

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 779 798**

51 Int. Cl.:

C12Q 1/6806 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.08.2016 PCT/EP2016/069121**

87 Fecha y número de publicación internacional: **16.02.2017 WO17025594**

96 Fecha de presentación y número de la solicitud europea: **11.08.2016 E 16760659 (9)**

97 Fecha y número de publicación de la concesión europea: **30.10.2019 EP 3334841**

54 Título: **Métodos para estudiar ácidos nucleicos**

30 Prioridad:

12.08.2015 EP 15180705

14.10.2015 EP 15189788

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

19.08.2020

73 Titular/es:

**CEMM - FORSCHUNGSZENTRUM FÜR
MOLEKULARE MEDIZIN GMBH (100.0%)
Lazarettgasse 14 / AKH BT 25.3
1090 Wien , AT**

72 Inventor/es:

**BOCK, CHRISTOPH y
SCHMIDL, CHRISTIAN**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 779 798 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 estudiar ácidos nucleicos

La presente invención proporciona un método novedoso para preparar una biblioteca de secuenciación y estudiar las interacciones moleculares que implican a un ácido nucleico. En particular, la invención se refiere a un método para preparar una biblioteca de secuenciación, comprendiendo el método añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada; aislar el ácido nucleico de la cromatina; y obtener una biblioteca de secuenciación. Además, la presente invención se refiere a un método para el mapeo de interacciones moleculares que implican a un ácido nucleico, comprendiendo el método añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada; aislar el ácido nucleico de la cromatina; amplificar el ácido nucleico; secuenciar el ácido nucleico amplificado; e identificar interacciones moleculares.

El conocimiento de las interacciones entre los ácidos nucleicos y otras sustancias químicas y/o biomoléculas es de gran interés para la investigación y la medicina. Un método bien conocido para estudiar las interacciones proteína-ácido nucleico es la inmunoprecipitación de cromatina (ChIP, por sus siglas en inglés), opcionalmente seguida de secuenciación masiva en paralelo (ChIP-seq). Un método para estudiar las interacciones de moléculas pequeñas con ácidos nucleicos y/o proteínas, por ejemplo en cromatina, es Chem-Seq, que se describe más adelante.

El método ChIP permite estudiar las interacciones ADN-proteína en todo el genoma. Contribuyó sustancialmente a la comprensión por parte de los inventores de la organización de la cromatina, la modificación de histonas, así como los patrones de unión de factor de transcripción (usando X-ChIP) y su influencia sobre la regulación génica en la salud y la enfermedad; véase por ejemplo Nature (2012) 489, págs. 57-74 o Ernst *et al.* (2011) Nature 473, págs. 43-49. Sin embargo, ChIP sigue siendo un protocolo relativamente tedioso, especialmente cuando se aplica a muestras de baja entrada (véase por ejemplo, Greenleaf, W.J. (2014) Methods). Para preparar ácidos nucleicos a partir de ChIP para la secuenciación de nueva generación (= para preparar una biblioteca), el enfoque clásico comprende varias etapas laboriosas: (i) reparación de extremos de las secuencias de ADN purificadas para generar fragmentos de ADN bicatenarios de extremos romos con un extremo 3' fosforilado; (ii) adición de una proyección A; (iii) ligamiento de adaptadores que tienen una proyección T complementaria a los fragmentos de ChIP-ADN bicatenarios y con reparación de extremos con proyección A. Los adaptadores permiten la amplificación de los fragmentos de ADN, lo que garantiza una cantidad suficiente de fragmentos para el control de calidad y la secuenciación posterior, y también prepara los fragmentos para el procedimiento de secuenciación mediante la introducción de extremos de celdas de flujo para la generación de agrupaciones y secuencias de códigos de barras para experimentos de secuenciación múltiplex. El método clásico tiene varias limitaciones: (i) generalmente se necesitan 5-10 ng de material de entrada para generar bibliotecas que no pueden recuperarse de ChIP en pequeñas cantidades de células. Así, la cantidad recomendada de células para un experimento de ChIP-seq está en el intervalo de 10⁶ células. (ii) El procedimiento de la biblioteca se basa en varias reacciones enzimáticas y purificaciones de ADN, que hacen que la generación de la biblioteca sea un procedimiento relativamente laborioso. Las reacciones enzimáticas imperfectas así como las purificaciones de ADN también reducen la cantidad de fragmentos de biblioteca recuperados, lo que explica los altos requisitos de entrada. (iii) Los adaptadores pueden autoligarse y deben excluirse de la amplificación y secuenciación. Así, es necesaria una selección por tamaño para seleccionar frente al exceso de adaptadores y dímeros de adaptador. Como alternativa al ligamiento de adaptador, se desarrollaron otros protocolos que generan bibliotecas de ChIP-seq mediante la transcripción inversa de fragmentos de ADN; vea el kit de ChIP-seq de Clontech, Mountain View, CA. Además, se desarrollaron protocolos de ChIP-seq para bajas cantidades de material de partida tales como iChIP (Lara-Astiaso *et al.* (2014) Science 345, págs. 943-9), linDA (Shankaranarayanan *et al.* (2011) Nature Methods 8, págs. 565-7) y ChIP asistido por transportador (Zwart *et al.* (2013) BMC Genomics 14, 232 o Jakobsen *et al.* (2015) BMC Genomics 16, 46), pero estos protocolos requieren reactivos adicionales, tiempo de preparación o requieren la agrupación de muchas muestras, lo que los hace costosos, que requieran mucho tiempo y/o sean inflexibles (véase la figura 1 para obtener una descripción general de los métodos de la técnica anterior y los inconvenientes conocidos).

Otro método para estudiar las interacciones entre moléculas pequeñas y sus dianas de proteína/ácido nucleico en la cromatina es Chem-seq. El método emplea la captura por afinidad química junto con una secuencia de ADN en paralelo de manera masiva para identificar sitios genómicos en los que las moléculas pequeñas interactúan con sus proteínas o ADN diana. Lo describieron por primera vez Anders *et al.* en Nature Biotechnology (2013), 32 (1), págs. 92-6.

Recientemente se describió un método adicional para la preparación de bibliotecas de ácidos nucleicos. El método hace uso del desarrollo de una transposasa Tn5 hiperactiva para simultáneamente fragmentar y marcar con etiqueta (tagmentación, "tagmentation" como contracción en inglés) adaptadores de ADN (véase Adey *et al.* (2010) Genome Biol 11, R119). Usa una transposasa, que está precargada con adaptadores compatibles con la secuenciación. La transposasa integra la carga de su adaptador en el ADN mientras lo fragmenta. Sólo se necesitan pequeñas cantidades de transposasa para generar bibliotecas de ADN genómico (Adey *et al.* (2010) Genome Biol 11, R119), ADN convertido con bisulfito para el análisis de metilación del ADN (Wang, Q. *et al.* (2013) Nature Protocols 8, 2022-2032), ADNc de sec. de ARN u otros ácidos nucleicos en bibliotecas listas para secuenciación (Picelli *et al.* (2014)

Genome Res vol. 24 (12) págs. 2033-2040). En la publicación Pirelli citada anteriormente, se preparan bibliotecas a partir de muestras de ADNc aisladas. Sin embargo, las reacciones de tagmentación de ácidos nucleicos purificados son extremadamente sensibles a las razones variables de transposasa con respecto a ácidos nucleicos ya que, entre otros parámetros, la cantidad de transposasa en la reacción de tagmentación determina la distribución de tamaño final de los fragmentos de ácido nucleico. Por tanto, los métodos resultantes requieren una experimentación que requiere mucho trabajo y resulta muy costosa para determinar las razones apropiadas de transposasa y ácido nucleico para lograr la distribución de tamaño deseable de los ácidos nucleicos para la secuenciación de nueva generación u otras aplicaciones posteriores. En algunos casos, la determinación de la distribución y abundancia de fragmentos de ácido nucleico ni siquiera es factible, por lo que es imposible encontrar las razones correctas de transposasa con respecto a ácido nucleico para preparar bibliotecas de secuenciación según, entre otros, aplicaciones de la técnica anterior revisadas en Furey *et al.* (2012) Nature Reviews Genetics 13 (12), págs. 840-852. Además, la adición de transposasa a núcleos celulares recupera regiones de cromatina abierta y proporciona información sobre la posición del nucleosoma, así como las huellas de factor de transcripción en las regiones reguladoras del genoma (Buenrostro JD. *et al.* (2013) Nat Meth, vol. 10 (12) págs. 1213-1218).

Sin embargo, no se describió sistemáticamente que la transposasa fuese adecuada para su uso en la generación de bibliotecas de secuenciación a partir de ácidos nucleicos posteriores a ChIP o Chem-Seq. Más bien, se discuten las posibles desventajas de este enfoque. Estas desventajas son el resultado de realizar la tagmentación en el ADN purificado posterior a ChIP, tal como se describe en los documentos WO 2013/078470 o WO 2014/205296. Por consiguiente, los principales inconvenientes de la combinación de ChIP y tagmentación son: (1) El ADN de ChIPed, que ya está sonificado en pequeños fragmentos (200-700 pb), se fragmenta adicionalmente en su totalidad. Por tanto, la tagmentación puede dar como resultado fragmentos de biblioteca muy pequeños hasta un tamaño mínimo de tan sólo ~40 pb (Adey *et al.* (2010) Genome Biol 11, R119) que pueden ser difíciles de secuenciar, ya que se recomiendan fragmentos de 150 pb a 200 pb como longitud mínima para la secuenciación en Illumina; (2) una mayor fragmentación mediante tagmentación genera probablemente múltiples lecturas de secuenciación por fragmento precipitado originalmente, lo que dificulta potencialmente el análisis posterior. Como ejemplo, un fragmento de ADN inmunoprecipitado de 600 pb puede producir el doble de la cantidad de fragmentos de biblioteca en comparación con un fragmento de 300 pb, aumentando de ese modo artificialmente la cantidad relativa de lecturas en la región de 600 pb. Esto puede ser problemático para la correcta asignación de picos cuando se analizan datos de ChIP-seq; (3) el enfoque para usar ADN de ChIP purificado para generar bibliotecas de secuenciación mediante tagmentación es inconveniente ya que se necesitan la determinación de tamaño y la cuantificación de ADN correctas para establecer la reacción de tagmentación. Como la sonicación puede variar entre las muestras y las cantidades de ADN, que varían generalmente en un orden de magnitud dependiendo del anticuerpo usado para la IP, estos parámetros deberían determinarse para cada muestra de ChIP antes de la tagmentación; (4) el enfoque para usar el ADN de ChIP purificado para generar bibliotecas de secuenciación mediante tagmentación tampoco suele ser posible ya que las cantidades de ADN de ChIP pueden ser demasiado bajas para una cuantificación y determinación de tamaño robustas, que son ambos parámetros críticos para establecer reacciones de tagmentación robustas; y/o (5) la tagmentación de ADN de ChIP purificado no conserva la información potencial de la estructura de cromatina local en las regiones diana inmunoprecipitadas. Un método adicional descrito en el documento WO 2013/078470, denominado TAM-ChIP, usa la transposasa Tn5 conjugada con anticuerpos para ChIP. Es decir, el factor limitante de TAM-ChIP, como también se describe en el documento WO 2014/190214, es la limitación al uso de conjugados anticuerpo-oligonucleótido que deben producirse antes de la aplicación. Esto impide el uso *ad hoc* de anticuerpos disponibles comercialmente que se usan principalmente para estudiar las interacciones proteína-ADN con inmunoprecipitación de cromatina. Aunque se usaron conjugados anticuerpo-oligonucleótido secundarios en TAM-ChIP para superar las limitaciones anteriores, deben usarse dos conjuntos de anticuerpos, lo que aumenta la complejidad del ensayo y al mismo tiempo aumenta los costes debido al uso de un anticuerpo secundario que no se usa normalmente en aplicaciones tales como ChIP. Por consiguiente, TAM-ChIP requiere optimizaciones extensas de complejos anticuerpo-oligonucleótido-transposasa para introducir las razones de cromatina, tal como se describe en el documento WO 2013/078470. Como la cantidad de complejos anticuerpo-oligonucleótido-transposasa reclutados en su sitio de reconocimiento determina el tamaño de biblioteca final, y debido a que el número de sitios de reconocimiento puede variar de unos cientos a cien mil dependiendo de los antígenos diana, las razones de complejos anticuerpo-oligonucleótido-transposasa con respecto a cromatina de entrada deben evaluarse para cada conjugado específico anticuerpo-transposasa. Por tanto, se requiere un método más robusto insensible con respecto a las razones de transposasa con respecto a cromatina de entrada. Una desventaja adicional de TAM-ChIP es la eficiencia de la reacción de tagmentación, que está en el intervalo del 0,5%-5% debido al hecho de que sólo se puede reclutar un transposoma para cada antígeno diana. Por consiguiente, se observa una baja conversión de ácidos nucleicos en fragmentos de biblioteca de secuenciación y probablemente un aumento de los requisitos de entrada. Por tanto, es deseable un método que pueda usar transposasa en exceso para marcar con etiqueta ácidos nucleicos en varios sitios cercanos a los sitios diana para introducir adaptadores de secuenciación junto a la mayoría de los fragmentos de ácido nucleico. Finalmente, TAM-ChIP requiere grandes cantidades de cromatina de entrada. En particular, la preparación de biblioteca de secuenciación satisfactoria que produce resultados de secuenciación comparables al ChIP-seq convencional sólo puede demostrarse usando 10 µg de cromatina de entrada, que corresponde a ~1,5 millones de células. Por tanto, no es factible investigar tipos de células de baja abundancia y poco comunes incluyendo células leucémicas o células tumorales de pacientes primarias, células de pacientes primarias de biopsias, pequeñas poblaciones de tipos de células hematopoyéticas, poblaciones de células de baja abundancia obtenidas de modelos de desarrollo o investigación en embriología. Por tanto, es deseable un método

que permita de manera robusta el uso de cantidades de baja entradas, es decir, menos números de células como entrada para estudiar las interacciones moleculares.

A la luz de lo anterior, existe la necesidad de métodos mejorados para preparar bibliotecas de secuenciación y/o mapear interacciones moleculares que impliquen a ácidos nucleicos, que sea independientes de las cantidades de entrada y que sea más rápidos, fáciles y/o más baratos.

Por tanto, la invención se refiere a los siguientes puntos:

1. Método para preparar una biblioteca de secuenciación, comprendiendo el método:

- (a) añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico;
- (b) aislar la cromatina unida por dicho agente;
- (c) añadir transposasa a la cromatina aislada de la etapa (b);
- (d) aislar el ácido nucleico de la cromatina; y
- (e) obtener una biblioteca de secuenciación.

2. Método para preparar una biblioteca de secuenciación, comprendiendo el método:

- (a) añadir un anticuerpo que se une a cromatina a una muestra que comprende un ácido nucleico;
- (b) aislar la cromatina unida por dicho anticuerpo;
- (c) añadir transposasa a la cromatina unida y aislada de la etapa (b);
- (d) aislar el ácido nucleico de la cromatina; y
- (e) obtener una biblioteca de secuenciación.

3. Método para el mapeo de interacciones moleculares que implican a un ácido nucleico, comprendiendo el método:

- (a) añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico;
- (b) aislar la cromatina unida por dicho agente;
- (c) añadir transposasa a la cromatina aislada de la etapa (b);
- (d) aislar el ácido nucleico de la cromatina;
- (e) amplificar el ácido nucleico;
- (f) secuenciar el ácido nucleico amplificado; y
- (g) identificar interacciones moleculares.

4. Método para el mapeo de interacciones moleculares que implican a un ácido nucleico, comprendiendo el método:

- (a) añadir un anticuerpo que se une a cromatina a una muestra que comprende un ácido nucleico;
- (b) aislar la cromatina unida por dicho anticuerpo;
- (c) añadir transposasa a la cromatina unida y aislada de la etapa (b);
- (d) aislar el ácido nucleico de la cromatina;
- (e) amplificar el ácido nucleico;
- (f) secuenciar el ácido nucleico amplificado; y
- (g) identificar interacciones moleculares.

5. El método de los puntos 1 a 4, en el que la muestra que comprende un ácido nucleico se ha preparado mediante

- (i) cultivar y recoger células;

- (ii) fijar las células;
 - (iii) lisar las células y obtener de ese modo una primera muestra que comprende un ácido nucleico; y
 - (iv) sonicar la primera muestra y obtener de ese modo una segunda muestra que comprende un ácido nucleico, en el que dicha segunda muestra va a usarse en el método de los puntos 1 ó 2.
- 5
6. El método de los puntos 1 a 5, en el que el método comprende además una etapa de revertir las reticulaciones introducidas durante la fijación de las células.
7. El método de uno cualquiera de los puntos 1 a 6, en el que el ácido nucleico es ADN.
- 10
8. El método de uno cualquiera de los puntos 5 a 7, en el que las células comprenden complejos ácido nucleico-proteína.
9. El método del punto 8, en el que las células son células humanas, células animales, células bacterianas, células de levadura, células de arqueas, células vegetales o virus.
10. El método del punto 9, en el que las células humanas o animales son células enfermas o células no enfermas o células derivadas de tejido enfermo o no enfermo.
- 15
11. El método de los puntos 9 o 10, en el que las células humanas o animales son células cancerosas, células inmunitarias, células sanguíneas o células madre.
12. El método del punto 11, en el que el cáncer es un cáncer sólido o cáncer de la sangre.
13. El método del punto 12, en el que el cáncer de la sangre es leucemia.
14. El método del punto 12, en el que el cáncer sólido es un tumor.
- 20
15. El método de los puntos 9, 10 u 11, en el que la célula es una célula embrionaria.
16. El método del punto 5, en el que la etapa (ii) comprende añadir una sustancia química y/o medios físicos.
17. El método del punto 16, en el que la sustancia química es formaldehído o paraformaldehído.
18. El método del punto 16, en el que los medios físicos comprenden luz UV o láser.
- 25
19. El método del punto 5, en el que la etapa (iv) comprende sonicar hasta que la mayoría de los fragmentos de ácido nucleico tienen 20-5000 pares de bases, preferiblemente 200-300 de longitud.
20. El método de los puntos 1 ó 2, en el que el agente que se une a cromatina es un anticuerpo o una sustancia química.
21. El método del punto 20, en el que el anticuerpo se une específicamente a histonas, factores de transcripción o proteínas que se unen a histonas y/o factores de transcripción.
- 30
22. El método del punto 21, en el que las proteínas que se unen a histonas y/o factores de transcripción son proteínas de remodelación de ácido nucleico o enzimas modificadoras de cromatina.
23. El método de los puntos 21 ó 22, en el que la histona es H3.3, H2A.Z, CENP-A, H3.2, H3.3A, H3.3B, H4 o H3.1.
- 35
24. El método de los puntos 21, 22 ó 23, en el que la histona es una histona modificada, en el que la modificación es metilación, acetilación, propionilación, butirilación, crotonilación, 2-hidroxiisobutirilación, malonilación, succinilación y/o ribosilación.
25. El método del punto 24, en el que la histona modificada es H3K4me1/2/3, H2BK5me1, H3K27me1/2/3, H3K9me1/2/3, H4K20me1, H3K79me1, H3K36me3, H2AK5ac, H2AK9ac, H2BK5ac, H2BK12ac, H2BK20ac, H2BK120ac, H3K4ac, H3K9ac, H3K14ac, H3K18ac, H3K23ac, H3K27ac, H3K36ac, H4K5ac, H4K8ac, H4K12ac, H4K16ac, H4K91ac, H2Aub o H2Bub.
- 40
26. El método del punto 20, en el que la sustancia química es un fármaco.
27. El método de los puntos 20 ó 26, en el que la sustancia química está biotilada.
28. El método de los puntos 1 a 4, en el que la transposasa comprende etiquetas de secuencia de ADN aleatorias o etiquetas de secuencia de ADN definidas.
- 45
29. El método del punto 28, en el que la transposasa es una transposasa Tn5.

Por consiguiente, la invención proporciona un método para preparar una biblioteca de secuenciación, comprendiendo el método añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada; aislar el ácido nucleico de la cromatina; y obtener una biblioteca de secuenciación.

- 5 Otras realizaciones se describen en el presente documento y se ejemplifican en la parte científica. Las figuras adjuntas proporcionan ilustraciones de la presente invención. Aunque los datos experimentales en los ejemplos y tal como se ilustra en las figuras adjuntas no se consideran limitativos. La información técnica comprendida en los mismos forma parte de esta invención.

- 10 Por tanto, la invención proporciona un método para preparar una biblioteca de secuenciación y un método para el mapeo de interacciones moleculares que implican a un ácido nucleico, en particular ADN. Tal como resulta evidente a partir de los ejemplos adjuntos, los métodos proporcionados en el presente documento comprenden en particular la preparación de una biblioteca de secuenciación o el mapeo de interacciones moleculares que implican a un ácido nucleico, en particular ADN, combinando las etapas de añadir un agente que se une a cromatina a una muestra, aislar la cromatina unida y añadir una transposasa a la cromatina aislada en un orden específico. Por consiguiente,
- 15 la invención proporciona un método para preparar una biblioteca de secuenciación, comprendiendo el método añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada; aislar el ácido nucleico de la cromatina; y obtener la biblioteca de secuenciación. La adición de transposasa debe realizarse después de aislar la cromatina unida. Se prefiere que el ácido nucleico sea ADN. Los métodos de la invención para el mapeo de interacciones moleculares que implican a un ácido nucleico comprenden añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada; aislar el ácido nucleico de la cromatina; amplificar el ácido nucleico; secuenciar el ácido nucleico amplificado; e identificar interacciones moleculares. La muestra que comprende un ácido nucleico puede ser una muestra de células primarias o una muestra obtenida mediante un método de cultivo. Cuando la muestra se obtiene mediante un método de cultivo, se prefiere que los métodos comprendan además cultivar y recoger células; fijar las células; lisar las células y obtener de ese modo una primera muestra que comprende un ácido nucleico; y sonicar la primera muestra y obtener de ese modo una segunda muestra que comprende un ácido nucleico. Se prefiere que dicha segunda muestra se use en los métodos de la invención. Además, se prefiere que dicho ácido nucleico sea ADN, en particular ADN bicatenario. Cuando la muestra que comprende un ácido nucleico es una muestra de células primarias, los métodos de la invención preferiblemente comprenden además fijar las células; lisar las células y obtener de ese modo una primera muestra que comprende un ácido nucleico; y sonicar la primera muestra y obtener de ese modo una segunda muestra que comprende un ácido nucleico.
- 20
- 25
- 30

- Por consiguiente, se descubrió de manera sorprendente e inesperada que la tagmentación directamente en la cromatina unida por un agente específico para la cromatina y aislado de la cromatina no unida en una reacción robusta de una única etapa conduce a un protocolo de propósito general muy robusto que es más rápido, más barato, más fácil, más robusto y más compatible con muestras de baja entrada que los métodos comprendidos en la técnica anterior. Por tanto, se prefiere que la adición de transposasa se realice después de aislar la cromatina unida por el agente que se une a cromatina, en particular el anticuerpo o la sustancia química. Por consiguiente, la presente invención se refiere preferiblemente a un método que comprende añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada después de aislar la cromatina unida por el agente que se une a cromatina; aislar el ácido nucleico de la cromatina; y obtener la biblioteca de secuenciación. Se prefiere que el ácido nucleico sea ADN. Los métodos de la invención para el mapeo de interacciones moleculares que implican a un ácido nucleico comprenden preferiblemente añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada después de aislar la cromatina unida por el agente que se une a cromatina; aislar el ácido nucleico de la cromatina; amplificar el ácido nucleico; secuenciar el ácido nucleico amplificado; e identificar interacciones moleculares.
- 35
- 40
- 45

- Tal como se muestra en la figura 2, el protocolo de ChIP-seq convencional comprende las etapas de fijación de células, lisis celular, sonicación de cromatina e inmunoprecipitación con un anticuerpo específico unido a perlas. La reticulación inversa va seguida por la purificación del ADN de ChIP, que luego se somete a la preparación de biblioteca en un procedimiento de múltiples etapas que comprende reparación de extremos, purificación, unión de colas A, ligamiento de adaptador y selección por tamaño. En una primera mejora del protocolo de ChIP-seq convencional, se halló un método denominado ChIP-tagmentación. En la ChIP-tagmentación, el ADN de ChIP purificado se usa para la preparación de biblioteca basada en tagmentación (véase la figura 11). El método es sensible a concentraciones variables de ADN, porque la tagmentación de ADN purificado es sensible a la razón de enzima de tagmentación con respecto a ADN, y las concentraciones de ADN pueden ser altamente variables y demasiado bajas como para cuantificarse en muchas aplicaciones de ChIP-seq.
- 50
- 55

- A este respecto, se halló de manera sorprendente e inesperada que, cuando los adaptadores de secuenciación se introducen en una única etapa posterior al aislamiento de la cromatina unida al agente usando tagmentación, por ejemplo, usando la transposasa Tn5 cargada en adaptador, el método resultante es insensible a concentraciones de ADN variables.
- 60

Por tanto, la robustez mejorada de los métodos de la invención se logra realizando la tagmentación directamente en la cromatina unida al agente, que se aísla de la cromatina no unida, donde las proteínas protegen al ácido nucleico frente a una tagmentación excesiva. El protocolo resultante de los métodos de la invención (por ejemplo, tal como se muestra en la figura 2, derecha y 4a) demostró ser muy robusto en una diferencia de 25 veces en las concentraciones de enzima de tagmentación, en términos de distribución de tamaño de las bibliotecas (figura 3), distribución de tamaño de las lecturas de secuenciación (figura 4b), rendimiento de mapeo (figura 4c), calidad de seguimiento (figura 4d) y concordancia (figura 4e). Los métodos de la invención también se mejoran porque no dan lugar a secuenciación de dímeros de adaptador y no requieren etapas de purificación de ácido nucleico más allá de la limpieza convencional después del protocolo de ChIP convencional.

Las ventajas anteriores de los métodos proporcionados en el presente documento los hacen superiores a los métodos conocidos en la técnica anterior. En particular, los métodos proporcionados en el presente documento son más flexibles, más baratos, más robustos, requieren menores cantidades de entrada de muestra y permiten obtener información adicional con respecto a los métodos conocidos en la técnica anterior, en particular TAM-ChIP tal como se proporciona en el documento WO 2014/190214. Específicamente, en TAM-ChIP se requieren conjugados oligonucleótido-anticuerpo o conjugados anticuerpo-oligonucleótido-transposasa específicos para preparaciones de biblioteca de secuenciación mediada por transposasa. En la presente invención, pueden usarse anticuerpos ChIP-seq disponibles comercialmente *ad hoc* sin reacciones de conjugación laboriosas y costosas. Además, el protocolo TAM-ChIP tal como se describe en el documento WO 2014/190214 requiere una optimización extensa de las razones de los conjugados oligonucleótido-transposasa-antibiótico con respecto a la cromatina de entrada, mientras que los métodos de la presente invención son robustos para razones de transposasa-cromatina variables.

La adición de transposasa de manera posterior al aislamiento de la cromatina de interés, como en los métodos proporcionados en el presente documento (por ejemplo, por un anticuerpo específico) tiene beneficios adicionales con respecto a los métodos conocidos en la técnica. En particular, TAM-ChIP requiere la determinación de razones optimizadas de conjugados anticuerpo-oligonucleótido para dirigirlos a sus sitios de reconocimiento en la cromatina. Así, sólo los ácidos nucleicos en estrecha proximidad de los sitios de reconocimiento pueden marcarse con etiqueta, lo que da como resultado una frecuencia de marcaje con etiqueta relativamente bajo (el 0,5-5%) cuando se usa una transposasa. En cambio, los métodos de la presente invención permiten la tagmentación independientemente del agente usado para aislar la cromatina de interés. Por tanto, en cuanto a la robustez de los métodos de la invención con respecto a las razones de transposasa-cromatina, puede usarse una cantidad en exceso de transposasa en los métodos de la invención para maximizar la eficiencia de generación de biblioteca de secuenciación. Este logro reduce sustancialmente los requisitos de entrada en los métodos de la presente invención en comparación con los métodos conocidos en la técnica.

Además, la adición de la transposasa de manera posterior al aislamiento de la cromatina de interés, permite inesperadamente una preparación eficiente de biblioteca de secuenciación usando bajas cantidades de enzima transposasa. Esto se debe a la presencia reducida de cromatina molde inespecífica para la reacción de tagmentación debido al aislamiento de la cromatina de interés, mientras que el resto se desecha. La reducción de las cantidades de transposasa requeridas es una ventaja de coste significativa de los métodos de la invención con respecto a los métodos conocidos en la técnica, en particular los métodos descritos en los documentos WO 2014/205296 y WO 2014/190214.

Con respecto a métodos alternativos conocidos en la técnica, por ejemplo, los descritos por Picelli *et al.* (tal como se citó anteriormente), que usa ADN purificado como molde para la preparación de biblioteca de secuenciación, el uso de cromatina como molde, como en los métodos de la presente invención, conserva información estructural de alta resolución del contexto de cromatina local; véase por ejemplo la figura 9.

Por tanto, los métodos de la presente invención, por primera vez, permiten la construcción y amplificación de bibliotecas de secuenciación a partir de cromatina para estudiar interacciones moleculares sin purificación o extracción previa de ácidos nucleicos, en particular como en el método ultrarrápido proporcionado en el presente documento. Además, los métodos de la invención permiten por primera vez de manera factible el mapeo de accesibilidad a cromatina a gran escala en enfermedades, en particular cáncer, cohortes e investigación clínica al proporcionar un flujo de trabajo racionalizado y de baja entrada para el mapeo en todo el genoma de marcas de histona y factores de transcripción. Dado que el análisis de obtención de perfil de cromatina proporcionado en el presente documento es suficientemente rápido y directo para su uso en un laboratorio de secuenciación clínica, la desregulación de la cromatina ahora es manejable como fuente de biomarcadores, por ejemplo, para la terapia del cáncer estratificado; véase también Rendeiro *et al.* (2016) Nature Comm. 7, número de artículo 11938.

Los métodos de la invención se validaron para cinco marcas de histona a modo de ejemplo (H3K4me3, H3K27ac, H3K4me1, H3K36me3 y H3K27me3) y cuatro factores de transcripción (PU.1, CTCF, GATA1 y REST). En todos los casos, los métodos de la invención mostraron una validez de datos similar en comparación con ChIP-Seq convencional (figura 5f). Sin embargo, los métodos de la invención permitieron la reducción significativa de la entrada de células. En particular, se obtuvieron datos de alta calidad para H3K4me3 y H3K27me3, así como para GATA1 y CTCF a partir de 10k y 100k células, respectivamente, sin ninguna amplificación previa (figura 4f). A este respecto, para el protocolo de ChIP-Seq convencional, se recomiendan al menos 500k/10M de células (modificaciones de histonas/factores de transcripción, respectivamente) como entrada, mientras que los datos que

usan los métodos de la invención se obtuvieron usando 10k o 100k células, respectivamente, mientras que el uso de 1k células es factible. A este respecto, 10k células producen aproximadamente 25 ng de cromatina según lo determinado mediante el fluorímetro Qubit® después de la purificación de ADN a partir de 10k células usando métodos convencionales conocidos en la técnica. Por tanto, los métodos novedosos e inventivos de la invención permiten una reducción de la entrada de células en al menos un factor de 50-100 en comparación con las recomendaciones de los requisitos de entrada para ChIP-seq clásico sobre modificaciones de histonas o factores de transcripción, respectivamente. Por tanto, en contraste con ChIP-seq convencional y preparaciones de biblioteca de secuenciación mediada por transposasa conocidas que usan ácidos nucleicos purificados o ADNc como molde para la preparación de biblioteca de secuenciación, los métodos proporcionados en el presente documento pueden usar cromatina como molde para generar bibliotecas de secuenciación.

Se observaron excelentes correlaciones y superposición de picos entre el protocolo de ChIP convencional y los métodos de la invención, y también entre réplicas biológicas y muestras de baja entrada preparadas con cualquiera de los métodos (figuras 4g, 4h, 5, 6 y 7). En total, se secuenciaron 52 bibliotecas usando los métodos de la invención, 24 bibliotecas con ChIP-seq convencional (figura 2, izquierda), y 9 bibliotecas con ChIP-tagmentación (figura 2, centro) y tasas de alineación observadas superiores al 95% y tasas de lecturas únicas de alrededor del 90% en la mayoría de los experimentos (figura 8). Dado que la tagmentación se realiza directamente en la cromatina, se investigó si la distribución de los eventos de tagmentación se ve influenciada por la estructura de cromatina local. Además, se observaron patrones que sugieren huellas de factor de transcripción (figura 9a) y unión a nucleosomas (figura 9b, 9c). Con algoritmos de normalización y análisis personalizados, es posible inferir huellas de factor de transcripción, y se prevé inferir también regiones de estabilidad de nucleosomas y/o posicionamiento de nucleosomas a partir de los datos obtenidos mediante los métodos de la presente invención, además de la lectura de ChIP-seq regular.

Los resultados, también mostrados en los ejemplos adjuntos, establecen los métodos de la invención como una mejora de propósito general de ChIP-seq convencional que es más rápida (10-20 minutos excluyendo ChIP y la etapa de amplificación de biblioteca final, figura 2i), más rentable (figura 10), más compatible con muestras de baja entrada (figura 6) y más fácil. Se descubrió que los métodos de la invención son extremadamente robustos en un amplio intervalo de números de células y concentraciones de enzimas, diversos agentes que se unen a cromatina y diferentes protocolos de ChIP, lo que minimiza la necesidad de adaptaciones y optimizaciones del protocolo. Por consiguiente, los métodos de la invención son muy adecuados para los usos de preparación de biblioteca de secuenciación y/o mapeo de interacciones moleculares que implican a un ácido nucleico que implican a una gran cantidad de muestras, es decir, como método de alto rendimiento, se centran en poblaciones de células poco comunes, y/o se benefician de un flujo de trabajo experimental rápido, rentable y robusto. A este respecto, también se contempla que los métodos de la presente invención se usen de manera de alto rendimiento, de manera automatizada y/o de manera paralela. Por consiguiente, se contempla que las instalaciones de alto rendimiento y/o los robots usados para facilitar el pipeteado/mejorar la reproducibilidad se usen para realizar los métodos de la invención. El experto en la técnica conocerá los medios adecuados para llevar a cabo los métodos de la invención de una manera de alto rendimiento, automatizada y/o paralela. Por ejemplo, pueden usarse placas de múltiples pocillos en los métodos de la invención para realizar múltiples experimentos de manera paralela. Dichas placas de múltiples pocillos pueden tener, por ejemplo, 96, 384 ó 1536 pocillos. Las placas de múltiples pocillos también pueden usarse en combinación con robótica adecuada para experimentos de alto rendimiento. Los sistemas robóticos conocidos permiten la ejecución simultánea de múltiples experimentos, lo que reduce el tiempo, los costes y/o aumenta la fiabilidad de los datos experimentales. Por ejemplo, puede usarse la estación de trabajo Sciclone NGS (P/N SG3-31020-0300, Perkin Elmer) para permitir la preparación de muestras de secuenciación de alto rendimiento automatizada. Cuando los experimentos se realizan de manera de alto rendimiento, se prefiere que el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, se una a perlas magnéticas, tal como se describe más adelante. Las perlas magnéticas son particularmente útiles en los métodos de alto rendimiento, ya que pueden usarse fácilmente para el aislamiento de partículas unidas a las perlas de sustancias no unidas. También se prevé que los métodos de la presente invención se usen en combinación con un dispositivo microfluídico. Por ejemplo, un dispositivo de poli(dimetilsiloxano) (PDMS), que presenta una cámara microfluídica simple, puede usarse en combinación con los métodos de la presente invención. Se prefiere que la cámara microfluídica tenga una entrada y una salida, y la salida tiene una microválvula neumática en chip que puede cerrarse parcialmente ejerciendo una presión en un orificio. Las perlas magnéticas recubiertas con el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, se hacen fluir hacia la cámara microfluídica y forman un lecho relleno mientras la microválvula neumática está parcialmente cerrada. Los fragmentos de cromatina sonificados se hacen fluir a entonces través de la cámara y se adsorben sobre la superficie de las perlas. Los espacios entre las perlas son menores de 2 µm y facilitan la adsorción rápida y de alta eficiencia de fragmentos de cromatina diana bajo la pequeña longitud de difusión. Se lavan las perlas entonces mediante lavado oscilatorio en dos tampones de lavado diferentes para retirar fragmentos de cromatina adsorbidos de manera inespecífica. Finalmente, las perlas se hacen fluir fuera de la cámara y se recogen para el procesamiento fuera del chip. Este enfoque, tal como se describe por Cao *et al.* (2015) *Nature Methods* (disponible en línea), en combinación con los métodos novedosos e inventivos proporcionados en el presente documento, permite una mayor reducción del tiempo y los costes experimentales. También se contempla la combinación de los métodos proporcionados en el presente documento con Drop-Seq, tal como se describe por Macosko *et al.* (2015) *Cell* 161 (5) págs. 1202-14 y Klein *et al.* (2015) *Cell* 161 (5) págs. 1187-2201.

Por consiguiente, la presente invención se refiere, entre otras cosas, a un método para preparar una biblioteca de secuenciación. En particular, un método para preparar una biblioteca de secuenciación que comprende añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada; aislar el ácido nucleico de la cromatina; y obtener la biblioteca de secuenciación.

Con respecto al método en el que el agente que se une a cromatina es una sustancia química distinta de un anticuerpo, se sabe que un número sustancial de ligandos de molécula pequeña, incluidos fármacos terapéuticos, provocan sus efectos al unirse a proteínas específicas asociadas con el genoma. El mapeo de las interacciones globales de estas entidades químicas con la cromatina en todo el genoma podría proporcionar información sobre los mecanismos mediante los cuales una molécula pequeña influye en las funciones celulares. Puede utilizarse Chem-seq para investigar los efectos en todo el genoma de las modalidades terapéuticas y para comprender los efectos de los fármacos sobre la arquitectura nuclear en diversos contextos biológicos. En un sentido más amplio, estos métodos son útiles para mejorar la comprensión de los mecanismos terapéuticos a través de los cuales las moléculas pequeñas modulan la función y la actividad de proteínas asociadas al genoma. A través de la identificación de las dianas celulares de un fármaco, se vuelve posible obtener una mayor comprensión de las causas de los efectos secundarios y la toxicidad en las primeras etapas del desarrollo del fármaco, lo que ayuda a reducir la tasa de exclusión en el desarrollo.

Chem-seq se basa en la capacidad de crear una versión biotinilada de una molécula pequeña de interés para permitir la captura por afinidad de manera posterior. Chem-seq puede llevarse a cabo *in vitro* o *in vivo*.

Durante Chem-seq *in vivo*, las células cultivadas en medio se tratan simultáneamente con o bien una versión biotinilada de la molécula pequeña en estudio o bien DMSO (como control) y formaldehído para la reticulación de ADN, proteínas y moléculas pequeñas. Entonces se extrae la cromatina de las células, se sonica y se enriquece para regiones que contienen la molécula biotinilada de interés mediante incubación con perlas magnéticas de estreptavidina, que tienen una afinidad muy alta por la biotina. La fracción de cromatina enriquecida se eluye entonces de las perlas, se revierten las reticulaciones, se purifica el ADN, se genera una biblioteca y se somete a secuenciación de nueva generación. Las regiones genómicas enriquecidas en la biblioteca de Chem-seq en relación con el control están asociadas con la molécula pequeña en estudio. Por consiguiente, la presente invención también se refiere a un método *in vivo* para el mapeo de interacciones entre moléculas pequeñas y ácidos nucleicos, en particular ADN.

Chem-seq *in vitro* comienza con la reticulación de células cultivadas en medio con formaldehído. Los núcleos celulares se recogen entonces de las células y se extrae su cromatina. Este extracto se sonica antes de incubarse con perlas magnéticas de estreptavidina que se unen a una forma biotinilada del presente compuesto de interés. Esto proporciona una oportunidad para que la molécula pequeña de interés interactúe con sus regiones de cromatina diana. Estas regiones de cromatina se aíslan luego usando un imán y se purifica el ADN. A partir del ADN, se prepara una biblioteca y se somete a la secuenciación de nueva generación, seguida de un análisis para determinar las regiones enriquecidas para la presente molécula pequeña de interés.

Por consiguiente, la presente invención también se refiere a un método para preparar una biblioteca de secuenciación o mapear interacciones moleculares que comprenden ácido nucleico que combina el enfoque Chem-Seq con la tagmentación, tal como se describió anteriormente.

Dentro del significado de la presente invención, el término “biblioteca de secuenciación” se refiere a una representación de ácido nucleico en la que cada ácido nucleico puede identificarse mediante, por ejemplo, el uso de una etiqueta de secuencia individual. Por consiguiente, “obtener una biblioteca de secuenciación” requiere un procedimiento que pueda garantizar que se añadan secuencias adaptadoras específicas a los extremos de los fragmentos de ácido nucleico que van a analizarse. Esta preparación de ácidos nucleicos se denomina frecuentemente “biblioteca de secuenciación”. La mayoría de las aplicaciones de secuenciación de nueva generación requieren la preparación de una biblioteca de secuenciación, ácidos nucleicos con adaptadores específicos en los extremos 5' y 3'. Por ejemplo, el flujo de trabajo de secuenciación en Illumina utiliza oligonucleótidos adaptadores parcialmente complementarios que se usan para cebar la amplificación por PCR e introducir las secuencias de nucleótidos específicas requeridas para la generación de agrupaciones mediante PCR de puente y facilitar las reacciones de secuenciación por síntesis. Por consiguiente, la biblioteca de secuenciación resultante de los métodos de la presente invención para preparar una biblioteca de secuenciación es adecuada para su uso en aplicaciones de secuenciación convencionales, por ejemplo secuenciación de nueva generación tal como se describe más adelante.

Un “ácido nucleico” dentro del significado de la presente invención es un polímero de cualquier longitud que se compone de nucleótidos, que tiene preferiblemente una longitud de más de aproximadamente 50 nucleótidos. Los métodos de la invención permiten la preparación de una biblioteca de secuenciación que comprende ácidos nucleicos y/o el mapeo de interacciones moleculares que implican a un ácido nucleico. El ácido nucleico comprendido en la muestra de partida de los métodos de la invención tiene preferiblemente una longitud de aproximadamente 50 a aproximadamente 5000 nucleótidos, preferiblemente de 100 a aproximadamente 1000, más preferiblemente de aproximadamente 200 a aproximadamente 700, incluso más preferiblemente de 200 a 700, lo

más preferiblemente 200 a 300 nucleótidos. La muestra de partida no debe confundirse con el ácido nucleico comprendido en las células usadas en los métodos de la invención que comprenden cultivar y recoger células; fijar las células; lisar las células; y sonicar. A este respecto, la sonicación se usa para fragmentar el ácido nucleico comprendido y obtenido de las células, obteniéndose de ese modo la muestra de partida, también denominada segunda muestra, cuando los métodos comprenden las etapas adicionales de cultivar y recoger células; fijar las células; lisar las células; y sonicar. A este respecto, se pretende que los “nucleótidos” incluyan aquellos restos que contienen no sólo las bases de purina y pirimidina conocidas, sino también otras bases heterocíclicas que se han modificado. Dichas modificaciones incluyen purinas o pirimidinas metiladas, purinas o pirimidinas aciladas, ribosas alquiladas u otros heterociclos. Además, el término “nucleótido” incluye aquellos restos que contienen marcadores de hapteno o fluorescentes y pueden contener no sólo azúcares de ribosa y desoxirribosa convencionales, sino también otros azúcares. Los nucleósidos o nucleótidos modificados también incluyen modificaciones en el resto de azúcar, por ejemplo, en las que uno o más de los grupos hidroxilo se reemplazan por átomos de halógeno o grupos alifáticos, se funcionalizan como éteres, aminas o similares. Sin embargo, en la mayoría de los casos, los ácidos nucleicos usados en los métodos de la invención comprenderán las bases de pirimidina y purina que se producen de manera natural como desoxirribonucleótidos, ribonucleótidos que pueden participar en las interacciones de apareamientos de bases de Watson-Crick. Los nucleótidos que se producen de manera natural incluyen guanina, citosina, adenina, timina y uracilo (G, C, A, T y U, respectivamente). El ácido nucleico puede ser ADN, ARN o cualquier otro tipo de ácidos nucleicos conocidos. Se prefiere que el ácido nucleico sea ADN, en particular ADN bicatenario.

El término “agente que se une a cromatina” incluye cualquier agente que sea miembro de un complejo de unión que comprende cromatina como pareja de unión. Por ejemplo, el agente que se une a cromatina puede ser un polipéptido, tal como una proteína o fragmentos de la misma, en particular un anticuerpo; un ácido nucleico, por ejemplo un oligonucleótido, polinucleótido y similares; o una molécula pequeña, por ejemplo una sustancia química. Por tanto, en una realización, el agente que se une a cromatina es un polipéptido que tiene un dominio de unión específico para la cromatina y/o moléculas adicionales que se unen a cromatina, en particular otros polipéptidos. Por ejemplo, los agentes que se unen a cromatina pueden tener un dominio de unión metil-CpG (MBD) que reconoce la cromatina. Se prefiere que el agente que se une a cromatina sea un polipéptido, en particular un anticuerpo que se une a cromatina, en el que el anticuerpo se une específicamente a cromatina, proteínas, por ejemplo factores de transcripción o histonas, asociados con cromatina y/o ADN.

A este respecto, la cromatina tal como se usa en el presente documento es un complejo de macromoléculas que se encuentra en las células, que comprende ADN, proteínas y/o ARN. Las funciones principales de la cromatina son 1) empaquetar el ADN en un volumen más pequeño para que quepa en la célula, 2) reforzar la macromolécula del ADN para permitir la mitosis, 3) prevenir el daño en el ADN y 4) controlar la expresión génica y la replicación del ADN. Los componentes proteicos primarios de la cromatina son las histonas que compactan el ADN. La estructura de la cromatina depende de varios factores. La estructura general depende de la etapa del ciclo celular. Durante la interfase, la cromatina está estructuralmente suelta para permitir el acceso a las ARN y ADN polimerasas que transcriben y replican el ADN. La estructura de cromatina local durante la interfase depende de los genes presentes en el ADN: los genes codificantes del ADN que se transcriben de manera activa (“encendidas”) se empaquetan de manera más suelta y se encuentran asociados con ARN polimerasas (denominado eucromatina) y factores de transcripción mientras que los genes inactivos codificantes de ADN (“apagados”) se encuentran asociados con proteínas estructurales y se empaquetan de manera más apretada (heterocromatina). La modificación química epigenética de las proteínas estructurales en la cromatina también altera la estructura de cromatina local, en particular las modificaciones químicas de las proteínas histonas mediante metilación y acetilación; véase más adelante.

El elemento de repetición básico de la cromatina es el nucleosoma, interconectado por secciones de ADN ligador, una disposición mucho más corta que el ADN puro en disolución. Además de las histonas centrales, existe la histona ligadora, H1, que entra en contacto con la salida/entrada de la hebra de ADN en el nucleosoma. La partícula central del nucleosoma, junto con la histona H1, se conoce como cromatosoma. Los nucleosomas, con aproximadamente de 20 a 60 pares de bases de ADN ligador, pueden formar, en condiciones no fisiológicas, una fibra a modo de “sartas en una cadena” de aproximadamente 10 nm. Los nucleosomas se unen al ADN de manera inespecífica, según lo requiera su función en el empaquetamiento general del ADN. Sin embargo, existen preferencias de secuencias de ADN grandes que rigen el posicionamiento de nucleosomas. Esto se debe principalmente a las propiedades físicas variables de las diferentes secuencias de ADN: por ejemplo, la adenina y la timina se comprimen más favorablemente en los surcos menores internos. Esto significa que los nucleosomas pueden unirse preferentemente en una posición aproximadamente cada 10 pares de bases (la repetición helicoidal del ADN), en la que el ADN se hace rotar para maximizar el número de bases A y T que se encontrarán en el surco menor interno. Los agentes que se unen a cromatina, tal como se hace referencia en la presente invención, pueden unirse a cualquier parte de la cromatina, eucromatina o heterocromatina. Por ejemplo, los agentes que se unen a cromatina pueden interactuar con ADN, ARN o proteínas comprendidas en la cromatina. En particular, los agentes que se unen a cromatina pueden interactuar con histonas o factores de transcripción comprendidos en la cromatina y/u otras proteínas asociadas con histonas, factores de transcripción o cromatina.

A este respecto, las histonas son proteínas altamente alcalinas que se encuentran en los núcleos de células eucariotas que empaquetan y ordenan el ADN en unidades estructurales denominadas nucleosomas (véase

anteriormente). Son los principales componentes proteicos de la cromatina, actúan como bobinas alrededor de las que se enrolla el ADN y desempeñan un papel en la regulación génica. Existen cinco familias principales de histonas: H1/H5, H2A, H2B, H3 y H4. Las histonas H2A, H2B, H3 y H4 se conocen como las histonas centrales, mientras que las histonas H1 y H5 se conocen como las histonas ligadoras. Dos de cada una de las histonas centrales se ensamblan para formar un centro de nucleosoma octamérico, de aproximadamente 63 ángstroms de diámetro (una partícula similar a un solenoide (ADN)). 147 pares de bases de ADN se envuelven alrededor de esta partícula central 1,65 veces en un giro superhelicoidal a la izquierda para dar una partícula de aproximadamente 100 ángstroms de anchura. La histona ligadora H1 se une al nucleosoma en los sitios de entrada y salida del ADN, bloqueando así el ADN en su lugar y permitiendo la formación de una estructura de orden superior. La formación más básica de este tipo es la fibra o sarta de 10 nm en una conformación de cadena. Esto implica la envoltura del ADN alrededor de los nucleosomas con aproximadamente 50 pares de bases de ADN que separan cada par de nucleosomas (también denominado ADN ligador). Las estructuras de orden superior incluyen la fibra de 30 nm (formando un zigzag irregular) y la fibra de 100 nm, siendo estas las estructuras que se encuentran en las células normales. Durante la mitosis y la meiosis, los cromosomas condensados se ensamblan a través de interacciones entre nucleosomas y otras proteínas reguladoras. Los agentes que se unen a cromatina, en particular el anticuerpo o la sustancia química, pueden interactuar con histonas, es decir, pueden unirse específicamente a histonas y/o unirse a otros polipéptidos y/o sustancias químicas asociados con las histonas. Se prefiere que los agentes que se unen a cromatina, en particular el anticuerpo o la sustancia química, interactúen directamente con histonas.

A este respecto, las histonas humanas conocidas incluyen cinco clases H1/H5, H2A, H2B, H3 y H4. La clase H1 incluye H1F0, H1FNT, H1FOO, H1FX, HIST1H1A, HIST1H1B, HIST1H1C, HIST1H1D, HIST1H1E e HIST1H1T. La clase H2A incluye H2AFB1, H2AFB2, H2AFB3, H2AFJ, H2AFV, H2AFX, H2AFY, H2AFY2, H2AFZ, HIST1H2AA, HIST1H2AB, HIST1H2AC, HIST1H2AD, HIST1H2AE, HIST1H2AG, HIST1H2AI, HIST1H2AJ, HIST1H2AK, HIST1H2AL, HIST1H2AM, HIST2H2AA3 e HIST2H2AC. La clase H2B incluye H2BFM, H2BFS, H2BFWT, HIST1H2BA, HIST1H2BB, HIST1H2BC, HIST1H2BD, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2BH, HIST1H2BI, HIST1H2BJ, HIST1H2BK, HIST1H2BL, HIST1H2BM, HIST1H2BN, HIST1H2BO e HIST2H2BE. La clase H3 incluye HISTH3A, HISTH3B, HISTH3C, HISTH3D, HISTH3E, HISTH3F, HISTH3G, HISTH3H, HISTH3I, HISTH3J, HIST2H3C e HIST3H3. La clase H4 incluye HIST1H4A, HIST1H4B, HIST1H4C, HIST1H4D, HIST1H4E, HIST1H4F, HIST1H4G, HIST1H4H, HIST1H4I, HIST1H4J, HIST1H4K, HIST1H4L e HIST4H4. Se prefiere que el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, se una a histonas de clase H3, en particular H3.3, H3.2, H3.3A, H3.3B o H3.1. Además, se prefiere que el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, se una a H4, H2A.Z o CENP-A (conteniendo estas dos últimas un plegamiento de histona relacionado con la histona H3).

Además, se prevé que los agentes que se unen a cromatina, en particular el anticuerpo o la sustancia química, sean específicos para versiones modificadas de las histonas conocidas. Se ha descrito un gran catálogo de modificaciones de histonas. Las modificaciones de histonas tienen significados y consecuencias específicos para la traducción genómica y la accesibilidad del ADN para proteínas de unión y/u otras sustancias químicas adicionales. Por consiguiente, se prevé que los métodos de la invención se usen para identificar regiones unidas por histonas modificadas que pueden experimentar alteraciones en la expresión génica, por ejemplo en células/tejidos enfermos tales como células cancerosas.

Las modificaciones de histonas conocidas incluyen metilación, acetilación, propionilación, butirilación, crotonilación, 2-hidroxiisobutirilación, malonilación, succinilación y ribosilación. En particular, metilación de lisina, metilación de arginina, acetilación de lisina, fosforilación de serina/treonina/tirosina. A este respecto, la adición de uno, dos o tres grupos metilo a lisina tiene poco efecto sobre la química de la histona; la metilación deja intacta la carga de la lisina y añade un número mínimo de átomos de modo que las interacciones estéricas no se vean afectadas en su mayor parte. Sin embargo, las proteínas que contienen dominios Tudor, cromodominios o dominios PHD, entre otros, pueden reconocer la metilación de lisina con una sensibilidad exquisita y diferenciar entre mono, di y tri-metil-lisina en la medida en que, para algunas lisinas (por ejemplo: H4K20) la mono, di y tri-metilación tienen diferentes significados. Debido a esto, la metilación de lisina es una marca muy informativa y domina las funciones de modificación de histonas conocidas. Por consiguiente, se prevé que los agentes que se unen a cromatina sean específicos para proteínas y/o histonas con lisina metilada que reconocen tales histonas modificadas, por ejemplo, proteínas que contienen dominios Tudor, cromodominios o dominios PHD. Con respecto a las histonas con arginina metilada, se aplica un razonamiento similar al anterior, es decir, algunos dominios de proteínas, por ejemplo dominios Tudor, pueden ser específicos para metil-arginina en lugar de metil-lisina. Se sabe que la arginina se mono o di-metila, y la metilación puede ser simétrica o asimétrica, potencialmente con diferentes significados. Con respecto a la acetilación de lisina, la adición de un grupo acetilo tiene un efecto químico importante sobre la lisina, ya que neutraliza la carga positiva. Esto reduce la atracción electrostática entre la histona y la estructura principal de ADN cargada negativamente, soltando la estructura de la cromatina; Las histonas altamente acetiladas forman una cromatina más accesible y tienden a asociarse con la transcripción activa. La acetilación de lisina parece tener un significado menos preciso que la metilación, ya que las histona acetiltransferasas tienden a actuar en más de una lisina; presumiblemente esto refleja la necesidad de alterar múltiples lisinas para tener un efecto significativo sobre la estructura de la cromatina. Por consiguiente, también se prevé que el agente que se une a cromatina sea específico para lisina acetilada y/o proteínas que interactúan con lisina acetilada. Además de lo anterior, la serina/treonina y/o tirosina comprendidas en histonas pueden modificarse mediante fosforilación. La adición de un grupo fosfato

cargado negativamente puede conducir a cambios importantes en la estructura de la proteína, lo que conduce al papel bien caracterizado de fosforilación en el control de la función de la proteína. La fosforilación de histonas tiene funciones claras como una modificación postraduccional, y se han caracterizado dominios de unión tales como BRCT (dominio de extremo C-terminal de BRCA1). Por tanto, también se prevé que tales histonas modificadas, es decir, modificadas mediante fosforilación, sean reconocidas por los agentes que se unen a cromatina.

Las modificaciones de las histonas descritas anteriormente y otras modificaciones descritas en la técnica tienen implicaciones para el control de la transcripción. A este respecto, dos modificaciones de histonas conocidas están asociadas particularmente con la transcripción activa: trimetilación de lisina 4 de H3 (H3K4Me3) y trimetilación de lisina 36 de H3 (H3K36Me3). H3K4Me3 se produce en el promotor de genes activos y lo realiza el complejo COMPASS. La modificación es una excelente marca de promotores activos y el nivel de esta modificación de histona en el promotor de un gen está correlacionado ampliamente con la actividad transcripcional del gen. La formación de esta marca está vinculada con la transcripción de una manera bastante enrevesada: al comienzo de la transcripción de un gen, la ARN polimerasa II experimenta un cambio de iniciación 'a' elongación', marcado por un cambio en los estados de fosforilación del dominio C-terminal (CTD) de la ARN polimerasa II. La misma enzima que fosforila el CTD también fosforila el complejo Rad6, que añade a su vez una marca de ubiquitina a K123 de H2B (K120 en mamíferos). H2BK123Ub se produce en todas las regiones transcritas, pero esta marca es necesaria para que COMPASS trimetile H3K4 en los promotores. Por tanto, en un aspecto preferido de la invención, el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, es específico para H3K4Me3. En un aspecto adicional, el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, es específico para H3K36Me3. Esta trimetilación se produce en el cuerpo de genes activos y se deposita por la metiltransferasa Set2. Esta proteína se asocia con la elongación de la ARN polimerasa II, y H3K36Me3 es indicativo de genes transcritos de manera activa. H3K36Me3 es reconocido por el complejo de histona desacetilasa Rpd3, que elimina las modificaciones de acetilo de las histonas circundantes, aumentando la compactación de la cromatina y reprimiendo la transcripción espuria. El aumento de la compactación de la cromatina impide que los factores de transcripción accedan al ADN y reduce la probabilidad de que se inicien nuevos eventos de transcripción dentro del cuerpo del gen. Por tanto, este proceso ayuda a garantizar que la transcripción no se interrumpa. Además, la acetilación de la lisina 27 de la histona H3 (H3K27ac) está presente en elementos reguladores activos como promotores y potenciadores. En genética, un potenciador es una región corta (50-1500 pb) de ADN que puede unirse con proteínas (activadores) para activar la transcripción de un gen. Estas proteínas se denominan generalmente factores de transcripción. Los potenciadores son generalmente de acción en cis, estando ubicados a una distancia de hasta 1 Mpb (1.000.000 pb) del gen y pueden estar en el sentido de 5' o de 3' del sitio de inicio, y o bien en sentido hacia adelante o bien hacia atrás. Hay cientos de miles de potenciadores en el genoma humano. En particular, se describió que H3K27ac distinguía elementos reguladores activos de listos para actuar. El enriquecimiento de H3K27ac en estos elementos es un buen indicador para la expresión del elemento genético asociado. Por consiguiente, el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, usado en los métodos de la presente invención puede ser específico para H3K27ac.

Las modificaciones de histonas también pueden estar asociadas con la represión de la expresión génica. Por ejemplo, se sabe que H3K27Me3, H3K9Me2/3 y H4K20Me3 están asociadas con genes reprimidos. H3K27Me3 se deposita por el complejo de Polycomb PRC2. Es un marcador claro de la represión génica, y es probable que otras proteínas lo obliguen a ejercer una función represora. Otro complejo de Polycomb, PRC1, puede unirse a H3K27Me3 y añade la modificación de histona H2AK119Ub que ayuda a la compactación de la cromatina. La di y la tri-metilación de la lisina 9 de H3 (H3K9Me2/3) es un marcador bien caracterizado de heterocromatina y, por tanto, está fuertemente asociado con represión génica. Lo mismo se aplica a H4K20Me3, que está asociado estrechamente con la heterocromatina. Esta marca la coloca la metiltransferasa Suv4-20h, que al menos en parte es reclutada por la proteína de heterocromatina 1. Por consiguiente, también se contempla que los agentes que se unen a cromatina usados en los métodos de la invención se unan específicamente a tales histonas modificadas asociadas con genes reprimidos y/o proteínas asociadas con los mismos.

Las modificaciones de histonas también desempeñan un papel en la reparación del ADN y la condensación cromosómica. Por ejemplo, marcar sitios de daño en el ADN es una función importante para las modificaciones de histonas. También protege el ADN frente a verse destruido por la radiación ultravioleta del sol. Por ejemplo, el H2AX fosforilado (también conocido como gamma-H2AX) es un marcador de roturas de doble cadena de ADN y forma parte de la respuesta al daño en el ADN. H2AX se fosforila temprano después de la detección de la rotura de doble cadena de ADN, y forma un dominio que se extiende muchas kilobases a ambos lados del daño. Gamma-H2AX actúa como sitio de unión para la proteína MDC1, que recluta a su vez proteínas de reparación de ADN clavar y, como tal, gamma-H2AX forma una parte vital de la maquinaria que garantiza la estabilidad del genoma. Además, se requiere H3K56Acx para la estabilidad del genoma. H3K56 se acetila por el complejo p300/Rtt109, pero se desacetila rápidamente alrededor de los sitios de daño en el ADN. La acetilación de H3K56 también se requiere para estabilizar las horquillas de replicación detenidas, impidiendo colapsos peligrosos de la horquilla de replicación. La fosforilación de H3 en la serina 10 (fosfo-H3S10) está asociada con condensación, pero la fosforilación de H3S10 también está presente en determinados sitios cromosómicos fuera de la mitosis, por ejemplo, en la heterocromatina pericéntrica de células durante G2. La fosforilación de H3S10 también se ha vinculado con el daño en el ADN provocado por la formación del bucle R en sitios altamente transcritos. La fosforilación de H2B en la serina 10 (levadura) o la serina 14 (mamíferos) también está vinculada con la condensación de cromatina, pero con

el propósito muy diferente de mediar en la condensación cromosómica durante la apoptosis. Esta marca no es simplemente un espectador de acción tardía en la apoptosis ya que las mutaciones portadoras de levadura de este residuo son resistentes a la muerte celular apoptótica inducida por peróxido de hidrógeno.

Por consiguiente, los agentes que se unen a cromatina, en particular el anticuerpo o la sustancia química, usados en los métodos de la invención pueden unirse específicamente a histonas, histonas modificadas y/u otros factores, en particular polipéptidos tales como enzimas, que interactúan con tales histonas y/o histonas modificadas. Cuando los agentes que se unen a cromatina, en particular el anticuerpo o la sustancia química, se unen a histonas modificadas, se prefiere que el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, se una a H3K4me1/2/3, H2BK5me1, H3K27me1/2/3, H3K9me1/2/3, H4K20me1, H3K79me1, H3K36me3, H2AK5ac, H2AK9ac, H2BK5ac, H2BK12ac, H2BK20ac, H2BK120ac, H3K4ac, H3K9ac, H3K14ac, H3K18ac, H3K23ac, H3K27ac, H3K36ac, H4K5ac, H4K8ac, H4K12ac, H4K16ac, H4K91ac, H2Aub o H2Bub.

En una realización adicional, los agentes que se unen a cromatina, en particular el anticuerpo o la sustancia química, usados en los métodos de la invención se unen específicamente a factores de transcripción. Un factor de transcripción (a veces denominado factor de unión a ADN específico de secuencia) es una proteína que se une a secuencias de ADN específicas, controlando de ese modo la tasa de transcripción de información genética del ADN al ARN mensajero. Los factores de transcripción a modo de ejemplo incluyen, entre otros, AAF, abl, ADA2, ADANF1, AF-1, AFP1, AhR, AIN3, ALL-1, alfa-CBF, alfa-CP 1, alfa-CP2a, alfaHo, alfaH2-alfaH3, Alx-4, aMEF-2, AML1, AML1a, AML1b, AML1c, AML1deltaN, AML2, AML3, AML3a, AML3b, AMY-1L, A-Myb, ANF, AP-1, AP-2alfaA, AP-2alfaB, AP-2beta, AP-2gamma, AP-3 (1), AP-3 (2), AP-4, AP-5, APC, AR, AREB6, Arnt, Arnt (forma de 774 M), ARP-1, ATBF1-A, ATBF1-B, ATF, ATF-1, ATF-2, ATF-3, ATF-3deltaZIP, ATF-a, ATF-adelta, ATPF1, Barhl1, Barhl2, Barxl, Barx2, Bcl-3, BCL-6, BD73, beta-catenina, Binl, B-Myb, BP1, BP2, brahma, BRCA1, Brn-3a, Brn-3b, Brn-4, BTEB, BTEB2, B-TFIID, C/EBPalfa, C/EBPbeta, C/EBPdelta, factor de unión a CACC, Cart-1, CBF (4), CBF (5), CBP, factor de unión a CCAAT, factor de unión a CCMT, CCF, CCG1, CCK-1a, CCK-1b, CD28RC, cdk2, cdk9, Cdx-1, CDX2, Cdx-4, CFF, ChxIO, CLIM1, CLIM2, CNBP, CoS, COUP, CPI, CPIA, CPIC, CP2, CPBP, proteína de unión a CPE, CREB, CREB-2, CRE-BPI, CRE-BPa, CREMalfa, CRF, Crx, CSBP-1, CTCF, CTF, CTF-1, CTF-2, CTF-3, CTF-5, CTF-7, CUP, CUTL1, Cx, ciclina A, ciclina T1, ciclina T2, ciclina T2a, ciclina T2b, DAP, DAX1, DB1, DBF4, DBP, DbpA, DbpAv, DbpB, DDB, DDB-1, DDB-2, DEF, deltaCREB, deltaMax, DF-1, DF-2, DF-3, Dlx-1, Dlx-2, Dlx-3, Dlx4 (isoforma larga), Dlx-4 (isoforma corta), Dlx-5, Dlx-6, DP-1, DP-2, DSIF, DSIF-pl4, DSIF-pl60, DTF, DUX1, DUX2, DUX3, DUX4, E, E1 2, E2F, E2F+E4, E2F+pl07, E2F-1, E2F-2, E2F-3, E2F-4, E2F-5, E2F-6, E47, E4BP4, E4F, E4F1, E4TF2, EAR2, EBP-80, EC2, EF1, EF-C, EGR1, EGR2, EGR3, EllaE-A, EllaE-B, EllaE-Calfa, EllaE-Cbeta, EivF, Elf-1, Elk-1, Emx-1, Emx-2, Emx-2, En-1, En-2, proteína de unión a ENH, ENKTF-1, EPAS1, épsilonFI, ER, Erg-1, Erg-2, ERR1, ERR2, ETF, Ets-1, Ets-1 deltaVil, Ets-2, Evx-1, F2F, factor 2, nombre de factor, FBP, f-EBP, FKBP59, FKHL18, FKHL1P2, Fli-1, Fos, FOXB1, FOXC1, FOXC2, FOXD1, FOXD2, FOXD3, FOXD4, FOXE1, FOXE3, FOXF1, FOXF2, FOXG1a, FOXG1b, FOXG1c, FOXH1, FOXI1, FOXJ1a, FOXJ1b, FOXJ2 (isoforma larga), FOXJ2 (isoforma corta), FOXJ3, FOXK1a, FOXK1b, FOXK1c, FOXL1, FOXM1a, FOXM1b, FOXM1c, FOXN1, FOXN2, FOXN3, FOXO1a, FOXO1b, FOXO2, FOXO3a, FOXO3b, FOXO4, FOXP1, FOXP3, Fra-1, Fra-2, FTF, FTS, factor G, factor G6, GABP, GABP-alfa, GABP-beta1, GABP-beta2, GADD 153, GAF, gammaCMT, gammaCAC1, gammaCAC2, GATA-1, GATA-2, GATA-3, GATA-4, GATA-5, GATA-6, Gbx-1, Gbx-2, GCF, GCMA, GCNS, GF1, GLI, GLI3, GR alfa, GR beta, GRF-1, Gsc, Gscl, GT-IC, GT-IIA, GT-IIBalfa, GT-IIBbeta, H1TF1, H1TF2, H2RIIBP, H4TF-1, H4TF-2, HAND1, HAND2, HB9, HDAC1, HDAC2, HDAC3, hDaxx, factor inducido por calor, HEB, HEBI-p67, HEBI-p94, HEF-1B, HEF-1T, HEF-4C, HEN1, HEN2, Hesxl, Hex, HIF-1, HIF-1alfa, HIF-1beta, HiNF-A, HiNF-B, HiNF-C, HiNF-D, HiNF-D3, HiNF-E, HiNF-P, HIP1, HIV-EP2, Hif, HLTf, HLTf (Met123), HLX, HMBP, HMG I, HMG I(Y), HMG Y, HMGI-C, HNF-1A, HNF-1B, HNF-1C, HNF-3, HNF-3alfa, HNF-3beta, HNF-3gamma, HNF4, HNF-4alfa, HNF-4alfal, HNF-4alfa2, HNF-4alfa3, HNF-4alfa4, HNF4gamma, HNF-6alfa, hnRNP K, HOX11, HOXA1, HOXAIO, HOXAIO PL2, HOXA1 I, HOXA13, HOXA2, HOXA3, HOXA4, HOXA5, HOXA6, HOXA7, HOXA9A, HOXA9B, HOXB-1, HOXB13, HOXB3, HOXB3, HOXB4, HOXB5, HOXB6, HOXA5, HOXB7, HOXB8, HOXB9, HOXC10, HOXC11, HOXC12, HOXC13, HOXC4, HOXC5, HOXC6, HOXC8, HOXC9, HOXD10, HOXD11, HOXD12, HOXD13, HOXD3, HOXD4, HOXD8, HOXD9, Hp55, Hp65, HPX42B, HrpF, HSF, HSF1 (largo), HSF1 (corto), HSF2, hsp56, Hsp90, IBP-1, ICER-II, ICER-ligamma, ICSBP, Idl, Idl H', Id2, Id3, Id3/Heir-1, IF1, IgPE-1, IgPE-2, IgPE-3, IkappaB, IkappaB-alfa, IkappaB-beta, IkappaBR, Il-1 RF, IL-6 RE-BP, 11-6 RF, INSAF, IPF1, IRF-1, IRF-2, B, IRX2a, Irx-3, Irx-4, ISGF-1, ISGF-3, ISGF3alfa, ISGF-3gamma, 1st-1, ITF, ITF-1, ITF-2, JRF, Jun, JunB, JunD, factor kappay, KBP-1, KER1, KER-1, Koxl, KRF-1, autoantígeno Ku, KUP, LBP-1, LBP-la, LBXI, LCR-FI, LEF-1, LEF-1B, LF-A1, LHX1, LHX2, LHX3a, LHX3b, LHX5, LHX6.1a, LHX6.1b, LIT-1, Lmol, Lmo2, LMX1A, LMX1B, L-Milo (forma larga), L-Milo (forma corta), L-My2, LSF, LXRA1a, LyF-1, Lil-I, factor M, Madl, MASH-1, Maxl, Max2, MAZ, MAZ1, MB67, MBF1, MBF2, MBF3, MBP-1 (1), MBP-1 (2), MBP-2, MDBP, MEF-2, MEF-2B, MEF-2C (forma de 433 AA), MEF-2C (forma de 465 AA), MEF-2C (forma de 473 M), MEF-2C/delta32 (forma de 441 AA), MEF-2D00, MEF-2D0B, MEF-2DA0, MEF-2DAO, MEF-2DAB, MEF-2DA'B, Meis-1, Meis-2a, Meis-2b, Meis-2c, Meis-2d, Meis-2e, Meis3, Meoxl, Meoxla, Meox2, MHox (K-2), Mi, MIF-1, Miz-1, MM-1, MOP3, MR, Msx-1, Msx-2, MTB-Zf, MTF-1, mtTFI, Mxil, Myb, Myc, Myc 1, Myf-3, Myf-4, Myf-5, Myf-6, MyoD, MZF-1, NCI, NC2, NCX, NELF, NER1, Net, NF III-a, NF NF NF-1, NF-1A, NF-1B, NF-1X, NF-4FA, NF-4FB, NF-4FC, NF-A, NF-AB, NFAT-1, NF-AT3, NF-Atc, NF-Atp, NF-Atx, Nf etaA, NF-CLEOA, NF-CLEOB, NFdeltaE3A, NFdeltaE3B, NFdeltaE3C, NFdeltaE4A, NFdeltaE4B, NFdeltaE4C, Nfe, NF-E, NF-E2, p45 de NF-E2, NFE3, NFE-6, NF-Gma, NF-GMb, NF-IL-2A, NF-IL-2B, NF-jun, NF-kappaB, (similar a) NF-kappaB, NF-kappaBI, NF-kappaB1, precursor, NF-kappaB2, NF-kappaB2 (p49), precursor de NF-kappaB2, NF-kappaEI, NF-kappaE2, NF-kappaE3, NFMHCIIA, NF-MHCIIIB, NF-muEI, NF-muE2, NF-muE3, NF-S, NF-X, NF-X1,

NF-X2, NF-X3, NF-Xc, NF-YA, NF-Zc, NFZz, NHP-1, NHP-2, NHP3, NHP4, NKX2-5, NKX2B, NKX2C, NKX2G, NKX3A, NKX3A v1, NKX3A v2, NKX3A v3, NKX3A v4, NKX3B, NKX6A, Nmi, N-Myc, N-Oct-2alfa, N-Oct-2beta, N-Oct-3, N-Oct-4, N-Oct-5a, N-Oct-5b, NP-TCII, NR2E3, NR4A2, Nrfl, Nrf-1, Nrf2, NRF-2beta, NRF-2gamma, NRL, forma 1 de NRSF, forma 2 de NRSF, NTF, O2, OCA-B, Oct-1, Oct-2, Oct-2.1, Oct-2B, Oct-2C, Oct-4A, Oct-4B, Oct-5, Oct-6, factor Octa, factor de unión a octámero, oct-B2, oct-B3, Otx1, Otx2, OZF, pl07, pl30, modulador p28, p300, p38erg, p45, p49erg, -p53, p55, p55erg, p65delta, p67, Pax-1, Pax-2, Pax-3, Pax-3A, Pax-3B, Pax-4, Pax-5, Pax-6, Pax-6/Pd-5a, Pax-7, Pax-8, Pax-8a, Pax-8b, Pax-8c, Pax-8d, Pax-8e, Pax-8f, Pax-9, Pbx-1a, Pbx-1b, Pbx-2, Pbx-3a, Pbx-3b, PC2, PC4, PC5, PEA3, PEBP2alfa, PEBP2beta, Pit-1, PITX1, PITX2, PITX3, PKNOX1, PLZF, PO-B, pontina 52, PPARalfa, PPARbeta, PPARgamma, PPARgamma2, PPUR, PR, PR A, pRb, PRD1-BF1, PRDI-BFc, Prop-1, PSE1, P-TEFb, PTF, PTFalfa, PTFbeta, PTFdelta, PTFgamma, factor de unión a caja Pu, factor de unión a caja Pu (B JA-B), PU.1, PuF, factor Pur, R1, R2, RAR-alfa, RAR-beta, RAR-beta2, RAR-gamma, RARgamma, RBP60, RBP-Jkappa, Rel, RelA, RelB, RFX, RFXI, RFX2, RFX3, RFXS, RF-Y, RORalfa, RORalfa2, RORalfa3, RORbeta, RORgamma, Rox, RPF1, RPGalfa, RREB-1, RSRFC4, RSRFC9, RVF, RXR-alfa, RXRbeta, SAP-la, SAPlb, SF-1, SHOX2a, SHOX2b, SHOXa, SHOXb, SHP, SIII-pl10, SIII-pl5, SIII-pl8, SIM', Six-1, Six-2, Six-3, Six-4, Six-5, Six-6, SMAD-1, SMAD-2, SMAD-3, SMAD-4, SMAD-5, SOX-11, SOX-12, Sox-4, Sox-5, SOX-9, Spl, Sp2, Sp3, Sp4, factor Sph, Spi-B, SPIN, SRCAP, SREBP-la, SREBP-lb, SREBP-lc, SREBP-2, SRE-ZBP, SRF, SRY, SRPI, Staf-50, STAT1alfa, STAT1beta, STAT2, STAT3, STAT4, STAT6, T3R, T3R-alfa, T3R-alfa2, T3Rbeta, TAF(I)110, TAF(I)48, TAF(I)63, TAF(II)100, TAF(II)125, TAF(II)135, TAF(II)170, TAF(II)18, TAF(II)20, TAF(II)250, TAF(II)250Delta, TAF(II)28, TAF(II)30, TAF(II)31, TAF(II)55, TAF(II)70-alfa, TAF(II)70-beta, TAF(II)70-gamma, TAFI, TAF-II, TAF-L, Tal-1, Tal-lbeta, Tal-2, factor TAR, TBP, TBX1A, TBX1B, TBX2, TBX4, TBXS (isoforma larga), TBXS (isoforma corta), TCF, TCF-1, TCF-1A, TCF-1B, TCF-1C, TCF-1D, TCF-1E, TCF-1F, TCF-1G, TCF-2alfa, TCF-3, TCF-4, TCF-4(K), TCF-4B, TCF-4E, TCFbeta, TEF-1, TEF-2, tel, TFE3, TFEB, TFIIA, precursor de TFIIA-alfa/beta, precursor de TFIIAalfa/beta, TFIIA-gamma, TFIIB, TFIID, TFIIE, TFIIE-alfa, TFIIE-beta, TFIIF, TFIIF-alfa, TFIIF-beta, TFIIH, TFIIH*, TFIIH-CAK, TFIIH-ciclina H, TFIIH-ERCC2/CAK, TFIIH-MAT1, TFIIH-M015, TFIIH-p34, TFIIH-p44, TFIIH-p62, TFIIH-p80, TFIIH-p90, TFII-I, Tf-LFI, Tf-LF2, TGIF, TGIF2, TGT3, THRAl, TIF2, TLE1, TLX3, TMF, TR2, TR2-11, TR2-9, TR3, TR4, TRAP, TREB-1, TREB-2, TREB-3, TREFl, TREF2, TRF (2), TTF-1, TXRE BP, TxREF, UBF, UBP-1, UEF-1, UEF-2, UEF-3, UEF-4, USF1, USF2, USF2b, Vav, Vax-2, VDR, vHNF-IA, vHNF-IB, vHNF-IC, VITF, WSTF, WT1, WT1I, WT1 I-KTS, WT1 I-del2, WT1-KTS, WTI-del2, X2BP, XBP-1, XW-V, XX, YAF2, YB-1, YEBP, YY1, ZEB, ZF1, ZF2, ZFX, ZHX1, ZIC2, ZID, ZNF 174 y similares.

Los factores de transcripción realizan esta función solos o con otras proteínas en un complejo, promoviendo (como activador) o bloqueando (como represor) el reclutamiento de ARN polimerasa (la enzima que realiza la transcripción de información genética de ADN a ARN) a genes específicos. Por consiguiente, el agente que se une a cromatina puede interaccionar directamente con un factor de transcripción, comprendiendo el complejo uno o más factores de transcripción y/o proteínas asociadas con factores de transcripción. Una característica definitoria de los factores de transcripción es que contienen uno o más dominios de unión a ADN (DBD, por sus siglas en inglés), que se unen a secuencias específicas de ADN adyacentes a los genes que regulan. Proteínas adicionales, tales como coactivadores, remodeladores de cromatina, histona acetilasas, desacetilasas, cinasas y metilasas, a la vez que desempeñan un papel crucial en la regulación génica, carecen de dominios de unión a ADN y, por tanto, no se clasifican como factores de transcripción. Sin embargo, los agentes que se unen a cromatina usados en los métodos de la invención también pueden interaccionar con tales proteínas. Los factores de transcripción se unen a regiones potenciadoras o promotoras de ADN adyacentes a los genes que regulan. Dependiendo del factor de transcripción, la transcripción del gen adyacente está regulada o bien por incremento o bien por disminución. Los factores de transcripción usan una variedad de mecanismos para la regulación de la expresión génica. Estos mecanismos incluyen: estabilizar o bloquear la unión de la ARN polimerasa al ADN; catalizar la acetilación o desacetilación de proteínas histonas. El factor de transcripción puede o bien hacer esto directamente o bien reclutar otras proteínas con esta actividad catalítica. Muchos factores de transcripción usan uno u otro de los dos mecanismos opuestos para regular la transcripción: la actividad histona acetiltransferasa (HAT): acetila las proteínas histonas, lo que debilita la asociación del ADN con histonas, lo que hace que el ADN sea más accesible para la transcripción, regulando por incremento de ese modo la transcripción; y/o actividad histona desacetilasa (HDAC): desacetila las proteínas histonas, lo que fortalece la asociación del ADN con histonas, lo que hace que el ADN sea menos accesible para la transcripción, regulando por disminución de ese modo la transcripción. Los mecanismos para la regulación de la expresión génica también incluyen el reclutamiento de proteínas coactivadoras o correpresoras al complejo ADN-factor de transcripción.

Muchos factores de transcripción tienen importancia clínica por al menos dos motivos: (1) las mutaciones pueden asociarse con enfermedades específicas y (2) pueden ser dianas de medicamentos. Por consiguiente, los agentes que se unen a cromatina usados en los métodos de la presente invención pueden ser relevantes en el diagnóstico y/o tratamiento de enfermedades asociadas con factores de transcripción. Por ejemplo, debido a sus importantes papeles en el desarrollo, la señalización intercelular y el ciclo celular, algunas enfermedades humanas se han asociado con mutaciones en factores de transcripción. Además, muchos factores de transcripción son o bien supresores tumorales o bien oncogenes y, por tanto, las mutaciones o la regulación aberrante de los mismos se asocia con cáncer. Se sabe que al menos tres grupos de factores de transcripción son importantes en el cáncer humano: (1) las familias de NF-kappaB y AP-1, (2) la familia de STAT y (3) los receptores de esteroides. Otros factores de transcripción implicados en enfermedades humanas se muestran en la siguiente tabla:

Tabla 1

Estado	Descripción	Locus
Síndrome de Rett	Se asocian mutaciones en el factor de transcripción MECP2 con el síndrome de Rett, un trastorno del neurodesarrollo.	Xq28
Diabetes	Una forma poco común de diabetes denominada MODY (diabetes hereditaria juvenil de tipo 2) puede estar provocada por mutaciones en los factores nucleares de hepatocitos (HNF) o el factor promotor de insulina-1 (IPF1/Pdx1).	múltiple
Dispraxia verbal del desarrollo	Se asocian mutaciones en el factor de transcripción FOXP2 con dispraxia verbal del desarrollo, una enfermedad en la que los individuos no pueden producir los movimientos coordinados con precisión necesarios para el habla.	7q31
Enfermedades autoinmunitarias	Mutaciones en el factor de transcripción FOXP3 provocan una forma poco común de enfermedad autoinmunitaria denominada IPEX.	Xp11.23-q13.3
Síndrome de Li-Fraumeni	Provocado por mutaciones en el supresor tumoral p53.	17p13.1
Cáncer de mama	La familia de STAT es relevante para el cáncer de mama.	múltiple
Múltiples cánceres	La familia de HOX está implicada en una variedad de cánceres.	múltiple

Por consiguiente, los agentes que se unen a cromatina, en particular el anticuerpo o la sustancia química, usados en los métodos de la presente invención pueden interaccionar con factores de transcripción que se sabe que están asociados con enfermedades, por ejemplo cáncer. A este respecto, los métodos de la invención pueden usarse para estudiar la interacción entre ADN y factores de transcripción en una célula enferma y/o células derivadas de tejido enfermo. Además, los métodos de la presente invención pueden usarse para estudiar interacciones entre fármacos y factores de transcripción/ADN. A este respecto, aproximadamente el 10% de los medicamentos recetados actualmente seleccionan como diana directamente la clase de factores de transcripción del receptor nuclear. Los ejemplos incluyen tamoxifeno y bicalutamida para el tratamiento del cáncer de mama y de próstata, respectivamente, y diversos tipos de esteroides antiinflamatorios y anabólicos. Además, los factores de transcripción a menudo se modulan indirectamente por fármacos a través de cascadas de señalización.

Según lo anterior, la presente invención se refiere a métodos para el mapeo de interacciones moleculares que implican a un ácido nucleico, en particular ADN, en los que el método proporciona información valiosa con respecto a la interacción de polipéptidos con un ácido nucleico, en particular ADN. El ácido nucleico puede derivarse de cualquier fuente, por ejemplo células. En particular, células que comprenden complejos ácido nucleico-proteína. Se prefiere que las células sean células humanas, células animales, células bacterianas, células de levadura, células de arqueas, células vegetales o virus. Se prefiere más que las células sean células humanas. Sin embargo, las células también pueden proceder de fuentes no nativas, por ejemplo células modificadas por ingeniería o células modificadas artificialmente, en particular células modificadas genéticamente. Además, las células humanas o animales pueden ser células enfermas o células no enfermas o células derivadas de tejido enfermo o no enfermo. A este respecto, las células humanas o animales pueden ser células cancerosas, células inmunitarias, células sanguíneas o células madre. Se prefiere que las células sean células cancerosas. El cáncer puede ser un cáncer sólido o de la sangre, en particular leucemia o un tumor. Los cánceres conocidos asociados con la transcripción alterada, es decir, la accesibilidad alterada al ADN, histonas modificadas, factores de transcripción modificados y similares, se resumen por Yeh *et al.* (2013) Curr. Opin. Oncol 25(6). Las células también pueden ser células embrionarias.

Debido a que los métodos de la invención son particularmente útiles para el análisis de bajos números de células, resulta evidente que se prevén en particular fuentes que tienen un número limitado de células disponibles como fuente del ácido nucleico que va a analizarse. En casos de enfermedades, en particular enfermedades humanas, los números de células pueden estar restringidos por la naturaleza de la enfermedad, por ejemplo, metástasis de cáncer, tumores primarios pequeños u órganos enfermos pequeños, tejidos poco comunes y tipos celulares poco comunes. Los números de células de muestras clínicas humanas pueden restringirse adicionalmente mediante el enfoque para obtener la muestra, por ejemplo, biopsias por punción o extracciones de sangre. Por consiguiente, también se contemplan muestras derivadas de tales fuentes para su uso en los métodos de la presente invención. Además, los números de células pueden estar limitados debido a otras restricciones, por ejemplo, animales protegidos, animales poco comunes, animales en peligro de extinción o similares. Además, los métodos de la invención son particularmente útiles en estudios de un único animal, en particular de animales pequeños, tales como *C. elegans* o pez cebra.

En los métodos de la invención, antes de la preparación de una biblioteca de secuenciación o el mapeo de interacciones moleculares que implican a un ácido nucleico, se prepara preferiblemente la muestra que comprende un ácido nucleico cultivando y recogiendo células; fijando las células; lisando las células y obteniéndose de ese modo una primera muestra que comprende un ácido nucleico; y sonicando la muestra y obteniéndose de ese modo una segunda muestra que comprende un ácido nucleico. Se prefiere que dicha segunda muestra se use en los métodos de la invención para la preparación de una biblioteca de secuenciación o el mapeo de interacciones moleculares que implican a un ácido nucleico. Cuando la muestra que comprende un ácido nucleico es una muestra de células primarias, por ejemplo una muestra derivada de un donante, puede omitirse la etapa de cultivo y recogida. Por consiguiente, cuando la muestra que comprende un ácido nucleico es una muestra de células primarias, los métodos de la invención preferiblemente comprenden además fijar las células; lisar las células y obtener de ese modo una primera muestra que comprende un ácido nucleico; y sonicar la primera muestra y obtener de ese modo una segunda muestra que comprende un ácido nucleico.

Por consiguiente, se prepara preferiblemente la muestra que comprende un ácido nucleico mediante un método que comprende el cultivo y la recogida de células. Esto puede realizarse usando métodos bien conocidos en la técnica. En particular, los métodos de cultivo deben ser adecuados para el tipo celular usado en el análisis. Tales métodos se describen, por ejemplo, en Helgason *et al.* (2005) Basic Cell Culture Protocols, Methods in Molecular Biology o Freshney (2010) Culture of Animal Cells, Wiley-Blackwell. La recogida de células también se realiza mediante métodos bien conocidos descritos en la técnica. Por ejemplo, las células pueden recogerse mediante centrifugación, mediante lo cual las células se encuentran en el sedimento celular resultante mientras el sobrenadante contiene el medio de cultivo usado.

De manera posterior a la recogida de células cultivadas, las células pueden fijarse. La fijación se usa para preservar una muestra frente a la descomposición. Por consiguiente, en este proceso, las estructuras se preservan en un estado (tanto química como estructuralmente) tan próximo al estado nativo, por ejemplo en tejido vivo, como sea posible. Esto requiere un fijador químico que pueda estabilizar proteínas y/o ácidos nucleicos del tejido al hacer que sean insolubles. Además de preservar dicho estado, los fijadores se usan para reticular macromoléculas, en particular proteínas y/o ácidos nucleicos, contenidos en la muestra.

Por consiguiente, los fijadores de reticulación actúan creando enlaces químicos covalentes entre macromoléculas, en particular proteínas y/o ácidos nucleicos. A este respecto, un fijador bien conocido es el formaldehído. Se usa preferiblemente como una formalina tamponada neutra al 10% (NBF), es decir, aprox. El 3,7%-4,0% de formaldehído en solución salina tamponada con fosfato. Debido a que el formaldehído es un gas a temperatura ambiente, se usa el gas de formaldehído-formalina en disolución en agua (al ~37% p/v) cuando se prepara el fijador anterior. El paraformaldehído es una forma polimerizada de formaldehído, obtenida generalmente como un polvo blanco fino, que se despolimeriza de vuelta a formalina cuando se calienta. El formaldehído fija el tejido mediante la reticulación de las proteínas, principalmente los residuos del aminoácido básico lisina. Sus efectos son reversibles mediante agua en exceso y evita la pigmentación de formalina. Otros beneficios incluyen: almacenamiento a largo plazo y buena penetración en el tejido. Otro aldehído popular para la fijación es el glutaraldehído. Funciona de manera similar al formaldehído al provocar la deformación de las estructuras de hélice alfa en las proteínas. Sin embargo, el glutaraldehído es una molécula más grande, y así su velocidad de difusión a través de las membranas es más lenta que la del formaldehído. Por consiguiente, la fijación de glutaraldehído en muestras más gruesas puede verse obstaculizada, pero este problema puede superarse reduciendo el tamaño de la muestra. Una de las ventajas de la fijación con glutaraldehído es que puede ofrecer un producto fijado más rígido o ligado estrechamente: su mayor longitud y dos grupos aldehído le permiten "forman un puente" y unir pares más distantes de moléculas de proteínas. Provoca cambios rápidos e irreversibles, se fija rápidamente, es muy adecuado para la microscopía electrónica, se fija bien a 4°C y proporciona el mejor detalle citoplasmático y nuclear global. Sin embargo, no es ideal para la tinción inmunohistoquímica.

Algunos protocolos de fijación exigen una combinación de formaldehído y glutaraldehído para que sus fortalezas respectivas se complementen entre sí.

Estos fijadores de reticulación, especialmente el formaldehído, tienden a preservar la estructura secundaria de las proteínas y también pueden proteger cantidades significativas de la estructura terciaria.

Sin embargo, la fijación también puede realizarse usando medios alternativos, por ejemplo fijación no química usando medios físicos, en particular luz UV tal como se describe, por ejemplo, en Zhang *et al.* (2004) Biochem Biophys Res Commun 322(3), 705-11. Alternativa o adicionalmente, la fijación puede realizarse usando un láser, en particular un láser UV, tal como se describe por ejemplo en Benedetti *et al.* (2014) Methods Mol Biol 1204:24-34.

Por consiguiente, se prefiere que la fijación se realice usando una sustancia química y/o medios físicos. A este respecto, se prefiere que los medios físicos comprendan luz UV o un láser UV. Se prefiere más que la fijación se realice usando una sustancia química, preferiblemente formaldehído o paraformaldehído.

Las reticulaciones introducidas pueden eliminarse de manera posterior a la preparación de biblioteca, es decir, de manera posterior a la adición de transposasa y de manera previa al aislamiento de ácido nucleico de la cromatina. La reversión de las reticulaciones puede realizarse usando métodos bien conocidos en la técnica.

Por ejemplo, las reticulaciones de formaldehído pueden eliminarse calentando la muestra. Preferiblemente, la muestra se calienta hasta aproximadamente 65°C, preferiblemente durante varias horas. En particular, la muestra puede calentarse hasta aproximadamente 65°C durante 4 horas o más, por ejemplo durante la noche. Alternativamente, la muestra puede calentarse hasta aproximadamente 95°C durante aproximadamente 10-15 minutos. Sin embargo, se prefiere calentar hasta menores temperaturas, en particular hasta aproximadamente 65°C, para retener la integridad de la muestra que comprende ácido nucleico. Además del calentamiento, pueden añadirse detergentes y/o sal (por ejemplo, SDS al 0,5-1% y/o NaCl aproximadamente 300 mM) para eliminar las reticulaciones. Además, pueden añadirse ARNasa y/o proteinasa K de manera posterior a la eliminación de las reticulaciones para retirar proteínas y/o ARN, respectivamente, de la muestra que comprende ácido nucleico, en particular ADN. Como ejemplo, las muestras pueden tratarse durante 30 minutos a 37°C con 0,5 µl de ARNasa libre de ADNasa-ARNasa A 10 mg/ml, y posteriormente con 1 µl de proteinasa K 20 mg/ml durante 1-2 horas a 55°C.

En una configuración ultrarrápida de los métodos de la presente invención, la muestra puede calentarse hasta altas temperaturas para revertir las reticulaciones. En particular, la muestra puede calentarse hasta aproximadamente 95°C para revertir las reticulaciones. Tales altas temperaturas reducen significativamente el tiempo requerido para revertir las reticulaciones. En particular, el tiempo requerido para revertir las reticulaciones puede reducirse desde varias horas, tal como aproximadamente 4 horas a aproximadamente 65°C, hasta aproximadamente 10-15 minutos a aproximadamente 95°C. Debido a que la transposasa usada en los métodos de la presente invención comprende preferiblemente oligonucleótidos que incluyen secuencias adaptadoras, tales secuencias adaptadoras pueden integrarse de manera previa a las reticulaciones inversas, ya que la reversión de las reticulaciones se realiza de manera posterior a la adición de transposasa en los métodos de la presente invención. Por consiguiente, no pueden usarse tales altas temperaturas usando protocolos de ChIP convencionales. Esto se debe a que el calentamiento hasta altas temperaturas desnaturaría el ADN de ChIP, y debido a la complejidad del ADN de ChIP, algunos fragmentos (especialmente las secuencias ricas en AT) no vuelven a aparearse correctamente. Cuando se prepara una biblioteca mediante el ligamiento de adaptadores bicatenarios, los fragmentos de ADN de ChIP que no vuelven a aparearse correctamente se excluyen probablemente de la biblioteca final e introducen un sesgo de secuenciación. Sin embargo, en los métodos de la presente invención, pueden emplearse altas temperaturas, como de aproximadamente 95°C, para revertir las reticulaciones. Esto reduce notablemente la duración global del ensayo; véase el ejemplo 14. Además, el uso de altas temperaturas para revertir reticulaciones, como de aproximadamente 95°C, evita la etapa de elución de las perlas. Evitar la elución de las perlas reduce adicionalmente la complejidad del método usado y reduce adicionalmente el tiempo global requerido para poner en práctica los métodos de la invención. Esto se debe a que la elución de las perlas comprende el uso de tampones incompatibles con la etapa de PCR posterior, que usa por ejemplo SDS y/o altas concentraciones de sal. Tales tampones hacen que la amplificación de biblioteca sea difícil o imposible sin una limpieza previa del ADN. Por consiguiente, la configuración ultrarrápida descrita en el presente documento hace que sea innecesaria la purificación de ADN. Cuando los métodos de la presente invención implican el uso de altas temperaturas para revertir reticulaciones, en particular temperaturas de aproximadamente 95°C, los métodos preferiblemente también comprenden una etapa de reparación de extremos de oligonucleótidos introducidos durante la reacción con transposasa de manera previa a la aplicación de altas temperaturas, es decir, una etapa de relleno de secuencias adaptadoras, en particular relleno de secuencias adaptadoras en la hebra inversa opuesta a la hebra que comprende el oligonucleótido introducido durante la reacción con transposasa. Por tanto, los métodos de la presente invención en los que se usan altas temperaturas, como de aproximadamente 95°C, para revertir las reticulaciones, comprenden preferiblemente una etapa de adición de componentes de PCR para la reparación de extremos de manera previa a la aplicación de alta temperatura. Preferiblemente, la reparación de extremos se realiza en las perlas usando PCR MM de manera previa al calentamiento en condiciones de reparación de extremos, por ejemplo 72°C durante 5 min usando una ADN polimerasa, como Taq polimerasa. Sin embargo, la reparación de extremos también puede realizarse usando una mezcla de reparación de extremos a menores temperaturas.

Por consiguiente, la presente invención proporciona un método ultrarrápido para preparar una biblioteca de secuenciación. El método ultrarrápido para preparar una biblioteca de secuenciación comprende añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico, en el que la muestra se ha fijado mediante reticulación; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada unida por dicho agente; rellenar los extremos de oligonucleótidos generados durante la reacción con transposasa; revertir las reticulaciones a altas temperaturas, preferiblemente a aproximadamente 95°C; y obtener una biblioteca de secuenciación. Además, la presente invención se refiere a un método ultrarrápido para el mapeo de interacciones moleculares que implican a un ácido nucleico. El método ultrarrápido para el mapeo de interacciones moleculares comprende añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico; aislar la cromatina unida por dicho agente; añadir transposasa a la cromatina aislada unida por dicho agente; rellenar los extremos oligonucleotídicos generados durante la reacción con transposasa; revertir las reticulaciones a altas temperaturas, preferiblemente a aproximadamente 95°C; amplificar el ácido nucleico; secuenciar el ácido nucleico amplificado; e identificar interacciones moleculares. Ultrarrápido a este respecto significa que los métodos ultrarrápidos de la invención reducen significativamente el tiempo de experimento global esperado para los métodos conocidos. En particular, los métodos ultrarrápidos de la presente invención permiten la preparación de una biblioteca de secuenciación o el mapeo de interacciones moleculares, respectivamente, en menos de un día hábil, es decir menos de aproximadamente 10 horas. El tiempo global requerido para preparar una biblioteca de secuenciación desde la obtención de una muestra, por ejemplo obtención de una muestra de sangre de un donante

hasta la obtención de una biblioteca de secuenciación está en el intervalo de aproximadamente 15 horas.

Los métodos de la invención pueden comprender además una etapa de lisis de células. La lisis se refiere a la descomposición de las membranas celulares. Esto puede lograrse mediante métodos bien conocidos en la técnica. En particular, la lisis puede lograrse mediante medios mecánicos o químicos. Por ejemplo, puede emplearse la perturbación mecánica de las membranas celulares, tal como mediante congelación y descongelación repetidas, sonicación, presión o filtración. Sin embargo, se prefiere que la lisis se logre mediante medios químicos usando, en particular, enzimas o detergentes u otros agentes caotrópicos. Se describen métodos preferidos de lisis celular en Thermo Scientific Pierce Cell Lysis Technical Handbook o Lottspeich, Engels (2012) Bioanalytik, Springer Spektrum. A este respecto, se realiza lisis tal como se usa en los métodos de la invención para aislar ácidos nucleicos de la muestra celular, obteniéndose de ese modo una primera muestra que comprende un ácido nucleico. Dicha primera muestra se usa para análisis adicional usando los métodos de la presente invención, es decir, dicha primera muestra o bien se usa para la preparación adicional de la muestra, en particular usando sonicación, o bien se analiza directamente usando los métodos de la invención para la preparación de una biblioteca de secuenciación o el mapeo de interacciones moleculares que implican a un ácido nucleico.

Por consiguiente, en una realización, de manera posterior a la lisis celular, los métodos de la invención pueden comprender una etapa de sonicación. La sonicación tiene numerosos efectos, tanto químicos como físicos. En aplicaciones biológicas, la sonicación se usa comúnmente para perturbar o desactivar un material biológico. Por ejemplo, la sonicación se usa a menudo para perturbar las membranas celulares y liberar contenido celular. Este proceso se llama sonoporación. La sonicación también se usa para fragmentar moléculas de ácidos nucleicos, en particular ADN, en las que el ácido nucleico, en particular ADN, sometido a breves periodos de sonicación, se corta en fragmentos más pequeños. La sonicación también se usa para fragmentar complejos de moléculas que contienen ácidos nucleicos y proteínas, en particular ácidos nucleicos que contienen cromatina, en particular ADN, en los que los complejos se someten a breves periodos de sonicación en los que el contenido de ácido nucleico en el complejo, en particular ADN, se corta en fragmentos más pequeños. A este respecto, se conoce bien cómo ajustar la intensidad de sonicación para generar fragmentos de ácidos nucleicos, en particular ADN, que tienen longitudes particulares y/o en los que la mayoría de los fragmentos contenidos en una muestra que comprende un ácido nucleico, en particular ADN, tienen longitudes particulares. A este respecto, se prefiere que la muestra que comprende un ácido nucleico, en particular ADN, comprenda fragmentos que tengan una longitud de 200 a 700 pares de bases. Por consiguiente, se prefiere que la sonicación se realice hasta que la mayoría de los fragmentos de ácido nucleico tengan una longitud de 200-700 pares de bases. Se conoce bien cómo ajustar la intensidad y la duración de la sonicación para generar tales fragmentos. Además, se conoce bien cómo determinar la longitud de tales fragmentos para verificar la configuración de sonicación.

A este respecto, la configuración de sonicación puede depender de las condiciones de fijación y la línea celular/tejido/tipo celular/organismo para obtener la muestra de ácido nucleico. Además, la configuración de sonicación puede depender del dispositivo de sonicación usado.

También se prevé usar técnicas alternativas para fragmentar la muestra de ácido nucleico, en particular la muestra que comprende ADN. Por ejemplo, puede usarse digestión enzimática para la fragmentación de ácidos nucleicos comprendidos en la cromatina. Enzimas a modo de ejemplo son fragmentasa (NEB) o MNasa (la nucleasa extracelular de *Staphylococcus aureus*). También pueden usarse agentes químicos u otros métodos físicos además del ultrasonido para fragmentar los ácidos nucleicos comprendidos en la cromatina.

Los resultados de la sonicación pueden verificarse mediante métodos bien conocidos en la técnica. Por ejemplo, para verificar si la mayoría de los fragmentos de ácido nucleico, en particular de ADN, tienen una longitud de 200-700 pares de bases, puede someterse a prueba la longitud de fragmento usando electroforesis en gel de agarosa.

Según lo anterior, los métodos de la invención comprenden como primera etapa, de manera posterior a las etapas preparatorias descritas anteriormente, la adición de un agente que se une a cromatina, en particular un anticuerpo o una sustancia química, a una muestra que comprende un ácido nucleico, en particular un ADN. Se prefiere que la muestra que comprende un ácido nucleico, en particular un ADN, se derive de una célula, tal como se describió anteriormente. De manera posterior a la adición del agente que se une a cromatina, se aísla la cromatina unida por dicho agente. En particular, la cromatina unida por dicho agente se aísla de la cromatina no unida. Al hacerlo así, la cantidad global de cromatina se reduce significativamente, lo que reduce los eventos de fragmentación. El aislamiento de la cromatina puede lograrse mediante diversas técnicas descritas en la técnica. Por ejemplo, el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, puede inmovilizarse sobre superficies mediante interacciones de afinidad. Se prefiere que estas superficies sean partículas (perlas). Sin embargo, también se prevén otras superficies, por ejemplo, columnas. Cuando el agente que se une a cromatina es un anticuerpo, la parte Fc del anticuerpo puede unirse a la superficie de las perlas a través de proteína A, proteína G, proteína L o similar. A este respecto, la proteína A es una proteína de superficie de 42 kDa que se encuentra originariamente en la pared celular de la bacteria *Staphylococcus aureus*. Está codificada por el gen spa y su regulación está controlada por la topología del ADN, la osmolaridad celular y un sistema bicomponente denominado ArlS-ArlR. Se usa comúnmente en la investigación bioquímica debido a su capacidad para unirse a inmunoglobulinas. Alternativamente, los anticuerpos pueden unirse a superficies a través de proteína G, que es una proteína de unión a inmunoglobulina expresada en las bacterias estreptocócicas del grupo C y G muy similar a la

proteína A pero con diferentes especificidades de unión. Es una proteína de superficie celular de 65 kDa (proteína G148 G) y 58 kDa (proteína G C40) usada comúnmente para purificar anticuerpos a través de su unión a la región Fab y Fc. Por consiguiente, el agente que se une a cromatina, en el que el agente es un anticuerpo, puede unirse a perlas a través de proteína A, proteína G, proteína L o similar para aislar la cromatina unida por dicho agente, en particular el anticuerpo, de la cromatina no unida.

En los métodos de la invención, la cromatina también puede aislarse por otros medios, por ejemplo, etiquetas de afinidad unidas al agente que se une a cromatina. Por ejemplo, una etiqueta de afinidad puede incluir biotina o His que puede unirse a estreptavidina o níquel, respectivamente. Otros ejemplos de complejos de etiqueta de afinidad de componentes múltiples incluyen ligandos y sus receptores, por ejemplo, avidina-biotina, estreptavidina-biotina y derivados de biotina, estreptavidina o avidina incluyendo, pero sin limitarse a, 2-iminobiotina, destiobiotina, NeutrAvidin, CaptAvidin y similares; proteínas/péptidos de unión incluyendo maltosa-proteína de unión a maltosa (MBP), calcio-proteína/péptido de unión a calcio (CBP); antígeno-anticuerpo, incluyendo etiquetas epitópicas, y sus anticuerpos anti-epítipo correspondientes; haptenos, por ejemplo, dinitrofenilo y digoxigenina, y sus anticuerpos correspondientes; aptámeros y sus dianas correspondientes; etiquetas de poli-His (por ejemplo, penta-His y hexa-His) y sus parejas de unión, incluidos los materiales de cromatografía de afinidad de iones metálicos inmovilizados (IMAC) correspondientes y anticuerpos anti-poli-His; fluoróforos y anticuerpos anti-fluoróforo; y similares.

Por consiguiente, se prefiere que el agente que se une a cromatina, en el que el agente que se une a cromatina es una sustancia química, se marque con etiqueta de biotina.

Las perlas pueden ser magnéticas, de material a base de agarosa o látex y similares. La cromatina diana inmovilizada puede aislarse entonces mediante aislamiento de las perlas. Esto puede lograrse mediante centrifugación por rotación usando columnas filtrantes que retienen las perlas con el agente que se une a cromatina en el filtro mientras la fracción de cromatina no unida pasa a través del filtro y puede desecharse. En el caso de las perlas magnéticas, se aplica fuerza magnética a las perlas para retenerlas en un recipiente de reacción, mientras que la fracción de cromatina no unida puede desecharse mediante pipeteo, por ejemplo. Dicho agente también puede acoplarse previamente a superficies/perlas antes de la adición a cromatina. El agente también puede reticularse químicamente a superficies cuando se acopla previamente, y no se basa exclusivamente en interacciones de afinidad para aislar la cromatina. Como ejemplo, puede usarse pimelimidato de dimetilo (DMP) para acoplar proteínas a perlas. El aislamiento de cromatina a menudo está soportado por etapas de lavado para eliminar interacciones inespecíficas de cromatina con dicho agente o interacciones inespecíficas de cromatina con el recipiente de reacción o la superficie del reactivo de aislamiento. El lavado de la cromatina aislada por dicho agente o dicha sustancia química que se aísla mediante los procedimientos mencionados anteriormente se logra mediante la adición y posterior retirada de disoluciones acuosas tamponadas que contienen productos químicos incluyendo sal y detergentes. Por consiguiente, los métodos de la invención pueden comprender además etapas de lavado de manera posterior al aislamiento de la cromatina unida por el agente que se une a cromatina.

De manera posterior al aislamiento de la cromatina unida por el agente que se une a cromatina, en particular el anticuerpo o la sustancia química, se añade una transposasa a la cromatina aislada. La transposasa es una enzima que se une al extremo de un transposón y cataliza el movimiento del transposón a otra parte del genoma mediante un mecanismo de cortar y pegar o un mecanismo de transposición replicativa. Las transposasas se clasifican según EC con el número EC 2.7.7. Los genes que codifican para transposasas están muy extendidos en los genomas de la mayoría de los organismos y son los genes más abundantes conocidos. Una transposasa preferida dentro del contexto de la presente invención es la transposasa (Tnp) Tn5. Tn5 es un miembro de la superfamilia de proteínas de ARNasa que incluye integrasas retrovirales. Tn5 puede encontrarse en las bacterias *Shewanella* y *Escherichia*. El transposón codifica para la resistencia a antibióticos a kanamicina y otros antibióticos aminoglicosídicos. Tn5 y otras transposasas son notablemente inactivas. Debido a que los eventos de transposición de ADN son inherentemente mutagénicos, la baja actividad de las transposasas es necesaria para reducir el riesgo de provocar una mutación letal en el huésped, y eliminar así el elemento transponible. Uno de los motivos por los que Tn5 es tan poco reactivo es porque los extremos N y C-terminales están ubicados en una proximidad relativamente estrecha entre sí y tienden a inhibirse mutuamente. Esto se elucidó mediante la caracterización de varias mutaciones que dieron como resultado formas hiperactivas de transposasas. Una de tales mutaciones, L372P, es una mutación del aminoácido 372 en la transposasa Tn5. Este aminoácido es generalmente un residuo de leucina en medio de una hélice alfa. Cuando esta leucina se reemplaza por un residuo de prolina, la hélice alfa se rompe, introduciendo un cambio conformacional en el dominio C-terminal, separándolo del dominio N-terminal lo suficiente como para promover una mayor actividad de la proteína. Por consiguiente, se prefiere que se use una transposasa modificada de este tipo, que tiene una mayor actividad que la transposasa Tn5 que se produce de manera natural. Además, se prefiere particularmente que la transposasa empleada en los métodos de la invención se cargue con oligonucleótidos, que se insertan en el ácido nucleico diana, en particular el ADN diana.

Por ejemplo, una transposasa codificada por la secuencia de ácido nucleico de SEQ ID NO:1 ó 2 o una secuencia de ácido nucleico que tiene una identidad de secuencia del 80, el 85, el 90, el 95, el 96, el 97, el 98 o el 99% con cualquiera de SEQ ID NO:1 ó 2 puede usarse en los métodos de la invención. A este respecto, la transposasa puede producirse usando un vector de expresión que tiene una secuencia de ácido nucleico tal como se muestra en la SEQ ID NO: 3 o usando un vector de expresión que comprende una secuencia que codifica para una transposasa correspondiente a una transposasa codificada por una secuencia de ácido nucleico que tiene una identidad de

secuencia del 80, el 85, el 90, el 95, el 96, el 97, el 98 o el 99% con cualquiera de SEQ ID NO:1 ó 2.

Por consiguiente, se prefiere usar una transposasa Tn5 hiperactiva y un sitio de reconocimiento de transposasa de tipo Tn5 (Goryshin y Reznikoff, J. Biol. Chem., 273:7367 (1998)), o transposasa MuA y un sitio de reconocimiento de transposasa Mu que comprende secuencias de extremo RI y R2 (Mizuuchi, K., Cell, 35: 785, 1983; Savilahti, H, *et al.*, EMBO J., 14: 4893, 1995). Más ejemplos de sistemas de transposición que pueden usarse en los métodos de la presente invención incluyen Tn552 de *Staphylococcus aureus* (Colegio *et al.*, J. Bacteriol, 183: 2384-8, 2001 ; Kirby C *et al.*, Mol. Microbiol, 43: 173-86, 2002), Tyl (Devine & Boeke, Nucleic Acids Res., 22: 3765-72, 1994 y la publicación internacional WO 95/23875), el transposón Tn7 (Craig, N L, Science. 271:1512, 1996; Craig, N L, revisión en: Curr Top Microbiol Immunol, 204: 27-48, 1996), Tn/O e IS 10 (Kleckner N, *et al.*, Curr Top Microbiol Immunol, 204: 49-82, 1996), la transposasa Mariner (Lampe D J, *et al.*, EMBO J., 15: 5470-9, 1996), Tel (Plasterk R H, Curr. Topics Microbiol. Immunol, 204:125-43, 1996), el elemento P (Gloor, G B, Methods Mol. Biol, 260: 97-1 14, 2004), Tn3 (Ichikawa y Ohtsubo, J Biol. Chem 265:18829-32, 1990), secuencias de inserción bacterianas (Ohtsubo y Sekine, Curr. Top. Microbiol Immunol 204:1-26, 1996), retrovirus (Brown, *et al.*, Proc Natl Acad Sci USA, 86: 2525-9, 1989) y el retrotransposón de levadura (Boeke y Corces, Annu Rev Microbiol. 43: 403-34, 1989). Más ejemplos incluyen IS5, TnIO, Tn903, IS91 1 y versiones modificadas por ingeniería de enzimas de la familia de transposasas (Zhang *et al.*, (2009) PLoS Genet. 5:e1000689. Publicación electrónica del 16 de octubre de 2009; Wilson C. *et al.* (2007) J. Microbiol. Methods 71: 332-5) y los descritos en las patentes estadounidenses n.ºs 5.925.545; 5.965.443; 6.437.109; 6.159.736; 6.406.896; 7.083.980; 7.316.903; 7.608.434; 6.294.385; 7.067.644, 7.527.966; y la publicación de patente internacional n.º WO20121 03545.

Aunque puede usarse cualquier tampón adecuado para la transposasa usada en los métodos de la presente invención, se prefiere usar un tampón particularmente adecuado para la reacción enzimática eficiente de la transposasa usada. A este respecto, se prefiere particularmente un tampón que comprende dimetilformamida para su uso en los métodos de la presente invención, en particular durante la reacción con transposasa. Además, pueden usarse tampones que comprenden sistemas de tamponamiento alternativos incluyendo TAPS, Tris-acetato o sistemas similares. Además, los reactivos de aglomeración como el polietilenglicol (PEG) son particularmente útiles para aumentar la eficiencia de tagmentación de cantidades muy bajas de ADN. Se describen condiciones particularmente útiles para la reacción de tagmentación por Picelli *et al.* (2014) Genome Res. 24: 2033-2040.

La enzima transposasa cataliza la inserción de un ácido nucleico, en particular un ADN en un ácido nucleico diana, en particular ADN diana. El ácido nucleico diana, en particular el ADN diana, para la inserción está comprendido en la cromatina aislada unida por el agente que se une a cromatina, en particular el anticuerpo o la sustancia química. La transposasa usada en los métodos de la presente invención se carga con oligonucleótidos, que se insertan en el ácido nucleico diana, en particular el ADN diana. El complejo de transposasa y oligonucleótido también se denomina transposoma. Preferiblemente, el transposoma es un heterodímero que comprende dos oligonucleótidos diferentes para la integración. A este respecto, los oligonucleótidos que se cargan en la transposasa comprenden múltiples secuencias. En particular, los oligonucleótidos comprenden, al menos, una primera secuencia y una segunda secuencia. La primera secuencia es necesaria para cargar el oligonucleótido en la transposasa. Se proporcionan secuencias a modo de ejemplo para cargar el oligonucleótido en la transposasa en el documento US 2010/0120098. La segunda secuencia comprende una secuencia ligadora necesaria para la unión de cebador durante la amplificación, en particular durante la amplificación por PCR. Por consiguiente, el oligonucleótido que comprende las secuencias primera y segunda se inserta en el ácido nucleico diana, en particular el ADN diana, por la enzima transposasa. El oligonucleótido puede comprender además secuencias que comprenden secuencias de código de barras. Las secuencias de código de barras pueden ser secuencias aleatorias o secuencias definidas. A este respecto, el término "secuencia aleatoria" según la invención debe entenderse como una secuencia de nucleótidos, en la que cada posición tiene una probabilidad independiente e igual de ser cualquier nucleótido. Los nucleótidos aleatorios pueden ser cualquiera de los nucleótidos, por ejemplo G, A, C, T, U, o análogos químicos de los mismos, en cualquier orden, en los que: G se entiende que representa nucleótidos guanílicos, A nucleótidos adenílicos, T nucleótidos timidílicos, C nucleótidos citidílicos y nucleótidos uracílicos U. El experto apreciará que los métodos de síntesis de oligonucleótidos conocidos pueden conducir inherentemente a una representación desigual de los nucleótidos G, A, C, T o U. Por ejemplo, la síntesis puede conducir a una representación excesiva de nucleótidos, tales como G en secuencias de ADN aleatorizadas. Esto puede conducir a un número reducido de secuencias aleatorias únicas tal como se espera basándose en una representación equitativa de nucleótidos. El oligonucleótido para la inserción en el ácido nucleico diana, en particular ADN, puede comprender además adaptadores de secuenciación, por ejemplo adaptadores adecuados para la secuenciación en nanoporos o secuenciación 454 de Roche. Además, el oligonucleótido puede comprender secuencias de etiqueta de biotina. Se prefiere que el oligonucleótido cargado en la transposasa comprenda dichas secuencias primera y segunda y una secuencia de código de barras para la indexación. La integración de secuencias de código de barras durante la reacción con transposasa permite la identificación única de cada fragmento de ácido nucleico, en particular fragmento de ADN, durante el análisis de secuenciación y/o mapeo de interacciones moleculares.

El experto en la técnica es consciente de que el tiempo requerido para que la transposasa usada integre eficientemente un ácido nucleico, en particular un ADN, en un ácido nucleico diana, en particular ADN diana, puede variar dependiendo de diversos parámetros, como componentes del tampón, temperatura y similares. Por consiguiente, el experto en la técnica es consciente de que pueden someterse a prueba/aplicarse diversos tiempos de incubación antes de que se encuentre un tiempo de incubación óptimo. Óptimo a este respecto se refiere al

tiempo óptimo teniendo en cuenta la eficiencia de integración y/o el tiempo requerido para realizar los métodos de la invención. Aunque los tiempos de incubación variables no están correlacionados necesariamente con la integración eficiente de dicho ácido nucleico, en particular dicho ADN, en dicho ácido nucleico diana, en particular el ADN diana, se prefiere usar tiempos de incubación de menos de 10 minutos, menos de 5 minutos, preferiblemente, menos de 2 minutos. Lo más preferido es emplear un tiempo de reacción de 1 minuto para la reacción de tagmentación. Además, pueden optimizarse parámetros como la temperatura y el volumen. A este respecto, la temperatura de incubación recomendada para la transposasa Tn5 es de aproximadamente 37°C. Por tanto, se prefiere que los métodos de la invención comprendan una etapa de adición de transposasa y posteriormente incubación para la tagmentación a aproximadamente 37°C, preferiblemente durante aproximadamente 1 minuto. Sin embargo, también pueden emplearse temperaturas de reacción alternativas, aunque se prefiere que se usen temperaturas superiores a aproximadamente 16°C e inferiores a aproximadamente 55°C para mantener la integridad de la muestra y la eficiencia de la transposasa.

De manera posterior a la adición de transposasa a la cromatina aislada, los ácidos nucleicos, en particular el ADN, se aíslan de la muestra que comprende cromatina, es decir los ácidos nucleicos, en particular el ADN, se separan de los componentes restantes de la cromatina. Esto puede lograrse mediante diversas técnicas conocidas en la técnica. Por ejemplo, los ácidos nucleicos, en particular el ADN, pueden purificarse usando purificación en columna, extracción con fenol-cloroformo seguida de precipitación con etanol, inmovilización reversible en fase sólida y Chelex® 100 y otras técnicas conocidas en la técnica. La purificación en columna se basa en la unión de ácidos nucleicos, en particular ADN (adsorción) a la fase sólida (sílice u otro) dependiendo del pH y el contenido de sal del tampón usado. Después de la centrifugación de la muestra, las proteínas desnaturalizadas permanecen en la fase orgánica, mientras que la fase acuosa que contiene ácido nucleico, en particular ADN, se mezcla con cloroformo, retirando los residuos fenólicos de la disolución. Para aislar el ADN de la fase acuosa, el fenol-cloroformo va seguido por precipitación con etanol o isopropanol. Dado que el ADN es insoluble en estos alcoholes, se agregará produciendo un sedimento tras centrifugación. La precipitación del ADN se mejora al aumentar la fuerza iónica, generalmente añadiendo acetato de sodio. Chelex® 100 es un material quelante distribuido por Bio-Rad, que se usa para purificar otros compuestos mediante intercambio iónico. También puede usarse para purificar el ADN. Las perlas SPRI (inmovilización reversible en fase sólida, por sus siglas en inglés) son paramagnéticas (magnéticas sólo en un campo magnético). Cada perla está compuesta por poliestireno rodeado por una capa de magnetita, que está recubierta con moléculas de carboxilo. Son estos los que se unen de manera reversible al ADN en presencia del "agente de aglomeración" polietilenglicol (PEG) y sal (comúnmente PEG al 20%, NaCl 2,5 M). El PEG hace que el ADN cargado negativamente se una con los grupos carboxilo en la superficie de las perlas. Como la inmovilización depende de la concentración de PEG y sal en la reacción, la razón volumétrica de perlas con respecto a ADN es crítica. La purificación de ADN a menudo está soportada por la retirada de ARN y proteínas mediante la adición de ARNasa y proteinasa a la disolución.

De manera posterior al aislamiento de los ácidos nucleicos, puede obtenerse una biblioteca de secuenciación tal como se describió anteriormente. En particular, una biblioteca de ácidos nucleicos, en particular ADN, compatible para la secuenciación comprende fragmentos de ácidos nucleicos, en particular ADN, que comprenden secuencias adaptadoras que son necesarias para la secuenciación. Por consiguiente, los métodos de la invención para preparar una biblioteca de secuenciación pueden comprender además una etapa de amplificación para integrar dichas secuencias adaptadoras. La amplificación se realiza tal como se describe a continuación. Las secuencias adaptadoras varían dependiendo del método de secuenciación usado de manera posterior a la preparación de biblioteca de secuenciación. Por ejemplo, cuando se usa la secuenciación en Illumina, los extremos i5 e i7 pueden unirse a los fragmentos de ácido nucleico. Esto también puede lograrse mediante la reacción con transposasa en la que los oligonucleótidos cargados en la enzima transposasa comprenden secuencias adaptadoras compatibles con la secuenciación.

Cuando los métodos de la invención son para el mapeo de interacciones moleculares que implican a un ácido nucleico, el ácido nucleico se amplifica de manera posterior al aislamiento. La amplificación puede lograrse mediante diversas técnicas conocidas en la técnica. La técnica más conocida para la amplificación de ácidos nucleicos, en particular ADN, es la reacción en cadena de la polimerasa (PCR), en la que una muestra se pone en contacto con un par de cebadores oligonucleotídicos en condiciones que permiten la hibridación de los cebadores a un molde de ácido nucleico en la muestra. Los cebadores se extienden en condiciones adecuadas, se disocian del molde, se vuelven a aparearse, se extienden y se disocian para amplificar el número de copias del ácido nucleico. Este ciclo puede repetirse. El producto de amplificación puede caracterizarse mediante técnicas tales como electroforesis, patrones de escisión de endonucleasas de restricción, hibridación o ligamiento de oligonucleótidos y/o secuenciación de ácidos nucleicos.

Los cebadores adecuados para su uso en los métodos de la invención comprenden secuencias hibridables con la segunda secuencia comprendida en los oligonucleótidos comprendidos en los transposomas usados en los métodos de la invención. Además, los cebadores pueden comprender secuencias necesarias para la secuenciación. Se prefiere que en los métodos de la invención se usen cebadores específicos que sean compatibles con el método de secuenciación usado posteriormente. A este respecto, la secuenciación en Illumina, como método preferido de secuenciación usado en los métodos de la invención, es compatible con cebadores que introducen extremos de la celda de flujo, que pueden hibridarse con la celda de flujo necesaria en la amplificación de agrupaciones. A este respecto, los cebadores pueden introducir los extremos i5 e i7 para la secuenciación en Illumina. Además, los

cebadores pueden introducir códigos de barras para la multiplexación. En particular, los códigos de barras comprendidos en las secuencias de cebador pueden usarse como identificadores moleculares únicos para discriminar entre duplicados de PCR y/o como códigos de barras definidos para combinar múltiples experimentos en una serie de secuenciación.

- 5 En los métodos de la invención para el mapeo de interacciones moleculares que implican a un ácido nucleico, en particular ADN, se secuencia el ADN amplificado. Existen diversos métodos de secuenciación conocidos en la técnica. En general, la secuenciación puede realizarse usando pirosecuenciación sobre un soporte sólido (secuenciación 454, Roche), secuenciación por síntesis con terminaciones reversibles (analyzer del genoma de ILLUMINA®) o tecnología de nanoporos (por ejemplo, MinION™ de Oxford Nanopore Technologies). En algunas realizaciones, los fragmentos sometidos a tagmentación aislados se analizan, por ejemplo determinando la
- 10 secuencia de nucleótidos. En algunos ejemplos, la secuencia de nucleótidos se determina usando técnicas de secuenciación o hibridación con o sin amplificación.

- Partiendo de la información de secuencia obtenida secuenciando el ácido nucleico, pueden identificarse interacciones moleculares usando herramientas conocidas en la técnica. Por ejemplo, pueden analizarse datos usando un software de comparación de secuencias que alinea los ácidos nucleicos secuenciados con secuencias genómicas. Las secuencias genómicas se conocen generalmente y pueden obtenerse de fuentes de datos de libre acceso. Una coincidencia de un ácido nucleico secuenciado, que se encuentra en la muestra que va a analizarse, y una secuencia genómica puede usarse como indicador de que dicho ácido nucleico secuenciado se une por una macromolécula, por ejemplo una histona o factor de transcripción, que es reconocido por el agente que se une a cromatina en los métodos de la invención.
- 15
- 20

Cuando el agente que se une a cromatina es una sustancia química, por ejemplo un fármaco, la coincidencia con una secuencia genómica comprendida en el fragmento de ácido nucleico que va a analizarse indica la unión de la sustancia química, por ejemplo fármaco, a una región de ácido nucleico particular comprendida en la cromatina.

- Basándose en la coincidencia de los ácidos nucleicos secuenciados con secuencias genómicas, pueden usarse métodos computacionales estadísticos para determinar regiones de unión significativa para distinguirlas de la “señal de fondo” inespecífica. Las regiones identificadas pueden usarse para inferir adicionalmente su papel biológico al correlacionarlas con otros conjuntos de datos, incluyendo la expresión génica, la anotación genómica, la ontología génica u otros conjuntos de datos biológicos de sistemas.
- 25

- Al acumular regiones derivadas de dicho ácido nucleico unido por el agente que se une a cromatina, en particular la sustancia química, por ejemplo un fármaco, también pueden usarse métodos computacionales para determinar características de secuencia significativas de dichas regiones. Tales enfoques pueden usarse para hallar un enriquecimiento para motivos de unión a ADN específicos que se sabe que se unen por un factor de transcripción específico.
- 30

- Los métodos de la invención también pueden usarse para identificar regiones comprendidas en el ácido nucleico diana, en particular ADN, que son inaccesibles para la enzima transposasa. En particular, cuando se preparan bibliotecas de secuenciación usando los métodos de la invención, pueden generarse fragmentos de biblioteca de secuenciación mediante la introducción de oligonucleótidos compatibles con la secuenciación por una transposasa en la cromatina diana. En la cromatina que comprende ácidos nucleicos y proteínas, las proteínas comprendidas en la cromatina pueden intervenir en cierta medida en la integración del adaptador en el ácido nucleico diana, en particular el ADN, en los sitios de interacciones ADN-proteína, sin interferir en la preparación de bibliotecas de secuenciación. Dichas regiones de ácido nucleico protegidas frente a la inserción de transposasa pueden identificarse mediante métodos computacionales. Dichas regiones son “huellas” de proteínas comprendidas en la cromatina, lo que revela por tanto interacciones de alta resolución.
- 35
- 40

- Aunque los métodos de la presente invención pueden llevarse a cabo en un tubo, se prefiere incluir transferencias de tubo durante la reacción. En particular, el tubo de reacción puede cambiarse de manera previa a la adición de transposasa. Adicional o alternativamente, el tubo puede cambiarse posteriormente a la adición de transposasa después de que la reacción con transposasa se termine. Esto último disminuye la tagmentación de fragmentos de cromatina inespecíficos que se adhieren a las paredes del tubo. Por consiguiente, se prefiere que los métodos de la presente invención comprendan al menos una transferencia de tubo, preferiblemente dos, en los que la primera transferencia de tubo se lleva a cabo de manera posterior a la reacción con transposasa, de manera previa al aislamiento de ácidos nucleicos de la cromatina, o revertir reticulaciones cuando se emplea el protocolo ultrarrápido, y en el que la segunda transferencia de tubo, si se emplea una segunda transferencia de tubo, se lleva a cabo de manera previa a la adición de transposasa.
- 45
- 50

- La presente invención también se refiere a kits, en particular kits de investigación. Los kits de la presente invención comprenden uno o más agente(s) que se une(n) a cromatina, como una o más sustancia(s) química(s) o uno o más anticuerpo(s) y transposasa. Los kits de la invención pueden comprender una transposasa hiperactiva, preferiblemente también cargada de oligonucleótidos. En una realización particular, los kits de la invención comprenden una transposasa codificada por la secuencia de ácido nucleico de SEQ ID NO:1 ó 2 o un vector de expresión que tiene una secuencia de ácido nucleico de SEQ ID NO:3. Los kits de la invención también pueden
- 55

comprender la enzima transposasa en una forma lista para su uso. Los kits de la invención pueden usarse en el diagnóstico de estados médicos como enfermedades. Dichos estados médicos, como enfermedades, pueden ser cualquier estado/enfermedad que implique la interacción del ADN con componentes adicionales como, por ejemplo, pero sin limitarse a, factores de transcripción/histonas y similares. Por ejemplo, las enfermedades que se sabe que están relacionadas con la interacción del ADN con factores de transcripción/histonas incluyen, pero no se limitan a, enfermedades proliferativas, como por ejemplo cáncer. Por consiguiente, los kits de la presente invención pueden usarse para diagnosticar enfermedades incluyendo, pero sin limitarse a, leucemia linfoblástica aguda de células T y similares, leucemia mieloide aguda, sarcoma de Ewing, leucemia promielocítica aguda, leucemia linfoblástica aguda, linfoma difuso de células B grandes, carcinoma de células de transición, cáncer colorrectal, cáncer de páncreas, cáncer de mama, síndrome mielodisplásico, carcinoma de línea media, cáncer de tiroides papilar, carcinoma renal, meduloblastoma, mieloma múltiple, síndrome mielodisplásico, cáncer de esófago, cáncer de ovario, cáncer de próstata, cáncer de pulmón, cáncer rabdoide, carcinoma hepatocelular, schwannomatosis familiar, condrosarcoma, sarcoma epitelioide, meningioma, cordoma, sarcoma indiferenciado, enfermedad de Parkinson, enfermedades de Huntington, distrofia miotónica congénita, artritis reumatoide, lupus eritematoso sistémico, diabetes tipo 1, síndrome de inmunodeficiencia, inestabilidad de la región centromérica y anomalías faciales y síndrome ATR-X entre otros.

Además, los kits pueden usarse para evaluar/determinar la interacción de un compuesto químico con el ADN. A este respecto, los kits de la presente invención pueden usarse para evaluar/determinar la probabilidad de respuesta de un individuo, como un paciente, al tratamiento con un compuesto químico que interacciona con el ADN. El tratamiento puede implicar el tratamiento de diversas enfermedades/estados usando un compuesto químico que se sabe que es eficaz, o cuando va a someterse a prueba la eficacia, en el que el estado médico que va a tratarse puede ser cualquier enfermedad/estado que implique la interacción del ADN. Por consiguiente, dichos estados médicos, como enfermedades, pueden ser cualquier estado/enfermedad que implique la interacción del ADN con componentes adicionales como, por ejemplo pero sin limitarse a, factores de transcripción/histonas y similares. Por ejemplo, las enfermedades que se sabe que están relacionadas con la interacción del ADN con factores de transcripción/histonas incluyen, pero no se limitan a, enfermedades proliferativas, como por ejemplo cáncer. Por consiguiente, los kits de la presente invención pueden usarse para diagnosticar enfermedades incluyendo, pero sin limitarse a, leucemia linfoblástica aguda de células T y similares, leucemia mieloide aguda, sarcoma de Ewing, leucemia promielocítica aguda, leucemia linfoblástica aguda, linfoma difuso de células B grandes, carcinoma de células de transición, cáncer colorrectal, cáncer de páncreas, cáncer de mama, síndrome mielodisplásico, carcinoma de línea media, cáncer de tiroides papilar, carcinoma renal, meduloblastoma, mieloma múltiple, síndrome mielodisplásico, cáncer de esófago, cáncer de ovario, cáncer de próstata, cáncer de pulmón, cáncer rabdoide, carcinoma hepatocelular, schwannomatosis familiar, condrosarcoma, sarcoma epitelioide, meningioma, cordoma, sarcoma indiferenciado, enfermedad de Parkinson, enfermedades de Huntington, distrofia miotónica congénita, artritis reumatoide, lupus eritematoso sistémico, diabetes tipo 1, síndrome de inmunodeficiencia, inestabilidad de la región centromérica y anomalías faciales y síndrome ATR-X entre otros.

En una realización particularmente preferida de la presente invención, los kits (que han de prepararse en contexto) de esta invención o los métodos y usos de la invención pueden comprender además o estar dotados de (un) manual(es) de instrucciones. Por ejemplo, dicho(s) manual(es) de instrucciones puede(n) guiar al experto (sobre cómo va) a emplear el kit de la invención en los usos de diagnóstico proporcionados en el presente documento y según la presente invención. En particular, dicho(s) manual(es) de instrucciones puede(n) comprender orientación para usar o aplicar los métodos o usos proporcionados en el presente documento.

El kit (que ha de prepararse en contexto) de esta invención puede comprender además sustancias/productos químicos y/o equipos adecuados/necesarios para llevar a cabo los métodos y usos de esta invención. Por ejemplo, tales sustancias/productos químicos y/o equipos son disolventes, diluyentes y/o tampones para estabilizar y/o almacenar (un) compuesto(s) requerido(s) para los usos proporcionados en el presente documento, como estabilizar y/o almacenar el/los agente(s) químico(s) y/o transposasa comprendidos en los kits de la presente invención.

A menos que se definan de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el entendido comúnmente por un experto en la técnica a la que pertenece esta invención. Aunque pueden usarse métodos y materiales similares o equivalentes a los descritos en el presente documento en la práctica o prueba de la presente invención, a continuación se describen materiales y métodos adecuados. En caso de conflicto, prevalecerá la presente memoria descriptiva, incluidas las definiciones. Además, los materiales, métodos y ejemplos son sólo ilustrativos y no pretenden ser limitativos.

Los métodos y las técnicas de la presente invención se realizan generalmente según métodos convencionales bien conocidos en la técnica y tal como se describe en diversas referencias generales y más específicas que se citan y comentan a lo largo de toda la presente memoria descriptiva a menos que se indique de otro modo. Véanse, por ejemplo, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York (1989).) y Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992) y Harlow and Lane Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1990)).

La invención también cubre todas las características adicionales mostradas en las figuras individualmente, aunque pueden no haberse descrito en la descripción anterior o siguiente. Además, las alternativas individuales de las

realizaciones descritas en las figuras y la descripción y las alternativas individuales de características de las mismas pueden excluirse del contenido del otro aspecto de la invención.

Además, en las reivindicaciones, la expresión “que comprende” no excluye otros elementos o etapas, y el artículo indefinido “un(o)” o “una” no excluye una pluralidad. Una única unidad puede cumplir las funciones de varias características citadas en las reivindicaciones. Los términos “esencialmente”, “aproximadamente”, “de manera aproximada” y similares en relación con un atributo o un valor en particular también definen exactamente el atributo o exactamente el valor, respectivamente. Cualquier símbolo de referencia en las reivindicaciones no debe interpretarse como limitativo del alcance.

La presente invención también se ilustra mediante las siguientes figuras.

10 Figura 1 Visión general de los métodos de la técnica anterior y sus principales inconvenientes

Figura 2 Visión general esquemática de ChIP-seq convencional, ChIP-tagmentación y ChIPmentación

15 Visión general esquemática de ChIPmentación en comparación con la preparación de ChIP-seq convencional y biblioteca mediante la tagmentación de ADN de ChIP purificado (ChIP-tagmentación). Los tres protocolos comienzan con la fijación de células con formaldehído, lisis celular, sonicación de cromatina e inmunoprecipitación con un anticuerpo específico unido a perlas (“inmunoprecipitación de cromatina”). Para ChIP-seq convencional (izquierda), la reticulación inversa va seguida por la purificación del ADN de ChIP, que luego se somete a la preparación de biblioteca en un procedimiento de múltiples etapas que comprende reparación de extremos, purificación, unión de colas A, ligamiento de adaptador y selección por tamaño. La ChIP-tagmentación (centro) usa ADN de ChIP purificado para la preparación de biblioteca basada en tagmentación, lo que tiene la desventaja de que el protocolo es sensible a concentraciones de ADN variables. En ChIPmentación (derecha), los adaptadores de secuenciación se introducen en una única etapa durante la inmunoprecipitación usando la tagmentación con transposasa Tn5 cargada en adaptador.

Figura 3 Figura 3 complementaria: Efecto de la concentración de la enzima Tn5 sobre las distribuciones de tamaño de biblioteca de ChIPmentación

25 Distribución de tamaño de fragmentos de ADN de las bibliotecas de ChIPmentación para H3K4me3 que se prepararon con diferentes cantidades de transposasa Tn5 (de 0,2 µl a 5 µl de enzima del kit de preparación de biblioteca de ADN Nextera de Illumina). Los tamaños de los fragmentos después de la reticulación inversa pero antes del enriquecimiento de biblioteca se muestran en rojo, los tamaños de los fragmentos después de la PCR de enriquecimiento se muestran en verde, y las bibliotecas finales seleccionadas por tamaño se muestran en azul.

30 Figura 4 ChIPmentación: ChIP-seq rápida, barata, de baja entrada para marcas de histona y factores de transcripción

(a) Visión general esquemática de ChIPmentación (véase la figura 2 para una comparación de ChIP-seq convencional, ChIP-tagmentación partiendo de ADN de ChIP purificado, y ChIPmentación).

35 (b) Distribución de tamaño de las longitudes de inserto mapeadas a partir de la secuenciación de extremos apareados de bibliotecas de ChIPmentación de H3K4me3 obtenidas con diferentes concentraciones de enzima Tn5.

(c) Porcentajes de lecturas mapeadas (parte superior) y únicas (parte inferior) para bibliotecas de ChIPmentación de H3K4me3 obtenidas con diferentes concentraciones de enzima Tn5.

40 (d) Captura de pantalla del navegador genómico que compara las bibliotecas de ChIPmentación de H3K4me3 obtenidas con diferentes concentraciones de enzima Tn5.

(e) Mapa de calor de correlación en todo el genoma (ventanas de 1.000 pb) para la ChIPmentación de H3K4me3 obtenida con diferentes concentraciones de enzima Tn5.

45 (f) Captura de pantalla del navegador genómico que muestra datos de ChIP-seq (“ChIP”) y ChIPmentación (“CM”) para cinco modificaciones de histonas y cuatro factores de transcripción a partir de diferentes cantidades de entrada. Se combinaron los datos de dos réplicas biológicas.

(g) Mapa de calor de correlación en todo el genoma (ventanas de 1000 pb) para datos de ChIP-seq convencional y ChIPmentación para diferentes marcas de histona y diferentes cantidades de entrada de células.

50 (h) Correlación en todo el genoma (ventanas de 1000 pb) de datos de ChIP-seq convencional y ChIPmentación para diferentes factores de transcripción y diferentes cantidades de entrada de células (alta: 10M células; baja: 100k o 500k células).

(i) Comparación del tiempo de preparación para ChIP-seq convencional (parte superior), kits de preparación

de bibliotecas de baja entrada disponibles comercialmente (centro) y ChIPmentación (parte inferior). El tiempo de preparación se calculó desde el comienzo del protocolo hasta (pero excluyendo) la reacción PCR de amplificación de biblioteca final.

Figura 5 Comparación global de datos de ChIP-seq convencional y ChIPmentación

- 5 (a) Gráfico compuesto para la distribución de marcas de histona a lo largo de todos los genes, que se muestra por separado para ChIPmentación (izquierda) y ChIP-seq convencional (derecha).
- (b) Fracción de lecturas en picos (FRiP) y número de picos asignados a partir de los datos de ChIPmentación (panel superior) y ChIP-seq (panel inferior) para todas las bibliotecas secuenciadas. Obsérvese que la profundidad de secuenciación varía entre las réplicas (figura 8).

Figura 6 El navegador genómico sigue la pista de datos de ChIPmentación de baja entrada

Captura de pantalla del navegador genómico que muestra datos de ChIPmentación ("CM") para réplicas biológicas individuales y diferentes cantidades de entrada de células (es decir, células 10M, 500k, 100k y 10k). La ChIP-seq convencional ("ChIP") obtenida a partir de datos de 10 millones de células se incluye como referencia.

Figura 7 Solapamiento de picos entre los experimentos de ChIP-seq convencional y ChIPmentación

- 15 El solapamiento de picos calculado como el porcentaje del X% superior de los picos en una réplica/método/cantidad de entrada que se solapan con un pico significativo en la otra réplica/método/cantidad de entrada.

Figura 8 Resumen de secuencia para 24 bibliotecas de ChIP-seq convencionales, 52 bibliotecas de ChIPmentación y 9 bibliotecas de ChIP-tagmentación

Figura 9 Patrones de alta resolución en datos de ChIPmentación

- 20 (a) Intensidad de señal (frecuencias de inserción de Tn5) para datos de ChIPmentación de CTCF, GATA1, PU.1 y REST alrededor de motivos del factor de transcripción respectivo bajo los picos asignados. Los paneles superiores muestran la señal sin procesar de ChIPmentación, ATAC-seq y ADNase-seq, mientras que los paneles inferiores muestran el fondo definido por la tagmentación de ADN genómico, así como las intensidades de señal de ChIPmentación y ATAC-seq normalizadas. Se realizó la normalización teniendo una señal sobre e para la puntuación Z de la señal de fondo para cada pico. Con fines de visualización, la señal normalizada se promedió a lo largo de todos los picos, se alisó con una ventana Hanning de 20 pb y se transformó la puntuación Z para comparación.
- 25 (b) Frecuencia de distancias por parejas entre eventos de inserción en los datos de ChIPmentación para H3K4me3. La frecuencia de oscilación periódica de 10 pb puede vincularse con la naturaleza rotacional del ADN alrededor de los nucleosomas¹.
- 30 (c) Intensidad de señal (frecuencias de inserción) para los datos de ChIPmentación de H3K4me1 alrededor de nucleosomas posicionados usando el software NucleoATAC y los datos de ATAC-seq para células GM12878. Obsérvese el patrón estructurado con inserciones más altas y periódicas en los bordes de nucleosoma.

Figura 10 Comparación de los costes de reactivos para ChIP-seq convencional y ChIPmentación

- 35 Comparación de los costes de reactivos para ChIP-seq convencional (parte superior), kits de preparación de bibliotecas de baja entrada disponibles comercialmente (centro) y ChIPmentación (parte inferior). Se calcularon estimaciones de costes para la preparación de biblioteca, incluyendo la amplificación e indexación, pero excluyendo los reactivos para la selección por tamaño, las purificaciones de reacción y la etapa de control de calidad final antes de la secuenciación.
- 40

Figura 11 Preparación de biblioteca mediante ChIP-tagmentación a partir de ADN de ChIP purificado

- (a) Captura de pantalla representativa del navegador genómico de UCSC de perfiles de ChIP-tagmentación para H3K4me3 en células mononucleares de sangre periférica (PBMC) usando diferentes cantidades de ADN de ChIP purificado como material de partida.
- 45 (b) Diagramas de dispersión por parejas que comparan ChIP-seq convencional (obtenida a partir de 10 millones de células) y la ChIP-tagmentación para H3K4me3 en células mononucleares de sangre periférica (PBMC) usando diferentes cantidades de ADN purificado como material de partida.

Figura 12 PU.1 lee los recuentos de diferentes experimentos en sitios de unión a PU.1 usando los métodos de la invención. Las líneas discontinuas y de puntos presentan visualmente datos de experimentos usando 500k o 10 millones de células, respectivamente. La línea de trazos muestra la señal de PU.1 derivada de un experimento usando 500k células y la configuración optimizada a modo de

50

ejemplo descrita en el ejemplo 13. El protocolo optimizado proporciona una mayor relación señal/ruido que un experimento con 10 millones de células usando el protocolo convencional.

Figura 13 H3K27ac lee los recuentos de diferentes experimentos que usan los métodos de la invención de una manera ultra rápida al comienzo de la transcripción anotada. Las líneas de puntos y discontinuas corresponden al protocolo ultrarrápido descrito en el ejemplo 14, mientras que la línea recta presenta visualmente el experimento que usa el protocolo convencional. El protocolo ultrarrápido (que también usa el protocolo optimizado del ejemplo 13) proporciona relaciones de señal/ruido iguales o mejores que un experimento con 10 millones de células que usa el protocolo convencional.

Figura 14 Generación de bibliotecas listas para secuenciación en un único día. Protocolo ultrarrápido que usa 500k células K562 que permite la generación de bibliotecas listas para secuenciación para marcas de histona y factores de transcripción en un único día. La pista superior a la derecha muestra señales densas para H3K27ac correspondientes a superpotenciadores en el locus de globina en células K562.

Figura 15 Comparación de las preparaciones de biblioteca de secuenciación obtenidas usando la transposasa Tn5 y una transposasa preparada internamente. Se sometieron células de una línea celular de leucemia a los métodos de la invención usando un anticuerpo contra H3K4me3 y o bien una transposasa Tn5 disponible comercialmente o bien una transposición interna que tiene una secuencia codificada por la secuencia de ácido nucleico tal como se muestra en las SEQ ID NO:1 y 2. Tal como puede observarse, los resultados no dependen de la transposasa usada.

La presente invención se describe adicionalmente por medio de los siguientes ejemplos ilustrativos no limitativos que proporcionan una mejor comprensión de la presente invención y de sus muchas ventajas. Los siguientes ejemplos se incluyen para demostrar realizaciones preferidas de la invención. Los expertos en la técnica deben apreciar que las técnicas dadas a conocer en los ejemplos que siguen representan técnicas usadas en la presente invención para funcionar bien en la práctica de la invención y, por tanto, pueden considerarse modos preferidos para su práctica.

A menos que se indique de otro modo, se usaron los métodos establecidos de tecnología de genes recombinantes tal como se describe, por ejemplo, en Sambrook, Russell "Molecular Cloning, A Laboratory Manual", Cold Spring Harbor Laboratory, Nueva York (2001)).

Ejemplo 1 Protocolo de ChIPmentación

Recogida de células y fijación

Se recogieron células, se lavaron una vez con PBS y se fijaron con paraformaldehído al 1% en hasta 1,5 ml de PBS durante 10 minutos a temperatura ambiente. Se añadió glicina hasta una cantidad final de 0,125 M durante 5 min a temperatura ambiente para detener la reacción. Se recogieron las células a 500 x g durante 10 minutos a 4°C y se lavaron dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 µM.

Lisis y sonicación

Se lisó el sedimento en tampón RIPA (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, NaCl 140 mM, Triton x-100 al 1%, SDS al 0,1%, DOC al 0,1%, inhibidores de proteasas 1x (Sigma)) y se sonica en un tubo MilliTUBE de 1 ml en un ultrasonificador S220 de Covaris durante 30 minutos hasta que la mayoría de los fragmentos tengan una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 5%, potencia incidente máxima de 140 vatios, ciclos por ráfaga 200 para células K562). Se centrifugaron los lisados a velocidad máxima durante 5 minutos a 4°C. Entonces se transfirió el sobrenadante que contenía la cromatina cortada a un tubo de PCR de 0,5 y se mantuvo en hielo.

Preparación de perlas para IP

Paralelamente a la sonicación, se bloquearon 50 µl de perlas magnéticas con proteínas A/G (10 µl para ChIPmentación de baja entrada) y se conjugaron con un anticuerpo mediante lavado y resuspensión de las mismas 2 veces en PBS, BSA al 0,5%, Tween-20 al 0,5%. Se añadió el anticuerpo y se unió a las perlas mediante su rotación > 1 h a temperatura ambiente (> 2 h a 4°C).

Inmunoprecipitación y lavados.

Por ChIP, se añadieron 50 µl de perlas magnéticas con proteína A conjugadas con anticuerpo bloqueado y se incubaron durante 3 horas a 4°C. Posteriormente se lavaron las perlas de inmunoprecipitación con 150 µl de RIPA frío (dos veces), RIPA-500 (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, NaCl 500 mM, Tritón x-100 al 1%, SDS al 0,1%, DOC al 0,1%) (dos veces) y RIPA-LiCl (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, LiCl 250 mM, Triton X-100 al 1%, DOC al 0,5%, NP40 al 0,5% (dos veces)).

Tagmentación - preparación de biblioteca

Se lavaron las perlas dos veces con Tris-Cl frío, pH 8,0 para retirar detergente, sales y EDTA. A continuación, se resuspendieron las perlas en 30 µl de la mezcla de reacción de tagmentación (Tris 10 mM, pH 8,0, MgCl 5 mM) que contenía 1 µl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina) y se incubaron a 37°C durante 10 minutos en un termociclador. Luego se colocaron las perlas sobre el imán para retirar la reacción de tagmentación seguido de 2 lavados con RIPA.

Lavado completo y elución de ADN, seguido de reticulación inversa

Finalmente, se lavaron las perlas dos veces con TE pH 8,0. Para eluir los complejos, se incubaron las perlas con 70 µl de tampón de elución (SDS al 0,5%, NaCl 300 mM, EDTA 5 mM, Tris-HCl 10 mM, pH 8,0) que contenía 2 µl de proteinasa K (NEB) durante 1 hora a 55°C y 8 horas a 65°C para revertir la reticulación con formaldehído, y se transfirió el sobrenadante a un nuevo tubo.

Purificación de ADN

Finalmente, se purificó el ADN con perlas AMPure XP (razón muestra:perlas 1:2) o columnas Qiagen MinElute.

Amplificación de bibliotecas

Se amplificó 1 µl de cada reacción de ChIPmentación en una reacción qPCR de 10 µl que contenía cebadores 0,15 µM (véase Buenrostro *et al.* Nature Methods (la publicación original de ATAC-seq) para las secuencias de cebador), SYBR green 1x y 5 µl de mezcla KAPA HIFI 2x ready mix para estimar el número óptimo de ciclos de enriquecimiento con el siguiente programa: 72°C 5 min, 98°C 30 s, 24 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se incubó la mezcla KAPA HIFI 2x ready mix a 98°C durante 45 s de manera previa a la preparación de la reacción PCR para activar la enzima de inicio en caliente para un traslado de mellas satisfactorio en la primera etapa de PCR. Se realizó el enriquecimiento final de las bibliotecas en una reacción de 50 µl usando cebadores 0,75 µM y 25 µl de mezcla KAPA HIF 2x ready mix. Se amplificaron las bibliotecas para N ciclos, donde N es igual al valor redondeado de Cq determinado en la reacción qPCR.

Purificación y selección por tamaño (opcional) de bibliotecas

Se purificaron las bibliotecas enriquecidas con un procedimiento de selección por tamaño usando perlas SPRI AMPure XP con una razón de 0,7:1 (perlas:muestra) para retirar fragmentos largos (> 600 pb), recuperando el ADN restante en la reacción con una razón de 2:1 (perlas:muestra). La secuenciación la realizó la Instalación de Secuenciación Biomédica en CeMM usando la plataforma HiSeq 2000/2500 de Illumina.

Ejemplo 2 ChIP-Seq, ChIP-tagmentación y ChIPmentación en comparación

Cultivo celular y recogida de muestras.

Se cultivaron células K562 en medio RPMI complementado con FCS al 10% y antibióticos. Se analizaron con un contador de células CASY para determinar el número de células. Se obtuvo sangre periférica de voluntarios sanos según lo aprobado por el comité de ética de la Universidad de Medicina de Viena. Se impidió la coagulación con EDTA o heparina, se diluyó la sangre periférica 1:1-1:3 en PBS y se aislaron células mononucleares de sangre periférica (PBMC) con gradiente de densidad de Lymphoprep (Axis-Shield) siguiendo las instrucciones del fabricante. Se suspendieron las células purificadas en RPMI complementado con FBS al 10% y penicilina-estreptomina.

Inmunoprecipitación de cromatina

Se sometió a prueba la ChIPmentación en combinación con tres protocolos diferentes para realizar la inmunoprecipitación de cromatina, que se describen en detalle en los ejemplos 3 a 5.

Preparación de biblioteca ChIP-seq convencional

Se sometió a reparación de extremos el ADN de ChIP purificado mediante el módulo de reparación de extremos NEBNext (NEB) según las instrucciones del fabricante. Se realizó una limpieza con perlas Ampure XP (Agencourt) según las instrucciones del fabricante. Se unieron colas A a los fragmentos usando Klenow de polimerasa (3'→ 5' exo-) (Enzymatics), y se ligaron adaptadores compatibles con TruSeq usando ADN ligasa de T4 (Enzymatics). Se seleccionó por tamaño la biblioteca final usando perlas Ampure XP para retirar los dímeros de adaptador.

Preparación de biblioteca de ChIPmentación

La ChIPmentación es compatible con diversos protocolos diferentes para ChIP, lo que facilita la aplicación de ChIPmentación a los anticuerpos que funcionan de la mejor manera con diferentes protocolos de ChIP. En general, se lleva a cabo el protocolo de ChIP de elección hasta que se lavan las perlas que transportan cromatina inmunoprecipitada con tampón de lavado que contiene LiCl (WBIII para ChIP como en el ejemplo 3, RIPA-LiCl para ChIP como en el ejemplo 4, y TF-WBIII para ChIP como en el ejemplo 5). Entonces se lavan las perlas dos veces

con Tris-Cl pH 8,0 para retirar detergente, sales y EDTA. Posteriormente, se resuspenden las perlas en 20-30 µl del tampón de reacción de tagmentación (Tris 10 mM, pH 8,0, MgCl 5 mM) que contiene 1 µl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina) y se incuban a 37°C durante 10-20 minutos en un termociclador. Tras la tagmentación, se lavan las perlas dos veces posteriormente con 150 µl de WBI (ChIP, ejemplo 3), RIPA (ChIP, ejemplo 4) o WBI (ChIP, ejemplo 5). Después de eso, se continúa con el protocolo de ChIP correspondiente con el último lavado de perlas, elución desde las perlas, reticulación inversa y purificación de ADN.

Las condiciones para la reacción de tagmentación varían dependiendo del agente usado para el aislamiento de cromatina. Las condiciones de tagmentación varían en cuanto a temperatura (por ejemplo, 4°C, 16°C, 55°C y similares), tiempo de tagmentación (por ejemplo, 1, 2, 3, 5, 15, 20, 30 ó 60 minutos y similares), concentraciones de enzima Tagment DNA (0,001, 0,01, 0,1, 0,2, 0,5, 1,5, 2, 3, 4, 5 ó 10 µl y similares) y volumen de reacción (0,001, 0,01, 0,1, 1, 5, 10, 15, 20, 50, 100 ó 200 µl y similares). Además, el tampón de reacción de tagmentación varía y también puede comprender aditivos incluyendo detergentes, sales, disolventes y similares (como ejemplo, el tampón de reacción de tagmentación puede contener dimetilformamida, polietilenglicol, acetato de manganeso (II) y similares).

Preparación de biblioteca de ChIP-tagmentación

Se midió el ADN de ChIP purificado a partir de una ChIP de H3K4me3 convencional en células mononucleares de sangre periférica (PBMC) usando un fluorímetro Qubit y luego se diluyó en Tris-Cl 10 mM, pH 8,5 complementado con Tween-20 al 0,1% hasta 100 pg, 10 pg o 2 pg de ADN total. Se realizó la reacción de tagmentación durante 5 minutos a 55°C en una reacción de 10 µl que contenía ADN diluido, 5 µl de tampón de tagmentación 2x (Illumina) y 1 µl (100 pg de ADN) o 0,5 µl (10 pg y 2 pg) de enzima de etiqueta de ADN Nextera diluida 1:10 (diluida en TE/glicerol al 50% enfriado previamente). Se amplificó el ADN sometido a tagmentación, con el kit de preparación de muestras de ADN Nextera (Illumina) según las instrucciones del fabricante con el siguiente programa: 72°C 5 min, 98°C 30 s, 14 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se purificaron las bibliotecas usando perlas SPRI AMPure XP con una razón de perlas:muestras de 1,5:1. Se preparó ADN de ChIP purificado o ADN de entrada desproteínizado a partir de ChIP de K562 como para las PBMC con ligeras modificaciones: se tomaron 5 ng de ADN de ChIP para la reacción de tagmentación usando 0,5 µl de una enzima Tn5 diluida 1:10 en una reacción de 5 µl a 55°C durante 5 minutos. Se purificó el ADN con el kit MinElute (Qiagen) y se amplificó con la mezcla KAPA HIFI 2x ready mix.

Amplificación y secuenciación de bibliotecas de ChIP-seq convencional, ChIP-tagmentación y ChIPmentación

Se amplificó 1 µl de cada reacción de ChIPmentación en una reacción qPCR de 10 µl que contenía cebadores 0,15 µM, SYBR green 1x y 5 µl de mezcla KAPA HIFI 2x ready mix para estimar el número óptimo de ciclos de enriquecimiento con el siguiente programa: 72°C 5 min, 98°C 30 s, 24 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se incubó la mezcla KAPA HIFI 2x ready mix a 98°C durante 45 s de manera previa a la preparación de la reacción PCR para activar la enzima de inicio en caliente para un traslado de mellas satisfactorio en la primera etapa de PCR. Se realizó el enriquecimiento final de las bibliotecas en una reacción de 50 µl usando cebadores 0,75 µM y 25 µl de mezcla KAPA HIF 2x ready mix. Se amplificaron las bibliotecas para N ciclos, donde N es igual al valor redondeado de Cq determinado en la reacción qPCR. Se purificaron las bibliotecas enriquecidas con un procedimiento de selección por tamaño usando perlas SPRI AMPure XP con una razón de 0,7:1 (perlas:muestra) para retirar fragmentos largos (> 600 pb), recuperando el ADN restante en la reacción con una razón de 2:1 (perlas:muestra). La secuenciación la realizó la Instalación de Secuenciación Biomédica en CeMM usando las plataformas HiSeq 2000/2500 de Illumina (véase la figura 8 para más detalles).

ATAC-seq

Se realizó el mapeo de cromatina abierta con el ensayo de cromatina accesible por transposasa (ATAC-seq) tal como se describió previamente con adaptaciones menores para las células K562. En cada experimento, se lavaron 1 x 10⁵ células una vez en 50 µl de PBS, se resuspendieron en 50 µl de tampón de lisis de ATAC-seq (Tris-HCl 10 mM, pH 7,4, NaCl 10 mM, MgCl₂ 3 mM e IGEPAL CA-630 al 0,01%), y se centrifugaron durante 10 min a 4°C. Tras la centrifugación, se lavó el sedimento brevemente en 50 µl de tampón de MgCl₂ (Tris 10 mM, pH 8,0, MgCl₂ 5 mM) antes de incubar en la mezcla de reacción con transposasa (12,5 µl de tampón TD 2x, 2 µl de transposasa (Illumina) y 10,5 µl de agua libre de nucleasas) durante 30 min a 37°C. Después de la purificación de ADN con el kit MinElute (Qiagen), se usó 1 µl del ADN eluido en una reacción qPCR para estimar el número óptimo de ciclos de amplificación. La amplificación de la biblioteca estuvo seguida por una selección por tamaño de SPRI para excluir fragmentos más grandes de 1200 pb. Se midió la concentración de ADN con un fluorímetro Qubit (Life Technologies).

Procesamiento de datos de secuenciación y análisis bioinformático

Se recortaron las lecturas usando el software Skewer. Se alinearon las lecturas recortadas con el conjunto hg19/GRCh37 del genoma humano usando el programa Bowtie2 con el parámetro “--very sensitive” (muy sensible). Para los datos de ChIPmentación y ATAC-seq, se ajustaron las posiciones de inicio de lectura para representar el

centro del evento de transposición. Las lecturas que se alinearon con la hebra positiva se desplazaron +4 pb, y las lecturas que se alinearon con la hebra negativa se desplazaron -5 pb tal como se describió anteriormente². Se usó MACS2 para asignar picos en muestras de ChIPmentación, ChIP-seq y ATAC-seq. Para los datos de ChIP y ChIPmentación, se ejecutó MACS2 usando un ancho de banda de 200 pb, y el control de IgG coincidente como fondo independientemente para las réplicas biológicas. Para las marcas de histona amplias (H3K27me3, H3K36me3) se proporcionaron las banderas y los argumentos "--broad" (amplio), "--nomodel" (sin modelo), "--extsize 73" (tamaño de ext. 73) y "--pvalue 1e-3" (valor de p 1e-3). Después de asegurarse de la sistematicidad entre réplicas, se realizó un análisis posterior en los picos asignados de las réplicas biológicas fusionadas de la misma manera que se describió. Para el análisis de correlación de las muestras tanto de ChIPmentación como de ChIP-seq, se calcularon los recuentos de lectura en ventanas de 1000 pb en todo el genoma y se normalizaron en relación con el número total de lecturas no duplicadas. Se calcularon los coeficientes de correlación de Pearson y se representó gráficamente el logaritmo de base 2 de la señal para todas las ventanas. Se realizaron comparaciones entre réplicas biológicas, entre diferentes técnicas (ChIP-seq frente a ChIPmentación), y entre diferentes números de células, en los últimos dos casos basándose en réplicas biológicas fusionadas. Se realizaron comparaciones entre picos asignados calculando la fracción de los picos del 5% o del 25% superior que se solapan con picos de la otra réplica. Se realizó la misma comparación entre datos de ChIP-seq y ChIPmentación, y entre muestras de ChIPmentación producidas con diferentes números de células usando muestras con ambas réplicas combinadas.

Ejemplo 3 Protocolo de ChIP a modo de ejemplo compatible con ChIPmentación

Se lavaron las células una vez con PBS y se fijaron con paraformaldehído al 1% en hasta 1 ml de PBS durante 5 minutos a temperatura ambiente. Se añadió glicina para detener la reacción. Se recogieron las células a 500 x g durante 10 minutos a 4°C (se realizó el trabajo posterior en hielo y se usaron disoluciones tampón y disoluciones frías a menos que se especifique de otro modo) y se lavaron dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 µM. Se lisó el sedimento en tampón de lisis celular (HEPES/KOH 50 mM, pH 7,4, NaCl 140 mM, EDTA 1 mM, EGTA 0,5 mM, glicerol al 10%, NP-40 al 0,5%, Triton X-100 al 0,25%, inhibidores de proteasa 1x (Sigma)) durante 10 minutos en hielo. Se aislaron los núcleos centrifugando las células lisadas durante 10 minutos a 1.000 x g a 4°C, se desechó el sobrenadante y se resuspendió el sedimento en tampón de sonicación (Tris-HCl 10 mM, pH 7,6, EDTA 1 mM, SDS al 0,1%) y se sonicó en un tubo microTUBE de 130 µl (para hasta 3 x 10⁶ células) en un ultrasonificador S220 de Covaris durante 12 minutos hasta que la mayoría de los fragmentos tenían una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 2%, potencia incidente máxima de 105 vatios, ciclos por ráfaga 200). Se centrifugaron los lisados a velocidad máxima durante 5 minutos a 4°C y se transfirió el sobrenadante a un nuevo tubo. Se ajustó el lisado a 200 µl por IP con una composición de tampón de HEPES 20 mM, SDS al 0,1%, Triton X-100 al 1%, NaCl 150 mM, EDTA 1 mM, EGTA 0,5 mM y se incubó con un anticuerpo contra H3K4me3 (1 µg/IP, pAb-003-050 de Diagenode) o H3K27me3 (1 µg/IP, pAb-195-050 de Diagenode) durante la noche a 4°C en un rotador. Se bloquearon 20 µl de perlas magnéticas con proteína A (o proteína G, dependiente del anticuerpo usado) durante la noche con BSA al 0,1% en PBS y se añadieron a la IP al día siguiente durante 2 horas en un rotador a 4°C para capturar los fragmentos inmunoprecipitados. Se lavó la cromatina inmunoprecipitada posteriormente con WBI (HEPES 20 mM, NaCl 150 mM, SDS al 0,1%, DOC al 0,1%, Triton X-100 al 1%, EDTA 1 mM, EGTA 0,5 mM) (dos veces), WBII (HEPES 20 mM, NaCl 500 mM, SDS al 0,1%, DOC al 0,1%, Triton X-100 al 1%, EDTA 1 mM, EGTA 0,5 mM) (una vez), WBIII (HEPES 20 mM, LiCl 250 mM, DOC al 0,5%, NP-40 al 0,5%, EDTA 1 mM, EGTA 0,5 mM) (una vez) y WBIV (HEPES 20 mM, EDTA 1 mM, EGTA 0,5 mM) (dos veces). Luego se incubaron las perlas con 70 µl de tampón de elución (SDS al 0,5%, NaCl 300 mM, EDTA 5 mM, Tris-HCl 10 mM, pH 8,0) que contenía 2 µl de proteinasa K (NEB) durante 1 hora a 55°C y 8 horas a 65°C para revertir la reticulación con formaldehído, y se transfirió el sobrenadante a un nuevo tubo. Se añadieron otros 30 µl de tampón de elución a las perlas durante 1 minuto, y se combinaron los eluatos e incubaron con otro 1 µl de proteinasa K durante 1 hora a 55°C. Finalmente, se purificó el ADN con perlas SPRI AMPure XP (razón muestra-perlas de 1:2) o columnas Qiagen MinElute.

Ejemplo 4 Protocolo de ChIP a modo de ejemplo compatible con ChIPmentación

Se lavaron las células una vez con PBS y se fijaron con paraformaldehído al 1% en hasta 1,5 ml de PBS durante 10 minutos a temperatura ambiente. Se añadió glicina para detener la reacción. Se recogieron las células a 500 x g durante 10 minutos a 4°C (se realizó el trabajo posterior en hielo y se usaron disoluciones tampón y disoluciones frías a menos que se especifique de otro modo) y se lavaron dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 µM. Se lisó el sedimento en tampón RIPA (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, NaCl 140 mM, Triton x-100 al 1%, SDS al 0,1%, DOC al 0,1%, inhibidores de proteasas 1x (Sigma)) y se sonicó en un tubo MilliTube de 1 ml en un ultrasonificador S220 de Covaris durante 30 minutos hasta que la mayoría de los fragmentos tenían una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 5%, potencia incidente máxima de 140 vatios, ciclos por ráfaga 200). Se centrifugaron los lisados a velocidad máxima durante 5 minutos a 4°C, y se transfirió el sobrenadante que contenía la cromatina sonicada a un nuevo tubo. Paralelamente, se bloquearon 50 µl (10 µl para ChIPmentación de baja entrada) de perlas magnéticas con proteína A o proteína G (dependiendo del anticuerpo usado) y se conjugaron con un anticuerpo mediante lavado y resuspensión dos veces en PBS, BSA al 0,5%, Tween-20 al 0,5%. Se añadió el anticuerpo y se unió a las perlas mediante rotación > 1 hora a temperatura ambiente. Los anticuerpos usados fueron contra H3K4me1 (1 µg/IP, pAb-194-050 de Diagenode), H3K36me3 (1 µg/IP, pAb-192-050 de Diagenode) y REST (10 µg/IP, 07-579 de Millipore). Las perlas conjugadas con

anticuerpos bloqueados se colocaron luego sobre un imán, se retiró el sobrenadante y se añadió el lisado sonificado a las perlas seguido de incubación durante 3 horas a 4°C en un rotador. Se lavaron las perlas posteriormente con 150 µl de RIPA (dos veces), RIPA-500 (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, NaCl 500 mM, Triton x-100 al 1%, SDS al 0,1%, DOC al 0,1%,) (dos veces), RIPA-LiCl (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, LiCl 250 mM, Triton X-100 al 1%, DOC al 0,5%, NP40 al 0,5%) y TE pH 8,0 (dos veces). Luego se incubaron las perlas con 70 µl de tampón de elución (SDS al 0,5%, NaCl 300 mM, EDTA 5 mM, Tris-HCl 10 mM, pH 8,0) que contenía 2 µl de proteinasa K (NEB) durante 1 hora a 55°C y 8 horas a 65°C para revertir la reticulación con formaldehído, y se transfirió el sobrenadante a un nuevo tubo. Finalmente, se purificó el ADN con perlas SPRI AMPure XP (razón muestra-perlas de 1:2) o columnas Qiagen MinElute.

Ejemplo 5 Protocolo de ChIP a modo de ejemplo compatible con ChIPmentación

Se lavaron las células una vez con PBS y se fijaron con paraformaldehído al 1% en hasta 1,5 ml de PBS durante 5-10 minutos a temperatura ambiente. Se añadió glicina para detener la reacción. Se recogieron las células a 500 x g durante 10 minutos a 4°C (se realizó el trabajo posterior en hielo y se usaron disoluciones tampón y disoluciones frías a menos que se especifique de otro modo) y se lavaron dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 µM. Se lisó el sedimento en tampón L3B (Tris-HCl 10 mM, pH 8,0, NaCl 100 mM, EDTA 1 mM, EGTA 0,5 mM, desoxicolato de Na al 0,1%, N-lauroilsarcosina al 0,5%, inhibidores de proteasas 1x (Sigma)) y se sonica en un tubo MilliTUBE de 1 ml en un ultrasonificador S220 de Covaris durante 20 minutos hasta que la mayoría de los fragmentos tengan una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 5%, potencia incidente máxima de 140 vatios, ciclos por ráfaga 200). Se complementaron los lisados con Triton-X-100 al 1% y se centrifugaron a velocidad máxima durante 5 minutos a 4°C, y se transfirió el sobrenadante que contenía la cromatina sonificada a un nuevo tubo. Paralelamente, se bloquearon las perlas y se conjugaron con un anticuerpo lavándolas dos veces en PBS con BSA al 0,5% y resuspendiendo 50 µl (10 µl de perlas para ChIPmentación de baja entrada) de perlas magnéticas con proteína A o proteína G (dependiendo del anticuerpo usado) por IP en 200 µl de PBS con BSA al 0,5%. Se añadió el anticuerpo y se unió a las perlas mediante rotación > 1 hora a temperatura ambiente o 2 h a 4°C en un rotador. Los anticuerpos usados fueron contra H3K27ac (2 µg, pAb-196-050 de Diagenode), PU.1 (5 µg/IP, sc-352 de Santa Cruz), CTCF (10 µl/IP, 07-729 de Millipore) y GATA1 (4 µg/IP y 2 µg para baja entrada, ab11852 de Abcam). Se añadieron perlas magnéticas conjugadas con anticuerpo bloqueado al tubo que contenía la cromatina y se incubaron durante 3 horas a 4°C. Se lavaron las perlas posteriormente con 150 µl de TF-WBI (Tris-HCl 20 mM/pH 7,4, NaCl 150 mM, SDS al 0,1%, Triton X-100 al 1%, EDTA 2 mM) (dos veces), TF-WBIII (LiCl 250 mM, Triton X-100 al 1%, DOC al 0,7%, Tris-HCl 10 mM, EDTA 1 mM (dos veces) y TET (Tween-20 al 0,2%, Tris-HCl 10 mM/pH 8,0, EDTA 1 mM) (dos veces). Luego se incubaron las perlas con 70 µl de tampón de elución (SDS al 0,5%, NaCl 300 mM, EDTA 5 mM, Tris HCl 10 mM, pH 8,0) que contenía 2 µl de proteinasa K (NEB) durante 1 hora a 55°C y 8 horas a 65°C para revertir la reticulación con formaldehído, y se transfirió el sobrenadante a un nuevo tubo. Se añadieron otros 30 µl de tampón de elución a las perlas durante 1 minuto y se combinaron los eluatos e incubaron con otro 1 µl de proteinasa K durante 1 hora a 55°C. Finalmente, se purificó el ADN con perlas SPRI AMPure XP (razón muestra-perlas de 1:2) o columnas Qiagen MinElute.

Ejemplo 6: Configuraciones de sonicación a modo de ejemplo

Para la línea celular leucémica K562, se realizó una fijación de 10 minutos a temperatura ambiente con formaldehído al 1%. Se sonificó la cromatina en tampón de sonicación (Tris-HCl 10 mM, pH 7,6, EDTA 1 mM, SDS al 0,1%) en un tubo microTUBE de Covaris de 130 µl (para hasta 3 x 10⁶ células) en un ultrasonificador S220 de Covaris (o versiones similares) durante 10-15 minutos con los ajustes: ciclo de trabajo del 2%, potencia incidente máxima de 105 vatios, ciclos por ráfaga 200, temperatura del agua recomendada máxima de 8°C, bomba de desgasificación encendida. Como segundo ejemplo, la cromatina puede estar en tampón RIPA (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, NaCl 140 mM, Triton x-100 al 1%, SDS al 0,1%, DOC al 0,1%, inhibidores de proteasas 1x (Sigma)) y sonificarse en un tubo MilliTUBE de Covaris de 1 ml en un ultrasonificador S220 de Covaris (o versiones similares de la máquina) durante 25-30 minutos con los ajustes: ciclo de trabajo del 5%, potencia incidente máxima de 140 vatios, ciclos por ráfaga 200, temperatura del agua recomendada máxima 8°C, bomba de desgasificación encendida.

Pueden usarse otros dispositivos de sonicación, como ejemplo el sistema Bioruptor (Diagenode): como ajuste de sonicación a modo de ejemplo adecuado para varias líneas celulares, puede sonificarse la cromatina en tampón de lisis (Tris-HCl 10 mM, pH 8,0, NaCl 100 mM, EDTA 1 mM, EGTA 0,5 mM, desoxicolato de Na al 0,1%, N-lauroilsarcosina al 0,5%, inhibidores de proteasas 1x (P8340 de Sigma)) con los ajustes "Alto" en tubos Eppendorf (50 µl-200 µl) 2x 15 minutos, ciclos de sonicación: 30 segundos ON/30 segundos OFF o en tubos cónicos de 15 ml (500 µl - 1,5 ml) con resonadores durante 15 minutos, ciclos de sonicación: 30 segundos ON/30 segundos OFF.

Otro dispositivo usado para la sonicación es un sonicador con sonda. Como ejemplo, puede sonificarse cromatina en tampón de lisis (Tris-HCl 10 mM, pH 8,0, NaCl 100 mM, EDTA 1 mM, EGTA 0,5 mM, desoxicolato de Na al 0,1%, N-lauroilsarcosina al 0,5%, inhibidores de proteasas 1x (P8340 de Sigma)) en hielo 6 veces 10 segundos a una salida de ~12 W con una pausa recomendada de 30 s entre ciclos de sonicación para impedir el sobrecalentamiento de la cromatina usando un sonicador Sonifier 450 de Branson con una sonda de micropunta.

Ejemplo 7 - Análisis de tumores primarios.

Se cortan en secciones trozos de tumor congelados de 50 µm usando un microtomo y se transfieren a un tubo de reacción en hielo (20-50 secciones son suficientes para múltiples reacciones de ChIPmentación de histonas dependiendo del tamaño del tumor). Se lavan las secciones una vez con PBS y se fijan usando paraformaldehído al 1% en hasta 1,5 ml de PBS durante 10 minutos a temperatura ambiente. Se añade glicina para detener la reacción.

5 Se recogen las células a 500 x g durante 10 minutos a 4°C (se realiza el trabajo posterior sobre hielo y las disoluciones tampón y las disoluciones se enfrían a menos que se especifique de otro modo) y se lavan dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 µM. Se lisó el sedimento en tampón L3B (Tris-HCl 10 mM, pH 8,0, NaCl 100 mM, EDTA 1 mM, EGTA 0,5 mM, desoxicolato de Na al 0,1%, N-lauroilsarcosina al 0,5%, inhibidores de proteasas 1x (Sigma)) y se sónica en un tubo milliTUBE de 1 ml en un ultrasonificador S220

10 de Covaris durante 35 minutos hasta que la mayoría de los fragmentos tengan una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 5%, potencia incidente máxima de 140 vatios, ciclos por ráfaga 200). Se complementan los lisados con Triton-X-100 al 1% y se centrifugan a velocidad máxima durante 5 minutos a 4°C, y se transfiere el sobrenadante que contiene la cromatina sonicada a un nuevo tubo. En paralelo, se bloquean las perlas y se conjugan con un anticuerpo lavándolas dos veces en PBS con BSA al 0,5% y resuspendiendo 50 µl (10 µl de perlas para ChIPmentación de baja entrada) de perlas magnéticas con proteína A o proteína G (dependiendo del anticuerpo usado) por IP en 200 µl de PBS con BSA al 0,5%. Se añade el anticuerpo y se une a las perlas mediante rotación > 1 hora a temperatura ambiente. Ejemplos de anticuerpos son contra H3K27ac (2 µg, PAb-196-050 de Diagenode), PU.1 (5 µg/IP, Sc-352 de Santa Cruz), CTCF (10 µl/IP, 07-729 de Millipore) y GATA1 (4 µg/IP y 2 µg para baja entrada, ab11852 de Abcam). Se añaden perlas magnéticas conjugadas con anticuerpo bloqueado al tubo

20 que contiene la cromatina y se incuban durante 3 horas a 4°C. Se lavan las perlas posteriormente con 150 µl de TF-WBI (Tris-HCl 20 mM/pH 7,4, NaCl 150 mM, SDS al 0,1%, Triton X-100 al 1%, EDTA 2 mM) (dos veces) y TF-WBIII (LiCl 250 mM, Triton X-100 al 1%, DOC al 0,7%, Tris-HCl 10 mM, EDTA 1 mM) (dos veces). Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 µl de la mezcla de reacción de tagmentación (Tris 10 mM, pH 8,0, MgCl 5 mM) que contiene 1 µl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina) y se incuban a 37°C durante 10 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con TF-WBI. Se lavan las perlas con TET (Tween-20 al 0,2%, Tris-HCl 10 mM/pH 8,0, EDTA 1 mM) (dos veces). Luego se incuban las perlas con 70 µl de tampón de elución (SDS al 0,5%, NaCl 300 mM, EDTA 5 mM, Tris HCl 10 mM, pH 8,0) que contiene 2 µl de proteinasa K (NEB) durante 1 hora

30 a 55°C y 8 horas a 65°C para revertir la reticulación con formaldehído, y se transfiere el sobrenadante a un nuevo tubo. Se añaden otros 30 µl de tampón de elución a las perlas durante 1 minuto y se combinan los eluatos y se incuban con otro 1 µl de proteinasa K durante 1 hora a 55°C. Finalmente, se purifica el ADN con perlas SPRI AMPure XP (razón muestra-perlas de 1:2) o columnas Qiagen MinElute, eluyendo en 11 µl de H₂O. Se amplifica 1 µl de cada reacción de ChIPmentación en una reacción qPCR de 10 µl que contiene cebadores 0,15 µM, SYBR green 1x y 5 µl de mezcla KAPA HIFI 2x ready mix para estimar el número óptimo de ciclos de enriquecimiento con el siguiente programa: 72°C 5 min, 98°C 30 s, 24 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se preincubó la mezcla KAPA HIFI 2x ready mix a 98°C durante 45 s de manera previa a la preparación de la reacción PCR para activar la enzima de inicio en caliente para un traslado de mellas satisfactorio en la primera etapa de PCR. Se realiza el enriquecimiento final de las bibliotecas (usando los 10 µl restantes de

40 ChIP) en una reacción de 50 µl usando cebadores 0,75 µM y 25 µl de mezcla KAPA HIF 2x ready mix. Se amplifican las bibliotecas para N ciclos, donde N es igual al valor redondeado de C_q determinado en la reacción qPCR. Las bibliotecas enriquecidas se purifican con un procedimiento de selección por tamaño usando perlas SPRI AMPure XP con una razón de 0,7:1 (perlas:muestra) para retirar fragmentos largos (> 600 pb), recuperando el ADN restante en la reacción con una razón de 2:1 (perlas:muestra). Se realiza secuenciación usando las plataformas HiSeq 2000/2500 de Illumina.

Ejemplo 8 - ChIPmentación en muestras incrustadas en parafina, fijadas en formalina (muestras FFPE) de muestras clínicas u otras fuentes

Se fijan las muestras en formalina y se incrustan en parafina con métodos conocidos en la técnica. Para usar la invención en muestras FFPE, se lleva a cabo el desparafinado de secciones de muestra de tejido mediante incubaciones secuenciales (10 min cada una) en 1 ml de disolución de histolemon (de seis a ocho veces) a

50 temperatura ambiente. Luego, se rehidratan las muestras disminuyendo las concentraciones de etanol partiendo del 100% (etanol absoluto) hasta el 95%, el 70%, el 50% y el 20%, con agua como etapa final (5 min a temperatura ambiente para cada etapa de rehidratación). Se incuban las secciones FFPE rehidratadas en 0,5 ml de tampón de permeabilización [1 x solución salina tamponada con Tris (TBS), Tween20 al 0,5%, PMSF 1 mM y ARNasa A

55 10 µg/ml] durante 30 minutos a temperatura ambiente en una plataforma rotatoria. Después de centrifugación a 18.000 x g durante 5 min a +4°C, se resuspenden las muestras en 200 µl de tampón de digestión [Tris-HCl 50 mM (pH 7,4), sacarosa 0,32 M, MgCl₂ 4 mM, CaCl₂ 1mM y PMSF 0,1 mM]. Se fragmentan parcialmente muestras derivadas de FFPE mediante sonicación leve, usando un sonicador Labsonic L (B. Braun, Biotech International) y luego se digieren durante 1 min a 37°C con nucleasa microcócica (N.70196Y; USB) a la concentración final de

60 1 U/10 µg de cromatina. Después de centrifugación a 18.000 x g durante 5 min a +4°C, se resuspenden las muestras en 200 µl de tampón de sonicación [1 x TBS, SDS al 0,1% y Na₂EDTA 1 mM (pH 8,0)] y se fragmentan adicionalmente. Después de centrifugación a 8.000 x g durante 5 min a temperatura ambiente, se recoge el primer sobrenadante (volumen de ~170 µl). Se lavan los sedimentos una vez con 50 µl de tampón de sonicación, se agitan

en vórtex durante 5 segundos y se centrifugan de nuevo para obtener el segundo sobrenadante (para alcanzar un volumen final de $\approx 220 \mu\text{l}$). Se cuantifica la cromatina fluorimétricamente mediante Qubit (Invitrogen). Se lleva a cabo la inmunoselección de cromatina en tampón ChIP [Tris-HCl 30 mM (pH 7,4), NaCl 50 mM, Na₂EDTA 5 mM y PMSF 0,1 mM] usando 260-600 ng de cromatina para cada ensayo (dependiendo o bien de la cantidad de cromatina extraída de muestras FFPE en cada experimento o bien del número de ensayos de ChIP que van a realizarse) y se incuban durante 16 h a +4°C en una plataforma rotatoria con el anticuerpo deseado. Se añaden veinte microlitros de suspensión al 50% vol/vol de conjugado rec-proteína G-Sepharose 4B (preincubado 16 h a +4°C con 1 mg/ml de BSA en tampón ChIP; Zymed) a cada ensayo de ChIP y se incuban durante 3 h a +4°C. Después de centrifugación a 2.000 x g durante 5 min a +4°C, se lavan los sedimentos secuencialmente con 2 ml de tampón de lavado A [Tris-HCl 50 mM (pH 7,4), TritonX-100 al 1%, NaCl 50 mM, Na₂EDTA 5 mM y PMSF 0,1 mM] y 2 ml de tampón de lavado B [Tris-HCl 50 mM (pH 7,4), TritonX-100 al 1%, NaCl 100 mM, Na₂EDTA 5 mM y PMSF 0,1 mM]. Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 μl de la mezcla de reacción de tagmentación (Tris 10 mM, pH 8,0, MgCl 5 mM) que contiene 1 μl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina) y se incuban a 37°C durante 10 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con tampón de lavado A y 10 ml de tampón de lavado C [Tris-HCl 50 mM (pH 7,4), TritonX-100 al 1%, NaCl 150 mM, Na₂EDTA 5 mM y PMSF 0,1 mM]. Se lleva a cabo la elución añadiendo 200 μl de tampón de elución [1 x Tris-EDTA (TE)/SDS al 1%] e incubando durante 30 min a temperatura ambiente en una plataforma rotatoria. Después de centrifugación a 1.200 x g durante 5 min a temperatura ambiente, se guarda el sobrenadante y se repite la elución para obtener un volumen final de 400 μl (fracción unida). Aislamiento de ADN. Se realizó la desreticulación a través de una incubación durante la noche a 65°C en tampón de elución/NaCl 0,2 M, seguido de digestión con proteinasa K 80 $\mu\text{g/ml}$ (3 h a +45°C). Se aisló el ADN mediante extracciones secuenciales con un volumen de un tercio de fenol:cloroformo (1:1), un volumen de fenol:cloroformo (1:1) y un volumen de cloroformo. Se precipita el ADN durante la noche a -20°C. Después de centrifugación, se resuspenden los sedimentos de ADN en 11 μl de tampón TE (almacenado a -20°C). Se amplifica 1 μl de cada reacción de ChIPmentación en una reacción qPCR de 10 μl que contiene cebadores 0,15 μM , SYBR green 1x y 5 μl de mezcla KAPA HIFI 2x ready mix para estimar el número óptimo de ciclos de enriquecimiento con el siguiente programa: 72°C 5 min, 98°C 30 s, 24 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se preincubó la mezcla KAPA HIFI 2x ready mix a 98°C durante 45 s de manera previa a la preparación de la reacción PCR para activar la enzima de inicio en caliente para un traslado de mellas satisfactorio en la primera etapa de PCR. Se realiza el enriquecimiento final de las bibliotecas (usando los 10 μl restantes de ChIP) en una reacción de 50 μl usando cebadores 0,75 μM y 25 μl de mezcla KAPA HIF 2x ready mix. Se amplifican las bibliotecas para N ciclos, donde N es igual al valor redondeado de Cq determinado en la reacción qPCR. Las bibliotecas enriquecidas se purifican con un procedimiento de selección por tamaño usando perlas SPRI AMPure XP con una razón de 0,7:1 (perlas:muestra) para retirar fragmentos largos (> 600 pb), recuperando el ADN restante en la reacción con una razón de 2:1 (perlas:muestra). Se realiza secuenciación usando las plataformas HiSeq 2000/2500 de Illumina.

Ejemplo 9 - ChIPmentación en leucemias humanas

La leucemia linfocítica crónica de células B (B-CLL), también conocida como leucemia linfocítica crónica (CLL), es el tipo más común de leucemia (un tipo de cáncer de los glóbulos blancos) en adultos. La CLL afecta a los linfocitos de células B, que se originan en la médula ósea, se desarrollan en los ganglios linfáticos y normalmente combaten las infecciones produciendo anticuerpos.

Aislamiento de muestras de CLL de pacientes primarios y selección negativa de células CD2+ mediante Robosep

Se añade butirato de sodio (NaB) a sangre periférica fresca hasta una concentración final de 5 mM. Esto se coloca en capas sobre un volumen igual de Ficoll (GE Healthcare, Amersham, R.U.) a temperatura ambiente y se centrifuga a 13,8 g durante 20 min. sin freno a 4°C. Se extrae la fase de PBMC y se lava dos veces en 20 ml de PBS que contiene NaB 5 mM. Se lavan las células obtenidas en medios completos y se resuspenden en 250 μl de tampón Robosep y se configura la máquina Robosep para la selección negativa según las instrucciones del fabricante (kit de selección positiva de CD2 humano EasySep, número de catálogo 18657; StemCell Technologies, Grenoble, Francia). Se recogen las células que no se unen a la columna (es decir, la población CD2-), se resuspenden en medios y se usan directamente para el procedimiento de inmunoprecipitación descrito en la invención.

En la práctica, pueden congelarse muestras de leucemia derivadas de pacientes usando métodos conocidos en la técnica.

Congelación de células

Las células deben congelarse a una conc. de 1×10^8 células/ml. Añadir gota a gota y mezclar repetidamente un vol. igual de RPMI/FCS al 50% + DMSO al 20% a lo largo de un periodo de 5 min en las células leucémicas (concentración máxima de 1×10^8 células/ml, concentración final de medios de congelación RPMI/FCS al 50% + DMSO al 10%). Transferir 1 ml de células en medios de congelación a criotubos marcados estériles. Transferir los criotubos a una caja de crioconservación que tenga temperatura ambiente y colocar la caja de crioconservación en

un congelador a -80°C. Transferir los criotubos congelados a un tanque de nitrógeno en el plazo de 24 horas.

Descongelación de células

Descongelar las células muy rápidamente en un baño de agua a 37°C. Limpiar el vial con etanol al 70% antes de abrirlo. Transferir inmediatamente 1 ml de células a un tubo de 15 ml que contiene RPMI precalentado a 37°C/FCS al 10%. Centrifugar las células a 250 x g 5 minutos a temperatura ambiente. Resuspender el sedimento celular cuidadosamente en 2 ml de RPMI precalentado a 37°C/FCS al 10%. Proceder a la inmunoprecipitación.

Inmunoprecipitación

Se lavan las células una vez con PBS y se fijan con paraformaldehído al 1% en hasta 1 ml de PBS durante 5 minutos a temperatura ambiente. Se añade glicina para detener la reacción. Se recogen las células a 500 x g durante 10 minutos a 4°C (se realizó el trabajo posterior en hielo y se usan disoluciones tampón y disoluciones frías a menos que se especifique de otro modo) y se lava dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 µM. Se lisa el sedimento en tampón de lisis celular (HEPES/KOH 50 mM, pH 7,4, NaCl 140 mM, EDTA 1 mM, EGTA 0,5 mM, glicerol al 10%, NP-40 al 0,5%, Tritón X-100 al 0,25%, inhibidores de proteasas 1x (Sigma)) durante 10 minutos en hielo. Se aíslan los núcleos centrifugando las células lisadas durante 10 minutos a 1.000 x g a 4°C, se desecha el sobrenadante y se resuspende el sedimento en tampón de sonicación (Tris-HCl 10 mM, pH 7,6, EDTA 1 mM, SDS al 0,1%) y se sonica en un tubo microTUBE de 130 µl (para hasta 3 x 10⁶ células) en un ultrasonificador S220 de Covaris durante 15 minutos hasta que la mayoría de los fragmentos tengan una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 2%, potencia incidente máxima de 105 vatios, ciclos por ráfaga 200). Se centrifugan los lisados a velocidad máxima durante 5 minutos a 4°C y se transfiere el sobrenadante a un nuevo tubo. Se ajusta el lisado hasta 200 µl por IP con una composición de tampón de HEPES 20 mM, SDS al 0,1%, Triton X-100 al 1%, NaCl 150 mM, EDTA 1 mM, EGTA 0,5 mM y se incuba con un anticuerpo de elección durante la noche a 4°C en un rotador. Se bloquean 20 µl de perlas magnéticas con proteína A (o proteína G, dependiendo del anticuerpo usado) durante la noche con BSA al 0,1% en PBS y se añaden a la IP al día siguiente durante 2 horas en un rotador a 4°C para capturar los fragmentos inmunoprecipitados. La cromatina inmunoprecipitada se lava posteriormente con WBI (HEPES 20 mM, NaCl 150 mM, SDS al 0,1%, DOC al 0,1%, Triton X-100 al 1%, EDTA 1 mM, EGTA 0,5 mM) (dos veces), WBII (HEPES 20 mM, NaCl 500 mM, SDS al 0,1%, DOC al 0,1%, Triton X-100 al 1%, EDTA 1 mM, EGTA 0,5 mM) (una vez) y WBIII (HEPES 20 mM, LiCl 250 mM, DOC al 0,5%, NP-40 al 0,5%, EDTA 1 mM, EGTA 0,5 mM) (una vez). Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 µl de la mezcla de reacción de tagmentación (Tris 10 mM, pH 8,0, MgCl 5 mM) que contiene 1 µl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina) y se incuban a 37°C durante 10 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con WBI. Se lavan las perlas con WBIV (HEPES 20 mM, EDTA 1 mM, EGTA 0,5 mM) (dos veces). Luego se incuban las perlas con 70 µl de tampón de elución (SDS al 0,5%, NaCl 300 mM, EDTA 5 mM, Tris-HCl 10 mM, pH 8,0) que contiene 2 µl de proteinasa K (NEB) durante 1 hora a 55°C y 8 horas a 65°C para revertir la reticulación con formaldehído, y se transfiere el sobrenadante a un nuevo tubo. Se añaden otros 30 µl de tampón de elución a las perlas durante 1 minuto, y se combinan los eluatos y se incuban con otro 1 µl de proteinasa K durante 1 hora a 55°C. Finalmente, se purifica el ADN con perlas SPRI AMPure XP (razón muestra-perlas de 1:2) o columnas Qiagen MinElute. Se eluye el ADN en 11 µl de tampón EB (Tris-HCl 10 mM, pH 8,5). Se amplifica 1 µl de cada reacción de ChIPmentación en una reacción qPCR de 10 µl que contiene cebadores 0,15 µM, SYBR green 1x y 5 µl de mezcla KAPA HIFI 2x ready mix para estimar el número óptimo de ciclos de enriquecimiento con el siguiente programa: 72°C 5 min, 98°C 30 s, 24 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se preincuba la mezcla KAPA HIFI 2x ready mix a 98°C durante 45 s de manera previa a la preparación de la reacción PCR para activar la enzima de inicio en caliente para un traslado de mellas satisfactorio en la primera etapa de PCR. Se realiza el enriquecimiento final de las bibliotecas (usando los 10 µl restantes de ChIP) en una reacción de 50 µl usando cebadores 0,75 µM y 25 µl de mezcla KAPA HIF 2x ready mix. Se amplifican las bibliotecas para N ciclos, donde N es igual al valor redondeado de Cq determinado en la reacción qPCR. Las bibliotecas enriquecidas se purifican con un procedimiento de selección por tamaño usando perlas SPRI AMPure XP con una razón de 0,7:1 (perlas:muestra) para retirar fragmentos largos (> 600 pb), recuperando el ADN restante en la reacción con una razón de 2:1 (perlas:muestra). Se realiza secuenciación usando las plataformas HiSeq 2000/2500 de Illumina.

Ejemplo 10 - ChIPmentación en líneas celulares

Las células K562 fueron la primera línea de leucemia mielógena inmortalizada humana que se estableció. Las células K562 son del tipo de eritroleucemia, y se deriva la línea de una paciente con LMC de 53 años en crisis blástica. Las células son no adherentes y redondeadas, son positivas para el gen de fusión bcr:abl y tienen cierta semejanza proteómica tanto con granulocitos como con eritrocitos no diferenciados.

Se lavan las células una vez con PBS y se fijan con paraformaldehído al 1% en hasta 1 ml de PBS durante 5 minutos a temperatura ambiente. Se añade glicina para detener la reacción. Se recogen las células a 500 x g durante 10 minutos a 4°C (se realizó el trabajo posterior en hielo y se usan disoluciones tampón y disoluciones frías

a menos que se especifique de otro modo) y se lava dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 μ M. Se lisa el sedimento en tampón de lisis celular (HEPES/KOH 50 mM, pH 7,4, NaCl 140 mM, EDTA 1 mM, EGTA 0,5 mM, glicerol al 10%, NP-40 al 0,5%, Tritón X-100 al 0,25%, inhibidores de proteasas 1x (Sigma)) durante 10 minutos en hielo. Se aíslan los núcleos centrifugando las células lisadas durante 10 minutos a 1.000 x g a 4°C, se desecha el sobrenadante y se resuspende el sedimento en tampón de sonicación (Tris-HCl 10 mM, pH 7,6, EDTA 1 mM, SDS al 0,1%) y se sonica en un tubo microTUBE de 130 μ l (para hasta 3 x 10⁶ células) en un ultrasonificador S220 de Covaris durante 15 minutos hasta que la mayoría de los fragmentos tengan una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 2%, potencia incidente máxima de 105 vatios, ciclos por ráfaga 200). Se centrifugan los lisados a velocidad máxima durante 5 minutos a 4°C y se transfiere el sobrenadante a un nuevo tubo. Se ajusta el lisado hasta 200 μ l por IP con una composición de tampón de HEPES 20 mM, SDS al 0,1%, Triton X-100 al 1%, NaCl 150 mM, EDTA 1 mM, EGTA 0,5 mM y se incuba con un anticuerpo de elección durante la noche a 4°C en un rotador. Se bloquean 20 μ l de perlas magnéticas con proteína A (o proteína G, dependiendo del anticuerpo usado) durante la noche con BSA al 0,1% en PBS y se añaden a la IP al día siguiente durante 2 horas en un rotador a 4°C para capturar los fragmentos inmunoprecipitados. La cromatina inmunoprecipitada se lava posteriormente con WBI (HEPES 20 mM, NaCl 150 mM, SDS al 0,1%, DOC al 0,1%, Triton X-100 al 1%, EDTA 1 mM, EGTA 0,5 mM) (dos veces), WBII (HEPES 20 mM, NaCl 500 mM, SDS al 0,1%, DOC al 0,1%, Triton X-100 al 1%, EDTA 1 mM, EGTA 0,5 mM) (una vez) y WBIII (HEPES 20 mM, LiCl 250 mM, DOC al 0,5%, NP-40 al 0,5%, EDTA 1 mM, EGTA 0,5 mM) (una vez). Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 μ l de la mezcla de reacción de tagmentación (Tris 10 mM, pH 8,0, MgCl 5 mM) que contiene 1 μ l de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina) y se incuban a 37°C durante 10 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con WBI. Se lavan las perlas con WBIV (HEPES 20 mM, EDTA 1 mM, EGTA 0,5 mM) (dos veces). Luego se incuban las perlas con 70 μ l de tampón de elución (SDS al 0,5%, NaCl 300 mM, EDTA 5 mM, Tris-HCl 10 mM, pH 8,0) que contiene 2 μ l de proteinasa K (NEB) durante 1 hora a 55°C y 8 horas a 65°C para revertir la reticulación con formaldehído, y se transfiere el sobrenadante a un nuevo tubo. Se añaden otros 30 μ l de tampón de elución a las perlas durante 1 minuto, y se combinan los eluatos y se incuban con otro 1 μ l de proteinasa K durante 1 hora a 55°C. Finalmente, se purifica el ADN con perlas SPRI AMPure XP (razón muestra-perlas de 1:2) o columnas Qiagen MinElute. Se eluye el ADN en 11 μ l de tampón EB (Tris-HCl 10 mM, pH 8,5). Se amplifica 1 μ l de cada reacción de ChIPmentación en una reacción qPCR de 10 μ l que contiene cebadores 0,15 μ M, SYBR green 1x y 5 μ l de mezcla KAPA HIFI 2x ready mix para estimar el número óptimo de ciclos de enriquecimiento con el siguiente programa: 72°C 5 min, 98°C 30 s, 24 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se preincubó la mezcla KAPA HIFI 2x ready mix a 98°C durante 45 s de manera previa a la preparación de la reacción PCR para activar la enzima de inicio en caliente para un traslado de mellas satisfactorio en la primera etapa de PCR. Se realiza el enriquecimiento final de las bibliotecas (usando los 10 μ l restantes de ChIP) en una reacción de 50 μ l usando cebadores 0,75 μ M y 25 μ l de mezcla KAPA HIF 2x ready mix. Se amplifican las bibliotecas para N ciclos, donde N es igual al valor redondeado de C_q determinado en la reacción qPCR. Las bibliotecas enriquecidas se purifican con un procedimiento de selección por tamaño usando perlas SPRI AMPure XP con una razón de 0,7:1 (perlas:muestra) para retirar fragmentos largos (> 600 pb), recuperando el ADN restante en la reacción con una razón de 2:1 (perlas:muestra). Se realiza secuenciación usando las plataformas HiSeq 2000/2500 de Illumina.

Ejemplo 11 - ChIPmentación en organismos modelo o partes de organismo modelo que pueden tener bajos números de células

Los métodos de la invención sólo requieren cantidades de baja entrada para analizar las interacciones histona-ADN en todo el genoma. Se prevé que la invención permita el análisis de etapas tempranas del desarrollo de animales individuales que consisten en bajos números de células.

Un ejemplo es el pez cebra (*Danio rerio*), que es un pez tropical de agua dulce perteneciente a la familia de las carpas (*Cyprinidae*) del orden *Cypriniformes*. Originario de la región del Himalaya, es un pez de acuario popular, vendido frecuentemente con el nombre comercial de danio cebra. El pez cebra es también un importante organismo modelo de vertebrados en la investigación científica. Tras la fecundación, los huevos se dividen y después de 4 horas el embrión ya consiste en varios miles de células, lo que puede ser suficiente para analizar una modificación de histona en un único embrión usando la invención. La agrupación de embriones de etapas tempranas del desarrollo también puede usarse para aumentar los números de células de modo que la invención pueda usarse en las células. El protocolo sugerido puede adaptarse a otros organismos de todo tipo, por ejemplo, ratón, cuando se lleva a cabo el aislamiento de los tipos de células/tejidos/etapas de desarrollo deseados con métodos bien conocidos en la técnica. En determinadas circunstancias, puede ser difícil obtener la gran cantidad de embriones requeridos para esta técnica de una vez, es decir, en experimentos de sobreexpresión o silenciamiento. En estos casos, pueden realizarse el procesamiento y la fijación de embriones en lotes que pueden congelarse en nitrógeno líquido y almacenarse a -80°C hasta que se recoja el número total de embriones requeridos. Para garantizar que todos los embriones se recojan en la misma etapa de desarrollo, se aparean las hembras y los machos de pez cebra sólo durante 15 minutos, se recogen los embriones en placas Petri con medio de embrión (medio E3: NaCl 5 mM, KCl 0,17 mM, CaCl₂ 0,4 mM, MgSO₄ 0,16 mM) y se reproducen a 28°C hasta que alcancen la etapa de desarrollo

deseada. Recoger los embriones en 0,50 ml de E3 y añadir 5 µl de Pronase (Roche, ref. 10165921001) a 30 mg/ml. Agitar suavemente e incubar los embriones a 28°C. El ablandamiento del corion lleva aproximadamente 15 minutos. Examinarlos con un microscopio estereoscópico hasta que se detecten los primeros embriones sin corion. Lavar inmediatamente los embriones meticulosamente con medio E3 (tres veces) para retirar Pronase completamente.

5 Para liberar los embriones de sus coriones, pipetearlos con cuidado en el medio E3 con una pipeta. Transferir los embriones con una pipeta a un tubo de 0,5 ml y retirar todo el E3. Añadir 0,46 ml de E3 y 0,4 ml de PFA al 4% (PFA al 4% (SIGMA P6148), tampón fosfato 200 mM, pH 7,4, NaOH 0,02 N) a los embriones y agitarlos suavemente a temperatura ambiente durante 15 minutos. Añadir glicina (Merck, 1.00590.1000) hasta una concentración final de 0,125 M para extinguir el formaldehído y agitar suavemente durante 5 minutos a temperatura ambiente. Retirar el sobrenadante y enjuagar los embriones tres veces en PBS 1X enfriado en hielo. Retirar el PBS y proceder con la lisis celular o congelar en nitrógeno líquido y almacenar los sedimentos a -80°C. Trabajar en una sala fría a 4°C de ahora en adelante. Mezclar bien las perlas magnéticas con proteína G Dyna. Tomar 10 µl (por anticuerpo) y lavarlos en 1 ml de nueva disolución de bloqueo (BSA al 0,5% en PBS 1X; puede mantenerse a 4°C durante una semana) en un tubo de cierre seguro de 1,5 ml. Recoger las perlas centrifugando a 3000 rpm durante 3 min. Lavar las perlas en 1,5 ml de disolución de bloqueo dos veces más. Resuspender las perlas vigorosamente después de cada lavado. Recoger las perlas con el soporte magnético (DYNAMag-Spin, Invitrogen 123.20D) y desechar el sobrenadante. Resuspender las perlas en 10 µl de disolución de bloqueo y añadir el anticuerpo. Cantidades posibles para algunos anticuerpos: 1 µl de Ac anti-H3K4me1 (Diagenode, n.º de cat. CS-037-100, concentración no determinada), 1 µl de Ac anti-H3K4me3 (Diagenode, n.º de cat. pAb-003-050, 1,1 µg/µl), 1 µl de Ac anti-H3K27ac (Abcam, número de catálogo ab4729, 0,80 mg/ml) y 1 µl de Ac anti-H3K27me3 (Millipore 07-449, 1 mg/ml). Incubar el anticuerpo a 4°C durante un mínimo de 4 h o durante la noche en una plataforma rotatoria. Recoger las perlas con el soporte magnético y retirar el sobrenadante. Lavar las perlas en 0,2 ml de disolución de bloqueo en la sala fría. Repetir estas etapa dos veces más. Resuspender las perlas en 10 µl de disolución de bloqueo. Añadir inhibidores de proteasa (pastilla Complete, Roche 11 697 498 001) a todos los tampones de lisis justo antes de su uso. (Puede mantenerse una disolución madre 50X de Complete (1 pastilla/ml de PBS 1X) a -20°C durante dos meses). Resuspender los embriones reticulados en 0,13 ml de tampón de lisis celular (Tris-HCl 50 mM, pH 7,5, EDTA 10 mM, SDS al 1%). Pipetear aspirando y dispensando y expulsar los embriones para la perturbación. Disponer el tubo sobre hielo e incubar durante 10 minutos. Dejar las muestras 15 minutos en hielo y renovar el baño de agua con hielo. Sonicar en un tubo microTUBE de 130 µl (Covaris, hasta 3 x 10⁶ células) en un ultrasonificador S220 de Covaris durante 10-60 minutos (dependiendo de la cantidad de embriones usados, debe determinarse empíricamente) hasta que la mayoría de los fragmentos tengan una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 2%, potencia incidente máxima de 105 vatios, ciclos por ráfaga 200).

Añadir 2 volúmenes de tampón de dilución de IP (Tris-HCl 16,7 mM, pH 7,5, NaCl 167 mM, EDTA 1,2 mM, SDS al 0,01%). Añadir Triton X-100 al 1% a la cromatina sonicada. Centrifugar a 14.000 rpm durante 10 minutos a 4°C y transferir la cromatina a un nuevo tubo. Añadir 10 µl de mezcla de perlas magnéticas/anticuerpo a cada alícuota de cromatina sonicada. Incubar los tubos durante la noche en una plataforma rotatoria a 4°C. Recoger las perlas de los tubos con el soporte magnético y retirar el sobrenadante. Añadir 0,2 ml de tampón de lavado RIPA (HEPES 50 mM pH 7,6, EDTA 1 mM, DOC al 0,7%, Igepal al 1%, LiCl 0,5 M) a cada tubo. Agitar los tubos suavemente para resuspender las perlas. Recoger las perlas con el soporte magnético y retirar el sobrenadante. Repetir la etapa anterior tres veces más. Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 µl de la mezcla de reacción de tagmentación (en este caso: 2 µl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina), 15 µl de tampón Tagment DNA 2X del kit de preparación de muestras de ADN Nextera y 13 µl de agua libre de nucleasas) y se incuban a 37°C durante 3 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con RIPA. Lavar una vez con 1 ml de TBS 1X (Tris 50 mM, pH 7,5, NaCl 150 mM), recoger las perlas con el soporte magnético y retirar el sobrenadante. Resuspender las perlas en 200 µl 1X TBS. Centrifugar a 3000 rpm durante 3 minutos y aspirar cualquier cantidad de TBS residual. Añadir 60 µl de tampón de elución (NaHCO₃ 50 mM, SDS al 1%). Eluir los complejos ADN-proteína de las perlas a 65°C durante 10-15 minutos con una breve agitación en vórtex cada 2 minutos. Centrifugar las perlas a 14.000 rpm durante 1 min. Transferir 650 µl de sobrenadante a un tubo de cierre seguro de 1,5 ml. Añadir NaCl 300 mM. Revertir las reticulaciones de formaldehído durante 6 h o durante la noche a 65°C. Añadir ARNasa A hasta una concentración final de 0,33 µg/µl e incubar a 37°C durante 2 h. Añadir 1 volumen de fenol/cloroformo/alcohol isoamílico (25:24:1, AMRESCO 0883), mezclar y centrifugar durante 5 min. Transferir la fase superior a un nuevo tubo de cierre seguro de 1,5 ml. Añadir 1 µg de glucógeno. Añadir 1/10 de NaAc 3 M y dos volúmenes de EtOH al 100%. Para precipitar el ADN, centrifugar durante 10 minutos a 14.000 rpm. Lavar el sedimento con 500 µl de EtOH frío al 75% y centrifugar durante 5 min a 14.000 rpm a 4°C. Secar al aire los sedimentos a temperatura ambiente y resuspender en 70 µl de Tris-HCl 10 mM, pH 8. Purificar el ADN (tanto de la entrada como de la reacción de ChIP) usando el kit de purificación de PCR QIAquick (Qiagen 28104) (seguir las instrucciones proporcionadas con el equipo). Eluir en 11 µl de tampón EB (Tris-HCl 10 mM, pH 8,5). Se amplifica 1 µl de cada reacción de ChIPmentación en una reacción qPCR de 10 µl que contiene cebadores 0,15 µM, SYBR green 1x y 5 µl de mezcla KAPA HIFI 2x ready mix para estimar el número óptimo de ciclos de enriquecimiento con el siguiente programa: 72°C 5 min, 98°C 30 s, 24 ciclos de 98°C 10 s, 63°C 30 s, 72°C 30 s, y una elongación final a 72°C durante 1 min. Se preincuba la mezcla KAPA HIFI 2x ready mix a 98°C durante 45 s de manera previa a la preparación de la reacción PCR para activar la enzima de inicio en caliente para un traslado de mellas satisfactorio en la primera etapa de PCR.

Se realiza el enriquecimiento final de las bibliotecas (usando los 10 µl restantes de ChIP) en una reacción de 50 µl usando cebadores 0,75 µM y 25 µl de mezcla KAPA HIF 2x ready mix. Se amplifican las bibliotecas para N ciclos, donde N es igual al valor redondeado de Cq determinado en la reacción qPCR. Las bibliotecas enriquecidas se purifican con un procedimiento de selección por tamaño usando perlas SPRI AMPure XP con una razón de 0,7:1 (perlas:muestra) para retirar fragmentos largos (> 600 pb), recuperando el ADN restante en la reacción con una razón de 2:1 (perlas:muestra). Se realiza secuenciación usando las plataformas HiSeq 2000/2500 de Illumina.

Ejemplo 12 - Métodos de la invención que usan una sustancia química como agente que se une a cromatina

Análisis de ocupación en todo el genoma *in vivo* de JQ1 biotinilado (Chem-seq *in vivo*). Se tratan células MM1.S con crecimiento exponencial (2 × 10⁸ células por muestra) simultáneamente con JQ1 biotinilado 5 µM (Bio-JQ1) o DMSO (vehículo) y formaldehído al 1% durante 20 min en medio de cultivo celular. La reticulación química se termina mediante la adición de tampón TRIS, pH 7,5, hasta una concentración final de TRIS 300 mM. Se recogen las células usando un raspador de silicio, se centrifugan y se lavan los sedimentos derivados tres veces con PBS. Se preparan núcleos celulares de la siguiente manera: se lisan las células en HEPES 50 mM, pH 7,5, NaCl 140 mM, EDTA 1 mM, glicerol al 10%, NP-40 al 0,5%, Triton X-100 al 0,25% más cóctel de inhibidores de proteasas 'Complete' (Roche), y se lavan los núcleos celulares una vez con Tris-HCL 10 mM, pH 8,0, NaCl 200 mM, EDTA 1 mM, EGTA 0,5 mM e inhibidores de proteasas. Se resuspenden los núcleos y se sonicen en HEPES-KOH 50 mM, pH 7,5, NaCl 140 mM, EDTA 1 mM, EGTA 1 mM, Triton X-100 al 1%, desoxicolato de Na al 0,1%, SDS al 0,1% (tampón de sonicación) y cóctel de inhibidores de proteasas a 18 W para 10 ciclos (30 s cada uno) en hielo con intervalos de 30 s entre ciclos. Se aclaran los lisados sonicados mediante centrifugación y se incuban durante 16-20 h a 4°C con perlas magnéticas con estreptavidina Dynabeads (MyOne Streptavidin T1, Invitrogen) (se bloquean las perlas en PBS que contiene BSA al 0,5% antes de esta etapa de incubación). Tras la incubación en lisado nuclear sonificado, se lavan las perlas dos veces en tampón de sonicación, una vez en tampón de sonicación que contiene NaCl 500 mM, una vez en tampón de LiCl (Tris-HCL 20 mM, pH 8,0, EDTA 1 mM, LiCl 250 mM, NP-40 al 0,5%, desoxicolato de Na al 0,5%). Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 µl de la mezcla de reacción de tagmentación (en este caso: 2 µl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina), 15 µl de tampón Tagment DNA 2X del kit de preparación de muestras de ADN Nextera y 13 µl de agua libre de nucleasas) y se incuban a 37°C durante 3 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con tampón de sonicación. Se lavan entonces las perlas una vez en TRIS 10 mM, pH 7,5, EDTA 0,1 mM. Posteriormente se eluyen los complejos proteína-ADN unidos en Tris-HCL 50 mM, pH 8,0, EDTA 10 mM, SDS al 1% a 65°C durante 15 min, y se revierten las reticulaciones mediante incubación durante la noche del eluato a 65°C. El ARN y las proteínas contaminantes se digieren mediante la adición de ARNasa y proteinasa K, respectivamente, y se purifica el ADN tal como se describió previamente³⁴. Finalmente, los fragmentos de ADN purificados se secuencian masivamente en paralelo.

Análisis de ocupación en todo el genoma *in vitro* de JQ1 biotinilado (Chem-seq *in vitro*). Se fijan células MM1.S no tratadas con crecimiento exponencial con formaldehído al 1% durante 20 min en medio de cultivo celular. Se termina la reticulación química, se preparan núcleos celulares y se obtiene un lisado nuclear sonificado tal como se describió anteriormente. Sin embargo, a diferencia del protocolo *in vivo*, las perlas Dynabeads de estreptavidina se incuban previamente en PBS que contiene BSA al 0,5% y fármaco biotinilado 200 µM o vehículo (DMSO) durante 6 h. Posteriormente se lavan las perlas con fármaco unido cuatro veces en PBS/BSA al 0,5% para retirar el fármaco no unido, y se incuban en lisado nuclear sonificado durante 16-20 h a 4°C. Todas las siguientes etapas son idénticas a las descritas anteriormente (método de Chem-seq *in vivo*).

Análisis de la ocupación en todo el genoma *in vitro* usando AT7519 biotinilado (Chem-seq *in vitro*). Se fijan células MM1.S no tratadas con crecimiento exponencial con formaldehído al 0,5% durante 5 min en medio de cultivo celular. La reticulación química se termina mediante la adición de tampón TRIS, pH 7,5, hasta una concentración final de TRIS 300 mM. Se lavan las células 3 veces en PBS y se preparan núcleos celulares de la siguiente manera: se lisan los núcleos celulares en HEPES 50 mM, pH 7,5, NaCl 140 mM, EDTA 1 mM, glicerol al 10%, NP-40 al 0,5%, Triton X-100 al 0,25% más cóctel de inhibidores de proteasas 'Complete' (Roche), y se lavan los núcleos celulares una vez con Tris-HCL 10 mM, pH 8,0, NaCl 200 mM, EDTA 1 mM, EGTA 0,5 mM e inhibidores de proteasas. Se resuspenden los núcleos y se sonicen en HEPES-KOH 50 mM, pH 7,5, NaCl 140 mM, EDTA 1 mM, EGTA 1 mM, NP-40 al 0,5%, Triton-X al 0,5% (tampón de sonicación). Se sonicen los sedimentos a 9-12 W para 4 ciclos (30 s cada uno) en un sonicador Misonix en hielo con intervalos de descanso de 1 minuto entre ciclos. Se añaden perlas con fármaco unido al producto sonificado aclarado y se permite que avance la precipitación durante 12-18 h. Posteriormente se lavan las perlas con fármaco unido tres veces en tampón de sonicación. Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 µl de la mezcla de reacción de tagmentación (en este caso: 2 µl de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina), 15 µl de tampón Tagment DNA 2X del kit de preparación de muestras de ADN Nextera y 13 µl de agua libre de nucleasas) y se incuban a 37°C durante 3 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con tampón de sonicación. Se eluyen las proteínas en SDS al 1%, y se revierten las reticulaciones mediante incubación durante la noche del eluato a 65°C en SDS al 1%. El ARN y las proteínas contaminantes se digieren mediante incubación secuencial con ARNasa A y proteinasa K, y se purifica el ADN tal como se describió

previamente. Los fragmentos de ADN purificados se someten a secuenciación masivamente en paralelo.

Análisis de ocupación en todo el genoma de psoraleno biotinilado mediante Chem-seq

Se preparan núcleos celulares a partir de células MM.S con crecimiento exponencial usando el kit Nuclei EZ prep (Sigma). Se resuspenden los núcleos luego en PBS enfriado con hielo y se incuban directamente con psoraleno biotinilado 5 μ M o vehículo (DMSO) durante 30 min a 4°C. Se lavan los núcleos una vez en PBS y se irradian inmediatamente a 360 nm durante 30 min (Stratalinker) en hielo. Se resuspenden los núcleos y se sonicen en HEPES-KOH 50 mM, pH 7,5, NaCl 140 mM, EDTA 1 mM, EGTA 1 mM, Triton X-100 al 1%, desoxicolato de Na al 0,1%, SDS al 0,1% (tampón de sonicación) y cóctel de inhibidores de proteasas a 18 W para 10 ciclos (30 s cada uno) en hielo con intervalos de 30 s entre ciclos. Los lisados sonicados se aclaran mediante centrifugación y se incuban durante 16-20 h a 4°C con perlas magnéticas con estreptavidina Dynabeads (MyOne Streptavidin T1, Invitrogen) (se bloquean las perlas en PBS que contiene BSA al 0,5% antes de esta etapa de incubación). Tras la incubación en lisado nuclear sonicado, se lavan las perlas dos veces en tampón de sonicación, una vez en tampón de sonicación que contiene NaCl 500 mM, una vez en tampón de LiCl (Tris-HCL 20 mM, pH 8,0, EDTA 1 mM, LiCl 250 mM, NP-40 al 0,5%, desoxicolato de Na al 0,5%). Se lavan las perlas dos veces con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Se resuspenden las perlas cuidadosamente en 30 μ l de la mezcla de reacción de tagmentación (en este caso: 2 μ l de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina), 15 μ l de tampón Tagment DNA 2X del kit de preparación de muestras de ADN Nextera y 13 μ l de agua libre de nucleasas) y se incuban a 37°C durante 3 minutos en un termociclador. Se retira la reacción de tagmentación colocando la reacción sobre un imán y retirando el sobrenadante, y se lavan las perlas dos veces con tampón de sonicación, seguido de lavado una vez en TRIS 10 mM, pH 7,5, EDTA 0,1 mM. Posteriormente se eluyen los complejos proteína-ADN unidos en Tris-HCL 50 mM, pH 8,0, EDTA 10 mM, SDS al 1% y biotina 10 mM, y se incuban el eluato a temperatura ambiente a 65°C. El ARN y las proteínas contaminantes se digieren mediante la adición de ARNasa y proteinasa K, respectivamente, y se purifica el ADN tal como se describió previamente. Finalmente, las muestras de ADN purificadas se irradian a 254 nm durante 5 min (Stratalinker) para revertir las reticulaciones de psoraleno-ADN, seguido de la preparación de bibliotecas, secuenciación de ADN masivamente en paralelo.

Ejemplo 13 - Protocolo de parámetros optimizados para mejorar la relación señal/ruido

Pueden optimizarse múltiples parámetros para mejorar la relación señal/ruido en experimentos que usan los métodos de la presente invención. Un protocolo a modo de ejemplo proporcionado en el presente documento reduce el tiempo de tagmentación a aproximadamente 1 minuto mediante el uso de una transferencia de tubos y un tampón de tagmentación específico. El protocolo es compatible con diversos protocolos para ChIP, como los protocolos descritos en los ejemplos 3, 4 y 5, respectivamente. Esto facilita la aplicación del protocolo a los anticuerpos que funcionan de la mejor manera con un determinado protocolo de ChIP. Se observó la mejor relación señal/ruido siguiendo los protocolos de los ejemplos 3, 4 ó 5 hasta que las perlas que transportaban cromatina inmunoprecipitada se lavaron con tampón de lavado que contenía LiCl (WBIII para el ejemplo 3, RIPA-LiCl para el ejemplo 4 y TF-WBIII para el ejemplo 5). Luego, se lavaron las perlas una vez con Tris-Cl frío pH 8,0 para retirar detergente, sales y EDTA. Posteriormente, se lavaron las perlas de nuevo con Tris-Cl frío pH 8,0 pero la reacción no se colocó sobre un imán para desechar el sobrenadante inmediatamente. En cambio, se transfirió toda la reacción, incluidas las perlas, a un nuevo tubo y luego se colocó sobre un imán para retirar el sobrenadante. Esto disminuyó la tagmentación de fragmentos de cromatina inespecíficos que se adhieren a la pared del tubo. Luego, se resuspendieron cuidadosamente las perlas en 25 μ l de la mezcla de reacción de tagmentación (Tris 10 mM, pH 8,0, MgCl₂ 5 mM, dimetilformamida al 10% p/v) que contenía 1 μ l de enzima Tagment DNA del kit de preparación de muestras de ADN Nextera (Illumina) y se incubó a 37°C durante 1 minuto en un termociclador. La dimetilformamida como disolvente aprótico polar que puede potenciar las reacciones nucleófilas y, por tanto, puede ser beneficiosa para la reacción de transposición, ya que la transposasa usa una molécula de agua para un ataque nucleófilo como mecanismo para integrar oligonucleótidos en el ADN. Alternativamente, puede usarse el tampón Tagment DNA del kit Nexera. Se retiró la reacción de tagmentación y se lavaron las perlas dos veces con WBI (ejemplo 3), RIPA (ejemplo 4) o TF-WBI (ejemplo 5). Luego se siguió el protocolo de ChIP, pero se transfirió la reacción de nuevo a un nuevo tubo cuando se lavó por segunda vez con WBIV (ejemplo 3), TE (ejemplo 4) o TET (ejemplo 5) tal como ya se describió en una etapa anterior. Esto disminuyó el remanente de fragmentos inespecíficos sometidos a tagmentación que se adhieren a la pared del tubo. El experimento se realizó con 500k células y se comparó con el protocolo convencional, sin transferencia de tubos y sin dimetilformamida, con 500k células o 10 millones de células, respectivamente. Los resultados se muestran en la figura 12. El protocolo optimizado proporciona una relación señal/ruido igual o mayor que un experimento con 10 millones de células que usa el protocolo convencional de los métodos descritos en el presente documento.

Ejemplo 14 - Duración optimizada del ensayo

Para optimizar adicionalmente la duración del ensayo, puede aumentarse la temperatura de calentamiento para revertir las reticulaciones. Por consiguiente, se realizaron los métodos de la presente invención con la etapa adicional de "reparación de extremos" para rellenar la secuencia adaptadora en la hebra inversa, como por ejemplo, en el protocolo Nextera, en el que se recomienda una etapa de PCR de 5 minutos a 72°C antes de la desnaturalización inicial. Para hacerlo así, se añadió una mezcla maestra de PCR y se calentó la muestra hasta

72°C, seguido de reticulación inversa a 95°C. Como la transposasa podría “adherirse” al ADN, los protocolos sugieren una “separación” de la transposasa con EDTA antes de la reparación de extremos, extinguiendo el EDTA con MgCl₂ para permitir la reparación de extremos (el EDTA perturbaría cualquier PCR). Se realizaron varios protocolos ultrarrápidos, incluidos protocolos que comprenden ChIPmentación hasta la etapa de lavado final, en lugar de eluir inmediatamente la cromatina de las perlas con tampón de elución, se resuspendieron las perlas añadiendo una mezcla maestra de PCR directamente para la reparación de extremos (a continuación a); EDTA para asegurar la separación completa de la transposasa de la cromatina, seguido de extinción con MgCl₂, seguido de adición de la mezcla maestra de PCR para la reparación de extremos (a continuación b); o como anteriormente sólo con un procedimiento más rápido (a continuación c). Posteriormente, se incubó la reacción a 72°C durante 5 minutos para reparar los extremos (= rellenar los extremos del adaptador de la segunda cadena), se revirtieron las reticulaciones a 95°C durante 10-15 minutos y se recargaron las reacciones con nueva mezcla maestra de PCR (o no en el caso del protocolo 2c), se añadieron cebadores y se amplificó la biblioteca. En el caso de (a), se siguió el procedimiento de ChIPmentación hasta que las perlas que transportaban la cromatina se lavaron por última vez. Se desechó el sobrenadante y se resuspendieron las perlas en 14 µl de mezcla fría KAPA HiFi Hot Start ReadyMix 1x (precalentada hasta 95°C durante 30 s). Se incubó la reacción durante 5 minutos a 72°C, luego durante 10 minutos a 95°C. Después de eso, se enfrió la reacción hasta 4°C. Para la PCR posterior, se añadieron 1,5 µl de cada uno de los cebadores directo e inverso (25 µM cada uno), 15 µl de H₂O y 18 µl de mezcla KAPA HiFi Hot Start ReadyMix 2x (precalentada hasta 95°C durante 30 s) y se amplificó el ADN. En el caso de (b), se siguió el procedimiento de ChIPmentación hasta que las perlas que transportaban la cromatina se lavaron por última vez y se desechó el sobrenadante. Se resuspendieron las perlas en 8 µl de EDTA 50 mM y se incubaron 30 minutos a 50°C. Luego, se añadieron 2 µl de MgCl₂ 200 mM e incubaron a los 30 minutos a 50°C, de manera previa a la adición de 10 µl de KAPA HiFi Hot Start ReadyMix 2x (precalentada hasta 95°C durante 30 s) y se incubaron durante 5 minutos a 72°C. Posteriormente, se incubó la reacción durante 10 minutos a 95°C, y luego se enfrió hasta 4°C. Finalmente, se añadieron 1,5 µl de cada uno de los cebadores directo e inverso 25 µM, 12 µl de H₂O y 15 µl de mezcla KAPA HiFi Hot Start ReadyMix 2x (precalentada hasta 95°C durante 30 s) y se amplificó la reacción usando 12-18 ciclos según los parámetros de la amplificación de biblioteca de ChIPmentación. En el caso de (c), se siguió el procedimiento de ChIPmentación hasta que las perlas que transportaban la cromatina se lavaron por última vez y se desechó el sobrenadante. Luego, se añadieron 11 µl de EDTA 20 mM y se incubaron durante 10 minutos a 50°C. Posteriormente, se añadieron 11 µl de MgCl₂ 20 mM + 25 µl de mezcla KAPA HiFi Hot Start ReadyMix 2x (precalentada hasta 95°C durante 30 s) y se incubaron durante 5 minutos a 72°C. Se incubó la reacción durante 10 minutos a 95°C, y luego se enfrió hasta 4°C. Luego, se añadieron 1,5 µl de cada uno de los cebadores directo e inverso 25 µM y se amplificó la reacción en 12-18 ciclos según los parámetros de amplificación de biblioteca de ChIPmentación. Los resultados se muestran en la figura 15. Por consiguiente, el uso del procedimiento ultrarrápido tal como se describe en el presente documento da como resultado una validez de datos comparable al tiempo que se reduce significativamente el tiempo experimental. En particular, en el caso de (c), todo el procedimiento desde el cultivo de células hasta la biblioteca amplificada puede completarse en un único día hábil. Más específicamente, las etapas experimentales comprenden recoger células y preparar las perlas (siguiendo el ejemplo 5) (20 min), fijar las células con formaldehído y lavar el sedimento (45 min), lisar y sonificar las células (40 min), aislar la cromatina con un anticuerpo (= etapa de inmunoprecipitación, 3 horas cuando se sigue el ejemplo 5), lavar la cromatina (30 min), añadir la transposasa (10 min), lavado posterior (10 minutos), reparación de extremos y reticulación inversa (30 min), amplificación de biblioteca (45 min) y purificación de biblioteca (45 min), lo que da como resultado un tiempo total de <8 h. En este caso, el aislamiento de células específicas de los pacientes (por ejemplo, células T CD4+) todavía puede incorporarse en la línea temporal de un largo día de trabajo (+90 minutos para la extracción de sangre y el aislamiento de CD4+). Si un segundo operario está preparando una máquina de secuenciación adecuada (por ejemplo, MiSeq de Illumina), las muestras pueden secuenciarse durante la noche, lo que da como resultado un flujo de trabajo completo desde la extracción de sangre del paciente hasta las secuencias en ~24 h, lo que permite epigenomas personalizados en una escala de tiempo clínica. Más específicamente, cuando la muestra de ChIPmentación se secuenciaría en un sistema MiSeq de Illumina usando el Kit de reactivos MiSeq v3, pueden completarse 50 ciclos de secuencia de alrededor de 25 millones de agrupaciones en 5-6 horas, lo que da como resultado un intervalo de tiempo total de ~15 horas desde la extracción de sangre hasta el experimento secuenciado de ChIPmentación.

Ejemplo 15 - Ensayo ultrarrápido usando una transposasa alternativa

Se preparó una transposasa alternativa según Picelli *et al.* (2014) Genome Research 24:2033-40. En resumen, se produce la enzima Tn5 y se almacena según Picelli *et al.* (como anteriormente) que tiene la secuencia de aminoácidos codificada por la secuencia de SEQ ID NO:1 ó 2, en la que SEQ ID NO:1 se refiere al ácido nucleico que codifica para la transposasa de producción interna que contiene una etiqueta de inteína C-terminal y un dominio de unión a quitina y la SEQ ID NO: 2 se refiere a la enzima transposasa central, y usando un vector de expresión que tiene la secuencia de SEQ ID NO: 3, se diluye una alícuota de Tn5 en tampón de dilución de Tn5 (dependiendo de la actividad de Tn5), se preparan oligonucleótidos que van a cargarse en Tn5 y se carga la Tn5 diluida con oligonucleótidos (todas las etapas tal como se describieron previamente). La Tn5 “de producción interna” puede usarse como sustituto directo. Se usaron los siguientes tampones: tampón de dilución de Tn5 (Tris-HCl 50 mM a pH 7,5; NaCl 100 mM; EDTA 0,1 mM; glicerol al 50%; Triton X-100 al 0,1% y DTT 1 mM (siempre añadir nuevo antes de diluir - 1 µl de DTT 1 M por 1 ml de tampón). Para el apareamiento previo de los oligonucleótidos de extremo de

mosaico (Mosaic End), se siguió el siguiente procedimiento: (1) preparación de una mezcla equimolar 100 μ M de oligonucleótidos Tn5ME-A + Tn5MErev y Tn5ME-B + Tn5MErev; usando

Tn5MErev: 5'-

[phos]CTGTCTCTTATACACATCT-3' (SEQ ID NO:4) ; Tn5ME-A: 5'-

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3' (SEQ ID NO:5) and Tn5ME-

B: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3' (SEQ ID NO:6),

en los que Tn5MErev se fosforila en el extremo 5'; (2) incubación de ambas mezclas a 95°C durante 3 minutos e incubación durante la noche en bloque térmico apagado para enfriamiento, luego ambos se mezclaron 1:1 (3) se realizó la generación de transposomas mediante la adición de 0,143 volúmenes de oligonucleótidos apareados previamente 1:1 a la dilución de Tn5 e incubación durante 1 hora a T/A e incubación en hielo hasta que sea necesario.

Para cada experimento, se lavaron 300.000 células de línea celular de leucemia K562 una vez con PBS y se fijaron con paraformaldehído al 1% en hasta 1,5 ml de PBS durante 10 minutos a temperatura ambiente. Se añadió glicina para detener la reacción. Se recogieron las células a 500 x g durante 10 minutos a 4°C (se realizó un trabajo posterior en hielo y se usaron disoluciones tampón y disoluciones frías a menos que se especifique de otro modo) y se lavaron dos veces con hasta 1 ml de PBS enfriado con hielo complementado con PMSF 1 μ M. Se lisó el sedimento en tampón de sonicación (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, SDS al 0,25%, inhibidores de proteasas 1x (Sigma)) y se sonicó en un tubo milliTUBE de 1 ml en un ultrasonificador S220 de Covaris durante 20 minutos hasta la mayoría de los fragmentos tenían una longitud de 200-700 pares de bases (ajustes: ciclo de trabajo del 5%, potencia incidente máxima de 140 vatios, ciclos por ráfaga 200). Se ajustaron los lisados a las condiciones de RIPA (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, NaCl 140 mM, Triton x-100 al 1%, SDS al 0,1%, DOC al 0,1%, inhibidores de proteasas 1x (Sigma)). Se centrifugaron los lisados a velocidad máxima durante 5 minutos a 4°C, y se transfirió el sobrenadante que contenía la cromatina sonicada a un nuevo tubo. Paralelamente, se bloquearon 10 μ l de perlas magnéticas con proteína A o proteína G (dependiendo del anticuerpo usado) y se conjugaron con un anticuerpo lavando y resuspendiendo dos veces en PBS, BSA al 0,5%, Tween-20 al 0,5%. Se añadió el anticuerpo y se unió a las perlas mediante rotación > 1 hora a temperatura ambiente. Los anticuerpos usados fueron contra H3K4me3 (1 μ g/IP, Diagenode). Las perlas conjugadas con anticuerpos bloqueados se colocaron luego sobre un imán, se retiró el sobrenadante y se añadió el lisado sonicado a las perlas seguido de incubación durante 3 horas a 4°C en un rotador. Se lavaron las perlas posteriormente con 150 μ l de RIPA (dos veces), RIPA-500 (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, NaCl 500 mM, Triton x-100 al 1%, SDS al 0,1%, DOC al 0,1%,) (dos veces), RIPA-LiCl (Tris-HCl 10 mM, pH 8,0, EDTA 1 mM, pH 8,0, LiCl 250 mM, Triton X-100 al 1%, DOC al 0,5%, NP40 al 0,5%) y Tris pH 8,0 (dos veces). Se añadieron 11 μ l de EDTA 20 mM a las perlas y se incubaron durante 10 minutos a 50°C. Posteriormente, se añadieron 11 μ l de MgCl₂ 20 mM + 25 μ l de mezcla KAPA HiFi Hot Start ReadyMix 2x (precalentada hasta 95°C durante 30 s) y se incubaron durante 5 minutos a 72°C. Se incubó la reacción durante 10 minutos a 95°C, y luego se enfrió hasta 4°C. Luego, se añadieron 1,5 μ l de cada uno de los cebadores directo e inverso 25 μ M y se amplificó la reacción en 10-14 ciclos según los parámetros de amplificación de biblioteca de ChIPmentación. Los resultados se muestran en la figura 15, que demuestran las bibliotecas de secuenciación de ChIPmentación para H3K4me3 usando o bien los transposomas Tn5 de Illumina disponibles comercialmente o bien los transposomas Tn5 "de producción interna" de 2 fuentes diferentes. Este experimento demuestra resultados idénticos usando o bien la transposasa de Illumina (tn5) disponible comercialmente o bien la enzima transposasa de producción interna.

Lista de secuencias

<110> CeMM - Forschungszentrum für Molekulare Medizin GmbH

<120> Métodos para estudiar ácidos nucleicos

<130> Y2067 PCT S3

<150> Documento EP15 18 9788.1

<151> 14-10-2015

<150> Documento EP15 18 0705.4

<151> 12-08-2015

<160> 6

<170> BiSSAP 1.3

<210> 1

<211> 2211

<212> ADN

<213> Secuencia artificial

5 <220>

<223> secuencia de ácido nucleico que codifica para la transposasa de producción interna que contiene una etiqueta de inteína C-terminal y un dominio de unión a quitina (CBD)

<400> 1

atgattacca gtgcactgca tcgtgcggcg gattgggcga aaagcgtgtt ttctagtgtc	60
gcgctgggtg atccgcgtcg taccgcgcgt ctggtgaatg ttgcggcgca actggccaaa	120
tatagcggca aaagcattac cattagcagc gaaggcagca aagccatgca ggaaggcgcg	180
tatcgtttta ttcgtaatcc gaacgtgagc gcggaagcga ttcgtaaagc gggtgccatg	240
cagaccgtga aactggccca ggaatttcgg gaactgctgg caattgaaga taccacctct	300
ctgagctatc gtcacaggt ggcggaagaa ctgggcaaac tgggtagcat tcaggataaa	360
agccgtgggt ggtgggtgca tagcgtgctg ctgctggaag cgaccacctt tcgtaccgtg	420
ggcctgctgc atcaagaatg gtggatgctg ccgatgatc cggcgatgca ggatgaaaaa	480
gaaagcggca aatggctggc cgctgctgca acttcgcgtc tgagaatggg cagcatgatg	540
agcaacgtga ttgcggtgtg cgatcgtgaa gcgatattc atgcgtatct gcaagataaa	600
ctggcccata acgaacgttt tgtggtgctg agcaaacatc cgcgtaaaga tgtggaaagc	660
ggcctgtatc tgtatgatca cctgaaaaac cagccggaac tgggcggcta tcagattagc	720
attccgcaga aaggcgtggt ggataaacgt ggcaaacgta aaaaccgtcc ggcgcgtaaa	780
gcgagcctga gcctgcgtag cggccgtatt accctgaaac agggcaacat taccctgaac	840
gcggtgctgg ccgaagaaat taatccgccg aaaggcgaaa ccccgctgaa atggctgctg	900
ctgaccagcg agccggtgga aagtctggcc caagcgctgc gtgtgattga tatttatacc	960
catcgttggc gcattgaaga atttcacaaa gcgtggaaaa cgggtgcggg tgcggaacgt	1020
cagcgtatgg aagaaccgga taacctggaa cgtatggtga gcattctgag ctttgtggcg	1080

ES 2 779 798 T3

```

gtgctgtctgc tgcaactgcg tgaatctttt actccgccgc aagcactgcg tgcgcagggc 1140
ctgctgaaag aagcggaaca cgttgaaagc cagagcgcgg aaaccgtgct gaccccgat 1200
gaatgccaac tgctgggcta tctggataaa ggcaaacgca aacgcaaaga aaaagcgggc 1260
agcctgcaat gggcgatat ggcgattgcg cgtctgggcg gctttatgga tagcaaactg 1320
accggcattg cgagctgggg tgcgctgtgg gaaggttggg aagcgctgca aagcaaactg 1380
gatggctttc tgcccgcaa agacctgatg gcgcagggca ttaaaatctg catcacggga 1440
gatgcactag ttgccctacc cgaggcgag tcggtacgca tcgccgacat cgtgccgggt 1500
gcgcggccca acagtgaca cgccatcgac ctgaaagtcc ttgaccggca tggcaatccc 1560
gtgctcgccg accggctgtt ccactccggc gagcatccgg tgtacacggg gcgtacgggc 1620
gaaggtctgc gtgtgacggg caccgcgaac caccggtgt tgtgtttggg cgacgtcgcc 1680
ggggtgcga ccctgctgtg gaagctgatc gacgaaatca agccgggcga ttacgcgggtg 1740
attcaacgca gcgcattcag cgtcgactgt gcaggttttg cccgcgggaa acccgaattt 1800
gcgcccacaa cctacacagt cggcgctccct ggactgggtgc gtttcttgga agcacaccac 1860
cgagaccggg acgccaagc tatcgccgac gagctgaccg acgggcgggt ctactacgcg 1920
aaagtcgcca gtgtcaccga cgcggcggtg cagccgggtg atagccttcg tgtcgacacg 1980
gcagaccacg cgtttatcac gaacgggttc gtcagccacg ctactggcct caccggtctg 2040
aactcaggcc tcacgacaaa tcctgggtga tccgcttggc aggtcaacac agcttatact 2100
gcgggacaat tggtcacata taacggcaag acgtataaat gtttgacagc ccacacctcc 2160
ttggcaggat gggaaccatc caacgttcct gccttgtggc agcttcaatg a 2211

```

<210> 2

<211> 1428

<212> ADN

5 <213> Secuencia artificial

<220>

<223> secuencia de ácido nucleico que codifica para la transposasa de producción interna

<400> 2

```

atgattacca gtgcactgca tcgtgcggcg gattgggcga aaagcgtgtt ttctagtgtc 60
gcgctgggtg atccgcgtcg taccgcgcgt ctggtgaatg ttgcggcgca actggccaaa 120
tatagcggca aaagcattac cattagcagc gaaggcagca aagccatgca ggaaggcgcg 180
tatcgtttta ttcgtaatcc gaacgtgagc gcggaagcga ttcgtaaagc gggtgccatg 240
cagaccgtga aactggccca ggaatttccg gaactgctgg caattgaaga taccacctct 300
ctgagctatc gtcacaggt ggcggaagaa ctgggcaaac tgggtagcat tcaggataaa 360
agccgtggtt ggtgggtgca tagcgtgctg ctgctggaag cgaccacctt tcgtaccgtg 420
ggcctgctgc atcaagaatg gtggatgcgt ccgatgatc cggcggatgc ggatgaaaaa 480

```

ES 2 779 798 T3

```

gaaagcggca aatggctggc cgctgctgca acttcgcgtc tgagaatggg cagcatgatg      540
agcaacgtga ttgcgggtgtg cgatcgtgaa gcgcatattc atgcgtatct gcaagataaa      600
ctggcccata acgaacgttt tgtggtgctg agcaaaccatc cgcgtaaaga tgtggaaagc      660
ggcctgtatc tgtatgatca cctgaaaaac cagccggaac tgggcggcta tcagattagc      720
attccgcaga aaggcgtggt ggataaacgt ggcaaacgta aaaaccgtcc ggcgcgtaaa      780
gcgagcctga gcctgcgtag cggccgtatt accctgaaac agggcaacat taccctgaac      840
gcggtgctgg ccgaagaaat taatccgccg aaaggcgaaa ccccgctgaa atggctgctg      900
ctgaccagcg agccgggtga aagtctggcc caagcgctgc gtgtgattga tatttatacc      960
catcgttggc gcattgaaga atttcacaaa gcgtggaaaa cgggtgcggg tgcggaacgt     1020
cagcgtatgg aagaaccgga taacctggaa cgtatggtga gcattctgag ctttgtggcg     1080
gtgcgtctgc tgcaactgcg tgaatctttt actccgccgc aagcactgcg tgcgcagggc     1140
ctgctgaaag aagcgggaaca cgttgaaagc cagagcgcgg aaaccgtgct gaccccgcat     1200
gaatgccaac tgctgggcta tctggataaa ggcaaacgca aacgcaaaga aaaagcgggc     1260
agcctgcaat gggcgatatat ggcgattgcg cgtctgggcg gctttatgga tagcaaacgt     1320
accggcattg cgagctgggg tgcgctgtgg gaaggttggg aagcgctgca aagcaaactg     1380
gatggctttc tggccgcgaa agacctgatg gcgcagggca ttaaaatc      1428

```

<210> 3

<211> 8079

<212> ADN

5 <213> Secuencia artificial

<220>

<223> secuencia de ácido nucleico del vector usado para preparar transposasa de producción interna

<400> 3

```

aactacgtca ggtggcactt ttccgggaaa tgtgcgcgga acccctatct gtttatcttt      60
ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata      120
atattgaaaa aggaagagta tgagtattca acatttcctg tgcgccctta ttcccttttt      180
tgccggcattt tgcccttctg tttttgctca ccagaaaacg ctggtgaaag taaaagatgc      240
tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggttaagat      300
ccttgagagt ttctgccccg aagaacgttc tccaatgatg agcactttta aagttctgct      360
atgtggcgcg gtattatccc gtgttgacgc cgggcaagag caactcggtc gccgcataca      420
ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg      480
catgacagta agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa      540
cttactttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg      600

```

ES 2 779 798 T3

ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga	660
cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg	720
cgaactactt actctagctt cccggcaaca attaatagac tggatggagc cggataaagt	780
tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg	840
agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc	900
ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca	960
gatcgtgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc	1020
atatatactt tagattgatt taccocgggt gataatcaga aaagcccca aaacaggaag	1080
attgtataag caaatattta aattgtaaac gttaatatTT tgttaaaatt cgcgttaaatt	1140
ttttgttaaa tcagctcatt ttttaaccaa taggccgaaa tcggcaaaat cccttataaa	1200
tcaaaagaat agcccgagat agggttgagt gttgttccag tttggaacaa gagtccacta	1260
ttaaagaacg tggactccaa cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca	1320
ctacgtgaac catcacccaa atcaagtttt ttggggtcga ggtgccgtaa agcactaaat	1380
cggaaacctt aaggagagccc ccgatttaga gcttgacggg gaaagccggc gaacgtggcg	1440
agaaaggaag ggaagaaagc gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc	1500
acgctgcgcg taaccaccac acccgccgcg cttaatgcgc cgctacaggg cgcgtaaaag	1560
gatctaggtg aagatccttt ttgataatct catgaccaa atcccttaac gtgagttttc	1620
gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt	1680
tctgcgcgta atctgctgct tgcaaacaaa aaaccaccg ctaccagcgg tggtttggtt	1740
gccggatcaa gagctacca ctctttttcc gaaggtaact ggcttcagca gagcgcagat	1800
accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc	1860
accgcctaca tacctcgctc tgctaatacct gttaccagtg gctgctgcca gtggcgataa	1920
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg	1980
ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag	2040
atacctacag cgtgagctat gagaaagcgc cacgcttccc gaaggagaa aggcggacag	2100
gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa	2160
cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt	2220
gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgc cctttttacg	2280
gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat ccctgattc	2340
tgtggataac cgtattaccg cttttgagtg agctgatacc gctcgcgcga gccgaacgac	2400
cgagcgcagc gagtcagtga gcgaggaagc tatggtgcac tctcagtaca atctgctctg	2460

ES 2 779 798 T3

atgccgcata gtttaagccag tatacaactcc gctatcgcta cgtgactggg tcatggctgc	2520
gccccgacac ccgccaacac ccgctgacgc gccctgacgg gcttgtctgc tcccggcatc	2580
cgcttacaga caagctgtga ccgtctccgg gagctgcatt tgtcagaggt tttcacccgc	2640
atcaccgaaa cgcgcgaggg agctgcggta aagctcatca gcgtggctgt gcagcgattc	2700
acagatgtct gcctgttcat ccgctgccag ctcggtttagt ttctccagaa gcgttaatgt	2760
ctggcttctg ataaagcggg ccatgttaag ggccgttttt tcctgtttgg tcaactgatgc	2820
ctccgtgtaa gggggatttc tgttcatggg ggtaaatgata ccgatgaaac gagagaggat	2880
gctcacgata cgggttactg atgatgaaca tgcccgggta ctggaacgtt gtgagggtaa	2940
acaactggcg gtatggatgc ggccggacca gagaaaaatc actcagggtc aatgccagcc	3000
gaacgccagc aagacgtagc ccagcgcgtc ggccgccatg ccggcgataa tggcctgctt	3060
ctcccgaaa cgtttgggtg cgggaccagt gacgaaggct tgagcgaggg cgtgcaagat	3120
tccgaatacc gcaacgcaga ggccgatcat cgtcgcgtc cagcgaaagc ggtcctcgcc	3180
gaaaatgacc cagacgcgtg ccggcacctg tcctacgagt tgcatgataa agaagacagt	3240
cataagtgcg gcgacgatag tcatgccccg cggccaccgg aaggagctga ctgggttgaa	3300
ggctctcaag ggcacgggtc gagatccggg tgccctaatga gtgagctaac ttacattaat	3360
tgctgtgcgc tcaactgccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg	3420
aatcgcccaa cgcgcgggga gagggcggtt gcgtattggg cgcaggggtg gtttttcttt	3480
tcaccagtga gacgggcaac agctgattgc ccttcaccgc ctggccctga gagagttgca	3540
gcaagcggtc cagctgggtt tgccccagca ggcgaaaatc ctgtttgatg gtggttaacg	3600
gcgggatata acatgagctg tcttcgggat cgtcgtatcc cactaccgag atatccgcac	3660
caacgcgcag cccggactcg gtaatggcgc gcattgcgcc cagcgccatc tgatcgttgg	3720
caaccagcat cgcagtggga acgatgccct cattcagcat ttgcatggtt tgttgaaaac	3780
cggacatggc actccagtcg ccttcccggt ccgctatcgg ctgaatttga ttgcgagtga	3840
gataatttatg ccagccagcc agacgcagac gcgccgagac agaacttaat gggcccgcta	3900
acagcgcgat ttgctgggtg cccaatgcga ccagatgctc cagcgccagt cgcgtaccgt	3960
cttcatggga gaaaataata ctgttgatgg gtgtctggtc agagacatca agaaataacg	4020
ccggaacatt agtgcaggca gcttccacag caatggcatc ctgggtcatcc agcggatagt	4080
taatgatcag cccactgacg cgttgcgcga gaagattgtg caccgccgct ttacaggctt	4140
cgacgccgct tcgttctacc atcgacacca ccacgctggc acccagttga tcggcgcgag	4200
atttaatcgc cgcgacaatt tgcgacggcg cgtgcagggc cagactggag gtggcaacgc	4260
caatcagcaa cgactgtttg cccgccagtt gttgtgccac gcggttggga atgtaattca	4320
gctccgccat cgcgcgttcc actttttccc gcgttttcgc agaaacgtgg ctggcctggt	4380

ES 2 779 798 T3

tcaccacgcg	ggaacggtc	tgataagaga	caccggcata	ctctgcgaca	tcgtataacg	4440
ttactggttt	cacattcacc	accctgaatt	gactctcttc	cgggcgctat	catgccatac	4500
cgcgaaaggt	tttgcccat	tcgatggtgt	cgggatctc	gacgctctcc	cttatgcgac	4560
tcctgcatta	ggaagcagcc	cagtagtagg	ttgaggccgt	tgagcaccgc	cgccgcaagg	4620
aatggtgcat	gccggcatgc	cgccttttcg	tcttcaagaa	ttaattccca	attccccagg	4680
catcaaataa	aacgaaaggc	tcagtcgaaa	gactgggcct	ttcgttttat	ctgtgttttg	4740
tcggtgaacg	ctctcctgag	taggacaaat	cgcgcgggag	cggatttgaa	cgttgcaag	4800
caacggcccc	gaggtggtg	ggcaggacgc	cgcgcataaa	ctgccaggaa	ttaattcccc	4860
aggcatcaaa	taaaacgaaa	ggctcagtcg	aaagactggg	cctttcgttt	tatctgttgt	4920
ttgtcgtgta	acgctctcct	gagtaggaca	aatccgccgg	gagcggattt	gaacgttgcg	4980
aagcaacggc	cgggaggtg	gcgggcagga	cgcgcgcat	aaactgccag	gaattaattc	5040
cccaggcatc	aaataaaaacg	aaaggctcag	tcgaaagact	gggccttttcg	ttttatctgt	5100
tgtttgctcg	tgaacgctct	cctgagtagg	acaaatccgc	cgggagcgga	tttgaacgtt	5160
gcgaagcaac	ggcccggagg	gtggcgggca	ggacgcccgc	cataaactgc	caggaattaa	5220
ttccccaggc	atcaataaaa	acgaaaggct	cagtcgaaag	actgggcctt	tcgttttatc	5280
tgttgtttgt	cgggtgaacgc	tctcctgagt	aggacaaatc	cgcggggagc	ggatttgaac	5340
gttgcgaagc	aacggcccgg	aggtggcg	gcaggacgcc	cgcgcataaac	tgccaggaat	5400
taattcccca	ggcatcaaat	aaaacgaaag	gctcagtcga	aagactgggc	ctttcgtttt	5460
atctgttgtt	tgctgggtgaa	cgtctcctg	agtaggacaa	atccgcggg	agcggatttg	5520
aacgttgcca	agcaacggcc	cggaggggtg	cgggcaggac	gcccgccata	aactgccagg	5580
aattggggat	cggaaattaat	tcccggttta	aaccggggat	ctcgatccc	cgaattaat	5640
acgactcact	ataggggaat	tgtgagcgga	taacaattcc	cctctagata	ttttgtttac	5700
tttaagaagg	agatatacat	atgattacca	gtgcactgca	tcgtgcggcg	gattgggcga	5760
aaagcgtgtt	ttctagtgt	gcgctgggtg	atccgcgtcg	taccgcgcgt	ctggtgaatg	5820
ttgcggcgca	actggccaaa	tatagcggca	aaagcattac	cattagcagc	gaaggcagca	5880
aagccatgca	ggaaggcgcg	tatcgtttta	ttcgtaatcc	gaacgtgagc	gcggaagcga	5940
ttcgtaaagc	gggtgccatg	cagaccgtga	aactggccca	ggaatttccg	gaactgctgg	6000
caattgaaga	taccacctct	ctgagctatc	gtcatcaggt	ggcggagaa	ctgggcaaac	6060
tgggtagcat	tcaggataaa	agccgtggtt	ggtgggtgca	tagcgtgctg	ctgctggaag	6120
cgaccacctt	tcgtaccgtg	ggcctgctgc	atcaagaatg	gtggatgcgt	cgggatgatc	6180
cggcggatgc	ggatgaaaaa	gaaagcggca	aatggctggc	cgtgctgca	acttcgcgtc	6240

ES 2 779 798 T3

tgagaatggg cagcatgatg agcaacgtga ttgCGgtgtg cgatCGtgaa gcggatattc 6300
 atgcgtatct gcaagataaa ctggcccata acgaacgttt tgtggtgcgt agcaaacatc 6360
 cgCGtaaaga tgtggaaagc ggCctgtatc tgtatgatca cctgaaaaac cagccggaac 6420
 tgggcggcta tcagattagc attccgcaga aaggcgtggt ggataaacgt ggcaaacgta 6480
 aaaacCGtcc ggCGcgtaaa gcgagcctga gcctgcgtag cggccgtatt accctgaaac 6540
 agggcaacat taccctgaac gcggtgctgg ccgaagaaat taatccgCG aaaggcgaaa 6600
 cccCGctgaa atggctgctg ctgaccagcg agccggtgga aagtctggcc caagcgtgc 6660
 gtgtgattga tatttatacc catCGttggc gcattgaaga atttcacaaa gcgtggaaaa 6720
 cgggtgCGggg tgcggaacgt cagCGtatgg aagaaccgga taacctggaa cgtatggtga 6780
 gcattctgag ctttgtggcg gtgcgtctgc tgcaactcg tgaatctttt actccgCGc 6840
 aagcactgcg tgcgcagggc ctgctgaaag aagcggaaca cgttgaaagc cagagcCGcg 6900
 aaacCGtgct gacccCGgat gaatgccaac tgctgggcta tctggataaa ggcaaacgca 6960
 aacgcaaaga aaaagcgggc agcctgcaat gggCGtatat ggcgattgcg cgtctggcg 7020
 gctttatgga tagcaaacgt accggcattg cgagctgggg tgcgctgtgg gaaggttggg 7080
 aagcgtgca aagcaaactg gatggctttc tggccCGgaa agacctgatg gcgagggca 7140
 ttaaaatctg catcacggga gatgcactag ttgccctacc cgagggcgag tcggtacgca 7200
 tcgCCgacat cgtgccgggt gcgCGgcca acagtgacaa cgcctcgcac ctgaaagtcc 7260
 ttgaccggca tggcaatccc gtgctCGccg accggctgtt ccaactCGgc gagcatccgg 7320
 tgtacacggt gcgtacggtc gaaggtctgc gtgtgacggg caccgCGaac caccCGttgt 7380
 tgtgtttggt cGacgtCGcc ggggtGCCga ccctgctgtg gaagctgatc gacgaaatca 7440
 agccgggCGa ttacCGgtg attcaacgca gcgcattcag cgtCGactgt gcaggttttg 7500
 cccCGgggaa accCGaattt gcgcccacaa cctacacagt cggCGtccct ggactggtgc 7560
 gtttcttgga agcacaccac cGagaccCGg acgcccAagc tatCGccgac gagctgaccg 7620
 acggcgCGgtt ctactacgCG aaagtCGcca gtgtcaccga cgcCGgCGt cagccggtgt 7680
 atagccttcg tGtcGacacg gcagaccacg cgtttatcac gaacgggttc gtcagccacg 7740
 ctactggcct caccggtctg aactcaggcc tcacgacaaa tcctgggtgta tccgcttggc 7800
 aggtcaacac agcttatact gcgggacaat tggtcacata taacggcaag acgtataaat 7860
 gtttcgagcc ccacacctcc ttggcaggat gggaaccatc caacgttcct gccttgtggc 7920
 agcttcaatg actgcaggaa ggggatccg ctgctaacaa agccCGaaag gaagctgagt 7980
 tggctgctgc caccgctgag caataactag cataaccct tgggCGctct aaacgggtct 8040
 tgaggggttt tttgctgaaa ggaggaacta tatccggat 8079

<210> 4

<211> 19

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Tn5MErev

	<400> 4	
	ctgtctcttta tacacatct	19
	<210> 5	
	<211> 33	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Tn5ME-A	
	<400> 5	
10	tcgtcggcag cgtcagatgt gtataagaga cag	33
	<210> 6	
	<211> 34	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Tn5ME-B	
	<400> 6	
	gtctcgtggg ctcggagatg tgtataagag acag	34

REIVINDICACIONES

1. Método para preparar una biblioteca de secuenciación, comprendiendo el método:
 - (a) añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico;
 - (b) aislar la cromatina unida por dicho agente;
 - 5 (c) añadir transposasa a la cromatina aislada de la etapa (b);
 - (d) aislar el ácido nucleico de la cromatina; y
 - (e) obtener una biblioteca de secuenciación.
2. Método para el mapeo de interacciones moleculares que implican a ácido nucleico, comprendiendo el método:
 - 10 (a) añadir un agente que se une a cromatina a una muestra que comprende un ácido nucleico;
 - (b) aislar la cromatina unida por dicho agente;
 - (c) añadir transposasa a la cromatina aislada de la etapa (b);
 - (d) aislar el ácido nucleico de la cromatina;
 - (e) amplificar el ácido nucleico;
 - 15 (f) secuenciar el ácido nucleico amplificado; y
 - (g) identificar interacciones moleculares.
3. Método según las reivindicaciones 1 ó 2, en el que el agente es un anticuerpo.
4. Método según las reivindicaciones 1 a 3, en el que la muestra que comprende un ácido nucleico se ha preparado mediante
 - 20 (i) cultivar y recoger células;
 - (ii) fijar las células;
 - (iii) lisar las células y obtener de ese modo una primera muestra que comprende un ácido nucleico; y
 - (iv) sonicar la primera muestra y obtener de ese modo una segunda muestra que comprende un ácido nucleico, en el que dicha segunda muestra se usa en el método según las reivindicaciones 1 ó 2.
- 25 5. Método según las reivindicaciones 1 a 4, en el que el método comprende además una etapa de revertir las reticulaciones introducidas durante la fijación de las células.
6. Método según una cualquiera de las reivindicaciones 1 a 5, en el que el ácido nucleico es ADN.
7. Método según una cualquiera de las reivindicaciones 4 a 6, en el que las células comprenden complejos ácido nucleico-proteína.
- 30 8. Método según la reivindicación 7, en el que las células son células humanas, células animales, células bacterianas, células de levadura, células de arqueas, células vegetales o virus.
9. Método según la reivindicación 8, en el que las células humanas o animales son células enfermas o células no enfermas o células derivadas de tejido enfermo o no enfermo y/o, en el que las células humanas o animales son células cancerosas, células inmunitarias, células sanguíneas o células madre.
- 35 10. Método según la reivindicación 9, en el que el cáncer es un cáncer sólido o cáncer de la sangre, preferiblemente en el que el cáncer de la sangre es leucemia o en el que el cáncer sólido es un tumor.
11. Método según la reivindicación 4, en el que la etapa (ii) comprende añadir una sustancia química y/o medios físicos y/o en el que la etapa (iv) comprende sonicar hasta que la mayoría de los fragmentos de ácido nucleico tienen 20-5000, preferiblemente 200-300, pares de bases de longitud.
- 40 12. Método según las reivindicaciones 1 ó 2, en el que el anticuerpo se une específicamente a histonas, factores de transcripción o proteínas que se unen a histonas y/o factores de transcripción.
13. Método según la reivindicación 12, en el que las proteínas que se unen a histonas y/o factores de

- transcripción son proteínas de remodelación de ácido nucleico o enzimas modificadoras de cromatina, preferiblemente, en el que la histona es H3.3, H2A.Z, CENP-A, H3.2, H3.3A, H3.3B, H4 o H3.1, preferiblemente en el que la histona es una histona modificada, en el que la modificación es metilación, acetilación, propionilación, butirilación, crotonilación, 2-hidroxiisobutirilación, malonilación, succinilación y/o ribosilación, preferiblemente en el que la histona modificada es H3K4me1/2/3, H2BK5me1, H3K27me1/2/3, H3K9me1/2/3, H4K20me1, H3K79me1, H3K36me3, H2AK5ac, H2AK9ac, H2BK5ac, H2BK12ac, H2BK20ac, H2BK120ac, H3K4ac, H3K9ac, H3K14ac, H3K18ac, H3K23ac, H3K27ac, H3K36ac, H4K5ac, H4K8ac, H4K12ac, H4K16ac, H4K91ac, H2Aub o H2Bub.
- 5
14. Método según las reivindicaciones 1 ó 2, en el que el agente es una sustancia química, preferiblemente en el que la sustancia química es un fármaco, preferiblemente en el que la sustancia química está biotinilada.
- 10
15. Método según las reivindicaciones 1 ó 2, en el que la transposasa comprende etiquetas de secuencia de ADN aleatorias o etiquetas de secuencia de ADN definidas, preferiblemente en el que la transposasa es una transposasa Tn5.

Figura 1

	Nano ChIP-seq	linDA	ChIP a pequeña escala	ChIP-seq de transportador	ChIP-seq de transportador de ADN bacteriano	iChIP	ChIP con números celulares bajos	ChIP-exo	ChIP-nexus	ChIPmentación
Título de la publicación	Genome-wide chromatin maps from limited numbers of hematopoietic progenitors	Single-tube linear DNA amplification (LinDA) for robust ChIP-seq	In vivo epigenomic profiling of germ cells reveals germ cell molecular signatures	A Carrier-assisted ChIP-seq method for estrogen receptor-chromatin interactions from breast cancer core needle biopsy samples	Amplification of pico-scale DNA mediated by bacterial carrier DNA for small-cell-number transcription factor ChIP-seq	Chromatin state dynamics during blood formation	Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline in vivo	Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution	ChIP-nexus enables improved detection of in vivo transcription factor binding footprints	ChIPmentation: fast, cheap, low-input ChIP-seq for histones and transcription factors
Referencia	Adli <i>et al.</i> , Nature Methods 7, 615-618, 2010	Shankararayanan <i>et al.</i> , Nature Methods 8, 565-567, 2011	Ng <i>et al.</i> , Developmental Cell 24, 324-333, 2013	Zwart <i>et al.</i> , BMC Genomics 14, 232, 2013	Jakobsen <i>et al.</i> , BMC Genomics 16, 46, 2015	Lara-Astiaso <i>et al.</i> , Science 345, 943-949, 2014	Sachs <i>et al.</i> , Cell Reports 3, 1777-1784, 2013	Rhee <i>et al.</i> , Cell 147, 1408-1419, 2011	He <i>et al.</i> , Nature Biotechnology 33, 395-401, 2015	este estudio
Números de células requeridos para ChIP-seq en marcas de histona	10k	5k	1k (H3K27me3)	n/a	10k	500 (en reacciones múltiples)	12.5k	n/a	n/a	10k

	Nano ChIP-seq	linDA	ChIP a pequeña escala	ChIP-seq de transportador	ChIP-seq de transportador de ADN bacteriano	iChIP	ChIP con números celulares bajos	ChIP-exo	ChIP-nexus	ChIPmentación
Números de células requeridos para ChIP-seq en factores de transcripción	n/a	5k	n/a	10k (1 factor sometido a prueba)	10k	10k (1 factor sometido a prueba; protocolo requiere multiplexación de las muestras)	n/a	2.5M	10M	100k
Protocolo	Preparación de biblioteca de múltiples etapas; el protocolo incluye preamplificación	Preparación de biblioteca de múltiples etapas; el protocolo incluye preamplificación	Preparación de biblioteca de múltiples etapas; el protocolo incluye preamplificación	El protocolo usa ARN transportador e histona recombinantes	El protocolo usa ADN transportador bacteriano	El protocolo usa inmunoprecipitación adicional con anticuerpo anti-histona H3	Kit de preparación de biblioteca de entrada comercial (Rubicon Genomics)	Protocolo de preparación de biblioteca de múltiples etapas complejo	Protocolo de preparación de biblioteca de etapas individual, simple	
Tiempo requerido para preparación de biblioteca	2 días	1.5 días de transcripción <i>in vitro</i> antes de la preparación de biblioteca mediada por ligamiento de adaptador convencional	4-5 horas adicionales antes de la preparación de biblioteca mediada por ligamiento de adaptador convencional	Preparación de biblioteca mediada por ligamiento de adaptador convencional	Preparación de biblioteca mediada por ligamiento de adaptador convencional	1 día adicional para inmunoprecipitación, después de preparación de biblioteca mediada por ligamiento de adaptador usando protocolo personalizado	~70 minutos	Preparación de biblioteca de 5-6 horas + purificaciónes de ADN	7-8 horas (tratamiento de ChIP-exo y preparación de biblioteca nexus)	Preparación de biblioteca de 10-20 minutos

	Nano ChIP-seq	linDA	ChIP a pequeña escala	ChIP-seq de transportador	ChIP-seq de transportador de ADN bacteriano	iChIP	ChIP con números celulares bajos	ChIP-exo	ChIP-nexus	ChIPmentación
Ventajas	Números de células bajos para modificación de histonas	Números de células bajos para modificación de histonas y factores de transcripción	Números de células muy bajos para modificaciones de histonas	Números de células muy bajos para factores de transcripción	Números de células muy bajos para modificaciones de histonas y factores de transcripción	Números de células muy bajos, muy adecuados para muchas ChIP con el mismo anticuerpo	Números de células muy bajos para modificaciones de histonas	Baja señal de fondo, vista en alta resolución de la unión del factor de transcripción	Baja señal de fondo, vista en alta resolución de la unión del factor de transcripción (mejorada con respecto a ChIP-exo)	Rápida, fácil; barata; números de células bajos para modificaciones de histonas y números de células relativamente bajos para factores de transcripción; potencial de análisis de alta resolución de regiones diana
Inconvenientes	Protocolo que requiere mucho tiempo; requisitos de reactivos adicionales; sesgo potencial de PCR debido a la preamplificación	Protocolo que requiere mucho tiempo; requisitos de reactivos adicionales; sesgo potencial de PCR debido a la preamplificación	Protocolo que requiere mucho tiempo; requisitos de reactivos adicionales; sesgo potencial de PCR debido a la preamplificación	Requisitos de reactivos adicionales	Costes de secuenciación aumentados debido a incremento del ADN bacteriano	Protocolo que requiere mucho tiempo; necesidad de agrupar ChIP (no son posibles ChIP individuales); subóptimo para análisis de TF sin sesgo debido al pre-	Se basan en el kit de preparación de biblioteca comercial caro	Protocolo que requiere mucho tiempo; requiere altos números de células; requisitos de reactivos adicionales	Protocolo que requiere mucho tiempo; requiere altos números de células; requisitos de reactivos adicionales	Requisitos de reactivos adicionales (transposasa + cebador de PCR)

	Nano ChIP-seq	ChIP-linDA	ChIP a pequeña escala	ChIP-seq de transportador	ChIP-seq de transportador de ADN bacteriano	iChIP	ChIP con números celulares bajos	ChIP-exo	ChIP-nexus	ChIPmentación
						enriquecimiento de fragmentos de cromatina unidos a histona H3				

Figura 2

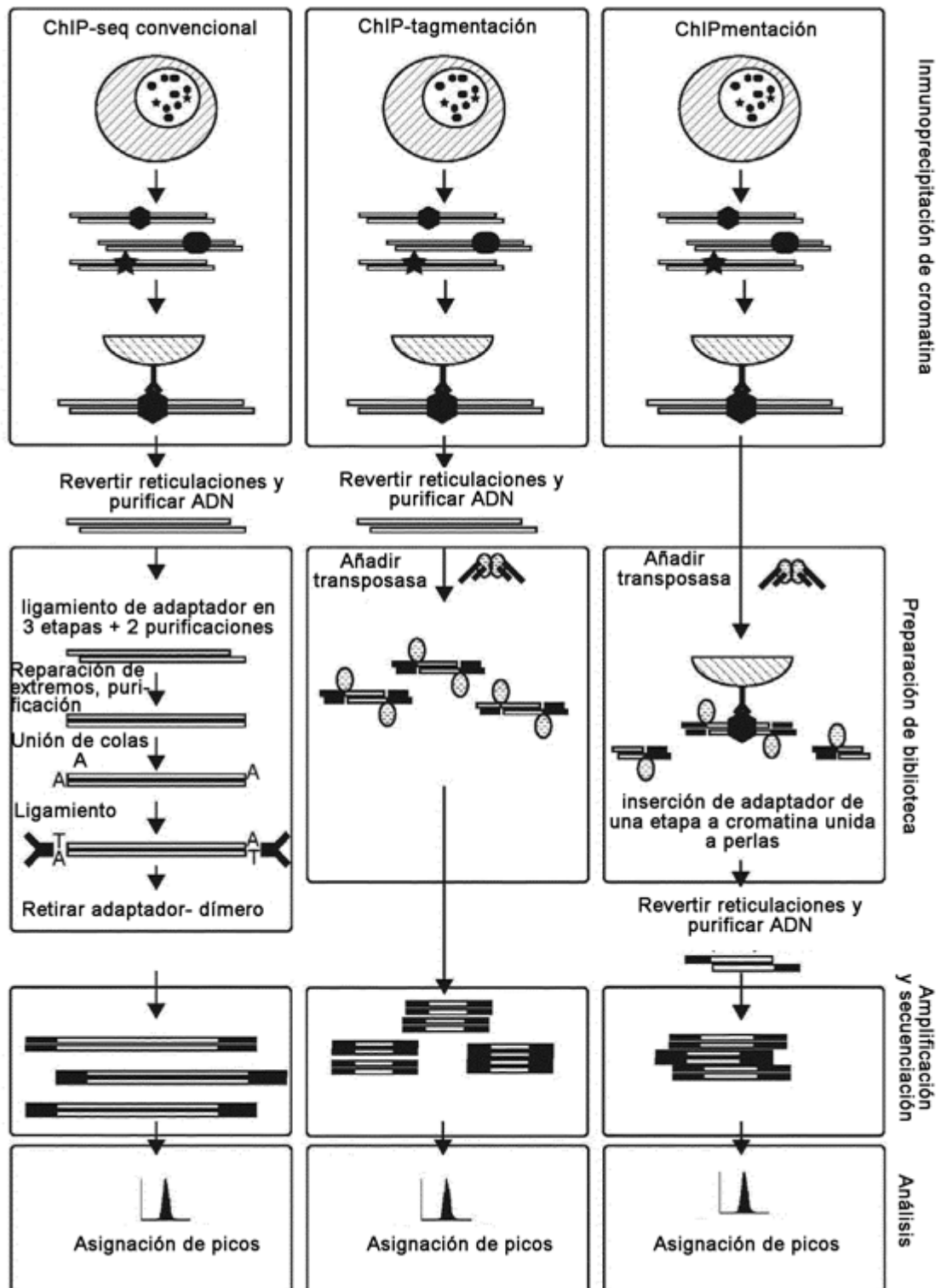
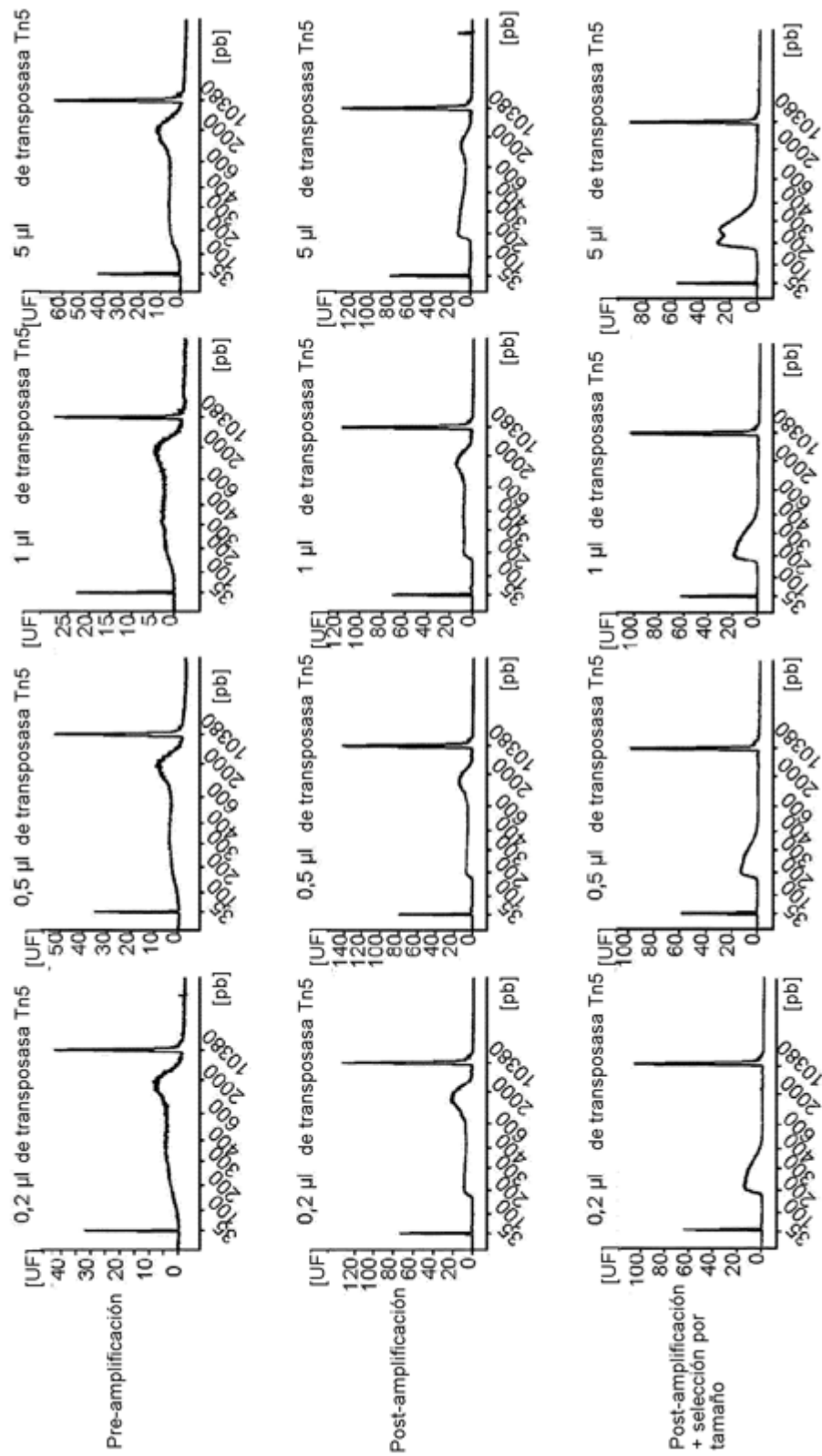


Figura 3

Distribución de tamaño de fragmentos antes y después de PCR de bibliotecas de ChIPmentación de H3K4me3 preparadas con diferentes cantidades de enzima transposasa Tn5



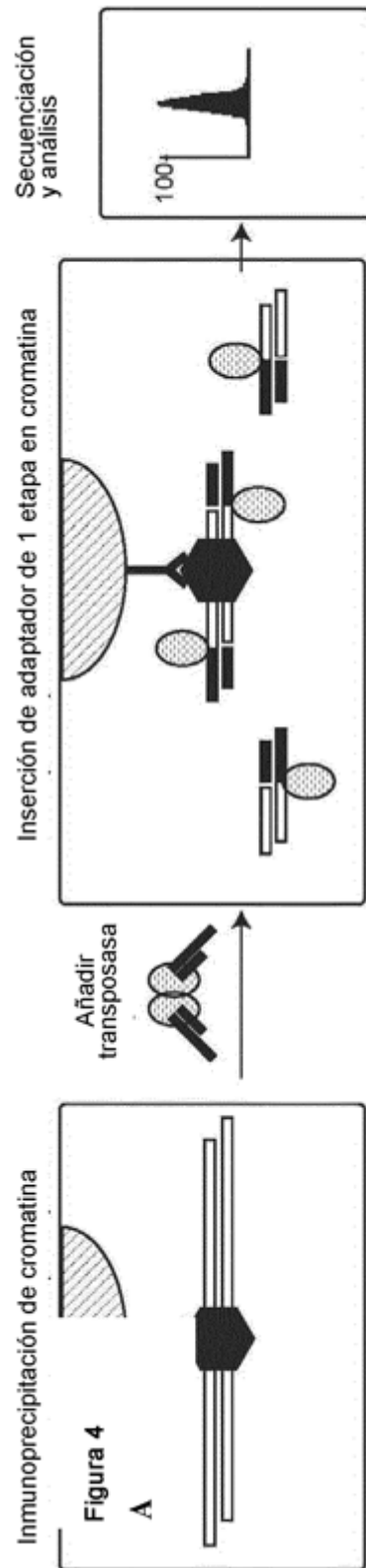


Figura 4 cont.

B

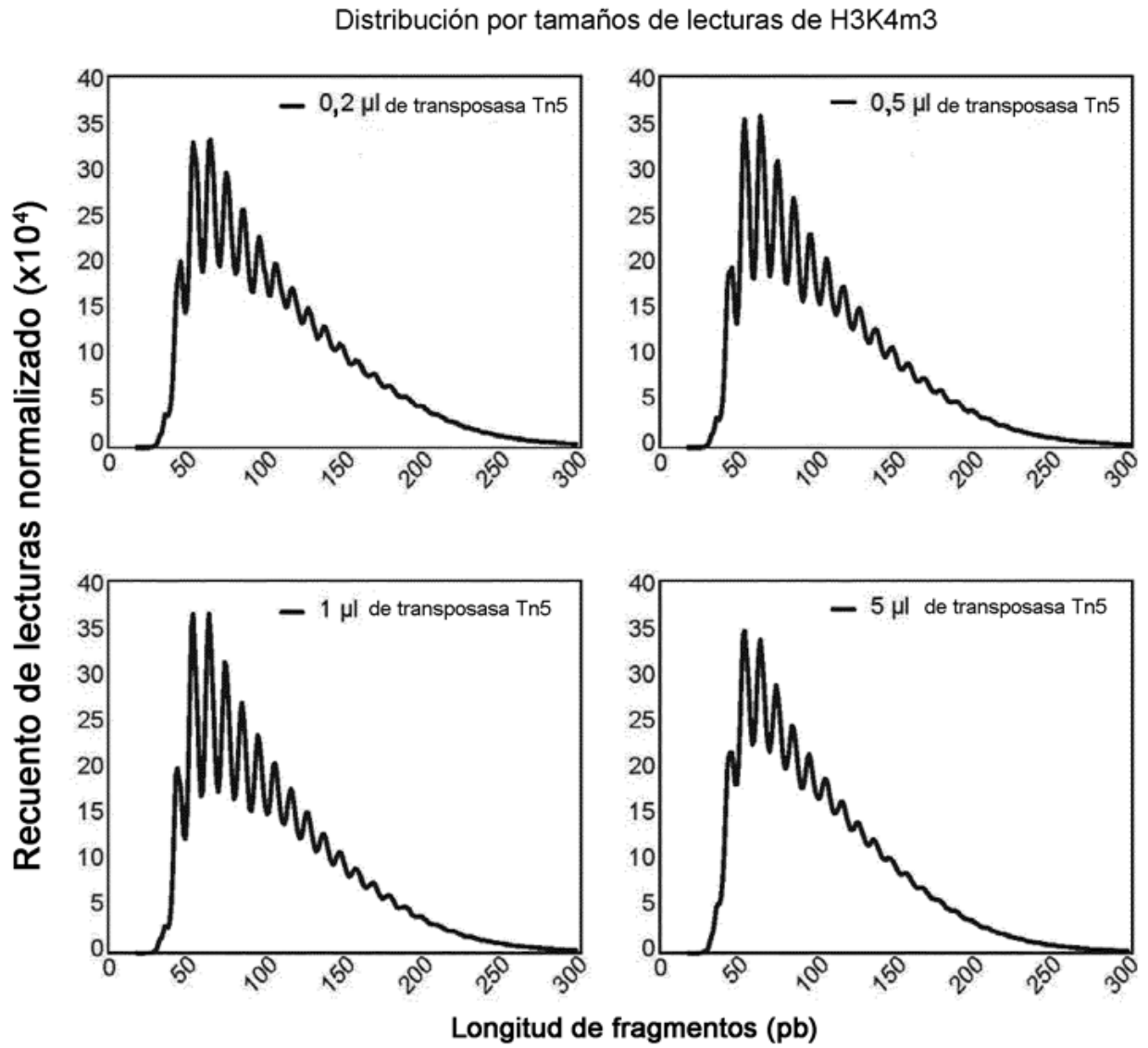
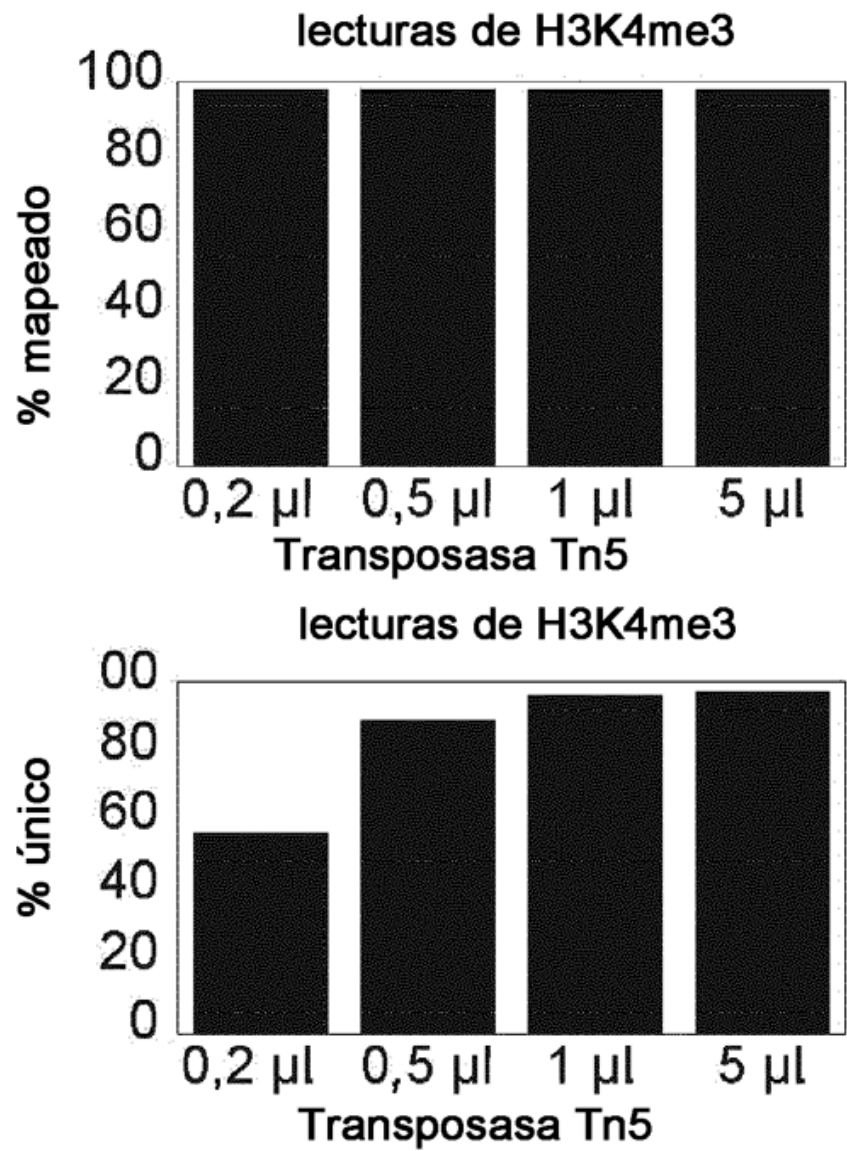


Figura 4 cont.

C



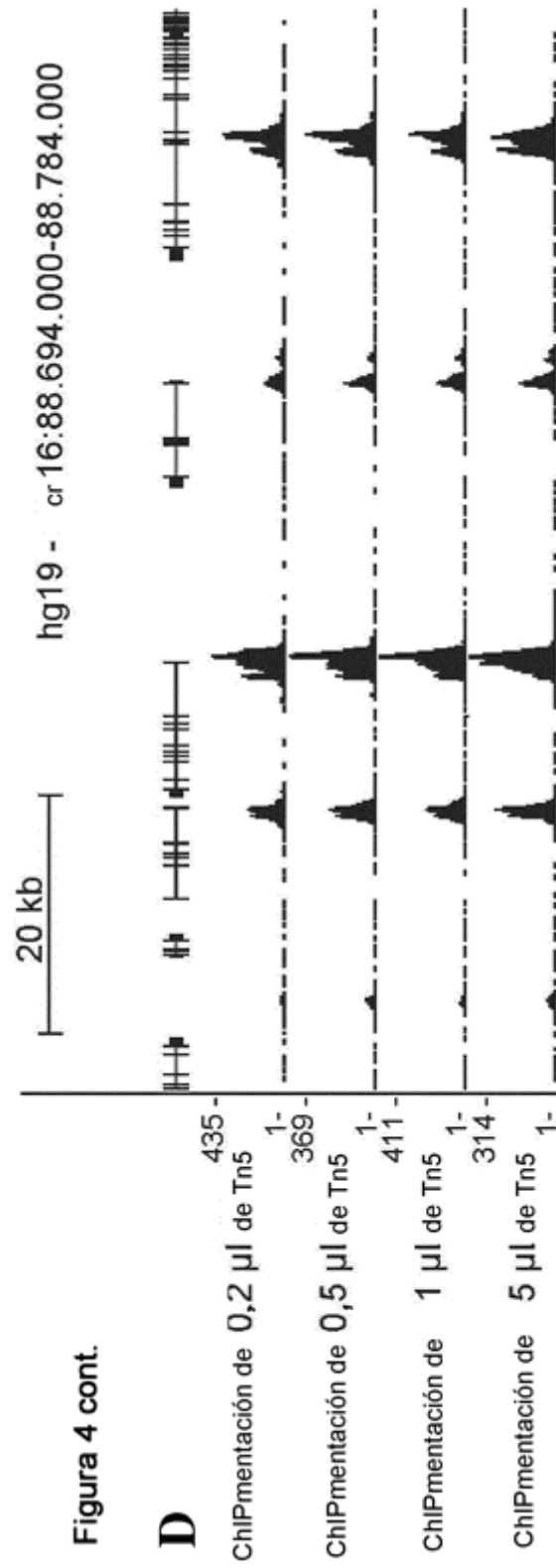


Figura 4 cont.

Figura 4 cont.
E

ChIPmentación de H3K4me3


	1	0,98	0,98	0,98	0,2 µl de Tn5
	0,98	1	0,99	0,99	5 µl de Tn5
	0,98	0,99	1	0,99	0,5 µl de Tn5
	0,98	0,99	0,99	1	1 µl de Tn5
	0,2 µl de Tn5	5 µl de Tn5	0,5 µl de Tn5	1 µl de Tn5	

Figura 4 cont.

F

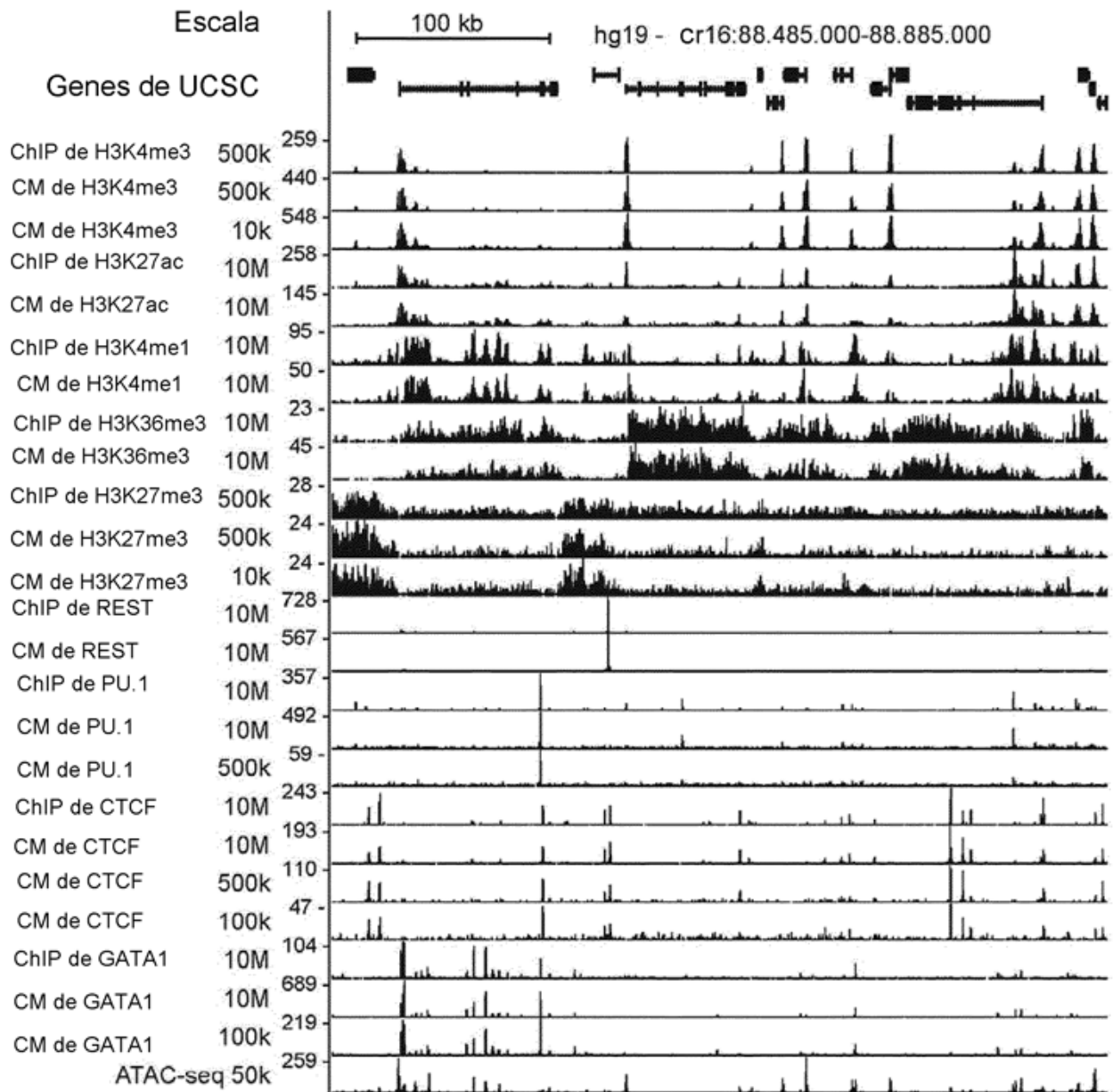


Figura 4 cont.

G

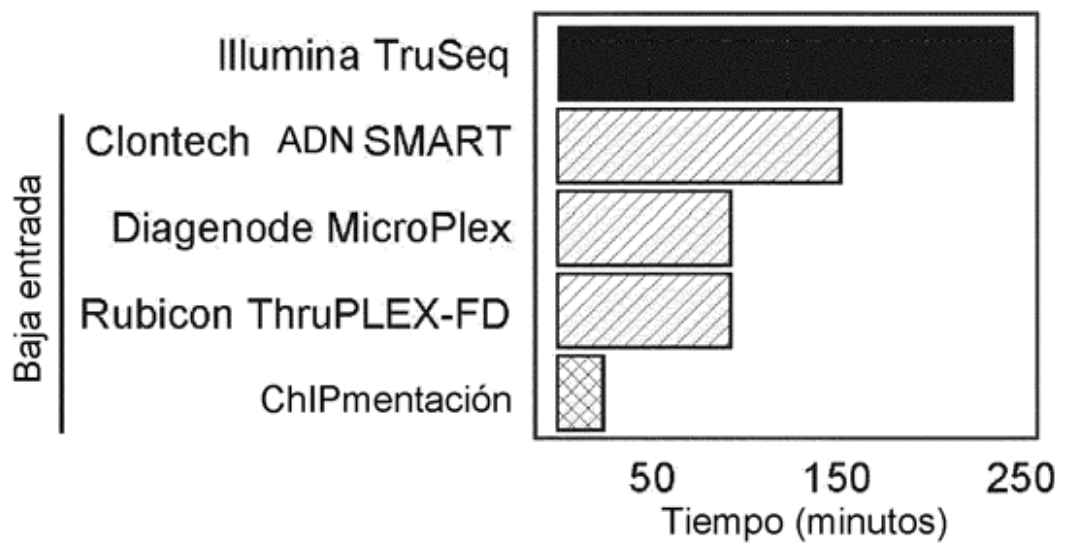
K27ac		K4me3		K27me3		K4me1		K36me3											
1	0,91	0,89	0,72	0,71	0,75	0,76	0,78	0,76	0,35	0,37	0,37	0,36	0,36	0,51	0,38	0,35	0,33	0,24	0,18
0,91	1	0,94	0,59	0,58	0,58	0,64	0,62	0,66	0,21	0,21	0,22	0,22	0,21	0,21	0,54	0,43	0,37	0,28	0,2
0,89	0,94	1	0,58	0,57	0,59	0,64	0,64	0,66	0,19	0,19	0,19	0,19	0,19	0,19	0,53	0,41	0,36	0,27	0,19
0,72	0,59	0,58	1	0,98	0,95	0,97	0,94	0,94	0,75	0,75	0,76	0,76	0,76	0,76	0,57	0,51	0,54	0,56	0,51
0,71	0,58	0,57	0,98	1	0,94	0,96	0,93	0,96	0,76	0,76	0,75	0,75	0,74	0,74	0,54	0,54	0,59	0,57	0,58
0,75	0,58	0,59	0,95	0,94	1	0,96	0,98	0,94	0,65	0,65	0,68	0,68	0,67	0,67	0,47	0,39	0,42	0,49	0,41
0,76	0,64	0,64	0,97	0,96	0,96	1	0,97	0,98	0,63	0,63	0,64	0,64	0,64	0,64	0,52	0,45	0,48	0,49	0,43
0,78	0,62	0,64	0,94	0,93	0,98	0,97	1	0,96	0,56	0,56	0,58	0,58	0,57	0,57	0,44	0,37	0,39	0,43	0,37
0,76	0,66	0,66	0,94	0,96	0,94	0,98	0,96	1	0,56	0,56	0,55	0,55	0,54	0,54	0,46	0,45	0,48	0,44	0,45
0,35	0,21	0,19	0,75	0,76	0,65	0,63	0,56	0,56	1	1	0,97	0,97	0,97	0,97	0,54	0,57	0,64	0,66	0,69
0,35	0,21	0,19	0,75	0,76	0,65	0,63	0,56	0,56	1	1	0,97	0,97	0,97	0,97	0,54	0,55	0,63	0,66	0,68
0,37	0,22	0,19	0,76	0,75	0,68	0,64	0,58	0,55	0,97	0,97	1	1	0,99	0,99	0,56	0,51	0,56	0,67	0,6
0,37	0,22	0,19	0,76	0,75	0,68	0,64	0,58	0,55	0,97	0,97	1	1	0,98	0,98	0,56	0,5	0,56	0,67	0,6
0,36	0,21	0,19	0,76	0,74	0,67	0,64	0,57	0,54	0,97	0,97	0,99	0,98	1	1	0,54	0,47	0,53	0,64	0,56
0,36	0,21	0,19	0,76	0,74	0,67	0,64	0,57	0,54	0,97	0,97	0,99	0,98	1	1	0,54	0,47	0,52	0,63	0,56
0,51	0,54	0,53	0,57	0,54	0,47	0,52	0,44	0,46	0,54	0,54	0,56	0,56	0,54	0,54	1	0,86	0,77	0,58	0,46
0,38	0,43	0,41	0,51	0,54	0,39	0,45	0,37	0,45	0,57	0,55	0,51	0,5	0,47	0,47	0,86	1	0,95	0,6	0,71
0,35	0,37	0,36	0,54	0,59	0,42	0,48	0,39	0,48	0,64	0,63	0,56	0,56	0,53	0,52	0,77	0,95	1	0,63	0,82
0,33	0,28	0,27	0,56	0,57	0,49	0,49	0,43	0,44	0,66	0,66	0,67	0,67	0,64	0,63	0,58	0,6	0,63	1	0,8
0,24	0,2	0,19	0,51	0,58	0,41	0,43	0,37	0,45	0,69	0,68	0,6	0,6	0,56	0,56	0,46	0,71	0,82	0,8	1
0,18	0,17	0,17	0,41	0,49	0,32	0,34	0,29	0,38	0,58	0,57	0,48	0,48	0,44	0,43	0,39	0,67	0,78	0,78	0,98
ChIP 10M (1)																			
CM 10M (1)																			
CM 10M (2)																			
CM 500k (1)																			
CM 500k (2)																			
ChIP 500k (1)																			
ChIP 500k (2)																			
CM 10k (1)																			
CM 10k (2)																			
CM 500k (1)																			
CM 500k (2)																			
ChIP 500k (1)																			
ChIP 500k (2)																			
CM 10k (1)																			
CM 10k (2)																			
ChIP 10M (1)																			
CM 10M (1)																			
CM 10M (2)																			
ChIP 10M (1)																			
CM 10M (1)																			
CM 10M (2)																			

Figura 4 cont.

H

	Réplica 1 frente a 2		CM frente a ChIP
	Números de células altos	Números de células bajos	
CTCF	0,94	0,98	0,96
GATA1	0,99	0,90	0,80
PU1	0,96	0,97	0,94
REST	0,97	0,98	0,96

I



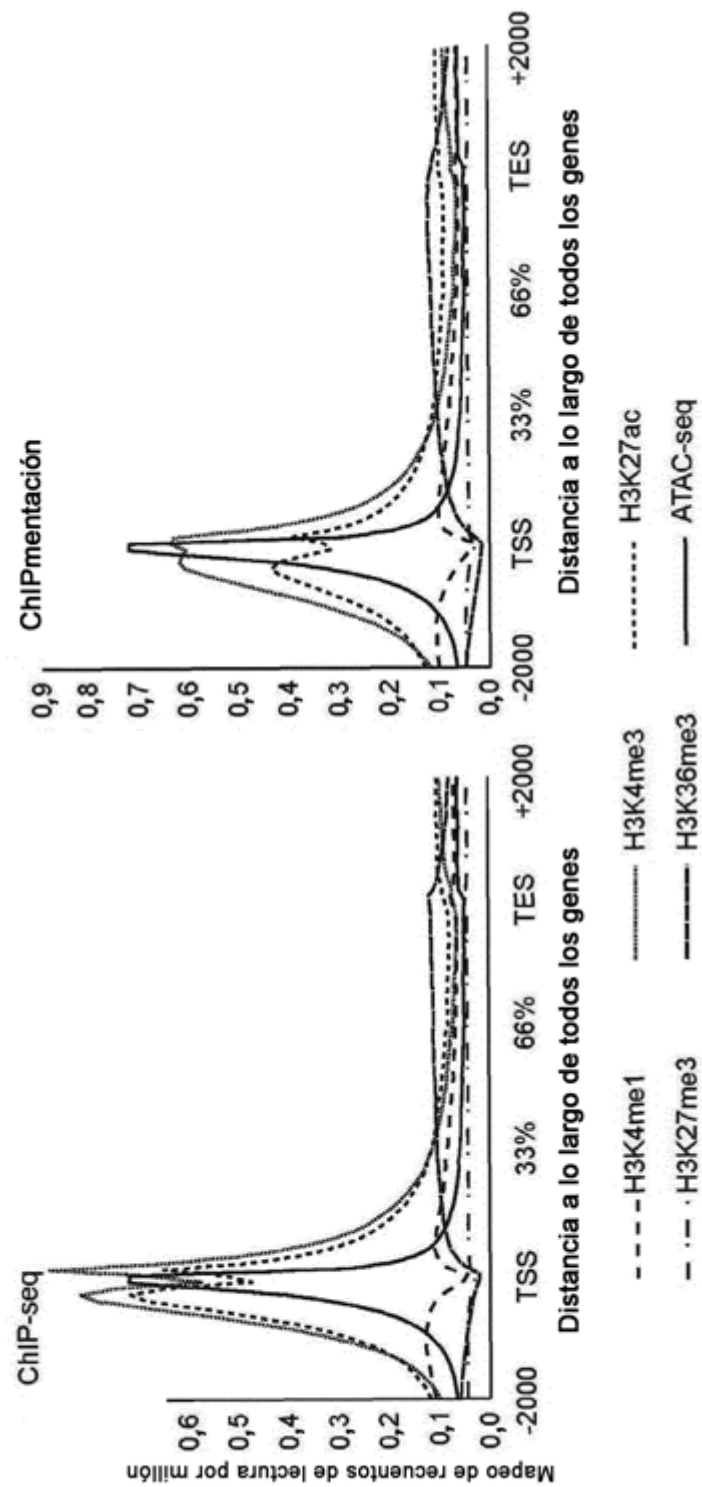
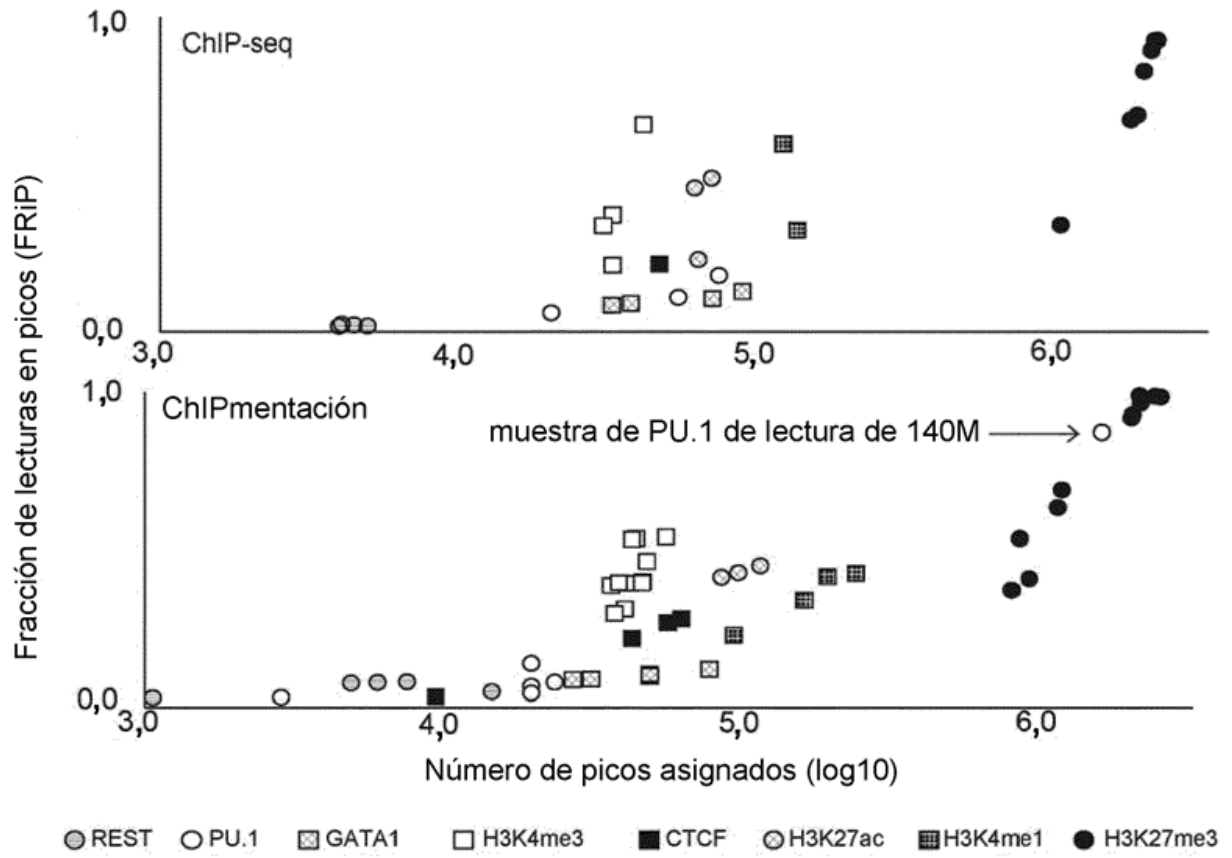


Figura 5
A

Figura 5 cont.

B



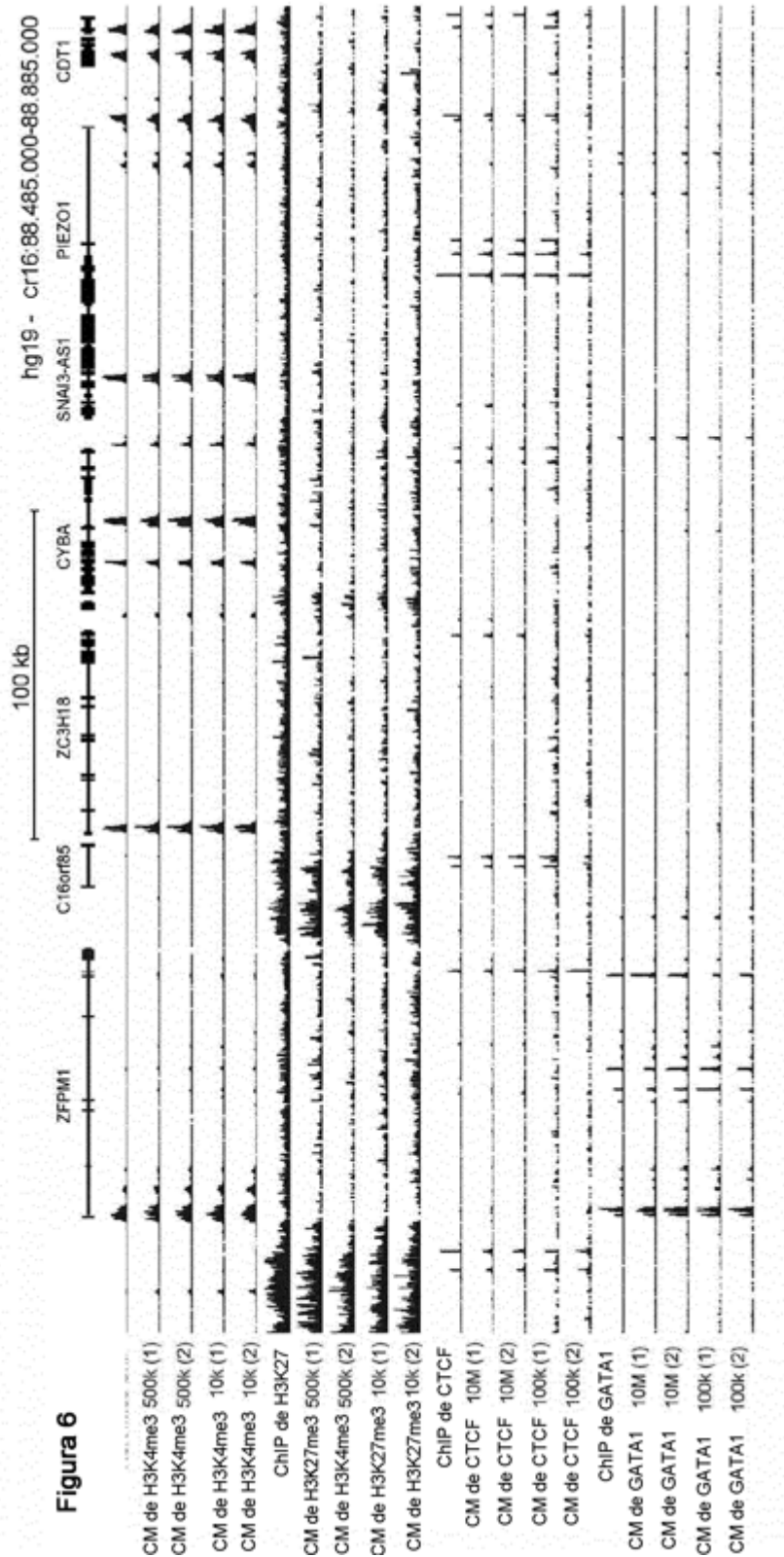


Figura 7

Solapamiento de picos

CM frente a ChIP			Réplica 1 frente a 2				Números de células		
	5% superior	25% superior	números de células altos	números de células altos	números de células bajos	números de células bajos	altos frente a bajos	altos frente a bajos	altos frente a bajos
	5% superior	25% superior	5% superior	25% superior	5% superior	25% superior	5% superior	25% superior	5% superior
CTCF	100 %	100 %	100 %	99 %	91 %	74 %	98 %	88 %	
GATA1	100 %	98 %	99 %	94 %	90 %	71 %	98 %	83 %	
PU1	100 %	92 %	99 %	95 %	100 %	92 %	100 %	92 %	
REST	100 %	91 %	100 %	99 %	82 %	58 %	96 %	67 %	

Figura 8

id	tipo de célula	cantidad de entrada	protocolo	anticuerpo	réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	tasa única	id. de muestra interna
1	K562	Células 10M	ChIP-seq	H3K4me1	1	HiSeq 2000	lectura individual	50	24.523.943	hg19	99	94	K562_10M_ChIP-seq_H3K4ME1_nan_1_hg19
2	K562	Células 10M	ChIP-seq	H3K4me1	2	HiSeq 2000	lectura individual	50	25.448.971	hg19	99	97	K562_10M_ChIP-seq_H3K4ME1_nan_2_hg19
3	K562	Células 500k	ChIP-seq	H3K4me3	1	HiSeq 2000	lectura individual	50	18.610.366	hg19	99	94	K562_500K_ChIP-seq_H3K4ME3_nan_1_hg19
4	PBMC	Células 500k	ChIP-seq	H3K4me3	1	HiSeq 2000	lectura individual	50	39.543.850	hg19	86	31	PBMC_nan_ChIP-seq_H3K4ME3_nan_1_hg19
5	K562	Células 500k	ChIP-seq	H3K4me3	2	HiSeq 2000	lectura individual	50	37.576.419	hg19	99	99	K562_500K_ChIP-seq_H3K4ME3_nan_2_hg19
6	K562	Células 10M	ChIP-seq	H3K27ac	1	HiSeq 2000	lectura individual	50	27.503.214	hg19	98	45	K562_10M_ChIP-seq_H3K27AC_nan_1_hg19
7	K562	Células 500k	ChIP-seq	H3K27ac	1	HiSeq 2000	lectura individual	50	7.714.041	hg19	98	95	K562_500K_ChIP-seq_H3K27AC_nan_1_hg19
8	K562	Células 500k	ChIP-seq	H3K27me3	1	HiSeq 2000	lectura individual	50	33.279.554	hg19	98	96	K562_500K_ChIP-seq_H3K27ME3_nan_1_hg19
9	K562	Células 500k	ChIP-seq	H3K27me3	2	HiSeq 2000	lectura individual	50	31.605.383	hg19	98	95	K562_500K_ChIP-seq_H3K27ME3_nan_2_hg19
10	K562	Células 10M	ChIP-seq	H3K36me3	1	HiSeq 2000	lectura individual	50	24.453.845	hg19	98	96	K562_10M_ChIP-seq_H3K36ME3_nan_1_hg19
11	K562	Células 10M	ChIP-seq	CTCF	1	HiSeq 2000	lectura individual	50	16.731.630	hg19	96	80	K562_10M_ChIP-seq_CTCF_nan_1_hg19

id	tipo de célula	cantidad de entrada	protocolo	anticuerpo	réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	tasa única	id. de muestra interna
12	K562	Células 10M	ChIP-seq	CTCF	2	HiSeq 2000	lectura individual	50	13.292.391	hg19	96	65	K562_10M_ChIP-seq_CTCF_nan_2_hg19
13	K562	Células 10M	ChIP-seq	GATA1	1	HiSeq 2000	lectura individual	50	20.908.847	hg19	98	59	K562_10M_ChIP-seq_GATA1_nan_1_hg19
14	K562	Células 10M	ChIP-seq	GATA1	2	HiSeq 2000	lectura individual	50	17.734.841	hg19	98	92	K562_10M_ChIP-seq_GATA1_nan_2_hg19
15	K562	Células 10M	ChIP-seq	PU.1	1	HiSeq 2000	lectura individual	50	13.663.300	hg19	97	93	K562_10M_ChIP-seq_PU1_nan_1_hg19
16	K562	Células 10M	ChIP-seq	PU.1	2	HiSeq 2000	lectura individual	50	13.453.300	hg19	97	82	K562_10M_ChIP-seq_PU1_nan_2_hg19
17	K562	Células 10M	ChIP-seq	REST	1	HiSeq 2000	lectura individual	50	20.078.470	hg19	98	80	K562_10M_ChIP-seq_REST_nan_1_hg19
18	K562	Células 10M	ChIP-seq	REST	2	HiSeq 2000	lectura individual	50	18.994.605	hg19	98	91	K562_10M_ChIP-seq_REST_nan_2_hg19
19	K562	Células 10M	ChIP-seq	Inmunoglobulina G	1	HiSeq 2000	lectura individual	50	31.818.701	hg19	95	59	K562_10M_ChIP-seq_IGG_nan_1_hg19
20	K562	Células 500k	ChIP-seq	Inmunoglobulina G	1	HiSeq 2000	lectura individual	50	32.545.198	hg19	98	86	K562_500K_ChIP-seq_IGG_nan_1_hg19
21	K562	Células 10M	ChIP-seq	Inmunoglobulina G	2	HiSeq 2000	lectura individual	50	12.999.527	hg19	93	34	K562_10M_ChIP-seq_IGG_nan_2_hg19
22	K562	Células 500k	ChIP-seq	Inmunoglobulina G	2	HiSeq 2000	lectura individual	50	36.366.641	hg19	90	59	K562_500K_ChIP-seq_IGG_nan_2_hg19
23	K562	Células 10M	ChIP-seq	Inmunoglobulina G	3	HiSeq 2000	lectura individual	50	14.298.197	hg19	96	38	K562_10M_ChIP-seq_IGG_nan_3_hg19
24	PBMC	Células 500k	ChIP-seq	Entrada	1	HiSeq 2000	lectura individual	50	4.674.578	hg19	93	96	PBMC_nan_ChIP-seq_Input_nan_1_hg19
25	K562	Células 10M	ChIPmentación	H3K4me1	1	HiSeq 2000	lectura individual	50	14.683.831	hg19	97	87	K562_10M_ChIPmentation_H3K4ME1_nan_1_hg19

id	tipo de célula	cantidad de entrada	protocolo	anticuerpo	réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	tasa única	id. de muestra interna
26	K562	Células 10M	ChIPmentación	H3K4me1	1	HiSeq 2000	extremos apareados	100	25.372.361	hg19	98	91	K562_10M_ChIPmentación_H3K4ME1_nan_1_hg19
27	K562	Células 10M	ChIPmentación	H3K4me1	1	HiSeq 2000	extremos apareados	100	17.989.151	hg19	98	92	K562_10K_ChIP_Tagmentación_H3K4ME1_nan_1_hg19
28	K562	Células 10M	ChIPmentación	H3K4me1	2	HiSeq 2000	lectura individual	50	13.216.471	hg19	97	89	K562_10M_ChIPmentación_H3K4ME1_nan_2_hg19
29	K562	Células 500k	ChIPmentación	H3K4me3	1	HiSeq 2000	lectura individual	50	28.139.774	hg19	98	79	K562_500K_ChIPmentación_H3K4ME1_nan_1_hg19
30	K562	Células 500k	ChIPmentación	H3K4me3	2	HiSeq 2000	lectura individual	50	14.306.608	hg19	99	87	K562_500K_ChIPmentación_H3K4ME3_nan_2_hg19
31	K562	Células 10k	ChIPmentación	H3K4me3	3	HiSeq 2000	lectura individual	50	37.048.339	hg19	98	34	K562_10K_ChIPmentación_H3K4ME3_nan_3_hg19
32	K562	Células 10k	ChIPmentación	H3K4me3	4	HiSeq 2000	lectura individual	50	35.357.419	hg19	97	40	K562_10K_ChIPmentación_H3K27ME3_nan_4_hg19
33	K562	Células 10M	ChIPmentación	H3K27ac	1	HiSeq 2000	lectura individual	50	19.377.579	hg19	98	83	K562_10K_ChIPmentación_H3K27AC_nan_1_hg19
34	K562	Células 10M	ChIPmentación	H3K27ac	2	HiSeq 2500	lectura individual	50	10.356.848	hg19	99	92	K562_10M_ChIPmentación_H3K27AC_nan_2_hg19
35	K562	Células 10M	ChIPmentación	H3K27me3	1	HiSeq 2000	lectura individual	50	23.965.157	hg19	98	86	K562_10K_ChIPmentación_H3K27ME3_nan_1_hg19
36	K562	Células 500k	ChIPmentación	H3K27me3	1	HiSeq 2000	lectura individual	50	22.064.368	hg19	97	86	K562_500K_ChIPmentación_H3K27ME3_nan_1_hg19

id	tipo de célula	cantidad de entrada	protocolo	anticuerpo	réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	tasa única	id. de muestra interna
37	K562	Células 10k	ChIPmentación	H3K27me3	2	HiSeq 2000	lectura individual	50	26.019.592	hg19	98	88	K562_10K_ChIPmentación_H3K27ME3_nan_2_hg19
38	K562	Células 500k	ChIPmentación	H3K27me3	2	HiSeq 2000	lectura individual	50	24.286.257	hg19	98	88	K562_500K_ChIPmentación_H3K27ME3_nan_2_hg19
39	K562	Células 10M	ChIPmentación	H3K36me3	1	HiSeq 2000	lectura individual	50	21.628.106	hg19	97	81	K562_10M_ChIPmentación_H3K36ME3_nan_1_hg19
40	K562	Células 10M	ChIPmentación	H3K36me3	2	HiSeq 2500	lectura individual	50	18.738.407	hg19	98	90	K562_10M_ChIPmentación_H3K36ME3_nan_2_hg19
41	K562	Células 10M	ChIPmentación	CTCF	1	HiSeq 2000	lectura individual	50	11.479.781	hg19	82	75	K562_100K_ChIPmentación_CTCF_nan_1_hg19
42	K562	Células 10M	ChIPmentación	CTCF	1	HiSeq 2000	lectura individual	50	11.830.997	hg19	98	93	K562_10M_ChIPmentación_CTCF_nan_1_hg19
43	K562	Células 500k	ChIPmentación	CTCF	1	HiSeq 2500	lectura individual	50	10.464.133	hg19	98	60	K562_500K_ChIPmentación_CTCF_nan_1_hg19
44	K562	Células 100k	ChIPmentación	CTCF	2	HiSeq 2500	lectura individual	50	10.725.153	hg19	98	64	K562_100K_ChIPmentación_CTCF_nan_2_hg19
45	K562	Células 10M	ChIPmentación	CTCF	2	HiSeq 2000	lectura individual	50	10.404.401	hg19	96	87	K562_10M_ChIPmentación_CTCF_nan_2_hg19
46	K562	Células 500k	ChIPmentación	CTCF	2	HiSeq 2000	lectura individual	50	10.772.645	hg19	97	92	K562_500K_ChIPmentación_CTCF_nan_2_hg19
47	K562	Células 100k	ChIPmentación	GATA1	1	HiSeq 2500	lectura individual	50	12.073.362	hg19	97	83	K562_100K_ChIPmentación_GATA1_nan_1_hg19
48	K562	Células 10M	ChIPmentación	GATA1	1	HiSeq 2000	lectura individual	50	18.086.796	hg19	97	90	K562_10M_ChIPmentación_GATA1_nan_1_hg19
49	K562	Células 100k	ChIPmentación	GATA1	2	HiSeq 2500	lectura individual	50	10.182.490	hg19	98	79	K562_100K_ChIPmentación_GATA1_nan_2_hg19

id	tipo de célula	cantidad de entrada	protocolo	anticuerpo	réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	tasa única	id. de muestra interna
50	K562	Células 10M	ChIPmentación	GATA1	2	HiSeq 2000	lectura individual	50	16.679.429	hg19	98	92	K562_10M_ChIPmentación_GATA1_nan_2_hg19
51	K562	Células 10M	ChIPmentación	PU.1	1	HiSeq 2000	lectura individual	50	65.539.103	hg19	98	84	K562_10M_ChIPmentación_PU1_nan_1_hg19
52	K562	Células 10M	ChIPmentación	PU.1	1	HiSeq 2000	extremos apareados	100	27.976.709	hg19	97	87	K562_10M_ChIPmentación_PU1_nan_1_hg19
53	K562	Células 500k	ChIPmentación	PU.1	1	HiSeq 2000	lectura individual	50	9.937.246	hg19	98	93	K562_500K_ChIPmentación_PU1_nan_1_hg19
54	K562	Células 10M	ChIPmentación	PU.1	1	HiSeq 2000	extremos apareados	100	13.476.614	hg19	96	95	K562_10M_ChIP-Tagmentation_PU1_nan_1_hg19
55	K562	Células 10M	ChIPmentación	PU.1	2	HiSeq 2000	lectura individual	50	13.231.293	hg19	98	93	K562_10M_ChIPmentación_PU1_nan_2_hg19
56	K562	Células 500k	ChIPmentación	PU.1	2	HiSeq 2000	lectura individual	50	4.121.056	hg19	96	95	K562_500K_ChIPmentación_PU1_nan_2_hg19
57	K562	Células 500k	ChIPmentación	PU.1	3	HiSeq 2500	lectura individual	50	10.438.095	hg19	98	87	K562_500K_ChIPmentación_PU1_nan_3_hg19
58	K562	Células 100k	ChIPmentación	REST	1	HiSeq 2500	lectura individual	50	10.260.884	hg19	94	91	K562_100K_ChIPmentación_REST_nan_1_hg19
59	K562	Células 10M	ChIPmentación	REST	1	HiSeq 2000	lectura individual	50	16.482.402	hg19	97	81	K562_10M_ChIPmentación_REST_nan_1_hg19
60	K562	Células 100k	ChIPmentación	REST	2	HiSeq 2500	lectura individual	50	10.233.072	hg19	92	83	K562_100K_ChIPmentación_REST_nan_2_hg19
61	K562	Células 10M	ChIPmentación	REST	2	HiSeq 2000	lectura individual	50	24.670.236	hg19	97	74	K562_10M_ChIPmentación_REST_nan_2_hg19
62	K562	Células 10M	ChIPmentación	Inmunoglobulina G	1	HiSeq 2000	lectura individual	50	13.561.441	hg19	98	86	K562_10M_ChIPmentación_IGG_nan_1_hg19
63	K562	Células 500k	ChIPmentación	Inmunoglobulina G	1	HiSeq 2000	extremos apareados	100	14.672.785	hg19	96	57	K562_500K_ChIPmentación_IGG_1ULTN5_1_hg19
64	K562	Células 500k	ChIPmentación	Inmunoglobulina G	1	HiSeq 2000	lectura individual	50	34.263.776	hg19	97	50	K562_500K_ChIPmentación_IGG_nan_1_hg19

id	tipo de célula	cantidad de entrada	protocolo	anticuerpo	réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	tasa única	id. de muestra interna
65	K562	Células 10M	ChIPmentación	Inmunoglobulina G	2	HiSeq 2000	lectura individual	50	12.858.450	hg19	96	74	K562_10M_ChIPmentación_IGG_nan_2_hg19
66	K562	Células 500k	ChIPmentación	Inmunoglobulina G	2	HiSeq 2000	lectura individual	50	18.412.113	hg19	96	49	K562_500K_ChIPmentación_IGG_nan_2_hg19
67	K562	Células 10k	ChIPmentación	Inmunoglobulina G	3	HiSeq 2000	lectura individual	50	17.612.836	hg19	94	22	K562_10K_ChIPmentación_IGG_nan_3_hg19
68	K562	Células 500k	ChIPmentación	Inmunoglobulina G	3	HiSeq 2000	lectura individual	50	7.474.797	hg19	83	65	K562_500K_ChIPmentación_IGG_nan_3_hg19
69	K562	Células 10k	ChIPmentación	Inmunoglobulina G	4	HiSeq 2000	lectura individual	50	45.937.406	hg19	92	8	K562_10K_ChIPmentación_IGG_nan_4_hg19
70	K562	Células 10M	ChIPmentación	Inmunoglobulina G	4	HiSeq 2000	lectura individual	50	5.670.213	hg19	91	69	K562_10M_ChIPmentación_IGG_nan_4_hg19
71	K562	Células 10k	ChIPmentación	Inmunoglobulina G	5	HiSeq 2000	lectura individual	50	12.200.893	hg19	87	28	K562_10K_ChIPmentación_IGG_nan_5_hg19
72	K562	Células 10M	ChIPmentación	Inmunoglobulina G	5	HiSeq 2000	lectura individual	50	10.716.675	hg19	97	67	K562_10M_ChIPmentación_IGG_nan_5_hg19
73	K562	Células 500k	ChIPmentación 0,1 µl de Tn5	Entrada	1	HiSeq 2000	extremos apareados	100	13.668.707	hg19	97	92	K562_500K_ChIP-Tagmentation_Input_01_ULTN5_1_hg19
74	K562	Células 500k	ChIPmentación 0,1 µl de Tn5	H3K4me3	1	HiSeq 2000	extremos apareados	100	11.615.337	hg19	98	91	K562_500K_ChIP-Tagmentation_H3K4ME3_01ULTN5_1_hg19
75	K562	Células 500k	ChIPmentación 0,2 µl de Tn5	H3K4me3	1	HiSeq 2000	extremos apareados	100	13.425.746	hg19	98	57	K562_500K_ChIPmentación_H3K4ME3_02ULTN5_1_hg19
76	K562	Células 500k	ChIPmentación 0,5 µl de Tn5	H3K4me3	1	HiSeq 2000	extremos apareados	100	13.052.416	hg19	98	89	K562_500K_ChIPmentación_H3K4ME3_05ULTN5_1_hg19
77	K562	Células 500k	ChIPmentación 1 µl de Tn5	H3K4me3	1	HiSeq 2000	extremos apareados	100	12.767.587	hg19	98	96	K562_500K_ChIPmentación_H3K4ME3_1ULTN5_1_hg19

id	tipo de célula	canti- dad de entrada	protocolo	anticuerpo	Réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	ta- sa uni- ca	id. de muestra interna
78	K562	Células 500k	ChIPmentación 5 µl de Tn5	H3K4me3	1	HiSeq 2000	extremos apareados	100	11.743.183	hg19	98	97	K562_500K_ChIPmentación_H3K4ME3_5ULTN5_1_hg19
79	PBMC	100 pg de ADN	ChIPmentación	Entrada	1	HiSeq 2000	lectura individual	50	18.751.591	hg19	96	73	PBMC_100pg DNA_ChIP-Tagmentation_Input_nan_1_hg19
80	PBMC	10 pg de ADN	ChIPmentación	Entrada	1	HiSeq 2000	lectura individual	50	9.031.248	hg19	90	99	PBMC_10pg DNA_ChIP-Tagmentation_Input_nan_1_hg19
81	PBMC	2 pg de ADN	ChIPmentación	Entrada	1	HiSeq 2000	lectura individual	50	12.031.802	hg19	90	99	PBMC_2pg DNA_ChIP-Tagmentation_Input_nan_1_hg19
82	PBMC	2 pg de ADN	ChIPmentación	H3K4me3	1	HiSeq 2000	lectura individual	50	11.960.898	hg19	90	99	PBMC_2pg DNA_ChIP-Tagmentation_H3K4ME3_nan_1_hg19
83	PBMC	100 pg de ADN	ChIPmentación	H3K4me3	1	HiSeq 2000	lectura individual	50	3.626.047	hg19	99	99	PBMC_100pg DNA_ChIP-Tagmentation_H3K4ME3_nan_1_hg19
84	PBMC	100 pg de ADN	ChIPmentación	H3K4me3	1	HiSeq 2000	lectura individual	50	8.393.968	hg19	43	90	PBMC_100pg DNA_ChIP-Tagmentation_H3K4ME3_nan_1_hg19
85	PBMC	10 pg de ADN	ChIPmentación	H3K4me3	1	HiSeq 2000	lectura individual	50	11.741.222	hg19	92	99	PBMC_10pg DNA_ChIP-Tagmentation_H3K4ME3_nan_1_hg19
86	K562	Células 50k	ATAC-seq	N/A	1	HiSeq 2000	lectura individual	50	51.962.519	hg19	98	34	K562_50K_ATAC-seq_nan_nan_1_hg19
87	K562	Células 50k	ATAC-seq	N/A	2	HiSeq 2000	lectura individual	50	32.581.411	hg19	98	66	K562_50K_ATAC-seq_nan_nan_1_hg19

id	tipo de célula	cantidad de entrada	protocolo	anticuerpo	réplica	modelo de secuenciador	tipo de lecturas	longitud de lecturas	recuento de lecturas	genoma	tasa de alineamiento	Tasa única	id. de muestra interna
													seq_nan_nan_2_hg19

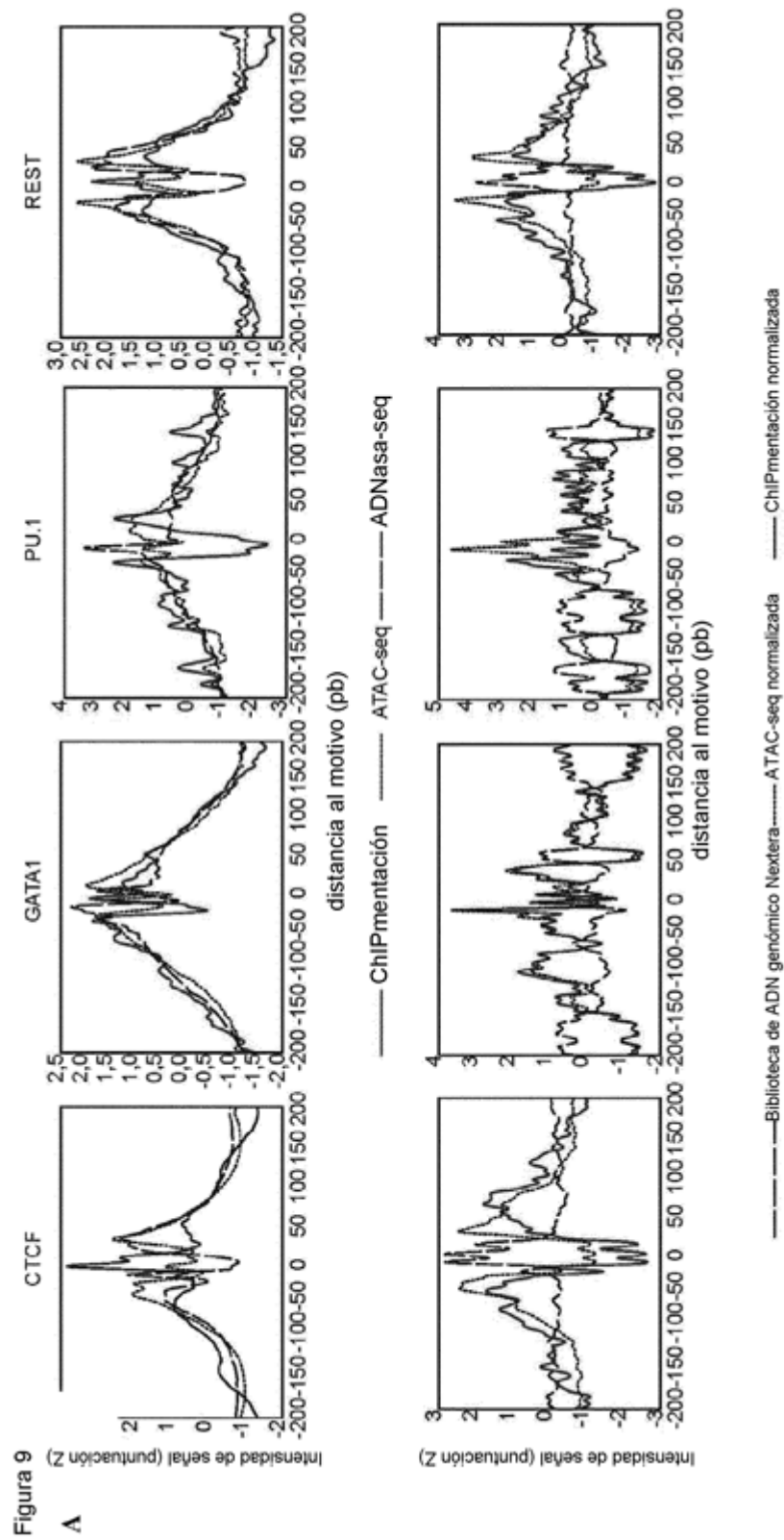
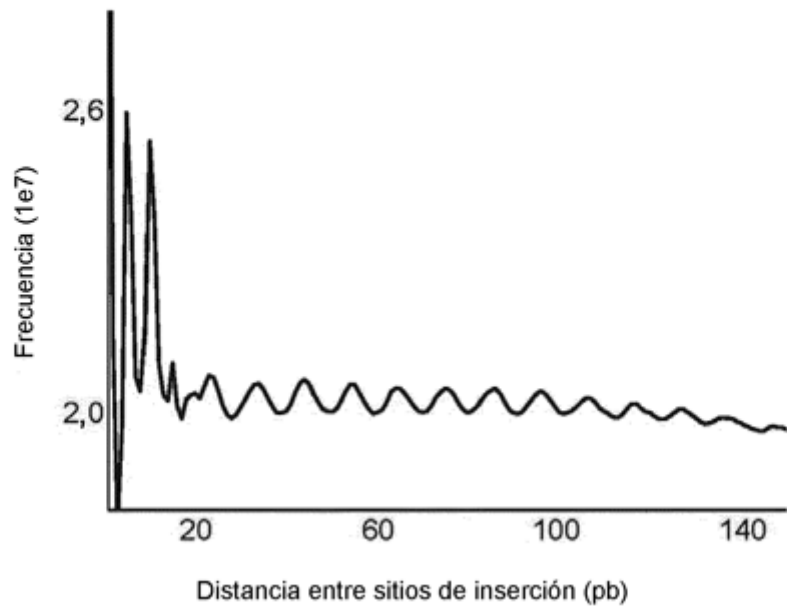


Figura 9 cont.

B



C

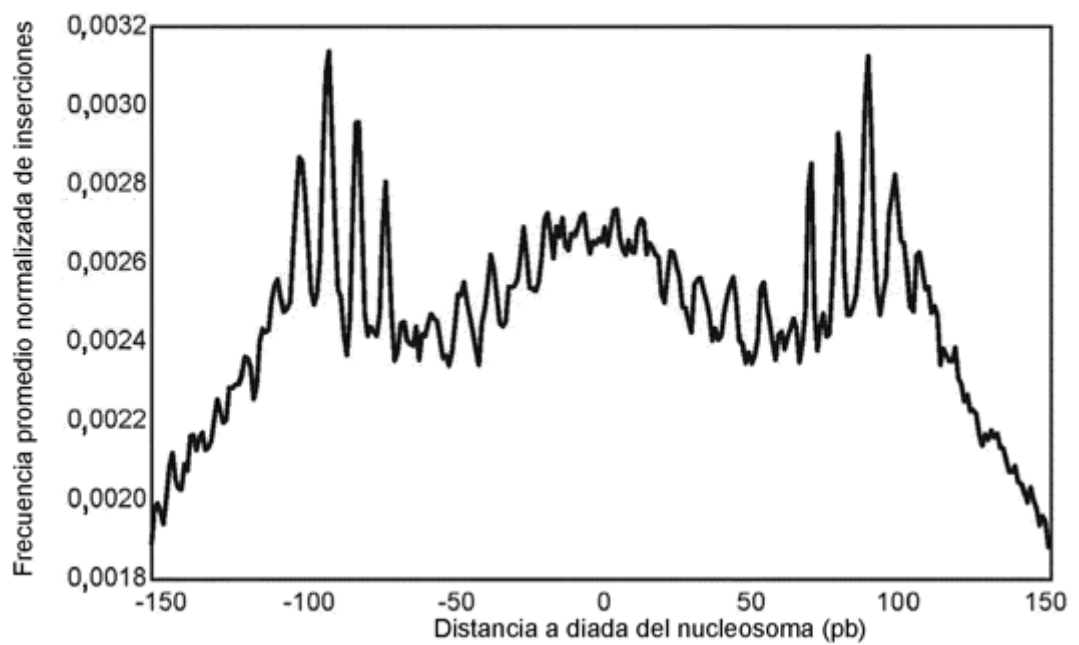


Figura 10

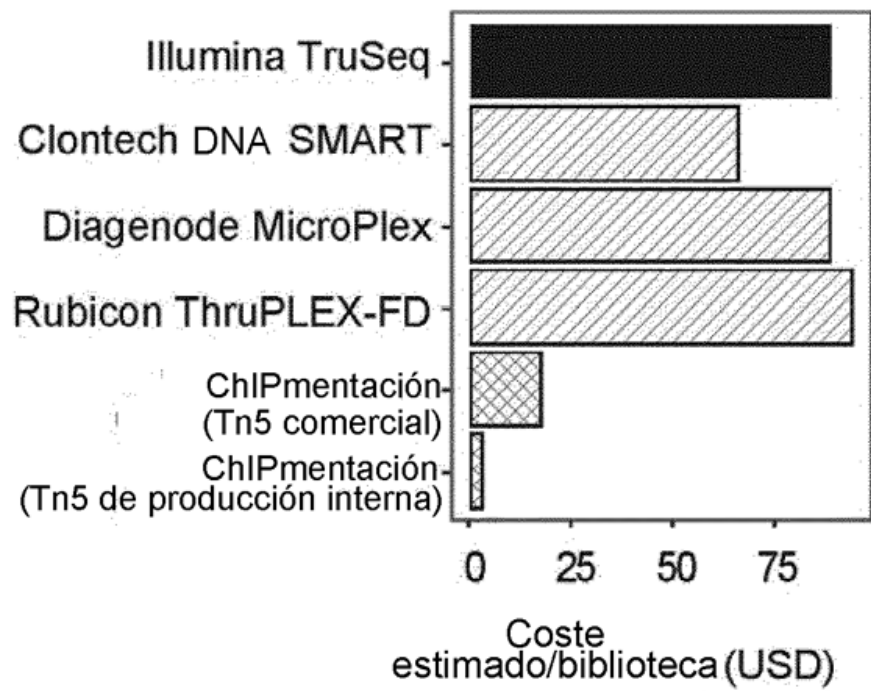


Figura 11 A

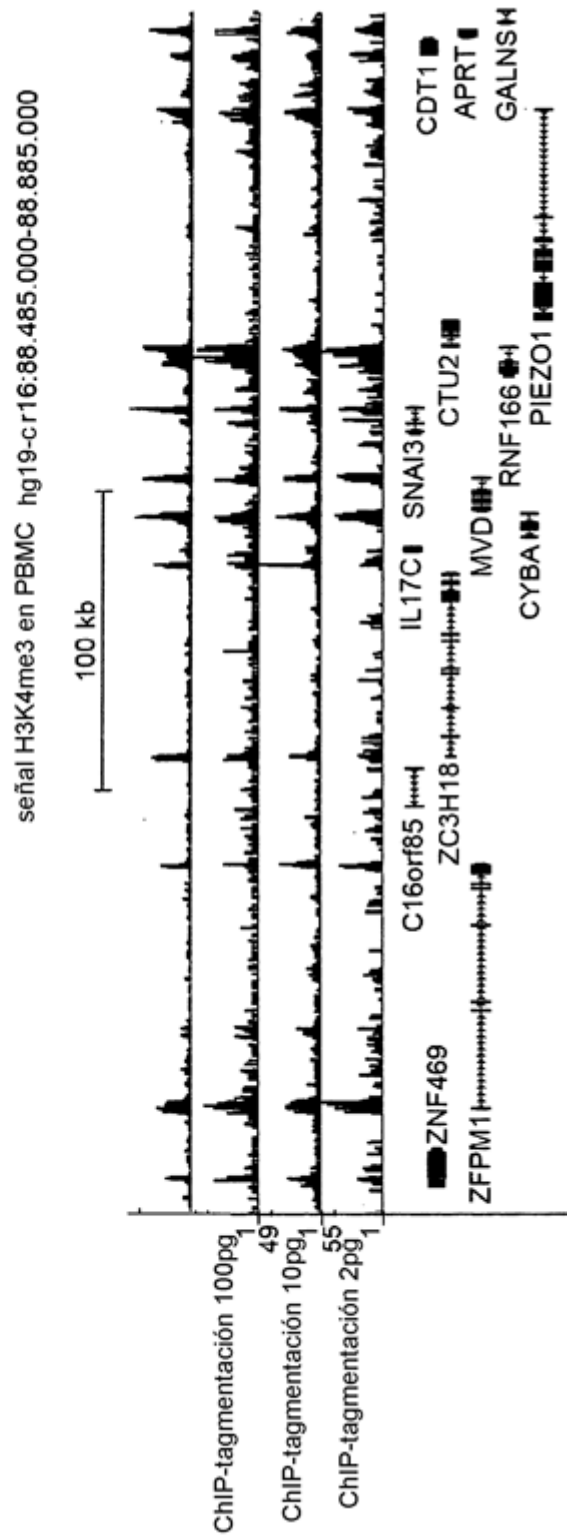


Figura 11 cont.

B

Correlaciones de secuenciación de H3K4me3 a partir de ChIP-seq convencional y ChIP-tagmentación con cantidades diferentes de ADN de entrada

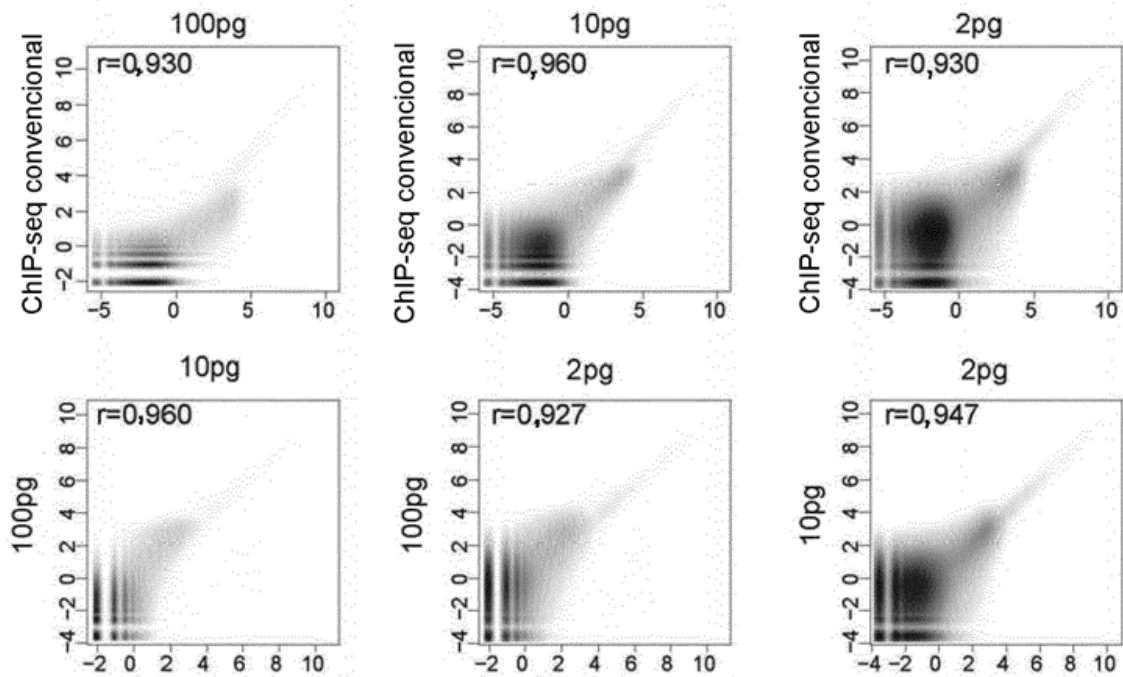


Figura 12

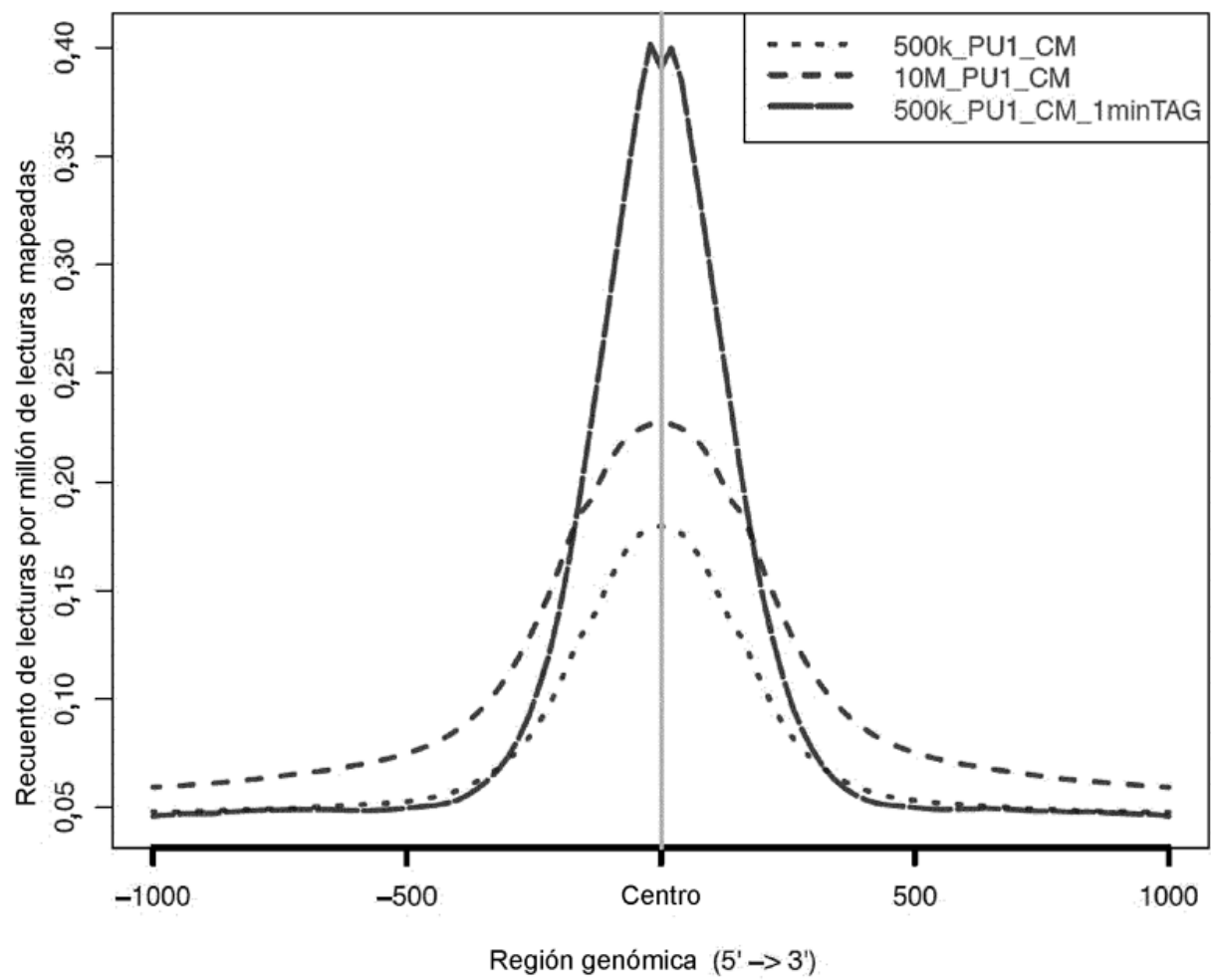


Figura 13

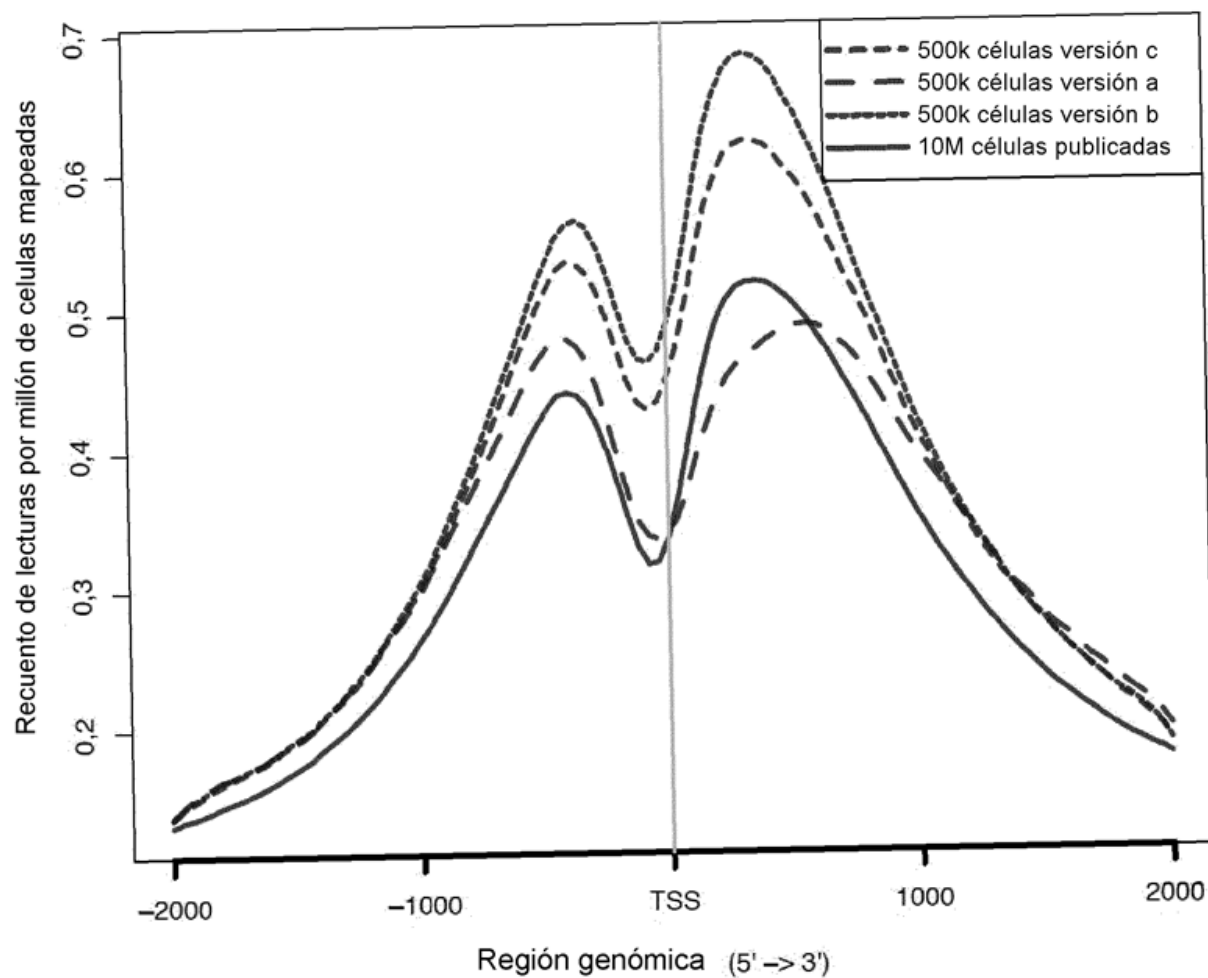


Figura 14

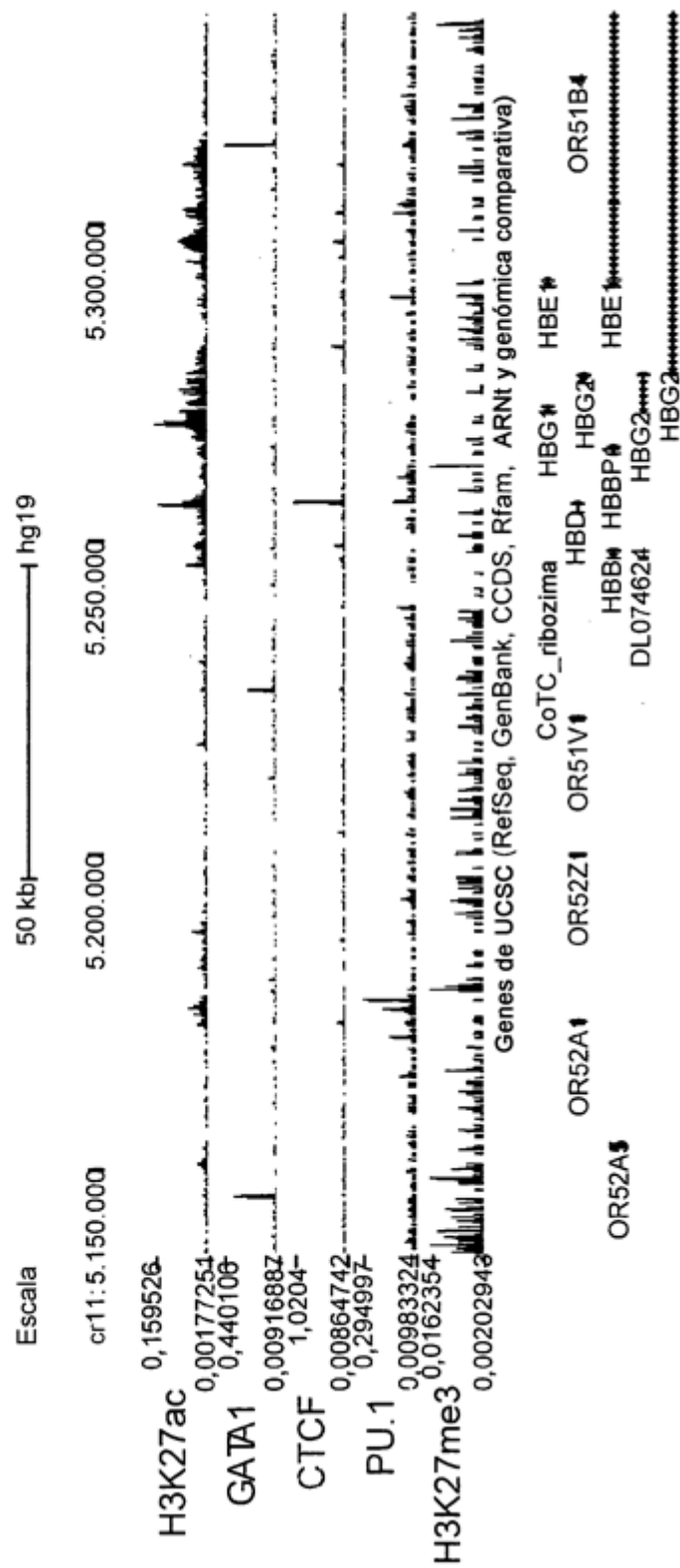


Figura 15

