

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 993 724**

51 Int. Cl.:

**C12Q 1/6806** (2008.01)

**C12Q 1/6841** (2008.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.01.2022** **PCT/US2022/011663**

87 Fecha y número de publicación internacional: **04.08.2022** **WO22164615**

96 Fecha de presentación y número de la solicitud europea: **07.01.2022** **E 22702549 (1)**

97 Fecha y número de publicación de la concesión europea: **14.08.2024** **EP 4284942**

54 Título: **Método para el marcaje espacial mediado por transposasa y el análisis de ADN genómico en una muestra biológica**

30 Prioridad:

**29.01.2021 US 202163143438 P**

**26.03.2021 US 202163166708 P**

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

**07.01.2025**

73 Titular/es:

**10X GENOMICS, INC. (100.00%)**

**6230 Stoneridge Mall Road**

**Pleasanton, CA 94588-3260, US**

72 Inventor/es:

**STAHL, PATRIK;**

**MARKLUND, MAJA;**

**LLORENS BOBADILLA, ENRIC y**

**FRISEN, JONAS**

74 Agente/Representante:

**CARVAJAL Y URQUIJO, Isabel**

### Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 993 724 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

Método para el marcaje espacial mediado por transposasa y el análisis de ADN genómico en una muestra biológica

Campo de la invención

5 La presente descripción se refiere a los métodos para determinar espacialmente la accesibilidad al ADN en una muestra biológica.

Antecedentes

10 Las células dentro de un tejido tienen diferencias en la morfología y/o función celular debido a niveles variados de analitos (por ejemplo, expresión génica y/o proteica) dentro de las diferentes células. La posición específica de una célula dentro de un tejido (por ejemplo, la posición de la célula con relación a las células vecinas o la posición de la célula con relación al microentorno del tejido) puede afectar, por ejemplo, la morfología, diferenciación, destino, viabilidad, proliferación, comportamiento, señalización y diafonía de la célula con otras células en el tejido.

15 La heterogeneidad espacial se ha estudiado previamente mediante el uso de técnicas que típicamente proporcionan datos para unos pocos analitos en el contexto de tejido intacto o una porción de un tejido (por ejemplo, una sección de tejido), o proporcionan datos significativos de analitos de células aisladas, individuales, pero no proporcionan información con respecto a la posición de las células aisladas de la muestra biológica de origen (por ejemplo, un tejido).

20 La estructura de la cromatina puede ser diferente entre las células en una muestra biológica o entre las muestras biológicas del mismo tejido. Las diferencias en los ensayo de la cromatina accesible pueden ser indicativas de secuencias transcripcionalmente activas, por ejemplo, genes, en una célula particular. Comprender mejor las regiones transcripcionalmente activas dentro de la cromatina permitirá la identificación de cuáles genes contribuyen a la función y/o fenotipo de una célula.

Resumen

La invención se expone en el conjunto de reivindicaciones adjuntas.

25 La presente descripción generalmente describe métodos para analizar espacialmente el ADN genómico presente en una muestra biológica.

30 Se han desarrollado métodos para estudiar epigenomas, por ejemplo, ensayos de accesibilidad a la cromatina (ATAC-seq) o identificar proteínas asociadas con la cromatina, por ejemplo, (ChIP-seq). Estos ensayos ayudan a identificar reguladores (por ejemplo, los reguladores cis y/o reguladores trans) que contribuyen a fenotipos celulares dinámicos. Aunque ATAC-seq y ChIP-seq han sido invaluable para definir la variabilidad epigenética dentro de una población celular, las aplicaciones convencionales de estos métodos se limitan en su capacidad de resolver espacialmente los genes asociados que promueven la variación celular. Los métodos espaciales ya se conocen, sin embargo, todavía se necesitan métodos adicionales y/o alternativos.

35 Por lo tanto, la presente descripción se refiere generalmente al marcaje y análisis espacial de los ácidos nucleicos. En algunas modalidades, se proporcionan en la presente descripción métodos que utilizan un transposoma para fragmentar el ADN genómico (por ejemplo, la cromatina abierta, cromatina accesible) y para capturar el ADN fragmentado en una matriz espacial, de esta manera se revelan conocimientos epigenómicos con respecto a las características estructurales que contribuyen a la regulación celular dentro del contexto espacial de una muestra biológica.

40 En la presente descripción se proporcionan métodos para determinar la accesibilidad al ADN genómico, el método incluye: (a) una muestra biológica en una matriz que incluye una pluralidad de sondas de captura, donde una sonda de captura de la pluralidad de sondas de captura incluye: (i) un código de barras espacial y (ii) un dominio de captura; (b) poner en contacto una pluralidad de oligonucleótidos de férula con la muestra biológica, donde un oligonucleótido de férula se hibrida con el dominio de captura; (c) poner en contacto un transposoma con la muestra biológica para insertar las secuencias de extremos del transposón en el ADN genómico accesible, de esta manera se genera ADN genómico fragmentado; (d) hibridar el ADN genómico fragmentado al oligonucleótido de férula y ligar el ADN genómico fragmentado a la sonda de captura; (e) liberar una o más secuencias de extremos del transposón no ligadas del ADN genómico fragmentado ligado; y (f) determinar (i) una secuencia del código de barras espacial o un complemento de la misma, y (ii) la totalidad o una porción de una secuencia del ADN genómico fragmentado, o un complemento de la misma, y usar las secuencias determinadas de (i) y (ii) para determinar la accesibilidad al ADN genómico en la muestra biológica.

55 En algunas modalidades, la matriz incluye uno o más elementos. En algunas modalidades, uno o más elementos incluyen una perla.

En algunas modalidades, la sonda de captura incluye además un dominio de escisión, uno o más dominios funcionales, un identificador molecular único o sus combinaciones.

En algunas modalidades, el método incluye una etapa de migración activa donde el ADN genómico fragmentado migra a la matriz por la aplicación de un campo eléctrico.

En algunas modalidades, la hibridación en la etapa (b) incluye hibridar el oligonucleótido de férula, o una porción del mismo, al dominio de captura, o una porción del mismo, de la sonda de captura. En algunas modalidades, la hibridación en la etapa (d) incluye hibridar el oligonucleótido de férula, o una porción de este, con una secuencia de extremo del transposón o una porción de este, de un ADN genómico fragmentado.

En algunas modalidades, la ligazón se realiza mediante el uso de una ADN ligasa.

En algunas modalidades, el método incluye extender un extremo 3' de la sonda de captura mediante el uso del ADN genómico fragmentado como una plantilla. En algunas modalidades, la etapa de extensión se realiza mediante el uso de una ADN polimerasa que tiene actividad de desplazamiento de la hebra.

En algunas modalidades, el método incluye realizar el relleno de espacios entre el oligonucleótido de férula y el ADN genómico fragmentado.

En algunas modalidades, el transposoma incluye una enzima transposasa, y donde la enzima transposasa es una enzima transposasa Tn5, una enzima transposasa Mu, una enzima transposasa Tn7, una transposasa de especies de *Vibrio*, o sus derivados funcionales. En algunas modalidades, la enzima transposasa Tn5 incluye una secuencia que es al menos un 80 % idéntica a la SEQ ID NO: 1.

En algunas modalidades, la determinación en la etapa (f) incluye la secuenciación (i) del código de barras espacial o un complemento del mismo, y (ii) la totalidad o una porción de la secuencia del ADN genómico fragmentado o un complemento de la misma y determinar además la ubicación del ADN genómico accesible en la muestra biológica.

En algunas modalidades, el método incluye obtener imágenes de la muestra biológica antes o después de ponerla en contacto con la matriz.

En algunas modalidades, la liberación en la etapa (d) incluye calentar la muestra biológica. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 65 °C a 85 °C. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 65 °C a aproximadamente 80 °C. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 75 °C.

En algunas modalidades, el método incluye teñir la muestra biológica. En algunas modalidades, la tinción incluye teñir con hematoxilina y eosina.

En algunas modalidades, poner en contacto el transposoma con la muestra biológica se realiza bajo una condición de permeabilización química, bajo una condición de permeabilización enzimática, o ambas. En algunas modalidades, la condición de permeabilización química incluye un detergente. En algunas modalidades, el detergente es uno o más de NP-40, Tween-20, Triton X-100 y Digitonina. En algunas modalidades, el detergente está a una concentración de aproximadamente 0,001 % (v/v) a aproximadamente 1,0 % (v/v).

En algunas modalidades, poner en contacto el transposoma con la muestra biológica se realiza después de una condición de prepermeabilización enzimática. En algunas modalidades, la condición de prepermeabilización enzimática incluye una proteasa. En algunas modalidades, la proteasa es una pepsina, una colagenasa, una proteinasa K y sus combinaciones. En algunas modalidades, la proteasa es una colagenasa.

También se proporcionan en la presente descripción métodos para determinar la accesibilidad al ADN genómico, el método incluye: (a) una muestra biológica en una matriz que incluye una pluralidad de sondas de captura, donde una sonda de captura de la pluralidad de sondas de captura incluye: (i) un código de barras espacial y (ii) un dominio de captura; (b) poner en contacto un transposoma con la muestra biológica para insertar secuencias de extremos del transposón en el ADN genómico accesible, de esta manera se genera ADN genómico fragmentado; (c) hibridar una secuencia de extremo del transposón del ADN genómico fragmentado al dominio de captura de la sonda de captura; (d) liberar secuencias de extremos del transposón no ligadas al dominio de captura; y (e) determinar (i) una secuencia del código de barras espacial o un complemento de la misma, y (ii) la totalidad o una porción de una secuencia del ADN genómico fragmentado, o un complemento de la misma, y usar las secuencias determinadas de (i) y (ii) para determinar la accesibilidad al ADN genómico en la muestra biológica.

En algunas modalidades, la matriz incluye uno o más elementos. En algunas modalidades, uno o más elementos incluyen una perla.

En algunas modalidades, la sonda de captura incluye además un dominio de escisión, uno o más dominios funcionales, un identificador molecular único o sus combinaciones.

En algunas modalidades, el método incluye además una etapa de migración activa donde el ADN genómico fragmentado migra a la matriz por la aplicación de un campo eléctrico.

En algunas modalidades, la hibridación en la etapa (c) incluye hibridar la secuencia de extremo del transposón, o una porción de esta, al dominio de captura, o una porción de esta, de la sonda de captura.

En algunas modalidades, el método incluye extender un extremo 3' de la sonda de captura mediante el uso del ADN genómico fragmentado como una plantilla. En algunas modalidades, la etapa de extensión se realiza mediante el uso de una ADN polimerasa que tiene actividad de desplazamiento de la hebra.

En algunas modalidades, el método incluye realizar el relleno de espacios entre la secuencia de extremo del transposón y el ADN genómico fragmentado.

En algunas modalidades, el transposoma incluye una enzima transposasa, y donde la enzima transposasa es una enzima transposasa Tn5, una enzima transposasa Mu, una enzima transposasa Tn7, una transposasa de especies de *Vibrio*, o sus derivados funcionales. En algunas modalidades, la enzima transposasa Tn5 incluye una secuencia que es al menos un 80 % idéntica a la SEQ ID NO: 1.

En algunas modalidades, la determinación en la etapa (e) incluye la secuenciación (i) de la secuencia del código de barras espacial o un complemento de la misma, y (ii) la totalidad o una porción de la secuencia del ADN genómico fragmentado o un complemento de la misma y determinar además la ubicación del ADN genómico accesible en la muestra biológica.

En algunas modalidades, el método incluye obtener imágenes de la muestra biológica antes o después de ponerla en contacto con la matriz.

En algunas modalidades, la liberación en la etapa (d) incluye calentar la muestra biológica. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 65 °C a 85 °C. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 65 °C a aproximadamente 80 °C. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 75 °C.

En algunas modalidades, el método incluye teñir la muestra biológica. En algunas modalidades, la tinción incluye teñir con hematoxilina y eosina.

En algunas modalidades, poner en contacto el transposoma con la muestra biológica se realiza después de una condición de permeabilización química, bajo una condición de permeabilización enzimática, o ambas. En algunas modalidades, la condición de permeabilización química incluye un detergente. En algunas modalidades, el detergente es uno o más de NP-40, Tween-20, Triton X-100 y Digitonina. En algunas modalidades, el detergente está a una concentración de aproximadamente 0,001 % (v/v) a aproximadamente 0,1 % (v/v). En algunas modalidades, el contacto del transposoma con la muestra biológica se realiza después de una condición de prepermeabilización enzimática. En algunas modalidades, la condición de prepermeabilización enzimática incluye una proteasa. En algunas modalidades, la proteasa es una pepsina, una collagenasa, una proteinasa K y sus combinaciones. En algunas modalidades, la proteasa es una collagenasa.

También se proporcionan en la presente descripción métodos para determinar la ubicación del ADN en una muestra biológica, el método incluye: (a) una muestra biológica en una matriz que incluye una pluralidad de sondas de captura, donde una sonda de captura de la pluralidad de sondas de captura incluye: (i) un código de barras espacial y (ii) un dominio de captura; (b) poner en contacto la muestra biológica con una proteasa, donde la proteasa es capaz de degradar una o más proteínas histonas, de esta manera libera el ADN; (c) poner en contacto un transposoma con la muestra biológica para insertar secuencias de extremos del transposón en el ADN genómico liberado, de esta manera se genera ADN genómico fragmentado; (d) hibridar una secuencia de extremo del transposón del ADN fragmentado al dominio de captura; (e) liberar secuencias de extremos del transposón no ligadas al dominio de captura; y (f) determinar (i) una secuencia del código de barras espacial o un complemento de la misma, y (ii) la totalidad o una porción de una secuencia del ADN, o un complemento del mismo, y usar las secuencias determinadas de (i) y (ii) para determinar la ubicación del ADN en la muestra biológica.

En algunas modalidades, la proteasa es capaz de degradar al menos una proteína histona enlazadora y al menos una proteína histona del núcleo de histonas en la muestra biológica. En algunas modalidades, la proteasa es capaz de degradar al menos una histona de cada familia de histonas del núcleo de histonas en la muestra biológica. En algunas modalidades, la proteasa es una serina proteasa, una aspartil proteasa, una enzima C1 de la familia de peptidasas, una proteasa que es inhibida por el inhibidor de diazometano Z-Phe-Phe-CHN(2) o el inhibidor de epóxido E-64, una proteasa lisosomal, una collagenasa o una enzima azurofílica. En algunas modalidades, la proteasa es una collagenasa.

En algunas modalidades, el dominio de captura incluye una secuencia homopolimérica. En algunas modalidades, el dominio de captura incluye una secuencia única.

En algunas modalidades, la sonda de captura incluye además un dominio de escisión, uno o más dominios funcionales, un identificador molecular único o sus combinaciones.

5 En algunas modalidades, el método incluye una etapa de migración activa donde el ADN genómico fragmentado migra a la matriz por la aplicación de un campo eléctrico.

10 En algunas modalidades, la hibridación en la etapa (d) incluye hibridar la secuencia de extremo del transposón, o una porción de esta, al dominio de captura, o una porción de esta, de la sonda de captura.

En algunas modalidades, el método incluye extender un extremo 3' de la sonda de captura mediante el uso del ADN genómico fragmentado como una plantilla. En algunas modalidades, la etapa de extensión se realiza mediante el uso de una ADN polimerasa que tiene actividad de desplazamiento de la hebra.

15 En algunas modalidades, el método incluye el relleno de espacios entre la secuencia de extremo del transposón y el ADN genómico fragmentado.

20 En algunas modalidades, el transposoma incluye una enzima transposasa, y donde la enzima transposasa es una enzima transposasa Tn5, una enzima transposasa Mu, una enzima transposasa Tn7, una transposasa de especies de *Vibrio*, o sus derivados funcionales. En algunas modalidades, la enzima transposasa Tn5 comprende una secuencia que es al menos un 80 % idéntica a la SEQ ID NO: 1.

25 En algunas modalidades, la determinación en la etapa (f) incluye la secuenciación (i) del código de barras espacial o un complemento del mismo, y (ii) la totalidad o una porción de la secuencia del ADN genómico fragmentado o un complemento de la misma.

En algunas modalidades, el método incluye además, obtener imágenes y/o teñir la muestra biológica. En algunas modalidades, la tinción incluye la tinción con hematoxilina y eosina.

30 En algunas modalidades, la proteasa se pone en contacto con la muestra biológica de aproximadamente 5 minutos a aproximadamente 15 minutos. En algunas modalidades, la proteasa se pone en contacto con la muestra biológica durante aproximadamente 10 minutos. En algunas modalidades, la proteasa se pone en contacto con la muestra biológica a una temperatura de aproximadamente 30 °C a aproximadamente 45 °C. En algunas modalidades, la proteasa se pone en contacto con la muestra biológica a una temperatura de aproximadamente 37 °C.

35 En algunas modalidades, la liberación en la etapa (d) incluye calentar la muestra biológica. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 65 °C a 85 °C.

En algunas modalidades, determinar la ubicación del ADN en una muestra biológica incluye además analizar espacialmente todo el genoma de la muestra biológica.

40 En algunas modalidades, la muestra biológica es una sección de tejido. En algunas modalidades, la sección de tejido es una sección de tejido fresco o congelado. En algunas modalidades, la sección de tejido es una sección de tejido fija. En algunas modalidades, la sección de tejido fijo es una sección de tejido fijo que se fija con formalina embebido en parafina, una sección de tejido fijo con acetona, una sección de tejido fijo con paraformaldehído o una sección de tejido fijo con metanol.

45 Cuando los valores se describen en términos de intervalos, debe entenderse que la descripción incluye la descripción de todos los posibles subintervalos dentro de tales intervalos, así como también valores numéricos específicos que caen dentro de tales intervalos independientemente de si se indica expresamente un valor numérico específico o un subintervalo específico.

50 El término "cada uno", cuando se usa en referencia a una recopilación de artículos, pretende identificar un artículo individual en la recopilación, pero no necesariamente se refiere a cada artículo en la recopilación, a menos que se indique expresamente de cualquier otra manera, o a menos que el contexto del uso indique claramente de cualquier otra manera.

Descripción de los dibujos

55 Las siguientes figuras ilustran ciertas modalidades de las características y ventajas de esta descripción. Estas modalidades no pretenden limitar el alcance de las reivindicaciones adjuntas de ninguna manera. Los símbolos de referencia en los dibujos acompañantes, los cuales ilustran una o más modalidades ilustrativas indican elementos similares.

La Figura 1 muestra una sonda de captura ilustrativa.

La Figura 2 muestra un ensayo espacial ilustrativo del flujo de trabajo de la cromatina accesible a la transposasa (spATAC).

Las Figuras 3A-B muestran A) la tinción con hematoxilina y eosina (HyE) y B) los patrones de expresión génica de un cerebro de ratón con una sección de tejido de xenoinjerto de glioma humano.

Las Figuras 4A-B muestran A) la tinción con HyE y B) los patrones de expresión génica de diferentes especies en una sección de tejido cerebral de ratón para probar la resolución espacial del flujo de trabajo descrito en la Figura 2.

Las Figuras 5A-B son gráficos de experimentos replicados que muestran el número de puntos (eje y) por el número de identificadores de molécula únicos (eje x) que se identifican en cada punto. Las imágenes correspondientes a los gráficos se muestran encima de cada gráfico.

Las Figuras 6A-B son gráficos ilustrativos que indican la recuperación de la periodicidad del nucleosoma al practicar los métodos descritos en la presente descripción; pb=pares de bases, FU=unidades fluorescentes.

Las Figuras 7A-H muestra dos secciones consecutivas de tejido de ratón inmunoteñidas con un anticuerpo SOX9 antes del flujo de trabajo de ATAC-seq espacial (Figuras 7A y 7E). Las Figuras 7B y 7F muestran el número total de fragmentos de ADN marcados que se capturan por punto. Las Figuras 7C y 7G son gráficos que muestran el enriquecimiento del sitio de inicio transcripcional (TSS) y la periodicidad del nucleosoma correspondiente al practicar los métodos descritos en la presente descripción (Figuras 7D y 7H).

La Figura 8 muestra el diagrama genómico de las densidades de lectura de ATAC-seq en un conjunto de datos de referencia de ratón (e13,5 parte superior) y un conjunto de datos de ATAC-seq espacial (e13,5, medio) de los embriones de ratón mostrados en las Figuras 7A y 7E. El enriquecimiento de la señal de ATAC-seq espacial y la designación de picos (parte inferior) muestran las posiciones coincidentes para el enriquecimiento de fragmentos.

Las Figuras 9A-B muestran un conglomerado no sesgado basado en los gráficos (Figura 9A) y la asignación de los conglomerados de cada punto en la sección de tejido (Figura 9B).

Las Figuras 10A-D muestra gráficos de UMAP (Figura 10A y 10C) que se colorean por la accesibilidad relativa de dos regiones génicas encontradas que son diferencialmente accesibles entre las regiones de la sección de tejido (Figura 10B y 10D).

Las Figuras 11A-F muestra el primer conglomerado de las regiones del panel basado en la expresión génica (Figura 11A). Los conglomerados espaciales se indican con números. Las Figuras 11B-11F muestran la accesibilidad de los genes marcadores para cada conglomerado en una sección adyacente.

La Figura 12 muestra el diagrama genómico del enriquecimiento de la señal de ATAC-seq espacial que muestran la accesibilidad de una de las regiones encontradas que son más accesibles en el conglomerado del tejido 7.

#### Descripción detallada

Las metodologías y composiciones de análisis espacial descritas en la presente descripción pueden proporcionar una gran cantidad de analito y/o datos de expresión para una variedad de analitos dentro de una muestra biológica con alta resolución espacial, lo que mantiene al mismo tiempo el contexto espacial nativo. Los métodos y composiciones de análisis espacial pueden incluir, por ejemplo, el uso de una sonda de captura que incluye un código de barras espacial (por ejemplo, una secuencia de ácido nucleico que proporciona información sobre la ubicación o posición de un analito dentro de una célula o una muestra de tejido (por ejemplo, una célula de mamífero o una muestra de tejido de mamífero) y un dominio de captura que es capaz de unirse a un analito (por ejemplo, una proteína y/o un ácido nucleico) producido por y/o presente en una célula. Los métodos y composiciones de análisis espacial también pueden incluir el uso de una sonda de captura que tiene un dominio de captura que captura un agente intermedio para la detección indirecta de un analito. Por ejemplo, el agente intermedio puede incluir una secuencia de ácido nucleico (por ejemplo, un código de barras) asociada con el agente intermedio. Por lo tanto, la detección del agente intermedio es indicativa del analito en la muestra de células o tejido.

Los aspectos no limitantes de las metodologías y composiciones de análisis espacial se describen en las patentes de Estados Unidos núms. 10,774,374, 10,724,078, 10,480,022, 10,059,990, 10,041,949, 10,002,316, 9,879,313, 9,783,841, 9,727,810, 9,593,365, 8,951,726, 8,604,182, 7,709,198, en las publicaciones de solicitud de patente de Estados Unidos núms. 2020/239946, 2020/080136, 2020/0277663, 2020/024641, 2019/330617, 2019/264268, 2020/256867, 2020/224244, 2019/194709, 2019/161796, 2019/085383, 2019/055594, 2018/216161, 2018/051322, 2018/0245142, 2017/241911, 2017/089811, 2017/067096, 2017/029875, 2017/0016053, 2016/108458, 2015/000854, 2013/171621, WO 2018/091676, WO 2020/176788, Rodriques y otros, Science 363(6434):1463-1467, 2019; Lee y otros, Nat. Protoc. 10(3):442-458, 2015; Trejo y otros, PLoS ONE 14(2):e0212031, 2019; Chen y otros, Science 348(6233):aaa6090, 2015; Gao y otros, BMC Biol. 15:50, 2017; y Gupta y otros, Nature Biotechnol. 36:1197-1202, 2018; la Guía del usuario de los kits de reactivos de expresión genética espacial de Visium (por ejemplo, Rev C, con fecha de junio de 2020) y/o la Guía del usuario de los kits de reactivos de optimización de tejidos espaciales de Visium (por ejemplo, Rev C, con fecha de julio de 2020), ambas disponibles en el sitio web de documentación de soporte de 10x Genomics y puede usarse en la presente descripción en cualquier combinación. El documento WO 2020/123309 describe un perfil resuelto espacialmente que se basa en el fragmento marcaje del ADN genómico accesible donde los extremos del transposón se hibridan con una férula en sondas de captura con código de barras espacial en una matriz. En la presente descripción se describen aspectos adicionales no limitantes de las metodologías y composiciones de análisis espacial.

Alguna terminología general que puede usarse en esta descripción puede encontrarse en la Sección (I)(b) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. Típicamente, un “código de barras” es un marcador o identificador que transmite o es capaz de transmitir información (por ejemplo, información sobre un analito en una muestra, una perla y/o una sonda de captura). Un código de barras puede ser parte de un analito o independiente de un analito. Un código de barras puede unirse a un analito. Un código de barras particular puede ser único con relación a otros códigos de barras. A los efectos de esta descripción, un “analito” puede incluir cualquier sustancia, estructura, resto o componente biológico a analizar. El término “diana” puede referirse de manera similar a un analito de interés.

Los analitos pueden clasificarse en términos generales en uno de dos grupos: analitos de ácidos nucleicos y analitos no ácidos nucleicos. Los ejemplos de analitos no ácidos nucleicos incluyen, pero sin limitarse a, lípidos, carbohidratos, péptidos, proteínas, glicoproteínas (unidas a N o unidas a O), lipoproteínas, fosfoproteínas, variantes específicas de proteínas fosforiladas o acetiladas, variantes de amidación de proteínas, variantes de hidroxilación de proteínas, variantes de metilación de proteínas, variantes de ubiquitinación de proteínas, variantes de sulfatación de proteínas, proteínas virales (por ejemplo, cápside viral, envoltura viral, cubierta viral, accesorio viral, glicoproteínas virales, pico viral, etc.), proteínas extracelulares e intracelulares, anticuerpos y fragmentos de unión a antígeno. En algunas modalidades, el(los) analito(s) puede(n) localizarse en ubicación(ones) subcelular(es), incluida(s), por ejemplo, orgánulos, por ejemplo, mitocondrias, aparato de Golgi, retículo endoplásmico, cloroplastos, vesículas endocíticas, vesículas exocíticas, vacuolas, lisosomas, etc. En algunas modalidades, el(los) analito(s) puede(n) ser péptidos o proteínas, incluidos sin limitación, anticuerpos y enzimas. Pueden encontrarse ejemplos adicionales de analitos en la Sección (I)(c) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm 2020/0277663. En algunas modalidades, un analito puede detectarse indirectamente, tal como mediante la detección de un agente intermedio, por ejemplo, un producto de ligazón o un agente de captura del analito (por ejemplo, un anticuerpo conjugado con oligonucleótido), tales como los descritos en la presente descripción.

Una “muestra biológica” se obtiene típicamente del sujeto para su análisis mediante el uso de cualquiera de una variedad de técnicas que incluyen, pero sin limitarse a, biopsia, cirugía y microscopía de captura láser (LCM), y generalmente incluye células y/u otro material biológico del sujeto. En algunas modalidades, una muestra biológica puede ser una sección de tejido. En algunas modalidades, una muestra biológica puede ser una muestra biológica fijada y/o teñida (por ejemplo, una sección de tejido fijada y/o teñida). Los ejemplos no limitantes de tinciones incluyen tinciones histológicas (por ejemplo, hematoxilina y/o eosina) y tinciones inmunológicas (por ejemplo, tinciones fluorescentes). En algunas modalidades, pueden obtenerse imágenes de una muestra biológica (por ejemplo, una muestra biológica fijada y/o teñida). Las muestras biológicas también se describen en la Sección (I)(d) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

En algunas modalidades, una muestra biológica se permeabiliza con uno o más reactivos de permeabilización. Por ejemplo, la permeabilización de una muestra biológica puede facilitar la captura del analito. Las condiciones y agentes de permeabilización ilustrativos se describen en la Sección (I)(d)(ii)(13) o la Sección de modalidades ilustrativas del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

Los métodos de análisis espacial basados en matrices implican la transferencia de uno o más analitos de una muestra biológica a una serie de características en un sustrato, donde cada característica está asociada con una ubicación espacial única en la matriz. El análisis subsecuente de los analitos transferidos incluye determinar la identidad de los analitos y la ubicación espacial de los analitos dentro de la muestra biológica. La ubicación espacial de un analito dentro de la muestra biológica se determina basándose en la característica a la que está unido el analito (por ejemplo, directa o indirectamente) en la matriz, y la ubicación espacial relativa de la característica dentro de la matriz.

Una “sonda de captura” se refiere a cualquier molécula capaz de capturar (directa o indirectamente) y/o marcar un analito (por ejemplo, un analito de interés) en una muestra biológica. En algunas modalidades, la sonda de captura es un ácido nucleico o un polipéptido. En algunas modalidades, la sonda de captura incluye un código de barras (por ejemplo, un código de barras espacial y/o un identificador molecular único (UMI)) y un dominio de captura. En algunas modalidades, una sonda de captura puede incluir un dominio de escisión y/o un dominio funcional (por ejemplo, un sitio de unión a cebador, tal como para secuenciación de nueva generación (NGS)). Ver, por ejemplo, la Sección (II)(b) (por ejemplo, las subsecciones (i)-(vi)) del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm 2020/0277663. La generación de sondas de captura puede lograrse mediante cualquier método apropiado, que incluye los descritos en la Sección (II)(d)(ii) del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

En algunas modalidades, puede detectarse más de un tipo de analito (por ejemplo, ácidos nucleicos y proteínas) de una muestra biológica (por ejemplo, simultánea o secuencialmente) mediante el uso de cualquier técnica en formato múltiple apropiada, tales como las descritas en la Sección (IV) del documento WO 2020/176788 y/o publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

En algunas modalidades, la detección de uno o más analitos (por ejemplo, analitos proteicos) puede realizarse mediante el uso de uno o más agentes de captura del analito. Como se usa en la presente, un “agente de captura del analito” se refiere a un agente que interactúa con un analito (por ejemplo, un analito en una muestra biológica) y con una sonda de captura (por ejemplo, una sonda de captura unida a un sustrato o una característica) para identificar el analito. En algunas modalidades, el agente de captura del analito incluye: (i) un resto de unión al analito (por ejemplo, que se une a un analito), por ejemplo, un anticuerpo o fragmento de unión a antígeno del mismo; (ii) código de barras del resto de unión al analito; y (iii) una secuencia de captura del analito. Como se usa en la presente, el término “código de barras del resto de unión al analito” se refiere a un código de barras que está asociado con, o de cualquier otra manera, identifica el resto de unión al analito. Como se usa en la presente, el término “secuencia de captura del analito” se refiere a una región o resto configurado para hibridarse, unirse, acoplarse o interactuar de cualquier otra manera con un dominio de captura de una sonda de captura. En algunos casos, un código de barras del resto de unión al analito (o una porción del mismo) puede eliminarse (por ejemplo, escindirse) del agente de captura del analito. Puede encontrarse una descripción adicional de los agentes de captura del analito en la Sección (II)(b)(ix) del documento WO 2020/176788 y/o en la Sección (II)(b)(viii) de la publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

Existen al menos dos métodos para asociar un código de barras espacial con una o más células vecinas, de manera que el código de barras espacial identifique la una o más células, y/o el contenido de la una o más células, como asociado con una ubicación espacial particular. Un método es promover analitos o sustitutos de analitos (por ejemplo, agentes intermedios) fuera de una célula y hacia una matriz con código de barras espacial (por ejemplo, incluyendo sondas de captura con código de barras espaciales). Otro método consiste en escindir sondas de captura con códigos de barras espaciales de una matriz y promover las sondas de captura con códigos de barras espaciales hacia y/o dentro o sobre la muestra biológica.

En algunos casos, las sondas de captura pueden configurarse para cebar, replicar y, en consecuencia, producir productos de extensión opcionalmente con códigos de barra a partir de una plantilla (por ejemplo, una plantilla de ADN o ARN, tal como un analito o un agente intermedio (por ejemplo, un producto de ligazón o un agente de captura de analito), o una porción del mismo), o derivados de los mismos (ver, por ejemplo, la Sección (II)(b)(vii) del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm 2020/0277663 con respecto a sondas de captura extendidas). En algunos casos, las sondas de captura pueden configurarse para formar productos de ligazón con una plantilla (por ejemplo, una plantilla de ADN o ARN, tal como un analito o un agente intermedio, o una porción del mismo), para crear de esta manera productos de ligazón que sirven como sustitutos de una plantilla.

Como se usa en la presente, una “sonda de captura extendida” se refiere a una sonda de captura que tiene nucleótidos adicionales añadidos al extremo (por ejemplo, extremo 3' o 5') de la sonda de captura, para extender de esta manera la longitud total de la sonda de captura. Por ejemplo, un “extremo 3' extendido” indica que se añadieron nucleótidos adicionales al nucleótido más 3' de la sonda de captura para extender la longitud de la sonda de captura, por ejemplo, mediante reacciones de polimerización usadas para extender moléculas de ácido nucleico, incluida la polimerización con plantilla catalizada por una polimerasa (por ejemplo, una ADN polimerasa o una transcriptasa inversa). En algunas modalidades, extender la sonda de captura incluye añadir a un extremo 3' de una sonda de captura una secuencia de ácido nucleico que es complementaria a una secuencia de ácido nucleico de un analito o agente intermedio unido específicamente al dominio de captura de la sonda de captura. En algunas modalidades, la sonda de captura se extiende mediante el uso de transcripción inversa. En algunas modalidades, la sonda de captura se extiende mediante el uso de una o más ADN polimerasas. Las sondas de captura extendidas incluyen la secuencia de la sonda de captura y la secuencia del código de barras espacial de la sonda de captura.

En algunas modalidades, las sondas de captura extendidas se amplifican (por ejemplo, en solución a granel o en la matriz) para producir cantidades que son suficientes para el análisis aguas abajo, por ejemplo, mediante secuenciación de ADN. En algunas modalidades, las sondas de captura extendidas (por ejemplo, moléculas de ADN) actúan como plantillas para una reacción de amplificación (por ejemplo, una reacción en cadena de la polimerasa).

Variantes adicionales de métodos de análisis espacial, que incluyen, en algunas modalidades, una etapa de obtención de imágenes, se describen en la Sección (II)(a) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm 2020/0277663. Análisis de analitos capturados (y/o agentes intermedios o porciones de los mismos), por ejemplo, incluyendo extracción de muestras, extensión de sondas de captura, secuenciación (por ejemplo, de una sonda de captura extendida escindida y/o una molécula de ADN complementaria a una sonda de captura extendida), la secuenciación en la matriz (por ejemplo, mediante el uso de, por ejemplo, enfoques de hibridación in situ o ligazón in situ), análisis temporal y/o captura de proximidad, se describe en la Sección (II)(g) del documento WO 2020/176788 y/o publicación de solicitud de patente de Estados Unidos núm 2020/0277663. Algunas mediciones de control de calidad se describen en la Sección (II)(h) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

La información espacial puede proporcionar información de importancia biológica y/o médica. Por ejemplo, los métodos y composiciones descritos en la presente descripción pueden permitir: la identificación de uno o más biomarcadores (por ejemplo, de diagnóstico, pronóstico y/o para la determinación de la eficacia de un tratamiento)



de una enfermedad o trastorno; identificación de una diana farmacológica candidato para el tratamiento de una enfermedad o trastorno; identificación (por ejemplo, diagnóstico) de un sujeto que tiene una enfermedad o trastorno; identificación de la etapa y/o pronóstico de una enfermedad o trastorno en un sujeto; identificación de un sujeto que tiene una mayor posibilidad de desarrollar una enfermedad o trastorno; seguimiento de la progresión de una enfermedad o trastorno en un sujeto; determinación de la eficacia de un tratamiento de una enfermedad o trastorno en un sujeto; identificación de una subpoblación de pacientes para la cual un tratamiento es efectivo para una enfermedad o trastorno; modificación de un tratamiento de un sujeto con una enfermedad o trastorno; selección de un sujeto para la participación en un ensayo clínico; y/o selección de un tratamiento para un sujeto con una enfermedad o trastorno.

La información espacial puede proporcionar información de importancia biológica. Por ejemplo, los métodos y composiciones descritos en la presente descripción pueden permitir: identificación de perfiles de expresión del transcriptoma y/o proteoma (por ejemplo, en tejido sano y/o enfermo); identificación de múltiples tipos de analitos muy cercanos (por ejemplo, análisis del vecino más cercano); determinación de genes y/o proteínas regulados positivamente y/o negativamente en tejido enfermo; caracterización de microambientes tumorales; caracterización de respuestas inmunitarias tumorales; caracterización de tipos celulares y su colocación en tejido; e identificación de variantes genéticas dentro de los tejidos (por ejemplo, basadas en los perfiles de expresión de genes y/o proteínas asociados con biomarcadores de enfermedades o trastornos específicos).

Típicamente, para los métodos basados en matrices espaciales, un sustrato funciona como soporte para la unión directa o indirecta de sondas de captura a características de la matriz. Un "elemento" es una entidad que actúa como soporte o depósito para varias entidades moleculares usadas en el análisis espacial. En algunas modalidades, algunas o todas las características de una matriz están funcionalizadas para la captura del analito. Se describen sustratos ilustrativos en la Sección (II)(c) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm 2020/0277663. Pueden encontrarse características y atributos geométricos ilustrativos de una matriz en las Secciones (II)(d)(i), (II)(d)(iii) y (II)(d)(iv) del documento WO 2020/176788 y/o publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

Generalmente, los analitos y/o agentes intermedios (o porciones de los mismos) pueden capturarse cuando se pone en contacto una muestra biológica con un sustrato que incluye sondas de captura (por ejemplo, un sustrato con sondas de captura incrustadas, manchadas, impresas, fabricadas sobre el sustrato, o un sustrato con elementos (por ejemplo, perlas, pocillos) que comprende sondas de captura). Como se usa en la presente descripción, "contacto", "contactado" y/o "que entra en contacto", una muestra biológica con un sustrato se refiere a cualquier contacto (por ejemplo, directo o indirecto) de manera que las sondas de captura pueden interactuar (por ejemplo, unirse covalente o no covalentemente (por ejemplo, hibridar)) con analitos de la muestra biológica. La captura puede lograrse activamente (por ejemplo, mediante el uso de electroforesis) o pasivamente (por ejemplo, mediante el uso de difusión). La captura de analitos se describe con más detalle en la Sección (II)(e) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

En algunos casos, el análisis espacial puede realizarse mediante la unión y/o introducción de una molécula (por ejemplo, un péptido, un lípido o una molécula de ácido nucleico) que tiene un código de barras (por ejemplo, un código de barras espacial) a una muestra biológica (por ejemplo, a una célula en una muestra biológica). En algunas modalidades, una pluralidad de moléculas (por ejemplo, una pluralidad de moléculas de ácido nucleico) que tienen una pluralidad de códigos de barras (por ejemplo, una pluralidad de códigos de barras espaciales) se introducen en una muestra biológica (por ejemplo, en una pluralidad de células en una muestra biológica) para su uso en análisis espacial. En algunas modalidades, después de unir y/o introducir una molécula que tiene un código de barras a una muestra biológica, la muestra biológica puede separarse físicamente (por ejemplo, disociarse) en células individuales o grupos de células para su análisis. Algunos de estos métodos de análisis espacial se describen en la Sección (III) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm 2020/0277663.

En algunos casos, el análisis espacial puede realizarse mediante la detección de múltiples oligonucleótidos que se hibridan con un analito. En algunos casos, por ejemplo, el análisis espacial puede realizarse mediante el uso de ligazón con plantilla de ARN (RTL). Los métodos de RTL se han descrito anteriormente. Ver, por ejemplo, Credle y otros, *Nucleic Acids Res.* 21 de agosto de 2017;45(14):e128. Típicamente, RTL incluye la hibridación de dos oligonucleótidos con secuencias adyacentes en un analito (por ejemplo, una molécula de ARN, tal como una molécula de ARNm). En algunos casos, los oligonucleótidos son moléculas de ADN. En algunos casos, uno de los oligonucleótidos incluye al menos dos bases de ácido ribonucleico en el extremo 3' y/o el otro oligonucleótido incluye un nucleótido fosforilado en el extremo 5'. En algunos casos, uno de los dos oligonucleótidos incluye un dominio de captura (por ejemplo, una secuencia poli(A), una secuencia no homopolimérica). Después de la hibridación con el analito, una ligasa (por ejemplo, la ligasa SplintR) liga los dos oligonucleótidos juntos, creando un producto de ligazón. En algunos casos, los dos oligonucleótidos se hibridan con secuencias que no son adyacentes entre sí. Por ejemplo, la hibridación de los dos oligonucleótidos crea un espacio entre los oligonucleótidos hibridados. En algunos casos, una polimerasa (por ejemplo, una ADN polimerasa) puede extender uno de los oligonucleótidos antes de la ligazón. Después de la ligazón, el producto de ligazón se libera del analito. En algunos casos, el producto de ligazón se libera mediante el uso de una endonucleasa (por ejemplo, ARNasa H). El producto de ligazón liberado puede capturarse a continuación mediante sondas de captura (por ejemplo, en

lugar de la captura directa de un analito) en una matriz, opcionalmente amplificarse y secuenciarse, para determinar de esta manera la ubicación y opcionalmente la abundancia del analito en la muestra biológica.

Durante el análisis de información espacial, se obtiene información de secuencia para un código de barras espacial asociado con un analito, y la información de secuencia puede usarse para proporcionar información sobre la distribución espacial del analito en la muestra biológica. Pueden usarse varios métodos para obtener la información espacial. En algunas modalidades, sondas de captura específicas y los analitos que capturan se asocian con ubicaciones específicas en una serie de características en un sustrato. Por ejemplo, pueden asociarse códigos de barras espaciales específicos con ubicaciones de matriz específicas antes de la fabricación de la matriz, y las secuencias de los códigos de barras espaciales pueden almacenarse (por ejemplo, en una base de datos) junto con información de ubicación de matriz específica, de modo que cada código de barras espacial se asigne de forma única a una ubicación particular de la matriz.

Alternativamente, pueden depositarse códigos de barras espaciales específicos en ubicaciones predeterminadas en una serie de características durante la fabricación, de manera que en cada ubicación, solo esté presente un tipo de código de barras espacial, de modo que los códigos de barras espaciales estén asociados de forma única con una única característica de la matriz. Cuando sea necesario, las matrices pueden decodificarse mediante el uso de cualquiera de los métodos descritos en la presente descripción de modo que los códigos de barras espaciales se asocien de forma única con las ubicaciones de los elementos de la matriz, y este mapeo puede almacenarse como se describió anteriormente.

Cuando se obtiene información de secuencia para sondas y/o analitos de captura durante el análisis de información espacial, las ubicaciones de las sondas y/o analitos de captura pueden determinarse haciendo referencia a la información almacenada que asocia de forma única cada código de barras espacial con una ubicación característica de matriz. De esta manera, las sondas de captura específicas y los analitos capturados se asocian con ubicaciones específicas en la matriz de elementos. Cada ubicación de elemento de matriz representa una posición con relación a un punto de referencia de coordenadas (por ejemplo, una ubicación de matriz, un marcador fiducial) para la matriz. En consecuencia, cada ubicación de característica tiene una "dirección" o ubicación en el espacio de coordenadas de la matriz.

Algunos flujos de trabajo de análisis espacial ilustrativos se describen en la sección modalidades ilustrativas del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm 2020/0277663. Ver, por ejemplo, la modalidad ilustrativa que comienza con "En algunos ejemplos no limitantes de los flujos de trabajo descritos en la presente descripción, la muestra puede sumergirse..." del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm 2020/0277663. Ver también, por ejemplo, la Guía del usuario de los kits de reactivos de expresión genética espacial de Visium (por ejemplo, Rev C, con fecha de junio de 2020) y/o la Guía del usuario de los kits de reactivos de optimización de tejidos espaciales de Visium (por ejemplo, Rev C, con fecha de julio de 2020).

En algunas modalidades, el análisis espacial puede realizarse mediante el uso de hardware y/o software dedicado, tal como cualquiera de los sistemas descritos en las Secciones (II)(e)(ii) y/o (V) de los documentos WO 2020/176788 y/o publicación de solicitud de patente de Estados Unidos núm 2020/0277663, o cualquiera de uno o más de los dispositivos o métodos descritos en las Secciones *Control Slide for Imaging, Methods of Using Control Slides and Substrates for, Systems of Using Control Slides and Substrates for Imaging, y/o Sample and Array Alignment Devices and Methods, Informational labels* del documento WO 2020/123320.

Los sistemas adecuados para realizar análisis espacial pueden incluir componentes tales como una cámara (por ejemplo, una celda de flujo o una cámara hermética de fluido) para contener una muestra biológica. La muestra biológica puede montarse, por ejemplo, en un soporte para muestras biológicas. Pueden conectarse una o más cámaras de fluido a la cámara y/o al portamuestras a través de conductos de fluido, y los fluidos se pueden suministrar a la cámara y/o al portamuestras a través de bombas fluidicas, fuentes de vacío u otros dispositivos acoplados a los conductos de fluido que crean un gradiente de presión para impulsar el flujo de fluido. También pueden conectarse una o más válvulas a conductos de fluido para regular el flujo de reactivos desde los depósitos hasta la cámara y/o el portamuestras.

Los sistemas pueden incluir opcionalmente una unidad de control que incluye uno o más procesadores electrónicos, una interfaz de entrada, una interfaz de salida (tal como una pantalla) y una unidad de almacenamiento (por ejemplo, un medio de almacenamiento de estado sólido tal como, pero sin limitarse a, un medio de almacenamiento magnético, óptico u otro de estado sólido, persistente, grabable y/o regrabable). Opcionalmente, la unidad de control puede conectarse a uno o más dispositivos remotos a través de una red. La unidad de control (y sus componentes) generalmente puede realizar cualquiera de las etapas y funciones descritos en la presente descripción. Cuando el sistema está conectado a un dispositivo remoto, el dispositivo (o dispositivos) remoto(s) puede(n) realizar cualquiera de las etapas o características descritas en la presente descripción. Los sistemas pueden incluir opcionalmente uno o más detectores (por ejemplo, CCD, CMOS) usados para capturar imágenes. Los sistemas también pueden incluir opcionalmente una o más fuentes de luz (por ejemplo, basadas en LED, basadas en diodos, láseres) para iluminar una muestra, un sustrato con elementos, analitos de una muestra biológica capturada en un sustrato y diversos medios de control y calibración.

Los sistemas pueden incluir opcionalmente instrucciones de software codificadas y/o implementadas en uno o más medios de almacenamiento tangibles y componentes de hardware tales como circuitos integrados de aplicaciones específicas. Las instrucciones de software, cuando son ejecutadas por una unidad de control (y en particular, un procesador electrónico) o un circuito integrado, pueden hacer que la unidad de control, circuito integrado u otro componente que ejecuta las instrucciones de software realice cualquiera de las etapas o funciones del método descritos en la presente descripción.

En algunos casos, los sistemas descritos en la presente descripción pueden detectar (por ejemplo, registrar una imagen) la muestra biológica en la matriz. Los métodos ilustrativos para detectar la muestra biológica en una matriz se describen en la solicitud PCT núm. 2020/061064 y/o en la solicitud de patente de los Estados Unidos con número de serie 16/951,854.

Antes de transferir los analitos desde la muestra biológica a la matriz de características en el sustrato, la muestra biológica puede alinearse con la matriz. La alineación de una muestra biológica y una matriz de características que incluyen sondas de captura puede facilitar el análisis espacial, que puede usarse para detectar diferencias en la presencia y/o el nivel de analito dentro de diferentes posiciones en la muestra biológica, por ejemplo, para generar un mapa tridimensional de la presencia y/o el nivel de analito. Los métodos ilustrativos para generar un mapa bidimensional y/o tridimensional de la presencia y/o el nivel de analito se describen en la solicitud PCT núm. 2020/053655 y los métodos de análisis espacial se describen generalmente en el documento WO 2020/061108 y/o en la solicitud de patente de los Estados Unidos con número de serie 16/951,864.

En algunos casos, un mapa de presencia y/o el nivel de analito puede alinearse con una imagen de una muestra biológica mediante el uso de uno o más marcadores fiduciales, por ejemplo, objetos que se colocan en el campo de visión de un sistema de obtención de imágenes que aparecen en la imagen producida, como se describe en la sección *Substrate Attributes*, la sección *Control Slide for Imaging* del documento WO 2020/123320, solicitud PCT núm. 2020/061066 y/o en la solicitud de patente de los Estados Unidos con número de serie 16/951,843. Los marcadores fiduciales pueden usarse como punto de referencia o escala de medición para la alineación (por ejemplo, para alinear una muestra y una matriz, para alinear dos sustratos, para determinar la ubicación de una muestra o matriz en un sustrato con relación a un marcador fiducial) y/o para mediciones cuantitativas de tamaños y/o distancias.

#### Ensayo espacial para la cromatina accesible a transposasas

El cuerpo humano incluye una gran colección de diversos tipos de células, cada una de las cuales proporciona una función especializada y específica del contexto. Comprender la estructura de la cromatina de una célula puede revelar información sobre la función de la célula. La cromatina abierta, o la cromatina accesible, o el ADN genómico accesible, a menudo es indicativo de secuencias transcripcionalmente activas, por ejemplo, genes, en una célula particular. Comprender mejor las regiones transcripcionalmente activas dentro de la cromatina permitirá la identificación de cuáles genes contribuyen a la función y/o fenotipo de una célula.

Se han desarrollado métodos para estudiar epigenomas, por ejemplo, ensayos de accesibilidad a la cromatina (ATAC-seq) o identificar proteínas asociadas con la cromatina, por ejemplo, (ChIP-seq). Estos ensayos ayudan a identificar, por ejemplo, los reguladores (por ejemplo, los reguladores cis y/o reguladores trans) que contribuyen a fenotipos celulares dinámicos. Aunque ATAC-seq y ChIP-seq son inestimables para definir la variabilidad epigenética dentro de una población celular, las aplicaciones convencionales de estos métodos se limitan en su capacidad de resolver espacialmente las estructuras tridimensionales y los genes asociados que promueven la variación celular.

Por lo tanto, la presente descripción se refiere generalmente al marcaje y análisis espacial de los ácidos nucleicos. En algunas modalidades, se proporcionan en la presente descripción métodos que utilizan una enzima transposasa para acoplarse y fragmentar, por ejemplo, el ADN genómico accesible (por ejemplo, la cromatina abierta) y permitir la captura simultánea del ADN y ARN a partir de una muestra biológica, de esta manera se revelan conocimientos epigenómicos con respecto a las características estructurales que contribuyen a la regulación celular.

En la presente descripción se proporcionan métodos para determinar la accesibilidad al ADN genómico que incluyen (a) una muestra biológica en una matriz que comprende una pluralidad de sondas de captura, donde una sonda de captura de la pluralidad de sondas de captura comprende: (i) un código de barras espacial y (ii) un dominio de captura; (b) poner en contacto una pluralidad de oligonucleótidos de férula con la muestra biológica, en donde un oligonucleótido de férula se une al dominio de captura; (c) poner en contacto un transposoma con la muestra biológica para insertar secuencias de extremos del transposón en el ADN genómico accesible, de esta manera se genera ADN genómico fragmentado; (d) hibridar el ADN genómico fragmentado con el oligonucleótido de férula y ligar las secuencias de extremos del transposón del ADN genómico fragmentado a la sonda de captura, de esta manera se generan secuencias de extremos del transposón ligadas; (e) liberar una o más secuencias de extremos del transposón no ligadas de las secuencias de extremos del transposón ligadas; (f) determinar (i) la totalidad o una porción de una secuencia del código de barras espacial o un complemento de la misma, y (ii) la totalidad o una porción de una secuencia del ADN genómico fragmentado, o un complemento de la misma, y usar las secuencias determinadas de (i) y (ii) para determinar la accesibilidad al ADN genómico en la muestra biológica.

En algunas modalidades, las etapas (d) y (e) se realizan secuencialmente. En algunas modalidades, las etapas (d) y (e) se realizan simultáneamente. Por ejemplo, algunos fragmentos de ADN marcados pueden capturarse con secuencias de extremos del transposón no ligadas aún hibridadas. En tales ejemplos, las secuencias de extremos del transposón no ligadas se liberan después de la captura del ADN marcado. En algunas modalidades, las secuencias de extremos del transposón no ligadas se liberan antes de la captura por el dominio de captura.

También se proporcionan en la presente descripción métodos para determinar la accesibilidad al ADN genómico que incluyen (a) una muestra biológica en una matriz que comprende una pluralidad de sondas de captura, en donde una sonda de captura de la pluralidad de sondas de captura comprende: (i) un código de barras espacial y (ii) un dominio de captura; (b) poner en contacto un transposoma con la muestra biológica para insertar secuencias de extremos del transposón en el ADN genómico accesible, de esta manera se genera ADN genómico fragmentado; (c) hibridar las secuencias de extremos del transposón del ADN genómico fragmentado al dominio de captura de la sonda de captura; (d) liberar una o más secuencias de extremos del transposón no ligadas al dominio de captura; (e) determinar (i) la totalidad o una porción de una secuencia del código de barras espacial o un complemento de esta, y (ii) la totalidad o una porción de una secuencia del ADN genómico fragmentado, o un complemento de la misma, y usar las secuencias determinadas de (i) y (ii) para determinar la accesibilidad al ADN genómico en la muestra biológica.

En algunas modalidades, las etapas (c) y (d) se realizan secuencialmente. En algunas modalidades, las etapas (c) y (d) se realizan simultáneamente. Por ejemplo, algunos fragmentos del ADN marcado pueden capturarse con una o más secuencias de extremos del transposón aún hibridadas. En tales ejemplos, una o más secuencias de extremos del transposón se liberan después de la captura del ADN marcado. En algunas modalidades, una o más secuencias de extremos del transposón se liberan antes de la captura por el dominio de captura.

En algunas modalidades, se proporcionan en la presente descripción métodos para el análisis espacial de los ácidos nucleicos (por ejemplo, el ADN genómico, ARNm) en una muestra biológica. En algunas modalidades, se proporciona una matriz, en donde la matriz comprende una pluralidad de sondas de captura. En algunas modalidades, las sondas de captura pueden unirse directamente al sustrato (por ejemplo, una matriz que comprende un sustrato que comprende una pluralidad de sondas de captura). En algunas modalidades, las sondas de captura pueden unirse indirectamente al sustrato. Por ejemplo, las sondas de captura pueden unirse a elementos en el sustrato. En algunas modalidades, un elemento es una perla. En algunas modalidades, las sondas de captura comprenden un código de barras espacial y un dominio de captura. En algunas modalidades, la sonda de captura puede ser parcialmente de doble hebra. En algunas modalidades, la sonda de captura puede unirse a un oligonucleótido complementario. En algunas modalidades, el oligonucleótido complementario (por ejemplo, un oligonucleótido de férula) puede tener una porción de una sola hebra. En algunas modalidades, la porción de una sola hebra puede hibridarse con el ADN fragmentado (por ejemplo, el marcado). En algunas modalidades, una muestra biológica se trata bajo condiciones suficientes para producir en las células de la muestra biológica ácidos nucleicos (por ejemplo, ADN genómico) accesibles para la inserción del transposón (por ejemplo, para marcar los fragmentos de ADN con extremos del transposón). En algunas modalidades, una secuencia de extremo del transposón y una enzima transposasa (colectivamente, un transposoma) se proporcionan a la muestra biológica de manera que la secuencia de extremo del transposón pueda insertarse en el ADN genómico accesible de las células presentes en la muestra biológica. En algunas modalidades, la enzima transposasa de los fragmentos complejos del transposoma, el ADN genómico y los extremos del transposón, se unen a los extremos de los fragmentos del ADN genómico (por ejemplo, el "marcado").

En algunas modalidades, la muestra biológica que comprende ácidos nucleicos (por ejemplo, ADN genómico, ARNm) se pone en contacto con el sustrato de manera que una sonda de captura puede interactuar con el ADN genómico fragmentado y marcado (por ejemplo, el marcado). En algunas modalidades, la muestra biológica que comprende ácidos nucleicos (por ejemplo, ADN genómico, ARNm) se pone en contacto con el sustrato de manera que la sonda de captura puede interactuar tanto con el ADN genómico marcado como con el ARNm presente en la muestra biológica (por ejemplo, una primera sonda de captura puede unirse al ADN genómico marcado, una segunda sonda de captura puede unirse al ARNm).

En algunas modalidades, la ubicación de la sonda de captura en el sustrato puede correlacionarse con una ubicación en la muestra biológica, de esta manera se determina espacialmente la ubicación del ADN genómico marcado en la muestra biológica. En algunas modalidades, la ubicación de la sonda de captura en el sustrato puede correlacionarse con una ubicación en la muestra biológica, de esta manera se determina espacialmente la ubicación del ADN genómico marcado y el ARNm en la muestra biológica.

#### ATAC-seq espacial

En algunas modalidades, de cualquiera de los métodos de análisis espacial descritos en la presente descripción, se usa ATAC-seq para generar mapas de accesibilidad a la cromatina de todo el genoma. Estos mapas de accesibilidad a todo el genoma pueden integrarse con datos adicionales del perfil de todo el genoma (por ejemplo, una RNA-seq, ChIP-seq, Metil-seq) para producir mapas de interacción reguladora de genes que facilitan la comprensión de la regulación transcripcional. Por ejemplo, la interrogante de los mapas de accesibilidad a todo el genoma puede revelar los factores de transcripción subyacentes y los motivos de los factores de transcripción

responsables de la accesibilidad a la cromatina en una ubicación genómica dada. La correlación de los cambios en la accesibilidad a la cromatina con los cambios en la expresión génica (RNA-seq), los cambios en la unión del factor de transcripción (por ejemplo, ChIP-seq) y/o los cambios en los niveles de metilación del ADN (por ejemplo, Metil-seq) pueden identificar la regulación de la transcripción que impulsa estos cambios. En los estados de enfermedad, a menudo existe un desequilibrio en la regulación transcripcional. Por lo tanto, analizar tanto la accesibilidad a la cromatina como, por ejemplo, la expresión génica mediante el uso de métodos de análisis espacial permite la identificación de los desequilibrios subyacentes en la regulación transcripcional y potencialmente las causas de los mismos.

En algunas modalidades, donde la determinación espacial de la ubicación de los analitos incluye un análisis concurrente de diferentes tipos de analitos de una célula única o una subpoblación de células dentro de una muestra biológica (por ejemplo, una sección de tejido), puede integrarse una capa adicional de información espacial en los mapas de interacción reguladores del genoma. En algunas modalidades, la determinación espacial de analitos puede realizarse en genomas completos. En algunas modalidades, el perfil espacial puede hacerse en una muestra biológica fijada.

En algunas modalidades, los mapas de accesibilidad de la cromatina a todo el genoma generados por ATAC-seq espacial pueden usarse para la identificación del tipo de célula. Por ejemplo, la clasificación tradicional del tipo de célula se basa en los niveles de expresión de ARNm, pero la accesibilidad a la cromatina puede ser más eficiente en capturar la identidad celular. Además, en algunas modalidades, las correlaciones entre regiones transcripcionalmente activas (por ejemplo, la cromatina abierta, accesible) con perfiles de expresión (por ejemplo, los perfiles de expresión de ARNm) pueden determinarse de manera espacial.

#### Permeabilización de la muestra biológica

La presente descripción describe generalmente métodos para marcar el ADN genómico para generar fragmentos de ADN en una muestra biológica. En algunos ejemplos, puede emplearse una "prepermeabilización" química o enzimática de las muestras biológicas inmovilizadas sobre un sustrato para permitir que el ADN en la muestra biológica sea accesible a una enzima transposasa (por ejemplo, en un complejo transposoma). En algunas modalidades, la permeabilización de la muestra biológica puede ser un proceso de dos etapas (por ejemplo, tratamiento de prepermeabilización, seguido de un tratamiento de permeabilización). En algunas modalidades, la permeabilización de la muestra biológica puede ser un proceso de una etapa (por ejemplo, un único tratamiento de permeabilización es suficiente para permeabilizar las membranas celulares y nucleares en la muestra biológica).

En algunas modalidades, la prepermeabilización puede incluir una condición enzimática o química. En algunas modalidades, la prepermeabilización puede realizarse con una enzima (por ejemplo, una proteasa). En algunas modalidades, de manera no limitante, la proteasa puede incluir una tripsina, pepsina, dispasa, papaína o colagenasa. En algunas modalidades, la prepermeabilización puede incluir un tratamiento enzimático con una pepsina. En algunas modalidades, la prepermeabilización puede incluir pepsina en 0,5 M de ácido acético. En algunas modalidades, la prepermeabilización puede incluir una pepsina en un tampón de Exonucleasa-1. En algunas modalidades, el pH del tampón puede ser ácido. En algunas modalidades, la prepermeabilización puede incluir un tratamiento enzimático con una colagenasa. En algunas modalidades, la prepermeabilización puede incluir una colagenasa en un tampón HBSS. En algunas modalidades, el tampón HBSS puede incluir una albúmina sérica bovina (BSA). En algunas modalidades, la prepermeabilización puede durar aproximadamente de 1 minuto a aproximadamente 20 minutos. En algunas modalidades, la prepermeabilización puede durar aproximadamente de 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 11, aproximadamente 12, aproximadamente 13, aproximadamente 14, aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18 o aproximadamente 19 minutos. En algunas modalidades, la prepermeabilización puede durar aproximadamente de 10 minutos a aproximadamente una hora. Por ejemplo, en algunas modalidades, la prepermeabilización puede durar aproximadamente de 20, aproximadamente 30, aproximadamente 40 o aproximadamente 50 minutos.

En algunas modalidades, la permeabilización de la muestra biológica comprende un tratamiento enzimático. En algunas modalidades, el tratamiento enzimático puede ser una enzima pepsina, o un tratamiento enzimático similar a la pepsina. En algunas modalidades, el tratamiento enzimático puede ser un tratamiento con una proteasa. En algunas modalidades, el tratamiento enzimático puede realizarse en presencia de reactivos. En algunas modalidades, el tratamiento enzimático (por ejemplo, la prepermeabilización) puede incluir poner en contacto la muestra biológica con una solución ácida que incluye una enzima proteasa. En algunas modalidades, el reactivo puede ser HCl. En algunas modalidades, el reactivo puede ser ácido acético. En algunas modalidades, la concentración de HCl puede ser de aproximadamente 100 mM. En algunas modalidades, aproximadamente 100 mM de HCl puede tener un pH de alrededor de, o aproximadamente 1,0. En algunas modalidades, un reactivo adicional puede ser 0,5 M de ácido acético, que tiene un pH de alrededor de, o aproximadamente 2,5. Se observa que el tratamiento enzimático de la muestra biológica puede tener diferentes efectos sobre el fragmento marcaje. Por ejemplo, el tratamiento enzimático con pepsina y 100 mM de HCl puede resultar en el fragmento marcaje de la cromatina independientemente de la accesibilidad a la cromatina. En algunas modalidades, el tratamiento

enzimático con pepsina y 0,5 M de ácido acético puede resultar en el fragmento marcaje de la cromatina que puede retener un patrón nucleosomal indicativo de la accesibilidad a la cromatina.

En algunas modalidades, el tratamiento enzimático puede comprender poner en contacto la muestra biológica con una mezcla de reacción (por ejemplo, una solución) que comprende una proteasa de aspartil (por ejemplo, pepsina) en un tampón ácido, por ejemplo, un tampón con un pH de aproximadamente 4,0 o menos, tal como de aproximadamente 3,0 o menos, por ejemplo, de aproximadamente 0,5 a aproximadamente 3,0, o de aproximadamente 1,0 a aproximadamente 2,5. En algunas modalidades, la proteasa de aspartil es una enzima pepsina, una enzima similar a la pepsina o un equivalente funcional de esta. Por lo tanto, cualquier enzima o combinación de enzimas en el número de comisión enzimática 3.4.23.1.

En algunas modalidades, el tratamiento enzimático (por ejemplo, la prepermeabilización) puede realizarse mediante el uso de una collagenasa. En algunas modalidades, el tratamiento enzimático con collagenasa puede proporcionar acceso al ADN genómico para la transposasa mientras se preserva la integridad nuclear. En algunas modalidades, la prepermeabilización (por ejemplo, el tratamiento enzimático) con collagenasa produce patrones nucleosomales generalmente asociados con la accesibilidad a la cromatina. Las collagenasas pueden aislarse a partir de *Clostridium histolyticum*. En algunas modalidades, el tratamiento enzimático con una endopeptidasa de zinc (por ejemplo, una collagenasa) con reactivos y bajo condiciones adecuadas para la actividad proteolítica comprende una solución tamponada con un pH de aproximadamente 7,0 a aproximadamente 8,0 (por ejemplo, de aproximadamente 7,4). Las collagenasas son endopeptidasas de zinc y pueden inhibirse mediante EDTA o EGTA, o ambos. Por lo tanto, en algunas modalidades, la muestra biológica puede ponerse en contacto con una endopeptidasa de zinc (por ejemplo, una collagenasa) en ausencia de un quelante de cationes divalentes (por ejemplo, EDTA, EGTA). En algunas modalidades, puede ser útil detener la endopeptidasa de zinc (por ejemplo, una collagenasa) y la etapa de permeabilización puede detenerse (por ejemplo, inhibirse) al poner en contacto la muestra biológica con un quelante de cationes divalentes (por ejemplo, EDTA, EGTA).

En algunas modalidades, la endopeptidasa de zinc es una enzima collagenasa, una enzima similar a la collagenasa o un equivalente funcional de esta. En tales modalidades, cualquier enzima o combinación de enzimas en el número de comisión enzimática 3.4.23.3 puede usarse de acuerdo con los materiales y métodos descritos en la presente descripción. En algunas modalidades, la collagenasa es una o más collagenasas del siguiente grupo (números de acceso de UniProtKB/Swiss-Prot): P43153/COLA\_CLOPE; P43154/COLA\_VIBAL; Q9KRJ0/COLA\_VIBCH; Q56696/COLA\_VIBPA; Q8D4Y9/COLA\_VIBVU; Q9X721/COLG\_HATHI; Q46085/COLH\_HATHI; Q899Y1/COLT\_CLOTE URSTH y variantes funcionales y derivados de estas (descritas en la presente descripción), o una de sus combinaciones.

Los métodos para permeabilizar muestras biológicas se conocen bien en la técnica. Se conocerá por un experto en la técnica que diferentes fuentes de muestras biológicas pueden tratarse con diferentes reactivos (por ejemplo, proteasas, ARNasas, detergentes, tampones) y bajo diferentes condiciones (por ejemplo, presión, temperatura, concentración, pH, tiempo). En algunas modalidades, la permeabilización de la muestra biológica puede comprender reactivos y condiciones para interrumpir suficientemente la membrana celular de la muestra biológica para capturar los ácidos nucleicos (por ejemplo, el ARNm). En algunas modalidades, la permeabilización de la muestra biológica puede comprender reactivos y condiciones para interrumpir suficientemente la membrana nuclear de la muestra biológica para capturar los ácidos nucleicos (por ejemplo, el ADN genómico). En algunas modalidades, pueden usarse proteasas disponibles comercialmente aisladas a partir de su origen nativo (por ejemplo, de fuente animal, microbiana). En algunas modalidades, pueden usarse proteasas producidas de manera recombinante (por ejemplo, de un sistema de expresión bacteriana, sistema de expresión viral). En algunas modalidades, prepermeabilizar y permeabilizar una muestra biológica puede ser un proceso de una etapa (por ejemplo, un tratamiento enzimático). En algunas modalidades, prepermeabilizar y permeabilizar una muestra biológica puede ser un proceso de dos etapas (por ejemplo, tratamiento enzimático, seguido de un tratamiento químico o detergente).

En algunas modalidades, las condiciones de permeabilización química comprenden poner en contacto la muestra biológica con una solución alcalina, por ejemplo, una solución tamponada con un pH de aproximadamente 8,0 a aproximadamente 11,0, tal como de aproximadamente 8,5 a aproximadamente 10,5 o de aproximadamente 9,0 a aproximadamente 10,0, por ejemplo de aproximadamente 9,5. En algunas modalidades, el tampón es un tampón de glicina-KOH. Otros tampones se conocen en la técnica.

En algunas modalidades, una muestra biológica puede tratarse con un detergente después de un tratamiento enzimático (por ejemplo, una permeabilización después de una etapa de prepermeabilización). Los detergentes se conocen en la técnica. Puede usarse cualquier detergente adecuado, que incluye, de manera no limitante NP-40 o equivalente, Digitonina, Tween-20, IGEPAL-40 o equivalente, Saponina, SDS, Pitsop2, Triton X-100 o sus combinaciones. En algunas modalidades, una muestra biológica puede tratarse con otros productos químicos que se conocen por permeabilizar las membranas celulares. Como se ejemplifica adicionalmente en los ejemplos más abajo, los detergentes descritos en la presente descripción pueden usarse a una concentración de entre aproximadamente 0,001 % (v/v) a aproximadamente 5 % (v/v). En algunas modalidades, los detergentes descritos en la presente descripción pueden usarse a una concentración de aproximadamente 0,01 % (v/v), aproximadamente 0,02 % (v/v), aproximadamente 0,03 % (v/v), aproximadamente 0,04 % (v/v), aproximadamente

0,05 % (v/v), aproximadamente 0,06 % (v/v), aproximadamente 0,07 % (v/v), aproximadamente 0,08 % o aproximadamente 0,09 %. En algunas modalidades, los detergentes descritos en la presente descripción pueden usarse a una concentración de aproximadamente 0,1 % (v/v), aproximadamente 0,2 % (v/v), aproximadamente 0,3 % (v/v), aproximadamente 0,4 % (v/v), aproximadamente 0,5 % (v/v), aproximadamente 0,6 % (v/v), aproximadamente 0,7 % (v/v), aproximadamente 0,8 % (v/v), aproximadamente 0,9 % (v/v), aproximadamente 1,0 % (v/v), o aproximadamente 1,1 % a aproximadamente 10 % (v/v) o más. En algunas modalidades, los detergentes descritos en la presente descripción pueden usarse a una concentración de aproximadamente 2 % (v/v), aproximadamente 3 % (v/v), aproximadamente 4 % (v/v), aproximadamente 5 % (v/v), aproximadamente 6 % (v/v), aproximadamente 7 % (v/v), aproximadamente 8 % (v/v), o aproximadamente 9 % (v/v), o aproximadamente 10 % (v/v).

Se describen métodos adicionales para la permeabilización de las muestras, por ejemplo, en Jamur y otros, *Method Mol. Biol.* 588:63-66, 2010. Generalmente puede usarse cualquier método adecuado para la permeabilización de muestras biológicas en relación con las muestras biológicas descritas en la presente descripción.

Pueden tratarse diferentes fuentes de muestras biológicas con diferentes reactivos (por ejemplo, proteasas, ARNasas, detergentes, tampones) y bajo diferentes condiciones adecuadas (por ejemplo, presión, temperatura, concentración, pH, tiempo) para lograr suficiente prepermeabilización y permeabilización para capturar los ácidos nucleicos (por ejemplo, el ADN genómico, ARNm).

En algunas modalidades, la mezcla de reacción (por ejemplo, una solución) que incluye las proteasas descritas en la presente descripción puede contener otros reactivos (por ejemplo, tampón, sal, etc.) suficientes para asegurar que las proteasas sean funcionales. Por ejemplo, la mezcla de reacción puede incluir además una proteína albúmina (por ejemplo, BSA). En algunas modalidades, la mezcla de reacción (por ejemplo, una solución) que incluye la enzima colagenasa (o una variante funcional o derivado de esta) incluye una proteína albúmina (por ejemplo, BSA).

En algunas modalidades, hay una o más etapas de lavado entre la prepermeabilización y la permeabilización de una muestra biológica. Por ejemplo, puede ser preferible lavar la mayor cantidad de veces la solución de prepermeabilización de la muestra biológica antes de añadir una solución de permeabilización. Por lo tanto, en algunas modalidades la muestra biológica se lava, por ejemplo con una solución de lavado de SSC, después de la prepermeabilización para eliminar los reactivos de prepermeabilización y antes de aplicar los reactivos de permeabilización a la muestra biológica. En algunas modalidades, la solución de permeabilización se elimina de la muestra biológica antes de la adición de los reactivos transposómicos para el fragmentomarcaje del ADN genómico liberado. También se pueden realizar uno o más lavados después de la permeabilización y antes del fragmentomarcaje, por ejemplo, mediante el uso de una solución SSC. En algunas modalidades, no hay etapas de lavado entre la permeabilización de la muestra biológica y el fragmentomarcaje del ADN genómico.

#### Fragmentomarcaje

Las enzimas transposasas y los transposones pueden usarse en métodos de análisis genómico espacial. Generalmente, la transposición es el proceso mediante el cual una secuencia genética específica (por ejemplo, una secuencia del transposón) se reubica de un lugar en un genoma a otro. Muchos métodos de transposición y elementos transponibles se conocen en la técnica (por ejemplo, transposones de ADN, retrotransposones, transposones autónomos, transposones no autónomos). Un ejemplo no limitante de un evento de transposición es la transposición conservadora. La transposición conservadora es un modo de transposición no replicativo en el que el transposón se elimina completamente del genoma y se reintegra en un nuevo locus, de manera que la secuencia del transposón se conserva, (por ejemplo, un evento de transposición conservador puede considerarse un evento de "corte y pega") (ver, por ejemplo, Griffiths A. J., y otros, *Mechanism of transposition in prokaryotes. An Introduction to Genetic Analysis* (7ma Ed.). Nueva York: W. H. Freeman (2000)).

En un ejemplo, la transposición de corte y pega puede ocurrir cuando una enzima transposasa se une a una secuencia que flanquea los extremos del transposoma (por ejemplo, una secuencia de reconocimiento, por ejemplo, una secuencia de extremo del mosaico, una secuencia del transposón). Un transposoma (por ejemplo, un complejo de transposición) se forma y el ADN endógeno puede manipularse en un complejo de preescisión de manera que dos enzimas transposasas puedan interactuar. En algunas modalidades, cuando las transposasas interactúan se introducen rupturas de doble hebra en el ADN. Las enzimas transposasas pueden localizar y unirse a un sitio diana en el ADN, crear una ruptura de doble hebra, e insertar la secuencia de extremo del transposón (ver, por ejemplo, Skipper, K.A., y otros, DNA transposon-based gene vehicles-scenes from an evolutionary drive, *J Biomed Sci*, 20: 92 (2013) doi:10.1186/1423-0127-20-92). Las transposasas de corte y pega alternativas incluyen Tn552 (College, y otros, *J. Bacteriol.*, 183: 2384-8, 2001; Kirby C y otros, *Mol. Microbiol.*, 43: 173-86, 2002), Tyl (Devine y Boeke, *Nucleic Acids Res.*, 22: 3765-72, 1994 y la Publicación Internacional WO 95/23875), Tn7 (Craig, N L, *Science*. 271 : 1512, 1996; Craig, N L, Review in: *Curr Top Microbiol Immunol*, 204:27-48, 1996), Tn/O e IS10 (Kleckner N, y otros, *Curr Top Microbiol Immunol*, 204:49-82, 1996), transposasa Mariner (Lampe D J, y otros, *EMBO J.*, 15: 5470-9, 1996), Tel (Plasterk R H, *Curr. Topics Microbiol. Immunol*, 204: 125-43, 1996), P Element (Gloor, G B, *Methods Mol. Biol.*, 260: 97- 114, 2004), Tn3 (Ichikawa & Ohtsubo, *J Biol. Chem.* 265: 18829-32, 1990), secuencias de inserción

bacterianas (Ohtsubo & Sekine, Curr. Top. Microbiol. Immunol. 204: 1-26, 1996), retrovirus (Brown, y otros, Proc Natl Acad Sci USA, 86:2525-9, 1989), y retrotransposón de levadura (Boeke & Corces, Annu Rev Microbiol. 43:403-34, 1989). Otros ejemplos incluyen IS5, TnIO, Tn903, IS911 y versiones modificadas genéticamente de enzimas de la familia de transposasas (Zhang y otros, (2009) PLoS Genet. 5:e1000689. Epub 2009 Oct. 16; Wilson C. y otros (2007) J. Microbiol. Methods 71:332-5).

La fragmentación mediada por transposomas y el marcaje ("fragmentomarcaje") es un proceso de fragmentación mediado por transposasas y de marcaje del ADN. Un transposoma es un complejo de una enzima transposasa y ADN que comprende una secuencia de extremo del transposón (también conocida como "secuencia de reconocimiento de transposasa" o "extremo del mosaico" (ME)). En algunos métodos de análisis genómico espacial, el ADN se fragmenta de tal manera que una secuencia funcional tal como una secuencia complementaria a un dominio de captura de una sonda de captura (por ejemplo, dominio de captura de un oligonucleótido de férula) se inserta en el ADN fragmentado (por ejemplo, el ADN fragmentado se "marca"), de manera que la secuencia (por ejemplo, un adaptador, por ejemplo, secuencia Nextera) puede hibridarse con la sonda de captura. En algunas modalidades, la sonda de captura está presente en un sustrato. En algunas modalidades, la sonda de captura (por ejemplo, una sonda de captura y un oligonucleótido de férula) está presente en un elemento. Un dímero de transposasa (por ejemplo, en el caso del sistema de transposasa Tn5) junto con una secuencia del transposón (por ejemplo, el transposoma) es capaz de fragmentar simultáneamente el ADN en base a sus secuencias de reconocimiento de transposones y ligar el ADN del transposoma (por ejemplo, la secuencia del transposón) al ADN fragmentado (por ejemplo, el ADN marcado). Este sistema se adapta mediante el uso de enzimas de transposasa hiperactivas y moléculas de ADN modificadas (adaptadores) que comprenden ME para fragmentar el ADN y marcar ambas hebras de fragmentos dúplex del ADN con moléculas de ADN funcionales (por ejemplo, sitios de unión de cebadores). Por ejemplo, la transposasa Tn5 puede producirse como monómeros proteicos purificados. La transposasa Tn5 también está disponible comercialmente (por ejemplo, el fabricante Illumina, Illumina.com, núm. de catálogo 15027865, Tampón de marcaje de ADN TD núm. de catálogo 15027866). Estos pueden cargarse subsecuentemente con los oligonucleótidos de interés, por ejemplo, los oligonucleótidos de ADNcc que contienen ME (por ejemplo, las secuencias del transposón) para el reconocimiento de Tn5 y secuencias funcionales adicionales (por ejemplo, los adaptadores Nextera, por ejemplo, los sitios de unión a cebadores) se hibridan para formar un extremo del oligonucleótido del mosaico (MEDS) de ADNc que se reconoce por Tn5 durante el ensamble del dímero (por ejemplo, dimerización del transposoma). En algunas modalidades, una transposasa hiperactiva Tn5 puede cargarse con adaptadores (por ejemplo, oligonucleótidos de interés) que pueden fragmentar y marcar simultáneamente un genoma.

Como se usa en la presente descripción, el término "fragmentomarcaje" se refiere a una etapa en el Ensayo de Cromatina Accesible a Transposasa mediante el uso de secuenciación (ATAC-seq). (Ver, por ejemplo, Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., Greenleaf, W. J., Transposition of native chromatin for Fast and sensitive epi genomic profiling of open chromatin, DNA-binding proteins and nucleosome position, *Nature Methods*, 10 (12): 1213-1218 (2013)). ATAC-seq identifica regiones de cromatina abierta mediante el uso de una transposasa hiperactiva Tn5 procariota, que se inserta preferentemente en la cromatina accesible y marca los sitios con adaptadores (Buenrostro, J.D., y otros, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*, 10: 1213—1218 (2013)).

Como se usa en la presente descripción, "cromatina accesible" o "cromatina abierta" o "ADN genómico accesible" se refiere a porciones de un genoma que son regiones empobrecidas en los nucleosomas que pueden unirse a proteínas y desempeñar varios papeles en la organización nuclear, la transcripción génica y generalmente se consideran regiones transcripcionalmente activas del ADN (Zhang, Q., y otros, Genome-wide open chromatin regions and their effects on the regulation of silk protein genes in *Bombyx mori*, *Scientific Reports*, 7: 12919 (2017)).

En algunas modalidades, la etapa de fragmentar el ADN genómico en las células de la muestra biológica comprende poner en contacto la muestra biológica que contiene el ADN genómico con la enzima transposasa (por ejemplo, un transposoma, por ejemplo, una mezcla de reacción (por ejemplo, una solución)) que incluye una transposasa), bajo cualquier condición adecuada. En algunas modalidades, tales condiciones adecuadas resultan en el fragmentomarcaje del ADN genómico de las células presentes en la muestra biológica. Las condiciones típicas dependerán de la enzima transposasa usada y pueden determinarse mediante el uso de métodos rutinarios que se conocen en la técnica. Por lo tanto, las condiciones adecuadas pueden ser condiciones (por ejemplo, tampón, sal, concentración, pH, temperatura, condiciones del tiempo) bajo las que la enzima transposasa es funcional, por ejemplo, en las que la enzima transposasa muestra actividad transposasa, particularmente actividad de fragmentomarcaje, en la muestra biológica.

El término "funcional", como se usa en la presente descripción en referencia a las enzimas transposasas, pretende incluir modalidades en las que la enzima transposasa puede mostrar cierta actividad reducida con relación a la actividad de la enzima transposasa en condiciones que son óptimas para la enzima, por ejemplo, en las condiciones de tampón, sal y temperatura que recomienda el fabricante. Por lo tanto, la transposasa puede considerarse "funcional" si tiene al menos aproximadamente el 50 %, por ejemplo, al menos aproximadamente el 60 %, aproximadamente 70 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 96 %, aproximadamente 97 %, aproximadamente 98 %, aproximadamente 99 %, o aproximadamente 100 % de la actividad de la enzima transposasa en condiciones que son óptimas para la enzima.



aproximadamente 99 % o aproximadamente el 100 % de actividad con relación a la actividad de la transposasa en condiciones que son óptimas para la enzima transposasa.

En un ejemplo no limitante, la mezcla de reacción comprende un transposoma en una solución tamponada (por ejemplo, Tris-acetato) que tiene un pH de aproximadamente 6,5 a aproximadamente 8,5, por ejemplo, de aproximadamente 7,0 a aproximadamente 8,0 tal como de aproximadamente 7,5. Alternativamente o adicionalmente, la mezcla de reacción puede usarse a cualquier temperatura adecuada, tal como aproximadamente 10 °C a aproximadamente 55 °C, por ejemplo., aproximadamente 10 °C a aproximadamente 54 °C, aproximadamente 11 °C a aproximadamente 53 °C, aproximadamente 12 °C a aproximadamente 52 °C, aproximadamente 13 °C a aproximadamente 51 °C, aproximadamente 14 °C a aproximadamente 50 °C, aproximadamente 15 °C a aproximadamente 49 °C, aproximadamente 16 °C a aproximadamente 48 °C, aproximadamente 17 °C a aproximadamente 47 °C, por ejemplo, aproximadamente 10 °C, aproximadamente 12 °C, aproximadamente 15 °C, aproximadamente 18 °C, aproximadamente 20 °C, aproximadamente 22 °C, aproximadamente 25 °C, aproximadamente 28 °C, aproximadamente 30 °C, aproximadamente 33 °C, aproximadamente 35 °C, o aproximadamente 37 °C, preferentemente aproximadamente 30 °C a aproximadamente 40 °C, por ejemplo, aproximadamente 37 °C. En algunas modalidades, el transposoma puede ponerse en contacto con la muestra biológica durante aproximadamente 10 minutos a aproximadamente una hora. En algunas modalidades, el transposoma puede ponerse en contacto con la muestra biológica durante aproximadamente 20, aproximadamente 30, aproximadamente 40 o aproximadamente 50 minutos. En algunas modalidades, el transposoma puede ponerse en contacto con la muestra biológica durante aproximadamente 1 hora a aproximadamente 4 horas.

En algunas modalidades, la enzima transposasa del complejo transposoma es una transposasa Tn5, o un derivado funcional o variante de esta. (Ver, por ejemplo, Reznikoff y otros, WO 2001/009363, las patentes de Estados Unidos núms. 5,925,545, 5,965,443, 7,083,980 y 7,608,434, y Goryshin y Reznikoff, J. Biol. Chem. 273:7367, (1998)). En algunas modalidades, la transposasa Tn5 es una hipertransposasa Tn5, o un derivado funcional o variante de esta (patente de Estados Unidos 9,790,476). Por ejemplo, la transposasa Tn5 puede ser una proteína de fusión (por ejemplo, una proteína de fusión Tn5). La Tn5 es un miembro de la superfamilia de proteínas RNasa. El transposón Tn5 es un transposón compuesto en el cual dos secuencias de inserción casi idénticas (IS50L e IS50R) flanquean tres genes de resistencia a antibióticos. Cada IS50 contiene dos secuencias de extremos (ES) de 19 pb invertidas, un extremo exterior (OE) y un extremo interior (IE). La enzima transposasa Tn5 de tipo salvaje generalmente está inactiva (por ejemplo, tiene baja actividad del evento de transposición). Sin embargo, las sustituciones de aminoácidos pueden resultar en variantes o derivados hiperactivos. En un ejemplo no limitante, la sustitución de aminoácidos, L372P, sustituye un aminoácido de leucina por un aminoácido de prolina que resulta en una ruptura de la hélice alfa, de esta manera se induce un cambio conformacional en el dominio C-terminal. La ruptura de la hélice alfa separa suficientemente el dominio C-terminal y el dominio N-terminal como para promover una mayor actividad del evento de transposición (Ver, Reznikoff, W.S., Tn5 as a model for understanding DNA transposition, *Mol Microbiol*, 47(5): 1199-1206 (2003)). Otras sustituciones de aminoácidos que resultan en la Tn5 hiperactiva se conocen en la técnica. Por ejemplo, la avidez mejorada de la enzima transposasa modificada (por ejemplo, la enzima transposasa Tn5 modificada) para las secuencias repetitivas para los extremos OE (mutación de clase (1)) puede lograrse al proporcionar un residuo de lisina en el aminoácido 54, que es el ácido glutámico en la enzima transposasa Tn5 de tipo salvaje (ver patente de Estados Unidos núm. 5.925,545). La mutación altera fuertemente la preferencia de la enzima transposasa modificada (por ejemplo, la enzima transposasa Tn5 modificada) para los extremos OE, como opuesto a los extremos IE. La unión más alta de esta mutación, conocida como EK54, a los extremos de OE resulta en una velocidad de transposición que es aproximadamente 10 veces mayor que la que se observa con la enzima transposasa de tipo salvaje (por ejemplo, la enzima transposasa Tn5 de tipo salvaje). Un cambio similar en la posición 54 a valina (por ejemplo, EV54) también resulta en algún aumento de la unión/transposición para los extremos OE, al igual que un cambio de treonina a prolina en la posición 47 (por ejemplo, TP47; aproximadamente 10 veces mayor) (ver patente de Estados Unidos núm. 5.925,545).

Se conocen otros ejemplos de enzimas transposasas modificadas (por ejemplo, las enzimas transposasas Tn5 modificadas). Por ejemplo, una enzima transposasa Tn5 modificada que se diferencia de la enzima transposasa Tn5 de tipo salvaje en que se une a las secuencias repetidas del ADN donante con mayor avidez que la enzima transposasa Tn5 de tipo salvaje y también es menos probable que la enzima transposasa de tipo salvaje asuma una forma multimérica inactiva (patente de Estados Unidos núm. 5,925,545). Además, se conocen en la técnica, técnicas que describen generalmente la introducción de cualquier elemento transponible (por ejemplo, Tn5) a partir de un ADN donante (por ejemplo, la secuencia adaptadora, por ejemplo, los adaptadores Nextera (por ejemplo, adaptador superior y parte inferior) en una diana. (Ver, por ejemplo, la patente de Estados Unidos núm. 5,925,545). Un estudio adicional identifica clases de mutaciones que resultan en una enzima transposasa modificada (por ejemplo, la enzima transposasa Tn5 modificada) (ver la patente de Estados Unidos núm. 5,965,443). Por ejemplo, una enzima transposasa modificada (por ejemplo, la enzima transposasa Tn5 modificada) con una "mutación de clase 1" se une a secuencias repetidas del ADN donante con mayor avidez que la enzima transposasa Tn5 de tipo salvaje. Adicionalmente, una enzima transposasa modificada (por ejemplo, la enzima transposasa Tn5 modificada) con una "mutación de clase 2" es menos probable que la enzima transposasa Tn5 de tipo salvaje asuma una forma multimérica inactiva. Se ha demostrado que una enzima transposasa modificada que contiene tanto una mutación

de clase 1 como una de clase 2 puede inducir al menos aproximadamente 100 veces (+10 %) más transposición que la enzima transposasa de tipo salvaje, cuando se prueba en combinación con un ensayo de conjugación in vivo como se describió en Weinreich, M.D., "Evidence that the cis Preference of the Tn5 Transposase is Caused by Nonproductive Multimerization," *Genes and Development* 8:2363-2374 (1994) (Ver, por ejemplo, la patente de Estados Unidos núm. 5,965,443). Además, bajo condiciones suficientes, la transposición mediante el uso de la enzima transposasa modificada (por ejemplo, la enzima transposasa Tn5 modificada) puede ser mayor. Una enzima transposasa modificada que contiene solo una mutación de clase 1 puede unirse a las secuencias repetidas con una avidez suficientemente mayor que la enzima transposasa Tn5 de tipo salvaje de manera que una enzima transposasa Tn5 induce aproximadamente de 5 a aproximadamente 50 veces más transposición que la enzima transposasa de tipo salvaje, cuando se mide in vivo. Una enzima transposasa modificada que contiene solo una mutación de clase 2 (por ejemplo, una mutación que reduce la enzima transposasa Tn5 de asumir una forma inactiva) es lo suficientemente menos probable que la enzima transposasa Tn5 de tipo salvaje para asumir la forma multimérica tal que la enzima transposasa Tn5 también induce aproximadamente de 5 a aproximadamente 50 veces más transposición que la enzima transposasa de tipo salvaje, cuando se mide in vivo (ver patente de Estados Unidos núm. 5,965,443)

Otros métodos para usar una enzima transposasa modificada (por ejemplo, la enzima transposasa Tn5 modificada) se describen además generalmente en la patente de Estados Unidos núm. 5,965,443 y la patente de Estados Unidos núm. 9,790,476. Por ejemplo, una enzima transposasa modificada podría proporcionar marcadores selectivos para el ADN diana, para proporcionar regiones portátiles de homología a un ADN diana, para facilitar la inserción de secuencias de ADN especializadas en el ADN diana, para proporcionar sitios de unión a cebadores o marcadores para la secuenciación del ADN, o para facilitar la producción de fusiones genéticas para la expresión génica. Estudios y mapeo del dominio proteico, así como también, para reunir otras combinaciones convenientes de secuencias de ADN (genética combinatoria) (patente de Estados Unidos núm. 5,965,443).

Aún se conocen otros métodos para insertar un elemento transponible (por ejemplo, un transposón) en ubicaciones aleatorias o semialeatorias en un ácido nucleico cromosómico o extracromosómico. Por ejemplo, los métodos incluyen una etapa de combinación en un ácido nucleico de una muestra biológica (por ejemplo, ADN genómico) con un complejo sináptico que comprende un complejo de una enzima transposasa Tn5 con una secuencia que comprende un par de secuencias de nucleótidos adaptadas para interactuar operativamente con la enzima transposasa Tn5 y un elemento transponible (por ejemplo, un transposón) bajo condiciones que median eventos de transposición en el ADN genómico. En este método, un complejo sináptico puede formarse in vitro bajo condiciones que desfavorezcan o eviten que los complejos sinápticos se sometan a un evento de transposición. La frecuencia de transposición (por ejemplo, los eventos de transposición) puede aumentarse mediante el uso de una enzima transposasa hiperactiva (por ejemplo, una enzima transposasa mutante) o un elemento transponible (por ejemplo, un transposón) que contiene secuencias bien adaptadas para eventos de transposición eficientes en presencia de una enzima transposasa hiperactiva (por ejemplo, una enzima transposasa Tn5 hiperactiva), o ambas (patente de Estados Unidos núm. 6,159,736).

Los métodos, composiciones y kits para tratar el ácido nucleico, y en particular, los métodos y composiciones para fragmentar y marcar el ADN mediante el uso de composiciones de transposón se describen en detalle en la publicación de solicitud de patente de Estados Unidos núm. US 2010/0120098, la publicación de solicitud de patente de Estados Unidos núm. US2011/0287435, y Satpathy, A.T., y otros, Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T-cell exhaustion, *Nat Biotechnol.*, 37, 925-936 (2019).

Puede usarse cualquier enzima transposasa con actividad de fragmentomarcaje, por ejemplo, cualquier enzima transposasa capaz de fragmentar el ADN e insertar oligonucleótidos (por ejemplo, adaptadores, por ejemplo, adaptadores de índice Nextera) a los extremos del ADN fragmentado (por ejemplo, el marcado). En algunas modalidades, la transposasa es cualquier transposasa capaz de la transposición conservadora. En algunas modalidades, la transposasa es una transposasa de corte y pega. Otros tipos de transposasas se conocen en la técnica y están dentro del alcance de esta descripción. Por ejemplo, las enzimas transposasa adecuadas incluyen, sin limitación, Mos-1, HyperMu<sup>TM</sup>, Ts-Tn5, Ts-Tn5059, Hermes, Tn7, una transposasa de las especies de *Vibrio* (ver, por ejemplo, la solicitud de patente de Estados Unidos núm. 20120301925A1 y WO 2015/069374), o cualquier variante funcional o derivada de las enzimas transposasa enumeradas anteriormente.

En algunas modalidades, una variante hiperactiva de la enzima transposasa Tn5 es capaz de mediar la fragmentación del ADN de doble hebra y la ligazón de oligonucleótidos sintéticos (por ejemplo, los adaptadores Nextera) en ambos extremos 5' del ADN en una reacción que toma un corto período de tiempo (por ejemplo, aproximadamente 5 minutos). Sin embargo, como las secuencias de extremos de tipo salvaje tienen una actividad relativamente baja, a veces se reemplazan in vitro por secuencias de extremos del mosaico (ME) hiperactivas. Un complejo de la transposasa Tn5 con ME de 19 pb facilita la transposición, siempre y cuando el ADN intermedio sea lo suficientemente largo como para acercar dos de estas secuencias para formar un homodímero de la enzima transposasa Tn5 activa.

En algunas modalidades, la enzima transposasa Tn5, o variante funcional o derivado de esta, comprende una secuencia de aminoácidos que tiene al menos un 80 % de identidad de secuencia con la SEQ ID NO: 1. En algunas

modalidades, la enzima transposasa Tn5, o variante funcional o derivado de esta, comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos aproximadamente un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia de aminoácidos con la SEQ ID NO: 1. En algunas modalidades, la enzima transposasa es un complejo con un adaptador que incluye una secuencia de extremo del transposón. En algunas modalidades, la secuencia de extremo del transposón Tn5 comprende una secuencia que tiene al menos un 80 % de identidad de secuencia con la SEQ ID NO: 2. En algunas modalidades, la secuencia de extremo del transposón Tn5 comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos aproximadamente un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia de aminoácidos con la SEQ ID NO: 2.

En algunas modalidades, la enzima transposasa es una enzima transposasa Mu, o una variante funcional o derivado de esta. En algunas modalidades, la enzima transposasa Mu, o una variante funcional o derivado de esta, comprende una secuencia de aminoácidos que tiene al menos un 80 % de identidad con la SEQ ID NO: 3. En algunas modalidades, la enzima transposasa Mu, o una variante funcional o derivado de esta, comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos aproximadamente un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia de aminoácidos con la SEQ ID NO: 3. En algunas modalidades, la secuencia de extremo del transposón Mu (por ejemplo, una secuencia de reconocimiento de la transposasa) comprende una secuencia que tiene al menos 80 % de identidad de secuencia con cualquiera de las SEQ ID NO: 4-9. En algunas modalidades, la secuencia de extremo del transposón Mu (por ejemplo, una secuencia de reconocimiento de la transposasa Mu) comprende una secuencia que tiene al menos aproximadamente un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia de aminoácidos con cualquiera de las SEQ ID NO: 4-9.

En algunas modalidades, la enzima transposasa es una transposasa de la familia ISR, o una variante funcional o derivado de esta. Por ejemplo, la transposasa de la familia ISR puede ser una transposasa de la familia ISR que se describe en la secuencia de referencia NCBI: WP\_012128611.1 y/o en la patente de Estados Unidos núm. 9,005,935. En algunas modalidades, la transposasa de la familia ISR, o una variante funcional o derivado de esta, comprende una secuencia de aminoácidos que tiene al menos un 80 % de identidad con la SEQ ID NO: 10. En algunas modalidades, la transposasa de la familia ISR, o una variante funcional o derivado de esta, comprende una secuencia de aminoácidos que tiene una identidad de secuencia de al menos aproximadamente un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia de aminoácidos con la SEQ ID NO: 10. En algunas modalidades, la secuencia de extremo del transposón de la familia ISR (por ejemplo, una secuencia de reconocimiento de la transposasa) comprende una secuencia que tiene al menos un 80 % de identidad con cualquiera de las SEQ ID NO: 11-13. En algunas modalidades, la secuencia de extremo del transposón de la familia ISR (por ejemplo, una secuencia de reconocimiento de la transposasa) comprende una secuencia que tiene al menos aproximadamente un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o 100 % de identidad de secuencia de aminoácidos con SEQ ID NO: 11-13.

Los adaptadores (por ejemplo, los adaptadores Nextera) en el complejo con la enzima transposasa (por ejemplo, que forman parte del transposoma, por ejemplo, los MEDS descritos en la presente descripción) pueden incluir parcialmente oligonucleótidos de doble hebra. En algunas modalidades, hay un primer adaptador y un segundo adaptador. En algunas modalidades, el primer adaptador puede ser un complejo con un primer monómero. En algunas modalidades, el segundo adaptador puede ser un complejo con un segundo monómero. En algunas modalidades, el complejo del primer monómero con el primer adaptador y el complejo del segundo monómero con el segundo monómero pueden ensamblarse para formar un dímero. En algunas modalidades, la porción de doble hebra de los adaptadores contiene secuencias de extremos del transposón (por ejemplo, las secuencias de extremos del mosaico (ME)). En algunas modalidades, la porción de una sola hebra de los adaptadores (por ejemplo, los adaptadores de índice Nextera) (saliente 5') contiene el dominio funcional o la secuencia que se incorpora en el ADN marcado. En algunas modalidades, los adaptadores pueden ser adaptadores Nextera (por ejemplo, el adaptador de índice) (por ejemplo, los reactivos que incluyen el kit Nextera ADN Library Prep para ATAC-seq (ya no disponible), la enzima de marcaje de ADN TDE-1 (núm. de catálogo 15027865), el tampón de marcaje de ADN TD (núm. de catálogo 15027866), disponible en Illumina, Illumina.com). En algunas modalidades, la secuencia que se incorpora en el ADN marcado es una secuencia complementaria a un dominio de captura de una sonda de captura. En algunas modalidades, la secuencia complementaria al dominio de captura de la sonda de captura es una secuencia de extremo del transposón. En tales modalidades, el dominio funcional está en la hebra del adaptador que se ligará a la sonda de captura. En otras palabras, el dominio funcional puede ubicarse aguas arriba (por ejemplo, hacia 5') de la secuencia ME, por ejemplo, en el saliente 5' del adaptador.

Los adaptadores (por ejemplo, los adaptadores de índice Nextera, por ejemplo, los primeros y segundos adaptadores) ligados al ADN marcado pueden ser cualquier secuencia adecuada. Por ejemplo, la secuencia puede ser una secuencia viral. En algunas modalidades, la secuencia puede ser una secuencia CRISPR. En algunas modalidades, el adaptador (por ejemplo, los oligonucleótidos) ligado al ADN marcado puede ser una secuencia guía CRISPR. En algunas modalidades, la secuencia guía CRISPR puede dirigirse a una secuencia de interés (por ejemplo, un locus genómico de interés, por ejemplo, un gen específico).

En algunas modalidades, la secuencia ME es una secuencia de reconocimiento de la transposasa Tn5. En algunas modalidades, la secuencia de extremo del mosaico (por ejemplo, ME) es una secuencia de reconocimiento

de la transposasa Mu. En algunas modalidades, la secuencia ME es una secuencia de reconocimiento de la transposasa de especies de *Vibrio*.

En algunas modalidades, una composición que comprende una enzima transposasa (por ejemplo, cualquier enzima transposasa descrita en la presente descripción) en un complejo con adaptadores (por ejemplo, los primeros y segundos adaptadores en un complejo con los primeros y segundos monómeros, respectivamente) que comprenden secuencias de extremos del transposón (por ejemplo, secuencias de extremos del mosaico) se usa en un método para marcar espacialmente ácidos nucleicos en una muestra biológica. En algunas modalidades, una composición que comprende una enzima transposasa comprende además un dominio que se une a una sonda de captura como se describe en la presente descripción (por ejemplo, el adaptador Nextera, por ejemplo, el primer adaptador) y un segundo adaptador se usa en un método para marcar espacialmente los ácidos nucleicos de una muestra biológica, tal como cualquiera de los métodos descritos en la presente descripción.

En algunas modalidades, la enzima transposasa puede tener la forma de un transposoma que comprende adaptadores (MEDS) en los que el saliente 5' puede fosforilarse. En algunas modalidades, los adaptadores (por ejemplo, los adaptadores Nextera, por ejemplo, los primeros y segundos adaptadores) pueden fosforilarse antes de su ensamble con la enzima transposasa para formar el transposoma. En algunas modalidades, la fosforilación de los adaptadores puede ocurrir cuando se forman complejos con una enzima transposasa (por ejemplo, una fosforilación in situ en el transposoma).

En algunas modalidades, el saliente 5' del adaptador no se fosforila antes de su ensamble en el transposoma. En tales modalidades, el saliente 5' puede tener grupos hidroxilo 5' accesibles fuera de la secuencia de extremo de la transposasa del mosaico. En algunas modalidades, la fosforilación del saliente 5' de los complejos transposoma ensamblados puede lograrse al exponer estos extremos 5' de los complejos transposomas a una polinucleótido cinasa (por ejemplo, la polinucleótido cinasa T4 (PNK-T4)) en presencia de ATP.

En algunas modalidades, el marcaje del ADN genómico de la muestra biológica con un transposoma (por ejemplo, cualquiera de los transposomas descritos en la presente descripción) puede comprender una etapa adicional de fosforilación de los extremos 5' de los adaptadores (por ejemplo, los salientes 5' de los adaptadores Nextera, por ejemplo, MEDS) en el complejo transposoma.

En algunas modalidades, los métodos proporcionados en la presente descripción comprenden una etapa de proporcionar un transposoma tratado para fosforilar los extremos 5' de los adaptadores (por ejemplo, los salientes 5' de los adaptadores Nextera (por ejemplo, los primeros y segundos adaptadores), por ejemplo, MEDS) en el complejo transposoma, de esta manera se fragmenta la muestra biológica con un transposoma tratado para fosforilar los extremos 5' de los adaptadores en el complejo transposoma.

Cualquier enzima y/o condiciones adecuadas pueden usarse para fosforilar los extremos 5' de los adaptadores (por ejemplo, los salientes 5' de los adaptadores, por ejemplo, MEDS) en el complejo transposoma, por ejemplo, PNK-T4 o PNK-T7. En algunas modalidades, la reacción de fosforilación puede llevarse a cabo al poner en contacto el transposoma con una polinucleótido cinasa (por ejemplo, PNK-T4 o PNK-T7) en una solución tamponada (por ejemplo, Tris-HCl, a un pH de aproximadamente 7,0 a aproximadamente 8,0, por ejemplo, de aproximadamente 7,6) en aproximadamente de 20 a aproximadamente 40 °C, por ejemplo, aproximadamente 25 a aproximadamente 37 °C, durante aproximadamente 1 a aproximadamente 60 minutos, por ejemplo, de aproximadamente 5 a aproximadamente 50, de aproximadamente 10 a aproximadamente 40, de aproximadamente 20 a aproximadamente 30 minutos.

En algunas modalidades, pueden realizarse el relleno de espacios y ligazón de rupturas en el ADN fragmentado (por ejemplo, el marcado). Por ejemplo, el evento de transposición de Tn5 da como resultado un espacio de 9 pares de bases entre una secuencia de extremo del transposón insertada y el ADN genómico. En algunas modalidades, el relleno de espacios se realiza entre la secuencia de extremo del transposón insertado y el ADN genómico fragmentado posteriormente se liga la secuencia del espacio relleno al ADN genómico.

En algunas modalidades, se liberan las secuencias de extremos del transposón adyacentes al espacio de 9 pares de bases seguido del ADN genómico fragmentado. En algunos ejemplos, las secuencias de extremos del transposón adyacentes al espacio se liberan (por ejemplo, se eliminan) del ADN genómico fragmentado (por ejemplo, se liberan de la secuencia de extremo del transposón complementaria). En algunas modalidades, las secuencias de extremos del transposón liberadas no se ligan al oligonucleótido de férula (por ejemplo, las secuencias de extremos del transposón no ligadas). En algunas modalidades, las secuencias de extremos del transposón no ligadas se liberan con un gradiente de calor. En algunas modalidades, las secuencias de extremos del transposón ligadas se ligan a la sonda de captura. En algunas modalidades, el oligonucleótido de férula se hibrida con el dominio de captura de la sonda de captura, o una porción del mismo, y la secuencia de extremo del transposón restante (por ejemplo, la secuencia de extremo del transposón ligada). En algunas modalidades, se realiza una reacción de relleno de espacios. En algunas modalidades, el relleno de espacios se produce entre el oligonucleótido de férula y el ADN genómico fragmentado. Por ejemplo, una polimerasa que rellena el espacio facilita la extensión del ácido nucleico entre el oligonucleótido de férula y el extremo fragmentado del ADN genómico, de esta manera "rellena el espacio: entre el oligonucleótido de férula y el ADN genómico fragmentado (por ejemplo, una porción en la cual se incluye la secuencia de extremo del transposón liberada).

En algunas modalidades, las secuencias de extremos del transposón no ligadas (por ejemplo, las secuencias de extremos del transposón adyacentes a un espacio de 9 pares de bases) se liberan con un gradiente de calor de aproximadamente 20 °C a aproximadamente 90 °C, de aproximadamente 25 °C a aproximadamente 85 °C, de aproximadamente 30 °C a aproximadamente 80 °C, de aproximadamente 35 °C a aproximadamente 75 °C, de aproximadamente 40 °C a aproximadamente 75 °C, de aproximadamente 45 °C a aproximadamente 75 °C, 50 °C a aproximadamente 75 °C o aproximadamente 50 °C a aproximadamente 70 °C. En algunas modalidades, liberar las secuencias de extremos del transposón no ligadas ocurre a aproximadamente 25 °C, aproximadamente 26 °C, aproximadamente 27 °C, aproximadamente 28 °C, aproximadamente 29 °C, aproximadamente 30 °C, aproximadamente 31 °C, aproximadamente 32 °C, aproximadamente 33 °C, aproximadamente 34 °C, aproximadamente 35 °C, aproximadamente 36 °C, aproximadamente 37 °C, aproximadamente 38 °C, aproximadamente 39 °C, aproximadamente 40 °C, aproximadamente 41 °C, aproximadamente 42 °C, aproximadamente 43 °C, aproximadamente 44 °C, aproximadamente 45 °C, aproximadamente 46 °C, aproximadamente 47 °C, aproximadamente 48 °C, aproximadamente 49 °C, aproximadamente 50 °C, aproximadamente 51 °C, aproximadamente 52 °C, aproximadamente 53 °C, aproximadamente 54 °C, aproximadamente 55 °C, aproximadamente 56 °C, aproximadamente 57 °C, aproximadamente 58 °C, aproximadamente 59 °C, aproximadamente 60 °C, aproximadamente 70 °C, aproximadamente 71 °C, aproximadamente 72 °C, aproximadamente 73 °C, aproximadamente 74 °C, aproximadamente 75 °C, aproximadamente 76 °C, aproximadamente 77 °C, aproximadamente 78 °C, aproximadamente 79 °C, aproximadamente 80 °C, aproximadamente 81 °C, aproximadamente 82 °C, aproximadamente 83 °C, aproximadamente 84 °C, aproximadamente 85 °C, aproximadamente 86 °C, aproximadamente 87 °C, aproximadamente 88 °C, aproximadamente 89 °C, o aproximadamente 90 °C.

En algunas modalidades, liberar las secuencias de extremos del transposón no ligadas (por ejemplo, las secuencias de extremos del transposón adyacentes al espacio de 9 pares de bases) se liberan con un gradiente de calor durante aproximadamente 10 minutos a aproximadamente 150 minutos, de aproximadamente 20 minutos a aproximadamente 140 minutos, de aproximadamente 30 minutos a aproximadamente 130 minutos, de aproximadamente 40 minutos a aproximadamente 120 minutos, de aproximadamente 40 minutos a aproximadamente 110 minutos, de aproximadamente 50 minutos a aproximadamente 110 minutos, de aproximadamente 60 minutos a aproximadamente 100 minutos, de aproximadamente 70 minutos a aproximadamente 90 minutos, de aproximadamente 10 minutos, aproximadamente 15 minutos, aproximadamente 20 minutos, aproximadamente 25 minutos, aproximadamente 30 minutos, aproximadamente 35 minutos, aproximadamente 40 minutos, aproximadamente 45 minutos, aproximadamente 50 minutos, aproximadamente 55 minutos, aproximadamente 60 minutos, aproximadamente 65 minutos, aproximadamente 70 minutos, aproximadamente 75 minutos, aproximadamente 80 minutos, aproximadamente 85 minutos, aproximadamente 90 minutos, aproximadamente 95 minutos, aproximadamente 100 minutos, aproximadamente 105 minutos, aproximadamente 110 minutos, aproximadamente 115 minutos, aproximadamente 120 minutos, aproximadamente 125 minutos, aproximadamente 130 minutos, aproximadamente 135 minutos, aproximadamente 140 minutos, aproximadamente 145 minutos, aproximadamente 150 minutos, aproximadamente 155 minutos, aproximadamente 160 minutos, aproximadamente 165 minutos, aproximadamente 170 minutos, aproximadamente 175 minutos, aproximadamente 180 minutos, aproximadamente 185 minutos, aproximadamente 190 minutos, aproximadamente 195 minutos, aproximadamente 200 minutos, aproximadamente 205 minutos, aproximadamente 210 minutos, aproximadamente 215 minutos, aproximadamente 220 minutos, aproximadamente 225 minutos, aproximadamente 230 minutos, aproximadamente 235 minutos, aproximadamente 240 minutos, aproximadamente 245 minutos, o aproximadamente 250 minutos.

En algunas modalidades, el marcado espacial del ADN genómico puede realizarse mediante la inserción de la secuencia del transposón en el ADN genómico con los adaptadores descritos en la presente descripción. Puede realizarse una etapa de amplificación a los adaptadores con cebadores (por ejemplo, los adaptadores insertados en el ADN genómico). Los productos amplificados pueden contener ADN genómico accesible que puede marcarse espacialmente mediante los métodos descritos en la presente descripción.

En algunas modalidades, el marcado espacial del ADN genómico puede realizarse mediante los complejos transposoma inmovilizados en la superficie del sustrato. En algunas modalidades, el marcaje espacial del ADN genómico puede realizarse mediante los complejos transposoma inmovilizados en un elemento (por ejemplo, una perla). En algunas modalidades, los complejos transposoma se ensamblan antes de añadir la muestra biológica al sustrato o los elementos. En algunas modalidades, los complejos transposoma se ensamblan después de añadir la muestra biológica al sustrato o elementos en un sustrato. Por ejemplo, un sustrato con código de barras espacial (por ejemplo, una matriz) puede incluir una pluralidad de sondas de captura que incluyen una secuencia de extremo del mosaico (por ejemplo, una secuencia de reconocimiento de transposasa). La secuencia de extremo del mosaico puede estar en el extremo 3' de la sonda de captura (por ejemplo, la sonda de captura se inmoviliza por su extremo 5' y la secuencia de extremo del mosaico está en el extremo 3' más lejano de la sonda de captura). La secuencia de extremo del mosaico puede ser una secuencia de extremo del mosaico para cualquiera de las enzimas transposasa descritas en la presente descripción. La secuencia de extremo del mosaico (por ejemplo, una secuencia de reconocimiento de transposasa) puede hibridarse con una secuencia complementaria inversa (por ejemplo, un oligonucleótido). Por ejemplo, la secuencia de complemento inversa (por ejemplo, el complemento inverso a la secuencia de extremo del mosaico) puede hibridarse con la secuencia de extremo del mosaico de esta

manera se genera una porción de un ADN de doble hebra en la sonda de captura. El complemento inverso a la secuencia de extremo del mosaico (por ejemplo, un oligonucleótido) puede proporcionarse a la matriz con código de barras espacial antes de que la muestra biológica se proporcione al sustrato. En algunas modalidades, el complemento inverso a la secuencia de extremo del mosaico puede proporcionarse después de que la muestra biológica se proporcione al sustrato. Las enzimas transposasas pueden proporcionarse al sustrato y ensamblarse en la porción de doble hebra de la sonda de captura (por ejemplo, un oligonucleótido inverso del complemento y la secuencia de extremo del mosaico hibridadas entre sí) de esta manera se genera un complejo transposoma. Por ejemplo, puede formarse un homodímero del transposoma en la porción de doble hebra de la sonda de captura. Una muestra biológica puede proporcionarse al sustrato de manera que la posición de la sonda de captura en el sustrato puede correlacionarse con una posición (por ejemplo, la ubicación) en la muestra biológica. Los complejos transposomas pueden fragmentar (por ejemplo, marcar) y marcar espacialmente el ADN genómico.

En algunas modalidades, el marcaje espacial del ADN genómico puede realizarse al hibridar una sonda de captura de una sola hebra con el ADN marcado. En algunas modalidades, la sonda de captura de una sola hebra puede ser una secuencia degenerada. En algunas modalidades, la captura de la secuencia de una sola hebra puede ser a una secuencia aleatoria. La sonda de captura de una sola hebra puede tener un dominio funcional, un código de barras espacial, un identificador molecular único, un dominio de escisión o sus combinaciones. La sonda de captura de una sola hebra (por ejemplo, una secuencia aleatoria, una secuencia degenerada) puede hibridarse inespecíficamente con el ADN genómico marcado, de esta manera captura espacialmente el ADN marcado. Los métodos para las reacciones de extensión se conocen en la técnica y puede realizarse cualquier método de reacción de extensión adecuado descrito en la presente descripción.

#### Oligonucleótidos de férula

Como se usa en la presente descripción, el término "oligonucleótido de férula" se refiere a un oligonucleótido que, cuando se hibrida con otros polinucleótidos, actúa como una "férula" para reclutar y colocar los polinucleótidos uno al lado del otro de manera que puedan ligarse entre sí. En algunas modalidades, el oligonucleótido de férula es ADN o ARN. El oligonucleótido de férula puede incluir una secuencia de nucleótidos que es parcialmente complementaria a las secuencias de nucleótidos de dos o más oligonucleótidos diferentes. En algunas modalidades, el oligonucleótido de férula participa en la ligazón de un oligonucleótido "donante" y un oligonucleótido "aceptador". En algunas modalidades, pueden usarse una ARN ligasa, una ADN ligasa, u otra variedad de ligasa para ligar dos secuencias de nucleótidos entre sí.

En algunas modalidades, el oligonucleótido de férula puede tener entre aproximadamente 10 y aproximadamente 50 nucleótidos de longitud, por ejemplo, entre aproximadamente 10 y aproximadamente 45, aproximadamente 10 y aproximadamente 40, aproximadamente 10 y aproximadamente 35, aproximadamente 10 y aproximadamente 30, aproximadamente 10 y aproximadamente 25, o aproximadamente 10 y aproximadamente 20 nucleótidos de longitud. En algunas modalidades, el oligonucleótido de férula puede tener entre aproximadamente 15 y aproximadamente 50, aproximadamente 15 y aproximadamente 45, aproximadamente 15 y aproximadamente 40, aproximadamente 15 y aproximadamente 35, aproximadamente 15 y aproximadamente 30, aproximadamente 15 y aproximadamente 25, o aproximadamente 15 y aproximadamente 20 nucleótidos de longitud. En algunas modalidades, el ADN fragmentado puede incluir una secuencia que se añade (por ejemplo, se liga) durante la fragmentación del ADN. Por ejemplo, durante un evento de transposición (por ejemplo, un evento de transposición de Tn5) puede unirse una secuencia adicional (por ejemplo, secuencias de extremos del transposón) (por ejemplo, unidas covalentemente, por ejemplo, a través de un evento de ligazón) al ADN fragmentado (por ejemplo, el ADN genómico fragmentado, por ejemplo, el ADN genómico marcado). En algunas modalidades, el oligonucleótido de férula puede tener una secuencia que es complementaria (por ejemplo, un dominio de captura) al ADN fragmentado (por ejemplo, un ADN genómico fragmentado, por ejemplo, un ADN genómico fragmentado que incluye una secuencia que se añade durante la fragmentación del ADN, por ejemplo, un primer adaptador que se une durante la fragmentación del ADN, por ejemplo, una secuencia de extremo del transposón) y una secuencia que es complementaria al dominio de captura de la sonda de captura. En algunas modalidades, el oligonucleótido de férula puede verse como parte de la sonda de captura. Por ejemplo, la sonda de captura puede ser parcialmente de doble hebra donde una porción de la sonda de captura puede funcionar como un oligonucleótido de férula que se hibrida con una porción de la sonda de captura (por ejemplo, una porción de ADNdh) y puede tener una porción de simple hebra que puede hibridarse con (por ejemplo, un dominio de captura) el ADN fragmentado (por ejemplo, el ADN genómico fragmentado, por ejemplo, el marcado, por ejemplo, un adaptador que se une durante la fragmentación del ADN, por ejemplo, un adaptador Nextera). La primera secuencia adaptadora (por ejemplo, la secuencia unida al ADN fragmentado complementario al dominio de captura, por ejemplo, un adaptador Nextera) puede ser cualquier secuencia adecuada. En algunas modalidades, la secuencia adaptadora puede tener entre aproximadamente 15 y 25 nucleótidos de longitud. En algunas modalidades, la secuencia del adaptador puede ser de aproximadamente 16, aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 21, aproximadamente 22, aproximadamente 23 o aproximadamente 24 nucleótidos de longitud.

En algunas modalidades, un oligonucleótido de férula puede incluir una secuencia que es complementaria (por ejemplo, un dominio de captura) al primer adaptador unido al ADN fragmentado (por ejemplo, el ADN marcado). En algunas modalidades, el oligonucleótido de férula incluye una secuencia que no es perfectamente

complementaria al primer adaptador (por ejemplo, adaptador Nextera) unido al ADN fragmentado (por ejemplo, el ADN marcado), pero aún es capaz de hibridar la primera secuencia adaptadora (por ejemplo, la secuencia complementaria al dominio de captura) ligada al ADN fragmentado (por ejemplo, un adaptador Nextera).

Puede usarse cualquier variedad de sondas de captura que tenga dominios de captura que se hibriden con un oligonucleótido de férula de acuerdo con los materiales y métodos descritos en la presente descripción. Como se describe en la presente descripción, un dominio de captura es un dominio en una sonda de captura capaz de hibridar el oligonucleótido de férula para formar una sonda de captura parcialmente de doble hebra. Por ejemplo, una sonda de captura de una sola hebra puede tener una secuencia complementaria (por ejemplo, un dominio de captura) a una porción del oligonucleótido de férula, de manera que una sonda de captura parcialmente de doble hebra se forme con una porción de una sola hebra capaz de hibridar con las secuencias de extremos del transposón insertadas. En algunas modalidades, un oligonucleótido de férula incluye una secuencia que es complementaria (por ejemplo, al menos parcialmente complementaria) al dominio de captura de la sonda de captura.

En algunas modalidades, el oligonucleótido de férula incluye una secuencia que no es perfectamente complementaria al dominio de captura de la sonda de captura, pero aún así es capaz de hibridar el dominio de captura de la sonda de captura. En algunas modalidades, el oligonucleótido de férula puede hibridarse tanto con la secuencia de extremo del transposón, por ejemplo, a través de la secuencia adicional unida al ADN marcado y al dominio de captura de la sonda de captura a través de su secuencia complementaria al dominio de captura. En tales modalidades, donde el oligonucleótido de férula puede hibridarse tanto con la secuencia de extremo del transposón (por ejemplo, adaptador Nextera, la secuencia adicional unida al ADN fragmentado por ejemplo, el ADN marcado), como con el dominio de captura de la sonda de captura, el oligonucleótido de férula puede verse como parte de la sonda de captura.

En algunas modalidades, el oligonucleótido de férula puede tener un dominio de captura que es un homopolímero. Por ejemplo, el dominio de captura puede ser un dominio de captura poli(T).

En algunas modalidades, un oligonucleótido de férula puede facilitar que se ligen el ADN marcado y la sonda de captura. Puede usarse cualquier variedad de ligasas adecuadas que se conocen en la técnica o descritas en la presente descripción. En algunas modalidades, la ligasa es una ADN ligasa T4. En algunas modalidades, la reacción de ligazón puede durar aproximadamente de 1 a aproximadamente 5 horas. En algunas modalidades, la reacción de ligazón puede durar aproximadamente 2, aproximadamente 3 o aproximadamente 4 horas. En algunas modalidades, después de la ligazón, puede realizarse la polimerización de desplazamiento de hebra. En algunas modalidades, puede usarse una ADN polimerasa para realizar la polimerización de desplazamiento de hebra. En algunas modalidades, la ADN polimerasa es ADN polimerasa I.

#### Análisis multiplex

La presente descripción describe métodos para permeabilizar muestras biológicas bajo condiciones suficientes para permitir el fragmentomarcado del ADN genómico. El ADN marcado puede capturarse a través de una sonda de captura (por ejemplo, una sonda de captura y un oligonucleótido de férula), sin embargo, a veces puede ser útil capturar simultáneamente el ADN marcado y otros ácidos nucleicos (por ejemplo, el ARNm). Por ejemplo, los perfiles de expresión de los transcritos pueden correlacionarse (o no) con la cromatina abierta. Dicho de otra manera, la presencia de transcritos puede correlacionarse con la cromatina abierta (por ejemplo, la cromatina accesible) correspondiente a los genes (por ejemplo, el ADN genómico) a partir de los cuales se transcribieron los transcritos.

La presente descripción describe métodos con respecto a la captura simultánea de ADN y ARNm marcados en matrices con código de barras espacial. Por ejemplo, una matriz con código de barras espacial puede tener una pluralidad de sondas de captura inmovilizadas en una superficie de sustrato. Alternativamente, una matriz con código de barras espacial puede tener una pluralidad de sondas de captura inmovilizadas en un elemento. En algunas modalidades, el elemento con una pluralidad de sondas de captura puede estar en un sustrato. Las sondas de captura pueden tener códigos de barras espaciales correspondientes a una posición (por ejemplo, una ubicación) en el sustrato. En algunas modalidades, las sondas de captura pueden tener además un identificador molecular único, uno o más dominios funcionales, y un dominio de escisión, o sus combinaciones. En algunas modalidades, la sonda de captura incluye un dominio de captura. En algunas modalidades, la sonda de captura puede ser una secuencia homopolimérica. Por ejemplo, de una manera no limitante, la secuencia homopolimérica puede ser una secuencia poli(T). En algunas modalidades, el ácido nucleico (por ejemplo, ARNm) puede capturarse por el dominio de captura mediante la unión (por ejemplo, la hibridación) de colas poli(A) de transcritos de ARNm. En algunas modalidades, el ADN marcado puede capturarse por el dominio de captura de la sonda de captura mediante la unión (por ejemplo, la hibridación) de un ADN marcado con una cola poli(A). Por ejemplo, después de fragmentar el ADN genómico, las polimerasas de relleno de espacios (por ejemplo, sin desplazamiento de hebra) y las ligasas pueden reparar los espacios y ligar la rupturas en el ADN marcado. En algunas modalidades, puede introducirse una secuencia complementaria al dominio de captura en el ADN fragmentado. Por ejemplo, una cola poli(A) puede añadirse al ADN marcado, de manera que el dominio de captura (por ejemplo, una secuencia poli(T)) de la sonda de captura pueda unirse (por ejemplo, hibridar) al ADN marcado con la cola poli(A) (ver, por ejemplo, el documento WO 2012/140224). En algunas modalidades, una cola poli(A) se añade al ADN marcado mediante una enzima



transferasa terminal. En algunas modalidades, la enzima transferasa terminal es una desoxinucleotidiltransferasa terminal (TdT), o una variante mutante de esta. La TdT es una polimerasa independiente (por ejemplo, no requiere una molécula como plantilla) que puede catalizar la adición de desoxinucleótidos al extremo 3' del extremo hidroxilo de las moléculas de ADN. Otras polimerasas independientes de una molécula como plantilla se conocen en la técnica. Por ejemplo, la polimerasa  $\pi$ , o una variante mutante de esta, puede usarse como una enzima de transferasa terminal (ver, por ejemplo, Kent, T., Polymerase  $\pi$  is a robust terminal transferase that oscillates between three different mechanisms during end-joining, *eLIFE*, 5: e13740 doi: 10.7554/eLife.13740, (2016)). Otros métodos para introducir una cola de poli(A) se conocen en la técnica. En algunas modalidades, una cola poli(A) puede introducirse en el ADN marcado mediante una polimerasa de no lectura de prueba. En algunas modalidades, una cola poli(A) puede introducirse en el ADN fragmentado mediante una polinucleótido cinasa.

En algunas modalidades, la enzima TDT generará marcapies con una cola poli(A) 3', de esta manera imita la cola poli(A) de un ARNm. En algunas modalidades, el dominio de captura (por ejemplo, la secuencia poli(T)) de la sonda de captura interactuaría con la cola poli(A) del ARNm y la cola poli(A) generada (por ejemplo, la sintetizada) añadida al ADN fragmentado (por ejemplo, el marcado), de esta manera captura simultáneamente el ADN fragmentado (por ejemplo, el ADN marcado) y el transcrito de ARNm. La cola poli(A) generada (por ejemplo, sintetizada) en el ADN fragmentado (por ejemplo, el ADN marcado) podría tener entre aproximadamente 10 nucleótidos a aproximadamente 30 nucleótidos de longitud. La cola poli(A) generada (por ejemplo, sintetizada) en el ADN fragmentado (por ejemplo, el ADN marcado) podría ser de aproximadamente 11, aproximadamente 12, aproximadamente 13, aproximadamente 14, aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18, aproximadamente 19, aproximadamente 20, aproximadamente 21, aproximadamente 22, aproximadamente 23, aproximadamente 24, aproximadamente 25, aproximadamente 26, aproximadamente 27, aproximadamente 28 o aproximadamente 29 nucleótidos de longitud.

Alternativamente o adicionalmente, en lugar de una reacción secuencial (por ejemplo, de dos etapas) (por ejemplo, relleno de espacios y ligazón, seguido de una transferasa terminal) el ADN fragmentado (por ejemplo, el marcado) puede ponerse en contacto con una polimerasa. Por ejemplo, la polimerasa puede ser una ADN polimerasa que puede realizar una reacción de extensión en el fragmento (por ejemplo, el ADN marcado). Puede usarse cualquier variedad de ADN polimerasas que se conocen en la técnica o descritas en la presente descripción. Los productos extendidos pueden capturarse y procesarse (por ejemplo, amplificarse y secuenciarse) mediante cualquier método descrito en la presente descripción.

Las etapas posteriores a la hibridación son idénticas como se describe en Ståhl P.L., y otros, Visualization and analysis of gene expression in tissue sections by spatial transcriptomics *Science*, vol. 353, 6294, pp. 78-82 (2016)). Composiciones

También se proporcionan en la presente descripción composiciones que incluyen sondas de captura, ADN marcado, oligonucleótidos de férula y una o más polimerasas. En algunas modalidades, un oligonucleótido de férula se hibrida con el dominio de captura de una sonda de captura. En algunas modalidades, un oligonucleótido de férula se hibrida con una secuencia de extremo del transposón de ADN genómico fragmentado (por ejemplo, el marcado). En algunas modalidades, un oligonucleótido de férula se hibrida con el dominio capturado de una sonda de captura y una secuencia de extremo del transposón del ADN genómico fragmentado. En algunas modalidades, la composición comprende una o más secuencias de extremos del transposón. En algunas modalidades, una o más secuencias de extremos del transposón se ligan a la sonda de captura. En algunas modalidades, una o más secuencias del transposón se liberan del ADN fragmentado ya sea antes o después de la ligazón del ADN genómico fragmentado a la sonda de captura. En algunas modalidades, la composición incluye una ligasa (por ejemplo, la ADN ligasa T4). En algunas modalidades, la composición incluye una polimerasa de relleno de espacios. En algunas modalidades, la composición incluye una ADN polimerasa. En algunas modalidades, la composición incluye una o más transposasas. En algunas modalidades, la composición comprende complejos transposomas.

En algunas modalidades, el dominio de captura de la sonda de captura se une al extremo del transposón (por ejemplo, sin la ayuda de un oligonucleótido de férula). En algunas modalidades, la composición incluye una polimerasa que desplaza la hebra. En algunas modalidades, la composición incluye una polimerasa de relleno de espacios.

#### qPCR y análisis

También se proporcionan en la presente descripción métodos y materiales para cuantificar la eficiencia de la captura. En algunas modalidades, la cuantificación de la eficiencia de la captura incluye la cuantificación de fragmentos capturados (por ejemplo, fragmentos de ADN genómico, por ejemplo, fragmentos de ADN marcados) a partir de cualquiera de los métodos de análisis espacial descritos en la presente descripción. En algunas modalidades, la cuantificación incluye PCR, qPCR, electroforesis, electroforesis capilar, espectroscopía de fluorescencia y/o espectrofotometría UV. En algunas modalidades, la qPCR incluye intercalar colorantes fluorescentes (por ejemplo, verde SYBR) y/o sondas marcadas fluorescentes (por ejemplo, sin limitación, sondas Taqman o sondas PrimeTime). En algunas modalidades, se usa un kit de cuantificación de la biblioteca NGS para la cuantificación. Por ejemplo, la cuantificación puede realizarse mediante el uso de un kit de cuantificación de



biblioteca KAPA (KAPA Biosystems), kit de cuantificación de biblioteca qPCR NGS (Agilent), sistema de cuantificación de biblioteca GeneRead (Qiagen) y/o un kit de cuantificación PerfeCTa NGS (Quantabio). En algunas modalidades que usan la qPCR para la cuantificación, la qPCR puede incluir, sin limitación, PCR digital, digital de gotas (ddPCR) y ddPCR-Tail. En algunas modalidades que usan electroforesis para la cuantificación, la electroforesis puede incluir, sin limitación, electroforesis automatizada (por ejemplo, TapeStation System, Agilent y/o Bioanalyzer, Agilent) y la electroforesis capilar (por ejemplo, Fragment Analyzer, Applied Biosystems). En algunas modalidades que usan espectroscopía para la cuantificación, la espectroscopía puede incluir, sin limitación, espectroscopía de fluorescencia (por ejemplo, Qubit, Thermo Fisher). En algunas modalidades, la NGS puede usarse para cuantificar la eficiencia de captura.

En algunas modalidades, la PCR cuantitativa (qPCR) se realiza en los marcados que se capturan. En algunas modalidades, el ADN fragmentado (por ejemplo, el ADN marcado) se amplifica, mediante cualquier método descrito en la presente descripción, antes de la captura. Por ejemplo, después de la captura del ADN fragmentado (por ejemplo, el ADN marcado), puede realizarse la qPCR de hibridación por ligazón y desplazamiento de la hebra. En algunas modalidades, puede usarse una ADN polimerasa para realizar la polimerización de desplazamiento de hebra. Puede usarse cualquier polimerasa de desplazamiento de hebra adecuada que se conoce en la técnica. En algunas modalidades, la ADN polimerasa es ADN polimerasa I. Como se ejemplifica en los ejemplos, la ADN polimerasa I puede incubarse para el desplazamiento de la hebra del ADN fragmentado (por ejemplo, ADN marcado) con reactivos (por ejemplo, BSA, dNTP, tampón). En algunas modalidades, la ADN polimerasa I puede incubarse con reactivos en el sustrato (por ejemplo, en un elemento, por ejemplo, un pocillo) durante aproximadamente 30 minutos a aproximadamente 2 horas. En algunas modalidades, la ADN polimerasa I puede incubarse con reactivos en el sustrato durante aproximadamente 40 minutos, aproximadamente 50 minutos, aproximadamente 60 minutos, aproximadamente 70 minutos, aproximadamente 80 minutos, aproximadamente 90 minutos, aproximadamente 100 minutos o aproximadamente 110 minutos. En algunas modalidades, la ADN polimerasa I puede incubarse con reactivos en el sustrato (por ejemplo, en un elemento, por ejemplo, un pocillo) a aproximadamente 35 °C a aproximadamente 40 °C. En algunas modalidades, la ADN polimerasa I puede incubarse con reactivos en el sustrato a aproximadamente 36 °C, aproximadamente 37 °C, aproximadamente 38 °C, o aproximadamente 39 °C. En algunas modalidades, la ADN polimerasa I puede incubarse con reactivos en el sustrato durante aproximadamente 1 hora a aproximadamente 37 °C.

Después de completar la hibridación del desplazamiento de la hebra, puede realizarse una reacción de qPCR. En algunas modalidades, las sondas de captura ligadas al ADN fragmentado (por ejemplo, el ADN marcado), pueden liberarse de la superficie del sustrato (por ejemplo, un elemento). En algunas modalidades, una solución (por ejemplo, una mezcla de liberación) puede incubarse con el sustrato para liberar las sondas de captura de la superficie del sustrato. La mezcla de liberación puede contener reactivos (por ejemplo, BSA, enzimas, tampón, etc.). Los métodos para liberar las sondas de captura del sustrato (por ejemplo, un elemento) se describen en la presente descripción. En algunas modalidades, una enzima puede escindir la sonda de captura. En algunas modalidades, la enzima puede ser la enzima USER (reactivo de escisión específico de uracilo). En algunas modalidades, la enzima USER puede incubarse con reactivos en el sustrato (por ejemplo, un elemento, por ejemplo, un pocillo) durante aproximadamente 30 minutos a aproximadamente 2 horas. En algunas modalidades, la enzima USER puede incubarse con reactivos en el sustrato durante aproximadamente 40 minutos, aproximadamente 50 minutos, aproximadamente 60 minutos, aproximadamente 70 minutos, aproximadamente 80 minutos, aproximadamente 90 minutos, aproximadamente 100 minutos o aproximadamente 110 minutos. En algunas modalidades, la enzima USER con reactivos en el sustrato (por ejemplo, un elemento, por ejemplo, un pocillo) a aproximadamente 35 °C a aproximadamente 40 °C. En algunas modalidades, la enzima USER puede incubarse con reactivos en el sustrato a aproximadamente 36 °C, aproximadamente 37 °C, aproximadamente 38 °C o aproximadamente 39 °C. En algunas modalidades, la enzima USER puede incubarse con reactivos en el sustrato durante aproximadamente 1 hora a aproximadamente 37 °C.

Después de la incubación con la enzima USER, pueden recolectarse las muestras (por ejemplo, las sondas de captura liberadas ligadas al ADN fragmentado (por ejemplo, el ADN marcado) en la mezcla de liberación, o una porción de esta). En algunas modalidades, el volumen de la muestra puede reducirse. Los métodos para reducir el volumen de la muestra se conocen en la técnica y puede usarse cualquier método adecuado. En algunas modalidades, la reducción del volumen de la muestra puede realizarse con un vacío rápido (por ejemplo, un SpeedVac). En algunas modalidades, la reducción del volumen de la muestra puede ser de aproximadamente un 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 % o aproximadamente 90 % de reducción del volumen de la muestra. En algunas modalidades, la reducción del volumen de la muestra puede estar aproximadamente entre un 80 % y 90 % de reducción del volumen de la muestra. En algunas modalidades, la reducción del volumen de la muestra puede ser aproximadamente un 81 %, aproximadamente 82 %, aproximadamente 83 %, aproximadamente 84 %, aproximadamente 85 %, aproximadamente 86 %, aproximadamente 87 %, aproximadamente 88 o aproximadamente 89 % de reducción del volumen de la muestra. En algunas modalidades, la reducción del volumen de la muestra puede ser aproximadamente un 85 % (por ejemplo, aproximadamente 10 µl después de la reducción del volumen de la muestra).

En algunas modalidades, puede realizarse una reacción de qPCR con el volumen de la muestra reducido. Como se describe en la presente descripción, puede realizarse cualquier método adecuado de qPCR. En algunas

modalidades, pueden usarse un 1x de KAPA HiFi HotStart Ready, 1x de EVA verde y cebadores. La amplificación puede realizarse de acuerdo con los métodos que se conocen en la técnica. Por ejemplo, la amplificación puede realizarse en consecuencia: 72 °C durante 10 minutos, 98 °C durante 3 minutos, seguido de un ciclo a 98 °C durante 20 segundos, 60 °C durante 30 segundos y 72 °C durante 30 segundos.

En algunas modalidades, pueden usarse uno o más pares de cebadores durante la reacción de la qPCR. En algunos ejemplos, un par de cebadores pueden cubrir la porción ligada (por ejemplo, el sitio de ligazón donde la sonda de captura y la secuencia adaptadora (por ejemplo, la secuencia unida al ADN fragmentado por ejemplo, el ADN marcado)). Por ejemplo, un par de cebadores cubre la porción ligada y la sonda de captura. Un producto de amplificación solo se detectará si se produce la ligazón, y no solo la hibridación. En algunas modalidades, un cebador diferente puede cubrir solo el ADN fragmentado (por ejemplo, el ADN marcado). En algunas modalidades, el par de cebadores que cubre el ADN fragmentado (por ejemplo, el ADN marcado) solo puede ser un control para la ligazón. En algunas modalidades, la qPCR puede realizarse con cualquiera de los nucleótidos marcados descritos en la presente descripción.

En algunas modalidades, las muestras pueden purificarse. En algunas modalidades, las muestras pueden purificarse de acuerdo con Lundin y otros, Increased Throughput by Parallelization of Library Preparation for Massive Sequencing, *PLOS ONE*, 5(4), doi.org/10.1371/journal.pone.0010029 (2010).

En algunas modalidades, puede determinarse la longitud promedio del ADN fragmentado que se captura (por ejemplo, el ADN marcado). En algunas modalidades, puede usarse un bioanalizador (por ejemplo, un bioanalizador 2100 (Agilent)). Puede usarse cualquier bioanalizador adecuado que se conoce en la técnica. En algunas modalidades, la qPCR y el análisis del bioanalizador pueden realizarse en genomas completos (por ejemplo, el ADN fragmentado que se purifica, por ejemplo, el ADN marcado). En algunas modalidades, la qPCR y el análisis del bioanalizador pueden realizarse en una muestra biológica inmovilizada (por ejemplo, una muestra biológica fija). Por ejemplo, los métodos descritos en la presente descripción (por ejemplo, prepermeabilización, permeabilización) pueden realizarse para capturar el ADN fragmentado (por ejemplo, el ADN marcado) y para optimizar la qPCR y el análisis del bioanalizador para diferentes muestras biológicas.

En algunas modalidades, después de la ligazón, puede realizarse una etapa de desnaturalización basada en la superficie. Dicho de otra manera, después de la ligazón del ADN fragmentado (por ejemplo, el ADN marcado) a la sonda de captura, seguido del desplazamiento de la hibridación de la hebra descrita en la presente descripción (por ejemplo, la ADN polimerasa I), puede realizarse una etapa de desnaturalización basada en la superficie en un flujo de trabajo paralelo. En algunas modalidades, una solución básica puede realizar la desnaturalización basada en la superficie. Por ejemplo, la solución básica puede desnaturalizar el ADN de doble hebra fragmentado que se captura (por ejemplo, el ADN marcado), de esta manera se generan las sondas de captura de una sola hebra ligadas al ADN fragmentado que se captura (por ejemplo, el ADN marcado). En algunas modalidades, la solución básica es KOH. En algunas modalidades, la solución básica es NaOH. En algunas modalidades, la solución básica puede ser de aproximadamente 1M de NaOH. Pueden usarse otras soluciones básicas en los métodos descritos en la presente descripción. En algunas modalidades, la solución básica puede aplicarse durante aproximadamente 1 minuto a aproximadamente 1 hora. En algunas modalidades, la solución básica puede aplicarse durante aproximadamente 10, aproximadamente 20, aproximadamente 30, aproximadamente 40 o aproximadamente 50 minutos. En algunas modalidades, la solución básica puede aplicarse durante aproximadamente 1 a aproximadamente 20 minutos. En algunas modalidades, la solución básica aproximadamente puede aplicarse durante aproximadamente 2, aproximadamente 3, aproximadamente 4, aproximadamente 5, aproximadamente 6, aproximadamente 7, aproximadamente 8, aproximadamente 9, aproximadamente 10, aproximadamente 11, aproximadamente 12, aproximadamente 13, aproximadamente 14 aproximadamente 15, aproximadamente 16, aproximadamente 17, aproximadamente 18 o aproximadamente 19 minutos. En algunas modalidades, la solución básica puede aplicarse a una temperatura de entre aproximadamente 30 °C a aproximadamente 40 °C. En algunas modalidades, la solución básica puede aplicarse a aproximadamente 31 °C, aproximadamente 32 °C, aproximadamente 33 °C, aproximadamente 34 °C, aproximadamente 35 °C, aproximadamente 36 °C, aproximadamente 37 °C, aproximadamente 38 °C o aproximadamente 39 °C. En algunas modalidades, la solución básica puede aplicarse durante aproximadamente 10 minutos a aproximadamente 37 °C.

En algunas modalidades, la etapa de desnaturalización puede exponer el ADN fragmentado (por ejemplo, el ADN marcado) a la hibridación mediante una sonda. En algunas modalidades, la sonda puede ser una sonda de oligonucleótidos. En algunas modalidades, la sonda de oligonucleótidos puede tener un marcador detectable (por ejemplo, cualquiera de la variedad de marcadores detectables descritos en la presente descripción). En algunas modalidades, el marcador detectable puede ser Cy5. En algunas modalidades, la sonda de oligonucleótidos puede marcarse con Cy5. En algunas modalidades, la sonda de oligonucleótidos marcada con Cy5 puede hibridarse con una secuencia complementaria en el ADN fragmentado (por ejemplo, el ADN marcado). En algunas modalidades, el oligonucleótido marcado con Cy5 puede hibridarse con la secuencia unida (por ejemplo, adaptador Nextera, por ejemplo, el primer adaptador o segundo adaptador) al ADN fragmentado (por ejemplo, el ADN marcado). En algunas modalidades, puede detectarse el marcador Cy5. Por ejemplo, detectar el marcador Cy5 en la sonda de oligonucleótidos puede revelar la ubicación espacial de los marcadores del ADN. En algunas modalidades, la muestra biológica se puede tefir (por ejemplo, tinción con hematoxilina y eosina). Los métodos de tinción de una

muestra biológica se conocen en la técnica y se describen en la presente descripción. En algunas modalidades, se pueden obtener imágenes de la muestra biológica.

#### Análisis del genoma completo

El análisis del genoma completo (por ejemplo, genómica espacial) también puede realizarse en muestras biológicas. Por ejemplo, los métodos de ATAC espacial descritos en la presente descripción se diseñan para capturar áreas accesibles (por ejemplo, "abiertas" o transcripcionalmente activas) de un genoma, sin embargo, también es posible capturar espacialmente todo el genoma (por ejemplo, el ADN). Para permitir la captura de todo el genoma, la estructura de la cromatina se interrumpe mediante la degradación de las histonas

Por lo tanto, se proporcionan en la presente descripción métodos para determinar la ubicación del ADN en una muestra biológica, el método incluye: (a) una muestra biológica en una matriz que incluye una pluralidad de sondas de captura, donde una sonda de captura de la pluralidad de sondas de captura incluye: (i) un código de barras espacial y (ii) un dominio de captura; (b) poner en contacto la muestra biológica con una proteasa, donde la proteasa es capaz de degradar una o más proteínas histonas, de esta manera libera el ADN; (c) poner en contacto un transposoma con la muestra biológica para insertar secuencias de extremos del transposón en el ADN genómico liberado, de esta manera genera ADN genómico fragmentado; (d) hibridar una secuencia de extremo del transposón del ADN fragmentado al dominio de captura; (e) liberar secuencias de extremos del transposón no unidas al dominio de captura; y (f) determinar (i) una secuencia del código de barras espacial o un complemento de la misma, y (ii) la totalidad o una porción de una secuencia del ADN, o un complemento del mismo, y usar las secuencias determinadas de (i) y (ii) para determinar la ubicación del ADN en la muestra biológica.

Por lo tanto, se proporcionan en la presente descripción métodos que incluyen tratar una muestra biológica con una proteasa capaz de degradar proteínas histonas lo que resulta en la generación del ADN genómico fragmentado. El ADN genómico fragmentado puede capturarse mediante el dominio de captura de una sonda de captura, donde, por ejemplo, la secuencia de extremo del transposón insertada en el ADN genómico incluye una secuencia complementaria a un dominio de captura de una sonda de captura. En algunas modalidades, el dominio de captura comprende una secuencia homopolimérica. En algunas modalidades, el dominio de captura comprende una secuencia única.

En algunas modalidades, una muestra biológica se permeabiliza al exponer la muestra a una proteasa capaz de degradar proteínas histonas. Como se usa en la presente descripción, el término "proteína histona" se refiere típicamente a una proteína histona enlazadora (por ejemplo, H1) y/o una proteína histona del núcleo (por ejemplo, H2A, H2B, H3 y H4). En algunas modalidades, una proteasa degrada proteínas histonas enlazadoras, proteínas histonas del núcleo o proteínas histonas enlazadoras y proteínas histonas del núcleo. Puede usarse cualquier proteasa adecuada capaz de degradar proteínas histonas en una muestra biológica. Los ejemplos no limitantes de proteasas capaces de degradar proteínas histonas incluyen proteasas inhibidas por leupeptina y TLCK (clorhidrato de tosil-L-lisil-clorometano), collagenasa, una proteasa codificada por el gen EUO de *Chlamydia trachomatis* serovar A, granzima A, una serina proteasa (por ejemplo, una tripsina o proteasa similar a tripsina, serina proteasa neutra, elastasa, catepsina G), una aspartil proteasa (por ejemplo, catepsina D), una enzima C1 de la familia de peptidasas (por ejemplo, catepsina L), una proteasa que es inhibida por el inhibidor de diazometano Z-Phe-Phe-CHN(2) o el inhibidor de epóxido E-64, una proteasa lisosomal, o una enzima azurofílica (por ejemplo, catepsina G, elastasa, proteinasa 3, serina proteasa neutra). En algunas modalidades, una serina proteasa es una enzima tripsina, una enzima similar a tripsina o una variante funcional o derivado de esta (por ejemplo, P00761; C0HK48; Q8IYP2; Q8BW11; Q6IE06; P35035; P00760; P06871; Q90627; P16049; P07477; P00762; P35031; P19799; P35036; Q29463; P06872; Q90628; P07478; P07146; P00763; P35032; P70059; P29786; P35037; Q90629; P35030; P08426; P35033; P35038; P12788; P29787; P35039; P35040; Q8NHM4; P35041; P35043; P35044; P54624; P04814; P35045; P32821; P54625; P35004; P35046; P32822; P35047; C0HKA5; C0HKA2; P54627; P35005; C0HKA6; C0HKA3; P52905; P83348; P00765; P35042; P81071; P35049; P51588; P35050; P35034; P35051; P24664; P35048; P00764; P00775; P54628; P42278; P54629; P42279; Q91041; P54630; P42280; C0HKA4) o una combinación de estas. En algunas modalidades, una enzima tripsina es P00761, P00760, Q29463, o una de sus combinaciones. En algunas modalidades, una proteasa capaz de degradar una o más proteínas histonas comprende una secuencia de aminoácidos con al menos un 80 % de identidad de secuencia con P00761, P00760 o Q29463. En algunas modalidades, una proteasa capaz de degradar una o más proteínas histonas comprende una secuencia de aminoácidos con al menos un 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con P00761, P00760 o Q29463. Una proteasa puede considerarse una variante funcional si tiene al menos un 50 %, por ejemplo, al menos un 60 %, 70 %, 80 %, 90 %, 95 %, 96 %, 97 %, 98, 99 % o 100 %, de la actividad con relación a la actividad de la proteasa en condición óptima para la enzima.

Adicionalmente, la proteasa puede contenerse en una mezcla de reacción (una solución), que también incluye otros componentes (por ejemplo, un tampón (por ejemplo, Tris-HCl, una sal, un quelante (por ejemplo, EDTA), detergente (por ejemplo, SDS)). La mezcla de reacción puede amortiguarse, con un pH de aproximadamente 6,5-8,5, por ejemplo, de aproximadamente 7,0-8,0. Adicionalmente, la mezcla de reacción puede usarse a cualquier temperatura adecuada, tal como aproximadamente 10-45 °C, por ejemplo aproximadamente 10-44 °C, 11-43 °C, 12-42 °C, 13-41 °C, 14-40 °C, 15-39 °C, 16-38, °C 17-37 °C, por ejemplo aproximadamente 10 °C, 12 °C, 15 °C,

18 °C, 20 °C, 22 °C, 25 °C, 28 °C, 30 °C, 33 °C, 35 °C o 37 °C, preferentemente aproximadamente de 30-40 °C, por ejemplo, aproximadamente 37 °C.

En algunas modalidades, la mezcla de reacción puede incubarse con la muestra biológica de aproximadamente 1 minuto a aproximadamente 30 minutos, de aproximadamente 5 minutos a aproximadamente 25 minutos, de aproximadamente 10 minutos a aproximadamente 20 minutos, o durante aproximadamente 15 minutos. En algunas modalidades, la mezcla de reacción se incuba con la muestra biológica durante 37 °C por aproximadamente 10 minutos.

En algunas modalidades, el método incluye obtener imágenes y/o teñir la muestra biológica, tal como por ejemplo, la tinción con hematoxilina y eosina.

En algunas modalidades, la sonda de captura incluye un dominio de escisión, uno o más dominios funcionales, un identificador molecular único o sus combinaciones.

En algunas modalidades, el método incluye una etapa de migración activa en donde el ADN genómico fragmentado migra a la matriz por la aplicación de un campo eléctrico.

En algunas modalidades, la hibridación en la etapa (d) comprende hibridar la secuencia de extremo del transposón, o una porción de esta, al dominio de captura, o una porción de esta, de la sonda de captura. En algunas modalidades, el método extiende un extremo 3' de la sonda de captura mediante el uso del ADN genómico fragmentado como una plantilla. En algunas modalidades, la etapa de extensión se realiza mediante el uso de una ADN polimerasa que tiene actividad de desplazamiento de la hebra. En algunas modalidades, el método incluye el relleno de espacios (por ejemplo, como se describe en la presente descripción) entre la secuencia de extremo del transposón y el ADN genómico fragmentado. En algunas modalidades, el transposoma incluye una enzima transposasa (por ejemplo, cualquiera de las enzimas transposasas descritas en la presente descripción), tal como una enzima transposasa Tn5, una enzima transposasa Mu, una enzima transposasa Tn7, una transposasa de las especies de *Vibrio*, o un derivado funcional o variantes de estas.

En algunas modalidades, la liberación en la etapa (d) comprende calentar la muestra biológica. En algunas modalidades, el calentamiento incluye calentar a una temperatura de aproximadamente 65 °C a 85 °C.

En algunas modalidades, determinar la ubicación del ADN en una muestra biológica incluye además analizar espacialmente todo el genoma de la muestra biológica.

## Ejemplos

### Ejemplo 1. Flujo de trabajo de ATAC para las matrices espaciales

Preparación de tejidos y flujo de trabajo de ATAC espacial

Las muestras de tejido se congelan instantáneamente en bloques de temperatura de corte óptima (OCT), se crioseccionan en secciones de 10 µm, se colocan en un portaobjetos de matriz espacial de Visium (10X Genomics, Inc.) y se calientan a 37 °C durante 1 minuto. Los tejidos se fijan en formalina (formaldehído al 1 % en PBS) a temperatura ambiente durante 10 minutos y se enjuagan varias veces en PBS. Después de la fijación, se inunda el portaobjetos en isopropanol (80 %), se retira y el portaobjetos con el tejido se seca al aire. El tejido se prepermeabiliza con collagenasa en tampón HBSS (0,2 U de collagenasa, 0,2 mg de BSA) a 37 °C durante 20 minutos y se lava con un tampón de lavado SSC 0,1X. Después del lavado, el tejido se permeabiliza a RT durante 10 minutos con una solución amortiguada que contiene NP40 al 0,1 %, Tween-20 al 0,1 % y digitonina al 0,01 %.

Los reactivos de fragmentomarcaje (Illumina, Inc.) se depositan en el tejido permeabilizado durante 1 hora a 37 °C. Los reactivos incluyen 2,5 µl de Tn5 en una mezcla de fragmentomarcaje (tampón TD 1x, PBS al 31 %, digitonina al 0,01 %, Tween-20 al 0,1 %, H<sub>2</sub>O). El portaobjetos se lava (tampón de lavado 3) y el fragmentomarcaje se detiene con un tampón de parada de fragmentomarcaje (SDS al 0,01 %, 50 mM de EDTA, 10 mM de Tris-HCl y H<sub>2</sub>O) a 37 °C durante 10 minutos.

Una mezcla de 2 µM de oligonucleótidos de férula (NEB2.1 1x, H<sub>2</sub>O) y reactivos asociados se añaden al tejido y se incuban a 75 °C durante 15 minutos. La temperatura disminuye lentamente a 20 °C durante toda la noche, con un enfriamiento de velocidad de rampa de 0,1 °C/min para permitir la difusión del marcado, la hibridación de los oligonucleótidos de férula con los marcados, y la captura subsecuente de los complejos marcado y oligonucleótidos de férula.

La polimerización y la ligazón se realizan mediante el uso de una mezcla de ADN T4 (5U de ADN polimerasa T4, 0,4U de ADN ligasa T4, 1 µM de dNTP, 3 mM de ATP, H<sub>2</sub>O, NEB2.1 1x) y se incuba a 20 °C durante 3 horas. El tejido se extrae mediante el uso de 2 mg/ml de proteinasa K en un tampón PKD, y la incubación se lleva a cabo a 56 °C durante 1 hora. Los portaobjetos se lavan durante 10 minutos a 50 °C en el tampón de lavado 1 (SSC 0,3x, SDS) posteriormente se lavan con el tampón de lavado 2 (SSC 0,2x) a temperatura ambiente durante 1 minuto, se finaliza en el tampón de lavado 3 (SSC 0,1x) y se seca por centrifugación.

## Preparación y secuencia de bibliotecas

Para permitir la preparación masiva de bibliotecas de los marcados unidos a la superficie con código de barras espacial, la desnaturalización de las sondas de captura en los portaobjetos se realiza con 0,08 N de KOH a temperatura ambiente durante 10 minutos. La mezcla se transfiere desde el portaobjetos a tubos LoBind que contienen 1M de Tris a pH 7,0. La etapa de desnaturalización se repite y se realiza la limpieza (Kit de limpieza por reacción MiniElute, Qiagen) seguida por la elución en 20 µl de tampón de elución. Las muestras se procesan en bibliotecas de secuenciación y se secuencian en un instrumento Illumina Nextseq 500 mediante el uso de lecturas extremos emparejadas según los protocolos del fabricante.

## Procedimientos de la tinción

Después del fragmentomarcage, se añade una capa de isopropanol al 80 % a los tejidos en los portaobjetos, se retira y se seca al aire como se describió anteriormente. Las secciones de tejido se tiñen con hematoxilina a temperatura ambiente durante 7 minutos, se enjuagan con agua, se incuban en tampón de azuleamiento a temperatura ambiente durante 2 minutos, se enjuagan con agua y se tiñen con eosina a temperatura ambiente durante 10 o 20 segundos en dependencia de la muestra de tejido (por ejemplo, cerebro de ratón y tejido de cáncer de próstata, respectivamente). Los portaobjetos se enjuagan con agua, se secan por centrifugación, se cubren con glicerol y un cubreobjetos y se obtienen imágenes.

## Inmunohistoquímica

Las secciones de tejido en los portaobjetos se bloquean mediante incubación con un tampón de tinción durante 5 minutos. El tampón de tinción se elimina y la dilución del anticuerpo primario se añade y se incuba a temperatura ambiente durante 30 minutos. Se realizan dos lavados con tampón de tinción durante 3 minutos cada uno, seguido de la adición de la dilución secundaria del anticuerpo y la incubación a temperatura ambiente durante 15 minutos. Se realizan tres lavados con tampón de tinción durante 3 minutos cada uno, y finalmente se lavan con una pipeta con PBS una vez. Los cortes se secan por centrifugación, se cubren con glicerol y un cubreobjetos y se obtienen imágenes.

La Figura 1 es un diagrama esquemático que muestra un ejemplo de una sonda de captura, como se describe en la presente descripción. Como se muestra, la sonda de captura 102 está opcionalmente acoplada a un elemento 101 mediante un enlazador escindible 103, tal como un enlazador fotoescindible. La sonda de captura puede incluir secuencias funcionales que son útiles para el procesamiento subsecuente, tales como la secuencia funcional 104, que puede incluir una secuencia de unión a la celda de flujo específica del secuenciador, por ejemplo, una secuencia P5 o P7, así como también la secuencia funcional 105, que puede incluir secuencias de cebadores de secuenciación, por ejemplo, un sitio de unión al cebador R1, un sitio de unión al cebador R2. En algunas modalidades, la secuencia 104 es una secuencia P7 y la secuencia 105 es un sitio de unión al cebador R2. Puede incluirse un código de barras espacial 106 dentro de la sonda de captura para su uso en la codificación del analito diana. Las secuencias funcionales generalmente pueden seleccionarse para compatibilidad con cualquiera de una variedad de sistemas de secuenciación diferentes, por ejemplo, Ion Torrent Proton o PGM, instrumentos de secuenciación Illumina, PacBio, Oxford Nanopore, etc., y los requisitos de los mismos. En algunas modalidades, las secuencias funcionales pueden seleccionarse por compatibilidad con sistemas de secuenciación no comercializados. Ejemplos de tales sistemas y técnicas de secuenciación, para los cuales pueden usarse secuencias funcionales adecuadas, incluyen (pero no se limitan a) secuenciación Ion Torrent Proton o PGM, secuenciación Illumina, secuenciación PacBio SMRT y secuenciación Oxford Nanopore. Además, en algunas modalidades, las secuencias funcionales pueden seleccionarse para que sean compatibles con otros sistemas de secuenciación, incluidos sistemas de secuenciación no comercializados.

En algunas modalidades, el código de barras espacial 106, las secuencias funcionales 104 (por ejemplo, la secuencia de unión a la celda de flujo) y 105 (por ejemplo, secuencias de cebadores de secuenciación) pueden ser comunes a todas las sondas unidas a un elemento dado. El código de barras espacial también puede incluir un dominio de captura 107 para facilitar la captura de un analito diana. En algunas modalidades, pueden incluirse secuencias adicionales en la sonda de captura. Por ejemplo, puede incluirse un identificador molecular único entre la secuencia funcional 105 y el código de barras espacial 106, o puede insertarse un identificador molecular único entre el código de barras espacial 106 y el dominio de captura 107. También pueden incluirse secuencias funcionales adicionales, por ejemplo, la secuenciación de cebadores (por ejemplo, secuencias de unión de cebadores de secuenciación, secuencias de unión de cebadores de amplificación) en la sonda de captura, por ejemplo, una segunda secuencia funcional que es diferente de la primera secuencia(s) funcional puede encontrarse entre el código de barras espacial 106 y el dominio de captura 107. En algunas modalidades, la sonda de captura puede incluir un identificador molecular único y una segunda secuencia funcional, de uno o ambos.

La Figura 2 muestra un flujo de trabajo ilustrativo de ATAC espacial (spATAC). El flujo de trabajo incluye poner en contacto una muestra biológica con una matriz que comprende una pluralidad de sondas de captura y marcar el ADN genómico accesible con un transposoma que resulta en un ADN fragmentado con secuencias de extremos del transposón insertadas (por ejemplo, Nextera A, Nextera B). El ADN marcado obtenido resulta en un espacio entre las secuencias de extremos del transposón y el ADN genómico fragmentado. En algunos ejemplos, las secuencias de extremos del transposón adyacentes al espacio se liberan (por ejemplo, se eliminan) del ADN

genómico fragmentado. En algunos ejemplos, tales secuencias de extremos del transposón se liberan con un gradiente de calor. Un oligonucleótido de férula se hibrida con el dominio de captura de la sonda de captura, o una porción de esta, y la secuencia de extremo del transposón restante. Se realiza el relleno de espacios, la extensión del ADN (por ejemplo, la extensión con una ADN polimerasa) y la ligazón de la secuencia de extremo del transposón a la sonda de captura. La polimerasa de relleno de espacios rellena el espacio entre el oligonucleótido de férula y el ADN genómico fragmentado (por ejemplo, una porción en la cual se incluye la secuencia de extremo del transposón liberada) con los ácidos nucleicos complementarios del espacio que necesita rellenarse. La extensión del ADN resulta en la generación de una segunda hebra con una secuencia complementaria al código de barras espacial y un complejo de ADN de doble hebra. En algunos ejemplos, la segunda hebra generada se libera (por ejemplo, se desnaturaliza), se recoge, se amplifica y se procesa para indexado y construcción de biblioteca. La biblioteca resultante se usa para determinar la ubicación de la cromatina accesible en la muestra biológica.

## Ejemplo 2. Análisis ATAC espacial

Los siguientes resultados se generaron mediante el flujo de trabajo de spATAC ilustrativo descrito en el Ejemplo 1. La Figura 3A muestra un cerebro de ratón teñido con hematoxilina y eosina (HyE) con una sección de tejido de xenoinjerto de glioma humano. La Figura 3B muestra el conglomerado espacial (por ejemplo, las posiciones puntuales) de la sección de tejido cerebral de ratón y los datos muestran una amplia coincidencia espacial entre el ADN genómico capturado y la estructura morfológica de la sección de tejido cerebral de ratón. Las casillas de ambas figuras indican áreas que sirven como áreas internas de control de calidad para demostrar que se obtuvo la resolución espacial de los métodos de spATAC (por ejemplo, en dos casillas no hay tejido y no se observa ningún conglomerado, en una casilla donde hay un poco de una pieza de tejido que se agrupa).

Las Figuras 4A-B muestran imágenes representativas de un experimento con tejido donde las células humanas que se expresan en ratón se usan para probar la resolución espacial del flujo de trabajo descrito en la Figura 2. La Figura 4A muestra un cerebro de ratón teñido con HyE con sección de tejido de xenoinjerto de glioma humano. Las porciones más oscuras rodeadas con un círculo en la imagen son áreas de la sección de tejido con células humanas. La Figura 4B muestra la identificación del ADN humano identificado dentro de la sección de tejido de ratón indicada por los círculos. Los datos muestran que el conglomerado refleja la morfología de la sección de tejido y es sensible a la detección y distinción entre el ADN humano frente al de ratón en la sección de tejido de muestra.

Tabla 1. Lecturas por punto en la escala de logaritmo natural

Índice	Promedio	Est.
1	14,197071531445	11,3092012410549
2	34,9050753875557	32,2738289613216
3	36,3514060287275	23,8985761010508
4	4,52903081732917	3,35271103443555
5	24,6123792800702	31,4306497970632
6	4,91824739513759	4,25066145637371

Tabla 2. Genes por punto en la escala de logaritmo natural

Índice	Promedio	Est.
1	12,8216514642343	9,68067308393996
2	22,8551709492461	14,954558667119
3	32,3910580619057	20,1521778991193
4	4,16033943724877	2,92892111421054
5	17,9218612818262	17,9319400455273
6	4,50601122094577	3,76867917965407

Las Figuras 5A-B muestran gráficos replicados del número de puntos por el número total de identificadores moleculares únicos (UMI) identificados en dos secciones de tejido cerebral de ratón (índice 6 e índice 5, respectivamente en la Tabla 1 y la Tabla 2; (datos del índice 1-4 no se muestran)). Los datos que se muestran se mapearon a cromosomas en lugar de genes. La línea discontinua indica la media (Figura 5A-5B, parte inferior). Las posiciones puntuales dentro de la muestra biológica se muestran para el índice 6 y 5 (Figura 5A-5B, parte superior). Los puntos más oscuros indican más UMI que se detectan en esa ubicación.

Las Figuras 6A-B muestran gráficos que indican la recuperación de la periodicidad del nucleosoma cuando la muestra biológica no se trató con proteinasa K después del fragmentomarcage y la tinción con HyE también se realizó después del fragmentomarcage. Los datos indican que la estructura de la cromatina permanece intacta sin el tratamiento con proteinasa K y la tinción después del fragmentomarcage.

En conjunto, los datos demuestran que el flujo de trabajo de spATAC ilustrativo descrito en la Figura 2 puede usarse para analizar espacialmente el ADN genómico accesible a partir de muestras biológicas. Además, el método

también es sensible, ya que el ADN humano y el de ratón se identificaron por separado mediante el flujo de trabajo en el cerebro de ratón con una sección de tejido de xenoinjerto de glioma humano.

### Ejemplo 3. Flujo de trabajo del análisis del genoma completo

El análisis del genoma completo (por ejemplo, genómica espacial) también puede realizarse en muestras de tejido. Por ejemplo, los métodos ATAC espaciales descritos en la presente descripción se diseñan para capturar áreas accesibles (por ejemplo, "abiertas" o transcripcionalmente activas) de un genoma, sin embargo, también es posible capturar espacialmente todo el genoma. Para permitir la captura de todo el genoma, la estructura de la cromatina se interrumpe por la degradación de las histonas después del tratamiento con colagenasa como se describió en el Ejemplo 1, seguido de la incubación con la proteinasa K (2 mg/ml) en un amortiguador de proteinasa K (SDS al 1 %, 50 mM de EDTA, 10 mM de Tris-HCl, H<sub>2</sub>O) a 37 °C durante 10 minutos. El ADN fragmentado resultante puede capturarse en la matriz espacial seguido de la preparación de la biblioteca y el análisis de las secuencias como se describe en la presente descripción.

### Ejemplo 4. Análisis ATAC espacial

Las Figuras 7A-H muestran un análisis replicado que se realizó en las dos secciones consecutivas de tejido embrionario de un ratón inmunotefidadas con un anticuerpo SOX9 (Figuras 7A y 7E) antes de los flujos de trabajo de ATAC-seq espacial descritos en la presente descripción. Las Figuras 7B y 7F muestran el número total de fragmentos de ADN marcados que se capturan por punto, lo que es indicativo de la resolución espacial de la captura de ADN genómico fragmentado. Las Figuras 7C y 7G son gráficos que muestran el enriquecimiento del sitio de inicio de la transcripción (TSS) para los puntos bajo las secciones de tejido para las secciones consecutivas y Figuras 7D y 7H son gráficos que muestran la periodicidad del nucleosoma correspondiente que se refleja en la distribución del tamaño del marcado que se captura reconstruido después de la secuenciación.

La Figura 8 muestra el diagrama genómico de las densidades de lectura de ATAC-seq en un conjunto de datos de referencia de ratón (e13,5 ENCODE) y de ATAC-seq espacial (e13,5) de los embriones de ratón que se muestran en las Figuras 7A y 7E. El enriquecimiento de la señal de ATAC-seq espacial y la designación de picos (parte inferior) muestran posiciones coincidentes para el enriquecimiento de fragmentos de los siguientes genes: Gga1, Mir6955, Sh3bp1, Pdxp, Lgals1, Noli2 y Triobp, por lo tanto, demuestra que los flujos de trabajo de ATAC espacial descritos en la presente descripción pueden usarse para detectar el ADN genómico accesible a partir de muestras biológicas (por ejemplo, secciones de embriones de ratón).

### Ejemplo 5. Análisis ATAC espacial

Los datos mostrados en las Figuras 9A-12B se generaron de acuerdo con la etapa de preparación y secuenciación de la biblioteca, los procedimientos de tinción y la inmunohistoquímica se realizaron como se describió en el Ejemplo 1. Sin embargo, la preparación de tejidos y el flujo de trabajo de ATAC espacial se prepararon de acuerdo con el protocolo más abajo.

#### Preparación de tejidos y flujo de trabajo de ATAC espacial

Las muestras de tejido se congelan instantáneamente en bloques de temperatura de corte óptima (OCT), se crio-seccionan en secciones de 10 µm, se colocan en un portaobjetos de matriz espacial de Visium y se calientan a 37 °C durante 1 minuto. Los tejidos se fijan en formalina (formaldehído al 1 % en PBS) a temperatura ambiente durante 10 minutos y se enjuagan varias veces en PBS. Después de la fijación, se inunda el portaobjetos en isopropanol (80 %), se retira y el portaobjetos con el tejido se seca al aire. El tejido se prepermeabiliza con colagenasa en tampón HBSS (0,2 U de colagenasa, 0,2 mg de BSA) a 37 °C durante 20 minutos y se lava con un tampón de lavado SSC 0,1X. Después del lavado, el tejido se permeabiliza a RT durante 10 minutos con una solución amortiguada que contiene NP40 al 0,1 %, Tween-20 al 0,1 % y digitonina al 0,01 %.

Los reactivos de fragmentomarcaje (Illumina) se depositan en el tejido permeabilizado durante 1 hora a 37 °C. Los reactivos incluyeron 2,5 µl de Tn5 en una mezcla de fragmentomarcaje (tampón TD 1x, PBS al 31 %, digitonina al 0,01 %, Tween-20 al 0,1 %, H<sub>2</sub>O). El portaobjetos se lava (tampón de lavado 3) y el fragmentomarcaje se detiene con un tampón de parada de fragmentomarcaje (SDS al 0,01 %, 50 mM de EDTA, 10 mM de Tris-HCl y H<sub>2</sub>O) a 37 °C durante 10 minutos.

Una mezcla de 2 µM de oligonucleótido de férula en una solución salina (a base de NEC 2,1 o SSC) que contiene 0,2 mg/µl de proteinasa K y una dilución de Triton X-100 (0,2-0,01 %) se añadieron a la sección de tejido y se incubaron a 30 °C durante aproximadamente 2-8 horas para permitir la liberación del marcador mediante la digestión de la proteasa, difusión del marcador y la hibridación de los oligonucleótidos de férula a los marcadores, y la captura subsecuente por las sondas de captura de los complejos de oligonucleótidos de férula-marcador.

La polimerización y la ligazón se realizan mediante el uso de una mezcla de ADN T4 (5 U de ADN polimerasa T4, 0,4 U de ADN ligasa T4, 1 µM de dNTP, 3 mM de ATP, H<sub>2</sub>O, NEB2.1 1x suplementado con Triton X-100 (0,2-0,01 %)), y se incubó a 20 °C durante 3 horas. El tejido se extrae mediante el uso de 2 mg/ml de proteinasa K en tampón PKD, y las incubaciones se llevaron a cabo a 56 °C durante 1 hora. Los portaobjetos se lavan durante 10

minutos a 50 °C en el tampón de lavado 1 (SSC 0,3x, SDS) posteriormente se lavan con el tampón de lavado 2 (SSC 0,2x) a temperatura ambiente durante 1 minuto, se finaliza en el tampón de lavado 3 (SSC 0,1x) y se seca por centrifugación.

La Figura 9A es una aproximación y proyección uniforme del colector (UMAP) del conglomerado no sesgada basada en gráficos y la asignación del conglomerado de cada punto en la sección de tejido que se muestra en la Figura 9B. La Figura 9B muestra unos gráficos (UMAP) de accesibilidad génica a través de los puntos tisulares en una sección de ratón procesada de acuerdo con ATAC-seq espacial. En conjunto, los datos muestran que ATAC-seq espacial captura una variación biológica significativa en la accesibilidad génica a través de las regiones de tejido.

Las Figuras 10A-D muestran gráficos de UMAP (Figura 10A y Figura 10C) que se colorean por la accesibilidad relativa de dos regiones génicas, la glicoforina C (Gypc) y un receptor acoplado a la proteína G de adhesión 1 (Bai1), que se encuentra que es diferencialmente accesible entre las regiones de la sección de tejido que corresponden a la accesibilidad relativa que se muestra en la Figura 10B y Figura 10D.

La Figura 11A muestra el conglomerado basado en la expresión génica. Los conglomerados, es decir, las regiones de tejido con perfiles de expresión génica característicos se indican mediante números. Las Figuras 11B-F muestran la accesibilidad de algunos de los genes marcadores principales (Figura 11B 28 genes; Figura 11C 61 genes; Figura 11D 132 genes; Figura 11E 180 genes; Figura 11F 807 genes) para cada conglomerado en una sección adyacente, de acuerdo con ATAC-seq espacial, que manifiesta una alta concordancia.

Se ensayó una muestra de tejido de ratón para determinar la accesibilidad a la región genómica, donde se encontró que las regiones eran diferencialmente accesibles cuando se usan los métodos de spATAC-seq descritos. La Figura 12 muestra el diagrama genómico del enriquecimiento de la señal de ATAC-seq espacial y la accesibilidad de una de las regiones que se encuentran que son más accesibles en una sección de tejido de ratón.

En conjunto, los datos de las Figuras 9A-12 demuestran que el flujo de trabajo ilustrativo, que incluye la preparación de tejidos y el flujo de trabajo de ATAC espacial, puede usarse para detectar el ADN genómico accesible a partir de muestras biológicas.

#### Apéndice del listado de secuencias

##### SEQ ID NO: 1 – Transposasa Tn5

MITSALHRAADWAKSVFSSAALGDPRRTARLVNVAQAQKYSKGSITISSEGSEAMQEGAYRFIRNPNVSAEAIKKA  
GAMQTVKLAQEFPELLAIEDTTSLSYRHQVAEELGKLGSIQDKSRGWWVHVSLLLEATTFRVTGLLHQEWWMRP  
DDPADADEKESGKWLAAAATSRRLMGSMMSNVIAVCDREADIHAYLQDKLAHNERFVVRSKHPRKDVESGLYLY  
DHLKNQPELGGYQISIPQKGVVDKRGKRKNRPARKASLSLRSGRITLQGNITLNAVLAEINPPKGETPLKWLILLT  
SEPVESLAQALRVIDIYTHRWRIEFHKAWKTGAGAERQRMEEPDNLERMVSILSFVAVRLLQLRESFTLPQALRA  
QGLLKEAEHVESQSAETVLTPEDECQLLGYLDKGKRKRKEKAGSLQWAYMAIARLGGFMDSKRTGIASWGALWEG  
WEALQSKLDGFLAAKDLMAQGIKI

##### SEQ ID NO: 2 secuencia de extremo del mosaico Tn5

CTGTCTCTTA TACACATCT

##### SEQ ID NO: 3 – Transposasa del bacteriófago Mu

MKEWYTAKELLGLAGLPKQATNITRKAQREGWEFRQVAGTKGVSFENIKSFPVALRAEILLQQGRIETSQGYFEIA  
RPTLEAHDYDREALWSKWDNASDSQRRRLAEKWLPVQAADMLNQGISTKTAFATVAGHYQVSASTLRDKYYQV  
QKFAKPDWAAALVDGRGASRRNVHKSEFDEDAWQFLIADYLRPEKPAFRKCYERLELAAREHGWSIPSRAAFRR  
IQQLDEAMVVACREGEHALMHLIPAQQRTVEHLDAMQWINGDGYLHNVFVRWFNGDVIRPKTWFWQDVKTRKIL  
GWRCDSVENSIRLSFMDVVTRYGIPEDFHITIDNTRGAANKWLTGGAPNRYRFKVKEDDPKGLFLLMGAKMHW  
TSVVAGKGWGQAKPVERAFGVGGLEEVYDKHPALAGAYTGPNPQAKPDNYGDRAVDAELFLKTLAEGVAMFNA  
RTGRETEMCGGKLSFDDVFEREYARTIVRKPTTEEQKRMILLPAEAVNVSRKGEFALKVGGSLKGAKNVYYNMAL  
MNAGVKKVVVRFDPPQLHSTVYCYTLDGRFICEAECLAPVAFNDAAAGREYRRRQKQLKSATKAAIKAQKQMDA  
LEVAELLPIAEPEAPESRIVGIFRPSGNTERVKNQERDDEYETERDEYLNHSLDILEQNRRKKAI

##### SEQ ID NO: 4 - Secuencia de reconocimiento de la transposasa Mu

TGAAGCGGCG CACGAAAAAC GCGAAAG

##### SEQ ID NO: 5 - Secuencia de reconocimiento de la transposasa Mu

GCGTTTCACG ATAAATGCGAAA

##### SEQ ID NO: 6 - Secuencia de reconocimiento de la transposasa Mu

CTGTTTCATT TGAAGCGCGA AAG



SEQ ID NO: 7 - Secuencia de reconocimiento de la transposasa Mu

TGTATTGATT CACTTGAAGT ACGAAAA

SEQ ID NO: 8 - Secuencia de reconocimiento de la transposasa Mu

CCTTAATCAA TGAAACGCGA AAG

5 SEQ ID NO: 9 - Secuencia de reconocimiento de la transposasa Mu

TTGTTTCATT GAAAATACGA AAA

SEQ ID NO: 10 - Familia de la Transposasa IS4

10 MTHSDAKLWAQEQFGQAQLKDPRTQRLISLATSIAQPGVSVAKLPFSPADMEGAYRFIRNENINAEDIAEAGFQ  
STVSRANEHKELLALEDTTTTLSFPHRSIKEELGHTNQGDRTALHVVHSTLLFAPQSQTIVGLIEQQRWSEDITKRGQ  
KHQHATRPYKEKESYKWEQASRRVVERLGDKMLDVISVCDREADLFEYLYTKRQHQQRFVVRSMQSRCLEEHA  
QKLYDYAALPSVETKALTIPQKGGRKARNVKLDVKYQVTLKAPANKKEHAGIPVYYYVGCLEQGTSKDLAWHLLT  
SEPINNVDAMRIIGYYERRWLIEDFHKVWKSEGTDESRLQSKDNLERLSVIYAFVATRLALRFMKEVDELTK  
SCEKVLGQKAWKLLWLKLESKTLPEVPDMGWAYKNLAKLGGWKDTKRTGRASIKVLWEGWFKLQTILEGYELA  
MSLDH

15 SEQ ID NO: 11- Secuencia de reconocimiento de la transposasa de *V. harveyi*

CTGTCTCTTGATCACAAGT

SEQ ID NO: 12- Secuencia de reconocimiento de la transposasa de *V. harveyi*

AGATGTGATCAAGAGACAG

SEQ ID NO: 13- Secuencia de reconocimiento de la transposasa de *V. harveyi*

20 CTGTCTCTTGATCACATCT

## REIVINDICACIONES

1. Un método para determinar la accesibilidad al ADN genómico, el método comprende:
  - (a) una muestra biológica en una matriz que comprende una pluralidad de sondas de captura, en donde una sonda de captura de la pluralidad de sondas de captura comprende: (i) un código de barras espacial y (ii) un dominio de captura;
  - (b) poner en contacto una pluralidad de oligonucleótidos de férula con la muestra biológica, en donde un oligonucleótido de férula se hibrida con el dominio de captura;
  - (c) poner en contacto un transposoma con la muestra biológica para insertar secuencias de extremos del transposón en el ADN genómico accesible, de esta manera se genera ADN genómico fragmentado;
  - (d) hibridar el ADN genómico fragmentado al oligonucleótido de férula y ligar el ADN genómico fragmentado a la sonda de captura;
  - (e) liberar una o más secuencias de extremos del transposón no ligadas del ADN genómico fragmentado ligado; y
  - (f) determinar (i) una secuencia del código de barras espacial o un complemento de la misma, y (ii) la totalidad o una porción de una secuencia del ADN genómico fragmentado, o un complemento de la misma, y usar las secuencias determinadas de (i) y (ii) para determinar la accesibilidad al ADN genómico en la muestra biológica.
2. El método de la reivindicación 1, en donde la matriz comprende uno o más elementos, opcionalmente en donde el uno o más elementos comprenden una perla.
3. El método de la reivindicación 1 o 2, en donde la sonda de captura comprende además un dominio de escisión, uno o más dominios funcionales, un identificador molecular único, o sus combinaciones.
4. El método de cualquiera de las reivindicaciones 1-3, que comprende además una etapa de migración activa en donde el ADN genómico fragmentado migra a la matriz por la aplicación de un campo eléctrico.
5. El método de cualquiera de las reivindicaciones 1-4, en donde la hibridación en la etapa (b) comprende hibridar el oligonucleótido de férula, o una porción del mismo, al dominio de captura, o una porción del mismo, de la sonda de captura, opcionalmente en donde la hibridación en la etapa (d) comprende hibridar el oligonucleótido de férula, o una porción del mismo, a una secuencia de extremo del transposón o una porción de la misma, de un ADN genómico fragmentado.
6. El método de cualquiera de las reivindicaciones 1-5, en donde la ligazón se realiza mediante el uso de una ADN ligasa, opcionalmente en donde la etapa de ligazón resulta en la generación de una molécula de ADN.
7. El método de cualquiera de las reivindicaciones 1-6, que comprende además extender un extremo 3' de la sonda de captura mediante el uso del ADN genómico fragmentado como una plantilla, opcionalmente en donde la etapa de extensión se realiza mediante el uso de una ADN polimerasa que tiene actividad de desplazamiento de hebra.
8. El método de cualquiera de las reivindicaciones 1-7, que comprende además realizar el relleno de espacios entre el oligonucleótido de férula y el ADN genómico fragmentado.
9. El método de cualquiera de las reivindicaciones 1-8, en donde el transposoma comprende una enzima transposasa, y en donde la enzima transposasa es una enzima transposasa Tn5, opcionalmente en donde la enzima transposasa Tn5 comprende una secuencia que es al menos 80 % idéntica a la SEQ ID NO: 1, una enzima transposasa Mu, una enzima transposasa Tn7, una transposasa de las especies de *Vibrio*, o derivados funcionales de esta.
10. El método de cualquiera de las reivindicaciones 1-9, en donde la determinación en la etapa (f) comprende secuenciar (i) el código de barras espacial o un complemento del mismo, y (ii) la totalidad o una porción de la secuencia del ADN genómico fragmentado o un complemento de la misma, y determinar además la ubicación del ADN genómico accesible en la muestra biológica.
11. El método de cualquiera de las reivindicaciones 1-10, que comprende además obtener imágenes de la muestra biológica antes o después de poner en contacto la muestra biológica con la matriz, opcionalmente que comprende además teñir la muestra biológica, opcionalmente en donde la tinción comprende teñir con hematoxilina y eosina.
12. El método de cualquiera de las reivindicaciones 1-11, en donde la liberación en la etapa (d) comprende calentar la muestra biológica, opcionalmente en donde el calentamiento comprende calentar hasta una temperatura de aproximadamente 65 °C a 85 °C.

13. El método de cualquiera de las reivindicaciones 1-12, en donde poner en contacto el transposoma con la muestra biológica se realiza bajo una condición de permeabilización química, bajo una condición de permeabilización enzimática, o ambas.
- 5 14. El método de la reivindicación 13, en donde la condición de permeabilización química comprende un detergente, opcionalmente en donde el detergente es uno o más de NP-40, Tween-20, Triton X-100 y Digitonina, opcionalmente en donde el detergente está a una concentración de aproximadamente 0,001 % (v/v) a aproximadamente 1,0 % (v/v).
- 10 15. El método de cualquiera de las reivindicaciones 1-14, en donde poner en contacto el transposoma con la muestra biológica se realiza después de una condición de prepermeabilización enzimática, opcionalmente en donde la condición de prepermeabilización enzimática comprende una proteasa, opcionalmente en donde la proteasa es una pepsina, una collagenasa, una Proteinasa K y sus combinaciones.

DIBUJOS

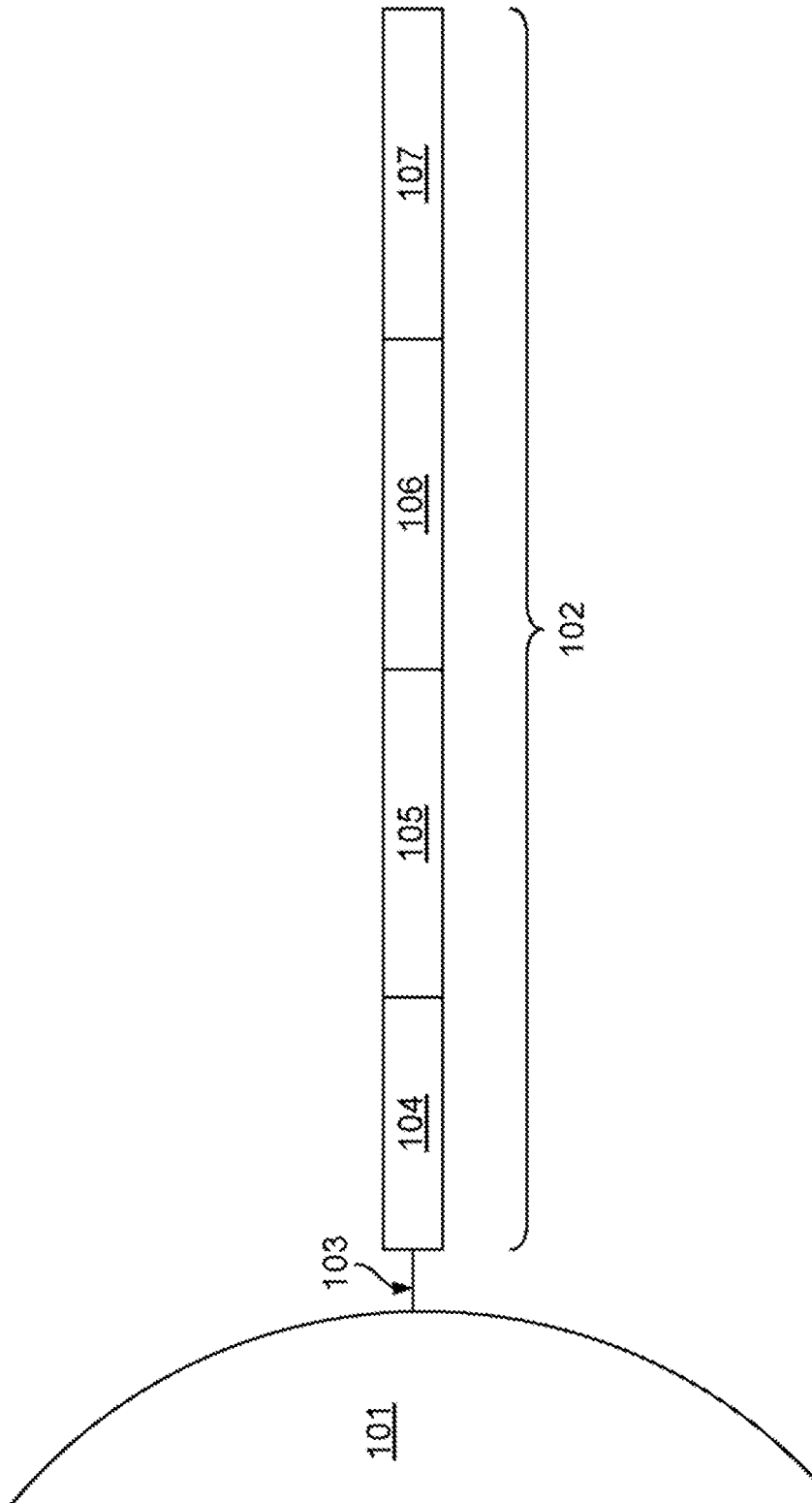
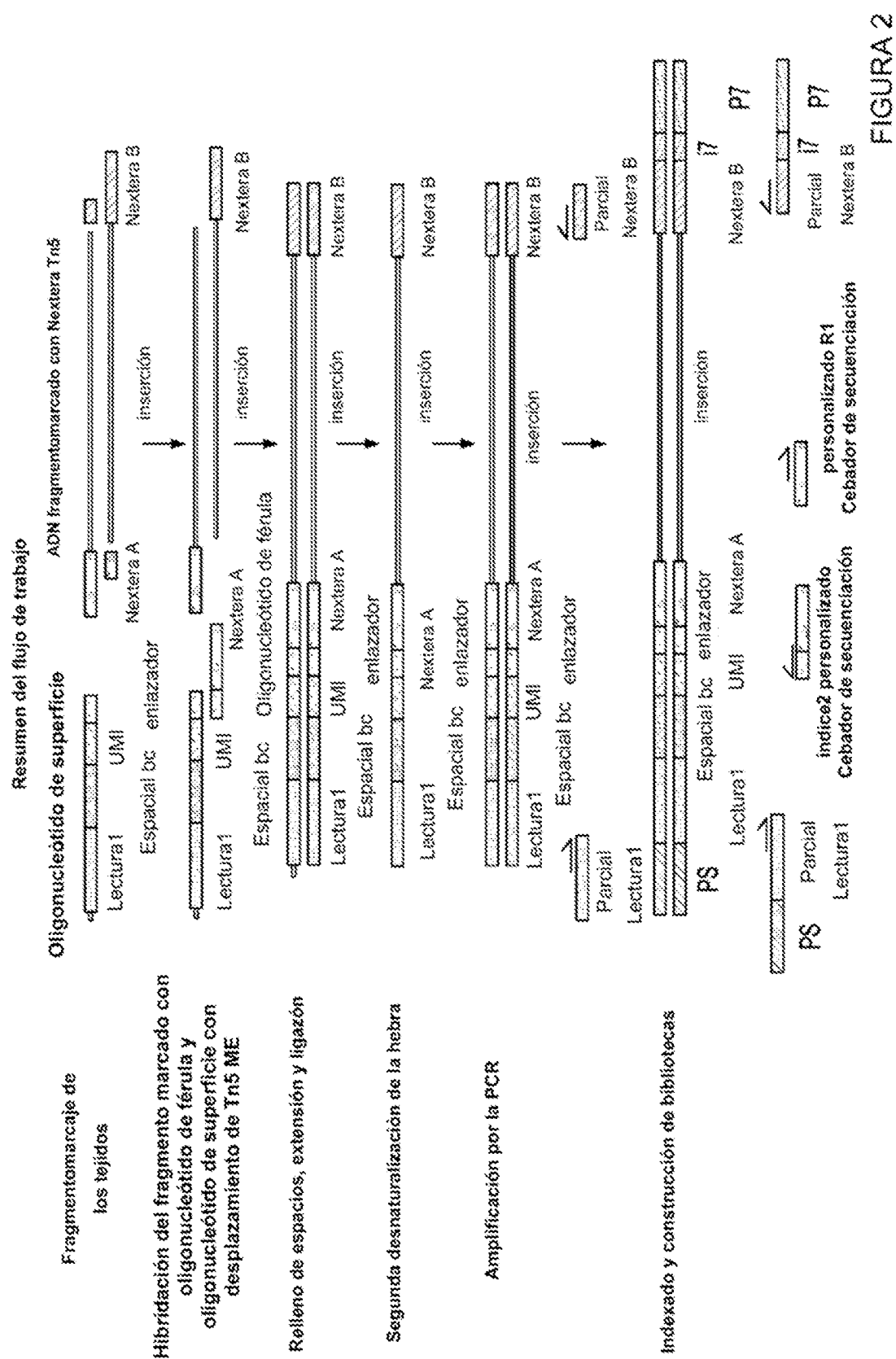


FIGURA 1



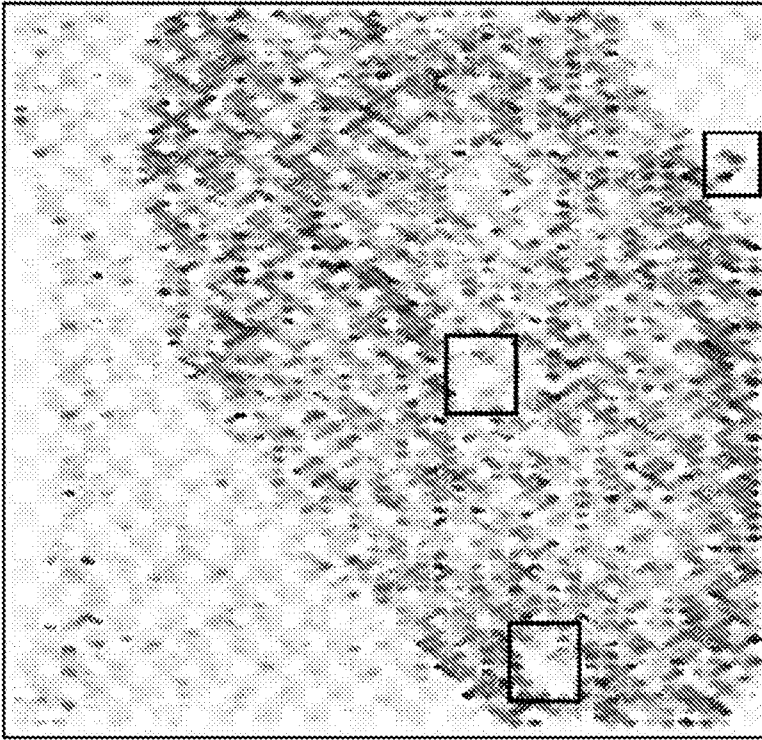


FIGURA 3B

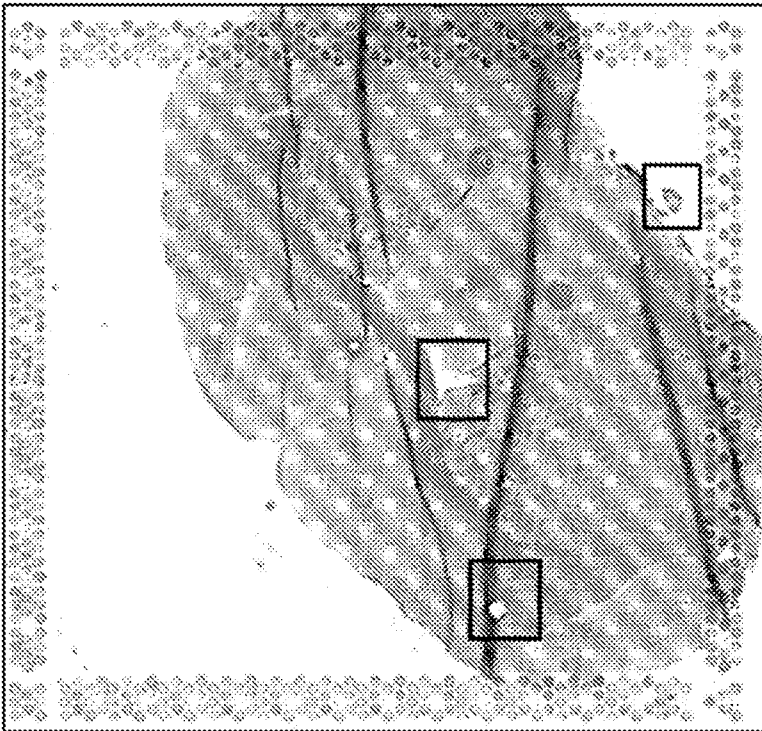


FIGURA 3A



FIGURA 4B

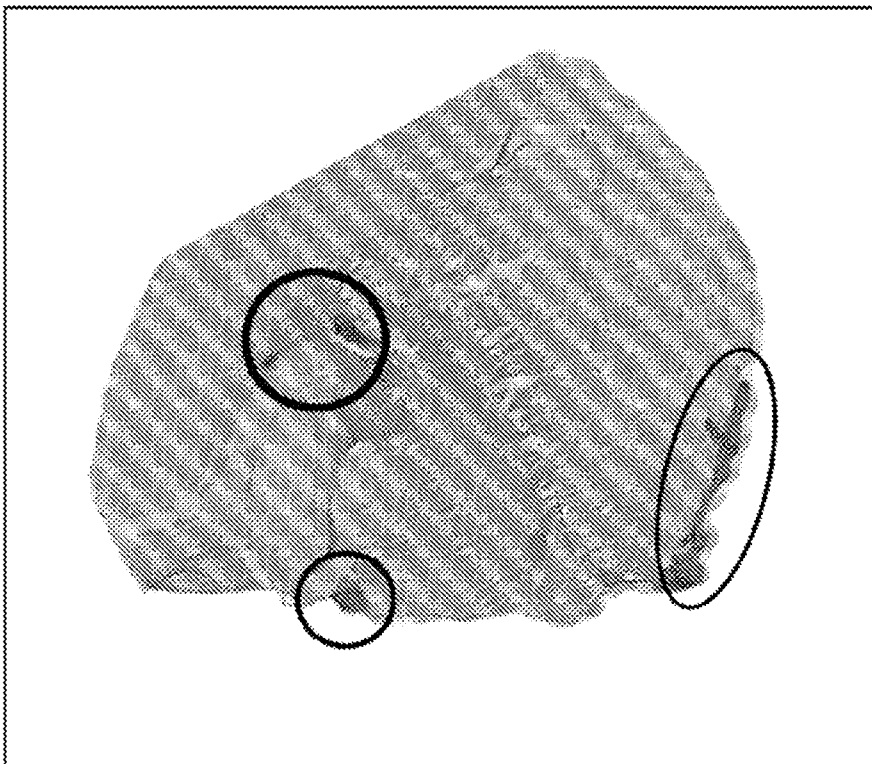
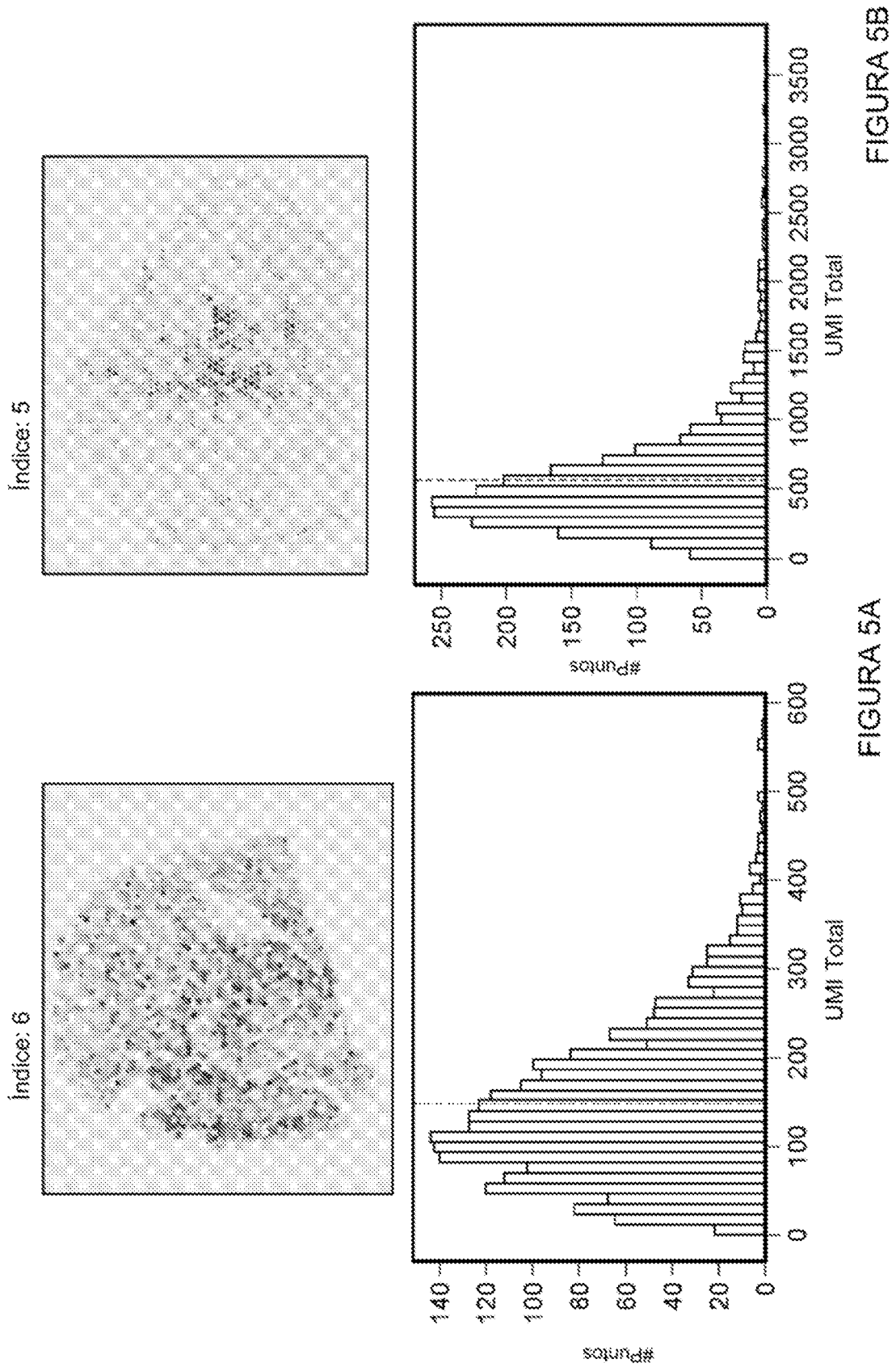


FIGURA 4A





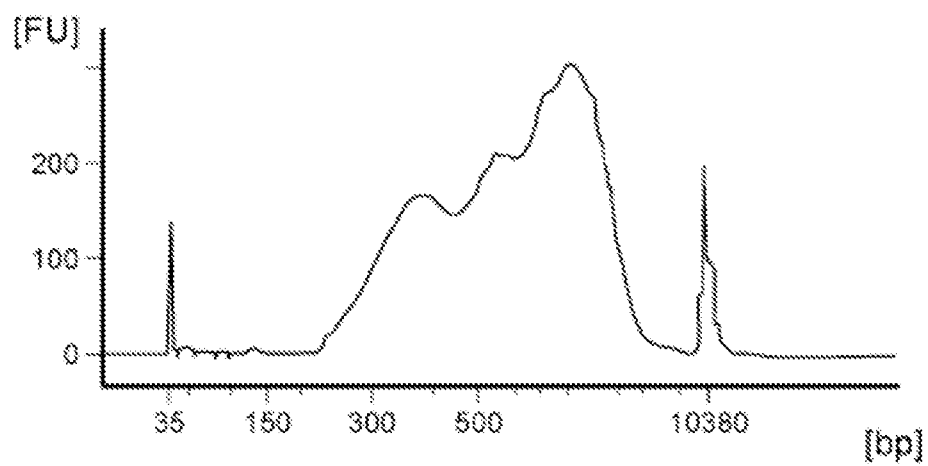


FIGURA 6A

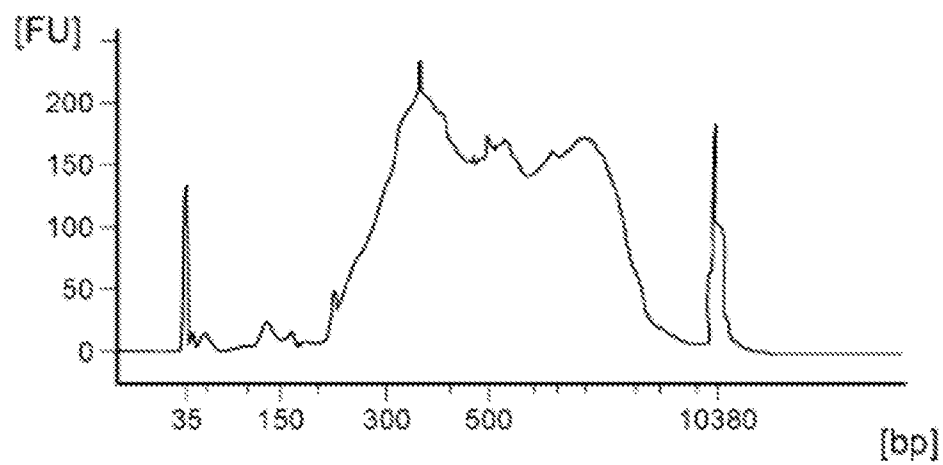


FIGURA 6B

FIGURA 7B

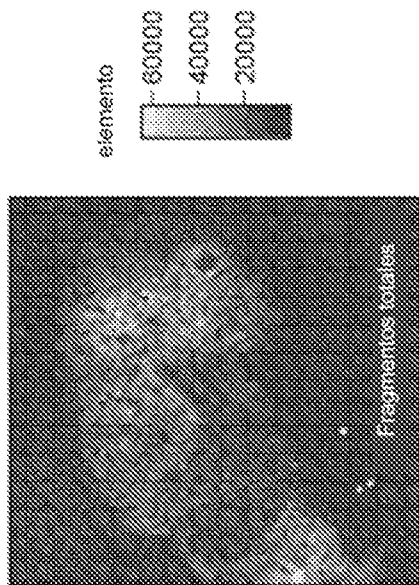


FIGURA 7A



FIGURA 7F

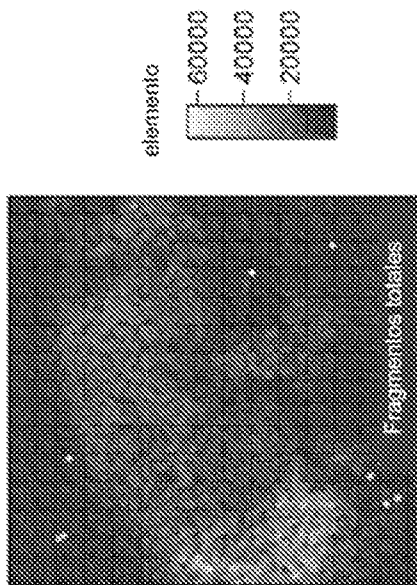
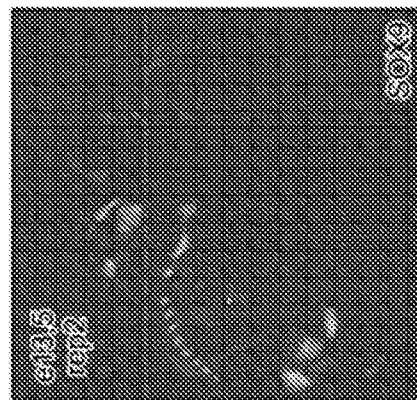
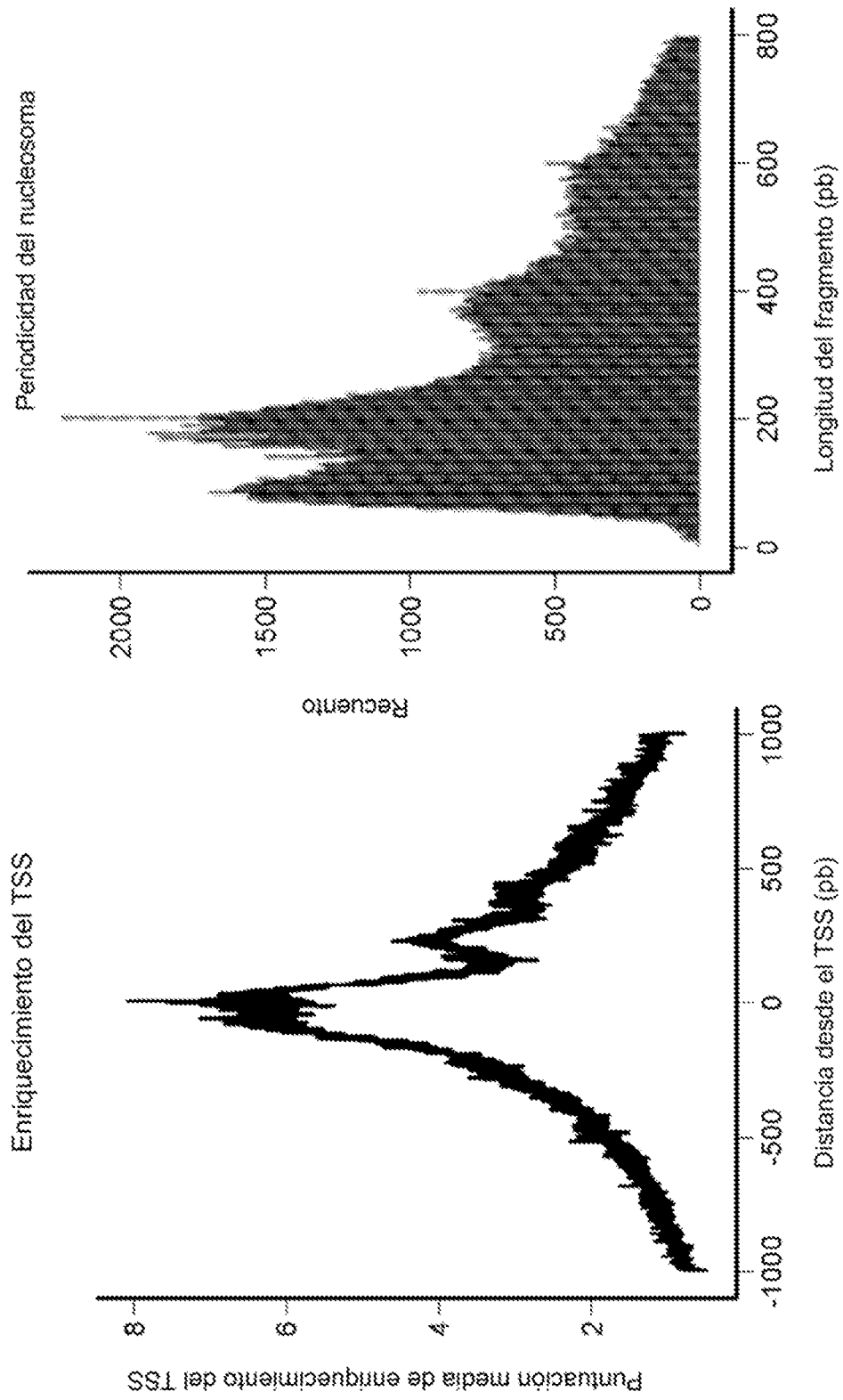


FIGURA 7E





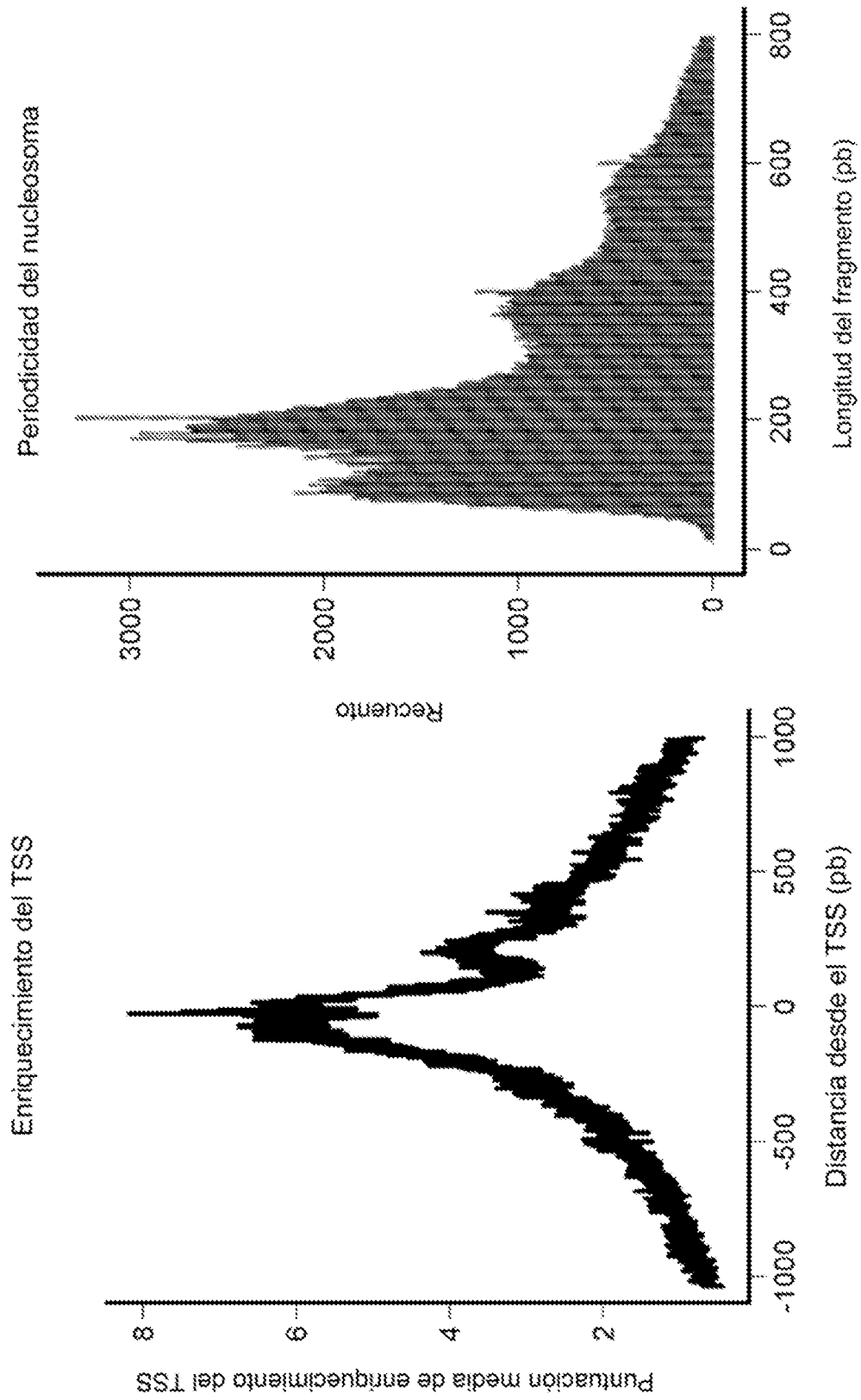


FIGURA 7H

FIGURA 7G

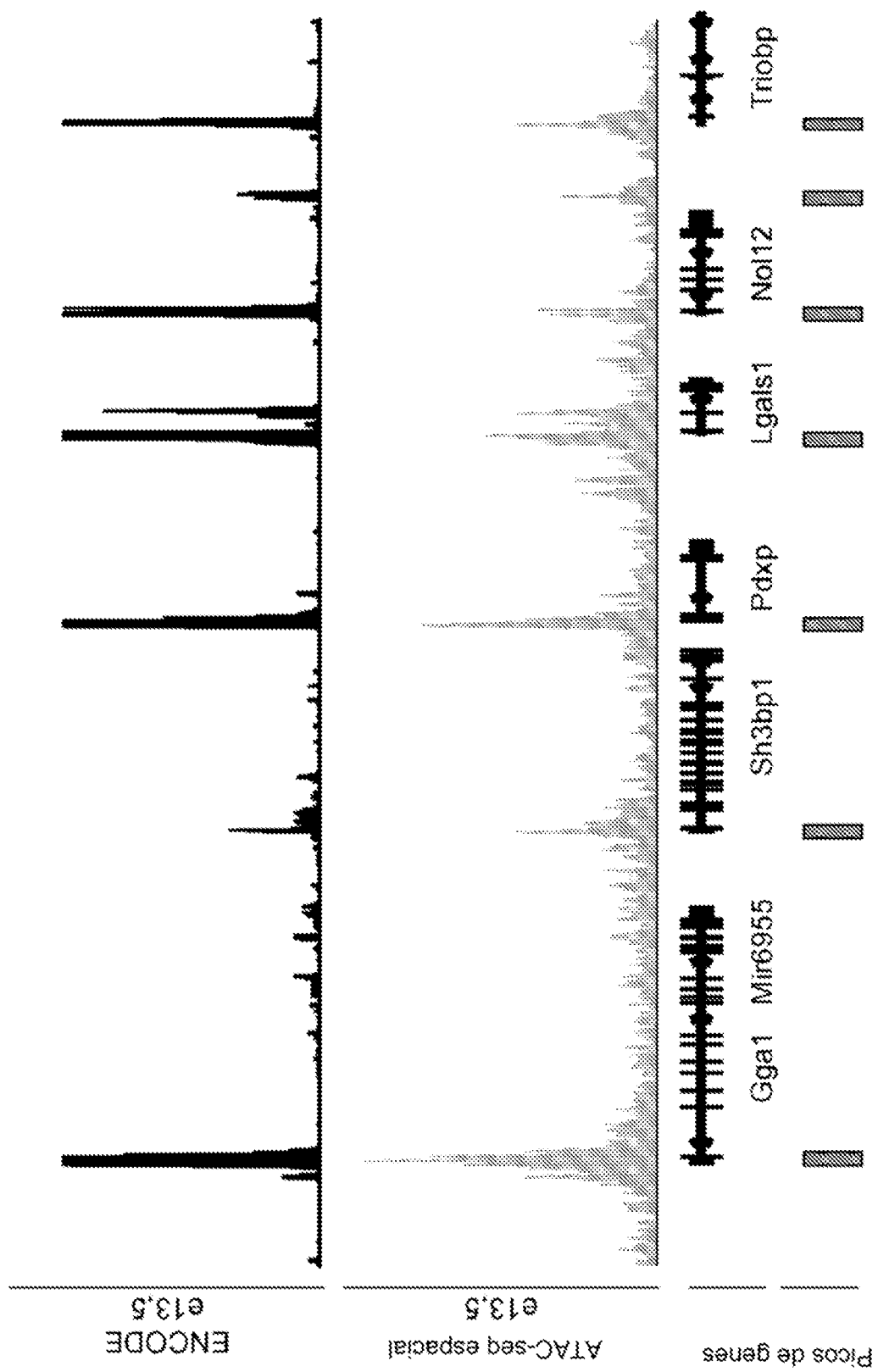


FIGURA 8

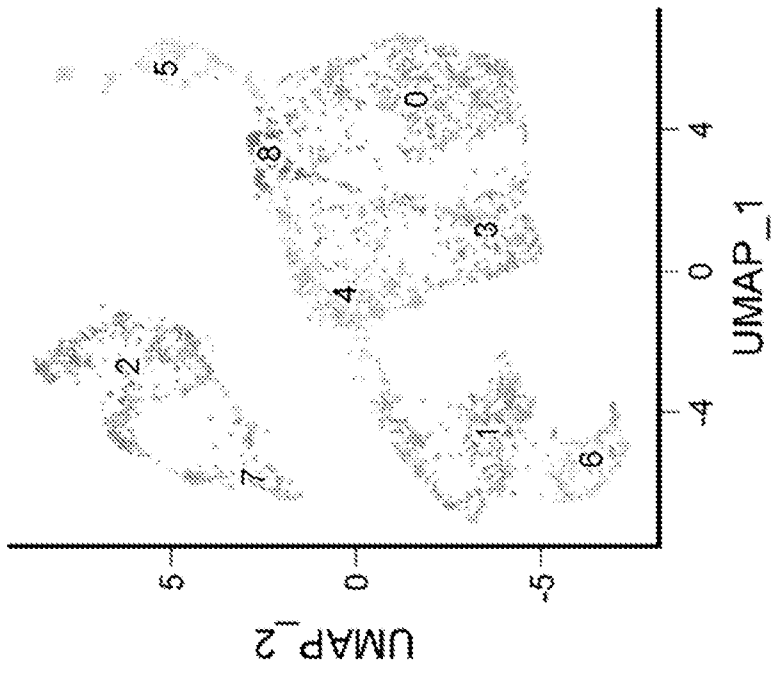


FIGURA 9A

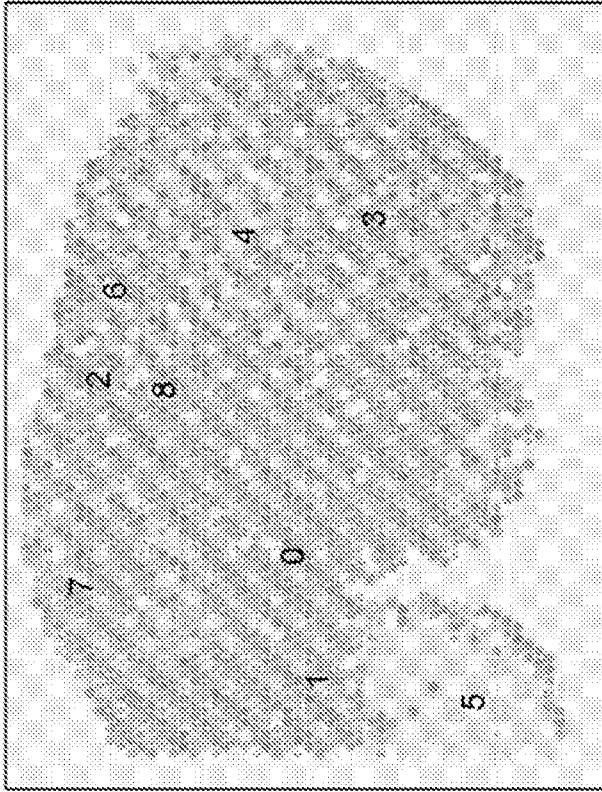


FIGURA 9B

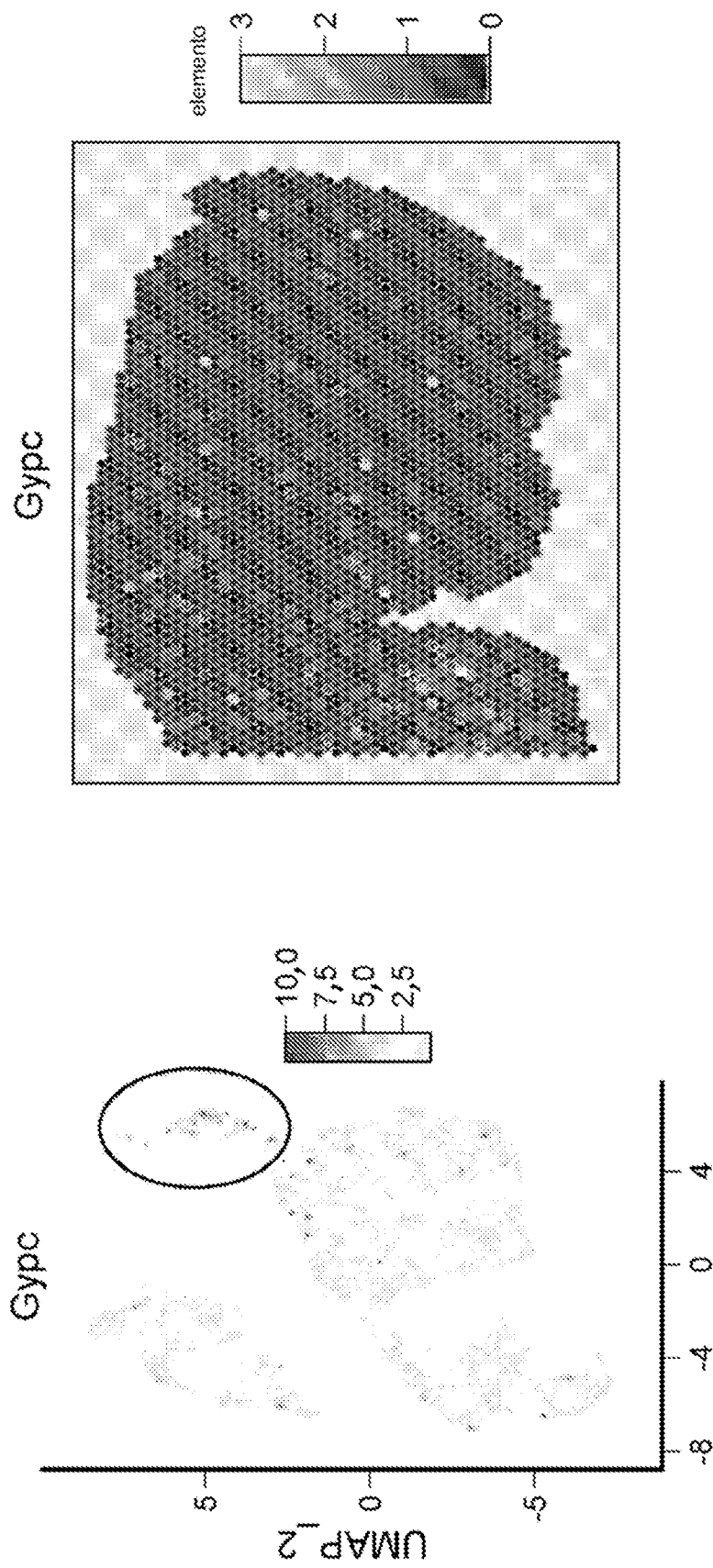


FIGURA 10B

UMAP\_1  
FIGURA 10A

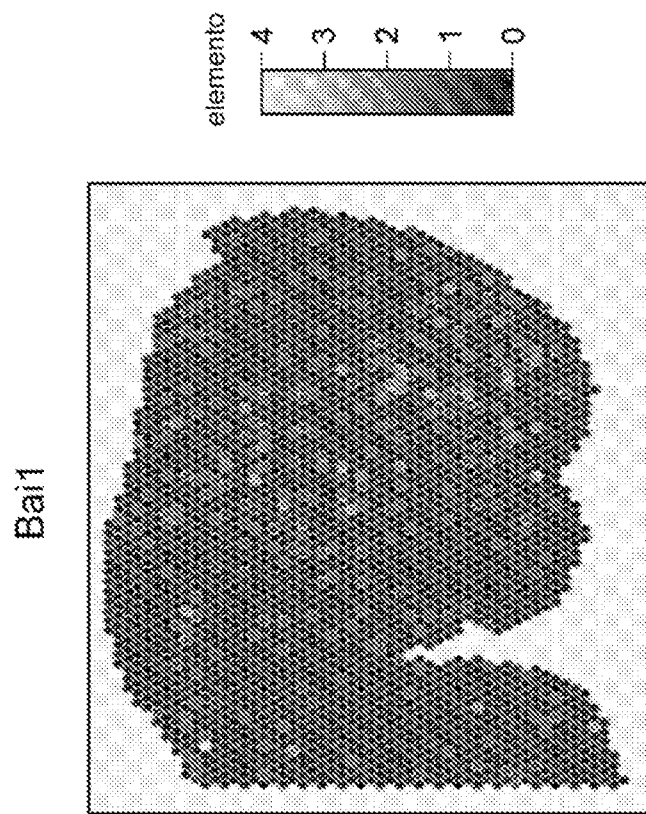


FIGURA 10D

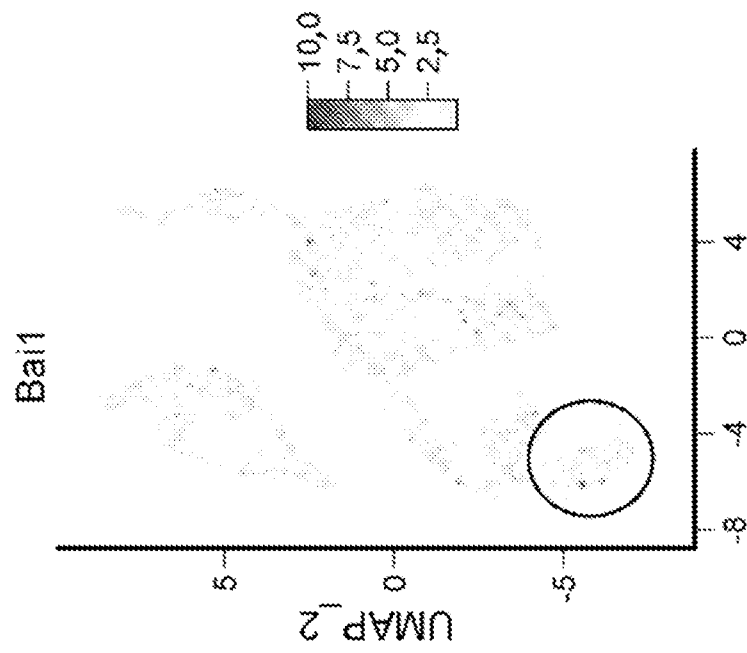


FIGURA 10C



FIGURA 11A

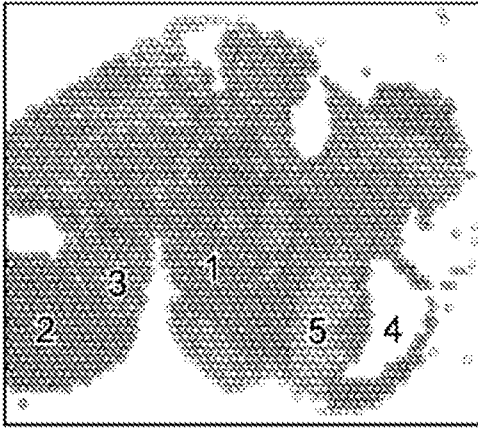


FIGURA 11B

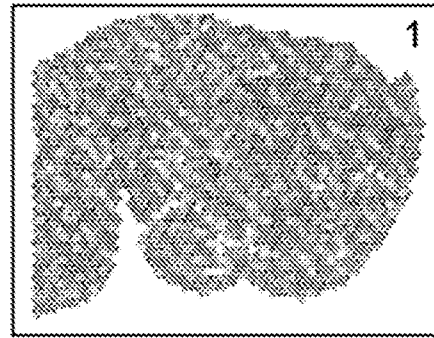


FIGURA 11C

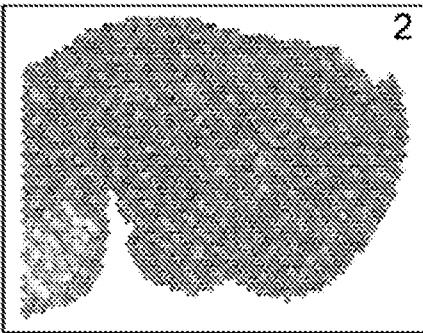


FIGURA 11D

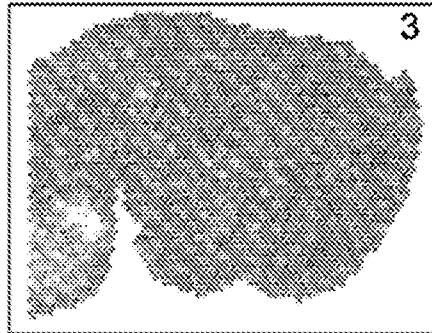


FIGURA 11E

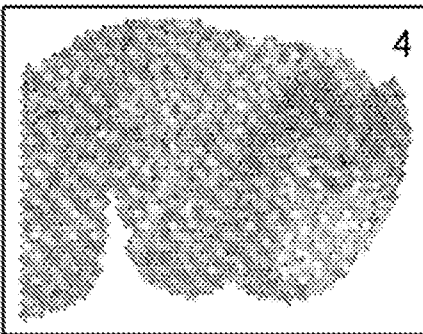
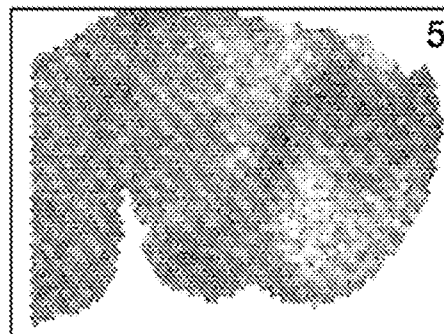


FIGURA 11F



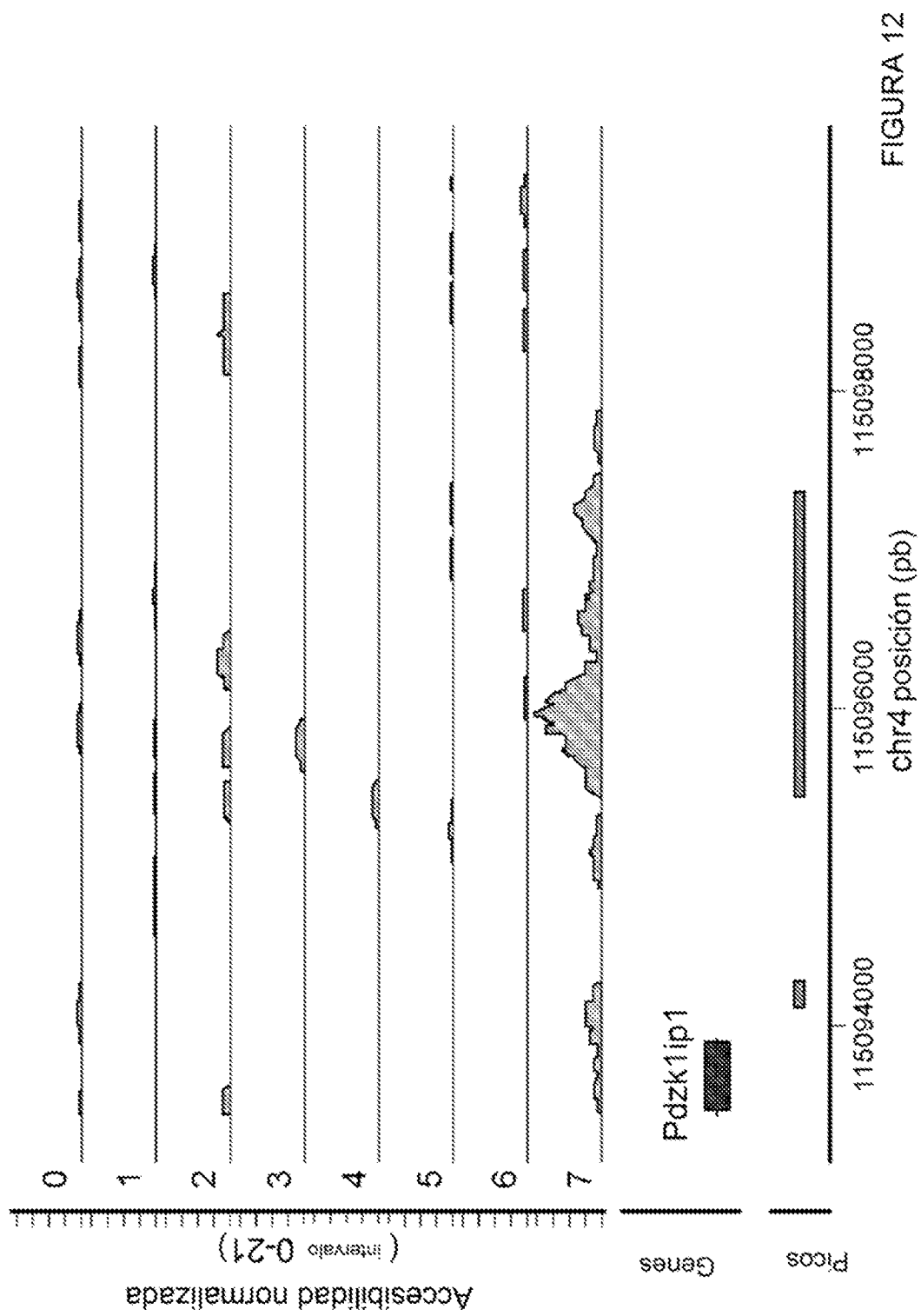


FIGURA 12