluye modificaciones de la cromatina mediadas por ARNi.

La persona experta estará familiarizada con los métodos para aislar proteínas asociadas al ADN. En general, el contenido de 5-metil-citosina del ADN o las MPT de las histonas pueden examinarse utilizando electroforesis capilar de alta resolución, cromatografía líquida de alto rendimiento, o espectrofotometría de masas. Los análisis epigenéticos, tales como la exploración genómica por sitios de restricción, pueden usarse para examinar el material genético haploide, sin embargo, los procedimientos que utilizan la amplificación (normalmente, por reacción en cadena de la polimerasa) de las secuencias diana, tales como la amplificación de sitios intermetilados (AIMS, *amplification of intermethylated sites*) o la hibridación de metilación diferencial (DMH, *differential methylation hybridization*) son de particular interés. Adicionalmente, los que utilizan la amplificación de una señal para la detección, tales como los episodios de detección de P-LISA (discutido más adelante), que facilitan la amplificación en círculo rodante de una diana de hibridación.

A fin de miniaturizar el análisis de metilación de cromosomas aislados, es posible que puedan usarse técnicas tales como el ensayo de ligamiento por proximidad *in situ* (P-LISA, *proximity-ligation in situ assay*), que son capaces de detectar cantidades zeptomolares (40×10^{-21} mol) de proteína (Fredriksson S y cols. (2002) Protein detection using proximity-dependent DNA ligation assays. *Nat Biotechnol.* 20(5):473-7). Este procedimiento utiliza anticuerpos frente a las proteínas de interés (en un ejemplo, anticuerpos contra proteínas metiladas) que, cuando se llevan a proximidad mediante la unión a sus proteínas diana, permiten la hibridación de oligonucleótidos conjugados que sirven como molde para la amplificación en círculo rodante tras el ligamiento enzimático. Dado que se necesitan dos episodios de reconocimiento para la detección de las proteínas diana, la detección es muy específica y capaz de detectar complejos de proteína individuales (Soderberg O (2006) Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods.* 2006 3(12):995-1000). Este procedimiento se ha adaptado recientemente al estudio de las interacciones de proteína-ADN, y se ha adaptado para permitir el análisis *in situ* de las interacciones de ADN-proteína para la detección localizada (Gustafsdottir SM y cols. (2007) In vitro analysis of DNA-protein interactions by proximity ligation. *Proc Natl Acad Sci U S A.* 104(9):3067-72).

Inmunoprecipitación de cromatina acoplada a tecnología de matriz multigénica (CHIP sobre chip). En este procedimiento, se usa un anticuerpo frente a la proteína de unión al ADN modificada postraduccionalmente de interés para inmunoprecipitar, tanto la proteína modificada como su ADN diana. Brevemente, se utiliza la reticulación para fijar las proteínas de unión a ADN al ADN, después de la fragmentación del ADN se usa la inmunoprecipitación para purificar los fragmentos de proteína-ADN con especificidad determinada por el anticuerpo usado. Después de la hidrólisis para revertir la reticulación, se realiza la amplificación y el marcaje de ADN, y el ADN se hibrida a continuación con una micromatriz y se analiza para identificar las regiones de unión por la proteína de unión al ADN. Alternativamente, los genes de interés pueden examinarse mediante PCR usando oligonucleótidos específicos de genes (conocidos como chips de un solo gen, *single-gene CHIP*).

La hibridación de metilación diferencial también puede usarse para identificar ADN metilado. En este procedimiento se utilizan micromatrices de ADN especializadas que comprenden islas CpG clonadas. En el caso de la presente invención, las muestras de ADN a comparar - dos muestras de ADN haploide (cromosomas, cromátidas, etc.) - se digieren cada una con una enzima de restricción sensible a metilación, y los amplicones de la reacción en cadena de la polimerasa obtenidos a partir de cada muestra se hibridan después en la matriz de isla CpG. Los elementos de la matriz con una señal de hibridación más fuerte en cada muestra representan islas CpG hipermetiladas de forma diferencial, protegidas de la escisión por la enzima de restricción sensible a la metilación, y amplificadas, por tanto, mediante PCR en la muestra.

Otros avances en el análisis epigenético permiten el examen de marcas epigenéticas, tales como metilación de ADN, usando cantidades decrecientes de ADN. Para el análisis de la metilación de ADN se ha usado el ensayo de recuperación de islas CpG metiladas (*methylated-CpG island recovery assay* MIRA), basado en la alta afinidad del complejo MBD2/MBD3L1 por el ADN metilado, para purificar el ADN metilado. Brevemente, el ADN tratado con ultrasonido se incuba con una matriz que contiene glutatión-S-transferasa (GST)-MBD2b en presencia de MBD3L-1, un compañero de unión de MBD2 que aumenta la afinidad de MBD2 por el ADN metilado. El ADN unido específicamente se eluye de la matriz y se realizan PCR específicas del gen para detectar la isla de metilación CpG. Alternativamente, el ADN unido puede examinarse usando análisis de micromatriz. Esta técnica ya puede detectar metilación usando 1 ng de ADN. Las fusiones a GST de otras proteínas implicadas en la epigénesis también pueden utilizarse para dichos procedimientos desplegables (*pull-down*), o fusiones que implican a la biotina para aumentar la afinidad por las matrices desplegables. Cuando se acoplan con la secuenciación con bisulfito, los avances en el análisis cuantitativo de la metilación, tales como MethyLight y *Pryosequencing*, permiten la detección de cantidades aún menores de ADN metilado, permitiendo un procedimiento, HeavyMethyl, la detección de 30 pg de ADN metilado.

Además, los avances recientes en la proteómica "de arriba a abajo" (el análisis de proteínas intactas en vez de digerirlas primero hacia péptidos) podrían permitir el examen de los patrones de modificación intactos de diferentes histonas en un nucleosoma dado.

Aunque ha habido numerosos estudios sobre los mecanismos de epigénesis, la técnica previa no ha apreciado adecuadamente los efectos de confusión de considerar simultáneamente la modificación epigenética de una molécula de ADN o de proteína de procedencia materna al mismo tiempo que una molécula de ADN o de proteína de procedencia paterna. Esto se ha producido porque los métodos de la técnica previa analizan las modificaciones epigenéticas utilizando material diploide.

El solicitante propone en el presente documento que una caracterización definitiva de estado epigenético de un sujeto sólo puede proporcionarse mediante el análisis por separado de la modificación epigenética del ADN y de las proteínas asociadas en el estado haploide. Por ejemplo, donde la metilación es la modificación epigenética, el uso de material haploide resuelve los efectos específicos de fase para el reconocimiento de fenómenos heterocigóticos y proporciona el análisis de la variación cuantitativa sobre los múltiples sitios CpG que comprenden una "isla" (unitariamente funcional), incluso para la variación en un único sitio CpG. Esta caracterización definitiva conduce a una información fenotípica más precisa, tal como la presencia y/o predisposición de un individuo a ciertas enfermedades o a la capacidad de un individuo para responder a ciertas moléculas terapéuticas.

Considerando un ejemplo más específico, la metilación del ADN es una característica epigenética que se piensa que está implicada en la patogenia del cáncer. La siguiente tabla describe los genes relacionados con el cáncer que se piensa que son sensibles a la metilación.

Gen	Mapa	Metilado	Notas
14-3-3 Sigma	1p	Cánceres de mama y gástricos	
ABL1 (P1)	9q34.1	50-100 % en LMC, algunas LLA	Metilado sólo cuando forma parte de la traslocación bcr-abl
ABO	9q34	Líneas celulares	
APC	5q21	Cáncer de colon, gástrico y de esófago	Un promotor solamente Correlación con expresión no establecida. Tipo A

ES 2 560 529 T3

Gen	Mapa	Metilado	Notas
AR (<i>androgen receptor</i> , receptor de andrógeno)	Xq11-12	Líneas celulares de cáncer de próstata, FCA de colon	
BLT1 (receptor de leucotrieno B4)		Líneas celulares diversas	
BRCA1	17q21	10-20 % en cánceres de mama, algunos de ovario	Causa de silenciamiento de la transcripción en estas células
CALCA (Calcitonina)	11p15	25-75 % en colon, pulmón, neoplasias hematopoyéticas	Una de las primeras islas CpG del promotor que se demostró que estaban metiladas en el cáncer
CASP8 (CASPASA 8)	2q33-34	Neuroblastoma	Se correlaciona con amplificación de MycN.
Caveolina 1	7q31.1	Líneas celulares de cáncer de mama	
CD44	11pter-p13	Cáncer de próstata	
CFTR	7q31.2	Líneas celulares	No se ha informado de tumores primarios.
COX2	1q25.2-q25.3	Líneas celulares de colon, mama y próstata 15 % en cánceres primarios de colon	Se correlaciona con la expresión cuando está completamente metilado.
CSPG2 (versicano)	5q12-14	ENVEJECIMIENTO del colon. 70% en colon	Proteoglicano secretado, regulado por Rb
CX26 (Conexina 26)	13q11-q12	Líneas celulares de cáncer de mama	
Ciclina A1	13q12.3-q13	Líneas celulares diversas	
DBCCR1	9q32-33	50 % en cáncer de vejiga	Metilación leve en la vejiga normal
ECAD (cadherina E)	16q22.1	20-70 % en ca. de mama, gástricos, de tiroides, epidermoides, leucemias y de hígado	La metilación suele ser heterogénea y no siempre se correlaciona con el silenciamiento. Presente también en algunas muestras de estómago e hígado normales.
Receptor B de endotelina	13q22	60-70 % en cáncer de próstata	
EPHA3	3p11.2	Leucemias	
EPO (eritropoyetina)	7q21	Células HeLa	Normal y tumores primarios
ER (<i>estrogen receptor</i> , receptor de estrógeno)	6q25.1	ENVEJECIMIENTO del colon, hígado, músculo cardíaco, AoSMC (cultivadas), cerebro, AoEC. 100 % en cáncer de colon, 20-30 % en cáncer de mama ER-, 60-70 % en LMA/LLA, 20-50 % en LMC en CB, 20 % en pulmón (CNMP), 60 % en GBM	Hacia el 5' del promotor, no en la isla CpG. No obstante, hay una buena correlación con la pérdida de expresión.
FHIT	3p14.2	10-20 % en carcinoma epidermoide de esófago	
GPC3 (glipicano 3)	Xq26	Líneas celulares de mesotelioma y cáncer de ovario	
GST-pi	11q13	80-100 % en próstata, hígado. 30-60% en colon, mama, riñón	Enzima de detoxificación/reparación del ADN
H19	11p15.5	20-50 % en tumores de Wilm	Gen sellado. La hipermetilación está asociada con la pérdida aparente del sellado del gen IGF2 en el tumor de Wilm, pero no en otros.

ES 2 560 529 T3

Gen	Mapa	Metilado	Notas
Cadherina H (CDH13)	16q24.1-24.2	45 % en cáncer de pulmón, algunos cánceres de ovario	
HIC1	17p13.3	Próstata, mama y cerebro, 80-100% en cáncer de colon, Próstata, mama, GBM. 20-50% en pulmón, riñón, tumores líquidos	Candidato a gen supresor tumoral. Primer gen clonado sobre la base del hallazgo de una isla CpG hipermetilada en cáncer.
hMLH1	2p22	10-20 % en colon, cánceres endometriales y gástricos 0 % en pulmón, mama, GBM, tumores líquidos, etc.	Asociado casi siempre con inestabilidad de microsatélites y, en líneas celulares, con reparación deficiente de errores de apareamiento.
HOXA5	7p15-p14.2	Cáncer de mama	
IGF2 (<i>Insulin-Like Growth Factor II</i> , factor de crecimiento similar a insulina)	11p15.5	ENVEJECIMIENTO del colon, 100 % en cáncer de colon, 50 % en LMA	IGF2 tiene una gran isla CpG que contiene los promotores sellados P2-4.
IGFBP7	4q12	Hepatocarcinogénesis inducida por antígeno T/t de SV40 murino	Normal y tumores primarios
IRF7	11	Líneas celulares diversas	
LKB1	19p13.3	Unos cuantos tumores primarios de colon, testículo y mama (medulares)	
LRP-2 (megalina)	2q24-31	Líneas celulares diversas	
MDGI (<i>mammary-derived growth inhibitor</i> , inhibidor del crecimiento procedente de la mama)	1p35-33	50-70% en cánceres de mama	
MDR1	7q21.1	Líneas celulares de leucemia sensibles a fármacos	Tumores primarios
MDR3 (PGY3)	7q21.1	Líneas celulares diversas	
MGMT (O6-metil-guanina-metil-transferasa)	10q26	25-50% en cerebro, colon, pulmón, mama, LNH, etc.	Asociado con el fenotipo MER
MT1a (metalotioneína 1)	16q13	Hepatoma de rata	Normal y tumores primarios
MUC2	11p15.5	Línea celular de cáncer de colon	Tumores primarios
MYOD1	11 p15.4	ENVEJECIMIENTO del colon. 100 % en colon, 30 % en mama, también vejiga, pulmón, tumores líquidos,	
N33	8p22	ENVEJECIMIENTO del colon. 60-80% en colon, próstata, cerebro	Oligosacaril-transferasa
NEP (<i>neutral endopeptidase</i> , endopeptidasa neutra 24.1) / CALLA	3q21-27	Cáncer de próstata (~10%)	
NF-L (gen codificante del neurofilamento ligero)	8p21	Línea celular de glioma de rata	
NIS (gen del cotransportador unidireccional de yoduro sódico)	19p13.2-p12	Líneas celulares de cáncer de tiroides	Metilación heterogénea en tumores primarios
P14/ARF	9p21	Líneas celulares de cáncer de colon (infrecuente)	Menos frecuente que la metilación de P16, pero asociado normalmente con la metilación de P16

Gen	Mapa	Metilado	Notas
P15 (CDKN2B)	9p21	80 % en LMA/LLA, 2-20 % en GBM, 0 % en colon/pulmón/mama	P15 está físicamente cerca de P16, pero la metilación simultánea de ambos genes es rara.
P16 (CDKN2A)	9p21	20-30 % en pulmón (CNMP), 25-35 % en colon, 5-25 % en linfomas (dependiendo del estadio), 0-5% en vejiga. Muchos otros (esófago, estómago, etc.)	La metilación es tan frecuente como las deleciones y más frecuente que las mutaciones.
P27KIP1	12p13	Líneas celulares de pituitaria de roedor	No se ha informado de tumores primarios.
p57 KIP2	11 p15.5	Líneas celulares de cáncer gástrico	
PAX6	11p13	Líneas celulares de cáncer de colon y 70 % de los tumores primarios	
PgR (<i>progesterone receptor</i> , receptor de progesterona)	11q22	10-20 % en cáncer de mama	Efecto sobre la transcripción
RAR-Beta2	3p24	Colon, mama, cáncer de pulmón	
RASSF1	3p21.3	Cáncer de pulmón	Un promotor solamente
RB1 (retinoblastoma)	13q14	10-20 % de retinoblastomas 0 % en pulmón/leucemia/colon, algunos adenomas de pituitaria	
TERT	5p15.33	Metilación heterogénea en numerosas líneas celulares	
TESTIN	7q31.2	Enfermedades hematopoyéticas malignas	Un promotor solamente
TGFBRI	9q33-q34	Líneas celulares de cáncer gástrico y tumores primarios (10 %)	
THBS1 (tromboespondina-1)	15q15	5-10 % en cáncer de colon 30-40 % en GBM 20-30% en LMA 0 % en endometrio/mama	Inhibidor de la angiogénesis Regulado por P53 y Rb en algunos sistemas
TIMP3	22q12.1-13.2	Cánceres humanos de cerebro (10-50 %) y riñón (20 %), modelo de ratón	
TLS3 (plastina T)	X	Líneas celulares de leucemia	
Urocinasa (uPA)	10q24	Líneas celulares de cáncer de mama	
VHL (Von-Hippell Lindau)	3p25-25	10-20 % en cánceres de células renales, 0 % en tumores comunes sólidos y líquidos	La misma selectividad tumoral que las mutaciones
WT1	11p13	90 % en cánceres de mama, 20-50% en colon, 5-10% en tumores de Wilm	Correlación con expresión
ZO2 (zona de oclusión 2)		Líneas celulares de cáncer de páncreas	

Como se observará en la tabla anterior, la técnica previa ha apreciado que la metilación es importante en la patogenia del cáncer, sin embargo, sólo se ha usado material diploide hasta la fecha. El solicitante propone que los métodos de la técnica previa son susceptibles de proporcionar información poco precisa. Por ejemplo, Al identificar un promotor dado en una muestra tumoral, puede notarse un grado moderado de metilación. Desde un punto de vista real, el grado de metilación medido es un promedio de la metilación para la secuencia promotor de origen materno y la secuencia del promotor de origen paterno, ya que el tumor es, por supuesto, diploide. Aunque este resultado promedio puede ser cierto en algunos casos (porque los promotores, tanto de origen materno como paterno están metilados en el mismo grado), esto no será una verdad universal. Por ejemplo, la secuencia del promotor materno podría estar completamente sin metilar, y la secuencia paterna puede estar muy metilada. Si la secuencia procedente de la madre es dominante sobre la secuencia paterna en cuanto a actividad del promotor, entonces, la importancia del análisis por separado de las secuencias materna y paterna se hace evidente.

Se entenderá, a partir de lo anterior, que la presencia o ausencia de una modificación epigenética se analiza con referencia al estado haploide de la célula. Por tanto, la presencia o ausencia de metilación en un sitio dado de ADN en consideración, puede determinar cuándo se ha separado sustancialmente la característica epigenética de origen

materno de las características equivalentes de origen paterno. De esta forma, la influencia de características epigenéticas procedentes de la madre puede determinarse en ausencia de características epigenéticas procedentes del padre, y viceversa. El solicitante propone que puede obtenerse un mapa de metilación más preciso mediante la eliminación de los resultados potencialmente imprecisos (o, al menos, poco definitivos) proporcionados cuando se analiza el material diploide.

El técnico experto apreciará que la presente invención se distingue del proceso natural de sellado genómico por el que el nivel de expresión de algunos genes depende de si se heredan o no del genoma materno o paterno. Por ejemplo, el gen del factor de crecimiento similar a la insulina 2 (IGF2) es un gen cuya expresión se requiere para el crecimiento y desarrollo fetal normales. La expresión de IGF2 se produce exclusivamente a partir de la copia paterna del gen. Los genes sellados están "marcados" por su estado de metilación. En el caso de IGF2, un elemento en el locus paterno, llamado elemento aislante, se metila, bloqueando su función. La función del aislante no metilado es unir una proteína que, cuando está unida, bloquea la activación de la expresión de IGF2. Cuando está metilada, la proteína no puede unirse al aislante, permitiendo así que un elemento potenciador distante dirija la expresión del gen IGF2. En el genoma materno, el aislante no está metilado, por tanto, la proteína se une a él bloqueando la acción del elemento potenciador distante.

Mediante "definitivo", se quiere decir que no está implicada ninguna estimación, inferencia o determinación de probabilidad, o probabilidad en la asignación de un cierto patrón de modificación epigenética a un cierto fenotipo. Los métodos para analizar modificaciones epigenéticas descritos en la técnica previa se confunden por la presencia de modificaciones paternas en combinación con modificaciones maternas. Por tanto, aunque, por ejemplo, pueden usarse ciertos algoritmos para deducir o inferir una asignación probable entre un patrón de metilación y un fenotipo dados, cuando se analice material no haploide, estas asignaciones serán necesariamente defectuosas. Se propone que la confusión que suele notarse al asignar un patrón de metilación a un fenotipo es el resultado, al menos en parte, de la influencia de la confusión por el material diploide.

De acuerdo con el presente método, un ADN se obtiene de una muestra biológica del sujeto humano, animal o vegetal. La muestra biológica puede ser cualquier material que contenga ADN, incluyendo, pero no limitándose a, sangre total, suero, una célula sanguínea, una célula de la piel, saliva, orina, pelo, uñas, lágrimas, y similar.

Cuando la etapa de análisis incluye el aislamiento sustancial de ADN, el técnico experto puede usar cualquier método adecuado que conozca él o ella. En el método, la etapa del aislamiento sustancial de ADN es por medios físicos. La ventaja que proporcionan los medios físicos sobre los no físicos (tales como el examen selectivo de material diploide) es que los problemas asociados con el discernimiento del material procedente de la madre del material procedente del padre se evitan. Por ejemplo, cuando se usan sondas selectivas, puede que la hibridación cruzada sea problemática, conduciendo a resultados inciertos. Aunque las condiciones de hibridación pueden variarse para limitar la hibridación cruzada, esto requiere más experimentación para realizar.

En una forma de la invención, el método físico para proporcionar ADN haploide que contenga exclusivamente ADN procedente del padre o de la madre es la microdissección cromosómica. La molécula de ADN sustancialmente aislada consiste en un único cromosoma, una cromátida o un fragmento de una cromátida.

Convenientemente, los cortes en el cromosoma pueden realizarse distales al centrómero para separar los brazos p y q, siendo cada uno una molécula de ADN haploide. La persona experta puede identificar la región física de interés en un cromosoma y adoptar un método apropiado para aislar ADN haploide a partir de una muestra diploide, triploide, tetraploide o cualquier otra que tenga un nivel de ploidía mayor.

La persona experta está familiarizada con las plataformas y herramientas usadas para la micromanipulación. Aunque técnicamente rigurosa, la microdissección es rutinariamente asequible. Los requisitos de equipamiento consisten en un microscopio (bien vertical, o bien invertido) provisto de un micromanipulador y una platina giratoria, y un estirador de pipetas (para producir microagujas). Se recomienda aislar el microscopio de la vibración. Aunque no se requiere una habitación especialmente limpia, los fragmentos cromosómicos microdisseccionados contienen sólo cantidades de ADN de femtogramo, y debe controlarse la contaminación con ADN externo.

En una forma del método, puede usarse un método sin contacto para aislar una molécula de ADN haploide. Un ejemplo de este procedimiento es el uso de un microhaz de láser. La microdissección con microhaz de láser puede implicar el uso de un láser ultravioleta pulsado de alta calidad de haz conectado con un microscopio. La microdissección con haz de láser puede realizarse usando, por ejemplo, un robot P.A.L.M.® Robot Microbeam (P.A.L.M. GmbH Bernried, Alemania), disponible en el mercado. La luz de láser es preferentemente de una longitud de onda que no daña o destruye el segmento genómico, tal como 337 nm, que está muy lejos del máximo de absorción de ácidos nucleicos tales como el ADN.

Otro sistema útil para la microdissección con láser de cromosomas es el microscopio de láser de Leica. El sistema usa un microscopio vertical DMLA que incluye revólver motorizado, platina motorizada, el elemento de control xyz motorizado y otras ventajas del nuevo microscopio DMLA. El láser usado es un láser UV de 337 nm de longitud de onda. El movimiento durante el corte lo realiza la óptica, mientras que la platina permanece estacionaria. La región

de interés puede marcarse sobre el monitor y se corta mediante control con el PC. La muestra cae dentro de los tubos de PCR sin fuerzas adicionales. El resultado del corte puede verificarse fácilmente mediante un modo de inspección automatizado.

5 El aislamiento de una molécula haploide de ADN puede conseguirse mediante, por ejemplo, microdissección usando catapulta por láser de un segmento cromosómico usando un láser PALM. En este caso, el proceso de no contacto implica la ablación por láser alrededor del elemento cromosómico diana, seguido de la catapulta por la fuerza del láser del elemento definido sobre la tapa de un tubo, tal como la tapa de un tubo de microcentrífuga, para el análisis posterior del ADN de un solo brazo.

10 La molécula de ADN haploide aislada resultante puede recuperarse usando catapulta de láser por presión. La catapulta de láser por presión puede lograrse enfocando un microhaz de laser bajo, por ejemplo, un segmento o segmentos genómicos haploides de interés, y generando una fuerza como resultado de la alta densidad fotónica que se desarrolla y que causa que el material haploide requerido sea catapultado desde el segmento genómico no requerido. La muestra viaja sobre la parte superior de una onda fotónica y se catapulta dentro de un tubo de recolección. Los tubos de recolección adecuados serán conocidos por los expertos en la técnica e incluyen tubos tales como tubos normales de reacción en cadena de la polimerasa (PCR) o un tubo de microcentrífuga.

15 El ADN puede aislarse sustancialmente mediante citometría de flujo preparatoria usando sondas capaces de discriminar entre ADN materno y paterno. Otro método es mediante el uso de híbridos de radiación, donde el desarrollo de material diploide implica cromosomas humanos como sólo uno de cada pareja de cromosomas. Otra estrategia es el uso de "tecnología de conversión", desarrollada por GMP Technologies Inc. La *GMP Conversion Technology*(R), utiliza un proceso para separar cromosomas apareados en cromosomas únicos. Cuando se separan, los alelos pueden analizarse individualmente usando sondas genéticas que identifican secuencias génicas. La tecnología es aplicable a un gen, un cromosoma, o a todo el genoma humano.

20 En otra forma del método, el material genético contaminante se inactiva o extirpa de modo que ya no realiza la función de contaminar el material genético. Por ejemplo, donde se desee aislar un ADN de procedencia materna, la contribución paterna puede inactivarse o extirparse. Esto puede lograrse mediante la destrucción de un cromosoma homólogo usando un haz de láser orientado cuidadosamente.

25 Otro método para el análisis selectivo de una molécula de ADN es la amplificación selectiva de una secuencia haploide usando PCR, de modo que el número de copias del ADN haploide esté en un gran exceso con respecto al ADN contaminante. La mezcla de moléculas de ADN podría digerirse después parcialmente con una nucleasa, de modo que se digiere sustancialmente todo el ADN contaminante, y queda un bajo nivel de ADN haploide.

30 También existe la posibilidad de amplificar selectivamente el ADN haploide mediante PCR larga usando cebadores que incorporan una etiqueta, y separar las copias usando la etiqueta.

35 Una vez que el ADN se ha aislado sustancialmente, se lleva a cabo el análisis para determinar la presencia o ausencia de una modificación epigenética.

40 Donde la modificación epigenética es la metilación del ADN, la metilación puede detectarse mediante el análisis de la presencia o ausencia de un grupo metilo en el carbono número 5 del anillo pirimidínico de la citosina.

45 El método puede ejecutarse usando cualquier metodología adecuada, sin embargo, típicamente, la etapa de análisis de la presencia o ausencia de metilación en dos o más sitios comprende un método seleccionado del grupo que consiste en secuenciación de ADN usando tratamiento con bisulfito, exploración genómica por sitios de restricción, PCR arbitraria sensible a metilación, análisis de *Southern* usando una enzima de restricción sensible a metilación, PCR específica de metilación, digestión con enzimas de restricción de productos de PCR amplificados a partir de ADN convertido con bisulfito y combinaciones de los mismos.

50 La secuenciación con bisulfito implica hacer reaccionar un ADN monocatenario con bisulfito sódico, que desamina selectivamente la citosina a uracilo pero no reacciona con metilcitosina. La secuencia modificada de ADN producida en la reacción con bisulfito se amplifica mediante PCR, y, después, el ADN amplificado se liga dentro de un vector plasmídico para su clonación y secuenciación. Cuando el ADN se secuencia, sólo los residuos intactos de citosina metilada se amplifican como citosina.

55 La secuenciación con bisulfito puede realizarse usando ADN aislado de menos de 100 células, lo cual es una de las principales ventajas de esta herramienta, porque las muestras tumorales son típicamente muy pequeñas. Otros beneficios de la secuenciación con bisulfito incluyen su capacidad para analizar tramos largos del genoma para determinar patrones muy claros de metilación en el ADN, y produce una demostración cuantitativa positiva de residuos de 5-metilcitosina.

60 La MSP (PCR específica de metilación, *methylation specific* PCR) es una técnica muy rápida y sensible para la exploración de metilación. La MSP se realiza usando bisulfito sódico para modificar el ADN y convertir las citosinas

65

no metiladas a uracilo. Se realiza una amplificación posterior con cebadores específicos para el ADN metilado frente al no metilado, y el análisis se realiza con una simple electroforesis en gel. MethyLight es la próxima generación del ensayo de MSP. La preparación de la muestra y la premisa de los ensayos son idénticos. El procedimiento MethyLight es un avance: El ensayo se hace más cuantitativo, y menos laborioso, a través de la incorporación de un formato "TaqMan" de PCR en tiempo real, mientras que se mantiene la sensibilidad exquisita proporcionada por la MSP convencional.

El parámetro más crítico que afecta a la especificidad de la PCR específica de metilación está determinado por el diseño de los cebadores. En la práctica, suele preferirse manejar sólo una cadena, por lo general, la cadena codificante. En una forma del método, los cebadores se diseñan para amplificar una región que tiene 20-30 pb de longitud, y que debería incorporar suficientes citosinas en la secuencia original para asegurar que el ADN no modificado no servirá como molde para los cebadores. Además, el número y posición de las citosinas dentro del dinucleótido CpG determina la especificidad de los cebadores para los moldes metilado o no metilado. Típicamente, se incluyen 1-3 sitios CpG en cada cebador, y se concentran en la región 3' de cada cebador. Esto proporciona una especificidad óptima y reduce al mínimo los falsos positivos debidos a cebado incorrecto. Para facilitar el análisis simultáneo de las reacciones U y M de un gen dado en el mismo termociclador, ajustamos la longitud de los cebadores para dar temperaturas de desnaturalización/ renaturalización casi iguales. Esto suele dar como resultado que el producto U sea unos cuantos pares de bases más largo que el producto M, lo cual proporciona una forma conveniente de reconocer cada carril después de la electroforesis.

Dado que la PCR específica de metilación utiliza reconocimiento específico por cebadores para discriminar entre los sitios metilados y no metilados, se prefiere que se mantengan condiciones rigurosas para la amplificación. Esto significa que las temperaturas de renaturalización deberían estar cerca de la máxima temperatura que permita la renaturalización y la posterior amplificación. En la práctica, los nuevos cebadores se analizan típicamente con una temperatura de renaturalización inicial 5-8 grados por debajo de la temperatura de desnaturalización calculada. La inespecificidad puede remediarse mediante leves incrementos en la temperatura de renaturalización, mientras que la falta o escasez de los productos de PCR puede mejorarse mediante un descenso de la temperatura de 1-3 grados Celsius. Al igual que con todos los protocolos de PCR, hay que tener cuidado para asegurarse de que los ADN moldes y los reactivos no se contaminen con ADN o productos de PCR exógenos.

La MSP utiliza las diferencias de secuencias entre alelos metilados y alelos no metilados que se producen después del tratamiento con bisulfito sódico. La frecuencia de sitios CpG en CpG facilita esta diferencia de secuencia. Los cebadores para un locus dado se diseñan, lo cual distingue el ADN metilado del no metilado en el ADN modificado con bisulfito. Dado que la distinción es parte de la amplificación con PCR, puede lograrse una extraordinaria sensibilidad, típicamente hasta la detección del 0,1 % de los alelos, mientras que se mantiene la sensibilidad. Los resultados se obtienen inmediatamente después de la amplificación con PCR y la electroforesis en gel, sin la necesidad de análisis de restricción o secuenciación posterior. La MSP también permite el análisis de muestras muy pequeñas, incluyendo muestras incluidas en parafina y microdisseccionadas.

En una realización del método, la etapa de análisis de la presencia o ausencia de metilación en dos o más sitios incluye: (a) reacción del ADN con bisulfito sódico para convertir los residuos de citosina no metilada a residuos de uracilo, mientras que se dejan sin modificar cualquiera de los residuos de 5-metilcitosina para crear una muestra de ADN expuesto convertido por bisulfito que tiene sitios de unión para cebadores específicos para la muestra de ADN convertido por bisulfito; (b) realización de un procedimiento de amplificación por PCR que utiliza cebadores específicos de la cadena codificante o no codificante; (c) aislamiento de los productos de amplificación por PCR; (d) realización de una elongación de los cebadores usando un cebador Ms-SNuPE, dNTP y polimerasaTaq, en la que el cebador Ms-SNuPE comprende una secuencia cebadora desde 15-mero hasta 22-mero de longitud que es complementaria a la muestra de ADN convertido por bisulfito y termina inmediatamente en el 5' del residuo de citosina de las dos o más secuencias de dinucleótido de CpG a analizar; y (e) determinación del estado de metilación de las dos o más secuencias de dinucleótido CpG mediante la determinación de la identidad de la primera base elongada a partir del cebador.

En una realización, los dNTP están marcados, y la determinación de la identidad de la primera base elongada a partir del cebador se analiza mediante la incorporación de los dNTP marcados.

El análisis de la metilación puede facilitarse mediante el uso de métodos de alto rendimiento para identificar sitios de metilación potencial. Las técnicas disponibles para la exploración incluyen exploración genómica por sitios de restricción (RLGS, *restriction landmark genome scanning*), matrices de expresión génica, que pueden usarse como sustitutas para ver qué genes se expresan después de la exposición a inhibidores de ADN-metil-transferasas como la azacitidina.

En la técnica se conocen ensayos de alto paralelismo del genoma completo, varios de ellos divulgados por Fan y cols. (Nat Rev Genet 2006, 7(8) 632-44). Muchos de estos métodos están disponibles para el experto como servicios contratados, siendo un ejemplo el "Golden Gate Methylation Solution" tal como el proporcionado por Illumina Inc (San Diego, CA). El sistema Illumina puede analizar hasta 1.536 sitios CpG con una resolución de un solo sitio sobre cientos de genes a lo largo de 96 muestras. Este sistema puede, por tanto, proporcionar hasta 147.456 mediciones

cuantitativas de metilación de ADN por ensayo.

Se prevé, además, la obtención de una biblioteca de islas CpG humanas y, posteriormente, de su inclusión en micromatrices. La construcción de la base de datos empieza con cromatogramas brutos. Después de eliminar los duplicados, los 6.800 elementos se someten a análisis con la herramienta *Basic Local Alignment Search Tool* (BLAST), contra las bases de datos *High-Throughput Genome Sequence* y varias bases de datos de secuencias de ADN de la Universidad de California, Santa Cruz. En la base de datos de islas CpG, se asigna a cada clon un número de identificación y se enumeran sus características, incluyendo información tal como su localización en el cromosoma, si es un promotor o se encuentra en la región adyacente en 5', su contenido en GC, y qué sitios de restricción existen en el clon. La última información es útil para determinar la utilidad del análisis con enzimas de restricción, así como la secuencia real.

La técnica de micromatriz de CGI (*de CpG Island*, islas CpG) también se prevé como útil en el contexto de los presentes métodos. Esta técnica permite la evaluación simultánea de miles de dianas potenciales de metilación de ADN en un único chip. Implica la disposición de los clones de islas CpG en micromatrices sobre portaobjetos de cristal, la preparación de los amplicones de la muestra diana, y la hibridación de los amplicones sobre las micromatrices de CGI. Como ejemplo, la técnica puede realizarse usando muestras de ADN genómico procedentes de tumores, que se obtienen usando una encima de restricción (Mse1), capaz de cortar inmediatamente fuera de la isla CpG. A continuación, se añaden engarzadores de captura que portan las secuencias de los cebadores de la PCR a los fragmentos Mse1. La muestra se divide en una porción de referencia, que sirve como denominador para determinar cómo ha funcionado de bien la amplificación genómica, y en una porción problema que se digiere con McrBC, una enzima de restricción sensible a metilación que sólo digiere ADN si la región está metilada. Las dos porciones se amplifican mediante PCR, y después, el ADN se marca directamente con los colorantes fluorescentes rojo Cy5 (problema) o verde Cy3 (referencia). Los fragmentos de ADN no digeridos con McrBC en la muestra problema producen producto de PCR marcado con Cy5, mientras que se no se produce ningún producto marcado si había fragmentos metilados digeridos con McrBC. Los productos de PCR marcados problema y de referencia se mezclan y se aplican sobre el portaobjetos de cristal. Los portaobjetos hibridados se escanean y las imágenes adquiridas se analizan para identificar las señales metiladas.

Como apreciará la persona experta, la presente invención tendrá numerosos usos en el campo de la biología y, particularmente, en medicina. Como se discute anteriormente, el cáncer se considera una afección epigenética. De hecho, los cambios epigenéticos, particularmente la metilación del ADN, son susceptibles de cambiar y son excelentes candidatos para explicar cómo ciertos factores ambientales pueden aumentar el riesgo de cáncer. La delicada organización de la metilación y de la cromatina regula la homeostasis celular normal de los patrones de expresión génica que pasan a ser irreconocibles en la célula cancerosa. El genoma de la célula transformada puede sufrir simultáneamente una hipometilación genómica global y una densa hipermetilación de las islas CpG asociadas con regiones reguladoras de genes. Estos drásticos cambios pueden conducir a inestabilidad cromosómica, activación de secuencias parasitarias endógenas, pérdida de sellado, expresión ilegítima, aneuploidía, y mutaciones, y puede contribuir al silenciamiento de la transcripción de los genes supresores de tumores. La inactivación asociada a hipermetilación puede afectar a muchas de las vías de la red celular, tales como la reparación del ADN (hMLH1, BRCA1, MGMT, líder em), el ciclo celular (p16 (INK4a), p14 (ARF), p15 (INK4b), y la apoptosis (DAPK, APAF-1). La metilación aberrante de islas CpG también usarse como un biomarcador de células malignas y como un indicador de su comportamiento.

En una forma de la invención, el ADN haploide está presente o se obtiene de una célula diploide. El método puede usar un cromosoma autosómico de una célula somática. La expresión "cromosoma autosómico" significa cualquier cromosoma dentro de una célula somática o germinal normal, excepto los cromosomas sexuales. Por ejemplo, en el ser humano, los cromosomas 1 a 22 son cromosomas autosómicos. El solicitante propone que la evitación del material naturalmente haploide (tal como el contenido en espermatozoides y óvulos) es ventajoso porque las modificaciones epigenéticas tales como la metilación se eliminan, al menos parcialmente, y suelen eliminarse totalmente. Esto es porque el patrón de metilación se reajusta típicamente durante la meiosis. Por tanto, el análisis de un espermatozoide u óvulo no proporcionará información útil que permita la definición de un fenotipo para el sujeto.

El uso de material naturalmente haploide tal como el de espermatozoides u óvulos también debe evitarse debido a los problemas con la obtención de estas células sexuales en la clínica. La obtención de óvulos es un procedimiento invasivo para las mujeres de cualquier edad, y la recolección de espermatozoides de un varón pediátrico también requiere intervención médica.

La evitación de células sexuales en la determinación del haplotipo epigenético tiene otra ventaja cuando se considera que durante el proceso de la recombinación meiótica, pueden darse sucesos tales que los loci que estaban inicialmente unidos en *cis* pasen a estar asociados en *trans*. Por tanto, el análisis del haplotipo epigenético de un gameto dará una información haplotípica diferente (es decir, incorrecta) a la del ADN haploide obtenido de una célula diploide.

En una realización de la invención, se analiza la metilación de dos o más sitios en la misma molécula de ADN. Esta forma del método es útil para determinar si los dos sitios metilados están presentes naturalmente en *cis* en la molécula de ADN de la célula. Esto puede lograrse considerando sólo los sitios de metilación en un solo cromosoma (de procedencia bien materna o bien paterna), por ejemplo, para demostrar que los dos sitios de metilación están presentes en *cis* porque ambos aparecen en el mismo cromosoma.

El método puede practicarse donde el ADN no esté físicamente separado en componentes de procedencia materna y paterna. Mediante el uso de un método *in situ*, será posible examinar selectivamente un ADN paterno o un ADN materno de modo que sea innecesario separar físicamente las dos moléculas. De esta forma, el material diploide puede usarse para proporcionar información sobre el estado haploide de la célula. Métodos tales como el marcaje cebado *in situ* (PRINS, *primed "in situ" labelling*) serán útiles a este respecto por identificar selectivamente ADN materno o paterno, permitiendo así la localización de una modificación epigenética en un cromosoma materno o paterno. Este método se ha usado con éxito con cebadores específicos para ciertos cromosomas, usando núcleos tanto en metafase como en interfase (Hindkjaer y cols., *Methods Mol Biol.* 1994; 33:95-107).

También se prevé que puedan utilizarse sondas de APN (ácido peptidonucleico) para la identificación *in situ* de cromosomas. La química del APN puede usarse para fabricar pequeños oligómeros. Cuando se marcan con colorante, estos oligómeros constituyen excelentes sondas y ofrecen diferentes ventajas sobre los ácidos nucleicos convencionales. Las sondas de APN son típicamente muy pequeñas (12-20-meros), y muestran tanto señales fuertes como bajo fondo. Las sondas de APN pueden usarse en el contexto de la hibridación *in situ* fluorescente (APN-FISH) como una poderosa tecnología para explorar la estructura cromosómica sobre el cromosoma en metafase e interfase. La APN-FISH es un método en el que se usan pequeñas sondas oligoméricas de APN para marcar cromosomas de una forma específica de secuencia, permitiendo la identificación de una secuencia subyacente en un cromosoma en particular. Los métodos para la APN-FISH se divulgan en Strauss 2002 ("PNA-FISH" in *FISH Technology*. Rautenstrauss/Liehr Eds. Springer Verlag. Heidelberg).

Se entenderá que los métodos descritos en el presente documento pueden encontrar uso en cualquier área de la biología en la que el efecto de la metilación y otras modificaciones de las proteínas tenga que considerarse en conexión con la expresión génica. Estos usos no se limitan a los animales (seres humanos y de otro tipo), y pueden aplicarse también a los vegetales.

La información fenotípica obtenida de los presentes métodos se selecciona del grupo que consiste en: La presencia o ausencia de una enfermedad, afección, o trastorno, o una predisposición a una enfermedad, afección, o trastorno.

La presente invención se describirá ahora completamente mediante la referencia a los siguientes ejemplos no limitantes.

Ejemplo 1: PCR específica de metilación sobre ADN haploide

Preparación del tejido

Se preparan extensiones de metafase de células de sangre periférica obtenidas de un sujeto humano, mediante métodos de cariotipado habituales sobre portas de membranas de PEN. Se prepara un colorante de Giemsa habitual y se usa para identificar cromosomas en la preparación.

Procedimiento de tinción

Etapa	Proceso	Tiempo
1.	Tampón fosfato a pH 6,8	1 minuto
2.	Tripsina al 0,025% en tampón fosfato a pH 6,8	2 minutos
3.	Tampón fosfato a pH 6,8	1 minuto
4.	Colorante de Giemsa, solución de trabajo	15 minutos
5.	Agua destilada	2-3 inmersiones

El exceso de agua se sacude del portaobjetos y la parte de atrás se seca con una toallita kimwipe y se deja secar al aire.

Microdissección con Láser

Se usa el sistema PALM™ de microdissección (P.A.L.M. Microlaser Technologies GmbH, Alemania). Las extensiones de metafase (preparadas como anteriormente) se usan para la captura con láser usando un microscopio P.A.L.M con el objetivo de 100 x. Los cromosomas únicos separados espacialmente de su cromosoma hermano, incluyendo cromosomas únicos se catapultan dentro de tapones de tubos de PCR de 200 µl que contienen 6 µl de Triton-X-100 al 0,1 % (v/v) usando protocolos habituales de microscopio P.A.L.M. Para el análisis de metilación, el material catapultado se transfirió al fondo del tubo mediante centrifugación. Se obtienen muestras de ADN de procedencia tanto materna como paterna.

Extracción de ADN

El ADN del fragmento de cromosoma microdisecionado del tapón CapSure Macro LCM se extrae usando el kit PicoPure de extracción de DNA (Arcturus Inc, CA), siguiendo el protocolo del *kit*.

5

Detección de metilación:

El ADN se modifica mediante tratamiento con bisulfito que convierte las citosinas no metiladas, pero no las metiladas, en uracilo. Después de la eliminación del bisulfito y de la finalización de la conversión química, este ADN modificado se usa como molde para la PCR. Se realizan dos reacciones de PCR para cada muestra de ADN, una específica para ADN metilado originalmente en el gen de interés, y una específica para ADN originalmente no metilado. Los productos de la PCR se separan en geles no desnaturizantes de poliacrilamida al 6-8 % y las bandas se visualizan mediante tinción con bromuro de etidio. La presencia de una banda del peso molecular apropiado indica la presencia de alelos no metilados y/o metilados en la muestra original.

10

15

Para preparar las mezclas, se descongelan el tampón de PCR 10 x, los NTP y los cebadores. Se determina el número de muestras a analizar, incluyendo un control positivo tanto para la reacción no metilada como para la metilada, y un control con agua. Se hace una mezcla maestra para cada reacción de PCR (metilada y no metilada). Para cada 50 µl de reacción deberían usarse las siguientes cantidades:

20

Tampón de PCR 10 x:	5 µl
Mezcla de 4 NTP 25 mM	2,5 µl
Cebador codificante (300 ng/µl)	1 µl
Cebador no codificante (300 ng/µl)	1 µl
Agua destilada estéril	28,5 µl

Se dispensan 38 µl de esta mezcla de PCR en tubos de PCR separados (tubos o tiras de 0,5 ml) marcados para cada muestra. Los componentes se mezclan bien antes de su dispensación.

25

Se añaden 2 µl de ADN molde modificado con bisulfito a cada tubo. Se incluyen una reacción metilada y no metilada para cada muestra, al igual que controles positivos, así como un control sin ADN.

Se añaden de una a dos gotas de aceite mineral (~25-50 µl) a cada tubo, antes de colocarlos en el termociclador. El aceite mineral cubre completamente la superficie de la mezcla de reacción para evitar la evaporación.

30

Los productos de PCR se amplifican después en el termociclador. La PCR se inicia con una desnaturalización de 5 minutos a 95 °C. La polimerasa Taq se añade después de la desnaturalización inicial: 1,25 unidades de polimerasa Taq diluidas en 10 µl de agua destilada estéril. Estos 10 µl se mezclan dentro de los 40 µl a través del aceite mediante pipeteo suave hacia arriba y hacia abajo. La amplificación se continúa durante 35 ciclos con los siguientes parámetros:

35

- 30 seg a 95 °C (desnaturalización)
- 30 seg a la temperatura específica del cebador (renaturalización)
- 30 seg a 72 °C (elongación)
- 40 Etapa final: 4 min a 72 °C (elongación)
- Mantener a 40 °C hasta el análisis.

Los productos de PCR se analizan mediante electroforesis en gel de la forma siguiente.

45

Se prepara un gel de poliacrilamida al 6-8 % no desnaturizante. El TBE 1 X proporciona mejor capacidad tamponadora y bandas más definidas para resolver estos productos. El tamaño de los productos generados típicamente mediante análisis de MSP está en el intervalo de los 80-200 pb, lo que hace a los geles de acrilamida óptimos para la resolución del tamaño. Los geles horizontales de alto porcentaje de agarosa pueden usarse como una alternativa.

50

Las reacciones de cada muestra se corren juntas para permitir la comparación directa entre alelos metilados y no metilados. Se incluyen controles positivos. Los geles verticales se corren a 10 V/cm durante 1 -2 horas. El gel se tiñe en bromuro de etidio y se visualiza bajo iluminación UV. Se hace una comparación del estado de metilación del ADN de procedencia materna y del ADN de procedencia paterna para una región dada de ADN.

55

REIVINDICACIONES

1. Un método para obtener información epigenética para una muestra biológica obtenida de un sujeto humano, animal o vegetal, incluyendo el método la etapa de:
- 5 analizar la muestra biológica, conteniendo la muestra una molécula de ADN sustancialmente aislada mediante:
- la determinación de la presencia o ausencia de dos o más bases metiladas en la molécula de ADN, en la que la presencia o ausencia de las bases metiladas puede modular la expresión de la molécula de ADN *in vivo*; y
- 10 la determinación de si cualquiera de las dos bases metiladas está presente en *cis* en la molécula de ADN;
- caracterizado por que** la molécula de ADN sustancialmente aislada, se ha aislado usando un método físico; y en la que la molécula de ADN sustancialmente aislada consiste en un cromosoma único, una cromátida, o un fragmento de la misma.
- 15 2. El método de acuerdo con la reivindicación 1, en el que el cromosoma único, la cromátida o el fragmento de la misma se aísla mediante disección mediada por láser.
3. El método de acuerdo con las reivindicaciones 1 o 2 en el que la molécula de ADN está presente en, o se obtiene de, una célula diploide.
- 20 4. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que el análisis incluye la determinación de la metilación de un dinucleótido de CpG.
5. El método de acuerdo con la reivindicación 4 o con la reivindicación 5, en el que el análisis incluye un método seleccionado del grupo que consiste en secuenciación de ADN usando tratamiento con bisulfito, exploración genómica por sitios de restricción, PCR arbitraria sensible a metilación, análisis de *Southern* usando una enzima de restricción sensible a metilación, PCR específica de metilación, digestión con enzimas de restricción de productos de PCR amplificados a partir de ADN convertido con bisulfito y combinaciones de los mismos.
- 25 6. El método de acuerdo con la reivindicación 5, en el que el análisis incluye secuenciación de ADN usando tratamiento con bisulfito, el análisis incluye:
- (a) reacción del ADN con bisulfito sódico para convertir los residuos de citosina no metilada a residuos de uracilo, mientras que se dejan sin modificar cualquiera de los residuos de 5-metilcitosina para crear una muestra de ADN expuesto convertido por bisulfito que tiene sitios de unión para cebadores específicos para la muestra de ADN convertido por bisulfito;
- 35 (b) realización de un procedimiento de amplificación por PCR que utiliza cebadores específicos de la cadena codificante o no codificante;
- (c) aislamiento de los productos de amplificación por PCR;
- 40 (d) realización de una reacción de elongación de los cebadores usando un cebador Ms-SNuPE, dNTP y polimerasaTaq, en la que el cebador Ms-SNuPE comprende una secuencia cebadora desde 15-mero hasta 22-mero de longitud que es complementaria a la muestra de ADN convertido por bisulfito y termina inmediatamente en el 5' del residuo de citosina de las dos o más secuencias de dinucleótido de CpG a analizar; y
- 45 (e) determinación del estado de metilación de las dos o más secuencias de dinucleótido de CpG mediante la determinación de la identidad de la primera base elongada a partir del cebador.
7. El método de la reivindicación 6, en el que los dNTP se marcan, y la determinación de la identidad de la primera base elongada a partir del cebador se mide mediante la incorporación de los dNTP marcados.
- 50 8. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en el que la información epigenética puede proporcionar información fenotípica para el sujeto a partir del cual se obtuvo la muestra biológica.
9. El método de acuerdo con la reivindicación 8, en el que la información fenotípica se selecciona del grupo que consiste en: presencia o ausencia de una enfermedad, afección, o trastorno; o una predisposición a una
- 55 enfermedad, afección o trastorno.