

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 431 944**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.06.2003 E 10009948 (0)**

97 Fecha y número de publicación de la concesión europea: **09.10.2013 EP 2339025**

54 Título: **Métodos para determinar los perfiles de metilación**

30 Prioridad:

26.06.2002 US 392071 P

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

28.11.2013

73 Titular/es:

**COLD SPRING HARBOR LABORATORY (50.0%)
Nichols Building, One Bungtown Road
Cold Spring Harbor, NY 11724, US y
WASHINGTON UNIVERSITY IN ST. LOUIS (50.0%)**

72 Inventor/es:

**MARTIENSSEN, ROBERT;
RICHARDS, ERIC J.;
LIPPMANN, ZACHARY y
COLOT, VINCENT**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 431 944 T3

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

DESCRIPCIÓN

Métodos para determinar los perfiles de metilación

5 **Campo de la invención**

La presente invención se refiere a la determinación de los perfiles de metilación.

10 **Estado de los derechos de las invenciones realizadas bajo investigación y desarrollo patrocinados por el gobierno federal**

Esta invención se realizó con el patrocinio del Gobierno bajo la Concesión Núm. NSF 0077774, emitido por la National Science Foundation. El gobierno tiene ciertos derechos en esta invención.

15 **Antecedentes de la invención**

20 La metilación del ADN es un proceso biológico ubicuo que se produce en diversos organismos que abarcan desde bacterias a seres humanos. Durante este proceso, las metiltransferasas de ADN catalizan la adición post-replicativa de un grupo metilo a la posición N6 de la adenina o al C5 o la posición N4 de la citosina, para lo cual la S-adenosilmetionina es el donante universal del grupo metilo. En eucariotas superiores, la metilación del ADN juega un papel en la impronta genómica. Además, las aberraciones en la metilación del ADN han sido implicadas en el envejecimiento y en diversas enfermedades incluyendo el cáncer.

25 Por lo tanto, existe la necesidad de métodos para determinar el perfil de metilación de una célula individual o un organismo. La presente invención aborda este y otros problemas.

Breve resumen de la invención

30 La presente invención proporciona métodos para determinar un perfil de metilación de ADN de una célula, tejido, u organismo. Los métodos comprenden las etapas de:

- 35 a. proporcionar una población de fragmentos de ADN escindidos al azar o sometidos a cizalla de la célula, tejido, u organismo, en donde el ADN comprende fragmentos metilados y no metilados;
- b. agotar el ADN metilado o no metilado de la población; y
- c. cuantificar la cantidad de al menos una secuencia del ADN metilado de la población agotada o del ADN no metilado de la población agotada con respecto a la cantidad de la al menos una secuencia del ADN genómico total.

40 En algunas realizaciones la etapa de agotamiento comprende la fragmentación de la población con una enzima de restricción sensible a la metilación o una enzima de restricción dependiente de la metilación para producir ADN digerido y ADN no digerido y separar el ADN digerido del ADN no digerido.

En algunas realizaciones:

45 etapa a. comprende proporcionar una población de fragmentos de ADN escindidos al azar o sometidos a cizalla de una célula, tejido, u organismo, en donde el ADN comprende una primera porción y una segunda porción y cada porción comprende nucleótidos metilados y no metilados;

50 etapa b. comprende agotar el ADN metilado o no metilado de la segunda porción;

etapa c. comprende cuantificar la cantidad relativa de al menos una secuencia específica en al menos dos muestras de ADN seleccionadas del grupo que consiste en la primera porción, el ADN metilado de la segunda porción, y el ADN no metilado de la segunda porción,

55 determinando de ese modo el perfil de metilación de varias de tales secuencias de ácido nucleico de la célula u organismo.

En algunas realizaciones, los métodos comprenden las etapas de:

60 marcar las al menos dos muestras de ADN con diferentes marcas, y

hibridar las al menos dos muestras de ADN a un ácido nucleico; y

determinar la hibridación relativa de las al menos dos muestras de ADN a la secuencia específica calculando la proporción de las dos marcas hibridantes.

En algunas realizaciones, la etapa de cuantificación comprende la amplificación cuantitativa.

5 En algunas realizaciones, las al menos dos muestras de ADN son el ADN metilado de la segunda porción y el ADN no metilado de la segunda porción. En algunas realizaciones, las al menos dos muestras de ADN son la primera porción y el ADN metilado de la segunda porción. En algunas realizaciones, las al menos dos muestras de ADN son las primera porción y el ADN no metilado de la segunda porción.

10 En algunas realizaciones, el ADN escindido al azar o sometido a cizalla comprende secuencias de reconocimiento metiladas y no metiladas de una enzima de restricción sensible a metilo y la etapa de agotamiento comprende la escisión de la segunda porción con una enzima de restricción sensible a metilo. En algunas realizaciones, el ADN escindido al azar o sometido a cizalla comprende secuencias de reconocimiento metiladas y no metiladas de una enzima de restricción dependiente de metilo y la etapa de agotamiento comprende la escisión de la segunda porción con la enzima de restricción dependiente de metilo.

15 En algunas realizaciones, el ácido nucleico está conectado a un soporte sólido. En algunas realizaciones, el soporte sólido es una micromatriz. En algunas realizaciones, el soporte sólido es una cuenta. En algunas realizaciones, el soporte sólido es una matriz.

20 En algunas realizaciones, el organismo es una planta. En algunas realizaciones, el organismo es un hongo. En algunas realizaciones, el organismo es un procariota. En algunas realizaciones, el procariota es un patógeno bacteriano. En algunas realizaciones, el patógeno bacteriano se selecciona del grupo que consiste en especies gram positivas y gram negativas y micobacterias. En algunas realizaciones, el organismo es un animal. En algunas realizaciones, el animal es un ser humano.

25 En algunas realizaciones, la célula es una célula pluripotencial. En algunas realizaciones, la célula es transgénica y el ácido nucleico corresponde al sitio de inserción de un transgén. En algunas realizaciones, el tejido es la sangre. En algunas realizaciones, el tejido es tejido de biopsia. En algunas realizaciones, el tejido es tejido resecado. En algunas realizaciones, el tejido es normal. En algunas realizaciones, el tejido es tejido tumoral. En algunas realizaciones, el tejido es precanceroso.

30 En algunas realizaciones, los métodos comprenden adicionalmente la comparación del perfil de metilación de un ácido nucleico con la transcripción del ácido nucleico, determinando de ese modo la relación entre la metilación y la transcripción del ácido nucleico. En algunas realizaciones, la transcripción del ácido nucleico se detecta por medio de una micromatriz.

35 En algunas realizaciones, los métodos comprenden adicionalmente la comparación del perfil de metilación de un ácido nucleico con el número de copias del ácido nucleico, determinando de ese modo la contribución a un fenotipo de la combinación de la metilación del ácido nucleico y el número de copias del ácido nucleico. En algunas realizaciones, el número de copias del ácido nucleico se detecta con una micromatriz.

40 En algunas realizaciones, los métodos comprenden adicionalmente la comparación del perfil de metilación de un espécimen de un patógeno bacteriano con una cepa de referencia del patógeno, en donde la similitud de los patrones de metilación indica el origen común del espécimen y la cepa de referencia.

45 La presente invención puede utilizar micromatrices de polinucleótidos. Las micromatrices de polinucleótidos pueden hibridar a una primera y una segunda porciones de ADN marcadas, en donde las porciones son de poblaciones de tamaño uniforme de ADN escindido al azar o sometido a cizalla de una célula, tejido, u organismo; en donde la primera porción de ADN comprende ADN no metilado y metilado marcado con una primera marca; y en donde la segunda porción de ADN tiene agotado el ADN no metilado o el ADN metilado y la segunda porción de ADN está marcada con una segunda marca diferente de la primera marca.

50 En algunas realizaciones, la segunda porción de ADN de ensayo tiene agotado el ADN metilado. En algunas realizaciones, la segunda porción de ADN de ensayo tiene agotado el ADN no metilado. En algunas realizaciones, la segunda porción de ADN se agota mediante tratamiento del ADN escindido al azar o sometido a cizalla con una enzima de restricción sensible a metilo o dependiente de metilo y la selección del ADN no escindido.

55 En algunas realizaciones, las poblaciones de ADN son de una planta. En algunas realizaciones, las poblaciones de ADN son de un animal. En algunas realizaciones, las poblaciones de ADN son de un hongo. En algunas realizaciones, las poblaciones de ADN son de un procariota. En algunas realizaciones, el procariota es un patógeno bacteriano. En algunas realizaciones, el patógeno bacteriano se selecciona del grupo que consiste en bacterias

gram negativas y gram positivas, que incluyen *Listeria*, *E. coli*, *Salmonella*, *Yersinia*, y *Neisseria*, y micobacterias. En algunas realizaciones, las poblaciones de ADN son de un organismo, célula, o tejido transgénicos.

5 En algunas realizaciones, la micromatriz de polinucleótidos comprende promotores génicos y/o secuencias de polinucleótidos que cuando se metilan, silencian la expresión del gen próximo.

Asimismo se describen en la presente memoria métodos para producir una población epigenéticamente uniforme o diversa de la progenie de uno o más individuos parentales. El método puede comprender las etapas de:

10 a. determinar el perfil de metilación genómica de la progenie propagada sexualmente o asexualmente de un individuo parental; y

b. seleccionar la progenie que muestra un perfil de metilación uniforme o diverso, produciendo de ese modo una población epigenéticamente uniforme de uno o más individuos parentales.

15 El método puede comprender adicionalmente la determinación del perfil de metilación de un individuo parental y la etapa de selección puede comprender la selección de la progenie que muestra el perfil de metilación del individuo parental. El parental puede ser un híbrido de la F1. La progenie se puede propagar sexualmente. La progenie se puede propagar asexualmente. El individuo parental puede ser una planta. El individuo parental puede ser un animal. El individuo parental puede ser un hongo. El individuo parental puede ser un procarionta. La progenie pueden ser clones del parental.

20 El perfil de metilación genómica se puede determinar sobre un soporte sólido. El soporte sólido puede ser una membrana. El soporte sólido puede ser una columna de unión a metilo. El soporte sólido puede ser una micromatriz. El soporte sólido puede ser una cuenta.

La etapa de determinación puede comprender

30 a. separar una población de ADN uniforme sometida a cizalla o escindida al azar en fracciones metiladas y no metiladas;

b. marcar las fracciones metiladas o no metiladas con una primera marca; e

35 c. hibridar las fracciones metiladas o no metiladas con un ácido nucleico.

El método puede comprender adicionalmente proporcionar el ADN genómico total marcado con una segunda marca e hibridar el ADN genómico total a un ácido nucleico, normalizando de ese modo la señal de la primera marca.

40 El perfil de metilación genómica de cada individuo o progenie se puede determinar mediante las etapas que comprenden:

45 a. proporcionar una población de tamaño uniforme de ADN escindido al azar o sometido a cizalla de la célula, tejido, u organismo, en donde el ADN comprende una primera porción y una segunda porción y cada porción comprende nucleótidos metilados y no metilados;

b. marcar la primera porción con una primera marca;

c. agotar el ADN metilado o no metilado de la segunda porción;

50 d. marcar la segunda porción agotada con una segunda marca que es diferente de la primera marca;

e. hibridar la primera porción y la segunda porción agotada a un ácido nucleico;

55 f. determinar la metilación relativa de los fragmentos de ácido nucleico complementarios en el ADN calculando la razón de las dos marcas hibridantes, determinando de ese modo el perfil de metilación de varias de tales secuencias de ácido nucleico de una célula, tejido, u organismo.

El método puede comprender las etapas de:

60 a. proporcionar una población de tamaño uniforme de ADN escindido al azar o sometido a cizalla de la célula, tejido, u organismo, en donde el ADN comprende una primera porción y una segunda porción y el ADN comprende secuencias de reconocimiento metiladas y no metiladas de una enzima de restricción sensible a metilo o dependiente de metilo;

- b. marcar la primera porción de la población de ADN con una primera marca;
- c. escindir la segunda porción con la enzima de restricción sensible a metilo o dependiente de metilo,
- 5 d. agotar el ADN metilado o no metilado de la segunda porción;
- e. marcar el ADN no escindido de la segunda porción con una segunda marca que es diferente de la primera marca;
- 10 f. hibridar el ADN marcado de la primera y la segunda porciones con un ácido nucleico; y
- g. determinar la metilación relativa de un ácido nucleico mediante la detección de la primera y la segunda marcas que hibridan con el ácido nucleico, determinando de ese modo el perfil de metilación de la célula, tejido, u organismo.

15 La segunda porción se puede escindir con una enzima de restricción dependiente de metilo. La segunda porción puede ser escindida con una enzima de restricción sensible a metilo. La progenie se puede escrutar en grupos.

20 También se describen en la presente memoria los métodos de asociación de la heterosis con los perfiles de metilación. El método puede comprender cruzamientos individuales para producir la progenie; determinación del perfil de metilación de los individuos y la progenie; y comparación de un rasgo de la progenie con los perfiles de metilación de los individuos, asociando de ese modo la aparición del rasgo con un perfil de metilación. Los individuos pueden ser de grupos heteróticos diferentes.

25 La presente invención proporciona métodos para determinar un perfil de metilación de una célula, tejido u organismo. En algunas realizaciones, el método comprende las etapas de:

- a. proporcionar una población de tamaño uniforme de ADN escindido al azar o sometido a cizalla de la célula, tejido u organismo, en donde el ADN comprende una primera porción y una segunda porción y cada porción comprende nucleótidos metilados y no metilados;
- 30 b. marcar la primera porción con una primera marca;
- c. agotar el ADN metilado o no metilado de la segunda porción;
- 35 d. marcar la segunda porción agotada con una segunda marca que es diferente de la primera marca;
- e. hibridar la primera porción y la segunda porción agotada a un ácido nucleico;
- 40 f. determinar la metilación relativa de los fragmentos de ácido nucleico complementarios en el ADN calculando la razón de las dos marcas hibridantes, determinando de ese modo el perfil de metilación de varias de tales secuencias de ácido nucleico de una célula, tejido, u organismo.

45 En algunas realizaciones, en la segunda porción se agota el ADN metilado. En algunas realizaciones, se agota el ADN no metilado de la segunda porción.

En algunas realizaciones, el método comprende las etapas de:

- a. proporcionar la población de tamaño uniforme de ADN escindido al azar o sometido a cizalla de la célula, tejido, u organismo, en donde el ADN comprende una primera porción y una segunda porción y el ADN comprende secuencias de reconocimiento metiladas y no metiladas de una enzima de restricción sensible a metilo o dependiente de metilo;
- 50 b. marcar la primera porción de la población de ADN con una primera marca;
- 55 c. escindir la segunda porción con la enzima de restricción sensible a metilo o dependiente de metilo,
- d. agotar el ADN metilado o no metilado de la segunda porción;
- 60 e. marcar el ADN no escindido de la segunda porción con una segunda marca que es diferente de la primera marca;
- f. hibridar el ADN marcado de la primera y la segunda porciones con un ácido nucleico; y

g. determinar la metilación relativa de un ácido nucleico detectando la primera y segunda marcas que hibridan con el ácido nucleico, determinando de ese modo el perfil de la célula, tejido, u organismo.

5 En algunas realizaciones, la segunda proporción es escindida con una enzima de restricción sensible a metilo. En algunas realizaciones, la segunda porción es escindida con una enzima de restricción dependiente de metilo.

En algunas realizaciones, el ácido nucleico se conecta con un soporte sólido. En algunas realizaciones, el soporte sólido es una micromatriz. En algunas realizaciones, el soporte sólido es una cuenta.

10 En algunas realizaciones, el organismo es una planta. En algunas realizaciones, el organismo es un hongo. En algunas realizaciones, el organismo es un procarionte. En algunas realizaciones, el procarionte es un patógeno bacteriano. En algunas realizaciones, el patógeno bacteriano se selecciona del grupo que consiste en especies gram negativas y gram positivas, que incluyen *Listeria*, *E. coli*, *Salmonella*, *Yersinia*, y *Neisseria*, y micobacterias. En algunas realizaciones, el organismo es un animal. En algunas realizaciones, el animal es un ser humano. En algunas realizaciones, la célula es una célula pluripotencial. En algunas realizaciones, la célula es transgénica y el ácido nucleico corresponde a la inserción de un transgén. En algunas realizaciones, la célula es una célula pluripotencial. En algunas realizaciones, la célula es transgénica y el ácido nucleico corresponde al sitio de inserción de un transgén. En algunas realizaciones, el tejido es la sangre. En algunas realizaciones, el tejido es tejido de biopsia. En algunas realizaciones, el tejido es tejido resecado. En algunas realizaciones, el tejido es normal. En algunas realizaciones, el tejido es tejido tumoral. En algunas realizaciones, el tejido es precanceroso.

25 En algunas realizaciones, el método comprende adicionalmente la comparación del perfil de metilación de un ácido nucleico con la transcripción del ácido nucleico, determinando de ese modo la relación entre la metilación y la transcripción del ácido nucleico. En algunas realizaciones, la transcripción del ácido nucleico se detecta con una micromatriz. En algunas realizaciones, comprende adicionalmente la comparación del perfil de metilación de un espécimen de un patógeno bacteriano con una cepa de referencia del patógeno, en donde la similitud de los patrones de metilación indica el origen común del espécimen y la cepa de referencia.

Definiciones

30 "Uniforme" hace referencia a un rasgo concreto que muestra poca o ninguna variación dentro de una población. Típicamente, los individuos de una población uniforme variarán en un rasgo concreto en no más de aproximadamente 500% y en algunos casos en tan poco como aproximadamente 300%, 200%, 100%, 75%, 50%, 25%, 10%, 5% o 1% del rasgo de un individuo concreto o la media individual en la población. De un modo similar, "uniforme" o "de tamaño uniforme", cuando se utiliza en el contexto de los fragmentos de ADN en una población de ADN, hace referencia a una población con una variación no mayor de aproximadamente 500% (y en algunos casos tan pequeña como una variación de aproximadamente 300%, 200%, 100%, 75%, 50%, 25%, 10%, 5% o 1%) en la longitud del fragmento. Por ejemplo, cuando la longitud media de un fragmento de ADN es de 1.000 pares de bases, una población uniforme con una variación de 500% tendría individuos con no más de 6.000 pares de bases.

40 "Uniforme epigenéticamente" hace referencia a una población cuyos miembros individuales tienen rasgos epigenéticos uniformes. Por ejemplo, los individuos epigenéticamente uniformes tendrán una variación pequeña o nula en los perfiles de metilación entre sus genomas.

45 "Metilación" hace referencia a la metilación de la citosina en las posiciones C⁵ o N⁴ de la citosina, la posición N⁶ de la adenina u otros tipos de metilación de ácido nucleico.

50 "Separar" en el contexto de la purificación de ácidos nucleicos entre sí, hace referencia a dividir ácidos nucleicos individuales de una mezcla en dos poblaciones físicamente distintas. Se reconoce que no es necesario que cada miembro de una población sea separado de la segunda población para que se produzca la separación. Por ejemplo, separar ADN no metilado no escindido de una segunda porción de ADN implica separar al menos algo de ADN no metilado en una población separada e implica típicamente separar la mayor parte del ADN no metilado. No es necesario separar cada especie de ADN no metilado de la segunda porción para que se produzca la separación. En otro ejemplo, separar ADN metilado de una segunda porción de ADN implica separar al menos algo de ADN metilado en una población separada e implica típicamente separar la mayor parte del ADN metilado. No es necesario separar cada especie de ADN metilado escindido de la segunda porción para que se produzca la separación. "Separar" no está limitado a la escisión por restricción y a la separación por tamaño, también incluye la purificación por afinidad como se describe en la presente memoria y otros métodos conocidos por los expertos en la técnica.

60 Un "híbrido individual" hace referencia a un individuo que es la progenie directa resultante del cruce sexual de dos parentales o es, de lo contrario, un material compuesto genético de al menos dos individuos.

Un "perfil de metilación del genoma" hace referencia a un conjunto de datos que representan el estado de metilación

del ADN dentro del genoma de un individuo. El perfil puede indicar el estado de metilación de cada par de bases en un individuo o puede comprender información referente a un subconjunto de pares de bases (p. ej., el estado de metilación de una secuencia de reconocimiento de una enzima de restricción específica) en un genoma. Se conocen en la técnica numerosos métodos para la determinación del estado de metilación del ADN y se describen en la presente memoria en la presente memoria.

El término "micromatriz" hace referencia a una disposición ordenada de elementos de una matriz hibridables. Los elementos de la matriz se disponen de manera que haya preferiblemente al menos uno o más elementos de la matriz diferentes, más preferiblemente al menos 100 elementos de la matriz, y lo más preferiblemente al menos 1.000 elementos de la matriz por cm^2 de superficie de sustrato. Además, la señal de hibridación de cada uno de los elementos de la matriz es típicamente distinguible de manera individual.

Una "enzima de restricción dependiente de metilo" hace referencia a una enzima de restricción (p. ej., McrBC) que escinde una secuencia de restricción metilada pero no escinde la misma secuencia cuando la secuencia no está metilada.

Una "enzima de restricción sensible a metilo" hace referencia a una enzima de restricción (p. ej., PstI) que escinde una secuencia de restricción no metilada pero no escinde la misma secuencia cuando la secuencia no está metilada.

Una muestra "con el ADN metilado agotado" hace referencia a fragmentos de ADN de los cuales la mayor parte de los fragmentos que contienen nucleótidos metilados en una secuencia de interés (p. ej., en un sitio de reconocimiento de una enzima de restricción dependiente de metilo) han sido eliminados. En algunas realizaciones, una población con el ADN metilado agotado no contiene más de 40%, 30%, 20%, 10%, 5% o 1% de fragmentos con al menos una secuencia metilada de interés. El resto de los fragmentos de la muestra agotada puede contener nucleótidos metilados en localizaciones diferentes de la de la secuencia de interés.

Una muestra "con el ADN no metilado agotado" hace referencia a fragmentos de ADN de los cuales la mayor parte de los fragmentos que contienen nucleótidos no metilados en una secuencia de interés (p. ej., en un sitio de reconocimiento de una enzima de restricción sensible a metilo) han sido eliminados. En algunas realizaciones, una población con el ADN no metilado agotado contiene no más de 40%, 30%, 20%, 10%, 5% o 1% de fragmentos con al menos una secuencia no metilada de interés. El resto de los fragmentos de la muestra agotada puede contener nucleótidos no metilados en localizaciones diferentes de la de la secuencia de interés.

"Anticuerpo" hace referencia a un polipéptido codificado sustancialmente por un gen de inmunoglobulina o genes de inmunoglobulina o fragmentos de los mismos, que se unen específicamente y se unen a un analito (antígeno). Los genes de inmunoglobulina reconocidos incluyen los genes de la región constante kappa, lambda, alfa, gamma, delta, epsilon y mu, así como la miríada de genes de la región variable de inmunoglobulina. Las cadenas ligeras se clasifican como kappa o lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta, o epsilon, que a su vez definen las clases de inmunoglobulinas, IgG, IgM, IgA, IgD e IgE, respectivamente.

Una unidad estructural de inmunoglobulina (anticuerpo) ilustrativa comprende un tetrámero. Cada tetrámero está compuesto por dos pares idénticos de cadenas polipeptídicas, teniendo cada par una cadena "ligera" (aproximadamente 25 kD) y una "pesada" (aproximadamente 50-70 kD). El extremo N de cada cadena define una región variable de aproximadamente 100 a 110 o más aminoácidos responsables principalmente del reconocimiento de antígenos. Los términos cadena ligera variable (V_L) y cadena pesada variable (V_H) hacen referencia a estas cadenas ligera y pesada, respectivamente.

Los anticuerpos existen, p. ej., en forma de inmunoglobulinas intactas o en forma de fragmentos bien caracterizados producidos por la digestión de diferentes peptidasas. De este modo, por ejemplo, la pepsina digiere un anticuerpo por debajo de los enlaces disulfuro de la región bisagra para producir $F(ab)'_2$, un dímero de Fab que en sí mismo es una cadena ligera unida a un $V_H\text{-C}_H1$ por un enlace disulfuro. El $F(ab)'_2$ puede ser reducido en condiciones suaves para romper la conexión disulfuro de la región bisagra, convirtiendo de ese modo el dímero $F(ab)'_2$ en un monómero Fab'. El monómero Fab' es esencialmente un Fab con parte de la región bisagra (véase, Paul (Ed.) Fundamental Immunology, Tercera Edición, Raven Press, NY (1993)). Si bien varios fragmentos de anticuerpo se definen en términos de la digestión de un anticuerpo intacto, un experto en la técnica apreciará que tales fragmentos pueden ser sintetizados *de novo* o bien químicamente o bien utilizando la metodología del ADN recombinante. De este modo, el término anticuerpo, según se utiliza en la presente memoria, también incluye fragmentos de anticuerpo producidos mediante la modificación de anticuerpos completos o aquellos sintetizados *de novo* utilizando metodologías de ADN recombinante (p. ej., Fv de cadena sencilla).

"Heterosis" o "vigor híbrido" se manifiesta como el funcionamiento mejorado de un híbrido de la F1 en comparación con sus dos parentales endogámicos diferentes. La heterosis se puede definir cuantitativamente como una desviación al alza del parental medio, basándose en la media de los valores de los dos parentales. Véanse, p. ej., Shull, G. 1909. Am Breed Assoc Rep 5:51-59/Johnson et al. Genetics 134(2): 465-474 (1993). Por ejemplo, supone

que se aparean dos individuos de diferentes razas que tienen pesos de 13,61 y 18,14 kg. Su progenie, si pesaba 22,68 kg, funcionaba a un nivel por encima de cada uno de los parentales. Se supone que el peso extra, definido como la diferencia entre el nivel de funcionamiento de la progenie y el de los parentales individuales es debido a la heterosis.

Un "grupo heterótico" es una población de genotipos que, cuando se cruza con individuos de otro grupo o población heteróticos, superan de manera regular los cruces intra-población. Véanse, *p. ej.*, Hallauer, et al. QUANTITATIVE GENETICS IN MAIZE BREEDING (Iowa State Univ., Ames, IA 1988); Hallauer, et al., "Corn Breeding" En (ed.) CORN AND CORN IMPROVEMENT 3ª. (ASA-CSSA-SSSA, Madison, WI. 1988), págs. 463-564; Lee et al., Crop. Sci. 29, págs. 1067-1071 (1989); Livini C., et al. Theor. Appl. Genet. 84, págs. 17-25 (1992); Smith et al., Maydica 37, págs. 53-60 (1992).

Descripción detallada de la invención

1. Introducción

El efecto de la metilación del ADN sobre la expresión génica es tanto regional como profundo. Las alteraciones en la metilación genómica dan lugar a la expresión inapropiada de los genes contiguos. Por consiguiente, la capacidad para inspeccionar los estados de metilación de múltiples regiones del genoma (o determinar un "perfil de metilación") permite la asociación de estados de metilación específicos con la expresión génica o con rasgos.

La presente invención proporciona métodos útiles para identificar y seleccionar individuos con una metilación genómica similar o idéntica y de ese modo seleccionar una población de individuos que tienen los fenotipos deseados. Estos métodos son útiles, por ejemplo, para seleccionar individuos de una población (*p. ej.*, progenie propagada sexualmente o asexualmente) que conservan los rasgos deseados. Además, los métodos son útiles para identificar parejas de apareamiento óptimo y grupos heteróticos óptimos para la generación de una progenie que tiene vigor híbrido. Además, los métodos y composiciones son útiles para el diagnóstico del cáncer, la identificación de biomarcadores predictivos, y el descubrimiento de nuevas dianas de fármacos. El papel del cáncer y la metilación es comentado por Jones y Baylin, Nat Rev Genet. 3(6): 415-28 (2002).

Esta invención se basa en técnicas rutinarias en el campo de la genética recombinante. Los textos básicos que describen los métodos generales de uso de esta invención incluyen Sambrook et al., Molecular Cloning, A Laboratory Manual (3ª ed. 2001); Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990); y Current Protocols in Molecular Biology (Ausubel et al., eds., 1994)).

2. Determinación de los perfiles de metilación

La presente invención proporciona métodos para determinar los perfiles de metilación de ácidos nucleicos, incluyendo perfiles de metilación de genomas completos. Los métodos de la invención comprenden generar una población de tamaño uniforme de ADN fragmentado (*p. ej.*, escindido al azar o sometido a cizalla) y generar muestras de ADN que consisten en ADN metilado y/o no metilado. Los perfiles de metilación de un ácido nucleico se pueden determinar a continuación cuantificando las cantidades relativas del ácido nucleico entre dos cualesquiera de los siguientes: ADN total, ADN metilado o ADN no metilado, esto es, muestras con ADN metilado o no metilado agotado, respectivamente.

Generalmente, estas muestras se generan dividiendo el ADN fragmentado en dos porciones iguales (una "primera porción" y una "segunda porción") y separando a continuación la segunda porción en sub-porciones de ADN metilado y no metilado. La cantidad relativa de un fragmento que contiene una secuencia de ácido nucleico se determina después en cualquiera de:

- a. la primera porción comparada con la sub-porción de ADN metilado;
- b. la primera porción comparada con la sub-porción de ADN no metilado; o
- c. la sub-porción de ADN no metilado comparada con la sub-porción de ADN metilado.
- d. la sub-porción de ADN no metilado comparada con la sub-porción de ADN metilado comparada con la primera porción.

ADN fragmentado

Como se comenta más arriba, en muchas realizaciones, el AD genómico de partida está fragmentado. La fragmentación se puede llevar a cabo mediante cualquier método conocido por los expertos en la técnica (*p. ej.*, sometido a cizalla mecánicamente, escindido con una enzima de restricción o una ADNasa I, etc.). En algunas realizaciones, se aísla una población de fragmentos de tamaño uniforme (*p. ej.*, mediante electroforesis en gel de agarosa y elución de un intervalo concreto de tamaños de fragmentos). Por ejemplo, el tamaño medio de los

fragmentos puede ser, p. ej., de aproximadamente 0,1, 0,5, 1, 2, 3, 4, 5 kb o más. En algunas realizaciones, el tamaño medio de los fragmentos oscila entre, p. ej., 0,1-1, 1-2, 1-3, 1-5, 2-4, o 1-10 kb.

Separación de ADN metilado y no metilado

Se pueden utilizar numerosos métodos para separar ADN en sub-porciones de ADN metilado o no metilado.

En algunas realizaciones, esto se puede lograr, por ejemplo, escindiendo el ADN genómico fragmentado de una longitud uniforme con una endonucleasa de restricción sensible a metilo (o alternativamente dependiente de metilo) para separar una o dos sub-porciones: una sub-porción de moléculas de ADN no escindido y una sub-porción de moléculas de ADN escindidas. Cuando se utilizan enzimas de restricción dependientes de metilo (que escinden secuencias metiladas pero no secuencias no metiladas), la sub-porción de fragmentos de ADN no escindido representará secuencias de restricción no metiladas y la sub-porción de fragmentos de ADN escindido representará secuencias de restricción metiladas. Por el contrario, cuando se utiliza una enzima de restricción sensible a metilo (que escinde secuencias no metiladas pero no secuencias metiladas), la sub-porción de fragmentos de ADN no escindido representará secuencias de restricción metiladas y la sub-porción de fragmentos de ADN escindido representará secuencias de restricción no metiladas.

Numerosas enzimas de restricción dependientes de metilo y sensibles a metilo son conocidas por los expertos en la técnica. Las enzimas de restricción se pueden obtener generalmente, p. ej., de New England Biolabs (Beverly, MA) o de Roche Applied Sciences (Indianapolis, IN). Las enzimas de restricción dependientes de metilo ilustrativas incluyen, p. ej., McrBC, McrA, MrrA, y DpnI. Las enzimas de restricción sensibles a metilo ilustrativas incluyen, p. ej., PstI, BstNI, FseI, MspI, CfoI, y HpaII. Véanse p. ej., McClelland, M. et al, *Nucleic Acids Res.* 1994 Sep;22(17):3640-59 y <http://rebase.neb.com>.

Las dos sub-porciones de moléculas de ADN (esto es, poblaciones escindidas y no escindidas) se pueden separar por el peso molecular utilizando numerosos métodos conocidos por los expertos en la técnica. Por ejemplo, se pueden utilizar la electroforesis en gel, cromatografía de exclusión por tamaños, centrifugación diferencial por tamaños (p. ej., en un gradiente de sacarosa) para separar fragmentos escindidos de fragmentos no escindidos más pesados.

Los expertos en la técnica reconocerán que también se pueden utilizar otros métodos de separación de poblaciones metiladas y no metiladas, agotando de ese modo la muestra de ADN metilado o no metilado. Por ejemplo, se pueden utilizar anticuerpos u otros agentes (p. ej., MeCP2) específicos para ácidos nucleicos o proteínas no metilados asociados con ácidos nucleicos metilados para purificar por afinidad los ácidos nucleicos metilados, separando de ese modo el ADN metilado del ADN no metilado. Véase, p. ej., Meehan, et al., *Nucleic Acids Res.* 20(19):5085-92 (1992). En este caso, el ADN puede, pero no necesita, ser escindido con una endonucleasa de restricción que percibe la metilación. En algunas realizaciones, por ejemplo, se utiliza una cromatografía de afinidad que comprende una proteína específica para ADN metilado para separar las fracciones metiladas y no metiladas. Una vez separado en fracciones, o bien cada fracción o bien ambas fracciones se pueden marcar para la hibridación.

En otras realizaciones, se utilizan agentes químicos, solos o en concierto con enzimas, capaces de escindir específicamente ácidos nucleicos metilados para generar poblaciones metiladas y no metiladas. Las poblaciones se pueden separar a continuación como se ha descrito anteriormente.

Pre-amplificación de las sub-porciones

Una vez que los fragmentos de ADN se han separado en una primera porción que comprende el ADN total y las sub-porciones de ADN metilado y ADN no metilado, existen numerosas maneras conocidas por los expertos en la técnica para amplificar uniformemente los fragmentos de cada sub-porción antes de que cualquier ácido nucleico sea cuantificado en las sub-porciones. Por ejemplo, la pre-amplificación de los fragmentos reforzará la señal de cualquier ácido nucleico específico dentro de una sub-porción y permitirá el perfilado de metilación de las secuencias específicas en cantidades vestigiales de ADN de partida presente. Tales técnicas son útiles cuando solamente se encuentran disponibles pequeñas muestras de ADN genómico, por ejemplo en muestras de tejido de biopsia o tumores resecados. En un ejemplo, se ligan adaptadores de ADN de doble hebra a los extremos del fragmento de ADN de las sub-porciones. A continuación se añaden oligonucleótidos específicos para los adaptadores de ADN a cada sub-porción y a continuación se amplifica la población de fragmentos de ADN de la sub-porción utilizando técnicas de amplificación de ADN lineal (p. ej., círculo rodante) o exponencial (p. ej. PCR), por ejemplo.

Cuantificación de ADN

La cuantificación de un ácido nucleico en las muestras de ADN (esto es, la primera porción, la sub-porción de ADN metilado y/o la sub-porción no metilada) se puede llevar a cabo mediante cualquier método conocido por los expertos en la técnica.

5 **Hibridación**

Por ejemplo, se puede utilizar la hibridación simple para cuantificar la secuencia de ácido nucleico en las muestras de ADN. En un ejemplo, las dos o más muestras se marcan con diferentes marcas (p. ej., marcas fluorescentes o detectables de otro modo) y las señales relativas de las diferentes marcas se determinan mediante métodos convencionales después de la hibridación de las muestras marcadas al ácido nucleico. La hibridación de una sonda de ADN a una secuencia metilada concreta indica que la secuencia está metilada. La ausencia de hibridación de la sonda indica que la secuencia no está metilada. De un modo similar, el ADN no metilado de un individuo se puede utilizar como diana, en donde la hibridación indica que la secuencia no está metilada en el individuo.

En algunas realizaciones, las muestras se hibridan a sondas sobre una micromatriz. Esta realización es particularmente útil para determinar el perfil de metilación de un gran número de secuencias, incluyendo, p. ej., un conjunto de sitios que comprenden el genoma completo.

En algunas realizaciones, se medirá la hibridación de ADN metilado a una sonda dada y de ADN no metilado a una sonda dada, y estas mediciones se compararán entre sí. Esto permite, por ejemplo, una determinación de la intensidad de hibridación relativa del ADN diana no metilado y metilado a una sonda dada.

En algunas realizaciones, la hibridación a una sonda dada de las sub-porciones metilada y no metilada se medirá y se comparará con la hibridación medida a una sonda dada del ADN genómico total, esto es, la "primera porción". El ADN genómico total actúa como una referencia para normalizar los datos de la hibridación de las sub-porciones de ADN metilado o no metilado. Esto permite, por ejemplo, la determinación de las intensidades de hibridación relativas a una sonda dada entre el ADN diana metilado y el ADN diana total. En los casos en los que el ADN diana hibrida con más de una secuencia en el genoma, la hibridación del ADN diana total permite la determinación de cuántas copias de la secuencia hibridan. Si la hibridación del ADN metilado solamente da como resultado una fracción de la señal producida por la diana de ADN total, el usuario puede calcular qué fracciones de las secuencias hibridantes están metiladas.

Un ácido nucleico sonda representa la secuencia de ácido nucleico a la cual se hibridan la diana o las dianas. Típicamente, el ácido nucleico sonda es al menos un fragmento de ADN genómico. La hibridación diferencial (que se determina controlando las dos o más marcas diferentes) indica la metilación relativa en una secuencia genómica concreta.

Los ácidos nucleicos de la sonda pueden ser cualquier secuencia. En algunas realizaciones, se sondean numerosos ácidos nucleicos diana diferentes, proporcionando de ese modo información sobre el estado de metilación de cada diana. En algunas realizaciones, los ácidos nucleicos de la sonda representan sitios de metilación conocidos u otras secuencias de ácido nucleico de interés (p. ej., una secuencia cuya metilación está asociada con un fenotipo tal como el cáncer). Alternativamente, los ácidos nucleicos de la sonda son secuencias aleatorias o expresadas.

Para procesar la información sobre un gran número de secuencias de ADN (esto es, sondas) en el genoma, puede ser conveniente hibridar las dos poblaciones marcadas a una micromatriz u otra matriz de sondas dirigidas. El número de sondas escrutadas puede ser, p. ej., de al menos aproximadamente 2, 5, 10, 20, 50, 100, 500, 1000, 10000 o más fragmentos. En algunas realizaciones, los ácidos nucleicos de la sonda se presentan sobre un soporte sólido. Los soportes sólidos ilustrativos incluyen, p. ej., cuentas o una micromatriz. En algunas realizaciones, las secuencias diana se presentan sobre un soporte sólido.

Para los fines de la siguiente discusión, las sondas hacen referencia a elementos de ácido nucleico sobre una micromatriz y los ácidos nucleicos diana hacen referencia a fracciones de ácidos nucleicos metilados o no metilados o a ácidos nucleicos genómicos totales. Cuando las sondas se emplean como elementos de una matriz hibridables sobre una micromatriz, los elementos de la matriz se organizan de una manera ordenada de manera que cada elemento esté presente en una localización especificada sobre el sustrato. Debido a que los elementos de la matriz están en localizaciones especificadas sobre el sustrato, los patrones de hibridación y las intensidades, incluyendo la hibridación diferencial de las dianas (que crean juntos un perfil de expresión único) pueden ser interpretados en términos de perfiles de metilación y se pueden correlacionar con un fenotipo (p. ej., vigor híbrido o cáncer).

Se puede analizar la hibridación diferencial del ADN total y el ADN metilado, el ADN total y el ADN no metilado, el ADN no metilado y el ADN metilado, o el ADN total y el ADN metilado y el ADN no metilado. Para la hibridación diferencial, se preparan al menos dos muestras de ADN diana diferentes y se marcan con radicales de marcaje diferentes. La mezcla de las dos o más muestras de ADN marcadas se añade a una micromatriz. La micromatriz se

examina a continuación en condiciones en las que las emisiones de cada una de las dos o más marcas diferentes es detectable individualmente.

5 En algunas realizaciones, las marcas son marcas fluorescentes con un espectro de emisión distinguible, tales como un análogo de nucleótido conjugado con lissamina y un análogo de nucleótido conjugado con fluoresceína. En otra realización, se emplean los fluoróforos Cy3/Cy5 (Amersham Pharmacia Biotech). Por ejemplo, para las aplicaciones en micromatrices, puede ser conveniente utilizar marcas fluorescentes (p. ej., Cy3 o Cy5) que sean fácilmente detectadas. Sin embargo, los expertos en la técnica reconocerán que se puede emplear cualquier tipo de marca detectable (p. ej., radioactiva, fluorescente, enzimática, u otros métodos conocidos por los expertos en la técnica).

10 Después de la hibridación, la micromatriz se lava para eliminar los ácidos nucleicos no hibridados, y se detecta la formación de complejos entre los elementos de la matriz hibridables y las sondas. Los métodos para detectar la formación de complejos son bien conocidos por los expertos en la técnica. Como se ha comentado anteriormente, en algunas realizaciones, los polinucleótidos diana se marcan con una marca fluorescente, y la medida de los niveles y los patrones de fluorescencia indicativos de la formación del complejo se completa mediante microscopía de fluorescencia, tal como microscopía de fluorescencia confocal. Un láser de ión argón excita la marca fluorescente, las emisiones se dirigen a un fotomultiplicador, y se detecta y se cuantifica la cantidad de luz emitida. La señal detectada debe ser proporcional a la cantidad de complejo de sonda/polinucleótido diana en cada posición de la micromatriz. El microscopio de fluorescencia puede estar asociado con un dispositivo de escáner dirigido por un ordenador para generar una imagen bi-dimensional cuantitativa de la intensidad de hibridación. La imagen escaneada se examina para determinar la abundancia de cada polinucleótido diana hibridado.

15 20 En un experimento de hibridación diferencial, los polinucleótidos diana de dos o más muestras biológicas diferentes se marcan con dos o más marcas fluorescentes diferentes con diferentes longitudes de onda de emisión. Las señales fluorescentes se detectan por separado con diferentes fotomultiplicadores ajustados para detectar longitudes de onda específicas. Se obtienen los niveles relativos de abundancia/expresión de los polinucleótidos diana en dos o más muestras.

25 30 Típicamente, las intensidades de fluorescencia de la micromatriz se pueden normalizar para tomar en consideración las variaciones en las intensidades de hibridación cuando se utiliza más de una micromatriz en condiciones de ensayo similares. En algunas realizaciones, las intensidades de hibridación del complejo de la sonda de polinucleótido/diana individual se normalizan utilizando intensidades derivadas de controles de normalización interna contenidos en cada micromatriz o a partir de la intensidad de hibridación del ADN genómico total.

35 **Amplificación cuantitativa**

Las secuencias de ácido nucleico de una muestra de ADN (p. ej., la primera porción de una sub-porción) también se pueden determinar mediante cualquiera de los numerosos mecanismos de amplificación cuantitativa conocidos por los expertos en la técnica (p. ej., PCR cuantitativa o amplificación lineal cuantitativa). Los métodos de amplificación cuantitativa son descritos, p. ej., en las Patentes de los Estados Unidos Núms. 6.180.349; 6.033.854; y 5.972.602, así como, p. ej., por Gibson et al., Genome Research 6:995-1001 (1996); DeGraves, et al., Biotechniques 34(1):106-10, 112-5 (2003); Deiman B, et al., Mol Biotechnol. 20(2):163-79 (2002).

45 50 Un método para la detección de productos de amplificación es el análisis mediante PCR con nucleasa 5' (también referido como análisis TaqMan™) (Holland et al., Proc. Natl. Acad. Sci. USA 88: 7276-7280 (1991); Lee et al., Nucleic Acids Res. 21: 3761-3766 (1993)). Este análisis detecta la acumulación de un producto de PCR específico mediante hibridación y escisión de una sonda fluorogénica doblemente marcada (la sonda "TaqMan™") durante la reacción de amplificación. La sonda fluorogénica consiste en un oligonucleótido marcado tanto con un colorante informador fluorescente como con un colorante extintor. Durante la PCR, esta sonda es escindida por la actividad exonucleasa 5' de la ADN polimerasa si, y solamente si, ésta hibrida con el segmento que se está amplificando. La escisión de la sonda genera un incremento en la intensidad de fluorescencia del colorante informador.

55 60 Otro método de detección de los productos de amplificación que se basa en el uso de la transferencia de energía es el método de la "sonda baliza" descrito por Tyagi y Kramer (Nature Biotech. 14:303-309 (1996)), que también es el tema de las Patentes de los Estados Unidos Núms. 5.119.801 y 5.312.728. Este método emplea sondas de hibridación a oligonucleótidos que pueden formar estructuras en horquilla. En un extremo de la sonda de hibridación (extremo 5' o 3'), existe un fluoróforo donador, y en el otro extremo, un radical aceptor. En el caso del método de Tyagi y Kramer, este radical aceptor es un extintor, esto es, el aceptor absorbe la energía liberada por el donador, pero en ese caso no fluoresce por sí mismo. De este modo, cuando la baliza está en la conformación abierta, la fluorescencia del fluoróforo donador es detectable, mientras que cuando la baliza está en conformación de horquilla (cerrada), la fluorescencia del fluoróforo donador es extinguida. Cuando se emplea en la PCR, la sonda baliza molecular, que hibrida con una de las hebras del producto de la PCR, está en "configuración abierta", y se detecta fluorescencia, mientras aquellas que permanecen no hibridadas no emitirán fluorescencia (Tyagi y Kramer, Nature Biotechnol. 14: 303-306 (1996)). Como resultado, la cantidad de fluorescencia aumentará a media que la cantidad de

producto de la PCR aumente, y de este modo se puede utilizar como una medida del progreso de la PCR. Los expertos en la técnica reconocerán que también se encuentran disponibles otros métodos de amplificación cuantitativa.

5 **Métodos adicionales de perfilado de la metilación**

Los perfiles de metilación se pueden detectar de numerosas maneras adicionales conocidas por los expertos en la técnica. Por ejemplo, el análisis de hibridación simple (p. ej., transferencia Southern) de ácidos nucleicos escindidos con endonucleasas de restricción sensibles a metilo o dependientes de metilo se puede utilizar para detectar los patrones de metilación. Típicamente, estos métodos implican el uso de una o más dianas que hibridan con al menos una secuencia que puede estar metilada. La presencia o ausencia de metilación de una secuencia de restricción se determina por medio de la longitud del polinucleótido que hibrida con la sonda. Este y otros métodos para detectar la metilación de ADN, tales como la secuenciación con bisulfito, son descritos, p. ej., por Thomassin et al., *Methods* 19(3):465-75 (1999).

3. Usos del perfilado de la metilación

A. Métodos generales

El perfilado de la metilación es útil para pronosticar cualquier fenotipo asociado con un patrón de metilación concreto. Una vez que se ha establecido semejante relación, el perfilado de la metilación es un método eficaz para identificar células individuales u organismos que tienen el mismo perfil o un similar al del individuo con el fenotipo deseado. Por ejemplo, los perfiles de metilación se pueden asociar con rasgos agronómicamente útiles en plantas o animales, o en el campo de la medicina, con tipos de cáncer específicos, permitiendo de ese modo el diagnóstico o el tratamiento.

La asociación de un fenotipo deseado con un patrón de metilación se puede producir de cualquier manera. En un ejemplo simple, el fenotipo concreto de un individuo (p. ej., una célula u organismo) es deseado en una población (p. ej., progenie o clones del individuo). En tales casos, se determina el perfil de metilación del individuo y se seleccionan los individuos de la población que tienen el mismo perfil o un similar al del individuo con el fenotipo deseado. Alternativamente, se puede evitar un perfil de metilación concreto seleccionando los individuos que carecen del perfil de metilación de una célula u organismo parentales. En otras realizaciones, se seleccionan la progenie o los clones para que tengan patrones de metilación diferentes de cualquier parental.

Un método útil para correlacionar los fenotipos deseados con los perfiles de metilación implica determinar una correlación de la metilación con la transcripción. De este modo, se determina la transcripción de uno o un conjunto de transcritos para diferentes individuos o células y después se correlaciona con un patrón de metilación concreto. La transcripción se puede determinar mediante cualquier método conocido por los expertos en la técnica. Los métodos particularmente útiles para determinar la transcripción de un gran número de genes implica el análisis con micromatrices (p. ej., "GeneChip").

Además, los métodos de perfilado de la metilación de la presente invención se pueden combinar con la hibridación comparativa del genoma (CGH). La CGH es un método para detectar deleciones y amplificaciones en una muestra de ADN genómico con respecto a otra muestra individual. Esto se realiza comparando la intensidad de hibridación de los rasgos de la micromatriz para cada muestra diana, marcada cada una con colorantes fluorescentes diferentes.

La CGH se puede combinar con los métodos de perfilado de la metilación de la presente invención debido a que una de las dianas marcadas para la hibridación a la micromatriz es una muestra de ADN total, que se puede comparar con otra de tales muestras de otro experimento de perfilado de la metilación.

En esta solicitud, el ADN genómico de dos individuos, líneas celulares, u organismos diferentes, por ejemplo, se somete a cizalla o se escinde al azar para crear fragmentos de ADN de tamaño uniforme. Una porción (p. ej., la mitad) de cada muestra se digiere a continuación con una enzima sensible a la metilación o dependiente de la metilación, como se describe en la presente memoria. Las cuatro muestras se vuelven a fragmentar a continuación para aislar el ADN total aislado y cualquiera de las sub-porciones de ADN metilado o no metilado de cada individuo. Estas cuatro muestras se pueden hibridar después con un ácido nucleico, p. ej., una micromatriz. En algunas realizaciones, las cuatro muestras se marcan con cuatro marcas diferentes (p. ej., colorantes fluorescentes). La razón de muestras de ADN total proporciona el perfil de CGH, mientras la razón de muestras agotadas y totales proporciona el perfil de metilación de cada individuo.

Alternativamente, las muestras marcadas se pueden hibridar en pares alternantes con las micromatrices. En semejante diseño, cada una de las muestras se marca con cada uno de los dos colorantes, lo que permite el análisis simultáneo de todas las muestras para las deleciones y los cambios por metilación. Este tipo de análisis es referido a veces como diseño en bucle (véase, p. ej., Craig, et al. 2001. "Designing microarray experiments: chips, dips, flips, and skips" en *PROCEEDINGS OF APPLIED STATISTICS IN AGRICULTURE*, 2001 (Ed. George Milliken)).

Un ejemplo de semejante diseño se ilustra a continuación. En el ejemplo de más abajo, los dos individuos se representan como A y B.

Número de Experimento de hibridación en matriz	Cy3	Cy5
1	A	A agotado
2	A agotado	B
3	B	B agotado
4	B agotado	A

5 En este diseño, cada muestra se marca con cada colorante, permitiendo tener en consideración la incorporación relativa de colorante. Solamente se utilizan cuatro matrices, minimizando de ese modo la cantidad de tiempo y los recursos necesarios para analizar las diferencias en el perfil de metilación a través de un genoma completo entre individuos. Alternativamente, la utilización de cuatro colorantes diferentes permitiría generar los mismos datos a partir de la hibridación de una única matriz.

10

B. Selección de las poblaciones deseadas

Las poblaciones epigenéticamente uniformes se pueden identificar y seleccionar determinando el perfil de metilación de los individuos y seleccionando a continuación aquellos individuos con perfiles similares. Estos métodos son útiles, por ejemplo, para aislar clones asexuales de un individuo con un fenotipo deseado, identificando de ese modo los clones con el mismo fenotipo. En algunas realizaciones, al menos aproximadamente 80%, 85%, 90%, 95%, 98%, o 99% de los clones seleccionados tienen un perfil de metilación sustancialmente idéntico al parental asexual.

15

Es probable que los clones o la progenie con un perfil de metilación similar o idéntico a un individuo con un fenotipo deseado (p. ej., vigor híbrido) tengan el fenotipo deseado. Cuando se produce una progenie asexual a partir de un individuo (p. ej., propagación por vía asexual de una planta o animal híbrido vigoroso, por medio de trasplante nuclear, micropropagación, por medio de división celular de células pluripotenciales, etc.), los clones son genéticamente idénticos al individuo pero difieren epigenéticamente. Seleccionando clones con perfiles de metilación similares o idénticos al híbrido, se pueden seleccionar clones que mantienen el fenotipo vigoroso del híbrido. La progenie propagada asexualmente es genéticamente idéntica al híbrido. Por lo tanto, los métodos descritos en la presente memoria son útiles para identificar las progenes asexuales que son genéticamente y epigenéticamente las mismas y por lo tanto tienen el mismo fenotipo que el parental.

20

25

Los expertos en la técnica reconocerán que la uniformidad de la población de clones seleccionada dependerá de cuán similar sea el perfil entre el híbrido y la población de clones seleccionada. Por ejemplo, el usuario puede decidir que no es necesaria la identidad absoluta con el híbrido en todos los loci y por lo tanto se puede seleccionar una progenie que tenga un porcentaje deseado de loci que sean idénticos. Por ejemplo, se pueden seleccionar los clones si al menos aproximadamente 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99% o 100% de los loci tienen el mismo estado de metilación que el híbrido. Alternativamente, se puede controlar la calidad de los loci individuales en una población (p. ej., progenie) para determinar la importancia relativa de loci concretos en el vigor híbrido. En este caso, el usuario puede seleccionar clones que tengan loci importantes de identidad completa conocidos por controlar o afectar al fenotipo deseado, a la vez que permiten al menos alguna variación, y a veces una variación completa, en otros loci.

30

35

Se pueden escrutar los individuos de acuerdo con los métodos de la invención individualmente, o en grupos (esto es, reservas). El agrupamiento de los individuos permite un rápido procesamiento de un gran número de individuos.

40

La invención se puede utilizar a lo largo de una amplia gama de organismos, incluyendo hongos, animales y plantas. Por ejemplo, se puede utilizar cualquier organismo agrícola. Los animales ilustrativos son aquellos en los que se ha empleado la cría de animales, incluyendo cerdos, bóvidos, volatería, otras aves, caballos, animales de zoológicos, especies casi extintas, y similares.

45

La invención tiene uso en una amplia gama de plantas, incluyendo especies de los géneros *Anacardium*, *Arachis*, *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Carthamus*, *Cocos*, *Coffea*, *Cucumis*, *Cucurbita*, *Daucus*, *Elaeis*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*,

Lolium, Lupinus, Lycopersicon, Malus, Manihot, Majorana, Medicago, Nicotiana, Olea, Oryza, Panieum, Pannesetum, Persea, Phaseolus, Pistachia, Pisum, Pyrus, Prunus, Raphanus, Ricinus, Secale, Senecio, Sinapis, Solanum, Sorghum, Theobromus, Trigonella, Triticum, Vicia, Vitis, Vigna, y Zea. Una vez que se seleccionan los clones o la progenie, éstos se pueden cultivar, produciendo de ese modo un cultivo de plantas que presentan un rasgo deseado uniforme.

En una realización, se utiliza el perfil de metilación de plantas propagadas asexualmente (p. ej., palmeras) para seleccionar una población uniforme que comprende plantas con patrones de metilación similares o idénticos.

La presente invención también se puede utilizar para escrutar poblaciones de células para determinar los fenotipos deseados. Las células ilustrativas incluyen células pluripotenciales, incluyendo células pluripotenciales adultas o fetales, o cualquier otra célula u organismo en el que se pueda producir una variación somaclonal dentro de una población. De este modo, la presente invención permite controlar la presencia de variación y seleccionar individuos que tengan la variación o carezcan de ella. De un modo similar, se pueden determinar los perfiles de metilación de las células cancerosas para su uso, p. ej., en el diagnóstico y el tratamiento.

En otra realización, se escrutan células u organismos transgénicos para determinar el efecto, si lo hubiera, de un transgén sobre la metilación. Los perfiles de metilación de estos individuos se pueden determinar a lo largo del genoma o se pueden determinar en una región del genoma que flanquea la inserción del transgén. Este método se puede utilizar, por ejemplo, para seleccionar eficazmente los transformantes que es más probable que no porten mutaciones deletéreas o efectos cromosómicos causados por la inserción del transgén.

C. Usos adicionales en la cría

Se pueden mejorar las técnicas de cría tradicionales determinando el perfil de metilación de pares de cría potenciales. Mediante la asociación de los patrones de metilación con rasgos heteróticos, los criadores pueden seleccionar parejas que generarán el patrón de metilación deseado en la progenie y por lo tanto darán como resultado una progenie vigorosa.

Además, los perfiles de metilación, o las secuencias individuales, se pueden identificar y utilizar para diseñar parejas óptimas. Los grupos heteróticos son poblaciones de individuos que, cuando se cruzan con individuos de otro grupo heterótico o población, superan de modo sistemático los cruces intra-población. Comparando los estados de metilación asociados (esto es, conectados) a cada grupo heterótico, y determinando los perfiles de la progenie que presentan un vigor híbrido, se pueden determinar los perfiles de metilación para las parejas de cría óptimas. Una vez que se ha determinado el perfil de metilación para un grupo heterótico concreto, se pueden determinar nuevos individuos dentro del grupo sin un cruzamiento extenso.

Ejemplos

Los siguientes ejemplos se proporcionan como ilustración solamente y no como limitación. Los expertos en la técnica reconocerán fácilmente una variedad de parámetros no críticos que se podrían cambiar o modificar para producir resultados esencialmente similares.

Ejemplo 1: Determinación de perfiles de metilación

Se someten a cizalla 20 microgramos de ADN genómico total a un intervalo de tamaño de 1-10 kb. El ADN sometido a cizalla se divide en dos partes iguales, una de las cuales se digiere por completo con 80 unidades de mcrBC durante al menos 12 horas a 37°C. El propósito de la nebulización es someter a cizalla al azar el ADN genómico de manera que el genoma completo esté igualmente representado en todos los tamaños de fragmentos de 1 kb-10 kb. Además, los tamaños pequeños de fragmentos también facilitan la purificación en gel y la incorporación de colorante Cy (véase más abajo). Las dos muestras se fragmentan a continuación sobre un gel de agarosa al 0,8% junto con los patrones de tamaño. A continuación se escinden las porciones del gel en el intervalo de tamaño de 1kb y más grandes, y el ADN se purifica utilizando columnas de centrifugación para la extracción en gel Qiagen Qiaquick de acuerdo con el protocolo de los fabricantes.

Las muestras purificadas del gel digeridas y no digeridas se marcan con Cy3 y Cy5 respectivamente utilizando el Kit de marcaje de ADN Megaprime, Amersham Pharmacia y se hibridan con micromatrices en 3xSSC al 0,3% SDS a 65°C durante la noche. Después los portaobjetos se lavan 12-16 horas en 1xSSC 0,1% SDS (1X) y 0,2X SSC (2X) a 50°C, y a continuación se escanea utilizando un escáner comercial, tal como Axon Instruments Genepix Pro 4000A Dual Laser Scanner.

Después del fondo y otras correcciones, se calcula la razón de la intensidad de señal en los canales Cy3 y Cy5. Se utiliza el análisis de intercambio de colorante para tomar en consideración la variación experimental, repitiendo la hibridación con muestras idénticas marcadas con Cy5 y Cy3, respectivamente.

La digestión con mcrBC elimina el ADN metilado de la fracción de ese tamaño, reduciendo la señal después del marcaje y la hibridación. De este modo la proporción de Cy5 con respecto a Cy3 representa la metilación relativa en la secuencia representada por la aplicación sobre la micromatriz. Este procedimiento permite analizar secuencias repetidas con un elevado número de copias en el mismo tiempo que las copias individuales, puesto que la proporción toma en consideración el número de copias.

Utilizando los métodos convencionales se aíslan aproximadamente 50 µg del ARN total. El ARN se convierte a continuación en un ADNc de la primera hebra utilizando cebadores oligo-dT o al azar. El ADNc convertido se marca con Cy3/Cy5 utilizando kits de cebado al azar disponibles en el mercado. El ADNc marcado fluorescentemente se hibrida después con una micromatriz que contiene los genes de interés, se lava y se escanea utilizando condiciones similares a las descritas anteriormente. Después del fono y otras correcciones, se calcula la proporción de la intensidad de la señal en los canales Cy3 y Cy5. Se utiliza el análisis de intercambio de colorante para tomar en consideración la variación experimental, repitiendo la hibridación con muestras idénticas marcadas con Cy5 y Cy3, respectivamente, y se pueden establecer las correlaciones con la metilación de ADN en las regiones idénticas comparando los conjuntos de datos obtenidos con el chip.

Ejemplo 2: Correlación de los perfiles de metilación con la transcripción

El perfil de metilación de un tramo de 1,7 Mb de la región de botón heterocromático de *Arabidopsis thaliana* sobre el cromosoma 4 se estableció utilizando las técnicas especificadas en el Ejemplo 1. Aquí, los autores de la presente invención demuestran que el patrón de metilación de ADN de esta región se corresponde con su patrón de expresión génica. Los autores de la presente invención también demuestran, utilizando el análisis de las mutaciones, que alterando genéticamente el patrón de metilación del ADN para la región, se altera profundamente el patrón de expresión génica para el locus. Finalmente, los autores de la presente invención demuestran que la modificación de la cromatina en la región (medida mediante un análisis diferente) confirma independientemente la correlación de los autores de la presente invención de la metilación del ADN y la expresión génica.

El cromosoma 4 de *Arabidopsis* contiene una región heterocromática o botón en el brazo corto. Los autores de la presente invención amplificaron amplicones de 1 kb sucesivos que abarcaban la región de botón y generaron una micromatriz de depósito que representaba la región. Los autores de la presente invención determinaron el perfil de metilación de la región, así como la asociación con histonas modificadas (H3K9) que están enriquecidas en ADN silencioso. Finalmente, los autores de la presente invención midieron la expresión génica y determinaron el perfil transcripcional para la región.

Se determinaron los perfiles de metilación como se describe en el Ejemplo 1. Brevemente, se nebulizó el ADN genómico a un tamaño constante y después se agotó el ADN metilado por medio de digestión con mcrBC. Se purificaron grandes fragmentos, que representaban el ADN no metilado. Se purificaron el ADN no digerido y el digerido y cada uno se marcó con una marca diferente y se hibridó a una micromatriz que representaba el genoma del cromosoma 4. A continuación se determinó la razón de las dianas de hibridación.

Se examinó el efecto de la pérdida de metilación dentro de nuestra región de botón tanto sobre la expresión génica como sobre el estado de la cromatina. Los autores de la presente invención seleccionaron las mutaciones *dm1* utilizadas para desmetilar la región de interés de los autores de la presente invención. Las mutaciones *dm1* hipometilan el ADN genómico por medio de la desorganización del complejo de remodelación de la cromatina relacionado con SWI/SNF. Véase, *p. ej.*, Jeddloh et al., Nat Genet. 22(1):94-7 (1999).

Los autores de la presente invención determinaron que la heterocromatina está des-reprimida en *dm1*. Utilizando los métodos descritos más arriba, los autores de la presente invención construyeron un mapa de la región de botón de 1,7 Mb completa del cromosoma 4. Se pronosticaron las posiciones y la dirección transcripcional de los genes pronosticados. La cantidad de metilación del ADN se detectó utilizando el perfilado de la metilación por la posición a lo largo de la región. Además, se detectó la cantidad de transcripción para cada posición a lo largo de toda la región, revelando solamente un puñado de genes activos que tampoco estaban metilados.

El perfil transcripcional por la posición también se obtuvo a partir de plantas mutantes para *dm1*. En contraste con las plantas de tipo salvaje, los mutantes para *dm1* expresaban un gran número de marcos de lectura abiertos en la región, la mayoría correspondientes a transposones.

También se llevó a cabo la inmunoprecipitación de la cromatina (ChIP) sobre plantas de tipo salvaje y mutantes. La inmunoprecipitación de la cromatina implicó las siguientes etapas. Se trataron plántulas jóvenes con formaldehído y se fijaron las interacciones proteína-ADN y proteína-proteína mediante entrecruzamiento *in vivo*. La cromatina se extrajo, se sometió a sonicación, se inmunoprecipitó, y se hizo eluir. El producto eluido se trató para revertir el entrecruzamiento y el ADN se purificó del eluyente. El ADN resultante se analizó mediante PCR y análisis Southern/micromatriz.

Utilizando estos métodos, se identificaron los constituyentes de la cromatina y se cuantificó su abundancia en una base por secuencia. La histona H3 (H3) estaba metilada en la Lisina 4 o la Lisina 9. La modificación por metilación en cualquiera específica diferentes destinos para la molécula. H3mK4 es abundante en genes expresados, y H3mK9 es abundante en genes silenciosos transcripcionalmente. Se ha demostrado que las regiones genómicas que contienen H3mK9 son altamente compactas, a diferencia de H3mH4 empaquetada de manera laxa.

Utilizando los datos obtenidos como se ha descrito anteriormente, los autores de la presente invención determinaron que la metilación de las histonas se corresponde con la metilación del ADN. La histona metilada H3 (mK9) se excluyó de la incorporación a genes expresados. De un modo similar, la histona metilada H3 (mK4) se excluyó de los genes silenciosos. En la planta de tipo salvaje, los genes expresados en la región estaban asociados con H3mK4. Los genes silenciosos solamente contenían H3mK9.

Los autores de la presente invención también determinaron que la metilación del ADN heterocromático se corresponde con la metilación de la lisina 9 de la histona H3. El estado de metilación del ADN coincidía estrechamente con el estado de empaquetamiento de la cromatina del ADN genómico.

La H3mK9 heterocromática y la metilación de ADN se perdían de manera coordinada en los mutantes *ddm1*. Los mutantes *ddm1* perdieron tanto la señal de metilación como la señal de empaquetamiento de la cromatina.

En total, estos datos demuestran que la identificación del perfil de metilación de un genoma concreto, o una región del gen confiere información de la expresión génica y del estado de empaquetamiento de la cromatina de los loci. El perfilado transcripcional ayudará a verificar los datos de expresión génica y a identificar el grado de correlación. ChIP puede proporcionar información similar, pero es mucho más laborioso y requiere la manipulación de más datos para determinar el grado de correlación.

Ejemplo 3: Detección de la hibridación ADN/ADN a micromatrices de un genoma complejo, grande

Se sometió a cizalla ADN genómico humano placentario femenino (Sigma) utilizando GeneMachines Hydroshear para un intervalo de tamaño uniforme entre 1 y 4 Kb. El ADN se aisló de las porciones de gel de agarosa de bajo punto de fusión después de la electroforesis. La concentración del ADN eluido se midió utilizando un espectrofotómetro de barrido Nanodrop. Aproximadamente 1 µg del ADN fue marcado por cebado al azar utilizando la incorporación directa de Cy3-dCTP con el kit Bioprime (Invitrogen), y una reacción paralela incorporó Cy5-dCTP. Los ADN marcados se purificaron después de la reacción de síntesis, y la incorporación de los colorantes Cy se verificó utilizando el barrido del espectro 1000 (NanoDrop). El cóctel de hibridación contenía reacciones de síntesis, ADN C0t-1 humano no marcado para suprimir el fondo (Invitrogen), oligonucleótidos de control positivo Agilent/Operon, y oligonucleótidos para suprimir la hibridación en cualquier poli-A (o T) presente en los ADNc. La matriz de ADNc del catálogo humano 1 de Agilent se hibridó durante la noche de acuerdo con las instrucciones de Agilent a 65°C.

Las matrices se lavaron a continuación tres veces, utilizando incubaciones de cinco minutos, en cantidades variables de SSC y SDS. Las matrices se secaron al aire en una centrífuga, y a continuación se escanearon utilizando el escáner y el programa Agilent (vA6.1). La intensidad de hibridación de cada rasgo se extrajo de los archivos TIFF utilizando el programa de extracción de rasgos de Agilent (v A6.1). Los archivos de datos se importaron a continuación a Genespring 5.0 (Silicon Genetics) para su visualización. El experimento se llevó a cabo después de tres días y se sometió a normalización de Lowess (normalización aplicación-a-aplicación y matriz-a-matriz) y se determinó el conjunto de datos promediados.

La matriz contiene 18.560 rasgos, que representan 12.814 EST (los -5000 rasgos restantes son controles patentados de Agilent). Los resultados demuestran que la detección de la secuencia codificante utilizando dianas de genoma humano completo marcadas es posible. En líneas generales, 5,5% (1.022/18.560 rasgos) de las sondas produjeron un escaso rendimiento, puesto que la intensidad de fluorescencia medida es menor de 100 U en cualquiera de los canales. No obstante, en este conjunto de 1.022 rasgos, solamente 260 fueron realmente EST, mientras los 762 restantes fueron rasgos de control de Agilent. Los rasgos génicos de bajo rendimiento representaban solamente un 2% de los de la matriz (260/12.814). Debido a que los datos siguen una esperanza de 1,0 a través de un intervalo de señal amplio (3 log), los experimentos se consideraron un éxito.

Ejemplo 4: Perfilado de la metilación utilizando micromatrices de ADN

Se aplicó la técnica de perfilado de la metilación al mismo ADN genómico placentario femenino humano. El perfilado de la metilación pronostica la metilación (rojo) cuando las proporciones se desvían de la línea de 1,0 (azul oscuro) y los test t revelaron que más del 75% de los rasgos tuvieron desviaciones típicas (SD) aceptablemente pequeñas. Estas observaciones sugieren que la técnica está detectando reproduciblemente secuencias que están metiladas en el genoma humano.

En los experimentos de perfilado de la metilación, solamente 845 rasgos produjeron un escaso rendimiento, como se ha definido previamente. Los rasgos de escaso rendimiento representaron solamente el 1,2% de los EST totales en la matriz (148/12814). Como antes, la mayor parte de la intensidad de la señal de EST cayó casi 3 logs, indicando grandes diferencias en el número de copias.

Ejemplo 5: relevancia biológica de los alelos metilados humanos detectados por los métodos de la invención

Si el perfilado de la metilación funciona como está previsto, en ese caso las señales de proporción de colorante deben ser alteradas por el pre-tratamiento *in vitro* del ADN con metilasas, incrementando de ese modo artificialmente el contenido de 5mC. El perfilado de la metilación se llevó a cabo sobre ADN genómico placentario femenino que había sido sometido a un tiempo de cóctel con metilasa de tres puntos (0 min, 3 min, y 15 min @ 1U/ μ g). Los ADN se expusieron a M.SssI, que transfiere un grupo metilo a las citosinas de las secuencias CpG, y M.MspI, que transfiere de un modo similar grupos metilo a la citosina más externa de los sitios CCGG genómicos (<http://rebase.neb.com>). La determinación cuantitativa se sometió a ensayo seleccionando los momentos puntuales antes de completarse la reacción con la metilasa. El procedimiento se llevó a cabo como se ha descrito previamente, aunque sin intercambio de colorante. La muestra no tratada se marcó de rojo (Cy5) para cada matriz. Los resultados de los mismos 1.500 rasgos sobre las cuatro matrices diferentes demuestran claramente que los métodos para el perfilado de la metilación de la invención detectaron una presencia creciente de metilación en una serie de muestras de ADN metiladas *in vitro*. Además, muchos loci contienen metilación endógena. La Tabla 1 enumera los mejores 16 loci que se pronosticaba que contenían metilación en el genoma placentario femenino.

Tabla 1. Predicciones de rasgos que contienen 5mC a partir de placenta femenina

Locus	Descripción	GenBank	Proporción de 5mC (NT/5mC agotado-)*	Valor p de test t-test**	Marcas 5mC
10760	Locus gamma del receptor de células T	AI972955	1,895 (1,417 a 2,38)	7,33E-08	P
16191	Cromosoma 3 humano, agrupación 1 de pseudogenes del receptor olfatorio, secuencia completa, y pseudogén de quinasa de la cadena ligera de miosina (MLCK), secuencia parcial.	AF042089	2,123 (1,898 a 2,34)	7,57E-08	P
5567	Proteína P311	NM_004772	1,834 (1,44 a 2,248)	2,57E-06	P
3258	Proteína DKFZP566C134	AB040922	1,862 (1,439 a 2,295)	3,48E-06	P
2192	Pseudogen del receptor olfatorio, familia 7, subfamilia E, miembro 12	AK021566	1,595 (1,13 a 1,879)	1,59E-05	P
3873	Proteína 6 de unión a retinoblastoma	X85133	1,625 (1,222 a 2,197)	1,89E-05	P
12149	Proteína ribosomal L10a	AI133371	1,577 (1,213 a 2,065)	3,21E-05	p
2317	Gen de glicoproteína CD44 de la superficie de células humanas (CD44) gene, exón 6.	L05411	1,566 (1,251 a 1,903)	6,44E-05	P
15420	Receptor de activina A, tipo I	L02911	1,642 (1,33 a 1,938)	7,89E-05	P

Locus	Descripción	GenBank	Proporción de 5mC (NT/5mC agotado-)*	Valor p de test t-test**	Marcas 5mC
2601	glutámico-oxaloacético transaminasa 1, soluble (aspartato aminotransferasa 1)	NM 002079	1,758 (1,335 a 2,224)	8,16E-05	P
18152	quimerina (quimerina) 1	X51408	1,542 (1,29 a 1,862)	8,78E-05	P
17302	Antígeno identificado por el anticuerpo monoclonal Ki-67	X65550	2,055 (1,354 a 3,351)	1,29E-04	P
1240	asparraginil-ARNt sintetasa	NM_004539	1,489 (1,237 a 1,621)	1,45E-04	P
13099	Proteína en dedo de cinc 236	AF085244	1,735 (1,285 a 2,398)	1,63E-04	P
12161	EST de Incyte	0	1,533 (1,254 a 2,091)	1,87E-04	P
1845	Cromosoma 3 humano, agrupación 1 del pseudogen del receptor olfatorio, secuencia completa, y pseudogen de quinasa de la cadena ligera de miosina (MLCK), secuencia parcial.	AF042089	1,471 (1,203 a 1,678)	1,93E-04	P
<p>* La Tabla 1 enumera las proporciones de intensidad de 5mC media (n=8 (4 intercambios de colorante)) obtenidas de cada rasgo clasificadas por el valor P del test T, y después por la proporción global. La proporción refleja la intensidad media obtenida del canal de colorante no tratado dividida por el canal con colorante con la metilación agotada. También se indica el intervalo de la proporción.</p> <p>** Se determinó el valor p del test t utilizando la precisión del canal de la señal, (la DT de la intensidad de hibridación/rasgo en píxeles) en GeneSpring (v5.0). Como tal, la tabla muestra la identidad del rasgo y la proporción de metilación correspondiente de las 16 mediciones obtenidas más reproduciblemente.</p>					

5 Examinando la anotación de los rasgos de la Tabla 1, se hicieron evidentes loci interesantes y biológicamente relevantes. Por ejemplo, la familia del gen del receptor gamma de células T y el locus de la agrupación del pseudogen del receptor olfatorio fueron fácilmente identificados. Los loci del receptor de células T contienen una gran cantidad de metilación de ADN que es esencial para la función apropiada del receptor de células T (Dennis, K., et al., Genes Dev 15(22):2940-4 (2001); Geiman, T.M., et al., Biochim Biophys Acta 1526(2):211-20 (2001); Geiman, T.M. y K. Muegge, Proc Natl Acad Sci USA 97(9): págs. 4772-7 (2000)).

10 El segundo locus interesante es la agrupación del pseudogen del receptor olfatorio. Cabría esperar que las agrupaciones de pseudogenes contuvieran secuencias metiladas debido a que los pseudogenes son transcripcionalmente silenciosos. Por otra parte, en ratones, se ha demostrado que los alelos del receptor olfatorio expresados y las agrupaciones de genes asociadas son susceptibles de silenciamiento génico epigenético de una manera específica del parental de origen (Ishii, T., et al., Genes Cells 6(1):71-8 (2001)). Una tercera secuencia correspondía a una isla CpG que se había clonado previamente en virtud de su metilación en sangre de varón (Cross, S.H., et al., Nat Genet, 6(3): p. 236-44 (1994)). Diferentes rasgos de la misma agrupación de pseudogenes presentaron diferentes proporciones, indicando la naturaleza cuantitativa del análisis. A continuación se verificó que los loci individuales pronosticados contenían metilación mediante un método independiente.

20 Para someter a ensayo estadísticamente el funcionamiento del rasgo, se promediaron las proporciones de la intensidad a partir del perfilado de la metilación mediante el funcionamiento del rasgo durante un experimento "self-self" como indicador del ruido inherente en el sistema. El modelado de este modo no es óptimo puesto que la

5 dirección del funcionamiento de la sonda puede ser diferente en cada experimento. Los datos se volvieron a trazar. Se seleccionaron dos loci a partir de este conjunto de datos que se pronosticó que contenían 5mC, el receptor de las células T, proporción = 1,38; valores de $p = 0,019$, y clon de la isla CpG Z62622, proporción = 1,26; valor de $p = 0,017$. También se seleccionaron dos loci que se había pronosticado que no contenían metilación, asimismo clones de la isla CpG, Z65427, proporción = 1,03; valor de $p = 0,99$, y Z59762, proporción = 0,91; valor de $p = 0,87$. El acceso GenBank para cada locus se recuperó, y se seleccionaron los cebadores de PCR que permitirían la amplificación del exón más 3' de cada diana genómica.

10 Para la verificación independiente, se sometió ADN genómico placentario femenino a digestión parcial con McrBC y se amplificó utilizando cebadores de regiones metiladas y no metiladas del genoma. Las reacciones de PCR se establecieron con concentraciones similares de molde digerido con McrBC y se amplificaron durante un número similar de ciclos. Teóricamente, el número de productos derivados de las secuencias no metiladas debería permanecer constante, mientras el número de productos derivados de las secuencias metiladas debería disminuir en proporción a la cantidad de digestión con McrBC. Los resultados de los análisis duplicados en paralelo confirmó
15 claramente que los loci metilados pronosticados por el perfilado de la metilación están, de hecho, metilados endógenamente. De un modo similar, se confirmó independientemente que los loci que el perfilado de metilación había pronosticado que no estaban metilados endógenamente no estaban metilados.

20 En conclusión, en el genoma humano, los métodos de perfilado de la metilación de la invención mostraron una detección cuantitativa de metilado *in vitro* y una detección cualitativa de metilación de ADN endógeno del genoma.

REIVINDICACIONES

1. Un método para determinar el perfil de metilación de ADN de una célula, tejido u organismo, comprendiendo el método las etapas de:
- 5 a. proporcionar una población de fragmentos de ADN escindidos al azar o sometidos a cizalla a partir de una célula, tejido, u organismo, en donde el DNA comprende fragmentos metilados y no metilados;
- b. agotar el ADN metilado o no metilado de la población; y
- c. cuantificar la cantidad de al menos una secuencia de ADN metilado en la población agotada o del ADN no metilado en la población agotada con respecto a la cantidad de la al menos una secuencia en el ADN genómico total.
- 10 2. El método de la reivindicación 1, en donde el tamaño medio de los fragmentos de ADN está entre 0,1-10 kb.
3. El método de la reivindicación 1 o 2, en donde la etapa de agotamiento comprende:
- 15 fragmentar la población con una enzima de restricción sensible a la metilación o una enzima de restricción dependiente de la metilación para producir ADN digerido y ADN no digerido; y separar el ADN digerido del ADN no digerido.
4. Un método de acuerdo con la reivindicación 1 o 2, en donde:
- 20 la etapa a comprende proporcionar una población de fragmentos de ADN escindidos al azar o sometidos a cizalla de una célula, tejido, u organismo, en donde el ADN comprende una primera porción y una segunda porción y cada porción comprende fragmentos metilados y no metilados;
- la etapa b comprende agotar el ADN metilado o no metilado de la segunda porción; y
- la etapa c comprende cuantificar la cantidad relativa de al menos una secuencia de al menos dos de las siguientes:
- 25 la primera porción,
ADN metilado de la segunda porción, y
ADN no metilados de la segunda porción.
- 30 5. El método de la reivindicación 4, en donde la etapa de cuantificación comprende la amplificación cuantitativa.
6. El método de la reivindicación 4, en donde la etapa de agotamiento comprende:
- fragmentar la segunda porción con una enzima de restricción sensible a la metilación o una enzima de restricción dependiente de la metilación para producir ADN digerido y ADN no digerido; y
- 35 separar el ADN digerido del ADN no digerido,
y en donde dicho método comprende adicionalmente:
marcar la primera porción con una marca;
marcar el ADN digerido o no digerido de la segunda porción con una marca;
- 40 hibridar el ADN marcado de la primera y segunda porciones a un ácido nucleico; y
determinar la metilación relativa de un ácido nucleico detectando la primera y la segunda marcas que hibridan con el ácido nucleico, hibridar con el ácido nucleico, determinando de ese modo el perfil de metilación de al menos una secuencia de ácido nucleico de la célula, tejido, u organismo.
7. El método de una cualquiera de las reivindicaciones 1 a 6, en donde el ADN metilado se agota a partir del ADN escindido al azar o sometido a cizalla.
- 45 8. El método de una cualquiera de las reivindicaciones 1 a 6, en donde el ADN no metilado se agota a partir del ADN escindido al azar o sometido a cizalla.
9. El método de la reivindicación 4, en donde la etapa de cuantificación comprende hibridar el ADN reducido del ADN metilado o no metilado con un ácido nucleico conectado a un soporte sólido.
10. El método de una cualquier de las reivindicaciones anteriores, en donde el organismo se selecciona entre una planta, un animal, un hongo, y un procarionta.
- 55 11. El método de una cualquiera de las reivindicaciones 1 a 10, en donde el método comprende adicionalmente comparar el perfil de metilación de un ácido nucleico con un perfil de transcripción del ácido nucleico, determinando de ese modo la relación entre el perfil de metilación y el perfil de transcripción del ácido nucleico.
- 60 12. El método de una cualquiera de las reivindicaciones 1 a 10, en donde el método comprende adicionalmente comparar el perfil de metilación de un ácido nucleico con un perfil del estado de empaquetamiento de la cromatina del ácido nucleico, determinando de ese modo la relación entre el perfil de metilación y el perfil del estado de empaquetamiento de la cromatina del ácido nucleico.

5 **13.** El método de una cualquiera de las reivindicaciones 1 a 10, que comprende adicionalmente comparar el perfil de metilación del ácido nucleico con el número de copias del ácido nucleico, determinando de ese modo la contribución a un fenotipo de la combinación de la metilación del ácido nucleico y el número de copias del ácido nucleico.

10 **14.** El método de la reivindicación 4, en donde:

se compara el perfil de metilación de al menos una primera y una segunda muestras de ADN,

la etapa a comprende proporcionar una población de fragmentos de ADN escindidos al azar o sometidos a cizalla a partir de las muestras de ADN;

15 la etapa b comprende:

(i) agotar el ADN metilado o no metilado del ADN escindido al azar o sometido a cizalla de la primera muestra de ADN, y

(ii) opcionalmente agotar el ADN metilado o no metilado del ADN escindido o sometido a cizalla de la segunda muestra de ADN; y

20 el método comprende adicionalmente:

c. comparar la cantidad de al menos una secuencia de la primera muestra de ADN agotada con la cantidad de la secuencia en la segunda muestra de ADN escindido o sometido a cizalla o la segunda muestra de ADN agotada.