

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 993 641**

51 Int. Cl.:

C12Q 1/6841 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.02.2021** **PCT/US2021/018826**

87 Fecha y número de publicación internacional: **26.08.2021** **WO21168287**

96 Fecha de presentación y número de la solicitud europea: **19.02.2021** **E 21711699 (5)**

97 Fecha y número de publicación de la concesión europea: **09.10.2024** **EP 4107285**

54 Título: **Métodos y composiciones para ensayo espacial integrado in situ**

30 Prioridad:

21.02.2020 US 202062980078 P
09.11.2020 US 202063111518 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
03.01.2025

73 Titular/es:

10X GENOMICS, INC. (100.00%)
6230 Stoneridge Mall Road
Pleasanton, CA 94588-3260, US

72 Inventor/es:

BAVA, FELICE ALESSIO y
BENT, ZACHARY

74 Agente/Representante:

PONS ARIÑO, Ángel

Observaciones:

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

ES 2 993 641 T3

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

DESCRIPCIÓN

Métodos y composiciones para ensayo espacial integrado *in situ*

5 Campo

La presente descripción se refiere en algunos aspectos a métodos para un ensayo espacial integrado *in situ* de dianas biológicas en una muestra.

10 Antecedentes

La caracterización de dianas biológicas en una muestra, tales como la caracterización genómica, transcriptómica o proteómica de células, son esenciales para muchos propósitos, tales como comprender las bases moleculares de la identidad celular y desarrollar un tratamiento para enfermedades. Las imágenes de microscopía, que pueden resolver múltiples analitos en una muestra, proporcionan información valiosa tal como la abundancia del analito y la información espacial de los analitos *in situ*. Los enfoques actuales basados en hibridación *in situ* y secuenciación tienen una baja eficiencia, pero el valor potencial de dicho análisis en el tejido podría ser enorme. Por lo tanto, existe una necesidad de métodos nuevos y mejorados para los análisis *in situ*. El documento US 2018/245142 describe métodos de mapeo espacial de información de secuencias de ácido nucleico.

20

Breve resumen de la invención

La invención se define en la reivindicación 1. Las modalidades ilustrativas de la invención se proporcionan en las reivindicaciones dependientes.

25

En la presente descripción se describe un método para analizar un ácido nucleico diana en una muestra biológica, que comprende poner en contacto una muestra biológica (por ejemplo, una muestra de tejido tal como una sección de tejido) con una o más sondas (por ejemplo, sondas de ácido nucleico) que se hibridan directa o indirectamente con un primer ácido nucleico diana en la muestra biológica. La muestra biológica se proporciona sobre un primer sustrato. La muestra biológica puede comprender células, por ejemplo, células aisladas a partir de una biopsia o un cultivo celular, que se depositan sobre una superficie. La muestra biológica puede reticularse opcionalmente de manera reversible, por ejemplo, en una matriz tal como una matriz polimérica o un hidrogel. En algunas modalidades, la muestra biológica es una muestra de tejido procesada o clarificada. En algunas modalidades, el método comprende analizar el primer ácido nucleico diana *in situ*, por ejemplo, en una ubicación espacial en la muestra biológica, mediante la detección del primer ácido nucleico diana y/o la una o más sondas o producto(s) de estos, mientras que el primer ácido nucleico diana y/o la una o más sondas o producto(s) de estos no se unen a un agente de captura que se inmoviliza directa o indirectamente sobre un segundo sustrato. En algunas modalidades, el segundo sustrato es el primer sustrato en el que se proporciona la muestra biológica. En algunas modalidades, el segundo sustrato se separa del primer sustrato, aunque los materiales del sustrato y/o moléculas sobre el mismo pueden ser los mismos o diferentes entre el primer sustrato y el segundo sustrato.

40

El segundo sustrato (y opcionalmente el primer sustrato) comprende una pluralidad de agentes de captura (por ejemplo, sondas de captura) que se inmovilizan directa o indirectamente sobre el mismo, y uno o más de la pluralidad de agentes de captura (por ejemplo, una o más sondas de captura tales como cada sonda de captura) comprenden (i) un dominio de captura capaz de capturar el producto de ligación y (ii) un código de barras espacial.

45

En algunas modalidades, el método comprende el análisis *in situ* de uno o más analitos y/o producto(s) de estos, y/o el análisis *in situ* de una o más sondas (y/o producto(s) de estas) para el uno o más analitos en la muestra. El método comprende detectar (mediante la obtención de imágenes de la muestra biológica) la una o más primeras sondas de ácido nucleico en una ubicación espacial de la muestra biológica. En algunas modalidades, al menos una de la una o más sondas comprende una secuencia de código de barras que corresponde a una secuencia del primer ácido nucleico diana. En algunas modalidades, la muestra biológica se captura en imágenes sobre el sustrato para determinar una secuencia de la secuencia de código de barras o una secuencia complementaria de esta, que puede estar en la al menos una sonda o un producto de esta, por ejemplo, mediante análisis *in situ* tal como secuenciación *in situ*. En cualquiera de las modalidades anteriores, previo a, o durante el análisis *in situ*, un analito (y/o un producto de este) y/o una sonda (y/o un producto de esta) que se unen al analito no se unen a un agente de captura que se inmoviliza sobre una matriz artificial (por ejemplo, una matriz que comprende sondas de ácido nucleico que comprenden códigos de barras espaciales), aunque el analito y/o producto de este y/o la sonda y/o producto de esta pueden inmovilizarse en la muestra biológica, por ejemplo, mediante hibridación de ácido nucleico o reticulación reversible o irreversible con otros componentes en la muestra biológica o con una matriz tal como una matriz polimérica o un hidrogel.

60

El método comprende además proporcionar condiciones para permitir que los agentes de captura sobre el segundo sustrato capturen un producto de ligación, como se define en las reivindicaciones. En algunas modalidades, el método comprende eliminar las reticulaciones de una muestra de tejido reticulado reversiblemente, lo que permite de esta manera que los agentes de captura capturen el producto de ligación.

65

El producto de ligación se captura por un agente de captura que se inmoviliza en una posición sobre el segundo sustrato (por ejemplo, una matriz artificial que comprende moléculas de ácido nucleico inmovilizadas), y un código de barras espacial del agente de captura corresponde a la posición de la sonda de captura sobre el segundo sustrato. En algunas modalidades, se genera un producto de la(s) molécula(s) capturada(s) en la posición sobre el segundo sustrato, por ejemplo, mediante el uso de una o más sondas que se hibridan con la(s) molécula(s) capturada(s). En algunas modalidades, la(s) molécula(s) capturada(s) y/o el producto de esta(s) no se detectan o se analizan mientras están presentes en la posición; en cambio, la(s) molécula(s) capturada(s) y/o el producto de esta(s) se eliminan del segundo sustrato y subsecuentemente se detectan o se analizan, por ejemplo, mediante la secuenciación de una mezcla de moléculas eliminadas que incluyen las que comprenden una secuencia del segundo ácido nucleico diana o una secuencia complementaria de este y una secuencia del dominio posicional (por ejemplo, código de barras espacial). En algunas modalidades, la(s) molécula(s) capturada(s) y/o el producto de esta(s) no se ponen en contacto con una sonda marcada de manera detectable (por ejemplo, una sonda marcada con fluorescencia) mientras permanecen en la posición sobre el segundo sustrato. En algunas modalidades, el producto de la(s) molécula(s) capturada(s) es un producto de extensión del cebador o un producto de transcripción inversa. En algunas modalidades, el producto de la(s) molécula(s) capturada(s) no es un producto de amplificación por círculo rodante. En algunas modalidades, el producto de la(s) molécula(s) capturada(s) es un producto de amplificación por círculo rodante (RCA), pero el método no comprende detectar el producto de RCA en la posición sobre el segundo sustrato.

En cualquiera de las modalidades descritas en la presente descripción, el primer y el segundo ácidos nucleicos diana pueden ser iguales o diferentes. En cualquiera de las modalidades descritas en la presente descripción, el segundo ácido nucleico diana puede comprender uno o más del primer ácido nucleico diana o un complemento de este o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este. En cualquiera de las modalidades descritas en la presente descripción, el primer y el segundo ácidos nucleicos diana pueden comprender secuencias de solapamiento, o pueden no solaparse en secuencia.

El método comprende además generar un polinucleótido marcado espacialmente que comprende (i) una secuencia del producto de la ligación y (ii) una secuencia del código de barras espacial o complemento de esta. En algunas modalidades, el polinucleótido marcado espacialmente se genera a partir del ácido nucleico capturado, donde todo o una porción del polinucleótido se libera del sustrato para su análisis. En algunas modalidades, el método descrito en la presente descripción se usa para analizar la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad del primer y el segundo ácidos nucleicos diana en la muestra biológica.

En cualquiera de las modalidades anteriores, los agentes de captura no necesitan proporcionarse sobre el mismo sustrato en el que está la muestra biológica. En otras palabras, la muestra biológica puede estar sobre un primer sustrato para el análisis *in situ*, y las moléculas en la muestra que han experimentado análisis *in situ* pueden ponerse en contacto con uno o más segundos sustratos. En el caso de múltiples segundos sustratos, la muestra puede ponerse en contacto con cada segundo sustrato secuencialmente, o en paralelo si dos o más segundos sustratos cubren solo una subregión de la muestra. Uno o más de los segundos sustratos pueden ponerse en contacto con uno o más terceros sustratos para producir réplicas del/de los segundo(s) sustrato(s).

En cualquiera de las modalidades anteriores, el código de barras espacial corresponde a la posición del agente de captura sobre el primer sustrato o un segundo sustrato.

En cualquiera de las modalidades anteriores, el primer ácido nucleico diana y el segundo ácido nucleico diana pueden comprender la misma secuencia de ácido nucleico. En cualquiera de las modalidades anteriores, el primer ácido nucleico diana y el segundo ácido nucleico diana pueden comprender diferentes secuencias de ácido nucleico. En cualquiera de las modalidades anteriores, el primer ácido nucleico diana y el segundo ácido nucleico diana pueden ser la misma molécula. En cualquiera de las modalidades anteriores, el primer ácido nucleico diana y el segundo ácido nucleico diana pueden ser moléculas diferentes.

En cualquiera de las modalidades anteriores, el segundo ácido nucleico diana puede ser o comprender el primer ácido nucleico diana, un complemento de este, un producto de hibridación de este, un producto de ligación de este, un producto de extensión de este, un producto de replicación de este, un producto de transcripción/transcripción inversa de este, y/o un producto de amplificación de este.

En cualquiera de las modalidades anteriores, el segundo ácido nucleico diana puede ser o comprender al menos una de la una o más sondas de ácido nucleico, un complemento de la(s) sonda(s) de ácido nucleico, un producto de hibridación de la(s) sonda(s) de ácido nucleico, un producto de ligación de la(s) sonda(s) de ácido nucleico, un producto de extensión de la(s) sonda(s) de ácido nucleico, un producto de replicación de la(s) sonda(s) de ácido nucleico, un producto de transcripción/transcripción inversa de la(s) sonda(s) de ácido nucleico, y/o un producto de amplificación de la(s) sonda(s) de ácido nucleico.

El primer ácido nucleico diana es el ARN. En cualquiera de las modalidades anteriores, el segundo ácido nucleico diana puede comprender además una secuencia de ARN. En cualquiera de las modalidades anteriores, el primer ácido nucleico diana y/o el segundo ácido nucleico diana pueden ser una molécula de ARN.

5 En cualquiera de las modalidades anteriores, la muestra biológica puede reticularse reversiblemente previo a, o durante el contacto de una muestra biológica sobre un primer sustrato con una o más sondas de ácido nucleico que se hibridan directa o indirectamente con un primer ácido nucleico diana o un complemento o un producto de amplificación de este en la muestra biológica.

10 En cualquiera de las modalidades anteriores, el método puede comprender eliminar la reticulación de la muestra biológica reticulada reversiblemente.

En cualquiera de las modalidades anteriores, el método puede comprender liberar la pluralidad de agentes de captura del sustrato (o el primer sustrato) o del segundo sustrato, y/o suministrar o conducir la pluralidad liberada de agentes de captura hacia y/o dentro de o sobre la muestra biológica.

15 En cualquiera de las modalidades anteriores, la pluralidad de agentes de captura puede comprender un agente de captura que se acopla de manera liberable a un agente de captura del analito.

20 En cualquiera de las modalidades anteriores, la detección de la una o más sondas de ácido nucleico en una ubicación espacial se realiza sobre una muestra biológica que se procesa o se clarifica. En cualquiera de las modalidades anteriores, la muestra biológica puede comprender una muestra de tejido. En cualquiera de las modalidades anteriores, la muestra de tejido puede ser un segmento de tejido entre aproximadamente 1 μm y aproximadamente 50 μm de grosor. En cualquiera de las modalidades anteriores, el segmento de tejido puede tener entre aproximadamente 5 μm y aproximadamente 35 μm de grosor. En cualquiera de las modalidades anteriores, la muestra de tejido puede incluirse en un hidrogel.

En cualquiera de las modalidades anteriores, el primer ácido nucleico diana, el complemento de este y/o el producto de amplificación de este pueden reticularse reversiblemente a una matriz, *por ejemplo*, un hidrogel.

30 Al menos una de la una o más primeras sondas de ácido nucleico comprende una o más secuencias de código de barras que corresponden a una secuencia del primer ácido nucleico diana.

La detección de la una o más primeras sondas de ácido nucleico comprende obtener imágenes de la muestra biológica.

35 En cualquiera de las modalidades anteriores, puede determinarse la detección de la una o más sondas de ácido nucleico de una secuencia de la secuencia de código de barras o una secuencia complementaria de esta o una secuencia amplificada de esta.

40 En cualquiera de las modalidades anteriores, la secuenciación *in situ* puede realizarse en la detección de la una o más sondas de ácido nucleico. En cualquiera de las modalidades anteriores, la detección de la una o más sondas de ácido nucleico puede comprender secuenciación *in situ* y/o hibridación secuencial de una pluralidad de sondas.

En cualquiera de las modalidades anteriores, la una o más sondas de ácido nucleico pueden comprender una sonda primaria que se hibrida directamente con el primer ácido nucleico diana o el complemento o el producto de amplificación de este.

45 En cualquiera de las modalidades anteriores, el primer ácido nucleico diana puede ser un ARNm, el complemento puede ser un ADNc y/o el producto de amplificación puede ser un producto de amplificación por círculo rodante (RCA). En cualquiera de las modalidades anteriores, la sonda primaria puede comprender una sonda candado, una sonda circular, o una sonda circularizada.

50 En cualquiera de las modalidades anteriores, la sonda primaria puede comprender una o más secuencias de código de barras que, opcionalmente corresponden a una secuencia del primer ácido nucleico diana.

55 En cualquiera de las modalidades anteriores, la etapa de detección puede comprender poner en contacto la muestra biológica con una o más sondas marcadas de manera detectable capaces de hibridar directa o indirectamente con la sonda primaria (o un complemento o producto de amplificación de esta), opcionalmente en donde la una o más sondas marcadas de manera detectable hibridan con una o más secuencias de código de barras de la sonda primaria (o un complemento o producto de amplificación de esta).

60 En cualquiera de las modalidades anteriores, la etapa de detección puede comprender poner en contacto la muestra biológica con una o más sondas secundarias capaces de hibridar directa o indirectamente con la sonda primaria (o un complemento o producto de amplificación de esta), opcionalmente en donde la una o más sondas secundarias hibridan con una o más secuencias de código de barras de la sonda primaria (o complemento o un producto de amplificación de esta).

65

En cualquiera de las modalidades anteriores, la etapa de detección puede comprender poner en contacto la muestra biológica con una o más sondas marcadas de manera detectable capaces de hibridar directa o indirectamente con la una o más sondas secundarias (o un complemento o producto de amplificación de estas), opcionalmente en donde la una o más sondas marcadas de manera detectable se hibridan con una o más secuencias de código de barras de la una o más sondas secundarias (o un complemento o producto de amplificación de estas).

En cualquiera de las modalidades anteriores, el método puede comprender obtener imágenes de la muestra biológica para secuenciar la una o más secuencias de código de barras de la sonda primaria (o un complemento o producto de amplificación de esta), y/o la una o más secuencias de código de barras de la una o más sondas secundarias (o un complemento o producto de amplificación de estas), opcionalmente en donde la secuenciación comprende secuenciación mediante ligación o secuenciación mediante hibridación.

En cualquiera de las modalidades anteriores, el método puede comprender además generar un producto de amplificación que comprende una secuencia de la secuencia de código de barras o una secuencia complementaria de esta. En cualquiera de las modalidades anteriores, el producto de amplificación puede generarse mediante amplificación por círculo rodante. En cualquiera de las modalidades anteriores, el producto de amplificación puede comprender uno o más nucleótidos modificados.

En cualquiera de las modalidades anteriores, la secuenciación *in situ* puede comprender secuenciación mediante ligación. En cualquiera de las modalidades anteriores, la secuenciación *in situ* puede comprender la secuenciación mediante hibridación.

En cualquiera de las modalidades anteriores, el método puede comprender además permeabilizar la muestra biológica sobre el sustrato después de la obtención de imágenes.

En cualquiera de las modalidades anteriores, el segundo ácido nucleico diana puede ser una molécula de ARNm, y el agente de captura comprende una sonda de captura.

En cualquiera de las modalidades anteriores, la sonda de captura puede comprender un extremo 3' libre de manera que la sonda de captura funciona como un cebador de transcriptasa inversa (RT) mediante el uso del segundo ácido nucleico diana como un molde para la extensión del cebador.

En cualquiera de las modalidades anteriores, el extremo 3' libre puede comprender un oligo dT, una secuencia aleatoria o una secuencia específica de un gen.

En cualquiera de las modalidades anteriores, la sonda de captura puede comprender además un dominio universal que está 5' con respecto al código de barras espacial, en donde el dominio universal comprende un dominio de amplificación; y/o un dominio de escisión para liberar el polinucleótido que se genera marcado espacialmente desde la superficie del sustrato.

En cualquiera de las modalidades anteriores, el polinucleótido que se genera marcado espacialmente puede ser un ADNc o producto de amplificación de este.

En cualquiera de las modalidades anteriores, el dominio de captura puede no capturar el primer ácido nucleico diana, el complemento de este, o el producto de amplificación de este previo a proporcionar condiciones para permitir una pluralidad de agentes de captura para capturar directa o indirectamente un segundo ácido nucleico diana o un complemento de este o un producto de amplificación de este.

En cualquiera de las modalidades anteriores, el dominio de captura puede no capturar el primer ácido nucleico diana, el complemento de este, o el producto de amplificación de este previo a eliminar las reticulaciones de la muestra biológica reticulada reversiblemente.

En cualquiera de las modalidades anteriores, la sonda de captura puede capturar el segundo ácido nucleico diana o complemento de este o producto de amplificación de este por hibridación, mediante ligación, o por hibridación seguido de ligación, *por ejemplo*, mediante ligación con soporte.

En cualquiera de las modalidades anteriores, el agente de captura puede capturar la una o más sondas de ácido nucleico o complementos de estas o productos de amplificación de estas.

En cualquiera de las modalidades anteriores, el agente de captura puede capturar la una o más sondas de ácido nucleico que se hibridan con el primer ácido nucleico diana.

En cualquiera de las modalidades anteriores, el método puede comprender además liberar la una o más sondas de ácido nucleico que se hibridan con el primer ácido nucleico diana previo a proporcionar condiciones para permitir que una pluralidad de agentes de captura capturen directa o indirectamente un segundo ácido nucleico diana o un complemento de este o un producto de amplificación de este.

En cualquiera de las modalidades anteriores, el polinucleótido marcado espacialmente puede comprender una secuencia de una sonda de ácido nucleico de la una o más sondas de ácido nucleico o un complemento de estas y una secuencia del código de barras espacial o complemento de este.

- 5 En cualquiera de las modalidades anteriores, el polinucleótido marcado espacialmente o una porción de este puede liberarse (por ejemplo, del primer o segundo sustrato) para su análisis.

En cualquiera de las modalidades anteriores, el método puede comprender además determinar al menos una porción del polinucleótido marcado espacialmente o un complemento de este.

- 10 En cualquiera de las modalidades anteriores, determinar puede comprender secuenciación mediante ligación, secuenciación mediante hibridación, secuenciación mediante síntesis y/o secuenciación mediante unión.

- 15 En cualquiera de las modalidades anteriores, el polinucleótido marcado espacialmente liberado puede analizarse mediante secuenciación directa o secuenciación indirecta, opcionalmente con amplificación previo a la secuenciación. En cualquiera de las modalidades anteriores, el método puede comprender además correlacionar el código de barras espacial del polinucleótido marcado espacialmente y la ubicación espacial detectada de la una o más sondas de ácido nucleico.

- 20 Breve descripción de los dibujos

La figura 1 muestra un flujo de trabajo ilustrativo de un ensayo espacial integrado *in situ* descrito en la presente descripción que comprende un módulo de ensayo *in situ* y un módulo de ensayo espacial.

- 25 Las figuras 2A-2B muestran flujos de trabajo ilustrativos del ensayo espacial integrado *in situ* descrito en la presente descripción.

La figura 3 muestra un flujo de trabajo ilustrativo del ensayo espacial integrado *in situ* descrito en la presente descripción.

- 30 La figura 4A muestra un flujo de trabajo de ensayo *in situ* ilustrativo, donde un conjunto de sondas primarias ilustrativas se une a un ARN diana, y los oligos de detección ilustrativos se hibridan con una secuencia de código de barras de la sonda primaria o un producto de amplificación (por ejemplo, producto de RCA) de esta. En el caso de una sonda candado, el candado puede ligarse mediante el uso del cebador como un ADN de soporte.

- 35 La figura 4B muestra un flujo de trabajo de ensayo *in situ* ilustrativo, donde un conjunto de sondas primarias ilustrativas se une a un ARN diana, las sondas secundarias ilustrativas cada una se hibrida con una secuencia de código de barras de la sonda primaria o un producto de amplificación (por ejemplo, producto de RCA) de esta, y los oligos de detección se hibridan con una secuencia de código de barras de una sonda secundaria. En el caso de una sonda candado, el candado puede ligarse mediante el uso del cebador como un ADN de soporte.

- 40 La figura 4C muestra un flujo de trabajo de ensayo *in situ* ilustrativo. Un ARN diana se transcribe inversamente para generar una molécula de ADN, y una sonda primaria ilustrativa después se hibrida con la molécula de ADN. En el caso de una sonda candado, el candado puede ligarse mediante el uso del ADN que se genera a partir del ARN como un soporte. Una sonda puente ilustrativa se hibrida con una secuencia de código de barras de la sonda primaria o un producto de amplificación (por ejemplo, producto de RCA) de esta, y los oligos de detección se hibridan con la sonda puente.

- 45 La figura 5 muestra flujos de trabajo de ensayo *in situ* ilustrativos. Diversas sondas primarias pueden hibridarse a un ARN diana y ligarse mediante el uso de ligación con molde de ARN y/o ligación con molde de ADN para formar una sonda circularizada que comprende una o más secuencias de código de barras. Una sonda secundaria o sonda puente puede hibridarse con la sonda circularizada o un producto de amplificación de esta (por ejemplo, como se muestra en las figuras 4A-4C). Los oligos de detección pueden hibridarse con la sonda circularizada o un producto de amplificación de esta, una sonda secundaria o un producto de amplificación de esta, o una sonda puente (por ejemplo, como se muestra en las figuras 4A-4C).

Las figuras 6A-6F muestran flujos de trabajo de ensayo *in situ* ilustrativos, por ejemplo, para un analito que no es ácido nucleico, mediante el uso de un agente de marcaje descrito en la presente descripción.

- 60 La figura 7 muestra un flujo de trabajo ilustrativo de ensayo espacial.

La figura 8 muestra un flujo de trabajo ilustrativo de ensayo espacial.

La figura 9 muestra un flujo de trabajo ilustrativo de ensayo espacial.

- 65 La figura 10 muestra un flujo de trabajo ilustrativo de ensayo espacial.

Las figuras 11A-11B muestran flujos de trabajo ilustrativos de ensayos espaciales.

La figura 12 es un diagrama esquemático que muestra un ejemplo de una sonda de captura con código de barras, como se describe en la presente descripción.

Las figuras 13A-13C ilustran esquemáticamente un método para analizar una molécula de ácido nucleico diana. La figura 13A ilustra una molécula de ácido nucleico diana, una primera sonda y una segunda sonda, y la figura 13B ilustra una molécula de ácido nucleico diana con la primera y la segunda sondas que se hibridan con esta. La figura 13C ilustra una molécula de ácido nucleico que se une a una sonda.

La figura 14 es un diagrama esquemático de un agente de marcaje ilustrativo (por ejemplo, un agente de captura del analito).

La figura 15 es un diagrama esquemático de un agente de marcaje ilustrativo (por ejemplo, un agente de captura del analito).

La figura 16 es un diagrama esquemático que representa una interacción ilustrativa entre una sonda de captura que se inmoviliza por un elemento y un agente de marcaje (por ejemplo, agente de captura del analito) con un oligonucleótido reportero asociado.

La figura 17 es un esquema que muestra el arreglo de elementos con código de barras dentro de una matriz.

La figura 18 es un esquema que ilustra una vista lateral de un medio resistente a la difusión, por ejemplo, una tapa.

Las figuras 19A-19B son esquemas que ilustran la vista expandida (figura 19A), y la vista lateral (figura 19B) de un sistema de transferencia electroforética configurado para dirigir los analitos transcritos hacia una matriz de sondas de captura con código de barras espacial.

La Figura 20 es un esquema que ilustra un flujo de trabajo ilustrativo que utiliza un sistema de transferencia electroforética.

Descripción detallada de la descripción

La información técnica que se expone más abajo puede, en algunos aspectos, ir más allá de la descripción de la invención, que se define por las reivindicaciones anexas. La información técnica adicional se proporciona para colocar la invención en un contexto técnico más amplio y para ilustrar posibles desarrollos técnicos relacionados. Dicha información técnica adicional que no cae dentro del alcance de las reivindicaciones anexas no es parte de la invención. En consecuencia, cualquiera de los aspectos, modalidades y ejemplos de la presente descripción que no caen dentro del alcance de las reivindicaciones anexas no forman parte de la invención y se proporcionan solamente con fines ilustrativos.

En la presente descripción, se proporcionan ensayos espaciales integrados *in situ* mediante el uso de una lectura de microscopía (por ejemplo, secuenciación óptica de una secuencia de código de barras de una sonda se une directa o indirectamente a un analito diana) y/o una lectura de secuenciación (por ejemplo, secuenciación NGS de una secuencia de ácido nucleico diana *per se* y/o una secuencia de código de barras de una sonda), por ejemplo, para analizar una célula en un tejido intacto. En algunas modalidades, el método comprende además determinar la caracterización espacial de los analitos tales como el transcriptoma o un subconjunto de estos en una muestra biológica. Se proporcionan métodos, composiciones, kits, dispositivos y sistemas para estos ensayos espaciales *in situ*, que incluyen ensayos espaciales de genómica y transcriptómica. En algunas modalidades, un método que se proporciona es cuantitativo y conserva la información espacial dentro de una muestra de tejido sin aislar físicamente las células o mediante el uso de homogeneizados. Además, en la presente descripción se proporcionan composiciones y métodos para detectar y/o cuantificar ácidos nucleicos en células, tejidos, órganos u organismos. En algunas modalidades, la presente descripción proporciona métodos para la caracterización con alto rendimiento de un gran número de dianas *in situ*, que incluyen información espacial de tales dianas, tales como transcritos de ARN y/o loci de ADN en una muestra de tejido.

Si una definición expuesta en la presente descripción es contraria o inconsistente de cualquier otra manera con una definición expuesta en las patentes, solicitudes, solicitudes publicadas y otras publicaciones a las que se hace referencia, la definición expuesta en la presente descripción prevalece sobre la definición que está en la referencia. Los encabezados de sección usados en la presente descripción son solamente para propósitos organizativos y no deben interpretarse como limitantes del objeto descrito.

I. Descripción general

En un aspecto, en la presente descripción se proporciona un método que comprende un módulo de ensayo *in situ* para uno o más analitos diana de interés en una muestra biológica, por ejemplo, una muestra de tejido. En algunas modalidades, el ensayo comprende analizar la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad de analitos (por ejemplo, moléculas de ácido nucleico) en una muestra de tejido *in situ*. Los analitos diana pueden incluir moléculas de ácido nucleico y moléculas que no son ácido nucleico, tales como proteínas y péptidos. Las moléculas de ácido nucleico diana pueden derivarse o analizarse en cualquier tipo específico de célula y/o una región subcelular específica, por ejemplo, a partir de citosol, a partir de núcleos celulares, a partir de mitocondrias, a partir de microsomas y, más generalmente, a partir de cualquier otro compartimiento, orgánulo o porción de una célula. Los ejemplos comprenden analitos de ADN tales como ADN genómico, ADN metilado, secuencias específicas de ADN metilado, ADN fragmentado, ADN mitocondrial e híbridos de ARN/ADN. Los ejemplos de moléculas de ácido nucleico diana comprenden además analitos de ARN tales como diversos tipos de ARN codificante y no codificante. Los ejemplos de los diferentes tipos de analitos de ARN comprenden ARN mensajero (ARNm), ARN ribosómico (ARNr), ARN de transferencia (ARNt), microARN (miARN), pre-ARNm y ARN viral. Los analitos de ARN pueden obtenerse o analizarse en células o compartimientos celulares (por ejemplo, núcleo).

En algunas modalidades, el método comprende analizar un transcriptoma o genoma, por ejemplo, el transcriptoma o genoma global, de una muestra de tejido. En algunas modalidades, el método comprende un proceso para realizar la transcriptómica espacial y/o la genómica espacial, que permite el análisis simultáneo de un patrón de expresión y/o un patrón de ubicación/distribución de los genes o loci genómicos expresados o presentes en una muestra de tejido.

En algunas modalidades, el ensayo *in situ* es un ensayo dirigido, por ejemplo, uno que analiza sondas prediseñadas que se unen directa o indirectamente a analitos biológicos diana de interés, por ejemplo, moléculas de ARNm en una muestra de tejido. En algunas modalidades, el ensayo *in situ* comprende proporcionar una o más sondas de ácido nucleico que se hibridan con un ácido nucleico diana (o un complemento, producto de amplificación o derivado de este) y detectar la una o más sondas de ácido nucleico en una ubicación espacial de la muestra biológica. En algunas modalidades, las sondas prediseñadas comprenden una o más sondas o conjuntos de sondas primarias, por ejemplo, una sonda que se une a un analito diana, y/o una o más sondas o conjuntos de sondas secundarias, por ejemplo, una sonda que se une a una sonda o complemento primario de esta o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este. En algunas modalidades, las sondas prediseñadas comprenden una o más sondas o conjuntos de sondas de orden superior, por ejemplo, una sonda de $(n+1)^{\text{ésimo}}$ orden que se une a una sonda de $n^{\text{ésimo}}$ orden o complemento de esta o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de esta, donde n es un número entero de 2 o más. En algunas modalidades, las sondas prediseñadas son sondas con código de barras. En algunas modalidades, una sonda de $n^{\text{ésimo}}$ orden comprende una secuencia de código de barras (una secuencia de código de barras de $n^{\text{ésimo}}$ orden) y una sonda de $(n+1)^{\text{ésimo}}$ orden se une a la secuencia de código de barras de $n^{\text{ésimo}}$ orden y comprende una secuencia de código de barras de $(n+1)^{\text{ésimo}}$ orden para unirse por otra sonda con código de barras o por una sonda marcada de manera detectable, por ejemplo, un oligo de detección marcado con fluorescencia, donde n es un número entero de 1 o más. En cualquiera de las modalidades descritas en la presente descripción, la unión de una sonda a otra sonda o a un analito diana puede ser directa (por ejemplo, mediante hibridación directa de secuencias de ácido nucleico o unión antígeno-anticuerpo) o indirecta (por ejemplo, hibridación indirecta mediante uno o más oligos puente o interacción de unión). Las interacciones de unión pueden analizarse mediante el uso de microscopía, tal como microscopía óptica de alta resolución, para proporcionar lecturas de la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad del analito diana. En algunas modalidades, el ensayo *in situ* comprende secuenciación *in situ* y/o hibridación *in situ*, tal como hibridación secuencial de sondas. En algunas modalidades, el ensayo *in situ* analiza aproximadamente 20, aproximadamente 50, aproximadamente 100, aproximadamente 200, aproximadamente 500, aproximadamente 1000, aproximadamente 2000, aproximadamente 5000 o aproximadamente 10 000 genes, por ejemplo, transcritos de ARNm de los genes en una muestra de tejido. En algunas modalidades, los transcritos de ARNm de entre aproximadamente 100 y aproximadamente 1000 genes en una muestra de tejido se analizan *in situ*.

En otro aspecto, un método descrito en la presente descripción comprende además un módulo de ensayo espacial, por ejemplo, un análisis cuantitativo y/o cualitativo de la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad de analitos (por ejemplo, moléculas de ácido nucleico) en una muestra de tejido en donde se retiene el patrón espacial de la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad de los analitos dentro de la muestra de tejido. En algunas modalidades, el método comprende analizar un transcriptoma o genoma, por ejemplo, el transcriptoma o genoma global, de una muestra de tejido. En algunas modalidades, el método comprende un proceso para realizar la transcriptómica espacial y/o la genómica espacial, que permite el análisis simultáneo de un patrón de expresión y/o un patrón de ubicación/distribución de los genes o loci genómicos expresados o presentes en una muestra de tejido. En algunas modalidades, el método acopla la captura basada en matrices de moléculas de ácido nucleico con tecnologías de secuenciación de ácido nucleico de alto rendimiento, lo que permite que las moléculas de ácido nucleico (por ejemplo, moléculas de ARN o ADN) en la muestra de tejido, se capturen y se marquen con una etiqueta posicional (por ejemplo, un código de barras espacial). En algunas modalidades, el método comprende además la síntesis de moléculas de ácido nucleico que se secuencian, por ejemplo, con resolución de nucleótidos, y se analizan para determinar qué genes se expresan en cualquiera y todas las partes de la muestra de tejido. En algunas modalidades, el transcriptoma individual, separado y específico de cada

célula en la muestra de tejido puede obtenerse al mismo tiempo, lo que proporciona firmas de transcriptoma completas altamente paralelas a partir de células individuales dentro de una muestra de tejido sin perder información espacial dentro de dicha muestra de tejido investigada.

- 5 En algunas modalidades, un método descrito en la presente descripción comprende realizar secuencialmente uno o más ensayos *in situ* y uno o más ensayos espaciales en cualquier orden adecuado. El análisis *in situ* de un primer analito puede realizarse antes, simultáneamente con o después de analizar un segundo analito diana con la matriz de sondas de captura en el ensayo espacial. En algunas modalidades, un método descrito en la presente descripción comprende proporcionar una lectura doble, por ejemplo, una lectura de microscopía y una lectura de secuenciación.
- 10 En algunas modalidades, un método descrito en la presente descripción comprende analizar una lectura de microscopía para el ensayo *in situ* y una lectura de secuenciación para el ensayo espacial. En algunas modalidades, un método descrito en la presente descripción comprende analizar un primer analito en el ensayo *in situ* y un segundo analito en el ensayo espacial. En algunas modalidades, el primer analito y el segundo analito son la misma molécula. En algunas modalidades, el primer analito y el segundo analito son moléculas diferentes. En algunas modalidades, el
- 15 primer analito y el segundo analito comprenden una secuencia de ácido nucleico común, o uno comprende una secuencia de ácido nucleico y el otro comprende una secuencia complementaria o un producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de esta. En algunos ejemplos, el segundo analito puede comprender una secuencia de ácido nucleico del primer analito o una secuencia de ácido nucleico de un agente de marcaje para el primer analito. En algunos
- 20 ejemplos, el segundo analito puede comprender una secuencia de ácido nucleico complementaria de una secuencia de ácido nucleico del primer analito o una secuencia de ácido nucleico de un agente de marcaje para el primer analito. En algunos ejemplos, el segundo analito puede ser un producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación,
- 25 un producto de transcripción/transcripción inversa y/o un producto de amplificación, por ejemplo, un producto de RCA) del primer analito o una sonda que se dirige al primer analito (por ejemplo, una sonda que se usa en el ensayo *in situ*).

- En algunas modalidades, el primer analito y el segundo analito comprenden diferentes secuencias de ácido nucleico. En algunas modalidades, un método descrito en la presente descripción comprende poner en contacto una muestra
- 30 biológica sobre un sustrato, por ejemplo, un portaobjeto de vidrio, con una o más sondas de ácido nucleico que se hibridan directa o indirectamente con un primer ácido nucleico diana o un complemento o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa, y/o un producto de amplificación) de este en la muestra biológica. En algunas modalidades, la muestra biológica sobre el sustrato se
- 35 somete a un análisis *in situ*, por ejemplo, secuenciación *in situ* para el primer ácido nucleico diana en la muestra. En algunas modalidades, la muestra biológica es una muestra de tejido tal como una sección de tejido aislada al momento o conservada. En algunas modalidades, la muestra biológica se fija mediante reticulación reversible. En algunas modalidades, la muestra biológica se procesa de manera que una o más moléculas del analito (por ejemplo, moléculas de ARN) se bloquean reversiblemente en su lugar para conservar un patrón espacial de la presencia/ausencia,
- 40 distribución, ubicación, cantidad, nivel, expresión o actividad del analito dentro de la muestra de tejido y/o relativamente a uno o más analitos en la muestra de tejido. En algunas modalidades, la muestra biológica se incluye en una matriz, tal como una matriz polimérica. En algunas modalidades, la muestra biológica se incluye en hidrogel. En algunas modalidades, la una o más moléculas del analito (por ejemplo, moléculas de ARN) se seleccionan por sondas y se analizan mediante el uso de imágenes *in situ*, por ejemplo, secuenciación mediante ligación, secuenciación mediante
- 45 hibridación, secuenciación mediante síntesis, secuenciación mediante unión y/o hibridación secuencial de sondas con código de barras seguido de decodificación. En algunas modalidades, las moléculas de una pluralidad de analitos en la muestra se analizan *in situ* en un enfoque altamente multiplexado.

- En algunas modalidades, una vez que se registran las imágenes del análisis *in situ*, el método comprende además
- 50 tratar la muestra de manera que se libere un segundo ácido nucleico diana o un complemento o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión, un producto de replicación, un producto de transcripción/transcripción inversa, derivado y/o un producto de amplificación) de este en la muestra. Por ejemplo, el método puede comprender además eliminar las reticulaciones de la muestra de manera que un segundo ácido nucleico diana o un complemento o producto (por ejemplo, un producto de hibridación, un producto de ligación, un
- 55 producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este en la muestra no se bloquea en su lugar (por ejemplo, a un hidrogel). En algunas modalidades, el segundo ácido nucleico diana (por ejemplo, ARNm) o un complemento de este o un producto de este puede capturarse directa o indirectamente mediante una pluralidad de agentes de captura. En algunas modalidades, la pluralidad de agentes de captura se proporciona sobre el sustrato
- 60 (por ejemplo, un portaobjetos de vidrio delgado) que soporta la muestra durante el análisis *in situ*. En algunas modalidades, el sustrato para la obtención de imágenes *in situ* es un primer sustrato, y la pluralidad de agentes de captura se proporcionan sobre un segundo sustrato que se proporciona a la muestra sobre el primer sustrato. En algunas modalidades, un agente de captura de la pluralidad de agentes de captura comprende (i) un dominio de captura capaz de capturar un ácido nucleico y (ii) un código de barras espacial que corresponde a la posición del
- 65 agente de captura en el primer sustrato y/o el segundo sustrato.

En algunas modalidades, el primer ácido nucleico diana y el segundo ácido nucleico diana son la misma molécula. En algunas modalidades, el primer ácido nucleico diana y el segundo ácido nucleico diana son moléculas diferentes. En algunas modalidades, el primer ácido nucleico diana y el segundo ácido nucleico diana comprenden una secuencia de ácido nucleico común, o uno comprende una secuencia de ácido nucleico y el otro comprende una secuencia complementaria o un producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de la secuencia de ácido nucleico. En algunas modalidades, el primer ácido nucleico diana y el segundo ácido nucleico diana comprenden diferentes secuencias de ácido nucleico. En algunas modalidades, el segundo ácido nucleico diana comprende una sonda que se usa en el análisis *in situ*, que puede liberarse de la muestra después del análisis *in situ*. En algunas modalidades, el segundo ácido nucleico diana comprende un producto de amplificación que se genera durante el análisis *in situ*, que puede liberarse de la muestra después del análisis *in situ*. En algunas modalidades, el segundo ácido nucleico diana comprende una secuencia de ácido nucleico que se selecciona por una o más sondas de ácido nucleico, que pueden liberarse de la muestra después del análisis *in situ*. En algunas modalidades, el segundo ácido nucleico diana comprende un ácido nucleico que se genera durante el análisis *in situ* (por ejemplo, mediante ligación o extensión), que puede liberarse de la muestra después del análisis *in situ*. En algunas modalidades, el segundo ácido nucleico diana comprende dos o más tipos de analitos. El segundo ácido nucleico diana puede generarse o procesarse antes, simultáneamente con, o después de que se realice el análisis *in situ* de un primer analito.

En algunas modalidades, un método descrito en la presente descripción comprende procesar una muestra, por ejemplo, mediante lisado o permeabilización de una muestra de tejido a la que se eliminan las reticulaciones, para permitir que una o más moléculas en la muestra se capturen directa o indirectamente por un agente de captura. En algunas modalidades, se permite que la una o más moléculas en la muestra migren, por ejemplo, hacia fuera de una o más células de la muestra y/o hacia fuera de la muestra. En algunas modalidades, el primer ácido nucleico diana o un complemento o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este puede migrar fuera de una célula de la muestra, para ser capturado por un agente de captura sobre un sustrato. En algunas modalidades, el segundo ácido nucleico diana o un complemento o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este puede migrar fuera de una célula de la muestra, para ser capturado por un agente de captura sobre un sustrato.

En algunas modalidades, un método descrito en la presente descripción comprende generar un polinucleótido marcado espacialmente que comprende (i) una secuencia del segundo ácido nucleico diana o complemento de este y (ii) una secuencia del código de barras espacial o complemento de este. En algunas modalidades, el segundo ácido nucleico diana es el mismo que el primer ácido nucleico diana que se analiza durante la obtención de imágenes *in situ*. En algunas modalidades, el segundo ácido nucleico diana es diferente del primer ácido nucleico diana que se analiza durante la obtención de imágenes *in situ*. En algunas modalidades, se analiza una pluralidad de los segundos ácidos nucleicos diana en el ensayo espacial. En algunas modalidades, la pluralidad de los segundos ácidos nucleicos diana son un transcriptoma o un subconjunto de este.

En algunas modalidades, el segundo ácido nucleico diana comprende una secuencia de ARNm, y el agente de captura comprende una sonda de captura. En algunas modalidades, la sonda de captura comprende un extremo 3' libre de manera que la sonda de captura funciona como un cebador de transcriptasa inversa (RT) mediante el uso del segundo ácido nucleico diana como un molde para la extensión del cebador. En algunas modalidades, el segundo ácido nucleico diana comprende ADN (por ejemplo, ADNc). En algunas modalidades, el segundo ácido nucleico diana se genera previo a, o durante el ensayo *in situ* mediante el uso de un molde de ARNm (por ejemplo, transcripción inversa). En algunas modalidades, la muestra se somete a una reacción de transcripción inversa, en donde una o más moléculas de ARN en la muestra se transcriben inversamente para generar moléculas de ADN (por ejemplo, ADNc) previo a, o durante un módulo de ensayo *in situ* descrito en la presente descripción. Las moléculas de ADN que se generan pueden analizarse en un módulo de ensayo *in situ* y/o un módulo de ensayo espacial de un método integrado descrito en la presente descripción. En algunas modalidades, el extremo 3' libre comprende un oligo dT, una secuencia aleatoria o una secuencia específica de un gen. En algunas modalidades, la sonda de captura comprende además un dominio universal que es 5' con respecto al código de barras espacial, en donde el dominio universal comprende: (i) un dominio de amplificación; y/o (ii) un dominio de escisión para liberar el polinucleótido que se genera marcado espacialmente de la superficie del sustrato. En algunas modalidades, el polinucleótido que se genera marcado espacialmente es un ADNc o producto de amplificación de este.

En algunas modalidades, un método descrito en la presente descripción comprende liberar el polinucleótido marcado espacialmente o una porción de este o un complemento o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este para su análisis. En algunas modalidades, el polinucleótido marcado espacialmente o una porción, complemento, o producto de este se libera de un sustrato, por ejemplo, del primer o segundo sustrato. En algunas modalidades, el método comprende una etapa para determinar una secuencia de al menos una porción del polinucleótido marcado

espacialmente o una porción, complemento, o producto de este. En algunas modalidades, la etapa para determinar comprende la secuenciación mediante ligación, secuenciación mediante hibridación, secuenciación mediante síntesis y/o secuenciación mediante unión. En algunas modalidades, el polinucleótido marcado espacialmente liberado o una porción, complemento o producto de este se analiza mediante secuenciación directa. En algunas modalidades, el polinucleótido marcado espacialmente liberado o una porción, complemento o producto de este se analiza mediante secuenciación indirecta. En algunas modalidades, el método comprende amplificar el polinucleótido marcado espacialmente o una porción, complemento o producto de este, por ejemplo, amplificar el polinucleótido marcado espacialmente o una porción, complemento o producto de este previo a, durante o después de la etapa de liberación. En algunas modalidades, la etapa para determinar comprende la amplificación del polinucleótido marcado espacialmente liberado o una porción, complemento, o producto de este previo a la secuenciación. En algunas modalidades, un método comprende correlacionar el código de barras espacial del polinucleótido marcado espacialmente y la ubicación espacial detectada de la una o más sondas de ácido nucleico.

La figura 1 proporciona un flujo de trabajo ilustrativo de un ensayo espacial integrado *in situ*. Un ensayo *in situ* típicamente adopta un enfoque dirigido que interroga un panel de analitos diana, por ejemplo, entre aproximadamente 100 y aproximadamente 1000 genes y transcritos de ARNm a partir de estos, mediante el uso de un microscopio y el análisis de señales ópticas para la determinación de las secuencias. En algunas modalidades, el ensayo *in situ* comprende la secuenciación *in situ* de una o más secuencias de código de barras de una sonda que se une directa o indirectamente a un analito. En algunas modalidades, el ensayo *in situ* comprende detectar señales que indican la hibridación *in situ* de uno o más oligos de detección a una o más secuencias de código de barras de una sonda que se hibrida directa o indirectamente con una porción de ácido nucleico de un analito, y se usa un patrón espacial y/o temporal de las señales de la hibridación secuencial para analizar el analito. En algunas modalidades, un ensayo espacial integrado *in situ* descrito en la presente descripción obtiene más información espacial de analitos a partir de una muestra, en comparación con un ensayo *in situ* sobre la misma muestra. En algunas modalidades, las sondas *in situ* (por ejemplo, sondas candado) se unen a los analitos diana (por ejemplo, moléculas de ácido nucleico) en una muestra para el análisis *in situ*. Las sondas *in situ* pueden comprender dos o más sondas que se ensamblan para formar una sonda más grande después de unirse a una molécula diana y conectar las dos o más sondas (por ejemplo, mediante ligación tal como ligación en molde de ARN (por ejemplo, como se muestra en las figuras 13A-13C) y/o ligación en molde de ADN). Las sondas *in situ* o productos de estas pueden amplificarse, por ejemplo, mediante el uso de amplificación por círculo rodante, para generar un producto de amplificación que puede secuenciarse *in situ*, por ejemplo, como se muestra en la figura 1. En algunas modalidades, al realizar un ensayo espacial integrado *in situ* descrito en la presente descripción, el ensayo *in situ* y la detección pueden dirigirse a detectar analitos diana (por ejemplo, mientras se evita el apilamiento óptico) que se complementa con el ensayo espacial que puede capturar información adicional (por ejemplo, información asociada con los analitos diana y/u otros analitos de interés) sobre la misma muestra. En algunas modalidades, un primer conjunto de analitos diana en una muestra biológica se analiza en el ensayo *in situ* y detección mientras que un segundo conjunto de analitos diana en la misma muestra biológica se analiza en el ensayo espacial.

En algunas modalidades, además de las sondas *in situ*, una o más sondas de marcaje y/o captura del analito descritas en la presente descripción también se ponen en contacto con la muestra previo a y/o durante el ensayo *in situ*. En algunas modalidades, un panel de sondas que se dirigen a uno o más analitos se pone en contacto con la muestra previo a y/o durante el ensayo *in situ* pero no se analizan en el análisis *in situ*. En su lugar, las sondas que se dirigen al uno o más analitos pueden analizarse en un análisis espacial como se describe en la presente descripción, tal como un ensayo espacial basado en matriz. En algunos casos, el panel de sondas es para capturar una pluralidad de analitos diana para el análisis en el análisis espacial. En algunas modalidades, los agentes de marcaje o agentes aglutinantes que se dirigen a uno o más analitos (por ejemplo, proteínas o elementos celulares) se ponen en contacto con la muestra previo a, o durante el ensayo *in situ* y/o el ensayo espacial.

Las figuras 1-3 proporcionan flujos de trabajo ilustrativos de un análisis espacial integrado *in situ*. Como se muestra en la figura 1, el análisis *in situ* puede no ser dirigido (por ejemplo, como en FISSEQ) y es preferentemente dirigido, por ejemplo, para entre aproximadamente 100 y aproximadamente 1000 genes y/o transcritos de ARNm. La misma muestra puede someterse a un análisis espacial, por ejemplo, un análisis espacial basado en matriz descrito en la presente descripción, que puede ser dirigido y/o no dirigido. En algunos aspectos, el análisis espacial puede proporcionar una lectura dirigida (por ejemplo, para un panel predeterminado de analitos diana) o una lectura no dirigida (por ejemplo, para la caracterización del transcriptoma de la muestra). En algunos aspectos, el análisis espacial usa tanto agentes de captura que tienen secuencias de captura específicas de genes para un panel predeterminado de transcritos de ARNm o sondas *in situ* (para lectura dirigida) como agentes de captura que tienen secuencias poli-T para la caracterización del transcriptoma no seleccionado. En algunas modalidades, los analitos, las sondas *in situ* y/o productos de estos, y/o los agentes de captura del analito (por ejemplo, sondas de captura del analito) migran a un sustrato que comprende elementos que comprenden sondas de captura.

En la figura 2A, una muestra se pone en contacto con sondas (por ejemplo, sondas candado) para un módulo de ensayo *in situ* en 201. Otros agentes, tales como sondas, agentes de marcaje, agentes de captura del analito y/o agentes de captura para el módulo de ensayo *in situ* y/o un módulo de ensayo espacial pueden ponerse en contacto con la muestra previo a, junto con o después de que la muestra se pone en contacto con sondas para el ensayo *in situ*. Opcionalmente, los analitos, sondas y/u otros reactivos (que incluye los del módulo de ensayo *in situ* y/o los del

módulo de ensayo espacial) pueden extenderse, ligarse, amplificarse, transcribirse de manera inversa y/o hacerse reaccionar o procesarse de cualquier otra manera, en cualquier orden y combinación adecuados, en 202. Por ejemplo, las moléculas de ARN en la muestra pueden transcribirse de manera inversa a moléculas de ADN que se generan, por ejemplo, en 201 y/o 202, previo a la detección *in situ*. Las señales del módulo de ensayo *in situ* se detectan en 203 y se analizan opcionalmente. Después del procesamiento y/o permeabilización opcional de la muestra, los analitos de la muestra después del módulo de ensayo *in situ* se capturan mediante el uso de sondas de captura en 204. Los agentes, tales como sondas (por ejemplo, sondas de ligación con molde), agentes de marcaje, agentes de captura del analito y/o agentes de captura para el módulo de ensayo espacial pueden ponerse en contacto con la muestra previo a, junto con o después de 201, 202, 203 y/o 204. La información del analito resuelta espacialmente se obtiene, por ejemplo, mediante el análisis de las sondas de captura y/o los analitos capturados, en 205. Las etapas del flujo de trabajo ilustrativo pueden realizarse en cualquier orden adecuado. En algunas modalidades, una cualquiera o más de las etapas 201, 202 y 203 se realizan previo a las etapas 204 y 205. Los analitos en el módulo de ensayo *in situ* y/o los del módulo de ensayo espacial pueden incluir uno o más analitos de ácido nucleico y/o uno o más analitos que no son ácido nucleico, tales como analitos de proteínas.

En la figura 2B, una muestra se pone en contacto con sondas (por ejemplo, sonda de ligación con molde) que se dirigen a analitos para un módulo de ensayo espacial en 206, y/o con sondas (por ejemplo, sondas candado) para un módulo de ensayo *in situ* 207. Otros agentes, tales como sondas, agentes de marcaje, agentes de captura del analito y/o agentes de captura para el módulo de ensayo *in situ* y/o el módulo de ensayo espacial pueden ponerse en contacto con la muestra previo a, junto con o después de la etapa 206 y/o 207. Opcionalmente, los analitos, sondas y/u otros reactivos (que incluye los del módulo de ensayo *in situ* y/o los del módulo de ensayo espacial) pueden extenderse, ligarse, amplificarse, transcribirse de manera inversa y/o hacerse reaccionar o procesarse de cualquier otra manera, en cualquier orden y combinación adecuados, en 208. Por ejemplo, las moléculas de ARN en la muestra pueden transcribirse de manera inversa a moléculas de ADN que se generan, por ejemplo, durante o previo a 206, 207 y/o 208, previo a la detección *in situ*. En algunos casos, 206 puede realizarse previo a, junto con o después de la etapa 207. Las señales del módulo de ensayo *in situ* se detectan en 209 y se analizan opcionalmente. Después del procesamiento y/o permeabilización opcional de la muestra, los analitos de la muestra después del módulo de ensayo *in situ* se capturan mediante el uso de sondas de captura en 210. Los agentes, tales como sondas (por ejemplo, sondas de ligación con molde), agentes de marcaje, agentes de captura del analito y/o agentes de captura para el módulo de ensayo espacial pueden ponerse en contacto con la muestra previo a, junto con o después de 207, 208, 209 y/o 210. La información del analito resuelta espacialmente se obtiene, por ejemplo, mediante el análisis de las sondas de captura y/o los analitos capturados, en 211. Las etapas del flujo de trabajo ilustrativo pueden realizarse en cualquier orden adecuado. En algunas modalidades, una cualquiera o más de las etapas 206, 207, 208 y 209 se realizan previo a las etapas 210 y 211. Los analitos en el módulo de ensayo *in situ* y/o los del módulo de ensayo espacial pueden incluir uno o más analitos de ácido nucleico y/o uno o más analitos que no son ácido nucleico, tales como analitos de proteínas.

La figura 3 muestra un flujo de trabajo ilustrativo adicional. Una muestra comprende o se pone en contacto con varias moléculas de ácido nucleico, que incluye opcionalmente una sonda con código de barras que se dirige al ácido nucleico 1; una sonda circularizable con código de barras (por ejemplo, sonda candado) que se dirige al ácido nucleico 2 y sondas con código de barras para ligación (por ejemplo, sondas de ligación con molde de ARN) que se dirigen al mismo ácido nucleico o a un ácido nucleico diferente; una sonda circularizable con código de barras (por ejemplo, sonda candado) que se dirige al ácido nucleico 3 y sondas con código de barras que usan una secuencia de la sonda circularizable (por ejemplo, una secuencia de código de barras de la sonda circularizable) como un molde para la ligación; el ácido nucleico 4 que puede ser parte de un agente de marcaje tal como un oligonucleótido reportero (el ácido nucleico 4 conjugado a un anticuerpo que reconoce un analito en la muestra se muestra como un ejemplo) o una molécula endógena en la muestra, tal como un ARN (por ejemplo, molécula de ARNm) que se somete a transcripción inversa para generar un ADNc *in situ*. Las diversas moléculas de ácido nucleico pueden presentarse en o ponerse en contacto con la muestra en cualquier combinación adecuada y en cualquier orden temporal adecuado. Las moléculas exógenas que incluyen las diversas sondas pueden ponerse en contacto con la muestra simultáneamente e incubarse en una reacción de unión/hibridación a un punto. A menos que se indique de otra forma, cualquiera de las diversas moléculas de ácido nucleico (que incluyen los ácidos nucleicos 1-4 y las diversas sondas) puede ser ADN o ARN o un complejo, conjugado o híbrido de estos, y puede ser endógeno a la muestra o añadirse exógenamente a la muestra. Las moléculas de ácido nucleico, que incluyen las del módulo de ensayo *in situ* y/o las del módulo de ensayo espacial pueden extenderse, ligarse, amplificarse, transcribirse de manera inversa y/o hacerse reaccionar o procesarse de cualquier otra manera, en cualquier orden y combinación adecuados. Las señales del módulo de ensayo *in situ* pueden detectarse y analizarse opcionalmente. Por ejemplo, la sonda circularizable con código de barras (por ejemplo, sonda candado) que se dirige al ácido nucleico 2 puede amplificarse mediante el uso de RCA, y el producto de RCA puede detectarse *in situ* mediante el uso de sondas de detección (por ejemplo, que incluyen sondas marcadas de manera detectable y/o sondas intermedias). Las sondas con código de barras para la ligación se ligan y pueden permanecer hibridadas al ácido nucleico 2, pero en algunos casos, no se detectan en el módulo de ensayo *in situ*. En el ejemplo del ácido nucleico 3, las sondas con código de barras para la ligación hibridan con la sonda candado y se ligan mediante el uso de la secuencia de código de barras como un molde. Sin embargo, durante la amplificación por círculo rodante de la sonda candado circularizada, la sonda ligada puede desplazarse de la sonda candado circularizada. La sonda ligada desplazada puede reticularse reversiblemente (por ejemplo, al ácido nucleico 3 o al producto de RCA o una o más moléculas en la muestra o una matriz en la que se incluye la muestra)

en la ubicación del ácido nucleico 3. Después del procesamiento y/o permeabilización opcional de la muestra, los analitos de la muestra después del módulo de ensayo *in situ* se capturan mediante los agentes de captura sobre una matriz espacial. La información del analito resuelta espacialmente se obtiene, por ejemplo, mediante el análisis de las sondas de captura y/o los analitos capturados, después de eliminar las moléculas capturadas de la matriz espacial y recolectar las moléculas capturadas y/o los productos de estas para la secuenciación.

Más abajo se proporcionan detalles adicionales de la descripción a manera de ejemplo.

II. Analitos

Los métodos, composiciones, aparatos y sistemas descritos en esta descripción pueden usarse para detectar y analizar una amplia variedad de analitos diferentes. En algunos aspectos, un analito descrito en la presente descripción puede incluir cualquier sustancia biológica, estructura, porción, o componente para analizar. En algunos aspectos, una diana descrita en la presente descripción puede incluir de manera similar cualquier analito de interés.

Los analitos pueden derivarse de un tipo específico de célula y/o de una región subcelular específica. Por ejemplo, los analitos pueden derivarse del citosol, de núcleos celulares, de mitocondrias, de microsomas, y más generalmente, de cualquier otro compartimiento, orgánulo, o porción de una célula. Los agentes permeabilizantes que se dirigen específicamente a determinados compartimientos y orgánulos celulares pueden usarse para liberar selectivamente analitos de las células para su análisis, y/o para permitir el acceso de uno o más reactivos (por ejemplo, sondas para el análisis *in situ* y/o sondas para el análisis espacial) a los analitos en la célula o compartimiento u orgánulo celular.

Los analitos pueden clasificarse en términos generales en uno de dos grupos: analitos de ácido nucleico y analitos que no son ácido nucleico. Un ensayo espacial integrado *in situ* descrito en la presente descripción puede usarse para analizar analitos de ácido nucleico y/o analitos que no son ácido nucleico en cualquier combinación adecuada. En algunos aspectos, uno o más módulos de ensayo *in situ* de un método descrito en la presente descripción pueden analizar una primera pluralidad de analitos de ácido nucleico en una muestra, y uno o más módulos de ensayo espacial del método pueden analizar una segunda pluralidad de analitos de ácido nucleico en la muestra, en donde la segunda pluralidad es opcionalmente un conjunto más grande de analitos de ácido nucleico que la primera pluralidad. En cualquiera de las modalidades en la presente descripción, uno o más módulos de ensayo *in situ* y/o ensayo espacial de un método descrito en la presente descripción pueden analizar una pluralidad de analitos que no son ácido nucleico, tales como analitos de proteínas.

Los ejemplos de analitos que no son ácido nucleico incluyen, pero no se limitan a, lípidos, carbohidratos, péptidos, proteínas, glucoproteínas (unidas a N o unidas a O), lipoproteínas, fosfoproteínas, variantes fosforiladas o acetiladas de proteínas específicas, 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 de recubrimiento viral, proteínas extracelulares e intracelulares, anticuerpos, y fragmentos de unión a antígeno. En algunas modalidades, el analito está dentro de una célula o sobre una superficie celular, tal como un analito transmembrana o uno que se une a la membrana celular. En algunas modalidades, el analito puede ser un orgánulo (por ejemplo, núcleos o mitocondrias). En algunas modalidades, el analito es un analito extracelular, tal como un analito secretado.

Los elementos de la superficie celular que se corresponden con los analitos pueden incluir, pero no se limitan a, un receptor, un antígeno, una proteína de superficie, una proteína transmembrana, un conglomerado de proteínas de diferenciación, una proteína canal, una proteína bomba, una proteína portadora, un fosfolípido, una glucoproteína, un glucolípido, un complejo de proteína de interacción célula-célula, un complejo presentador de antígenos, un complejo mayor de histocompatibilidad, un receptor de linfocitos T manipulado genéticamente, un receptor de linfocitos T, un receptor de linfocitos B, un receptor de antígeno quimérico, una proteína de la matriz extracelular, una modificación postraducciona (por ejemplo, fosforilación, glucosilación, ubiquitinación, nitrosilación, metilación, acetilación o lipidación) el estado de una proteína de la superficie celular, una unión en hendidura, y una unión adherente.

Los ejemplos de analitos de ácido nucleico incluyen analitos de ADN tales como ADN genómico, ADN metilado, secuencias específicas de ADN metilado, ADN fragmentado, ADN mitocondrial, productos de PCR sintetizados *in situ* e híbridos ADN/ARN.

Los ejemplos de analitos de ácido nucleico incluyen además analitos de ARN tales como diversos tipos de ARN codificante y no codificante. Ejemplos de ARN no codificantes (ARNnc) que no se traducen en una proteína incluyen ARN de transferencia (ARNt) y ARN ribosomal (ARNr), así como también ARN pequeños no codificantes tales como microARN (miARN), ARN de interferencia pequeño (ARNip), ARN de interacción con Piwi (ARNpi), ARN pequeño nucleolar (ARNpno), ARN pequeño nuclear (ARNpn), ARN extracelular (ARNex), ARN pequeño específico del cuerpo de Cajal (ARNpcc), y los ARNnc largos tales como Xist y HOTAIR. Ejemplos de los diferentes tipos de analitos de ARN incluyen ARN mensajero (ARNm), que incluye un ARN naciente, un pre-ARNm, un ARN de transcrito primario y un ARN procesado, tal como un ARNm con caperuza (por ejemplo, con una caperuza de 7-metil guanosina 5'), un ARNm poliadenilado (cola de poli-A en el extremo 3'), y un ARNm empalmado en el que se han eliminado uno o más intrones. También se incluyen en los analitos descritos en la presente descripción ARNm sin caperuza, un ARNm no poliadenilado, y un ARNm no empalmado. Los ejemplos adicionales de analitos de ARN incluyen ARNr, ARNt, miARN,

y ARN viral. El ARN puede ser un transcrito (por ejemplo, presente en una sección de tejido). El ARN puede ser pequeño (por ejemplo, de longitud menor que 200 bases de ácido nucleico) o grande (por ejemplo, ARN de longitud mayor que 200 bases de ácido nucleico). Ejemplos de ARN pequeños incluyen ARN ribosomal (ARNr) 5.8S, ARNr 5S, ARNt, miARN, ARNip, ARNpno, ARNpn, ARNpi, ARN pequeño derivado de ARNt (ARNpt), y ARN pequeño derivado de ADNr (ARNpr). El ARN puede ser ARN bicatenario o ARN monocatenario. El ARN puede ser ARN circular. El ARN puede ser un ARNr bacteriano (por ejemplo, ARNr 16s o ARNr 23s).

En determinadas modalidades, un analito puede extraerse a partir de una célula viva. Las condiciones de procesamiento pueden ajustarse para asegurar que una muestra biológica permanezca viable durante el análisis, y que los analitos se extraigan (o se liberen) de células vivas de la muestra. Los analitos derivados de células vivas pueden obtenerse solo una vez a partir de la muestra, o pueden obtenerse a intervalos a partir de una muestra que permanece en una condición viable.

En general, los sistemas, aparatos, métodos y composiciones pueden usarse para analizar cualquier número de analitos. Por ejemplo, la cantidad de analitos que se analizan puede ser al menos aproximadamente 2, al menos aproximadamente 3, al menos aproximadamente 4, al menos aproximadamente 5, al menos aproximadamente 6, al menos aproximadamente 7, al menos aproximadamente 8, al menos aproximadamente 9, al menos aproximadamente 10, al menos aproximadamente 11, al menos aproximadamente 12, al menos aproximadamente 13, al menos aproximadamente 14, al menos aproximadamente 15, al menos aproximadamente 20, al menos aproximadamente 25, al menos aproximadamente 30, al menos aproximadamente 40, al menos aproximadamente 50, al menos aproximadamente 100, al menos aproximadamente 1000, al menos aproximadamente 10 000, al menos aproximadamente 100 000 o más analitos diferentes presentes en una región de la muestra o dentro de un elemento individual del sustrato.

III. Muestras y procesamiento de muestras

A. Muestras biológicas

En algunas modalidades, en la presente descripción se describen los métodos y composiciones para analizar una muestra biológica, que puede obtenerse de un sujeto para analizar mediante el uso de cualquiera de una variedad de técnicas, que incluye, pero no se limita a, biopsia, cirugía y microscopía de captura por láser (LCM), e incluye generalmente células y/u otro material biológico del sujeto. Además de los sujetos descritos anteriormente, una muestra biológica puede obtenerse, además, a partir de un procariota tal como una bacteria, por ejemplo, *Escherichia coli*, *Staphylococci* o *Mycoplasma pneumoniae*; una archaea; un virus tal como el virus de la hepatitis C o el virus de la inmunodeficiencia humana; o un viroide. Una muestra biológica puede obtenerse a partir de organismos no mamíferos (por ejemplo, una planta, un insecto, un artrópodo, un nemátodo, un hongo, o un anfibio). Además, puede obtenerse una muestra biológica a partir de un eucariota, tal como un organoide derivado del paciente (PDO) o un xenoinjerto derivado del paciente (PDX). Los sujetos a partir de los cuales pueden obtenerse muestras biológicas pueden ser individuos sanos o asintomáticos, individuos que tienen o se sospecha que tienen una enfermedad (por ejemplo, un paciente con una enfermedad tal como cáncer) o una predisposición a una enfermedad, y/o individuos que necesitan terapia o se sospecha que necesitan terapia.

La muestra biológica puede incluir cualquier número de macromoléculas, por ejemplo, macromoléculas celulares y de orgánulos (por ejemplo, mitocondrias y núcleos). La muestra biológica puede ser una muestra de ácido nucleico y/o una muestra de proteína. La muestra biológica puede ser una muestra de carbohidratos o una muestra de lípidos. La muestra biológica puede obtenerse como una muestra de tejido, tal como una sección de tejido, biopsia, una biopsia con aguja gruesa, aspirado con aguja, o aspirado con aguja fina. La muestra puede ser una muestra de fluido, tal como una muestra de sangre, una muestra de orina, o una muestra de saliva. La muestra puede ser una muestra de piel, una muestra de colon, un raspado bucal, una muestra de histología, una muestra de histopatología, una muestra de plasma o suero, una muestra de tumor, células vivas, células cultivadas, una muestra clínica tal como, por ejemplo, sangre completa o productos derivados de la sangre, células sanguíneas, o tejidos o células cultivados, que incluye suspensiones celulares. En algunas modalidades, la muestra biológica puede comprender células que se depositan sobre una superficie.

Las muestras biológicas libres de células pueden incluir polinucleótidos extracelulares. Los polinucleótidos extracelulares pueden aislarse a partir de una muestra corporal, por ejemplo, sangre, plasma, suero, orina, saliva, excreciones mucosales, esputo, heces y lágrimas.

Las muestras biológicas pueden derivarse a partir de un cultivo homogéneo o población de los sujetos u organismos mencionados en la presente descripción, o alternativamente, de una colección de varios organismos diferentes, por ejemplo, en una comunidad o ecosistema.

Las muestras biológicas pueden incluir una o más células enfermas. Una célula enferma puede tener propiedades metabólicas, expresión génica, expresión proteica y/o elementos morfológicos alterados. Los ejemplos de enfermedades incluyen trastornos inflamatorios, trastornos metabólicos, trastornos del sistema nervioso y cáncer. Las

células cancerosas pueden derivarse de tumores sólidos, neoplasias hematológicas, líneas celulares u obtenerse como células tumorales circulantes.

Las muestras biológicas también pueden incluir células fetales. Por ejemplo, puede realizarse un procedimiento tal como la amniocentesis para obtener una muestra de células fetales de la circulación materna. La secuenciación de células fetales puede usarse para identificar cualquiera de una serie de trastornos genéticos, que incluyen, por ejemplo, aneuploidía tal como síndrome de Down, síndrome de Edwards y síndrome de Patau. Además, los elementos de la superficie celular de las células fetales pueden usarse para identificar cualquiera de una serie de trastornos o enfermedades.

Las muestras biológicas también pueden incluir células inmunitarias. El análisis de secuencias del repertorio inmunitario de tales células, que incluye elementos genómicos, proteómicos y de superficie celular, puede proporcionar una gran cantidad de información para facilitar la comprensión del estado y función del sistema inmunitario. A manera de ejemplo, determinar el estado (por ejemplo, negativo o positivo) de la enfermedad de residuos mínimos (EMR) en un paciente con mieloma múltiple (MM) después de un trasplante autólogo de células madre se considera un factor predictivo de EMR en el paciente con MM (véase, por ejemplo, la publicación de solicitud de patente de Estados Unidos núm. 2018/0156784).

Los ejemplos de células inmunitarias en una muestra biológica incluyen, pero no se limitan a, linfocitos B, linfocitos T (por ejemplo, linfocitos T citotóxicos, linfocitos T asesinos naturales, linfocitos T reguladores y linfocitos T colaboradores), células asesinas naturales, células asesinas inducidas por citocinas (CIK), células mieloides, tales como granulocitos (granulocitos basófilos, granulocitos eosinófilos, granulocitos neutrófilos/neutrófilos hipersegmentados), monocitos/macrófagos, mastocitos, trombocitos/megacariocitos y células dendríticas.

Como se discutió anteriormente, una muestra biológica puede incluir un único analito de interés, o más de un analito de interés. Se proporcionan métodos para realizar ensayos multiplexados para analizar dos o más analitos diferentes en una única muestra biológica.

B. Procesamiento de muestras biológicas

Puede realizarse una variedad de etapas para preparar o procesar una muestra biológica para y/o durante un análisis. Excepto cuando se indique de cualquier otra manera, las etapas de preparación o procesamiento descritas más abajo pueden combinarse generalmente de cualquier manera y en cualquier orden para preparar o procesar apropiadamente una muestra en particular para su análisis.

En algunas modalidades, la muestra puede tratarse con uno o más reactivos útiles para el ensayo *in situ* y/o el ensayo espacial. En algunas modalidades, la muestra puede tratarse con una o más enzimas. Por ejemplo, pueden añadirse una o más endonucleasas para fragmentar ADN, enzimas ADN polimerasa y dNTP usados para amplificar ácidos nucleicos. Otras enzimas que también pueden añadirse a la muestra incluyen, pero no se limitan a, polimerasa, transposasa, ligasa y ADNsa y ARNasa. En algunas modalidades, pueden añadirse enzimas transcriptasa inversa a la muestra, que incluyen enzimas con actividad de transferasa terminal, cebadores y oligonucleótidos que cambian de molde. El cambio de molde puede usarse para aumentar la longitud de un ADNc, por ejemplo, mediante la adición de una secuencia de ácido nucleico predefinida al ADNc. En algunas modalidades, el ADNc puede analizarse durante el ensayo espacial.

i. Corte en secciones de tejidos

Una muestra biológica puede recolectarse de un sujeto (por ejemplo, mediante biopsia quirúrgica, corte en secciones de sujeto completo) o cultivarse *in vitro* sobre un sustrato de crecimiento o placa de cultivo como una población de células, y prepararse para su análisis como un corte de tejido o sección de tejido. Las muestras cultivadas pueden ser suficientemente finas para el análisis sin etapas de procesamiento adicionales. Alternativamente, las muestras cultivadas, y las muestras que se obtienen mediante biopsia o cortes en secciones, pueden prepararse como secciones de tejido finas mediante el uso de un aparato de corte mecánico tal como un micrótopo de cuchillas vibratorias. Como otra alternativa, en algunas modalidades, puede prepararse una sección de tejido fina mediante la aplicación de una impresión táctil de una muestra biológica a un material de sustrato adecuado.

El grosor de la sección de tejido puede ser una fracción de (por ejemplo, menos de 0,9, 0,8, 0,7, 0,6, 0,5, 0,4, 0,3, 0,2, o 0,1) la dimensión de la sección transversal máxima de una célula. Sin embargo, también pueden usarse secciones de tejido que tengan un grosor que es mayor que la dimensión de la sección transversal máxima de la célula. Por ejemplo, pueden usarse secciones de criostato, que pueden tener, por ejemplo, 10-20 micrómetros de grosor.

Más generalmente, el grosor de una sección de tejido depende típicamente del método usado para preparar la sección y de las características físicas del tejido, y por lo tanto pueden prepararse y usarse secciones que tengan una amplia variedad de grosores diferentes. Por ejemplo, el grosor de la sección de tejido puede ser al menos 0,1, 0,2, 0,3, 0,4, 0,5, 0,7, 1,0, 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 20, 30, 40 o 50 micrómetros. Además, pueden usarse secciones más gruesas si se desea o es conveniente, por ejemplo, al menos 70, 80, 90, o 100 micrómetros o más.

Típicamente, el grosor de una sección de tejido está entre 1-100 micrómetros, 1-50 micrómetros, 1-30 micrómetros, 1-25 micrómetros, 1-20 micrómetros, 1-15 micrómetros, 1-10 micrómetros, 2-8 micrómetros, 3-7 micrómetros o 4-6 micrómetros, pero como se mencionó anteriormente, también pueden analizarse secciones con grosores mayores o menores que estos intervalos.

También pueden obtenerse varias secciones a partir de una única muestra biológica. Por ejemplo, pueden obtenerse varias secciones de tejido a partir de una muestra de biopsia quirúrgica mediante la realización de cortes en secciones seriados de la muestra de biopsia mediante el uso de una cuchilla para corte en secciones. La información espacial entre las secciones seriadas puede conservarse de esta manera, y las secciones pueden analizarse sucesivamente para obtener información tridimensional sobre la muestra biológica.

ii. Congelación

En algunas modalidades, la muestra biológica (por ejemplo, una sección de tejido como se describió anteriormente) puede prepararse mediante ultracongelación a una temperatura adecuada para mantener o preservar la integridad (por ejemplo, las características físicas) de la estructura del tejido. Tal temperatura puede ser, por ejemplo, menor que -20 °C, o menor que -25 °C, -30 °C, -40 °C, -50 °C, -60 °C, -70 °C, -80 °C, -90 °C, -100 °C, -110 °C, -120 °C, -130 °C, -140 °C, -150 °C, -160 °C, -170 °C, -180 °C, -190 °C o -200 °C. La muestra de tejido congelado puede cortarse en secciones, por ejemplo, cortarse finamente, sobre una superficie de sustrato mediante el uso de cualquier número de métodos adecuados. Por ejemplo, una muestra de tejido puede prepararse mediante el uso de un micrótopo de congelación (por ejemplo, un criostato) configurado a una temperatura adecuada para mantener tanto la integridad estructural de la muestra de tejido como las propiedades químicas de los ácidos nucleicos en la muestra. Dicha temperatura puede ser, por ejemplo, menor que -15 °C, menor que -20 °C, o menor que -25 °C.

iii. Fijación con formalina e inclusión en parafina

En algunas modalidades, la muestra biológica puede prepararse mediante el uso de fijación con formalina e inclusión en parafina (FFPE), que son métodos establecidos. En algunas modalidades, las suspensiones celulares y otras muestras no tisulares pueden prepararse mediante el uso de fijación con formalina e inclusión en parafina. Después de la fijación de la muestra y su inclusión en un bloque de parafina o resina, la muestra puede cortarse en secciones como se describió anteriormente. Antes del análisis, el material que se incluye en parafina puede eliminarse de la sección de tejido (por ejemplo, desparafinación) mediante la incubación de la sección de tejido en un solvente apropiado (por ejemplo, xileno) seguido de un enjuague (por ejemplo, etanol al 99,5 % durante 2 minutos, etanol al 96 % durante 2 minutos, y etanol al 70 % durante 2 minutos).

iv. Fijación

Como una alternativa a la fijación con formalina descrita anteriormente, una muestra biológica puede fijarse en cualquiera de una variedad de otros fijadores para conservar la estructura biológica de la muestra previo al análisis. Por ejemplo, una muestra puede fijarse mediante inmersión en etanol, metanol, acetona, paraformaldehído (PFA)-triton, y sus combinaciones.

En algunas modalidades, la fijación en acetona se usa con muestras congeladas frescas, que pueden incluir, pero no se limitan a, muestras de tejido de corteza, bulbo olfatorio de ratón, tumor cerebral humano, cerebro humano post mortem, y cáncer de mama. Cuando se realiza la fijación en acetona, pueden no realizarse las etapas de pre-permeabilización (descritas más abajo). Alternativamente, la fijación en acetona puede realizarse junto con las etapas de permeabilización.

En algunas modalidades, los métodos que se proporcionan en la presente descripción comprenden una o más etapas después de la fijación (también denominadas posfijación). En algunas modalidades, se realizan una o más etapas de posfijación después de poner en contacto una muestra con un polinucleótido descrito en la presente descripción, por ejemplo, una o más sondas tales como una sonda circular o candado. En algunas modalidades, se realizan una o más etapas de posfijación después de que se forma un complejo de hibridación que comprende una sonda y una diana en una muestra. En algunas modalidades, se realizan una o más etapas de posfijación previo a una reacción de ligación descrita en la presente descripción, tal como la ligación para circularizar una sonda candado.

En algunas modalidades, se realizan una o más etapas de posfijación después de poner en contacto una muestra con un agente de unión o marcaje (por ejemplo, un anticuerpo o fragmento de unión a antígeno de este) para un analito que no es un ácido nucleico tal como un analito de proteína. El agente de marcaje puede comprender una molécula de ácido nucleico (por ejemplo, oligonucleótido reportero) que comprende una secuencia que corresponde al agente de marcaje y por lo tanto corresponde al analito (por ejemplo, identifica de manera única). En algunas modalidades, el agente de marcaje puede comprender un oligonucleótido reportero que comprende una o más secuencias de código de barras. En algunas modalidades, la muestra se pone en contacto con un agente de marcaje durante el *ensayo in situ* y/o durante el ensayo espacial.

Puede realizarse una etapa de posfijación mediante el uso de cualquier reactivo de fijación adecuado descrito en la presente descripción, por ejemplo, paraformaldehído al 3 % (p/v) en DEPC-PBS.

v. Inclusión

Como una alternativa a la inclusión en parafina descrita anteriormente, una muestra biológica puede incluirse en cualquiera de una variedad de otros materiales de inclusión para proporcionar un sustrato estructural a la muestra previo al corte en secciones y a otras etapas de manipulación. En general, el material de inclusión se elimina previo al análisis de las secciones de tejido que se obtienen de la muestra. Los materiales de inclusión adecuados incluyen, pero no se limitan a, ceras, resinas (por ejemplo, resinas de metacrilato), epoxis, y agar.

En algunas modalidades, la muestra biológica puede incluirse en una matriz de hidrogel. La inclusión de la muestra de esta manera implica típicamente poner en contacto la muestra biológica con un hidrogel de manera que la muestra biológica queda rodeada por el hidrogel. Por ejemplo, la muestra puede incluirse al poner en contacto la muestra con un material polimérico adecuado, y activar el material polimérico para formar un hidrogel. En algunas modalidades, el hidrogel se forma de manera que el hidrogel se internaliza dentro de la muestra biológica.

En algunas modalidades, la muestra biológica se inmoviliza en el hidrogel mediante la reticulación del material polimérico que forma el hidrogel. La reticulación puede realizarse química y/o fotoquímicamente, o alternativamente, mediante cualquier otro método de formación de hidrogel conocido en la técnica.

La composición y aplicación de la matriz de hidrogel a una muestra biológica depende típicamente de la naturaleza y la preparación de la muestra biológica (por ejemplo, seccionada, no seccionada, tipo de fijación). Como un ejemplo, cuando la muestra biológica es una sección de tejido, la matriz de hidrogel puede incluir una solución de monómero y una solución de iniciador de persulfato de amonio (APS)/acelerador de tetrametiletildiamina (TEMED). Como otro ejemplo, cuando la muestra biológica consiste en células (por ejemplo, células cultivadas o células disociadas a partir de una muestra de tejido), las células pueden incubarse con la solución de monómero y las soluciones APS/TEMED. Para las células, los geles de matriz de hidrogel se forman en compartimientos, lo que incluye, pero no se limita a, dispositivos usados para cultivar, mantener, o transportar las células. Por ejemplo, las matrices de hidrogel pueden formarse con solución de monómero más APS/TEMED que se añade al compartimiento a una profundidad que varía de aproximadamente 0,1 μm a aproximadamente 2 mm.

Métodos y aspectos adicionales de la inclusión en hidrogel de muestras biológicas se describen, por ejemplo, en Chen y otros, *Science* 347(6221):543–548, 2015.

vi. Tinción

Para facilitar la visualización, las muestras biológicas pueden teñirse mediante el uso de una amplia variedad de tinciones y técnicas de tinción. En algunas modalidades, por ejemplo, una muestra puede teñirse mediante el uso de cualquier número de tinciones, que incluyen, pero no se limitan a, naranja acridina, marrón Bismarck, carmín, azul coomassie, violeta cresílica, DAPI, eosina, bromuro de etidio, fucsina ácida, hematoxilina, tinciones Hoechst, yodo, verde metilo, azul de metileno, rojo neutro, azul Nilo, rojo Nilo, tetróxido de osmio, yoduro de propidio, rodamina o safranina.

La muestra puede teñirse mediante el uso de técnicas de tinción con hematoxilina y eosina (H&E), mediante el uso de técnicas de tinción de Papanicolaou, técnicas de tinción tricrómica de Masson, técnicas de tinción con plata, técnicas de tinción con Sudán, y/o mediante el uso de técnicas de tinción con ácido peryódico de Schiff (PAS). La tinción con PAS se realiza típicamente después de la fijación en formalina o acetona. En algunas modalidades, la muestra puede teñirse mediante el uso de la tinción de Romanowsky, lo que incluye la tinción de Wright, la tinción de Jenner, la tinción de May-Grunwald, la tinción de Leishman, y la tinción de Giemsa.

En algunas modalidades, las muestras biológicas pueden destañarse. Los métodos para destañar o decolorar una muestra biológica se conocen en la técnica, y generalmente dependen de la naturaleza de la(s) tinción(ones) aplicada(s) a la muestra. Por ejemplo, en algunas modalidades, se aplican una o más tinciones inmunofluorescentes a la muestra mediante el acoplamiento de anticuerpos. Tales tinciones pueden eliminarse mediante el uso de técnicas tales como la escisión de enlaces disulfuro mediante el tratamiento con un agente reductor y el lavado con detergente, tratamiento con sal caotrópica, tratamiento con solución de recuperación de antígenos, y tratamiento con un tampón ácido de glicina. Los métodos para teñir y destañar multiplexados se describen, por ejemplo, en Bolognesi y otros, *J. Histochem. Cytochem.* 2017; 65(8): 431-444, Lin y otros, *Nat Commun.* 2015; 6:8390, Pirici y otros, *J. Histochem. Cytochem.* 2009; 57:567–75, y Glass y otros, *J. Histochem. Cytochem.* 2009; 57:899–905.

vii. Expansión isométrica

En algunas modalidades, una muestra biológica que se incluye en un hidrogel puede expandirse isométricamente. Los métodos de expansión isométrica que pueden usarse incluyen la hidratación, una etapa preparativa en la microscopía de expansión, como se describió en Chen y otros, *Science* 347(6221):543–548, 2015.

La expansión isométrica puede realizarse mediante el anclaje de uno o más componentes de una muestra biológica a un gel, seguido de la formación de gel, proteólisis, e hinchazón. La expansión isométrica de la muestra biológica puede producirse previo a la inmovilización de la muestra biológica sobre un sustrato, o después de que la muestra biológica se inmovilice en un sustrato. En algunas modalidades, la muestra biológica expandida isométricamente puede eliminarse del sustrato previo a ponerse en contacto el sustrato con las sondas de captura descritas en la presente descripción.

En general, las etapas usadas para realizar la expansión isométrica de la muestra biológica pueden depender de las características de la muestra (por ejemplo, grosor de la sección de tejido, la fijación, la reticulación), y/o del analito de interés (por ejemplo, condiciones diferentes para anclar el ARN, ADN y proteína a un gel).

En algunas modalidades, las proteínas en la muestra biológica se anclan a un gel hinchable tal como un gel de polielectrolitos. Un anticuerpo puede dirigirse a la proteína antes, después, o junto con el anclaje al gel hinchable. El ADN y/o ARN en una muestra biológica también pueden anclarse al gel hinchable mediante un enlazador adecuado. Ejemplos de tales enlazadores incluyen, pero no se limitan a, ácido 6-((acrilóil)amino)hexanoico (Acrilóil-X SE) (disponible en ThermoFisher, Waltham, MA), Label-IT Amine (disponible en MirusBio, Madison, WI) y Label X (descrito, por ejemplo, en Chen y otros, *Nat. Methods* 13:679-684, 2016).

La expansión isométrica de la muestra puede aumentar la resolución espacial del análisis subsecuente de la muestra. El aumento de la resolución en la caracterización espacial puede determinarse mediante la comparación de una muestra expandida isométricamente con una muestra que no se ha expandido isométricamente.

En algunas modalidades, una muestra biológica se expande isométricamente hasta un tamaño al menos 2x, 2,1x, 2,2x, 2,3x, 2,4x, 2,5x, 2,6x, 2,7x, 2,8x, 2,9x, 3x, 3,1x, 3,2x, 3,3x, 3,4x, 3,5x, 3,6x, 3,7x, 3,8x, 3,9x, 4x, 4,1x, 4,2x, 4,3x, 4,4x, 4,5x, 4,6x, 4,7x, 4,8x o 4,9x de su tamaño no expandido. En algunas modalidades, la muestra se expande isométricamente hasta al menos 2x y menos de 20x de su tamaño no expandido.

viii. Reticulaciones

En algunas modalidades, la muestra biológica se reticula de manera reversible previo a, o durante un módulo de ensayo *in situ*. En algunos aspectos, los polinucleótidos y/o el producto de la amplificación (por ejemplo, amplicón) de un analito o una sonda que se une a este pueden anclarse a una matriz polimérica. Por ejemplo, la matriz polimérica puede ser un hidrogel. En algunas modalidades, una o más de la sonda(s) de polinucleótidos y/o el producto de la amplificación (por ejemplo, amplicón) de estas pueden modificarse para contener grupos funcionales que pueden usarse como un sitio de anclaje para unir las sondas de polinucleótidos y/o el producto de la amplificación a una matriz polimérica. En algunas modalidades, puede usarse una sonda modificada que comprende oligo dT para unirse a las moléculas de ARNm de interés, seguido de la reticulación reversible de las moléculas de ARNm.

En algunas modalidades, las sondas candado con código de barras de dirigidas o aleatorias pueden usarse para unirse a una molécula de ácido nucleico de interés (por ejemplo, ADN o ARN tal como transcritos de ARNm) seguido de amplificación (por ejemplo, amplificación por círculo rodante) con nucleótidos modificados para bloquear el amplicón en su lugar. En algunos aspectos, durante la etapa de amplificación, pueden añadirse nucleótidos modificados a la reacción para incorporar los nucleótidos modificados en el producto de la amplificación. Los nucleótidos modificados ilustrativos comprenden nucleótidos modificados con amina. En algunas modalidades, el nucleótido modificado con amina comprende una modificación con una porción de N-hidroxisuccinimida de ácido acrílico. Los ejemplos de otros nucleótidos modificados con amina comprenden, pero no se limitan a, una modificación con la porción 5-aminoalil-dUTP, una modificación con la porción 5-propargilamino-dCTP, una modificación con la porción N6-6-aminohexil-dATP, o una modificación con la porción 7-deaza-7-propargilamino-dATP.

En algunas modalidades, la muestra biológica se inmoviliza en un hidrogel mediante la reticulación del material polimérico que forma el hidrogel. La reticulación puede realizarse química y/o fotoquímicamente, o alternativamente, mediante cualquier otro método de formación de hidrogel conocido en la técnica.

Cuando el sustrato incluye un gel (por ejemplo, un hidrogel o matriz de gel), los oligonucleótidos dentro del gel pueden unirse al sustrato. Los términos "hidrogel" y "matriz de hidrogel" se usan indistintamente en la presente descripción para referirse a un gel de polímero macromolecular que incluye una red. Dentro de la red, algunas cadenas de polímeros pueden reticularse opcionalmente, aunque no siempre se produce la reticulación.

En algunas modalidades, un hidrogel puede incluir subunidades de hidrogel. Una "subunidad de hidrogel" es un monómero hidrófilo, un precursor molecular, o un polímero que puede polimerizarse (por ejemplo, reticularse) para formar una red de hidrogel tridimensional (3D). Las subunidades de hidrogel pueden incluir cualquiera de las subunidades de hidrogel convenientes, tales como, pero no limitadas a, acrilamida, bis-acrilamida, poliacrilamida y derivados de estas, poli(etilenglicol) y derivados de este (por ejemplo, PEG-acrilato (PEG-DA), PEG-RGD), gelatina-metacrilóilo (GelMA), ácido hialurónico metacrilado (MeHA), poliuretanos polialifáticos, poliuretanos de poliéter, poliuretanos de poliéster, copolímeros de polietileno, poliamidas, alcoholes polivinílicos, polipropilenglicol, óxido de politetrametileno, polivinilpirrolidona, poliacrilamida, poli(acrilato de hidroxietilo) y poli(metacrilato de hidroxietilo),

colágeno, ácido hialurónico, quitosano, dextrano, agarosa, gelatina, alginato, polímeros de proteína, metilcelulosa, y similares, y sus combinaciones.

5 En algunas modalidades, un hidrogel incluye un material híbrido, por ejemplo, el material de hidrogel incluye elementos de polímeros tanto sintéticos como naturales. Los ejemplos de hidrogeles adecuados se describen, por ejemplo, en las patentes de Estados Unidos núms. 6,391,937, 9,512,422, y 9,889,422, y en las publicaciones de solicitud de patentes de Estados Unidos núms. 2017/0253918, 2018/0052081 y 2010/0055733.

10 En algunas modalidades, se añaden formadores de reticulaciones y/o iniciadores a las subunidades de hidrogel. Los ejemplos de formadores de reticulaciones incluyen, sin limitación, bis-acrilamida y diazirina. Los ejemplos de iniciadores incluyen, sin limitación, azobisisobutironitrilo (AIBN), riboflavina y L-arginina. La inclusión de formadores de reticulaciones y/o iniciadores puede conducir a una unión covalente aumentada entre macromoléculas biológicas que interactúan en etapas de polimerización posteriores.

15 En algunas modalidades, los hidrogeles pueden tener una estructura coloidal, tal como agarosa, o una estructura de malla polimérica, tal como gelatina.

20 En algunas modalidades, algunas subunidades de hidrogel se polimerizan (por ejemplo, experimentan "formación"), forman reticulaciones covalente o físicamente, para formar una red de hidrogel. Por ejemplo, las subunidades de hidrogel pueden polimerizarse mediante cualquier método que incluye, pero no se limita a, formación de reticulaciones térmica, formación de reticulaciones química, formación de reticulaciones física, formación de reticulaciones iónica, formación de reticulaciones inducidas por luz, formación de reticulaciones irradiativa (por ejemplo, rayos X, haz de electrones) y sus combinaciones. Las técnicas tales como la fotopolimerización litográfica también pueden usarse para formar hidrogeles.

25 Los métodos de polimerización para las subunidades de hidrogel pueden seleccionarse para formar hidrogeles con diferentes propiedades (por ejemplo, tamaño de los poros, propiedades de hinchazón, biodegradabilidad, conducción, transparencia y/o permeabilidad del hidrogel). Por ejemplo, un hidrogel puede incluir poros de tamaño suficiente para permitir el paso de macromoléculas (por ejemplo, ácidos nucleicos, proteínas, cromatina, metabolitos, ARNg, anticuerpos, carbohidratos, péptidos, metabolitos y/o moléculas pequeñas) hacia la muestra (por ejemplo, sección de tejido). Se conoce que el tamaño de los poros generalmente disminuye con el aumento de la concentración de las subunidades de hidrogel y generalmente aumenta con una relación creciente de las subunidades de hidrogel con respecto al formador de reticulaciones. Por lo tanto, puede prepararse una composición fijadora/de hidrogel que incluye una concentración de subunidades de hidrogel que permite el paso de tales macromoléculas biológicas.

35 En algunas modalidades, el hidrogel puede formar el sustrato. En algunas modalidades, el sustrato incluye un hidrogel y uno o más segundos materiales. En algunas modalidades, el hidrogel se coloca sobre uno o más segundos materiales. Por ejemplo, el hidrogel puede formarse previamente y después colocarse sobre, debajo, o en cualquier otra configuración con uno o más segundos materiales. En algunas modalidades, la formación del hidrogel se produce después de entrar en contacto con uno o más segundos materiales durante la formación del sustrato. La formación del hidrogel también puede producirse dentro de una estructura (por ejemplo, pocillos, crestas, proyecciones, y/o marcas) ubicada sobre un sustrato.

40 En algunas modalidades, la formación del hidrogel sobre un sustrato se produce antes, simultáneamente con, o después de que los elementos se unen al sustrato. Por ejemplo, cuando se une una sonda de captura (por ejemplo, directa o indirectamente) a un sustrato, la formación de hidrogel puede realizarse en el sustrato que ya contiene las sondas de captura.

45 En algunas modalidades, la formación del hidrogel se produce dentro de una muestra biológica. En algunas modalidades, una muestra biológica (por ejemplo, sección de tejido) se incluye en un hidrogel. En algunas modalidades, las subunidades de hidrogel se difunden en la muestra biológica, y la polimerización del hidrogel se inicia mediante un estímulo externo o interno.

50 En modalidades en las que se forma un hidrogel dentro de una muestra biológica, puede usarse una química de funcionalización. En algunas modalidades, la química de funcionalización incluye la química hidrogel-tejido (HTC). Puede usarse cualquier cadena principal de hidrogel-tejido (por ejemplo, sintética o natural) adecuada para la HTC para anclar macromoléculas biológicas y modular la funcionalización. Ejemplos no limitantes de métodos que usan variantes de la cadena principal de HTC incluyen CLARITY, PACT, ExM, SWITCH y ePACT. En algunas modalidades, la formación del hidrogel dentro de una muestra biológica es permanente. Por ejemplo, las macromoléculas biológicas pueden adherirse permanentemente al hidrogel permitiendo varias rondas de pesquiasaje. En algunas modalidades, la formación del hidrogel dentro de una muestra biológica es reversible.

55 En algunas modalidades, se añaden reactivos adicionales a las subunidades de hidrogel antes, simultáneamente con, y/o después de la polimerización. Por ejemplo, los reactivos adicionales pueden incluir, pero no se limitan a, oligonucleótidos (por ejemplo, sondas de captura), endonucleasas para fragmentar ADN, tampón de fragmentación para ADN, enzimas ADN polimerasa, dNTP usados para amplificar el ácido nucleico y unir el código de barras a los

fragmentos amplificados. Pueden usarse otras enzimas, lo que incluye, sin limitación, ARN polimerasa, transposasa, ligasa, proteinasa K, y ADNasa. Los reactivos adicionales pueden incluir, además, enzimas transcriptasas inversas, lo que incluye enzimas con actividad transferasa terminal, cebadores y oligonucleótidos que cambian de molde. En algunas modalidades, se añaden marcadores ópticos a las subunidades de hidrogel antes, simultáneamente con, y/o después de la polimerización.

En algunas modalidades, los reactivos para HTC se añaden al hidrogel antes, simultáneamente con, y/o después de la polimerización. En algunas modalidades, se añade un agente de marcaje de células al hidrogel antes, simultáneamente con, y/o después de la polimerización. En algunas modalidades, se añade un agente penetrante de células al hidrogel antes, simultáneamente con, y/o después de la polimerización.

Los hidrogeles que se incluyen dentro de las muestras biológicas pueden clarificarse mediante el uso de cualquier método adecuado. Por ejemplo, pueden usarse métodos electroforéticos para clarificar tejidos para eliminar macromoléculas biológicas de la muestra que se incluye en hidrogel. En algunas modalidades, una muestra que se incluye en hidrogel se almacena, antes o después de clarificar el hidrogel, en un medio (por ejemplo, un medio de montaje, metilcelulosa, u otros medios semisólidos).

Un “recubrimiento extraíble condicionalmente” es un recubrimiento que puede eliminarse de la superficie de un sustrato tras la aplicación de un agente de liberación. En algunas modalidades, un recubrimiento extraíble condicionalmente incluye un hidrogel como se describe en la presente descripción, por ejemplo, un hidrogel que incluye un material basado en polipéptidos. Los ejemplos no limitantes de un hidrogel que presenta un material basado en polipéptidos incluyen un material sintético basado en péptidos que presenta una combinación de seda araña y un segmento transmembrana del canal de calcio de tipo L del músculo humano (por ejemplo, PEPGEL®), un péptido anfifílico de 16 residuos que contiene una secuencia repetitiva de arginina-alanina-aspartato-alanina (RADARADARADARADA) (por ejemplo, PURAMATRIX®), EAK16 (AEAEAKAKAEAEAKAK), KLD12 (KLDLKLKLDL) y PGMATRIX™.

En algunas modalidades, el hidrogel en el recubrimiento extraíble condicionalmente es un hidrogel que responde al estímulo. Un hidrogel que responde al estímulo puede someterse a una transición de gel a solución y/o gel a sólido tras la aplicación de uno o más desencadenantes externos (por ejemplo, un agente de liberación). Véase, por ejemplo, Willner, *Acc. Chem. Res.* 50:657–658, 2017. Los ejemplos no limitantes de un hidrogel que responde al estímulo incluyen un hidrogel que responde al calor, un hidrogel que responde al pH, un hidrogel que responde a la luz, un hidrogel que responde al redox, un hidrogel que responde al analito, o una de sus combinaciones. En algunas modalidades, un hidrogel que responde a estímulos puede ser un hidrogel que responde a múltiples estímulos.

Un “agente de liberación” o “desencadenante externo” es un agente que permite la eliminación de un recubrimiento extraíble condicionalmente de un sustrato cuando el agente de liberación se aplica al recubrimiento extraíble condicionalmente. Un desencadenante externo o agente de liberación puede incluir activadores físicos tales como estímulos térmicos, magnéticos, ultrasónicos, electroquímicos y/o de luz, así como también desencadenantes químicos tales como pH, reacciones redox, complejos supramoleculares y/o reacciones conducidas biocatalíticamente. Véase, por ejemplo, Echeverria, y otros, *Gels* (2018), 4, 54; doi:10.3390/gels4020054. El tipo de “agente de liberación” o “desencadenante externo” puede depender del tipo de recubrimiento extraíble condicionalmente. Por ejemplo, un recubrimiento extraíble condicionalmente que presenta un hidrogel que responde al redox puede eliminarse tras la aplicación de un agente de liberación que incluye un agente reductor tal como ditioneitol (DTT). Como otro ejemplo, un hidrogel que responde al pH puede eliminarse tras la aplicación de un agente de liberación que cambia el pH.

En algunas modalidades, un método descrito en la presente descripción comprende eliminar las reticulaciones de la muestra biológica reticulada reversiblemente. En algunas modalidades, la eliminación de las reticulaciones se realiza previo al ensayo espacial. No es necesario completar la eliminación de las reticulaciones. En algunas modalidades, solo se eliminan las reticulaciones de una porción de las moléculas reticuladas en la muestra biológica reticulada reversiblemente y se permite que migren.

ix. Desagregación de células

En algunas modalidades, la muestra biológica corresponde a células (por ejemplo, derivadas de un cultivo celular, una muestra de tejido o células depositadas sobre una superficie). En una muestra de célula con una pluralidad de células, las células individuales pueden desagregarse naturalmente. Por ejemplo, las células pueden derivarse de una suspensión de células y/o células disociadas o desagregadas de un tejido o sección de tejido.

Alternativamente, las células en la muestra pueden agregarse, y pueden desagregarse en células individuales mediante el uso de, por ejemplo, técnicas enzimáticas o mecánicas. Los ejemplos de enzimas que se usan en la desagregación enzimática incluyen, pero no se limitan a, dispasa, colagenasa, tripsina, y sus combinaciones. La desagregación mecánica puede realizarse, por ejemplo, mediante el uso de un homogeneizador de tejidos. La muestra biológica puede comprender células desagregadas (por ejemplo, células no adherentes o suspendidas) que se depositan sobre una superficie y se someten a un ensayo *in situ* y un ensayo espacial descrito en la presente descripción.

x. Permeabilización y tratamiento de los tejidos

En algunas modalidades, una muestra biológica puede permeabilizarse para facilitar la transferencia de analitos fuera de la muestra, y/o para facilitar la transferencia de especies (tales como sondas de captura) en la muestra. Si una muestra no se permeabiliza lo suficiente, la cantidad del analito capturado en la muestra puede ser demasiado baja para permitir un análisis adecuado. Por el contrario, si la muestra de tejido es demasiado permeable, puede perderse la relación espacial relativa de los analitos dentro de la muestra de tejido. Por tanto, es conveniente un equilibrio entre permeabilizar la muestra de tejido lo suficiente como para obtener una buena intensidad de señal, al tiempo que se mantiene aún la resolución espacial de la distribución del analito en la muestra.

En general, una muestra biológica puede permeabilizarse al exponer la muestra a uno o más agentes permeabilizantes. Los agentes adecuados para este propósito incluyen, pero no se limitan a, solventes orgánicos (por ejemplo, acetona, etanol, y metanol), agentes reticulantes (por ejemplo, paraformaldehído), detergentes (por ejemplo, saponina, Tritón X-100™ o Tween-20™), y enzimas (por ejemplo, tripsina, proteasas). En algunas modalidades, la muestra biológica puede incubarse con un agente permeabilizante de células para facilitar la permeabilización de la muestra. 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 en relación con las muestras descritas en la presente descripción.

En algunas modalidades, cuando se usa un medio resistente a la difusión para limitar la migración de analitos u otras especies durante el procedimiento analítico, el medio resistente a la difusión puede incluir al menos un reactivo de permeabilización. Por ejemplo, el medio resistente a la difusión puede incluir pocillos (por ejemplo, micro, nano o picopocillos) que contienen un tampón o reactivos de permeabilización. En algunas modalidades, donde el medio resistente a la difusión es un hidrogel, el hidrogel puede incluir un tampón de permeabilización. En algunas modalidades, el hidrogel se empapa en tampón de permeabilización previo a poner en contacto el hidrogel con una muestra. En algunas modalidades, el hidrogel u otro medio resistente a la difusión puede contener reactivos o monómeros secos para suministrar reactivos de permeabilización cuando el medio resistente a la difusión se aplica a una muestra biológica. En algunas modalidades, el medio resistente a la difusión (es decir, hidrogel) se une covalentemente a un sustrato sólido (es decir, un portaobjetos de vidrio acilado). En algunas modalidades, el hidrogel puede modificarse tanto para contener sondas de captura como para suministrar reactivos de permeabilización. Por ejemplo, una película de hidrogel puede modificarse para incluir sondas de captura con código de barras espacial. La película de hidrogel con código de barras espacial se empapa después en tampón de permeabilización antes de poner en contacto la película de hidrogel con código de barras espacial con la muestra. La película de hidrogel con código de barras espacial por lo tanto suministra reactivos de permeabilización a una superficie de la muestra en contacto con el hidrogel con código de barras espacial, lo que mejora la migración y captura del analito. En algunas modalidades, el hidrogel con código de barras espacial se aplica a una muestra y se coloca en una solución de permeabilización a granel. En algunas modalidades, la película de hidrogel empapada en reactivos de permeabilización se intercala entre una muestra y una matriz con código de barras espacial. En algunas modalidades, los analitos diana son capaces de difundir a través del hidrogel empapado del reactivo de permeabilización e hibridar o unir las sondas de captura sobre el otro lado del hidrogel. En algunas modalidades, el grosor del hidrogel es proporcional a la pérdida de resolución. En algunas modalidades, los pocillos (por ejemplo, micro, nano o picopocillos) pueden contener sondas de captura con código de barras espacial y reactivos y/o tampón de permeabilización. En algunas modalidades, las sondas de captura con código de barras espacial y los reactivos de permeabilización se mantienen entre separadores. En algunas modalidades, la muestra se perfora, se corta o se transfiere al pocillo, en donde un analito diana difunde a través del reactivo/tampón de permeabilización y a las sondas de captura con código de barras espacial. En algunas modalidades, la pérdida de resolución puede ser proporcional al grosor del espacio (por ejemplo, la cantidad de tampón de permeabilización entre la muestra y las sondas de captura). En algunas modalidades, el medio resistente a la difusión (por ejemplo, hidrogel) tiene entre aproximadamente 50-500 micrómetros de grosor que incluyen 500, 450, 400, 350, 300, 250, 200, 150, 100 o 50 micrómetros de grosor, o cualquier grosor dentro de 50 y 500 micrómetros.

En algunas modalidades, la solución de permeabilización puede suministrarse a una muestra a través de una membrana porosa. En algunas modalidades, se usa una membrana porosa para limitar las pérdidas difusivas del analito, mientras que permite que los reactivos de permeabilización alcancen una muestra. La química de la membrana y el tamaño de los poros pueden manipularse para minimizar la pérdida del analito. En algunas modalidades, la membrana porosa puede fabricarse de vidrio, silicio, papel, hidrogel, monolitos de polímero u otro material. En algunas modalidades, el material puede ser naturalmente poroso. En algunas modalidades, el material puede tener poros o pocillos grabados en material sólido. En algunas modalidades, los reactivos de permeabilización se hacen fluir a través de una cámara o canal microfluídico sobre la membrana porosa. En algunas modalidades, el flujo controla el acceso de la muestra a los reactivos de permeabilización. En algunas modalidades, una membrana porosa se intercala entre una matriz con código de barras espacial y la muestra, en donde la solución de permeabilización se aplica sobre la membrana porosa. Los reactivos de permeabilización difunden a través de los poros de la membrana y dentro del tejido.

En algunas modalidades, la muestra biológica puede permeabilizarse mediante la adición de uno o más reactivos de lisis a la muestra. Los ejemplos de agentes de lisis adecuados incluyen, pero no se limitan a, reactivos bioactivos tales como enzimas de lisis que se usan para la lisis de diferentes tipos de células, por ejemplo, bacterias grampositivas o gramnegativas, plantas, levaduras, mamíferos, tales como lisozimas, acromopeptidasa, lisostafina, labiasa, kitalasa, litílica, y una variedad de otras enzimas de lisis disponibles comercialmente.

Pueden añadirse adicional o alternativamente otros agentes de lisis a la muestra biológica para facilitar la permeabilización. Por ejemplo, pueden usarse soluciones de lisis basadas en surfactantes para lisar las células de la muestra. Las soluciones de lisis pueden incluir surfactantes iónicos tales como, por ejemplo, sarcosilo y dodecilsulfato de sodio (SDS). Más generalmente, los agentes de lisis química pueden incluir, sin limitación, solventes orgánicos, agentes quelantes, detergentes, surfactantes y agentes caotrópicos.

En algunas modalidades, la muestra biológica puede permeabilizarse mediante métodos de permeabilización no químicos. Los métodos de permeabilización no químicos se conocen en la técnica. Por ejemplo, los métodos de permeabilización no químicos que pueden usarse incluyen, pero no se limitan a, técnicas de lisis física tales como electroporación, métodos de permeabilización mecánica (por ejemplo, agitación con perlas mediante el uso de un homogeneizador y bolas de trituración para romper mecánicamente las estructuras de tejido de la muestra), permeabilización acústica (por ejemplo, sonicación), y técnicas de lisis térmica tales como calentamiento para inducir la permeabilización térmica de la muestra.

Pueden añadirse reactivos adicionales a una muestra biológica para realizar diversas funciones previo al análisis de la muestra. En algunas modalidades, pueden añadirse a la muestra agentes inactivadores de ADNasa y ARNasa o inhibidores tales como proteinasa K, y/o agentes quelantes tales como EDTA. Por ejemplo, un método descrito en la presente descripción puede comprender una etapa para aumentar la accesibilidad de un ácido nucleico para la unión, por ejemplo, una etapa de desnaturalización para la abertura del ADN en una célula para la hibridación mediante una sonda. Por ejemplo, puede usarse el tratamiento con proteinasa K para liberar el ADN con proteínas que se unen a este.

xi. Enriquecimiento selectivo de especies de ARN

En algunas modalidades, donde el ARN es el analito, pueden enriquecerse selectivamente una o más especies del analito de ARN de interés. Por ejemplo, pueden seleccionarse una o más especies del ARN de interés mediante la adición de uno o más oligonucleótidos a la muestra. En algunas modalidades, el oligonucleótido adicional es una secuencia que se usa para cebar una reacción mediante una enzima (por ejemplo, una polimerasa). Por ejemplo, pueden usarse una o más secuencias de cebadores con complementariedad de secuencias con uno o más ARN de interés para amplificar el uno o más ARN de interés, de esta manera se enriquecen selectivamente estos ARN.

En algunas modalidades, una o más sondas de ácido nucleico pueden usarse para hibridar con un ácido nucleico diana (por ejemplo, ADNc o molécula de ARN, tal como un ARNm) y ligarse en una reacción de ligación con molde (por ejemplo, ligación con molde de ARN (RTL) o ligación con molde de ADN (por ejemplo, en ADNc)) para generar un producto para el análisis. En algunos aspectos, cuando se analizan dos o más analitos, se usa una primera y segunda sonda que es específica para (por ejemplo, se hibrida específicamente con) cada analito de ARN o de ADNc. Por ejemplo, en algunas modalidades de los métodos que se proporcionan en la presente descripción, la ligación con molde se usa para detectar la expresión génica en una muestra biológica. Un analito de interés (tal como una proteína), que se une por un agente de marcaje o agente aglutinante (por ejemplo, un anticuerpo o fragmento de unión a epítipo de este), en donde el agente aglutinante se conjuga o se asocia de cualquier otra manera con un oligonucleótido reportero que comprende una secuencia reportera que identifica el agente aglutinante, puede seleccionarse para el análisis. Las sondas pueden hibridarse al oligonucleótido reportero y ligarse en una reacción de ligación con molde para generar un producto para el análisis. En algunas modalidades, los espacios entre las sondas oligonucleotídicas pueden llenarse previo a la ligación, mediante el uso de, por ejemplo, Mu polimerasa, ADN polimerasa, ARN polimerasa, transcriptasa inversa, VENT polimerasa, Taq polimerasa y/o cualquier combinación, derivados y variantes de estos (por ejemplo, mutantes modificados genéticamente). En algunas modalidades, el ensayo puede incluir además la amplificación de productos de ligación con molde (por ejemplo, mediante PCR multiplex).

En algunas modalidades, un oligonucleótido con complementariedad de secuencias con la hebra complementaria del ARN capturado (por ejemplo, ADNc) puede unirse al ADNc. Por ejemplo, los oligonucleótidos biotinilados con secuencia complementaria con uno o más ADNc de interés se unen al ADNc y pueden seleccionarse mediante el uso de la afinidad de la estreptavidina en la biotinylation que usa cualquiera de una variedad de métodos conocidos en el campo (por ejemplo, las perlas de estreptavidina).

Alternativamente, pueden seleccionarse negativamente una o más especies de ARN (por ejemplo, eliminarse) mediante el uso de cualquiera de una variedad de métodos. Por ejemplo, las sondas pueden administrarse a una muestra que se hibrida selectivamente al ARN ribosomal (ARNr), de esta manera se reduce la mezcla y concentración de ARNr en la muestra. La aplicación subsecuente de las sondas de captura a la muestra puede resultar en una captura mejorada de otros tipos de ARN debido a la reducción del ARN inespecífico presente en la muestra. Adicionalmente y alternativamente, el tratamiento con nucleasas específicas de la doble cadena (DSN) puede eliminar

el ARNr (véase, por ejemplo, Archer, y otros, Selective and flexible depletion of problematic sequences from RNA-seq libraries at the cDNA stage, *BMC Genomics*, 15 401, (2014)). Además, la cromatografía de hidroxapatita puede eliminar especies abundantes (por ejemplo, ARNr) (véase, por ejemplo, Vandernoot, V.A., cDNA normalization by hydroxyapatite chromatography to enrich transcriptome diversity in RNA-seq applications, *Biotechniques*, 53(6) 373-80, (2012)).

IV. Módulos de ensayo *in situ*

En algunos aspectos, en la presente descripción se proporcionan métodos que comprenden ensayos *in situ* mediante el uso de microscopía como una lectura, por ejemplo, secuenciación de ácido nucleico, hibridación u otros métodos de detección o determinación que implican una lectura óptica. En algunos aspectos, la detección o determinación de una secuencia de uno, dos, tres, cuatro, cinco o más nucleótidos de un ácido nucleico diana se realiza *in situ* en una célula en un tejido intacto. En algunas modalidades, el ensayo comprende detectar la presencia o ausencia de un producto de amplificación (por ejemplo, producto de RCA). En algunas modalidades, la presente descripción proporciona métodos para la caracterización con alto rendimiento de un gran número de dianas *in situ*, tales como transcritos y/o loci de ADN, por ejemplo, para detectar y/o cuantificar ácidos nucleicos y/o proteínas en células, tejidos, órganos u organismos. En algunas modalidades, la hibridación de sondas con la muestra y/o las etapas de detección durante el ensayo *in situ* se realiza sobre los analitos en la muestra que no se capturan mediante sondas de captura o agentes de captura.

A. Análisis *in situ*

En algunos aspectos, en la presente descripción se proporciona un método que comprende analizar dianas biológicas en base a la hibridación *in situ* de sondas que comprenden secuencias de ácido nucleico. En algunas modalidades, el método comprende la hibridación secuencial de oligonucleótidos marcados de manera detectable con sondas con código de barras que se unen directa o indirectamente a dianas biológicas en una muestra. En algunas modalidades, un oligonucleótido marcado de manera detectable se une directamente a una o más sondas con código de barras. En algunas modalidades, un oligonucleótido marcado de manera detectable se une indirectamente a una o más sondas con código de barras, por ejemplo, mediante una o más moléculas de ácido nucleico puente.

En algunos aspectos, se usa un ensayo basado en hibridación *in situ* para localizar y analizar secuencias de ácido nucleico (por ejemplo, una molécula de ADN o ARN que comprende una o más secuencias específicas de interés) dentro de una muestra biológica natural, por ejemplo, una porción o sección de tejido o una célula única. En algunas modalidades, el ensayo *in situ* se usa para analizar la presencia, ausencia, una cantidad o nivel de transcritos de ARNm (por ejemplo, un transcriptoma o un subconjunto de este, o moléculas de ARNm de interés) en una muestra biológica, mientras se conserva el contexto espacial. En algunas modalidades, la presente descripción proporciona composiciones y métodos para la hibridación *in situ* mediante el uso de moléculas marcadas directa o indirectamente, por ejemplo, ADN o ARN complementario o ácidos nucleicos modificados, como sondas que se unen o hibridan con un ácido nucleico diana dentro de una muestra biológica de interés.

Las sondas de ácido nucleico, en algunos ejemplos, pueden marcarse con radioisótopos, epítopos, hapteno, biotina o fluoróforos, para permitir la detección de la ubicación de secuencias de ácido nucleico específicas en cromosomas y en tejidos. En algunas modalidades, las sondas son específicas del locus (por ejemplo, específicas del gen) y se unen o se acoplan a regiones específicas de un cromosoma. En modalidades alternativas, las sondas son sondas repetitivas alfoides o centroméricas que se unen o se acoplan a secuencias repetitivas dentro de cada cromosoma. Las sondas también pueden ser sondas cromosómicas completas (por ejemplo, múltiples sondas más pequeñas) que se unen o se acoplan a secuencias a lo largo de un cromosoma completo.

En algunas modalidades, en la presente descripción se proporciona un método que comprende hibridación *in situ* del ADN para medir y localizar el ADN. En algunas modalidades, en la presente descripción se proporciona un método de hibridación *in situ* de ARN para medir y localizar ARN (por ejemplo, ARNm, ARNinc y miARN) dentro de una muestra biológica (por ejemplo, una muestra de tejido fijado). En algunas modalidades, la hibridación *in situ* del ARN implica la hibridación *in situ* por fluorescencia del ARN de una única molécula (FISH). En algunas modalidades, las sondas de ácido nucleico marcadas con fluorescencia se hibridan con dianas de ARN predeterminadas, para visualizar la expresión génica en una muestra biológica. En algunas modalidades, un método FISH comprende usar una única sonda de ácido nucleico específica para cada diana, por ejemplo, FISH de molécula única (smFISH). El uso de smFISH puede producir una señal de fluorescencia que permite la medición cuantitativa de transcritos de ARN. En algunas modalidades, smFISH comprende un conjunto de sondas de ácido nucleico, de aproximadamente 50 pares de bases de longitud, en donde cada sonda se acopla a un conjunto de fluoróforos. Por ejemplo, el conjunto de sondas de ácido nucleico puede comprender cinco sondas, en donde cada sonda se acopla a cinco fluoróforos. En algunas modalidades, dichas sondas de ácido nucleico se acoplan cada una a un fluoróforo. Por ejemplo, un protocolo smFISH puede usar un conjunto de aproximadamente 40 sondas de ácido nucleico, de aproximadamente 20 pares de bases de longitud, cada una se acopla a un único fluoróforo. En algunas modalidades, la longitud de las sondas de ácido nucleico varía, lo que comprende de 10 a 100 pares de bases, tales como de 30 a 60 pares de bases. Alternativamente, puede usarse una pluralidad de sondas de ácido nucleico que se dirigen a diferentes regiones del mismo transcrito de ARN. Los expertos en la técnica apreciarán que el tipo de sondas de ácido nucleico, el número de sondas de ácido

nucleico, el número de fluoróforos que se acoplan a dichas sondas, y la longitud de dichas sondas, pueden variar para adaptarse a las especificaciones del ensayo individual.

En modalidades adicionales, smFISH se aplica a un flujo de trabajo multiplexado, en donde se usan hibridaciones consecutivas/secuenciales (por ejemplo, como en seqFISH o seqFISH+) para impartir un código de barras temporal en los transcritos diana. Las rondas secuenciales de hibridación *in situ* de fluorescencia pueden acompañarse de imágenes y extracción de sondas, detección de transcritos individuales (por ejemplo, transcritos de ARN) dentro de una muestra biológica de interés (por ejemplo, una muestra de tejido, una célula única o ARN extraído). En algunas modalidades, cada ronda de hibridación comprende un conjunto predefinido de sondas (por ejemplo, entre aproximadamente 10 y aproximadamente 50 sondas tales como 24 a 32 sondas) que se dirigen a transcritos de ARN únicos. En algunos ejemplos, el conjunto predefinido de sondas es de múltiples colores. Opcionalmente, múltiples sondas de ácido nucleico se unen a la muestra, en donde cada sonda comprende una secuencia de iniciación para la amplificación, que permite una disminución de la autofluorescencia (por ejemplo, como en la reacción en cadena de hibridación de una única molécula (smHCR)). En algunas modalidades, un método smFISH multiplexado descrito en la presente descripción puede multiplexarse de 10 a más de 10 000 ARNm, opcionalmente acompañado de imágenes, para caracterizar de manera eficiente y precisa todo el transcriptoma. Los métodos de hibridación *in situ* pueden comprender además el uso de dos sondas para unirse a transcritos diana (por ejemplo, transcritos de ARN), que sirven como dianas de unión para cebadores de amplificación. En algunas modalidades, este proceso resulta en la amplificación de la señal (por ejemplo, como en el ARNscope). En algunas modalidades, los métodos de hibridación *in situ* pueden emplear etiquetas metálicas en lugar de fluoróforos (por ejemplo, citometría de masa de imágenes). Los anticuerpos conjugados con metal pueden acoplarse a las etiquetas metálicas que se hibridan con transcritos sobre una muestra biológica. En algunas modalidades, la citometría de masa puede usarse para cuantificar abundancias de metales, lo que permite la evaluación simultánea del ARN y la proteína dentro de una muestra biológica.

En algunas modalidades, un método descrito en la presente descripción comprende un protocolo FISH multiplexado que es resistente al error (por ejemplo, MERFISH). En algunas modalidades, dicho protocolo comprende sondas de ácido nucleico no leídas (por ejemplo, sondas primarias) que comprenden una región de unión (por ejemplo, una región que se une a una diana tal como transcritos de ARN) que se acopla a una o más regiones flanqueantes. En algunas modalidades, cada sonda de ácido nucleico no leída se acopla a dos regiones flanqueantes. Las sondas de ácido nucleico no leídas pueden hibridarse con un transcrito (por ejemplo, transcrito de ARN) dentro de una muestra biológica (por ejemplo, muestra de tejido o una célula única), de manera que las sondas de ácido nucleico de lectura fluorescente pueden hibridarse subsecuentemente en serie con la región(s) flanqueante(s) de las sondas de ácido nucleico no leídas. En algunas modalidades, cada ronda de hibridación comprende imágenes sucesivas y extracción de sondas para inactivar las señales de las sondas de ácido nucleico de lectura de las rondas anteriores. De los ARN pueden obtenerse imágenes mediante FISH, y los errores acumulados durante múltiples rondas de obtención de imágenes (por ejemplo, hibridaciones imperfectas) se detectan y/o se corrigen. En algunas modalidades, la microscopía de expansión se emplea para aumentar el número de dianas de ARN detectadas sin solapamiento de señales. En modalidades similares, las sondas de ácido nucleico no leídas se reticulan con los transcritos diana previo a la obtención de imágenes. La reticulación puede realizarse mediante cualquier método conocido en la técnica. En modalidades preferidas, la reticulación se realiza mediante el uso de la inclusión de tejidos en el hidrogel. Después de dichas etapas de reticulación, puede realizarse la codificación mediante código de barras, que comprende hibridaciones secuenciales mediante el uso de sondas de lectura acopladas a colores predeterminados para generar códigos de barras únicos (por ejemplo, generar pseudocolores a partir de hibridaciones consecutivas).

En algunas modalidades, uno o más códigos de barras de una sonda se seleccionan mediante oligonucleótidos de detección marcados de manera detectable, tales como oligos marcados con fluorescencia. En algunas modalidades, se usan uno o más esquemas de descodificación para descodificar las señales, tales como la fluorescencia, para la determinación de la secuencia. En cualquiera de las modalidades en la presente descripción, los códigos de barras (por ejemplo, secuencias de código de barras primarias y/o secundarias) pueden analizarse mediante el uso de cualquier método o técnica adecuada, que comprende las descritas en la presente descripción, tales como sondaje secuencial de ARN de dianas (SPOT de ARN), hibridación *in situ* fluorescente secuencial (seqFISH o seqFISH+), hibridación *in situ* fluorescente de molécula única (smFISH) o hibridación *in situ* con fluorescencia resistente a error multiplexada (MERFISH). En algunas modalidades, los métodos que se proporcionan en la presente descripción comprenden analizar los códigos de barras mediante hibridación secuencial y detección con una pluralidad de sondas marcadas (por ejemplo, oligonucleótidos de detección). Esquemas de descodificación ilustrativos se describen en *Engel y otros*, "Transcriptome-scale Super-Resolved Imaging in Tissues by RNA SeqFISH+", *Nature* 568(7751):235-239 (2019); *Chen y otros*, "Spatially resolved, highly multiplexed RNA profiling in single cells", *Science*, 348(6233):aa6090 (2015); los documentos US 10,457,980 B2; US 2016/0369329 A1; WO 2018/026873 A1; y US 2017/0220733 A1. En algunas modalidades, estos ensayos permiten la amplificación de la señal, la descodificación combinatoria, y los esquemas de corrección de errores al mismo tiempo.

Las estrategias similares de hibridación *in situ* mediante el uso de variaciones de las técnicas FISH pueden adoptarse, además, mediante los métodos descritos en la presente descripción. En algunas modalidades, un método comprende protocolos FISH multiplexados sin codificación mediante código de barras (por ejemplo, ouroboros sm-FISH (osmFISH)). Los métodos no codificados mediante código de barras pueden limitarse para detectar un número

específico de dianas, que se definen por el número de rondas de hibridación que se realizan. En algunas modalidades, la obtención de imágenes se realizan después de cada ronda de hibridación, en donde la sonda se extrae después de la obtención de imágenes, lo que permite la hibridación subsecuente y las rondas de obtención de imágenes.

Las modalidades adicionales de la presente descripción pueden incluir el uso de protocolos de hibridación *in situ* que no dependen de la captura de sondas de transcritos desde ubicaciones predefinidas. En algunas modalidades, puede usarse un mapeo espacial libre de óptica de transcritos en una muestra biológica (por ejemplo, un sistema de microscopía químicamente codificado). En algunas modalidades, los transcritos se etiquetan primero *in situ* con etiquetas de nucleótidos únicos (por ejemplo, identificadores moleculares únicos). Esta primera reacción puede continuarse con una segunda reacción de amplificación *in situ*, marcado por un nuevo conjunto de etiquetas de nucleótidos únicos (por ejemplo, identificadores de eventos únicos). En algunas modalidades, la secuenciación de ARN o ADN puede usarse para identificar cada secuencia de cadena molecular (por ejemplo, concatenadores). En modalidades adicionales, puede usarse un algoritmo para evaluar las proximidades de las secuencias y producir imágenes de los transcritos diana, en combinación con la información de secuencia.

En algunas modalidades, en la presente descripción se proporciona un método que comprende vincular la información de secuenciación y la información espacial de las dianas dentro de los entornos endógenos. Por ejemplo, el análisis de las secuencias de ácido nucleico puede realizarse directamente sobre el ADN o ARN dentro de una muestra biológica intacta de interés, por ejemplo, mediante secuenciación *in situ*. En algunas modalidades, la presente descripción permite la identificación y cuantificación simultánea de una pluralidad de dianas, tales como 100, 1000 o más de transcritos (por ejemplo, transcritos de ARNm), además de la resolución espacial de dichos transcritos. En algunos aspectos, la resolución espacial de los transcritos puede ser subcelular. Opcionalmente, la resolución espacial puede aumentarse mediante el uso de estrategias de amplificación de señales descritas en la presente descripción. En algunas modalidades, los colorantes fluorescentes se usan para dirigirse a las bases de ácido nucleico, y las sondas candado se usan para dirigirse al ARN de interés *in situ*. En algunas modalidades, los ARNm se transcriben de manera inversa en ADNc, y las sondas candado son capaces de unirse o acoplarse al ADNc. En algunas modalidades, las sondas candado comprenden oligonucleótidos con extremos que son complementarios a una secuencia diana (por ejemplo, transcritos de ADNc diana). Después de la hibridación de las sondas candado con la secuencia diana, pueden usarse enzimas para ligar los extremos de las sondas candado y catalizar la formación de ADN circularizado.

En algunas modalidades, los extremos de las sondas candado están muy cerca después de la hibridación con el ARN o ADNc diana, para permitir la ligación y circularización de la sonda candado. Las sondas candado pueden comprender adicionalmente una o más secuencias de código de barras. En modalidades alternativas, puede haber un espacio entre los extremos de las sondas candado después de la hibridación con el ARN o ADNc diana, que debe llenarse con ácidos nucleicos (por ejemplo, mediante polimerización de ADN), previo a la ligación de los extremos de las sondas candado y la circularización. En algunas modalidades, el espacio entre los extremos de las sondas candado es de longitud variable, por ejemplo, hasta cuatro pares de bases, y puede permitir leer la secuencia real de ARN o ADNc. En algunas modalidades, la ADN polimerasa tiene actividad de desplazamiento de hebra. En algunas modalidades, la ADN polimerasa puede, en cambio, no tener actividad de desplazamiento de hebra, tal como la polimerasa que se usa en la secuenciación de dianas *in situ* con código de barras (BaristaSeq), que proporciona una longitud de lectura de hasta 15 bases mediante el uso de un enfoque de sonda candado con relleno de espacios. Véase, por ejemplo, Chen y otros, *Nucleic Acids Res.* 2018, 46, e22.

Un método descrito en la presente descripción puede comprender la circularización y la amplificación de ADN (por ejemplo, amplificación por círculo rodante), en la ubicación de las sondas candado. En algunas modalidades, la amplificación resulta en múltiples repeticiones de secuencias de sonda candado. La secuenciación y/o decodificación de las sondas candado amplificadas puede realizarse mediante el uso de secuenciación mediante ligación. En métodos alternativos, se usan secuenciación mediante hibridación o secuenciación mediante síntesis. En algunas modalidades, los amplicones se estabilizan mediante la formación de reticulaciones descrito en la presente descripción, durante el proceso de secuenciación. En algunas modalidades, los métodos de secuenciación *in situ* presentados en esta descripción pueden automatizarse en una plataforma microfluídica.

Los expertos en la técnica apreciarán enfoques adicionales para la secuenciación *in situ*. Por ejemplo, en algunas modalidades, las sondas candado con código de barras pueden no transcribirse de manera inversa. En su lugar, un segundo cebador se une (por ejemplo, se liga) directamente a una secuencia de ARN adyacente a la sonda candado. En algunas modalidades, se realiza la amplificación (por ejemplo, amplificación por círculo rodante), en donde el producto de amplificación se incluye dentro de un hidrogel mediante cualquier método adecuado conocido en la técnica (por ejemplo, química hidrogel-tejido), que después se limpia de proteínas y lípidos que no se unen. Los productos de amplificación incluidos pueden, por ejemplo, secuenciarse mediante el uso de variaciones del enfoque de secuenciación mediante ligación, para determinar la secuencia de código de barras de cada sonda candado. En algunas modalidades, las combinaciones de química y secuenciación descritas en la presente descripción pueden usarse para analizar la orientación espacial de los transcritos diana en 3D.

En algunas modalidades, los métodos de secuenciación *in situ* descritos en la presente descripción pueden ser no dirigidos. En algunas modalidades, la secuenciación *in situ* no dirigida puede comprender una caracterización amplia

del genoma/transcriptoma de la expresión génica dentro de una muestra biológica de interés, por ejemplo, como en la secuenciación fluorescente de ARN *in situ* (FISSEQ). En algunas modalidades, las especies de ARN se capturan y se convierten en amplicones de ADNc reticulados (por ejemplo, ADNc reticulado a la matriz de proteína celular de la muestra). En algunos ejemplos, la síntesis de ADNc se realiza mediante el uso de bases de aminas modificadas para promover el proceso de reticulación. La síntesis de amplicones de ADNc reticulados puede continuarse mediante amplificación (por ejemplo, amplificación por círculo rodante) como se describe en otra parte de la presente descripción. En algunas modalidades, la secuenciación mediante ligación puede usarse para secuenciar los productos de amplificación. En algunas modalidades, la etapa de secuenciación incluye la secuenciación de partición para secuenciar selectivamente los subconjuntos de productos de amplificación. En algunas modalidades, las estrategias descritas en la presente descripción permiten la detección de ARN, ADN y/o proteínas, en tándem. En algunas modalidades, la secuenciación *in situ* puede combinarse con secuenciación *ex situ*, por ejemplo, como en la secuenciación de accesibilidad al transcriptoma *in situ* (INSTA-Seq).

En algunas modalidades, la secuenciación *in situ* implica la incorporación de un nucleótido marcado (por ejemplo, mononucleótidos o dinucleótidos marcados con fluorescencia) de manera secuencial dependiente de molde o la hibridación de un cebador marcado (por ejemplo, un hexámero aleatorio marcado) a un molde de ácido nucleico de manera que pueden determinarse las identidades (es decir, la secuencia de nucleótidos) de los nucleótidos incorporados o productos de extensión del cebador marcados, y en consecuencia, la secuencia de nucleótidos del molde de ácido nucleico correspondiente. Aspectos de la secuenciación *in situ* se describen, por ejemplo, en Mitra y otros, (2003) *Anal. Biochem.* 320, 55-65, y Lee y otros, (2014) *Science*, 343(6177), 1360-1363. Además, ejemplos de métodos y sistemas para realizar la secuenciación *in situ* se describen en los documentos WO2014/163886, WO2018/045181, WO2018/045186, y en US 10,138,509 y US. 10,179,932. Las técnicas ilustrativas para la secuenciación *in situ* comprenden, pero no se limitan a, STARmap (descrita, por ejemplo, en Wang y otros, (2018) *Science*, 361(6499) 5691), MERFISH (descrita, por ejemplo, en Moffitt, (2016) *Methods in Enzymology*, 572, 1-49), y FISSEQ (descrita, por ejemplo, en el documento US 2019/0032121).

i. Sondas e hibridación de la sonda

En algunos aspectos, los métodos descritos en la presente descripción implican el uso de una o más sondas o conjuntos de sondas que se hibridan con un ácido nucleico diana, tal como una molécula de ARN. Las sondas o conjuntos de sondas ilustrativas pueden basarse en una sonda candado, una sonda candado espaciada, un conjunto de sondas SNAIL (ligación intramolecular asistida por nucleótidos soporte), un conjunto de sondas PLAYR (ensayo de ligación de proximidad para ARN), un conjunto de sondas PLISH (hibridación *in situ* por ligación de proximidad) y sondas de ligación con molde de ARN. El diseño específico de la sonda o conjunto de sondas puede variar. En algunas modalidades, una sonda primaria (por ejemplo, una sonda de ADN que se une directamente a un ARN diana) se amplifica a través de la amplificación por círculo rodante, por ejemplo, mediante el uso de una sonda circular o una sonda circularizada de la ligación con candado como un molde. En algunas modalidades, las sondas primarias, tales como una sonda candado o un conjunto de sondas que comprende una sonda candado, contienen uno o más códigos de barras. En algunas modalidades, uno o más códigos de barras son indicativos de una secuencia en el ácido nucleico diana, tal como un único nucleótido de interés (por ejemplo, SNP o mutaciones puntuales), una secuencia de dinucleótidos, una secuencia corta de aproximadamente 5 nucleótidos de longitud, o una secuencia de cualquier longitud adecuada.

En algunas modalidades, en la presente descripción se proporciona una sonda o conjunto de sondas capaz de realizar ligación con molde de ADN, tal como de una molécula de ADNc. Véase, por ejemplo, la patente de Estados Unidos 8,551,710. En algunas modalidades, en la presente descripción se proporciona una sonda o conjunto de sondas capaz de realizar ligación con molde de ARN. Véase, por ejemplo, el documento PCT App. PCT/EP2018/077161, publicado como el documento WO2019068880. En algunas modalidades, el conjunto de sondas es un conjunto de sondas SNAIL. Véase, por ejemplo, la publicación de patente de Estados Unidos 20190055594. En algunas modalidades, en la presente descripción se proporciona una sonda o conjunto de sondas capaz de realizar ligación por proximidad, por ejemplo, un conjunto de sondas de ensayo de ligación por proximidad para ARN (por ejemplo, PLAYR). Véase, por ejemplo, la publicación de patente de Estados Unidos 20160108458.

En algunas modalidades, una sonda circular puede hibridarse indirectamente con el ácido nucleico diana. En algunas modalidades, el constructo circular se forma a partir de un conjunto de sondas capaz de ligación por proximidad, por ejemplo, un conjunto de sondas de hibridación *in situ* de ligación por proximidad (PLISH). Véase, por ejemplo, el documento PCT App. PCT/US2018/023846, publicado como el documento WO2018175779.

Un conjunto de sondas ilustrativo y un complejo de hibridación se muestran en la figura 4A, donde un candado o sonda circular se hibrida directamente con un transcrito de ARN. Un cebador de soporte puede usarse para facilitar la ligación del candado con el molde de ADN. El candado o sonda circular puede comprender una secuencia de direccionamiento (por ejemplo, que se hibrida con la diana) y una o más regiones de código de barras diana, tales como las secuencias de código de barras primarias BC1 y BC2 que se muestran en la figura 4A. Después de la hibridación de la sonda y/o cualquier etapa de circularización para proporcionar una sonda circular, en algunas modalidades la sonda circular se amplifica, por ejemplo, en una reacción RCA, para generar una molécula amplificada que comprende los códigos de barras primarios/dianas (por ejemplo, BC1) o secuencias complementarias de estos. En algunas modalidades,

después de la amplificación, el método comprende además detectar el producto de amplificación mediante el uso de un oligonucleótido marcado de manera detectable (tal como un oligo de detección marcado con fluorescencia) que es capaz de hibridarse con una o más de las secuencias de código de barras (por ejemplo, BC1 o BC2) o secuencias complementarias de estas.

Otro conjunto de sondas y complejo de hibridación ilustrativos se muestran en la figura 4B, donde un candado o sonda circular se hibrida directamente con un transcrito de ARN. Un cebador de soporte puede usarse para facilitar la ligación del candado con el molde de ADN. El candado o sonda circular puede comprender una secuencia de direccionamiento (por ejemplo, que se hibrida con la diana) y una o más regiones de código de barras diana, tales como las secuencias de código de barras primarias BC1 y BC2 que se muestran en la figura 4B. Después de la hibridación de la sonda y/o cualquier etapa de circularización para proporcionar una sonda circular, en algunas modalidades la sonda circular se amplifica, por ejemplo, en una reacción RCA, para generar una molécula amplificada que comprende los códigos de barras primarios/dianas (por ejemplo, BC1) o secuencias complementarias de estos. En algunas modalidades, después de la amplificación, el método comprende además usar una sonda de detección (por ejemplo, una sonda secundaria) que comprende (1) una región de unión a código de barras que se hibrida con la región de código de barras primaria/diana de la sonda de direccionamiento directa o indirectamente, y (2) dos o más regiones de código de barras de detección (por ejemplo, SBC11 y SBC12) que cada una hibrida con un oligonucleótido marcado de manera detectable. Por ejemplo, la figura 4B muestra dos sondas de detección, la Sonda Secundaria 1 y la Sonda Secundaria 2. La Sonda secundaria 1 comprende una región de unión al código de barras que se hibrida con BC1 de la sonda de direccionamiento directa o indirectamente, y cuatro regiones de código de barras de detección, códigos de barras secundarios (SBC) 11, SBC12, SBC13 y SBC14. Cada uno de SBC11, SBC12, SBC13 y SBC14 es capaz de hibridarse con un oligonucleótido marcado de manera detectable, tal como un oligo de detección marcado con fluorescencia. Igualmente, la Sonda Secundaria 2 comprende una región de unión al código de barras que se hibrida con BC2 de la sonda de direccionamiento directa o indirectamente, y cuatro regiones de código de barras de detección, códigos de barras secundarios (SBC) 21, SBC22, SBC23 y SBC24, cada una de las cuales es capaz de hibridarse con un oligonucleótido marcado de manera detectable, tal como un oligo de detección marcado con fluorescencia. En algunas modalidades, dos o más de los códigos de barras secundarios son diferentes entre sí. Por ejemplo, todos los códigos de barras secundarios de las sondas secundarias que se unen a la misma sonda primaria pueden ser diferentes, por ejemplo, cada código de barras secundario puede hibridarse específicamente con un oligo de detección e identificarse de manera única por la secuencia del oligo de detección.

Aunque la figura 4B muestra sondas secundarias que se hibridan con la sonda primaria para facilitar la ilustración, se apreciará que en algunas modalidades, la sonda primaria o un conjunto de sondas que comprende la sonda primaria se hibrida primero con la diana de ARN, seguido de la amplificación de la sonda primaria que es circular o circularizada después de la hibridación de la diana, por ejemplo, mediante el uso de RCA. La figura 4B muestra un producto de RCA de la sonda primaria, donde se amplifican las secuencias de códigos de barras primarios BC1 y BC2 o secuencias complementarias de estas. Por lo tanto, el producto de RCA es capaz de hibridarse con una pluralidad de Sondas Secundarias 1 así como también con una pluralidad de Sondas Secundarias 2.

En cualquiera de las modalidades descritas en la presente descripción, en la presente descripción se describe un ensayo multiplexado donde múltiples ácidos nucleicos diana (por ejemplo, genes o transcritos de ARN) se sondan con múltiples sondas primarias (por ejemplo, sondas candado primarias) y opcionalmente múltiples sondas secundarias que se hibridan con los códigos de barras primarios (o secuencias complementarias de estos) se hibridan a la vez, seguido de detección secuencial de códigos de barras secundarios y decodificación de las señales.

La figura 4C proporciona otro flujo de trabajo de ensayo *in situ* ilustrativo. Una diana de ARN se transcribe inversamente para generar una molécula de ADN, y después una sonda primaria se hibrida con la molécula de ADN. En el caso de una sonda candado, el candado puede ligarse mediante el uso del ADN que se genera a partir del ARN como un soporte. En lugar de usar sondas secundarias con código de barras en la figura 4B, la figura 4C muestra una sonda puente capaz de hibridarse con una secuencia de código de barras de la sonda primaria o un producto de amplificación de esta (por ejemplo, producto de RCA). Una sonda puente puede comprender una secuencia que no se hibrida con una secuencia de código de barras (o complemento de esta) de la sonda primaria pero que es capaz de hibridarse con uno o más oligos de detección marcados de manera detectable. Un método ilustrativo para usar oligos de detección en un sistema de codificación mediante código de barras a través de la química de secuencia mediante hibridación para la detección espacial de transcritos de ARN puede encontrarse en Gyllborg y otros, "Hybridization-based *In Situ* Sequencing (HybISS): spatial transcriptomic detection in human and mouse brain tissue," bioRxiv 2020.02.03.931618.

La figura 5 proporciona otros flujos de trabajo de ensayo *in situ* ilustrativos. Diversas sondas primarias pueden hibridarse a un ARN diana y ligarse mediante el uso de ligación con molde de ARN y/o ligación con molde de ADN para formar una sonda circularizada que comprende una o más secuencias de código de barras. Una sonda secundaria o sonda puente puede hibridarse con la sonda circularizada o un producto de amplificación de esta (por ejemplo, como se muestra en las figuras 4A-4C). Los oligos de detección pueden hibridarse con la sonda circularizada o un producto de amplificación de esta, una sonda secundaria o un producto de amplificación de esta, o una sonda puente (por ejemplo, como se muestra en las figuras 4A-4C). Por ejemplo, la sonda candado puede circularizarse mediante el uso de ligación con molde de ARN, véase, por ejemplo, la primera y tercera sondas candado en la figura 5. Puede usarse

una ligasa con molde de ARN para cerrar el círculo de una sonda de ADN lineal para circularizar el candado, y la eficiencia de la ligación puede aumentarse mediante la incorporación de ribonucleótidos en las sondas candado de ADN. Véase, *por ejemplo*, el documento PCT/EP2018/077161. En algunas modalidades, la eficiencia de la ligación de la sonda candado puede aumentarse mediante el uso de oligonucleótidos de ADN soporte tales como los que se muestran en la segunda y cuarta sondas candado en la figura 5. En algunas modalidades, las dos mitades (por ejemplo, como se muestra en la cuarta sonda candado de la figura 5 antes de la ligación) pueden servir como un ADN soporte entre sí para la ligación. Cualquier método adecuado de ligación con molde de ARN o ligación con molde de ADN puede usarse en el ensayo *in situ* y están comprendidos en la presente descripción.

La figura 6A muestra agentes aglutinante o de marcaje ilustrativos (por ejemplo, un anticuerpo o fragmento de unión a antígeno de este) que comprende un oligonucleótido reportero. En algunas modalidades, el oligonucleótido reportero comprende una secuencia correspondiente al agente de marcaje y por lo tanto corresponde a (por ejemplo, que identifica de manera única) un analito (por ejemplo, un analito de proteína) o elemento celular que el agente de marcaje marca. En algunas modalidades, el agente de marcaje puede comprender un oligonucleótido reportero que comprende una o más secuencias de código de barras. Por lo tanto, el oligonucleótido reportero puede ser un analito de ácido nucleico descrito en la presente descripción, y puede analizarse mediante el uso de cualquiera de los métodos descritos en la presente descripción. Por ejemplo, como se muestra en la figura 6A, puede usarse una sonda tal como una sonda candado para el analito de un oligonucleótido reportero. En algunos ejemplos, el oligonucleótido reportero de un agente de marcaje que reconoce específicamente una proteína puede analizarse mediante el uso de hibridación *in situ* (por ejemplo, hibridación secuencial) y/o secuenciación *in situ* (por ejemplo, mediante el uso de sondas candado y la amplificación por círculo rodante de sondas candado ligadas). Además, el oligonucleótido reportero del agente de marcaje y/o un complemento de este y/o un producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este puede capturarse mediante un agente de captura descrito en la presente descripción y analizarse mediante el uso de un ensayo espacial.

En algunas modalidades, un analito (un analito de ácido nucleico o analito que no es ácido nucleico) puede unirse específicamente a uno o más oligonucleótidos reporteros (por ejemplo, anticuerpos) cada uno de los cuales se une a un oligonucleótido reportero. En algunas modalidades, un analito (un analito de ácido nucleico o analito que no es ácido nucleico) puede unirse específicamente por dos agentes de marcaje (por ejemplo, anticuerpos) cada uno de los cuales se une a un oligonucleótido reportero (por ejemplo, ADN) que puede participar en reacciones de ligación, replicación y decodificación de secuencias, por ejemplo, mediante el uso de una sonda o conjunto de sondas (por ejemplo, una sonda candado como se muestra en la figura 6B, un conjunto de sonda SNAIL como se muestra en la figura 6C, una sonda circular como se muestra en la figura 6D, o una sonda candado y un conector como se muestra en la figura 6E). En algunas modalidades, el conjunto de sondas puede comprender dos o más sondas oligonucleotídicas, cada una que comprende una región que es complementaria entre sí. Por ejemplo, una reacción de ligación por proximidad puede incluir oligonucleótidos reporteros que se unen a pares de anticuerpos que pueden unirse mediante ligación si los anticuerpos se han traído cerca entre sí, por ejemplo, mediante la unión a la misma proteína diana (complejo), y los productos de ligación de ADN que se forman se usan después como molde para la amplificación por PCR, como se describe por ejemplo en Söderberg y otros, *Methods*. (2008), 45(3): 227-32. En algunas modalidades, una reacción de ligación por proximidad puede incluir oligonucleótidos reporteros que se unen a anticuerpos que cada uno se une a un miembro de un par de unión o complejo, por ejemplo, para analizar una unión entre miembros del par de unión o complejo. Para la detección de analitos mediante el uso de oligonucleótidos cerca, véase, por ejemplo, la publicación de solicitud de patente de Estados Unidos núm. 2002/0051986. En algunas modalidades, dos analitos cerca pueden unirse específicamente por dos agentes de marcaje (por ejemplo, anticuerpos) cada uno de los cuales se une a un oligonucleótido reportero (por ejemplo, ADN) que puede participar, cuando está cerca, cuando se une a sus respectivas dianas, en reacciones de ligación, replicación y/o decodificación de secuencias.

En algunas modalidades, dos analitos (o dos regiones de un analito) pueden unirse específicamente por dos agentes de marcaje diferentes (por ejemplo, anticuerpos) cada uno de los cuales se une a un oligonucleótido reportero (por ejemplo, ADN) que puede ligarse si los dos agentes de marcaje están lo suficientemente cerca para permitir que los oligonucleótidos reporteros se unan mediante ligación. En algunos casos, una vez que se produce la ligación, el producto ligado (por ejemplo, oligonucleótidos reporteros ligados) o producto o derivado de este puede capturarse mediante un agente de captura y analizarse.

En algunas modalidades, dos analitos (o dos regiones de un analito) pueden unirse específicamente por dos agentes de marcaje (por ejemplo, anticuerpos) cada uno de los cuales se une a un oligonucleótido reportero (por ejemplo, ADN), y una sonda que comprende una primera región para hibridar con uno de los dos oligonucleótidos reporteros y una segunda región para hibridar con el otro oligonucleótido reportero se añade a la muestra. En algunas modalidades, la sonda es una sonda candado que comprende opcionalmente un código de barras que puede asociarse con los agentes de marcaje. En algunos aspectos, la sonda puede detectarse mediante la hibridación de dos o más sondas para la ligación a secuencias de la sonda (por ejemplo, sonda candado). En algunos casos, una de las dos o más sondas de ligación se liga mediante el uso de la sonda candado como molde y el producto ligado puede capturarse mediante un agente de captura y analizarse. En algunos casos, la sonda (por ejemplo, sonda candado) puede ligarse y usarse para el análisis o detección aguas abajo. Por ejemplo, la sonda ligada puede usarse para RCA y el producto

de RCA puede detectarse mediante el uso de cualquiera de los métodos adecuados. En algunas modalidades, la información de los agentes de marcaje puede ser útil para caracterizar células (por ejemplo, mediante el direccionamiento hacia un marcador celular y/o proteína con el/los agente(s) de marcaje. En algunos casos, la información del/de los agente(s) de marcaje puede asociarse con la información del ensayo *in situ* que se proporciona en la sección IV.

En las figuras 6A, 6B y 6E el uno o más oligonucleótidos reporteros (y opcionalmente una o más moléculas de ácido nucleico tales como un conector que se muestra en la figura 6E) ayudan en la ligación de la sonda. Después de la ligación, la sonda puede formar una sonda circularizada. En algunas modalidades, una o más sondas adecuadas pueden usarse y ligarse, en donde la una o más sondas comprenden una secuencia que es complementaria al uno o más oligonucleótidos reporteros (o porción de este). La sonda puede comprender una o más secuencias de código de barras. En algunas modalidades, el uno o más oligonucleótidos reporteros pueden servir como un cebador para la amplificación por círculo rodante (RCA) de la sonda circularizada. En algunas modalidades, un ácido nucleico diferente del uno o más oligonucleótidos reportero se usa como un cebador para la amplificación por círculo rodante (RCA) de la sonda circularizada. Por ejemplo, un ácido nucleico capaz de hibridarse con la sonda circularizada en una secuencia distinta de la(s) secuencia(s) que se hibridan con el uno o más oligonucleótidos reporteros puede usarse como el cebador para la RCA. En otros ejemplos, el cebador en un conjunto de sondas SNAIL (por ejemplo, como se muestra en la figura 6C) se usa como cebador para la RCA.

En la figura 6F, uno o más analitos pueden unirse específicamente por dos anticuerpos primarios, cada uno de los cuales a su vez se reconoce por un anticuerpo secundario cada uno de los cuales se une a un oligonucleótido reportero (por ejemplo, ADN). Los anticuerpos secundarios que se unen a oligonucleótidos, sondas y conjuntos de sondas, conectores y/o cebadores reporteros pueden incluir cualquiera de los que se muestran en las figuras 6A-6E. Cada molécula de ácido nucleico puede ayudar en la ligación de la sonda para formar una sonda circularizada. En algunos casos, la sonda puede comprender una o más secuencias de código de barras. Además, el oligonucleótido reportero puede servir como un cebador para la amplificación por círculo rodante de la sonda circularizada. Las moléculas de ácido nucleico, sondas circularizadas y productos de RCA pueden analizarse mediante el uso de cualquier método adecuado descrito en la presente descripción para el análisis *in situ* así como también el análisis espacial.

En algunas modalidades, una o más sondas que se dirigen directa o indirectamente a uno o más analitos (por ejemplo, ácidos nucleicos, proteínas o elementos celulares) se ponen en contacto con la muestra previo a, o durante un módulo de ensayo *in situ*. La una o más sondas pueden incluir un agente de marcaje (por ejemplo, un anticuerpo que comprende un oligonucleótido reportero), una sonda candado o conjunto de sondas, sondas de ligación con molde, un agente de captura del analito, o cualquiera de sus combinaciones. En algunas modalidades, una o más sondas que se dirigen directa o indirectamente a uno o más analitos (por ejemplo, ácidos nucleicos, proteínas o elementos celulares) se ponen en contacto con la muestra después de un módulo de ensayo *in situ* pero previo a, o durante un módulo de ensayo espacial, en donde la una o más sondas pueden incluir un agente de marcaje (por ejemplo, un anticuerpo que comprende un oligonucleótido reportero), sondas de ligación con molde, un agente de captura del analito, una sonda de captura o cualquiera de sus combinaciones.

ii. Ligación

En algunas modalidades, los métodos que se proporcionan implican ligar uno o más polinucleótidos que son parte de un complejo de hibridación que comprende un ácido nucleico diana para el análisis *in situ*. En algunas modalidades, la ligación implica una ligación química. En algunas modalidades, la ligación implica una ligación dependiente de molde. En algunas modalidades, la ligación implica una ligación independiente de molde. En algunas modalidades, la ligación implica una ligación enzimática.

En algunas modalidades, la ligación enzimática implica el uso de una ligasa. En algunos aspectos, la ligasa que se usa en la presente descripción comprende una enzima que se usa comúnmente para unir polinucleótidos entre sí o para unir los extremos de un solo polinucleótido. Pueden usarse una ARN ligasa, una ADN ligasa, u otra variedad de ligasa para ligar dos secuencias de nucleótidos entre sí. Las ligasas comprenden ligasas de polinucleótidos bicatenarios dependientes de ATP, ADN o ARN ligasas bicatenarias dependientes de NAD⁺ y ligasas de polinucleótidos monocatenarios, por ejemplo, cualquiera de las ligasas descritas en los documentos EC 6.5.1.1 (ligasas dependientes de ATP), EC 6.5.1.2 (ligasas dependientes de NAD⁺), EC 6.5.1.3 (ARN ligasas). Los ejemplos específicos de ligasas comprenden ligasas bacterianas tales como ADN ligasa de *E. coli*, ADN ligasa de *Tth*, ADN ligasa de *Thermococcus* sp. (cepa 9°N) (ADN ligasa 9°N™, New England Biolabs), ADN ligasa Taq, Ampligasa™ (Epicentre Biotechnologies) y ligasas de fagos tales como ADN ligasa T3, ADN ligasa T4 y ADN ligasa T7 y mutantes de estas. En algunas modalidades, la ligasa es una ARN ligasa T4. En algunas modalidades, la ligasa es una ligasa splintR. En algunas modalidades, la ligasa es una ADN ligasa monocatenaria. En algunas modalidades, la ligasa es una ADN ligasa T4. En algunas modalidades, la ligasa es una ligasa que tiene una actividad ADN ligasa con puente de ADN. En algunas modalidades, la ligasa es una ligasa que tiene una actividad ADN ligasa con puente de ARN.

En algunas modalidades, la ligación en la presente descripción es una ligación directa. En algunas modalidades, la ligación en la presente descripción es una ligación indirecta. "Ligación directa" significa que los extremos de los polinucleótidos se hibridan inmediatamente adyacentes entre sí para formar un sustrato para una enzima ligasa, lo

que resulta en su ligación entre sí (ligación intramolecular). Alternativamente, "indirecta" significa que los extremos de los polinucleótidos se hibridan de manera no adyacente entre sí, es decir, separados por uno o más nucleótidos intervinientes o "espacios". En algunas modalidades, dichos extremos no se ligan directamente entre sí, sino que se producen ya sea mediante la intermediación de uno o más (oligo)nucleótidos intervinientes (llamados (oligo)nucleótidos de "espacio" o de "relleno de espacios") o mediante la extensión del extremo 3' de una sonda para "rellenar" el "espacio" correspondiente a dichos nucleótidos intervinientes (ligación intermolecular). En algunos casos, el espacio de uno o más nucleótidos entre los extremos hibridados de los polinucleótidos puede ser "rellenado" mediante uno o más (oligo)nucleótidos de "espacio" que son complementarios a un soporte, sonda candado, o ácido nucleico diana. El espacio puede ser un espacio de 1 a 60 nucleótidos, o un espacio de 1 a 40 nucleótidos, o un espacio de 3 a 40 nucleótidos. En modalidades específicas, el espacio puede ser un espacio de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 o más nucleótidos, de cualquier número entero (o intervalo de números enteros) de nucleótidos entre los valores indicados. En algunas modalidades, el espacio entre dichas regiones terminales puede rellenarse mediante un oligonucleótido de espacio o al extender el extremo 3' de un polinucleótido. En algunos casos, la ligación implica ligar los extremos de la sonda a al menos un (oligo)nucleótido de espacio, de manera que el (oligo)nucleótido de espacio se incorpora en el polinucleótido resultante. En algunas modalidades, la ligación en la presente descripción está precedida por el relleno de espacios. En otras modalidades, la ligación en la presente descripción no requiere el relleno de espacios.

En algunas modalidades, la ligación de los polinucleótidos produce polinucleótidos con una temperatura de fusión mayor que la de los polinucleótidos no ligados. Por tanto, en algunos aspectos, la ligación estabiliza el complejo de hibridación que contiene los polinucleótidos ligados previo a las etapas subsecuentes, que comprenden la amplificación y la detección.

En algunos aspectos, se usa una ligasa de alta fidelidad, tal como una ADN ligasa termoestable (por ejemplo, una ADN ligasa *Taq*). Las ADN ligasas termoestables son activas a temperaturas elevadas, lo que permite una discriminación adicional mediante la incubación de la ligación a una temperatura cercana a la temperatura de fusión (T_m) de las hebras de ADN. Esto reduce selectivamente la concentración de sustratos asociados con errores de apareamiento (se espera que tengan una T_m ligeramente inferior alrededor del error de apareamiento) sobre los sustratos con apareamiento total de bases asociados. Por tanto, la ligación de alta fidelidad puede lograrse a través de una combinación de la selectividad intrínseca del sitio activo de la ligasa y condiciones equilibradas para reducir la incidencia del ADNbc asociado con errores de apareamiento.

iii. Amplificación

En algunas modalidades, los métodos de la invención comprenden la etapa de amplificar uno o más polinucleótidos, por ejemplo, la sonda candado o una sonda circular que se forma a partir de la sonda candado. En algunas modalidades, la amplificación se logra al realizar la amplificación por círculo rodante (RCA). En otras modalidades, un cebador que se hibrida con la sonda candado se añade y se usa como tal para la amplificación.

En algunas modalidades, se realiza una etapa de eliminación para eliminar moléculas que no se hibridan específicamente con el ácido nucleico diana y/o la sonda circular. En algunas modalidades, la etapa de eliminación se realiza para eliminar sondas no ligadas. En algunas modalidades, la etapa de eliminación se realiza después de la ligación y previo a la amplificación.

En algunas modalidades, la amplificación se realiza a una temperatura entre o entre aproximadamente 20 °C y aproximadamente 60 °C. En algunas modalidades, la amplificación se realiza a una temperatura entre o entre aproximadamente 30 °C y aproximadamente 40 °C. En algunos aspectos, la etapa de amplificación, tal como la amplificación por círculo rodante (RCA) se realiza a una temperatura entre a o aproximadamente 25 °C y a o aproximadamente 50 °C, tal como a o aproximadamente 25 °C, 27 °C, 29 °C, 31 °C, 33 °C, 35 °C, 37 °C, 39 °C, 41 °C, 43 °C, 45 °C, 47 °C, o 49 °C.

En algunas modalidades, después de la adición de una ADN polimerasa en presencia de precursores de dNTP apropiados y otros cofactores, un cebador se extiende para producir múltiples copias del molde circular. Esta etapa de amplificación puede utilizar amplificación isotérmica o amplificación no isotérmica. En algunas modalidades, después de la formación del complejo de hibridación y la asociación de la sonda de amplificación, el complejo de hibridación se amplifica por círculo rodante para generar una nanobola de ADNc (es decir, amplicón) que contiene múltiples copias del ADNc. Las técnicas para la amplificación por círculo rodante (RCA) se conocen en la técnica tal como RCA lineal, RCA ramificada, RCA dendrítica, o cualquiera de sus combinaciones. (Véase, por ejemplo, Baner y otros, *Nucleic Acids Research*, 26:5073-5078, 1998; Lizardi y otros, *Nature Genetics* 19:226, 1998; Mohsen y otros, *Acc Chem Res.* 2016 noviembre 15; 49(11): 2540-2550; Schweitzer y otros, *Proc. Natl Acad. Sci. USA* 97:101 13- 19, 2000; Faruqi y otros, *BMC Genomics* 2:4, 2000; Nallur y otros, *Nucl. Acids Res.* 29:el 18, 2001; Dean y otros, *Genome Res.* 11:1095- 1099, 2001; Schweitzer y otros, *Nature Biotech.* 20:359-365, 2002; patente de Estados Unidos Núms. 6,054,274, 6,291,187, 6,323,009, 6,344,329 y 6,368,801). Las polimerasas ilustrativas para usar en la RCA comprenden ADN polimerasa tal como polimerasa phi29 (ϕ 29), fragmento de Klenow, ADN polimerasa de *Bacillus stearothermophilus* (BST), ADN polimerasa T4, ADN polimerasa T7, o ADN polimerasa I. En algunos aspectos, pueden

emplearse ADN polimerasas que se han manipulado genéticamente o mutado para tener características convenientes. En algunas modalidades, la polimerasa es la ADN polimerasa phi29.

En algunos aspectos, durante la etapa de amplificación, pueden añadirse nucleótidos modificados a la reacción para incorporar los nucleótidos modificados en el producto de la amplificación (por ejemplo, nanobola). Los nucleótidos modificados ilustrativos comprenden nucleótidos modificados con amina. En algunos aspectos de los métodos, por ejemplo, para anclar o reticular el producto de la amplificación que se genera (por ejemplo, nanobola) a un andamio, a estructuras celulares y/o a otros productos de amplificación (por ejemplo, otras nanobolas). En algunos aspectos, los productos de la amplificación comprenden un nucleótido modificado, tal como un nucleótido modificado con amina. En algunas modalidades, el nucleótido modificado con amina comprende una modificación con una porción de N-hidroxisuccinimida de ácido acrílico. Los ejemplos de otros nucleótidos modificados con amina comprenden, pero no se limitan a, una modificación con la porción 5-aminoalil-dUTP, una modificación con la porción 5-propargilamino-dCTP, una modificación con la porción N6-6-aminohekil-dATP, o una modificación con la porción 7-deaza-7-propargilamino-dATP.

En algunos aspectos, los polinucleótidos y/o el producto de la amplificación (por ejemplo, amplicón) pueden anclarse a una matriz polimérica. Por ejemplo, la matriz polimérica puede ser un hidrogel. En algunas modalidades, una o más de las sondas de polinucleótidos pueden modificarse para contener grupos funcionales que pueden usarse como un sitio de anclaje para unir las sondas de polinucleótidos y/o el producto de amplificación a una matriz polimérica.

La modificación ilustrativa y la matriz polimérica que pueden emplearse de acuerdo con las modalidades que se proporcionan comprenden las descritas en, por ejemplo, los documentos WO 2014/163886, WO 2017/079406, US 2016/0024555, US 2018/0251833 y WO2014/025392. En algunos ejemplos, el andamio también contiene modificaciones o grupos funcionales que pueden reaccionar con o incorporar las modificaciones o grupos funcionales del conjunto de sondas o el producto de la amplificación. En algunos ejemplos, el andamio puede comprender oligonucleótidos, polímeros o grupos químicos, para proporcionar una matriz y/o estructuras de soporte.

Los productos de la amplificación pueden inmovilizarse dentro de la matriz generalmente en la ubicación del ácido nucleico que se amplifica, de esta manera se crea una colonia de amplicones localizada. Los productos de la amplificación pueden inmovilizarse dentro de la matriz mediante factores estéricos. Los productos de la amplificación también pueden inmovilizarse dentro de la matriz mediante unión covalente o no covalente. De esta manera, los productos de la amplificación pueden considerarse unidos a la matriz. Al inmovilizarse a la matriz, tal como mediante unión covalente o reticulación, se mantiene el tamaño y la relación espacial de los amplicones originales. Al inmovilizarse a la matriz, tal como mediante la unión covalente o reticulación, los productos de la amplificación son resistentes al movimiento o desenrollado bajo esfuerzo mecánico.

En algunos aspectos, los productos de la amplificación se polimerizan conjuntamente y/o se unen covalentemente a la matriz circundante, de esta manera se preserva su relación espacial y cualquier información inherente a la misma. Por ejemplo, si los productos de la amplificación son aquellos que se generan a partir de ADN o ARN dentro de una célula que se incluye en la matriz, los productos de la amplificación también pueden funcionalizarse para formar una unión covalente a la matriz, lo que conserva su información espacial dentro de la célula, de esta manera se proporciona un patrón de distribución de localización subcelular. En algunas modalidades, los métodos que se proporcionan implican incluir el uno o más conjuntos de sondas de polinucleótidos y/o los productos de la amplificación en presencia de subunidades de hidrogel para formar uno o más productos de la amplificación que se incluyen en hidrogel. En algunas modalidades, la química hidrogel-tejido descrita comprende unir covalentemente los ácidos nucleicos al hidrogel sintetizado in situ para la clarificación de tejidos, la difusión enzimática, y la secuenciación de múltiples ciclos mientras que un método de química hidrogel-tejido existente no lo hace. En algunas modalidades, para permitir la inclusión del producto de la amplificación en la configuración tejido-hidrogel, están comprendidos en la etapa de amplificación (por ejemplo, RCA) nucleótidos modificados con amina, funcionalizados con una porción de acrilamida mediante el uso de ésteres de N-hidroxisuccinimida del ácido acrílico, y copolimerizados con monómeros de acrilamida para formar un hidrogel.

iv. Detección y análisis

En algunas modalidades, la secuenciación puede realizarse mediante secuenciación por síntesis (SBS). En algunas modalidades, un cebador para secuenciación es complementario a secuencias en o cerca del uno o más códigos de barras. En tales modalidades, la secuenciación por síntesis puede comprender la transcripción inversa y/o amplificación para generar una secuencia molde a partir de la cual puede unirse una secuencia del cebador. Métodos de SBS ilustrativos comprenden los descritos, por ejemplo, pero no limitados a, los documentos US 2007/0166705, US 2006/0188901, US 7,057,026, US 2006/0240439, US 2006/0281109, WO 05/065814, US 2005/0100900, WO 06/064199, WO07/010,251, US 2012/0270305, US 2013/0260372, y US 2013/0079232.

En algunas modalidades, la secuenciación puede realizarse mediante hibridación fluorescente secuencial (por ejemplo, secuenciación mediante hibridación). La hibridación fluorescente secuencial puede implicar la hibridación secuencial de sondas de detección que comprenden un oligonucleótido y un marcador detectable.

En algunas modalidades, la secuenciación puede realizarse mediante el uso de secuenciación mediante ligación de molécula única. Tales técnicas utilizan una ADN ligasa para incorporar oligonucleótidos e identificar la incorporación de tales oligonucleótidos. Los oligonucleótidos tienen típicamente diferentes marcadores que se correlacionan con la identidad de un nucleótido en particular en una secuencia con la que se hibridan los oligonucleótidos. Los aspectos y elementos implicados en la secuenciación mediante ligación se describen, por ejemplo, en Shendure y otros, *Science* (2005), 309: 1728-1732, y en los documentos US 5,599,675; US 5,750,341; US 6,969,488; US 6,172,218; US y 6,306,597.

En algunas modalidades, los códigos de barras de las sondas de detección se seleccionan mediante los oligonucleótidos de detección marcados de manera detectable, tales como oligos marcados con fluorescencia. En algunas modalidades, se usan uno o más esquemas de decodificación para decodificar las señales, tales como la fluorescencia, para la determinación de la secuencia. En cualquiera de las modalidades en la presente descripción, los códigos de barras (por ejemplo, secuencias de código de barras primarias y/o secundarias) pueden analizarse (por ejemplo, detectarse o secuenciarse) mediante el uso de cualquier método o técnica adecuado, que comprende los descritos en la presente descripción, tales como sondeos secuenciales de ARN de dianas (SPOT de ARN), hibridación *in situ* fluorescente secuencial (seqFISH), hibridación *in situ* fluorescente de molécula única (smFISH), hibridación *in situ* con fluorescencia resistente a error multiplexada (MERFISH), secuenciación *in situ* basada en hibridación (HybISS), secuenciación *in situ*, secuenciación dirigida *in situ*, secuenciación *in situ* fluorescente (FISSEQ), o mapeo de lectura de amplicones de transcritos resueltos espacialmente (STARmap). En algunas modalidades, los métodos que se proporcionan en la presente descripción comprenden analizar los códigos de barras mediante hibridación secuencial y detección con una pluralidad de sondas marcadas (por ejemplo, oligonucleótidos de detección). Esquemas de decodificación ilustrativos se describen en *Enget* y otros, "Transcriptome-scale Super-Resolved Imaging in Tissues by RNA SeqFISH+", *Nature* 568(7751):235-239 (2019); Chen y otros, "Spatially resolved, highly multiplexed RNA profiling in single cells", *Science*, 348(6233):aa6090 (2015); los documentos US 10,457,980 B2; US 2016/0369329 A1; WO 2018/026873 A1; y US 2017/0220733 A1. En algunas modalidades, estos ensayos permiten la amplificación de la señal, la decodificación combinatoria, y los esquemas de corrección de errores al mismo tiempo. En algunas modalidades, puede usarse la hibridación de ácido nucleico para la secuenciación. Estos métodos utilizan sondas decodificadoras de ácido nucleico marcadas que son complementarias a al menos una porción de una secuencia de código de barras. La decodificación multiplexada puede realizarse con mezclas de muchas sondas diferentes con marcadores distinguibles. Ejemplos no limitantes de secuenciación mediante hibridación de ácido nucleico se describen, por ejemplo, en el documento US 8,460,865, y en Gunderson y otros, *Genome Research* 14:870-877 (2004).

En algunas modalidades, puede usarse el monitoreo en tiempo real de la actividad de la ADN polimerasa durante la secuenciación. Por ejemplo, las incorporaciones de nucleótidos pueden detectarse a través de transferencia de energía por resonancia de fluorescencia (FRET), como se describió, por ejemplo, en Levene y otros, *Science* (2003), 299, 682-686, Lundquist y otros, *Opt. Lett.* (2008), 33, 1026-1028, y el término "perfectamente y otros, *Proc. Natl. Acad. Sci. USA* (2008), 105, 1176-1181.

En algunos aspectos, el análisis y/o la determinación de secuencias pueden realizarse a temperatura ambiente para una mejor conservación de la morfología del tejido con bajo ruido de fondo y reducción de errores. En algunas modalidades, el análisis y/o la determinación de secuencias comprenden eliminar la acumulación de errores a medida que avanza la secuenciación.

En algunas modalidades, el análisis y/o la determinación de secuencias implica el lavado para eliminar los polinucleótidos no unidos, revelando posteriormente un producto fluorescente para la obtención de imágenes.

En algunos aspectos, la detección (que comprende la obtención de imágenes) se realiza mediante el uso de cualquiera de una serie de tipos diferentes de microscopía, por ejemplo, microscopía confocal, microscopía de dos fotones, microscopía de campo claro, microscopía de expansión de tejido intacto, y/o microscopía de lámina de luz optimizada con CLARITY™ (COLM).

En algunas modalidades, se usa la microscopía de fluorescencia para la detección y la obtención de imágenes de la sonda de detección. En algunos aspectos, un microscopio de fluorescencia es un microscopio óptico que usa fluorescencia y fosforescencia en lugar de, o además de, la reflexión y la absorción para estudiar las propiedades de sustancias orgánicas o inorgánicas. En la microscopía de fluorescencia, una muestra se ilumina con luz de una longitud de onda que excita la fluorescencia en la muestra. A continuación, se obtienen imágenes de la luz fluorescente, que está normalmente a una longitud de onda más larga que la iluminación, a través de un objetivo de microscopio. En esta técnica pueden usarse dos filtros; un filtro de iluminación (o excitación) que asegura que la iluminación esté cerca de la monocromática y a la longitud de onda correcta, y un segundo filtro de emisión (o barrera) que asegura que ninguna fuente de luz de excitación alcance el detector. Alternativamente, ambas funciones pueden llevarse a cabo mediante un único filtro dicróico. El "microscopio de fluorescencia" comprende cualquier microscopio que usa fluorescencia para generar una imagen, ya sea una configuración más simple como un microscopio de epifluorescencia, o un diseño más complicado tal como un microscopio confocal, que usa seccionamiento óptico para obtener una mejor resolución de la imagen fluorescente.

En algunas modalidades, la microscopía confocal se usa para la detección y obtención de imágenes de la sonda de detección. La microscopía confocal usa iluminación puntual y una apertura espacial en un plano óptico conjugado en frente del detector para eliminar la señal desenfocada. Como solo puede detectarse la luz producida por fluorescencia muy cercana al plano focal, la resolución óptica de la imagen, particularmente en la dirección de la profundidad de la muestra, es mucho mejor que la de los microscopios de campo amplio. Sin embargo, dado que gran parte de la luz de la fluorescencia de la muestra se bloquea en la apertura espacial, este aumento de la resolución se produce a costa de la disminución de la intensidad de la señal, por lo que a menudo se requieren exposiciones largas. Como solo se ilumina un punto en la muestra a la vez, la obtención de imágenes 2D o 3D requiere un barrido sobre una trama regular (es decir, un patrón rectangular de líneas de barrido paralelas) en el espécimen. El grosor alcanzable del plano focal se define principalmente por la longitud de onda de la luz usada dividida por la abertura numérica de la lente del objetivo, pero también por las propiedades ópticas del espécimen. El seccionamiento óptico fino hace posible que estos tipos de microscopios sean particularmente buenos en la obtención de imágenes 3D y la caracterización superficial de muestras. La microscopía de lámina de luz optimizada con CLARITY™ (COLM) proporciona una microscopía alternativa para la obtención rápida de imágenes 3D de muestras clarificadas grandes. COLM pesquiza tejidos inmunoteñidos grandes, permite el aumento de la velocidad de adquisición y da como resultado una mayor calidad de los datos generados.

Otros tipos de microscopía que pueden emplearse comprenden microscopía de campo claro, microscopía de iluminación oblicua, microscopía de campo oscuro, contraste de fases, microscopía de contraste de interferencia diferencial (DIC), microscopía de reflexión de interferencia (también conocida como contraste de interferencia reflejado, o RIC), microscopía de iluminación de plano único (SPIM), microscopía de superresolución, microscopía láser, microscopía electrónica (EM), microscopía electrónica de transmisión (TEM), microscopía electrónica de barrido (SEM), microscopía electrónica de reflexión (REM), microscopía electrónica de transmisión de barrido (STEM) y microscopía electrónica de bajo voltaje (LVEM), microscopía de sonda de barrido (SPM), microscopía de fuerza atómica (ATM), microscopía de emisión de electrones balísticos (BEEM), microscopía de fuerza química (CFM), microscopía de fuerza atómica conductora (C-AFM), microscopio de efecto túnel electroquímico de barrido (ECSTM), microscopía de fuerza electrostática (EFM), microscopio de fuerza de fluidos (FluidFM), microscopía de modulación de fuerza (FMM), microscopía de sonda de barrido orientada a funciones (FOSPM), microscopía de fuerza de sonda kelvin (KPFM), microscopía de fuerza magnética (MFM), microscopía de fuerza de resonancia magnética (MRFM), microscopía óptica de barrido de campo cercano (NSOM) (o SNOM, microscopía óptica de campo cercano de barrido, microscopía de fuerza de respuesta piezoeléctrica (PFM), microscopía de efecto túnel de barrido de fotones (PSTM), microscopía/microespectroscopía fototérmica (PTMS), microscopía de capacitancia de barrido (SCM), microscopía electroquímica de barrido (SECM), microscopía de puerta de barrido (SGM), microscopía de sonda Hall de barrido (SHPM), microscopía de conductancia iónica de barrido (SICM), microscopía de efecto túnel de barrido polarizado en espín (SPSM), microscopía de resistencia a la dispersión de barrido (SSRM), microscopía térmica de barrido (SThM), microscopía de efecto túnel de barrido (STM), potenciometría de efecto túnel de barrido (STP), microscopía de voltaje de barrido (SVM), y microscopía de efecto túnel de barrido con rayos X de sincrotrón (SXSTM), y microscopía de expansión de tejido intacto (exM).

V. Módulos de ensayo espacial

En algunos aspectos, un ensayo espacial integrado *in situ* descrito en la presente descripción comprende uno o más módulos de ensayo espacial. En algunas modalidades del ensayo espacial integrado *in situ*, se realizan uno o más módulos de ensayo *in situ* en una muestra. Por ejemplo, la muestra puede someterse solo a un módulo de ensayo *in situ*, tal como un ensayo de hibridación *in situ* fluorescente secuencial o un ensayo de secuenciación *in situ*. En algunas modalidades, la muestra puede someterse a dos o más módulos de ensayo *in situ*, tales como un ensayo de hibridación *in situ* fluorescente secuencial seguido de un ensayo de secuenciación *in situ*, antes de someter la misma muestra a uno o más módulos de ensayo espacial.

En un aspecto, en la presente descripción se proporcionan métodos, composiciones, aparatos y sistemas para el análisis espacial de una muestra biológica, por ejemplo, un análisis basado en matriz espacial. Los aspectos no limitantes de las metodologías del análisis espacial se describen en la publicación de patente de Estados Unidos núm. 10,308,982; publicación de patente de Estados Unidos núm. 9,879,313; publicación de patente de Estados Unidos núm. 9,868,979; Liu y otros, *bioRxiv* 788992, 2020; publicación de patente de Estados Unidos núm. 10,774,372; publicación de patente de Estados Unidos núm. 10,774,374; documento WO 2018/091676; publicación de patente de Estados Unidos núm. 10,030,261; publicación de patente de Estados Unidos núm. 9,593,365; patente de Estados Unidos núm. 10,002,316; patente de Estados Unidos núm. 9,727,810; publicación de patente de Estados Unidos núm. 10,640,816; Rodriques y otros, *Science* 363(6434):1463-1467, 2019; documento WO 2018/045186; Lee y otros, *Nat. Protoc.* 10(3):442-458, 2015; publicación de patente de Estados Unidos núm. 10,179,932; documento WO 2018/045181; publicación de patente de Estados Unidos núm. 10,138,509; Trejo y otros, *PLoS ONE* 14(2):e0212031, 2019; publicación de solicitud de patente de Estados Unidos núm. 2018/0245142; Chen y otros, *Science* 348(6233):aaa6090, 2015; Gao y otros, *BMC Biol.* 15:50, 2017; documentos WO 2017/144338; WO 2018/107054; WO 2017/222453; WO 2019/068880; WO 2011/094669; Patente de Estados Unidos núm. 7,709,198; Patente de Estados Unidos núm. 8,604,182; Patente de Estados Unidos núm. 8,951,726; Patente de Estados Unidos núm. 9,783,841; Patente de Estados Unidos núm. 10,041,949; documentos WO 2016/057552; WO 2017/147483; publicación de patente de Estados Unidos núm. 10,370,698; documento WO 2016/166128; publicación de patente de Estados Unidos

núm. 10,364,457; publicación de patente de Estados Unidos núm. 10,317,321; documentos WO 2018/136856; WO 2019/075091; Patente de Estados Unidos núm. 10,059,990; documentos WO 2018/057999; WO 2015/161173; y Gupta y otros, *Nature Biotechnol.* 36:1197-1202, 2018, y puede usarse en la presente descripción en cualquier combinación. Los aspectos adicionales no limitantes de las metodologías de análisis espacial se describen en la

En algunas modalidades, un método descrito en la presente descripción comprende la transferencia de uno o más analitos de una muestra biológica a una matriz de elementos sobre un sustrato, cada uno de los cuales se asocia con una ubicación espacial única sobre la matriz. Cada elemento puede comprender una pluralidad de agentes de captura capaces de capturar una o más moléculas de ácido nucleico, y cada uno de los agentes de captura del mismo elemento puede comprender un código de barras espacial correspondiente a una ubicación espacial única del elemento sobre 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. La ubicación espacial de cada analito dentro de la muestra se determina sobre la base del elemento al que se une cada analito en la matriz, y la ubicación espacial relativa del elemento dentro de la matriz.

En algunas modalidades, el ensayo espacial se realiza para analizar uno o más analitos (por ejemplo, el/los segundo(s) ácido(s) nucleico(s) diana). En algunas modalidades, el segundo ácido nucleico diana es o comprende el primer ácido nucleico diana (por ejemplo, ADN o ARN), una secuencia de este, un complemento de este, un producto de hibridación de este, un producto de ligación de este, un producto de extensión de este, un producto de replicación de este, un producto de transcripción/transcripción inversa de este, y/o un producto de amplificación de este (por ejemplo, un producto de amplificación por círculo rodante (RCA)). Por ejemplo, un producto de transcripción inversa que se genera mediante el uso del primer ácido nucleico diana como un molde previo a, o durante el ensayo *in situ* (por ejemplo, descrito en la sección IV) puede analizarse en el ensayo espacial. En algunos casos, un producto de ligación que se genera previo a, o durante el ensayo *in situ* puede analizarse en el ensayo espacial. En algunas modalidades, el segundo ácido nucleico diana es o comprende al menos una de la una o más sondas de ácido nucleico, un complemento de la(s) sonda(s) de ácido nucleico, un producto de hibridación de la(s) sonda(s) de ácido nucleico, un producto de ligación de la(s) sonda(s) de ácido nucleico, un producto de extensión de la(s) sonda(s) de ácido nucleico, un producto de replicación de la(s) sonda(s) de ácido nucleico, un producto de transcripción/transcripción inversa de la(s) sonda(s) de ácido nucleico, y/o un producto de amplificación de la(s) sonda(s) de ácido nucleico. Por ejemplo, una sonda o código de barras de esta que se proporciona previo a, o durante el ensayo *in situ* (por ejemplo, descrito en la sección IV) puede analizarse en el ensayo espacial.

En algunas modalidades, un método descrito en la presente descripción comprende la asociación de un código de barras espacial con uno o más analitos, por ejemplo, moléculas en una o más células tales como células vecinas, de manera que el código de barras espacial identifica el uno o más analitos, y/o el contenido de la una o más células, como asociado con una ubicación espacial particular.

En algunas modalidades, un método descrito en la presente descripción comprende conducir los analitos diana fuera de una célula y hacia una matriz con código de barras espacial. La figura 7 representa una modalidad ilustrativa, donde la matriz con código de barras espacial poblada con sondas de captura (como se describe adicionalmente en la presente descripción) se pone en contacto con una muestra en la etapa 701, y la muestra se permeabiliza, lo que permite que el analito diana migre lejos de la muestra y hacia la matriz. El analito diana interactúa con una sonda de captura sobre la matriz con código de barras espacial en la etapa 702. Una vez que el analito diana se une (por ejemplo, se hibrida) con la sonda de captura, la muestra se elimina opcionalmente de la matriz y las sondas de captura se analizan para obtener información del analito resuelto espacialmente en la etapa 703.

En algunas modalidades, un método descrito en la presente descripción comprende suministrar o conducir moléculas de ácido nucleico con código de barras espacial (por ejemplo, etiquetas de dirección o sondas de captura) a y/o hacia dentro o sobre una muestra. En algunas modalidades, un método descrito en la presente descripción comprende escindir moléculas de ácido nucleico con código de barras espacial (por ejemplo, etiquetas de dirección o sondas de captura) de una matriz y conducir las moléculas de ácido nucleico escindidas a y/o hacia dentro o sobre una muestra. La figura 8 representa una modalidad ilustrativa, donde la matriz con código de barras espacial poblada con sondas de captura (como se describe adicionalmente en la presente descripción) puede ponerse en contacto con una muestra en la etapa 801. Las sondas de captura con código de barras espacial se escinden y después interactúan con las células dentro de la muestra que se proporciona en la etapa 802. La interacción puede ser una interacción covalente o no covalente tal como una interacción de la superficie y la célula. La interacción puede ser una interacción intracelular facilitada por un sistema de suministro o un péptido de penetración celular. Alternativamente, la muestra puede permeabilizarse y fijarse/reticularse para restringir la movilidad de uno o más analitos diana, mientras que permite que las sondas de captura con código de barras espacial migren hacia y/o dentro o sobre la muestra. Una vez que la sonda de captura con código de barras espacial se asocia con un analito particular (por ejemplo, analitos en una o más células) la muestra puede eliminarse opcionalmente para el análisis. La muestra puede disociarse opcionalmente antes del análisis. Una vez que el analito o célula etiquetado se asocia con la sonda de captura con código de barras espacial, las sondas de captura pueden analizarse para obtener información resuelta espacialmente acerca del analito o la célula etiquetada en la etapa 803.

La figura 9 muestra un flujo de trabajo ilustrativo que incluye preparar una muestra sobre una matriz con código de barras espacial en la etapa 901. La preparación de la muestra puede incluir colocar la muestra sobre un portaobjetos, procesar la muestra (por ejemplo, fijar y/o permeabilizar la muestra), y/o poner en contacto la muestra con uno o más reactivos (por ejemplo, una o más sondas y/o anticuerpos para la tinción) para la obtención de imágenes. Después, opcionalmente puede obtenerse una imagen de la muestra sobre la matriz en la etapa 902 mediante el uso tanto de las modalidades de campo claro (para obtener una imagen de la tinción con hematoxilina y eosina) y de fluorescencia (para obtener una imagen de los elementos). En algunas modalidades, después los analitos diana se liberan de la muestra y las sondas de captura que forman la matriz con código de barras espacial hibridan o se unen a los analitos diana que se liberan en la etapa 903. Después, la muestra se elimina de la matriz en la etapa 904 y las sondas de captura se escinden de la matriz en la etapa 905. Después, de la muestra y de la matriz se obtienen imágenes opcionalmente una segunda vez en ambas modalidades en la etapa 905B mientras se procesan los analitos (por ejemplo, mediante transcripción inversa para convertir analitos de ARN tales como transcritos de ARNm en ADNc), y se prepara una biblioteca de amplicones en 906 y se secuencian en 907. En algunas modalidades, los dos conjuntos de imágenes se superponen espacialmente para correlacionar la información de la muestra identificada espacialmente en 908. Cuando no se obtienen imágenes de la muestra y de la matriz por segunda vez en 905B, se suministra un archivo de coordenadas de puntos en lugar de reemplazar la segunda etapa de obtención de imagen. En algunas modalidades, la preparación de la biblioteca de amplicones 906 puede realizarse con un adaptador de PCR (por ejemplo, un adaptador de PCR único) y secuenciarse en 907.

La figura 10 muestra otro flujo de trabajo ilustrativo que utiliza una matriz marcada espacialmente sobre un sustrato, donde las sondas de captura marcadas con códigos de barras espaciales se agrupan en las áreas (por ejemplo, elementos) sobre el sustrato. Las sondas de captura marcadas espacialmente pueden incluir un dominio de captura, un código de barras espacial, y opcionalmente un dominio de escisión, una o más secuencias funcionales, y/o un identificador molecular único. Las sondas de captura marcadas espacialmente pueden incluir, además, una modificación del extremo 5' para la unión reversible al sustrato. Por ejemplo, la matriz con código de barras espacial se pone en contacto con una muestra en 1001, y la muestra se permeabiliza mediante la aplicación de reactivos de permeabilización en 1002. En algunas modalidades, se realiza una etapa de permeabilización previo a, o durante un módulo de ensayo *in situ* y la muestra biológica no necesita permeabilizarse nuevamente para el módulo de ensayo espacial. Los reactivos de permeabilización pueden administrarse al colocar el conjunto de matriz/muestra dentro de una solución a granel. Alternativamente, los reactivos de permeabilización pueden administrarse a la muestra mediante un medio resistente a la difusión y/o una barrera física tal como una tapa, en donde la muestra se intercala entre el medio resistente a la difusión y/o la barrera y el sustrato que contiene la matriz. Los analitos se hacen migrar hacia la matriz de captura con código de barras espacial mediante el uso de cualquier número de técnicas descritas en la presente descripción. Por ejemplo, la migración de analitos puede ocurrir mediante el uso de una tapa de medio resistente a la difusión y una migración pasiva. Como otro ejemplo, la migración del analito puede ser una migración activa, mediante el uso de un sistema de transferencia electroforética, por ejemplo. Una vez que los analitos están muy cerca de las sondas de captura con código de barras espacial, las sondas de captura pueden hibridarse o unirse de cualquier otra manera a un analito diana en 1003. La muestra puede eliminarse opcionalmente de la matriz en 1004. Las sondas de captura pueden escindirse opcionalmente de la matriz en 1005, y los analitos capturados pueden etiquetarse espacialmente mediante la realización de una reacción de transcriptasa inversa del ADNc de la primera hebra. Una reacción de ADNc de la primera hebra puede realizarse opcionalmente mediante el uso de oligonucleótidos que cambian de molde. Por ejemplo, un oligonucleótido que cambia de molde puede hibridarse con una cola poli(C) que se añade a un extremo 3' del ADNc mediante una enzima transcriptasa inversa. Después, el molde de ARNm original y el oligonucleótido que cambia de molde pueden desnaturalizarse del ADNc y la sonda de captura con código de barras puede hibridarse después con el ADNc y puede generarse un complemento del ADNc. Después, el ADNc de la primera hebra puede purificarse y recolectarse para etapas de amplificación aguas abajo. El ADNc de la primera hebra puede amplificarse mediante el uso de PCR en 1006, en donde los cebadores directos e inversos flanquean el código de barras espacial y las regiones del analito diana de interés, lo que genera una biblioteca asociada con un código de barras espacial particular. En algunas modalidades, el ADNc comprende una secuencia de cebador de secuenciación por síntesis (SBS). Los amplicones de la biblioteca se secuencian y se analizan para decodificar la información espacial en 1007 y 1008.

Las figuras 11A-11B, representan un flujo de trabajo ilustrativo donde la muestra se elimina de la matriz con códigos de barras espaciales y las sondas de captura con códigos de barras espaciales se eliminan de la matriz para la amplificación del analito con códigos de barras y la preparación de bibliotecas. En algunas modalidades, un método descrito en la presente descripción comprende realizar la síntesis de la primera hebra, por ejemplo, mediante el uso de oligonucleótidos que cambian de molde sobre la matriz con códigos de barras espaciales sin escindir las sondas de captura.

En algunas modalidades, el segundo analito que se analiza en un módulo de ensayo espacial comprende un ARN tal como transcritos de ARNm. En algunas modalidades, un método descrito en la presente descripción comprende la preparación de la muestra 1101 (por ejemplo, que comprende una etapa de eliminación de las reticulaciones) y opcionalmente la permeabilización de la muestra 1102, ambas pueden realizarse como se describe en otra parte de la presente descripción, por ejemplo, en relación con un módulo de ensayo *in situ* y/o un módulo de ensayo espacial. Como se muestra en la figura 11A, una vez que las sondas de captura capturan ARN dianas (por ejemplo, transcritos de ARNm de una muestra que se ha analizado en un módulo de ensayo *in situ* descrito en la presente descripción), el

ADNc de la primera hebra que se crea mediante cambio de molde y transcriptasa inversa en 1103 se desnaturaliza después y la segunda hebra se extiende después en 1104. Después, el ADNc de la segunda hebra se desnaturaliza a partir del ADNc de la primera hebra, se neutraliza y se transfiere a un tubo en 1105. La cuantificación y amplificación del ADNc puede realizarse mediante el uso de técnicas estándar descritas en la presente descripción. Después, el ADNc puede someterse a la preparación de la biblioteca en 1106 e indexado en 1107, que incluye etapas de fragmentación, reparación de los extremos, cola A y/o PCR de indexado, seguido de una etapa opcional de QC de la biblioteca en 1108.

En algunas modalidades, el segundo analito que se analiza en un módulo de ensayo espacial comprende un ADN, tal como un ADNc de un ARNm que se genera en un módulo de ensayo *in situ*, una sonda que se hibrida con el ARNm o el ADNc en el módulo de ensayo *in situ*, un producto de ligación de la sonda en el módulo de ensayo *in situ* (por ejemplo, mediante el uso de una ligación con molde de ADN o una ligación con molde de ARN, por ejemplo, como se muestra en la figura 13), y/o un producto de amplificación del ARNm, ADNc o sonda en el módulo de ensayo *in situ*. En algunas modalidades, el segundo analito (por ejemplo, el segundo analito de ácido nucleico) comprende una sonda circular o una sonda circularizada de la ligación de una sonda candado que se usa en un módulo de ensayo *in situ* y/o un producto de RCA de la sonda. En algunas modalidades, un método descrito en la presente descripción comprende la preparación de muestras 1109 (por ejemplo, que comprende una etapa de eliminación de las reticulaciones para eliminar las reticulaciones de los ARNm, ADNc, sondas, productos de ligación y/o productos de RCA del módulo de ensayo *in situ*) y opcionalmente la permeabilización de la muestra 1110, ambos de los cuales pueden realizarse como se describe en otra parte de la presente descripción, por ejemplo, en relación con un módulo de ensayo *in situ* y/o un módulo de ensayo espacial. Como se muestra en la figura 11B, una vez que el segundo analito está cerca de una sonda de captura con código de barras espacial, las sondas de captura pueden hibridar o de cualquier otra manera unirse al segundo analito de ácido nucleico (por ejemplo, un ADNc, sonda, producto de ligación y/o productos de RCA del módulo de ensayo *in situ*). La muestra puede eliminarse opcionalmente de la matriz. La sonda de captura puede extenderse por una polimerasa mediante el uso de una secuencia del segundo analito de ácido nucleico capturado como molde en 1111. El dúplex que se forma mediante la sonda de captura extendida y el segundo analito de ácido nucleico capturado pueden desnaturalizarse y después se extiende una segunda hebra en 1112 mediante el uso de un segundo cebador de la hebra que se hibrida con la sonda de captura extendida. Después, la segunda hebra se desnaturaliza a partir de la sonda de captura extendida, se neutraliza y se transfiere a un tubo en 1113. La cuantificación y amplificación de ADN de la segunda hebra puede realizarse mediante el uso de técnicas estándar descritas en la presente descripción. Después, el ADN puede someterse a la preparación de la biblioteca en 1114 e indexado en 1115, que incluye etapas de fragmentación, reparación de los extremos, cola A y/o PCR de indexado, seguido de una etapa opcional de QC de la biblioteca en 1116.

Las etapas ilustrativas para la preparación de muestras, permeabilización, generación de ADN (por ejemplo, generación de ADNc de la primera hebra y generación de la segunda hebra), amplificación de ADN (por ejemplo, amplificación de ADNc) y control de calidad, y construcción de bibliotecas de expresión génica espacial se describen, por ejemplo, en los documentos WO 2020/047002, WO 2020/047004, WO 2020/047005, WO 2020/047007, y WO 2020/047010.

A. Analitos seleccionados

En algunas modalidades, los ensayos espaciales descritos en la presente descripción comprenden capturar un analito seleccionado. En algunos casos, los ensayos espaciales descritos en la presente descripción comprenden analizar un primer analito diana mediante el uso de análisis *in situ* y analizar un segundo analito diana mediante el uso de una matriz de sondas de captura. El análisis *in situ* del primer analito puede realizarse antes, simultáneamente con o después de analizar el segundo analito diana con la matriz de sondas de captura. En algunas modalidades, los analitos (o derivados de estos) se capturan mediante sondas de captura o agentes de captura después que se completa la hibridación de las sondas con la muestra y las etapas de detección durante el ensayo *in situ*. En algunas modalidades, el segundo ácido nucleico diana se selecciona por una o más sondas de ácido nucleico complementarias al segundo ácido nucleico diana (por ejemplo, ARN, tal como ARN fijado en formalina, o una molécula de ADNc que se genera a partir de una molécula de ARN), en donde la una o más sondas, o un producto que se genera a partir de la una o más sondas, se liberan de la muestra para interactuar con la matriz de sondas de captura (por ejemplo, después del análisis *in situ*). Por ejemplo, en algunas modalidades de los métodos que se proporcionan en la presente descripción, la ligación con molde se usa para detectar la expresión génica espacial (por ejemplo, del segundo analito diana) en una muestra biológica. En algunos aspectos, las etapas de la ligación con molde incluyen la hibridación de pares de sondas (por ejemplo, sondas de ADN) a la segunda molécula de ácido nucleico diana (por ejemplo, ARN, tal como ARN fijado en formalina, o una molécula de ADNc que se genera a partir de una molécula de ARN) dentro de una sección de tejido. En algunas modalidades, los pares de sondas que se hibridan adyacentemente pueden ligarse *in situ*. En algunas modalidades, la muestra puede tratarse con uno o más reactivos (tales como RNasa H o proteinasa K para liberar productos de ligación con molde de ARN del tejido (por ejemplo, en solución)) para el análisis aguas abajo (por ejemplo, hibridados o capturados de cualquier otra manera sobre la matriz de sondas de captura). En algunas modalidades, el ensayo puede incluir además la amplificación de productos de ligación con molde (por ejemplo, mediante PCR multiplex).

En algunos aspectos, la ligación con molde (por ejemplo, la ligación con molde de ADN o ARN) puede incluir una ADN ligasa. En algunos aspectos, la ligación con molde (por ejemplo, la ligación con molde de ARN) puede incluir una ARN ligasa. En algunos aspectos, la ligación con molde puede incluir ARN ligasa T4. En algunos aspectos, la ligación con molde se usa para la detección del ARN, la determinación de la identidad de la secuencia de ARN y/o el monitoreo de la expresión y el análisis de transcritos. En algunos aspectos, la ligación con molde permite la detección de un cambio particular en un ácido nucleico (por ejemplo, una mutación, adición, delección, o polimorfismo de un único nucleótido (SNP)), detección o expresión de un ácido nucleico particular, o detección o expresión de un conjunto particular de ácidos nucleicos (por ejemplo, en una vía celular similar o expresado en una patología particular). En algunas modalidades, los métodos que incluyen ligación con molde se usan para analizar ácidos nucleicos, por ejemplo, mediante genotipado, cuantificación del número de copias de ADN o transcritos de ARN, localización de transcritos particulares dentro de las muestras y similares. En algunos aspectos, los sistemas y métodos que se proporcionan en la presente descripción, que incluyen una ligación con molde, identifican polimorfismos de un único nucleótido (SNP). En algunos aspectos, tales sistemas y métodos utilizan ligación con molde para identificar mutaciones. En algunos aspectos, tales sistemas y métodos utilizan una ligación con molde para identificar isoformas de ARN, o variantes de corte y empalme. En algunos aspectos, dos o más analitos de ARN se analizan mediante el uso de ligación con molde (por ejemplo, ligación con molde de ARN (RTL) o ligación con molde de ADN (por ejemplo, en ADNc)).

En algunos aspectos, cuando se analizan dos o más analitos, se usa una primera y segunda sonda que es específica para (por ejemplo, se hibrida específicamente con) cada analito de ARN o de ADNc. En algunos casos, un analito de proteína se analiza mediante el uso de ligación con molde. Por ejemplo, en algunas modalidades, una muestra se pone en contacto con un agente aglutinante (por ejemplo, un anticuerpo o fragmento de unión a epítipo de este) específico de un analito de interés (tal como una proteína), en donde el agente aglutinante se conjuga o se asocia de cualquier otra manera con un oligonucleótido reportero que comprende una secuencia reportera que identifica el agente aglutinante. Después, las sondas pueden hibridarse al oligonucleótido reportero y ligarse en una reacción de ligación con molde para generar un producto para su análisis posterior (por ejemplo, mediante el uso de una matriz de sondas de captura).

En algunas modalidades, un par de sondas se diseñan para hibridarse con una secuencia diana de manera que las sondas se hibridan adyacentes entre sí de manera que los extremos 5' y 3' de dos sondas adyacentes pueden ligarse. En algunas modalidades, la presencia o ausencia de la secuencia diana en la muestra biológica puede determinarse al determinar si las dos sondas se ligaron o no en presencia de ligasa. La figura 13 ilustra esquemáticamente un método representativo para analizar una molécula de ácido nucleico (por ejemplo, ADNc o ARN, tal como ARNm). Aunque se muestra una molécula de ácido nucleico que tiene una cola 3' de poli-A, se apreciará que puede analizarse cualquier molécula de ácido nucleico adecuada, y la ligación puede comprender una ligación con molde de ARN (RTL) y/o una ligación con molde de ADN (por ejemplo, en ADNc) en la figura 13. La figura 13A muestra una molécula de ácido nucleico 1300 que comprende regiones diana 1302 y 1304. En algunos casos, las regiones diana 1302 y 1304 son adyacentes entre sí. La sonda 1306 comprende la secuencia de la sonda 1308, la secuencia de unión 1310 y la porción reactiva 1312. La sonda 1314 comprende secuencias de la sonda 1316, secuencia adaptadora 1348 y porción reactiva 1318. La secuencia de la sonda 1308 de la sonda 1306 es complementaria a la región diana 1302. De manera similar, la secuencia de la sonda 1316 de la sonda 1314 es complementaria a la región diana 1304. La figura 13B muestra la secuencia de la sonda 1308 de la sonda 1306 que se hibrida con la región diana 1302 y la secuencia de la sonda 1316 de la sonda 1314 que se hibrida con la región diana 1304. En algunos casos, la porción reactiva 1312 de la sonda 1306 y la porción reactiva 1318 de la sonda 1314 son adyacentes entre sí. La figura 13C muestra la porción de enlace 1320 que se produce a través de una reacción de porciones reactivas 1312 y 1318. En algunos casos, las porciones 1312 y 1318 se ligan químicamente (por ejemplo, química clic) y, en otros casos, enzimáticamente (por ejemplo, una ligasa, tal como SplintR, ligasa KOD o ligasa T4). Las sondas unidas 1306 y 1314 comprenden una molécula de ácido nucleico que se une a una sonda 1330 que comprende las secuencias 1310, 1308, 1316 y 1348. En algunos casos, la molécula de ácido nucleico que se une a la sonda (por ejemplo, 1230) puede comprender un dominio de unión a la sonda de captura que comprende una secuencia que es complementaria al dominio de captura 1207 de una sonda de captura 1202. En algunos casos, las sondas (por ejemplo, 1306 y/o 1314) pueden comprender un dominio de unión de la sonda de captura que comprende una secuencia que es complementaria a un dominio de captura particular presente en una sonda de captura (por ejemplo, 1207 de la sonda de captura 1202). En algunas modalidades, la ligación en la presente descripción se precede por un llenado de espacio (por ejemplo, si la sonda 1306 y la sonda 1314 no son adyacentes entre sí). En algunos casos, las regiones diana 1302 y 1304 no son adyacentes entre sí y pueden separarse por 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más nucleótidos.

En algunos aspectos, se usan tres o más sondas en los métodos de ligación con molde que se proporcionan en la presente descripción. En algunas modalidades, las tres o más sondas se diseñan para hibridarse con una secuencia diana de manera que las tres o más sondas se hibridan adyacentes entre sí de manera que los extremos 5' y 3' de las sondas adyacentes pueden ligarse. En algunas modalidades, la presencia o ausencia de la secuencia diana en la muestra biológica puede determinarse al determinar si las tres o más sondas se ligaron o no en la presencia de ligasa. En algunos aspectos, la sonda es una sonda de ADN. En algunos aspectos, se usan un par de sondas de ADN. En algunas modalidades, la sonda es una sonda de doble hebra o una sonda parcialmente de doble hebra. En algunos casos, la sonda es una sonda de ADN/ARN quimérica. En algunos aspectos, la sonda es una sonda de ADN modificada con ribonucleótidos. Para las sondas ilustrativas y las sondas de ADN modificadas con ribonucleótidos,

véase, por ejemplo, Zhang y otros, Chem Commun 2013 Nov 4;49(85):10013-5; patente de Estados Unidos núm. 9,371,598; publicación de solicitud de patente de Estados Unidos núms. 2019/0367997 y 2018/0237864.

En algunos aspectos, los métodos de ligación con molde de ARN utilizan la ARN ligasa T4 2 para unir eficientemente pares de sondas de ARN-ADN quiméricas adyacentes que se hibridan in situ sobre secuencias diana de ARN fijadas. El tratamiento subsecuente opcional (por ejemplo, con proteinasa K o con RNasa H) libera productos de ligación con molde para el análisis aguas abajo (por ejemplo, captura y análisis sobre sondas de captura que se unen a la matriz). En la presente descripción se proporcionan métodos para el análisis del ácido nucleico seleccionados (por ejemplo, moléculas de ARN o ADNc que se generan a partir de estos) que incluyen una primera sonda oligonucleotídica y una segunda sonda oligonucleotídica (por ejemplo, ligación con molde, tal como sondas RTL). La primera y segunda sondas oligonucleotídicas incluyen cada una secuencias que son complementarias a la secuencia de un analito de interés. En algunas modalidades, la primera y la segunda sondas oligonucleotídicas se unen a secuencias complementarias que son adyacentes entre sí o están sobre el mismo transcrito. En algunas modalidades, las secuencias complementarias a las que se unen la primera sonda oligonucleotídica y la segunda sonda oligonucleotídica son 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, aproximadamente 15, aproximadamente 20, aproximadamente 25, aproximadamente 30, aproximadamente 35, aproximadamente 40, aproximadamente 45, aproximadamente 50, aproximadamente 55, aproximadamente 60, aproximadamente 65, aproximadamente 70, aproximadamente 75, aproximadamente 80, aproximadamente 85, aproximadamente 90, aproximadamente 95, aproximadamente 100, aproximadamente 125, aproximadamente 150, aproximadamente 175, aproximadamente 200, aproximadamente 250, aproximadamente 300, aproximadamente 350, aproximadamente 400, aproximadamente 450, aproximadamente 500, aproximadamente 600, aproximadamente 700, aproximadamente 800, aproximadamente 900, o aproximadamente 1000 nucleótidos alejados entre sí. Los espacios entre las sondas oligonucleotídicas pueden llenarse previo a la ligación, mediante el uso de, por ejemplo, Mu polimerasa, ADN polimerasa, ARN polimerasa, transcriptasa inversa, VENT polimerasa, Taq polimerasa y/o cualquiera de las combinaciones, derivados y variantes (por ejemplo, mutantes modificados genéticamente) de estos. En algunas modalidades, cuando la primera y la segunda sonda oligonucleotídica se separan entre sí por uno o más nucleótidos, los ribonucleótidos se ligan entre la primera y la segunda sonda oligonucleotídica. En algunas modalidades, cuando la primera y la segunda sondas oligonucleotídicas se separan entre sí por uno o más nucleótidos, los desoxirribonucleótidos se ligan entre la primera y la segunda sondas oligonucleotídicas. Para sondas ilustrativas y reacciones y análisis mediados por ligación, véase, por ejemplo, la publicación de solicitud de patente de Estados Unidos núm. 20200239874.

En algunos aspectos, en la presente descripción se describen métodos para detectar la expresión de ARN que incluyen entrar en contacto con una primera sonda, una segunda sonda y ligasa (por ejemplo, ARN ligasa T4). En algunas modalidades, la primera sonda y la segunda sonda se diseñan para hibridarse con una secuencia diana de manera que el extremo 5' de la primera sonda y el extremo 3' de la segunda sonda son adyacentes y pueden ligarse, en donde al menos el nucleótido 5' terminal de la primera sonda y al menos el nucleótido 3' terminal de la segunda sonda son desoxirribonucleótidos (ADN), y en donde la secuencia diana es una molécula de ARN (tal como una molécula de ARNm). En algunas modalidades, una o más de las sondas pueden comprender uno o más ribonucleótidos. Después de la hibridación, una ligasa (por ejemplo, ARN ligasa T4) liga la primera sonda y la segunda sonda si la secuencia diana está presente en la muestra diana, pero no liga la primera sonda y la segunda sonda si la secuencia diana no está presente en la muestra diana. La presencia o ausencia y la localización espacial de la secuencia diana en la muestra biológica puede determinarse mediante el análisis de los productos de la sonda ligados mediante el uso, por ejemplo, de agentes de captura que se unen a la matriz como se describe en otra parte de la presente descripción. Cualquiera de una variedad de métodos puede usarse para determinar si la primera y la segunda sondas se ligaron o no en presencia de ligasa, que incluye pero no se limita a, secuenciar el producto ligado, hibridar el producto ligado con una sonda de detección que se hibrida solo cuando la primera y la segunda sondas se ligaron en la presencia de ligasa, análisis de enzimas de restricción y otros métodos conocidos en la técnica.

En algunas modalidades, la primera y/o segunda sonda, como se describe en la presente descripción, incluye una o más de las siguientes: una o más bases de ácido ribonucleico en el extremo 3' (por ejemplo, al menos dos bases de ácido ribonucleico en el extremo 3'); una o más secuencias funcionales; un nucleótido fosforilado en el extremo 5'; y/o un dominio de unión de la sonda de captura. En algunas modalidades, la secuencia funcional es una secuencia de cebador o una secuencia de unión al cebador. En algunas modalidades, la secuencia funcional incluye una o más secuencias funcionales que pueden usarse en el procesamiento subsecuente. En algunos casos, el dominio de unión de la sonda de captura es una secuencia que es complementaria a un dominio de captura particular presente en una sonda de captura. Por ejemplo, como se muestra en la figura 13, la secuencia de unión 1310 puede ser complementaria al dominio de captura 1207 de la sonda de captura 1202 en la figura 12. En algunas modalidades, el dominio de unión de la sonda de captura incluye una secuencia poli(A). En algunas modalidades, el dominio de unión de la sonda de captura incluye una secuencia de poliuridina, una secuencia de politimidina o ambas. En algunas modalidades, el dominio de unión de la sonda de captura incluye una secuencia aleatoria (por ejemplo, un hexámero u octámero aleatorio). En algunas modalidades, el dominio de unión de la sonda de captura es complementario a un dominio de captura en una sonda de captura que detecta una(s) diana(s) particular(es) de interés. En algunas modalidades, se proporciona una porción de bloqueo del dominio de unión de la sonda de captura que interactúa con el dominio de unión de la sonda de captura. En algunas modalidades, una porción de bloqueo del dominio de unión de la sonda de captura incluye una secuencia que es complementaria o sustancialmente complementaria a un dominio de unión de la sonda de captura. En algunas modalidades, una porción de bloqueo del dominio de unión de la sonda

de captura evita que el dominio de unión de la sonda de captura se una a la sonda de captura cuando está presente. En algunas modalidades, una porción de bloqueo del dominio de unión de la sonda de captura se elimina previo a la unión del dominio de unión de la sonda de captura (por ejemplo, presente en una sonda ligada) a una sonda de captura. En algunas modalidades, una porción de bloqueo del dominio de unión de la sonda de captura comprende una secuencia de poliuridina, una secuencia de politimidina o ambas.

En algunas modalidades, la primera sonda oligonucleotídica se hibrida con un analito. En algunas modalidades, la segunda sonda oligonucleotídica se hibrida con un analito. En algunas modalidades, tanto la primera sonda oligonucleotídica como la segunda sonda oligonucleotídica se hibridan con un analito. La hibridación puede ocurrir en una diana que tiene una secuencia que es 100 % complementaria a la(s) sonda(s) oligonucleotídica(s). En algunas modalidades, la hibridación puede ocurrir en una diana que tiene una secuencia que es al menos (por ejemplo, al menos aproximadamente) 80 %, al menos (por ejemplo, al menos aproximadamente) 85 %, al menos (por ejemplo, al menos aproximadamente) 90 %, al menos (por ejemplo, al menos aproximadamente) 95 %, al menos (por ejemplo, al menos aproximadamente) 96 %, al menos (por ejemplo, al menos aproximadamente) 97 %, al menos (por ejemplo, al menos aproximadamente) 98 %, o al menos (por ejemplo, al menos aproximadamente) 99 % complementaria a la(s) sonda(s) oligonucleotídica(s). Después de la hibridación, en algunas modalidades, se extiende la primera sonda oligonucleotídica. Después de la hibridación, en algunas modalidades, se extiende la segunda sonda oligonucleotídica. En algunas modalidades, los métodos descritos en la presente descripción incluyen una etapa de lavado después de hibridar la primera y segunda sondas oligonucleotídicas. La etapa de lavado elimina cualquier oligonucleótido no unido y puede realizarse mediante el uso de cualquier técnica conocida en la técnica. En algunas modalidades, se usa un tampón de prehibridación para lavar la muestra. En algunas modalidades, se usa un tampón fosfato. En algunas modalidades, se realizan múltiples etapas de lavado para eliminar oligonucleótidos no unidos.

En algunas modalidades, después de la hibridación de la sonda oligonucleotídica (por ejemplo, la primera y la segunda sonda oligonucleotídica) con el analito, las sondas oligonucleotídicas (por ejemplo, la primera sonda oligonucleotídica y la segunda sonda oligonucleotídica) se ligan juntas, lo que crea una única sonda ligada que es complementaria al analito. La ligación puede realizarse enzimática o químicamente, como se describe en la presente descripción.

En algunos casos, la primera y la segunda sonda oligonucleotídica se hibridan con la primera y la segunda regiones diana del analito, y las sondas oligonucleotídicas se someten a una reacción de ácido nucleico para ligarlas entre sí. Por ejemplo, las sondas pueden someterse a una reacción de ligación enzimática, mediante el uso de una ligasa (por ejemplo, ARN ligasa T4 (Rnl2), una ligasa splintR, una ADN ligasa monocatenaria o una DNA ligasa T4). Véase, por ejemplo, Credle y otros, *Nucleic Acids Research*, 2017, Vol. 45, No. 14 e128 para el análisis que usa ligación mediante hibridación in situ (LISH); Zhang L., y otros; *Archaeal RNA ligase from thermococcus kodakarensis for template dependent ligation* *Biol. 2017*; 14(1): 36–44 para una descripción de la ligasa KOD. Después de la reacción de ligación enzimática, la primera y la segunda sonda oligonucleotídica pueden considerarse ligadas.

En algunas modalidades, las sondas oligonucleotídicas (por ejemplo, la primera sonda oligonucleotídica y la segunda sonda oligonucleotídica) cada una pueden comprender una porción reactiva de manera que, después de la hibridación con la diana y la exposición a condiciones de ligación apropiadas, las sondas oligonucleotídicas pueden ligarse entre sí. En algunas modalidades, la sonda oligonucleotídica que incluye una porción reactiva se liga químicamente. Por ejemplo, una sonda oligonucleotídica capaz de hibridarse con una primera región diana de una molécula de ácido nucleico puede comprender una primera porción reactiva, y una sonda oligonucleotídica capaz de hibridarse con una segunda región diana de la molécula de ácido nucleico puede comprender una segunda porción reactiva. Cuando la primera y la segunda sonda oligonucleotídica se hibridan con la primera y la segunda regiones diana de la molécula de ácido nucleico, la primera y la segunda porciones reactivas pueden ser adyacentes entre sí. Una porción reactiva de una sonda puede seleccionarse del grupo no limitante que consiste en azidas, alquinos, nitrones (por ejemplo, 1,3-nitrones), alquenos deformados (por ejemplo, cicloalquenos trans tales como ciclooctenos u oxanorbornadieno), tetrazinas, tetrazoles, yoduros, tioatos (por ejemplo, fosfortioato), ácidos, aminos y fosfatos. Por ejemplo, la primera porción reactiva de una primera sonda oligonucleotídica puede comprender una porción de azida, y una segunda porción reactiva de una segunda sonda oligonucleotídica puede comprender una porción alquino. La primera y la segunda porciones reactivas pueden reaccionar para formar una porción de enlace. Una reacción entre la primera y la segunda porciones reactivas puede ser, por ejemplo, una reacción de cicloadición tal como una cicloadición de azida-alquino promovida por cepa, una cicloadición de azida-alquino catalizada por cobre, una cicloadición de alquilo-nitrón promovida por cepa, una reacción de Diels-Alder, una cicloadición [3+2], una cicloadición [4+2] o una cicloadición [4+1]; una reacción de tiol-eno; una reacción de subestación nucleófila; u otra reacción. En algunos casos, la reacción entre la primera y la segunda porciones reactivas puede producir una porción de triazol o una porción de isoxazolina. Una reacción entre la primera y la segunda porciones reactivas puede implicar someter las porciones reactivas a condiciones adecuadas tales como una temperatura, pH o presión adecuados y proporcionar uno o más reactivos o catalizadores para la reacción. Por ejemplo, una reacción entre la primera y la segunda porciones reactivas puede catalizarse mediante un catalizador de cobre, un catalizador de rutenio, o una especie deformada tal como un difluorooctino, dibencilciclooctino, o biarilazaciclooctinona. La reacción entre una primera porción reactiva de una primera sonda oligonucleotídica que se hibrida con una primera región diana de la molécula de ácido nucleico y una segunda porción reactiva de una tercera sonda oligonucleotídica que se hibrida con una segunda región diana de la molécula de ácido nucleico puede enlazar la primera sonda oligonucleotídica y la segunda sonda oligonucleotídica para proporcionar una sonda ligada. Después del enlace, la primera y la segunda sonda oligonucleotídica pueden

considerarse ligadas. En consecuencia, la reacción de la primera y la segunda porciones reactivas puede comprender una reacción de ligación química tal como una reacción química "clic" de azida 5' catalizada por cobre a alquino 3' para formar un enlace triazol entre dos sondas oligonucleotídicas. En otros ejemplos no limitantes, una porción de yoduro puede ligarse químicamente a una porción de fosforotioato para formar un enlace de fosforotioato, un ácido puede ligarse a una amina para formar un enlace de amida, y/o un fosfato y amina pueden ligarse para formar un enlace fosforamidato.

En algunas modalidades, dos sondas RTL (izquierda (LHS) y derecha (RHS)) se ponen en contacto e hibridan con el analito (por ejemplo, un ARNm diana). Después de la ligación de las dos sondas RTL mediante el uso de una ligasa (por ejemplo, Rnl2), el ARN diana puede digerirse enzimáticamente mediante el uso de ARNasa H. En algunas modalidades, las sondas RTL se proporcionan previo a, o durante el ensayo *in situ* descrito en la sección IV. En algunas modalidades, puede realizarse una etapa de permeabilización, la sonda RTL puede liberarse y capturarse sobre una matriz espacial. En algunos aspectos, la secuencia de captura de la sonda RTL se hibrida con una sonda de captura. Después de la hibridación, la sonda RTL se extiende en el extremo 3', lo que crea una sonda RTL que tiene además una secuencia de código de barras espacial. La amplificación y secuenciación identifica la secuencia de la sonda RTL y su ubicación (por ejemplo, mediante el uso de la secuencia de código de barras espacial).

B. Agentes de captura (por ejemplo, sondas de captura)

Una sonda de captura o agente de captura en la presente descripción puede comprender cualquier molécula capaz de capturar (directa o indirectamente) y/o marcar un analito de interés en una muestra biológica (por ejemplo, un segundo ácido nucleico diana). En algunas modalidades, la sonda de captura es un ácido nucleico o un polipéptido. En algunas modalidades, la sonda de captura es un conjugado (por ejemplo, un conjugado oligonucleótido-anticuerpo). 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, los analitos en una muestra biológica pueden procesarse previo a la interacción con una sonda de captura. Por ejemplo, previo a la interacción con las sondas de captura, las reacciones de polimerización catalizadas por una polimerasa (por ejemplo, ADN polimerasa o transcriptasa inversa) se realizan en la muestra biológica. En algunas modalidades, un cebador para la reacción de polimerización incluye un grupo funcional que mejora la hibridación con la sonda de captura. Las sondas de captura pueden incluir dominios de captura apropiados para capturar analitos biológicos de interés (por ejemplo, secuencia poli(dT) para capturar poli(A) del ARNm).

En algunas modalidades, una reacción catalizada por transcriptasa inversa (RT) tiene lugar durante la hibridación de una o más sondas de ácido nucleico a una primera diana de ácido nucleico en una muestra biológica para un módulo de ensayo *in situ*. En algunas modalidades, la reacción de RT convierte uno o más analitos de ARN en la muestra biológica en ADN para el módulo de ensayo *in situ* y/o un módulo de ensayo espacial. En algunas modalidades, la una o más sondas de ácido nucleico comprenden una sonda que se liga con otra sonda o a sí misma. Por ejemplo, una sonda candado se liga mediante el uso de una ligación con molde de ARN y/o con molde de ADN.

En algunas modalidades, una reacción catalizada por transcriptasa inversa (RT) tiene lugar después de la ligación de una sonda de ácido nucleico con otra sonda o a sí misma, en donde la sonda de ácido nucleico se hibrida con una primera diana de ácido nucleico en una muestra biológica para un módulo de ensayo *in situ*. En algunas modalidades, la reacción de RT convierte uno o más analitos de ARN en la muestra biológica en ADN para el módulo de ensayo *in situ* y/o un módulo de ensayo espacial.

En algunas modalidades, los analitos biológicos se procesan previamente para la generación de bibliotecas mediante secuenciación de nueva generación. Por ejemplo, los analitos pueden procesarse previamente mediante la adición de una modificación (por ejemplo, ligación de secuencias que permiten la interacción con sondas de captura). En algunas modalidades, los analitos (por ejemplo, ADN o ARN) se fragmentan mediante el uso de técnicas de fragmentación (por ejemplo, mediante el uso de transposasas y/o tampones de fragmentación).

La fragmentación puede continuarse con una modificación del analito. Por ejemplo, una modificación puede ser la adición a través de la ligación de una secuencia adaptadora que permite la hibridación con la sonda de captura. En algunas modalidades, donde el analito de interés es ARN, se realiza la adición de una cola poli(A). La adición de una cola poli(A) al ARN que no contiene una cola poli(A) puede facilitar la hibridación con una sonda de captura que incluye un dominio de captura con una cantidad funcional de secuencia poli(dT).

En algunas modalidades, previo a la interacción con las sondas de captura, las reacciones de ligación catalizadas por una ligasa se realizan en la muestra biológica. En algunas modalidades, la ligación puede realizarse mediante ligación química. En algunas modalidades, la ligación puede realizarse mediante el uso de la química clic como más abajo. En algunas modalidades, el dominio de captura incluye una secuencia de ADN que tiene complementariedad con una molécula de ARN, donde la molécula de ARN tiene complementariedad con una segunda secuencia de ADN, y donde la complementariedad de la secuencia de ARN-ADN se usa para ligar la segunda secuencia de ADN a la secuencia de ADN en el dominio de captura. En estas modalidades, es posible la detección directa de moléculas de ARN.

En algunas modalidades, previo a la interacción con las sondas de captura, se realizan reacciones específicas de la diana en la muestra biológica. Los ejemplos de reacciones específicas de la diana incluyen, pero no se limitan a, ligación de adaptadores, sondas y/u otros oligonucleótidos específicos de la diana, amplificación específica de la diana mediante el uso de cebadores específicos para uno o más analitos, y detección específica de la diana mediante el uso de hibridación in situ, microscopía de ADN y/o detección de anticuerpos. En algunas modalidades, una sonda de captura incluye dominios de captura dirigidos a productos específicos de la diana (por ejemplo, amplificación o ligación).

La Figura 12 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 1202 se acopla opcionalmente a un elemento 1201 mediante un dominio de escisión 1203, tal como un enlazador disulfuro. La sonda de captura puede incluir secuencias funcionales que son útiles para el procesamiento subsecuente, tal como la secuencia funcional 1204, que puede incluir una secuencia de unión a la celda de flujo específica del secuenciador, por ejemplo, una secuencia P5, así como también la secuencia funcional 1206, que puede incluir secuencias de cebadores de secuenciación, por ejemplo, un sitio de unión del cebador R1. En algunas modalidades, la secuencia 1204 comprende una secuencia P7 y una secuencia 1206 comprende un sitio de unión al cebador R2. Puede incluirse un código de barras espacial 1205 dentro de la sonda de captura para usar en la codificación mediante código de barras del analito diana. Las secuencias funcionales generalmente pueden seleccionarse para compatibilidad con cualquiera de una variedad de sistemas de secuenciación diferentes y los requisitos de estos. Los ejemplos de tales sistemas y técnicas de secuenciación, para los cuales pueden usarse secuencias funcionales adecuadas, incluyen (pero no se limitan a) secuenciación 454 de Roche, secuenciación Ion Torrent Proton o PGM, secuenciación Illumina X10, secuenciación PacBio SMRT, secuenciación Oxford Nanopore y secuenciación basada en detectores basados en CMOS (Semiconductor de óxido de metal complementario). Además, en algunas modalidades, las secuencias funcionales pueden seleccionarse para su compatibilidad con otros sistemas de secuenciación, que incluyen sistemas de secuenciación no comercializados, por ejemplo, sistemas basados en secuenciación mediante síntesis, secuenciación mediante hibridación, secuenciación mediante ligación y/o secuenciación mediante unión.

En algunas modalidades, el código de barras espacial 1205, las secuencias funcionales 1204 (por ejemplo, la secuencia de unión a la celda de flujo) y 1206 (por ejemplo, secuencias de cebadores de secuenciación) pueden ser comunes a todas las sondas que se unen a un elemento dado. El código de barras espacial puede incluir además un dominio de captura 1207 para facilitar la captura de un analito diana.

En algunas modalidades, las sondas de captura pueden comprender una o más sondas de captura escindible, en donde la sonda de captura escindida puede entrar en una célula no permeabilizada y unirse a analitos diana dentro de la muestra. La sonda de captura puede contener un dominio de escisión, un péptido que penetra en las células, una molécula reportera y un enlace disulfuro (-S-S-). En algunas modalidades, la sonda de captura puede incluir además un código de barras espacial y un dominio de captura.

i. Dominio de captura

En algunas modalidades, cada agente de captura (por ejemplo, una sonda de captura) comprende al menos un dominio de captura, que puede comprender un oligonucleótido, un polipéptido, una molécula pequeña o cualquiera de sus combinaciones, que se une específicamente a un analito deseado. En algunas modalidades, puede usarse un dominio de captura para capturar o detectar un analito deseado, tal como una molécula de ácido nucleico.

En algunas modalidades, el dominio de captura comprende una secuencia de ácido nucleico funcional que se configura para interactuar con uno o más analitos, tales como uno o más tipos diferentes de ácidos nucleicos (por ejemplo, moléculas de ARN y moléculas de ADN). En algunas modalidades, la secuencia de ácido nucleico funcional puede incluir una secuencia de N-mer (por ejemplo, una secuencia de N-mer aleatoria), cuyas secuencias de N-mer se configuran para interactuar con una pluralidad de moléculas de ácido nucleico, que incluyen moléculas de ARN y/o ADN. En algunas modalidades, la secuencia funcional puede incluir una secuencia poli(T), cuyas secuencias poli(T) se configuran para interactuar con moléculas de ARN mensajero (ARNm) mediante la cola poli(A) de un transcrito de ARNm. En algunas modalidades, la secuencia de ácido nucleico funcional comprende la diana de unión de una proteína (por ejemplo, un factor de transcripción, una proteína de unión a ADN o una proteína de unión a ARN), donde el analito de interés comprende una proteína. En algunas modalidades, un analito que no es ácido nucleico tal como un analito de proteína se conjuga directa o indirectamente a una molécula de ácido nucleico que es capaz de interactuar con un dominio de captura de un agente de captura.

Las sondas pueden incluir ribonucleótidos y/o desoxirribonucleótidos, así como también residuos de nucleótidos sintéticos que son capaces de participar en interacciones de tipo Watson-Crick o de pares de bases análogas. En algunas modalidades, el dominio de captura es capaz de imprimir una reacción de transcripción inversa para generar ADNc que es complementario a las moléculas de ARN capturadas. En algunas modalidades, el dominio de captura de la sonda de captura puede cebar una reacción de extensión de ADN (polimerasa) para generar ADN que es complementario a las moléculas de ADN capturadas. En algunas modalidades, el dominio de captura puede servir de molde a una reacción de ligación entre las moléculas de ADN capturadas y una sonda de superficie que se inmoviliza directa o indirectamente sobre el sustrato. En algunas modalidades, el dominio de captura puede ligarse a una hebra

de las moléculas de ADN capturadas. Por ejemplo, la ligasa SplintR junto con secuencias de ARN o ADN (por ejemplo, ARN degenerado) puede usarse para ligar un ADN o ARN monocatenario al dominio de captura. En algunas modalidades, las ligasas con actividad de ligasa con molde de ARN, por ejemplo, ligasa SplintR, ARN ligasa T4 2 o ligasa KOD, pueden usarse para ligar un ADN o ARN monocatenario al dominio de captura. En algunas modalidades, un dominio de captura incluye un oligonucleótido de soporte. En algunas modalidades, un dominio de captura, captura un oligonucleótido de soporte.

En algunas modalidades, el dominio de captura se ubica en el extremo 3' de la sonda de captura e incluye un extremo 3' libre que puede extenderse, por ejemplo, mediante polimerización dependiente del molde, para formar una sonda de captura extendida como se describe en la presente descripción. En algunas modalidades, el dominio de captura incluye una secuencia de nucleótidos que es capaz de hibridarse con ácido nucleico, por ejemplo ARN u otro analito, presente en las células de la muestra de tejido en contacto con la matriz. En algunas modalidades, el dominio de captura puede seleccionarse o diseñarse para unirse selectiva o específicamente a un ácido nucleico diana. Por ejemplo, el dominio de captura puede seleccionarse o diseñarse para capturar ARNm por medio de hibridación con la cola poli(A) de ARNm. Por lo tanto, en algunas modalidades, el dominio de captura incluye un oligonucleótido de ADN poli(T), es decir, una serie de residuos consecutivos de desoxitimidina que se unen por enlaces fosfodiéster, que es capaz de hibridarse con la cola de poli(A) del ARNm. En algunas modalidades, el dominio de captura puede incluir nucleótidos que son funcional o estructuralmente análogos a una cola poli(T). Por ejemplo, un oligonucleótido poli(U) o un oligonucleótido que se incluye de análogos de desoxitimidina. En algunas modalidades, el dominio de captura incluye al menos 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 nucleótidos. En algunas modalidades, el dominio de captura incluye al menos 25, 30 o 35 nucleótidos.

En algunas modalidades, pueden usarse secuencias aleatorias, por ejemplo, hexámeros aleatorios o secuencias similares, para formar todo o una parte del dominio de captura. Por ejemplo, las secuencias aleatorias pueden usarse junto con las secuencias poli(T) (o análogo poli(T)). Por lo tanto, cuando un dominio de captura incluye un oligonucleótido poli(T) (o un oligonucleótido "similar a poli(T)"), también puede incluir una secuencia de oligonucleótidos aleatoria (por ejemplo, sonda de "secuencia aleatoria de poli(T)"). Esto puede, por ejemplo, ubicarse 5' o 3' de la secuencia poli(T), por ejemplo, en el extremo 3' del dominio de captura. La sonda de secuencia aleatoria de poli(T) puede facilitar la captura de la cola de poli(A) del ARNm. En algunas modalidades, el dominio de captura puede ser una secuencia aleatoria completamente. En algunas modalidades, pueden usarse dominios de captura degenerados.

En algunas modalidades, un conjunto de dos o más sondas de captura forman una mezcla, donde el dominio de captura de una o más sondas de captura incluye una secuencia poli(T) y el dominio de captura de una o más sondas de captura incluye secuencias aleatorias. En algunas modalidades, un conjunto de dos o más sondas de captura forman una mezcla donde el dominio de captura de una o más sondas de captura incluye una secuencia similar a poli(T) y el dominio de captura de una o más sondas de captura incluye secuencias aleatorias. En algunas modalidades, un conjunto de dos o más sondas de captura forman una mezcla donde el dominio de captura de una o más sondas de captura incluye secuencias poli(T) aleatorias y el dominio de captura de una o más sondas de captura incluye secuencias aleatorias. En algunas modalidades, las sondas con dominios de captura degenerados pueden añadirse a cualquiera de las combinaciones anteriores enumeradas en la presente descripción. En algunas modalidades, las sondas con dominios de captura degenerados pueden sustituirse por una de las sondas en cada uno de los pares descritos en la presente descripción.

El dominio de captura puede basarse en una secuencia génica particular o secuencia de motivo particular o secuencia común/conservada, que se diseña para capturar (es decir, un dominio de captura específico de secuencia). Por lo tanto, en algunas modalidades, el dominio de captura es capaz de unirse selectivamente a un subtipo o subconjunto de ácido nucleico deseado, por ejemplo, un tipo particular de ARN, tal como ARNm, ARNr, ARNt, ARN SRP, ARNtm, ARNpn, ARNpno, ARN SmY, ARNacc, ARNg, RNasa P, RNasa MRP, TERC, ARN SL, ARNa, cis-NAT, ARNcr, ARNinc, miARN, ARNpi, ARNip, ARNc, ARNtasi, ARNrasi, 7SK, ARNp, ARNnc u otros tipos de ARN. En un ejemplo no limitante, el dominio de captura puede ser capaz de unirse selectivamente a un subconjunto deseado de ácidos ribonucleicos, por ejemplo, ARN de microbioma, tal como ARNr 16S.

En algunas modalidades, un dominio de captura incluye un "anclaje" o "secuencia de anclaje", que es una secuencia de nucleótidos que se diseña para garantizar que el dominio de captura se hibride con el analito biológico previsto. En algunas modalidades, una secuencia de anclaje incluye una secuencia de nucleótidos, que incluye una secuencia de 1-mer, 2-mer, 3-mer o más larga. En algunas modalidades, la secuencia corta es aleatoria. Por ejemplo, un dominio de captura que incluye una secuencia poli(T) puede diseñarse para capturar un ARNm. En tales modalidades, una secuencia de anclaje puede incluir un 3-mer aleatorio (por ejemplo, GGG) que ayuda a garantizar que el dominio de captura poli(T) se hibride con un ARNm. En algunas modalidades, una secuencia de anclaje puede ser VN, N o NN. Alternativamente, la secuencia puede diseñarse mediante el uso de una secuencia específica de nucleótidos. En algunas modalidades, la secuencia de anclaje está en el extremo 3' del dominio de captura. En algunas modalidades, la secuencia de anclaje está en el extremo 5' del dominio de captura.

En algunas modalidades, los dominios de captura de las sondas de captura se bloquean previo a poner en contacto la muestra biológica con la matriz, y las sondas de bloqueo se usan cuando el ácido nucleico en la muestra biológica

se modifica previo a su captura sobre la matriz. En algunas modalidades, se usa la sonda de bloqueo para bloquear o modificar el extremo 3' libre del dominio de captura. En algunas modalidades, las sondas de bloqueo pueden hibridarse con las sondas de captura para enmascarar el extremo 3' libre del dominio de captura, por ejemplo, sondas en horquilla o sondas parcialmente de doble hebra. En algunas modalidades, el extremo 3' libre del dominio de captura puede bloquearse mediante modificación química, por ejemplo, la adición de un grupo azidometilo como una porción protectora químicamente reversible, de manera que las sondas de captura no incluyen un extremo 3' libre. El bloqueo o modificación de las sondas de captura, particularmente en el extremo 3' libre del dominio de captura, previo a poner en contacto la muestra biológica con la matriz, evita la modificación de las sondas de captura, por ejemplo, evita la adición de una cola de poli(A) al extremo 3' libre de las sondas de captura.

Los ejemplos no limitantes de modificaciones 3' incluyen didesoxi C-3' (3'-ddC), dT invertido 3', espaciador C3 3', amino 3' y fosforilación 3'. En algunas modalidades, el ácido nucleico en la muestra biológica puede modificarse de manera que pueda capturarse por el dominio de captura. Por ejemplo, una secuencia adaptadora (que incluye un dominio de unión capaz de unirse al dominio de captura de la sonda de captura) puede añadirse al extremo del ácido nucleico, por ejemplo, ADN genómico fragmentado. En algunas modalidades, esto se logra mediante la ligación de la secuencia adaptadora o la extensión del ácido nucleico. En algunas modalidades, se usa una enzima para incorporar nucleótidos adicionales en el extremo de la secuencia de ácido nucleico, por ejemplo, una cola de poli(A). En algunas modalidades, las sondas de captura pueden enmascararse o modificarse reversiblemente de manera que el dominio de captura de la sonda de captura no incluya un extremo 3' libre. En algunas modalidades, el extremo 3' se elimina, se modifica o se hace inaccesible de manera que el dominio de captura no sea susceptible al proceso que se usa para modificar el ácido nucleico de la muestra biológica, por ejemplo, ligación o extensión.

En algunas modalidades, el dominio de captura de la sonda de captura se modifica para permitir la eliminación de cualquiera de las modificaciones de la sonda de captura que ocurra durante la modificación de las moléculas de ácido nucleico de la muestra biológica. En algunas modalidades, las sondas de captura pueden incluir una secuencia adicional aguas abajo del dominio de captura, es decir, 3' con respecto al dominio de captura, específicamente un dominio de bloqueo.

En algunas modalidades, el dominio de captura de la sonda de captura puede ser un dominio que no es ácido nucleico. Los ejemplos de dominios de captura adecuados que no se basan exclusivamente en ácido nucleico incluyen, pero no se limitan a, proteínas, péptidos, aptámeros, antígenos, anticuerpos y análogos moleculares que imitan la funcionalidad de cualquiera de los dominios de captura descritos en la presente descripción.

ii. Dominio de escisión

Cada sonda de captura puede incluir opcionalmente al menos un dominio de escisión. El dominio de escisión representa la porción de la sonda que se usa para unir reversiblemente la sonda a un elemento de la matriz, como se describirá más abajo. Además, uno o más segmentos o regiones de la sonda de captura pueden liberarse opcionalmente desde el elemento de la matriz mediante escisión del dominio de escisión. Como un ejemplo, los códigos de barras espaciales y/o los identificadores moleculares universales (UMI) pueden liberarse mediante escisión del dominio de escisión.

En algunas modalidades, el dominio de escisión que une la sonda de captura a un elemento es un enlace disulfuro. Puede añadirse un agente reductor para romper los enlaces disulfuro, lo que resulta en la liberación de la sonda de captura del elemento. Como otro ejemplo, el calentamiento también puede resultar en la degradación del dominio de escisión y la liberación de la sonda de captura unida del elemento de la matriz. En algunas modalidades, la radiación láser se usa para calentar y degradar los dominios de escisión de sondas de captura en ubicaciones específicas. En algunas modalidades, el dominio de escisión es un enlace químico fotosensible (es decir, un enlace químico que se disocia cuando se expone a la luz tal como la luz ultravioleta).

Otros ejemplos de dominios de escisión incluyen enlaces químicos lábiles tales como, pero no limitados a, enlaces éster (por ejemplo, escindible con un ácido, una base o hidroxilamina), un enlace diol vicinal (por ejemplo, escindible mediante periodato de sodio), un enlace Diels-Alder (por ejemplo, escindible mediante calor), un enlace de sulfona (por ejemplo, escindible mediante una base), un enlace de éter silílico (por ejemplo, escindible mediante un ácido), un enlace glucosídico (por ejemplo, escindible mediante una amilasa), un enlace peptídico (por ejemplo, escindible mediante una proteasa), o un enlace fosfodiéster (por ejemplo, escindible mediante una nucleasa (por ejemplo, DNAasa)).

En algunas modalidades, el dominio de escisión incluye una secuencia que se reconoce por una o más enzimas capaces de escindir una molécula de ácido nucleico, por ejemplo, capaz de romper el enlace fosfodiéster entre dos o más nucleótidos. Un enlace puede ser escindible mediante otras enzimas que se dirigen a moléculas de ácido nucleico, tales como enzimas de restricción (por ejemplo, endonucleasas de restricción). Por ejemplo, el dominio de escisión puede incluir una secuencia de reconocimiento de endonucleasa de restricción (enzima de restricción). Las enzimas de restricción cortan el ADN de doble hebra o de hebra única en secuencias de nucleótidos de reconocimiento específicas conocidas como sitios de restricción. En algunas modalidades, una enzima de restricción de corte raro, es

decir, enzimas con un sitio de reconocimiento largo (al menos 8 pares de bases de longitud), se usa para reducir la posibilidad de escindir en otra parte en la sonda de captura.

En algunas modalidades, el dominio de escisión incluye una secuencia de poli(U) que puede escindirse mediante una mezcla de uracil ADN glucosilasa (UDG) y la endonucleasa ADN glucosilasa-liasa VIII, conocida comercialmente como la enzima USER™. Las sondas de captura liberables pueden estar disponibles para la reacción una vez liberadas. Por lo tanto, por ejemplo, una sonda de captura activable puede activarse mediante la liberación de las sondas de captura de un elemento.

En algunas modalidades, donde la sonda de captura se une indirectamente a un sustrato, por ejemplo, mediante una sonda de superficie, el dominio de escisión incluye uno o más nucleótidos no apareados, de manera que las partes complementarias de la sonda de superficie y la sonda de captura no son complementarias al 100 % (por ejemplo, el número de pares de bases no apareadas puede ser uno, dos o tres pares de bases). Tal error de apareamiento se reconoce, por ejemplo, por las enzimas endonucleasa I de MutY y T7, lo que resulta en la escisión de la molécula de ácido nucleico en la posición del error de apareamiento.

En algunas modalidades, donde la sonda de captura se une a un elemento indirectamente, por ejemplo, mediante una sonda de superficie, el dominio de escisión incluye un sitio o secuencia de reconocimiento de nickasa. Las nickasas son endonucleasas que escinden solo una única hebra de un dúplex de ADN. Por lo tanto, el dominio de escisión puede incluir un sitio de reconocimiento de nickasa cerca del extremo 5' de la sonda de superficie (y/o el extremo 5' de la sonda de captura) de manera que la escisión de la sonda de superficie o sonda de captura desestabiliza el dúplex entre la sonda de superficie y la sonda de captura, de esta manera se libera la sonda de captura del elemento.

Las enzimas nickasas pueden usarse además en algunas modalidades donde la sonda de captura se une directamente al elemento. Por ejemplo, el sustrato puede ponerse en contacto con una molécula de ácido nucleico que se hibrida con el dominio de escisión de la sonda de captura para proporcionar o reconstituir un sitio de reconocimiento de nickasa, por ejemplo, una sonda auxiliar de escisión. Por lo tanto, el contacto con una enzima nickasa resultará en la escisión del dominio de escisión, de esta manera se libera la sonda de captura del elemento. Tales sondas auxiliares de escisión pueden usarse además para proporcionar o reconstituir sitios de reconocimiento de escisión para otras enzimas de escisión, por ejemplo, enzimas de restricción.

Algunas nickasas introducen mellas de hebra única solo en sitios particulares en una molécula de ADN, al unirse y reconocer una secuencia de reconocimiento de nucleótidos particular. Se ha descubierto un número de nickasas de origen natural, de las cuales se han determinado en la actualidad las propiedades de reconocimiento de secuencias para al menos cuatro. Las nickasas se describen en la patente de Estados Unidos núm. 6,867,028. En general, cualquier nickasa adecuada puede usarse para unirse a un sitio de reconocimiento de nickasa complementario de un dominio de escisión. Después del uso, la enzima nickasa puede eliminarse del ensayo o inactivarse después de la liberación de las sondas de captura para evitar la escisión no deseada de las sondas de captura.

Los ejemplos de dominios de captura adecuados que no se basan exclusivamente en ácido nucleico incluyen, pero no se limitan a, proteínas, péptidos, aptámeros, antígenos, anticuerpos y análogos moleculares que imitan la funcionalidad de cualquiera de los dominios de captura descritos en la presente descripción.

En algunas modalidades, un dominio de escisión está ausente de la sonda de captura. Los ejemplos de sustratos con sondas de captura unidas que carecen de un dominio de escisión se describen, por ejemplo, en Macosko y otros, (2015) Cell 161, 1202–1214.

En algunas modalidades, la región de la sonda de captura correspondiente al dominio de escisión puede usarse para alguna otra función. Por ejemplo, puede incluirse una región adicional para la extensión o la amplificación del ácido nucleico donde normalmente se colocaría el dominio de escisión. En tales modalidades, la región puede complementar el dominio funcional o incluso existir como un dominio funcional adicional. En algunas modalidades, el dominio de escisión está presente pero su uso es opcional.

iii. Dominio funcional

Cada sonda de captura puede incluir opcionalmente al menos un dominio funcional. Cada dominio funcional incluye, típicamente, una secuencia de nucleótidos funcional para una etapa analítica aguas abajo en el procedimiento de análisis general.

En algunos casos, la molécula de ácido nucleico puede comprender una o más secuencias funcionales. Por ejemplo, una secuencia funcional puede comprender una secuencia para la unión a una celda de flujo de secuenciación, tal como, por ejemplo, una secuencia P5 para la secuenciación de Illumina®. En algunos casos, la molécula de ácido nucleico o derivado de esta (por ejemplo, oligonucleótido o polinucleótido que se genera a partir de la molécula de ácido nucleico) puede comprender otra secuencia funcional, tal como, por ejemplo, una secuencia P7 para unirse a una celda de flujo de secuenciación para la secuenciación de Illumina. En algunos casos, la secuencia funcional puede comprender una secuencia de código de barras o múltiples secuencias de código de barras. En algunos casos, la

secuencia funcional puede comprender un identificador molecular único (UMI). En algunos casos, la secuencia funcional puede comprender una secuencia de cebador (por ejemplo, una secuencia de cebador R1 para la secuenciación de Illumina, una secuencia de cebador R2 para la secuenciación de Illumina, etc.). En algunos casos, una secuencia funcional puede comprender una secuencia parcial, tal como una secuencia de código de barras parcial, una secuencia de anclaje parcial, una secuencia de cebador de secuenciación parcial (por ejemplo, secuencia R1 parcial, secuencia R2 parcial, etc.), una secuencia parcial que se configura para unirse a la celda de flujo de un secuenciador (por ejemplo, secuencia P5 parcial, secuencia P7 parcial, etc.), o una secuencia parcial de cualquier otro tipo de secuencia descrita en otra parte en la presente descripción. Una secuencia parcial puede contener una porción o segmento contiguo o continuo, pero no todos, de una secuencia completa, por ejemplo. En algunos casos, un procedimiento aguas abajo puede extender la secuencia parcial, o derivado de esta, para lograr una secuencia completa de la secuencia parcial, o derivado de esta. Ejemplos de tales sondas de captura y usos de estas se describen en las publicaciones de patentes de Estados Unidos núms. 2014/0378345 y 2015/0376609. Los dominios funcionales pueden seleccionarse para su compatibilidad con una variedad de diferentes sistemas de secuenciación, por ejemplo, secuenciación 454, Ion Torrent Proton o PGM, Illumina X10, etc., u otras plataformas de Illumina, BGI, Qiagen, Thermo-Fisher, PacBio y Roche, y los requisitos de estos.

iv. Código de barras espacial

Como se discutió anteriormente, la sonda de captura puede incluir uno o más códigos de barras espacial (por ejemplo, dos o más, tres o más, cuatro o más, cinco o más códigos de barras espacial). Un "código de barras espacial" es un segmento de ácido nucleico contiguo o dos o más segmentos de ácido nucleico no contiguos que funcionan como una marca o identificador que transmite o es capaz de transmitir información espacial. En algunas modalidades, una sonda de captura incluye un código de barras espacial que posee un aspecto espacial, donde el código de barras se asocia con una ubicación particular dentro de una matriz o una ubicación particular sobre un sustrato. Los códigos de barras espaciales ilustrativos se describen en la patente de Estados Unidos núm. 10,030,261.

Un código de barras espacial puede ser parte de un analito o independiente de un analito (es decir, parte de una sonda de captura). Un código de barras espacial puede ser una etiqueta que se une a un analito (por ejemplo, una molécula de ácido nucleico) o una combinación de una etiqueta además de una característica endógena del analito (por ejemplo, tamaño del analito o secuencia(s) del/de los extremo(s)). Un código de barras espacial puede ser único. En algunas modalidades donde el código de barras espacial es único, el código de barras espacial funciona tanto como un código de barras espacial y como un identificador molecular único (UMI), asociado con una sonda de captura particular.

Los códigos de barras espaciales pueden tener una variedad de formatos diferentes. Por ejemplo, los códigos de barras espaciales pueden incluir códigos de barras espaciales de polinucleótidos, secuencias de aminoácidos y/o ácido nucleico aleatorias, y secuencias de aminoácidos y/o ácido nucleico sintéticas. En algunas modalidades, un código de barras espacial se une a un analito de manera reversible o irreversible. En algunas modalidades, se añade un código de barras espacial, por ejemplo, a un fragmento de una muestra de ADN o ARN antes, durante y/o después de la secuenciación de la muestra. En algunas modalidades, un código de barras espacial permite la identificación y/o cuantificación de lecturas de secuenciación individuales. En algunas modalidades, un código de barras espacial es un código de barras fluorescente para el que las sondas de oligonucleótidos marcadas con fluorescencia se hibridan con el código de barras espacial.

En algunas modalidades, el código de barras espacial es una secuencia de ácido nucleico que no se hibrida sustancialmente con moléculas de ácido nucleico del analito en una muestra biológica. En algunas modalidades, el código de barras espacial tiene menos del 80 % de identidad de secuencia (por ejemplo, menos del 70 %, 60 %, 50 % o menos del 40 % de identidad de secuencia) con las secuencias de ácido nucleico a través de una parte sustancial (por ejemplo, 80 % o más) de las moléculas de ácido nucleico en la muestra biológica.

Las secuencias de código de barras espaciales pueden incluir de aproximadamente 6 a aproximadamente 20 o más nucleótidos dentro de la secuencia de las sondas de captura. En algunas modalidades, la longitud de una secuencia de código de barras espacial puede ser de aproximadamente 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucleótidos o más larga. En algunas modalidades, la longitud de una secuencia de código de barras espacial puede ser al menos aproximadamente 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucleótidos o más larga. En algunas modalidades, la longitud de una secuencia de código de barras espacial es a lo máximo aproximadamente 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucleótidos o más corta.

Estos nucleótidos pueden ser completamente contiguos, es decir, en un único tramo de nucleótidos adyacentes, o pueden separarse en dos o más subsecuencias separadas que se separan por 1 o más nucleótidos. Las subsecuencias de código de barras espacial separadas pueden ser de aproximadamente 4 a aproximadamente 16 nucleótidos de longitud. En algunas modalidades, la subsecuencia del código de barras espacial puede ser aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 nucleótidos o más larga. En algunas modalidades, la subsecuencia del código de barras espacial puede ser al menos aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 nucleótidos o más larga. En algunas modalidades, la subsecuencia del código de barras espacial puede ser a lo máximo aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 nucleótidos o más corta.

Para múltiples sondas de captura que se unen a un elemento de matriz común, la una o más secuencias de códigos de barras espaciales de las múltiples sondas de captura pueden incluir secuencias que son las mismas para todas las sondas de captura acopladas al elemento, y/o secuencias que son diferentes en todas las sondas de captura acopladas al elemento. En algunas modalidades, una pluralidad de sondas de captura que se unen a un elemento de la matriz común puede poseer el mismo código de barras espacial, pero tener diferentes dominios de captura diseñados para asociar el código de barras espacial del elemento con más de un analito diana. Por ejemplo, un elemento puede acoplarse a dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, o más tipos diferentes de sondas de captura con códigos de barras espacial, cada tipo de sonda de captura con códigos de barras espacial posee el código de barras espacial. En algunos aspectos, los constructos de sonda de captura con códigos de barras pueden adaptarse para el análisis de cualquier analito dado asociado con un ácido nucleico y capaz de unirse con tal constructo. En algunas modalidades, el analito es un analito de ácido nucleico capaz de unirse con una sonda de captura con código de barras espacial descrita en la presente descripción. En algunas modalidades, el analito es un analito que no es ácido nucleico (por ejemplo, un analito de proteína) que se une covalente o no covalentemente a un ácido nucleico capaz de unirse con una sonda de captura con código de barras espacial descrita en la presente descripción. Por ejemplo, un método descrito en la presente descripción puede comprender analizar simultáneamente una pluralidad de analitos, que incluyen pero no se limitan a: ADN, ARN (por ejemplo, ARNm y/o ARN no codificante), superficie celular o proteínas y metabolitos intracelulares, un agente de marcaje con código de barras o cualquiera de sus combinaciones.

Las sondas de captura que se unen a un único elemento de matriz pueden incluir secuencias de códigos de barras espaciales idénticas (o comunes), secuencias de códigos de barras espaciales diferentes o una combinación de ambas. Las sondas de captura que se unen a un elemento pueden incluir múltiples conjuntos de sondas de captura. Las sondas de captura de un conjunto dado pueden incluir secuencias de códigos de barras espaciales idénticas. Las secuencias de códigos de barras espaciales idénticas pueden ser diferentes de las secuencias de códigos de barras espaciales de sondas de captura de otro conjunto.

La pluralidad de sondas de captura puede incluir secuencias de códigos de barras espaciales (por ejemplo, secuencias de códigos de barras de ácido nucleico) que están asociadas con ubicaciones específicas sobre una matriz espacial. Por ejemplo, una primera pluralidad de sondas de captura puede asociarse con una primera región, basándose en una secuencia de código de barras espacial común a las sondas de captura dentro de la primera región, y una segunda pluralidad de sondas de captura puede asociarse con una segunda región, basándose en una secuencia de código de barras espacial común a las sondas de captura dentro de la segunda región. La segunda región puede estar asociada o no con la primera región. Pueden asociarse pluralidades adicionales de sondas de captura con secuencias de códigos de barras espaciales comunes a las sondas de captura dentro de otras regiones. En algunas modalidades, las secuencias de códigos de barras espaciales pueden ser las mismas en una pluralidad de moléculas de sonda de captura.

En algunas modalidades, se incorporan múltiples códigos de barras espaciales diferentes en una única matriz de sonda de captura. Por ejemplo, un conjunto mixto pero conocido de secuencias de códigos de barras espaciales puede proporcionar una dirección o atribución más sólida de los códigos de barras espaciales a un lugar o ubicación determinados, al proporcionar una confirmación duplicada o independiente de la identidad de la ubicación. En algunas modalidades, los múltiples códigos de barras espaciales representan una especificidad creciente de la ubicación del punto de la matriz particular.

v. Identificador molecular único

La sonda de captura puede incluir uno o más (por ejemplo, dos o más, tres o más, cuatro o más, cinco o más) Identificadores Moleculares Únicos (UMI). Un identificador molecular único es un segmento de ácido nucleico contiguo o dos o más segmentos de ácido nucleico no contiguos que funcionan como una marca o identificador para un analito particular, o para una sonda de captura que se une a un analito particular (por ejemplo, mediante el dominio de captura).

Un UMI puede ser único. Un UMI puede incluir una o más secuencias de polinucleótidos específicas, una o más secuencias de aminoácidos y/o ácido nucleico aleatorias, y/o una o más secuencias de aminoácidos y/o ácido nucleico sintéticas.

En algunas modalidades, el UMI es una secuencia de ácido nucleico que no se hibrida sustancialmente con moléculas de ácido nucleico del analito en una muestra biológica. En algunas modalidades, el UMI tiene menos del 80 % de identidad de secuencia (por ejemplo, menos del 70 %, 60 %, 50 % o menos del 40 % de identidad de secuencia) con las secuencias de ácido nucleico a través de una parte sustancial (por ejemplo, 80 % o más) de las moléculas de ácido nucleico en la muestra biológica.

El UMI puede incluir de aproximadamente 6 a aproximadamente 20 o más nucleótidos dentro de la secuencia de las sondas de captura. En algunas modalidades, la longitud de una secuencia de UMI puede ser de aproximadamente 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucleótidos o más larga. En algunas modalidades, la longitud de una secuencia de UMI puede ser al menos aproximadamente 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucleótidos

o más larga. En algunas modalidades, la longitud de una secuencia de UMI es a lo máximo aproximadamente 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 nucleótidos o más corta.

Estos nucleótidos pueden ser completamente contiguos, es decir, en un único tramo de nucleótidos adyacentes, o pueden separarse en dos o más subsecuencias separadas que se separan por 1 o más nucleótidos. Las subsecuencias de UMI separadas pueden ser de aproximadamente 4 a aproximadamente 16 nucleótidos de longitud. En algunas modalidades, la subsecuencia del UMI puede ser de aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 nucleótidos o más larga. En algunas modalidades, la subsecuencia del UMI puede ser al menos aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 nucleótidos o más larga. En algunas modalidades, la subsecuencia del UMI puede ser a lo máximo aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 nucleótidos o más corta.

En algunas modalidades, un UMI se une a un analito de manera reversible o irreversible. En algunas modalidades, se añade un UMI, por ejemplo, a un fragmento de una muestra de ADN o ARN antes, durante y/o después de la secuenciación del analito. En algunas modalidades, un UMI permite la identificación y/o cuantificación de lecturas de secuenciación individuales. En algunas modalidades, un UMI se usa como un código de barras fluorescente para el cual las sondas de oligonucleótidos marcadas con fluorescencia se hibridan con el UMI.

vi. Otros aspectos de las sondas de captura

Para las sondas de captura que se unen a un elemento de la matriz, un elemento de la matriz individual puede incluir una o más sondas de captura. En algunas modalidades, un elemento de la matriz individual incluye cientos o miles de sondas de captura. En algunas modalidades, las sondas de captura se asocian con un elemento individual particular, donde el elemento individual contiene una sonda de captura que incluye un código de barras espacial único para una región o ubicación definida sobre la matriz.

En algunas modalidades, un elemento particular puede contener sondas de captura que incluyen más de un código de barras espacial (por ejemplo, una sonda de captura en un elemento particular puede incluir un código de barras espacial que es diferente al código de barras espacial que se incluye en otra sonda de captura en el mismo elemento particular, mientras que ambas sondas de captura incluyen un segundo código de barras espacial común), donde cada código de barras espacial corresponde a una región o ubicación definida particular sobre la matriz. Por ejemplo, múltiples secuencias de código de barras espaciales asociadas con un elemento particular sobre una matriz pueden proporcionar una dirección o atribución más fuerte a una ubicación dada al proporcionar una confirmación duplicada o independiente de la ubicación. En algunas modalidades, los múltiples códigos de barras espaciales representan una especificidad creciente de la ubicación del punto de la matriz particular. En un ejemplo no limitante, un punto de matriz particular puede codificarse con dos códigos de barras espaciales diferentes, donde cada código de barras espacial identifica una región definida particular dentro de la matriz, y un punto de matriz que posee ambos códigos de barras espaciales identifica la subregión donde dos regiones definidas se solapan, por ejemplo, tal como la porción de solapamiento de un diagrama de Venn.

En otro ejemplo no limitante, un punto de matriz particular puede codificarse con tres códigos de barras espaciales diferentes, donde el primer código de barras espacial identifica una primera región dentro de la matriz, el segundo código de barras espacial identifica una segunda región, donde la segunda región es una subregión completamente dentro de la primera región, y el tercer código de barras espacial identifica una tercera región, donde la tercera región es una subregión completamente dentro de la primera y la segunda subregiones.

En algunas modalidades, las sondas de captura que se unen a los elementos de la matriz se liberan de los elementos de la matriz para la secuenciación. Alternativamente, en algunas modalidades, las sondas de captura permanecen unidas a los elementos de la matriz, y las sondas se secuencian mientras permanecen unidas a los elementos de la matriz. Otros aspectos de la secuenciación de sondas de captura se describen en las secciones subsecuentes de esta descripción.

En algunas modalidades, un elemento de la matriz puede incluir diferentes tipos de sondas de captura que se unen al elemento. Por ejemplo, el elemento de la matriz puede incluir un primer tipo de sonda de captura con un dominio de captura diseñado para unirse a un tipo de analito, y un segundo tipo de sonda de captura con un dominio de captura diseñado para unirse a un segundo tipo de analito. En general, los elementos de la matriz pueden incluir uno o más (por ejemplo, dos o más, tres o más, cuatro o más, cinco o más, seis o más, ocho o más, diez o más, 12 o más, 15 o más, 20 o más, 30 o más, 50 o más) tipos diferentes de sondas de captura que se unen a un único elemento de la matriz.

En algunas modalidades, la sonda de captura es un ácido nucleico. En algunas modalidades, la sonda de captura se une al elemento de la matriz mediante su extremo 5'. En algunas modalidades, la sonda de captura incluye desde el extremo 5' hasta 3': uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI) y uno o más dominios de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 5' hasta 3': un código de barras (por ejemplo, un código de barras espacial o un UMI) y un dominio de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 5' hasta 3': un dominio de escisión, un dominio funcional, uno o más

códigos de barras (por ejemplo, un código de barras espacial y/o un UMI), y un dominio de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 5' hasta 3': un dominio de escisión, un dominio funcional, uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI), un segundo dominio funcional y un dominio de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 5' hasta 3': un dominio de escisión, un dominio funcional, un código de barras espacial, un UMI y un dominio de captura. En algunas modalidades, la sonda de captura no incluye un código de barras espacial. En algunas modalidades, la sonda de captura no incluye un UMI. En algunas modalidades, la sonda de captura incluye una secuencia para iniciar una reacción de secuenciación.

En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento mediante su extremo 3'. En algunos casos, la sonda de captura comprende: una secuencia adaptadora – un código de barras (por ejemplo, un código de barras espacial) – una secuencia del identificador molecular único (UMI) opcional – un dominio de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 3' hasta 5': uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI) y uno o más dominios de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 3' hasta 5': un código de barras (por ejemplo, un código de barras espacial o un UMI) y un dominio de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 3' hasta 5': un dominio de escisión, un dominio funcional, uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI), y un dominio de captura. En algunas modalidades, la sonda de captura incluye desde el extremo 3' hasta 5': un dominio de escisión, un dominio funcional, un código de barras espacial, un UMI y un dominio de captura.

En algunas modalidades, una sonda de captura incluye un oligonucleótido sintetizado in situ. En algunas modalidades, el oligonucleótido sintetizado in situ incluye una o más secuencias constantes, una o más de las cuales sirven como una secuencia de cebado (por ejemplo, un cebador para amplificar los ácidos nucleicos diana). En algunas modalidades, una secuencia constante es una secuencia escindible. En algunas modalidades, el oligonucleótido sintetizado in situ incluye una secuencia de código de barras, por ejemplo, una secuencia de código de barras variable. En algunas modalidades, el oligonucleótido sintetizado in situ se une a un elemento de una matriz.

En algunas modalidades, una sonda de captura es un producto de dos o más secuencias de oligonucleótidos, por ejemplo, dos o más secuencias de oligonucleótidos que se ligan entre sí. En algunas modalidades, una de las secuencias de oligonucleótidos es un oligonucleótido sintetizado in situ.

En algunas modalidades, la sonda de captura incluye un oligonucleótido de soporte. Dos o más oligonucleótidos pueden ligarse juntos mediante el uso de un oligonucleótido de soporte y cualquier variedad de ligasas conocidas en la técnica o descritas en la presente descripción (por ejemplo, ligasa SplintR).

En algunas modalidades, uno de los oligonucleótidos incluye: una secuencia constante (por ejemplo, una secuencia complementaria a una porción de un oligonucleótido de soporte), una secuencia degenerada y un dominio de captura (por ejemplo, como se describe en la presente descripción). En algunas modalidades, la sonda de captura se genera al tener una enzima que añade polinucleótidos en el extremo de una secuencia de oligonucleótidos. La sonda de captura puede incluir una secuencia degenerada, que puede funcionar como un identificador molecular único.

Una sonda de captura puede incluir una secuencia degenerada, que es una secuencia en la que algunas posiciones de una secuencia de nucleótidos contienen un número de bases posibles. Una secuencia degenerada puede ser una secuencia de nucleótidos degenerada que incluye aproximadamente o al menos 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45 o 50 nucleótidos. En algunas modalidades, una secuencia de nucleótidos contiene 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más posiciones degeneradas dentro de la secuencia de nucleótidos. En algunas modalidades, la secuencia degenerada se usa como un UMI.

En algunas modalidades, una sonda de captura incluye una secuencia de reconocimiento de endonucleasa de restricción o una secuencia de nucleótidos escindibles mediante actividades enzimáticas específicas. Por ejemplo, las secuencias de uracilo pueden escindirse mediante actividad enzimática específica. Como otro ejemplo, otras bases modificadas (por ejemplo, modificadas mediante metilación) pueden reconocerse y escindirse mediante endonucleasas específicas. Las sondas de captura pueden someterse a una escisión enzimática, que elimina el dominio de bloqueo y cualquiera de los nucleótidos adicionales que se añaden al extremo 3' de la sonda de captura durante el proceso de modificación. La eliminación del dominio de bloqueo revela y/o restaura el extremo 3' libre del dominio de captura de la sonda de captura. En algunas modalidades, pueden eliminarse nucleótidos adicionales para revelar y/o restaurar el extremo 3' del dominio de captura de la sonda de captura.

En algunas modalidades, un dominio de bloqueo puede incorporarse en la sonda de captura cuando se sintetiza, o después de su síntesis. El nucleótido terminal del dominio de captura es un nucleótido terminador reversible (por ejemplo, terminador reversible bloqueado en 3'-O y terminador reversible desbloqueado en 3'), y puede incluirse en la sonda de captura durante o después de la síntesis de la sonda.

vii. Sondas de captura extendidas

Una "sonda de captura extendida" es una sonda de captura con una secuencia de ácido nucleico ampliada. Por ejemplo, donde la sonda de captura incluye ácido nucleico, un "extremo 3' extendido" indica que se añadieron nucleótidos adicionales al nucleótido más cerca del 3' de la sonda de captura para extender la longitud de la sonda de captura, por ejemplo, mediante reacciones de polimerización estándar utilizadas para extender moléculas de ácido nucleico, que incluye la polimerización con molde catalizada por una polimerasa (por ejemplo, una ADN polimerasa o una transcriptasa inversa).

En algunas modalidades, la extensión de la sonda de captura incluye generar ADNc a partir del ARN capturado (hibridado). Este proceso involucra la síntesis de una hebra complementaria del ácido nucleico hibridado, por ejemplo, que genera ADNc en base al molde de ARN capturado (el ARN que se hibrida con el dominio de captura de la sonda de captura). Por lo tanto, en una etapa inicial de extensión de la sonda de captura, por ejemplo, la generación de ADNc, el ácido nucleico capturado (hibridado), por ejemplo ARN, actúa como un molde para la etapa de extensión, por ejemplo, transcripción inversa.

En algunas modalidades, la sonda de captura se extiende mediante el uso de transcripción inversa. Por ejemplo, la transcripción inversa incluye sintetizar ADNc (ADN complementario o copia) a partir de ARN, por ejemplo, (ARN mensajero), mediante el uso de una transcriptasa inversa. En algunas modalidades, la transcripción inversa se realiza mientras el tejido aún está en su lugar, lo que genera una biblioteca de analitos, donde la biblioteca de analitos incluye los códigos de barras espaciales de las sondas de captura adyacentes. En algunas modalidades, la sonda de captura se extiende mediante el uso de una o más ADN polimerasas.

En algunas modalidades, el dominio de captura de la sonda de captura incluye un cebador para producir la hebra complementaria del ácido nucleico hibridado a la sonda de captura, por ejemplo, un cebador para la ADN polimerasa y/o la transcripción inversa. Las moléculas de ácido nucleico, por ejemplo, ADN y/o ADNc, generadas por la reacción de extensión incorporan la secuencia de la sonda de captura. La extensión de la sonda de captura, por ejemplo, una reacción de ADN polimerasa y/o transcripción inversa, puede realizarse mediante el uso de una variedad de enzimas y protocolos adecuados.

En algunas modalidades, se genera una molécula de ADN de longitud completa, por ejemplo, ADNc. En algunas modalidades, una molécula de ADN de "longitud completa" se refiere a la totalidad de la molécula de ácido nucleico capturada. Sin embargo, si el ácido nucleico, por ejemplo, ARN, se degrada parcialmente en la muestra de tejido, entonces las moléculas de ácido nucleico capturadas no tendrán la misma longitud que el ARN inicial en la muestra de tejido. En algunas modalidades, se modifica el extremo 3' de las sondas extendidas, por ejemplo, las moléculas de ADNc de la primera hebra. Por ejemplo, un enlazador o adaptador puede ligarse al extremo 3' de las sondas extendidas. Esto puede lograrse mediante el uso de enzimas de ligación de hebras únicas tales como ARN ligasa T4 o CircLigase™ (disponible en Epicentre Biotechnologies, Madison, WI). En algunas modalidades, los oligonucleótidos que cambian de molde se usan para extender el ADNc para generar un ADNc de longitud completa (o lo más cerca posible de un ADNc de longitud completa). En algunas modalidades, una sonda auxiliar de síntesis de la segunda hebra (una molécula de ADN parcialmente de doble hebra capaz de hibridarse con el extremo 3' de la sonda de captura extendida), puede ligarse al extremo 3' de la sonda extendida, por ejemplo, molécula de ADNc de la primera hebra, mediante el uso de una enzima de ligación de doble hebra tal como ADN ligasa T4. Otras enzimas apropiadas para la etapa de ligación se conocen en la técnica e incluyen, por ejemplo, Tth ADN ligasa, Taq ADN ligasa, ADN ligasa de *Thermococcus* sp. (cepa 9°N) (ADN ligasa 9°N™ New England Biolabs), Ampligase™ (disponible en Lucigen, Middleton, WI), y SplintR (disponible en New England Biolabs, Ipswich, MA). En algunas modalidades, una cola de polinucleótido, por ejemplo, una cola de poli(A), se incorpora en el extremo 3' de las moléculas de sonda extendida. En algunas modalidades, la cola de polinucleótido se incorpora mediante el uso de una enzima activa terminal transferasa.

En algunas modalidades las sondas de captura extendidas de doble hebra se tratan para eliminar cualquier sonda de captura no extendida previo a la amplificación y/o el análisis, por ejemplo, el análisis de secuencias. Esto puede lograrse mediante una variedad de métodos, por ejemplo, mediante el uso de una enzima para degradar las sondas no extendidas, tales como, una enzima exonucleasa o columnas de purificación.

En algunas modalidades, las sondas de captura extendidas se amplifican para producir cantidades que son suficientes para el análisis, por ejemplo, mediante secuenciación de ADN. En algunas modalidades, la primera hebra de las sondas de captura extendidas, (por ejemplo moléculas de ADN y/o ADNc) actúa como un molde para la reacción de amplificación (por ejemplo, una reacción en cadena de la polimerasa).

En algunas modalidades, la reacción de amplificación incorpora un grupo de afinidad sobre la sonda de captura extendida (por ejemplo, el híbrido ARN-ADNc) mediante el uso de un cebador que incluye el grupo de afinidad. En algunas modalidades, el cebador incluye un grupo de afinidad y las sondas de captura extendida incluyen el grupo de afinidad. El grupo de afinidad puede corresponder a cualquiera de los grupos de afinidad descritos anteriormente.

En algunas modalidades, las sondas de captura extendida que incluyen el grupo de afinidad pueden acoplarse a un elemento de matriz específico para el grupo de afinidad. En algunas modalidades, el sustrato puede incluir un anticuerpo o fragmento de anticuerpo. En algunas modalidades, el elemento de matriz incluye avidina o estreptavidina.

y el grupo de afinidad incluye biotina. En algunas modalidades, el elemento de matriz incluye maltosa y el grupo de afinidad incluye proteína de unión a maltosa. En algunas modalidades, el elemento de matriz incluye proteína de unión a maltosa y el grupo de afinidad incluye maltosa. En algunas modalidades, la amplificación de las sondas de captura extendida puede funcionar para liberar las sondas extendidas del elemento de matriz, en la medida en que las copias de las sondas extendidas no se unan al elemento de matriz.

En algunas modalidades, la sonda de captura extendida o complemento o amplicón de esta se libera de un elemento de matriz. La etapa de liberación de la sonda de captura extendida o complemento o amplicón de esta de la superficie de un elemento de matriz puede lograrse de varias maneras. En algunas modalidades, una sonda de captura extendida o un complemento de esta se libera del elemento mediante escisión de ácido nucleico y/o por desnaturalización (por ejemplo, mediante calentamiento para desnaturalizar una molécula de doble hebra).

VI. Agentes de marcaje

En algunas modalidades, en la presente descripción se proporcionan métodos, composiciones, dispositivos y kits para usar agentes de captura del analito para la caracterización espacial de analitos biológicos (por ejemplo, ARN, ADN y superficie celular o proteínas y/o metabolitos intracelulares). En algunas modalidades, un agente de captura del analito (también denominado a veces "agente de marcaje") puede incluir un agente que interactúa con un analito (por ejemplo, un analito en una muestra) y con un agente de captura (por ejemplo, una sonda de captura que se une a un sustrato) para identificar el analito. En algunas modalidades, la muestra puede contraerse con uno o más agentes de marcaje previo a, durante o después de los ensayos *in situ* y/o los ensayos espaciales que se proporcionan en la presente descripción. En algunas modalidades, el método comprende una o más etapas de fijación posterior (también denominadas posfijación) después de entrar en contacto la muestra con uno o más agentes de marcaje. En algunas modalidades, el agente de captura del analito comprende una porción de unión al analito y un dominio de código de barras del agente de captura.

En los métodos y sistemas descritos en la presente descripción, pueden usarse uno o más agentes de marcaje capaces de unirse o acoplarse de cualquier otra manera a uno o más elementos para caracterizar analitos, células y/o elementos celulares. En algunos casos, los elementos celulares incluyen elementos de la superficie celular. Los analitos pueden incluir, pero no se limitan a, una proteína, un receptor, un antígeno, una proteína de superficie, una proteína transmembrana, un conglomerado de proteínas de diferenciación, una proteína canal, una proteína bomba, una proteína portadora, un fosfolípido, una glicoproteína, un glucolípido, un complejo de proteína de interacción célula-célula, un complejo presentador de antígenos, un complejo mayor de histocompatibilidad, un receptor de linfocitos T manipulado genéticamente, un receptor de linfocitos T, un receptor de linfocitos B, un receptor de antígeno quimérico, una unión en hendidura, una unión adherente, o cualquiera de sus combinaciones. En algunos casos, los elementos celulares pueden incluir analitos intracelulares, tales como proteínas, modificaciones de proteínas (por ejemplo, estado de fosforilación u otras modificaciones postraduccionales), proteínas nucleares, proteínas de la membrana nuclear, o cualquiera de sus combinaciones.

Un agente de marcaje puede incluir, pero no se limita a, una proteína, un péptido, un anticuerpo (o un fragmento de unión al epítipo de este), una porción lipófila (tal como colesterol), una molécula de unión al receptor de la superficie celular, un ligando del receptor, una molécula pequeña, un anticuerpo biespecífico, un activador de linfocitos T biespecífico, un activador de receptores de linfocitos T, un activador de receptores de linfocitos B, un procuerpo, un aptámero, un monocuerpo, un afímero, una darpina, y un andamio de proteínas, o cualquiera de sus combinaciones. Los agentes de marcaje pueden incluir (por ejemplo, se unen a) un oligonucleótido reportero que es indicativo del elemento de la superficie celular al que se une el grupo de unión. Por ejemplo, el oligonucleótido reportero puede comprender una secuencia de código de barras que permite la identificación del agente de marcaje. Por ejemplo, un agente de marcaje que es específico para un tipo de elemento celular (por ejemplo, un primer elemento de la superficie celular) puede tener acoplado a este un primer oligonucleótido reportero, mientras que un agente de marcaje que es específico para un elemento celular diferente (por ejemplo, un segundo elemento de la superficie celular) puede tener un oligonucleótido reportero diferente que se acopla a este. Para una descripción de agentes de marcaje ilustrativos, oligonucleótidos reporteros, y métodos de uso, véase, por ejemplo, la patente de Estados Unidos 10,550,429; la publicación de patente de Estados Unidos 20190177800; y la publicación de patente de Estados Unidos 20190367969. En otros casos, por ejemplo, para facilitar el multiplexado de muestras, un agente de marcaje que es específico para un elemento celular en particular puede tener una primera pluralidad del agente de marcaje (por ejemplo, un anticuerpo o porción lipófila) que se acopla a un primer oligonucleótido reportero y una segunda pluralidad del agente de marcaje que se acopla a un segundo oligonucleótido reportero.

En algunos aspectos, estos oligonucleótidos reporteros pueden comprender secuencias de código de barras de ácido nucleico que permiten la identificación del agente de marcaje al que se acopla el oligonucleótido reportero. La selección de oligonucleótidos como el reportero puede proporcionar la ventaja de ser capaz de generar una diversidad significativa en términos de secuencia, al tiempo que también puede unirse fácilmente a la mayoría de las biomoléculas, por ejemplo, anticuerpos, etc., así como también detectarse fácilmente, por ejemplo, mediante el uso de tecnologías de secuenciación o de matriz.

La unión (acoplamiento) de los oligonucleótidos reporteros a los agentes de marcaje puede lograrse a través de cualquiera de una variedad de asociaciones o uniones directas o indirectas, covalentes o no covalentes. Por ejemplo, los oligonucleótidos pueden unirse covalentemente a una porción de un agente de marcaje (tal como una proteína, por ejemplo, un anticuerpo o fragmento de anticuerpo) mediante el uso de técnicas de conjugación química (por ejemplo, kits de marcaje de anticuerpos Lightning-Link® disponibles en Innova Biosciences), así como también otros mecanismos de unión no covalente, por ejemplo, mediante el uso de anticuerpos y oligonucleótidos biotinilados (o perlas que incluyen uno o más enlazadores biotinilados, que se acoplan a oligonucleótidos) con un enlazador de avidina o estreptavidina. Se dispone de técnicas de biotinilación de anticuerpos y oligonucleótidos. Véase, por ejemplo, Fang, y otros, "Fluoride-Cleavable Biotinylation Phosphoramidite for 5'-end-Labeling and Affinity Purification of Synthetic Oligonucleotides," *Nucleic Acids Res.* Jan. 15, 2003; 31(2):708-715. Igualmente, se han desarrollado técnicas de biotinilación de proteínas y péptidos, y están fácilmente disponibles. Véase, por ejemplo, la patente de Estados Unidos núm. 6,265,552. Además, la química de la reacción clic tal como una reacción de éster Metiltetrazina-PEG5-NHS, una reacción de éster de TCO-PEG4-NHS, o similares, puede usarse para acoplar oligonucleótidos reporteros a agentes de marcaje. Pueden usarse kits disponibles comercialmente, tales como los de Thunderlink y Abcam, y técnicas comunes en la técnica, para acoplar oligonucleótidos reporteros a agentes de marcaje según sea apropiado. En otro ejemplo, un agente de marcaje se acopla indirectamente (por ejemplo, mediante hibridación) a un oligonucleótido reportero que comprende una secuencia de código de barras que identifica el agente de marcaje. Por ejemplo, el agente de marcaje puede acoplarse directamente (por ejemplo, unirse covalentemente) a un oligonucleótido de hibridación que comprende una secuencia que se hibrida con una secuencia del oligonucleótido reportero. La hibridación del oligonucleótido de hibridación al oligonucleótido reportero acopla el agente de marcaje al oligonucleótido reportero. En algunas modalidades, los oligonucleótidos reporteros son liberables del agente de marcaje, tal como tras la aplicación de un estímulo. Por ejemplo, el oligonucleótido reportero puede unirse al agente de marcaje a través de un enlace lábil (por ejemplo, lábil químicamente, fotolábil, lábil térmicamente, etc.) como se describe generalmente para liberar moléculas de soportes en cualquier otra parte en la presente descripción. En algunos casos, los oligonucleótidos reporteros descritos en la presente descripción pueden incluir una o más secuencias funcionales que pueden usarse en el procesamiento subsecuente, tal como una secuencia adaptadora, una secuencia de identificador molecular único (UMI), una secuencia de unión a celdas de flujo específicas del secuenciador (tal como un P5, P7, o secuencia parcial P5 o P7), un cebador o secuencia de unión al cebador, un cebador de secuenciación o secuencia de unión al cebador (tal como un R1, R2, o secuencia parcial R1 o R2).

En algunos casos, el agente de marcaje puede comprender un oligonucleótido reportero y un marcador. Un marcador puede ser un fluoróforo, un radioisótopo, una molécula capaz de una reacción colorimétrica, una partícula magnética, o cualquier otra molécula o compuesto adecuado capaz de ser detectado. El marcador puede conjugarse con un agente de marcaje (u oligonucleótido reportero) ya sea directa o indirectamente (por ejemplo, el marcador puede conjugarse a una molécula que puede unirse al agente de marcaje u oligonucleótido reportero). En algunos casos, un marcador se conjuga a un primer oligonucleótido que es complementario (por ejemplo, se hibrida) a una secuencia del oligonucleótido reportero.

La figura 14 es un diagrama esquemático de un agente de marcaje ilustrativo (por ejemplo, un agente de captura del analito) 1402 que comprende una porción de unión al analito 1404 y un reportero 1408. El agente de marcaje 1402, mediante la porción de unión al analito 1404, es capaz de unirse a un analito 1406. El agente de marcaje (por ejemplo, un agente de captura del analito) es capaz de interactuar, además, con una sonda de captura con código de barras espacial. La porción de unión al analito puede unirse al analito 1406 con alta afinidad y/o con alta especificidad. El reportero 1408 puede comprender un ácido nucleico (por ejemplo, un oligonucleótido reportero), que puede hibridarse con al menos una porción o la totalidad de un dominio de captura de un agente de captura (por ejemplo, una sonda de captura). La porción de unión al analito 1404 puede incluir un polipéptido y/o un aptámero (por ejemplo, un oligonucleótido o molécula peptídica que se une a un analito diana específico). La porción de unión al analito 1404 puede incluir un anticuerpo o fragmento de anticuerpo (por ejemplo, un fragmento de unión a antígeno).

La figura 15 es un diagrama esquemático de otro agente de marcaje ilustrativo (por ejemplo, un agente de captura del analito) 1502 que comprende una porción de unión al analito 1504 y un oligonucleótido reportero 1508. La porción de unión al analito 1504 puede comprender un anticuerpo o fragmento de unión a antígeno de este que se une específicamente a un analito 1506. El agente de marcaje (por ejemplo, un agente de captura del analito) que puede capturar el analito es capaz, además, de interactuar con una sonda de captura con código de barras espacial. Por ejemplo, el oligonucleótido reportero 1508 puede comprender uno o más dominios de código de barras 1510 y una secuencia 1512 que es capaz de unirse a un dominio de captura de una sonda de captura. El oligonucleótido reportero 1508 puede comprender opcionalmente una o más secuencias funcionales.

La figura 16 es un diagrama esquemático que representa una interacción ilustrativa entre una sonda de captura que se inmoviliza por el elemento 1624 y un agente de marcaje (por ejemplo, un agente de captura del analito) 1626. La sonda de captura que se inmoviliza por el elemento 1624 puede unirse a un elemento 1602 (opcionalmente mediante un enlazador 1604 que puede ser escindible) y comprende un código de barras espacial 1608 así como también una o más secuencias funcionales 1606 y 1610, como se describe en otra parte de la presente descripción. La sonda de captura puede comprender, además, un dominio de captura 1612 que es capaz de unirse a un agente de marcaje (por ejemplo, un agente de captura del analito) 1626. El agente de marcaje (por ejemplo, agente de captura del analito) 1626 puede incluir una o más secuencias funcionales 1618, uno o más dominios de código de barras del agente de

captura 1616 y una secuencia de captura del analito 1614 que es capaz de unirse al dominio de captura 1612 de la sonda de captura 1624. El agente de marcaje (por ejemplo, agente de captura del analito) puede incluir, además, un enlazador 1620 que permite que el dominio de código de barras del agente de captura 1616 se acople a la porción de unión al analito 1622. Juntos, la sonda de captura que se inmoviliza por el elemento 1624 y el agente de marcaje (por ejemplo, agente de captura del analito) 1626 forman un agente de captura que comprende el agente de marcaje (por ejemplo, agente de captura del analito) que se acopla de manera liberable a la sonda de captura. En algunos ejemplos, el reportero 1408 en la figura 14 y/o el oligonucleótido reportero en la figura 15 comprenden la secuencia de captura del analito 1614, el uno o más dominios de código de barras del agente de captura 1616 y la una o más secuencias funcionales 1618.

En algunas modalidades, una porción de unión al analito puede incluir cualquier molécula o porción capaz de unirse a un analito (por ejemplo, un analito biológico, por ejemplo, un constituyente macromolecular). En algunas modalidades, de cualquiera de los métodos de caracterización espacial descritos en la presente descripción, la porción de unión al analito del agente de captura del analito que se une a un analito biológico puede incluir, pero no se limita a, un anticuerpo, o un fragmento de unión al epítipo de este, una molécula de unión a un receptor de superficie celular, un ligando del receptor, una molécula pequeña, un anticuerpo biespecífico, un activador de linfocitos T biespecífico, un activador de receptores de linfocitos T, un activador de receptores de linfocitos B, un procuerpo, un aptámero, un monocuerpo, un afímero, una darpina, y un andamio de proteínas, o cualquiera de sus combinaciones. La porción de unión al analito puede unirse al constituyente macromolecular (por ejemplo, analito) con alta afinidad y/o con alta especificidad. La porción de unión al analito puede incluir una secuencia de nucleótidos (por ejemplo, un oligonucleótido), que puede corresponder a al menos una porción o a la totalidad de la porción de unión al analito. La porción de unión al analito puede incluir un polipéptido y/o un aptámero (por ejemplo, un polipéptido y/o un aptámero que se une a una molécula diana específica, por ejemplo, un analito). La porción de unión al analito puede incluir un anticuerpo o fragmento de anticuerpo (por ejemplo, un fragmento de unión a antígeno) que se une a un analito específico (por ejemplo, un polipéptido).

En algunas modalidades, los agentes de captura del analito son capaces de unirse a analitos presentes dentro de una célula. En algunas modalidades, los agentes de captura del analito son capaces de unirse a analitos de la superficie celular que pueden incluir, sin limitación, un receptor, un antígeno, una proteína de superficie, una proteína transmembrana, un conglomerado de proteínas de diferenciación, una proteína canal, una proteína bomba, una proteína portadora, un fosfolípido, una glucoproteína, un glucolípido, un complejo de proteína de interacción célula-célula, un complejo presentador de antígenos, un complejo mayor de histocompatibilidad, un receptor de linfocitos T manipulado genéticamente, un receptor de linfocitos T, un receptor de linfocitos B, un receptor de antígeno quimérico, una proteína de la matriz extracelular, una modificación postraduccional (por ejemplo, fosforilación, glucosilación, ubiquitinación, nitrosilación, metilación, acetilación o lipidación) del estado de una proteína de la superficie celular, una unión en hendidura, y una unión adherente. En algunas modalidades, los agentes de captura del analito son capaces de unirse a analitos de la superficie celular que se modifican postraduccionalmente. En tales modalidades, los agentes de captura del analito pueden ser específicos para los analitos de la superficie celular en base a un estado dado de modificación postraduccional (por ejemplo, fosforilación, glucosilación, ubiquitinación, nitrosilación, metilación, acetilación o lipidación), de manera que un perfil del analito de la superficie celular puede incluir información de modificación postraduccional de uno o más analitos.

En algunas modalidades, el agente de captura del analito incluye un dominio de código de barras del agente de captura que se conjuga o se une de cualquier otra manera a la porción de unión al analito. En algunas modalidades, el dominio de código de barras del agente de captura se une covalentemente a la porción de unión al analito. En algunas modalidades, un dominio de código de barras del agente de captura es una secuencia de ácido nucleico. En algunas modalidades, un dominio de código de barras del agente de captura incluye un código de barras de la porción de unión al analito y una secuencia de captura del analito.

Como se usa en la presente, el término "código de barras de la porción de unión al analito" se refiere a un código de barras que se asocia con, o de cualquier otra manera, identifica la porción de unión al analito. En algunas modalidades, al identificar una porción de unión al analito mediante la identificación de su código de barras de la porción de unión al analito asociado, también puede identificarse el analito al que se une la porción de unión al analito. Un código de barras de la porción de unión al analito puede ser una secuencia de ácido nucleico de una longitud y/o secuencia dadas que se asocia con la porción de unión al analito. Un código de barras de la porción de unión al analito puede incluir generalmente cualquiera de la variedad de aspectos de códigos de barras descritos en la presente descripción. Por ejemplo, un agente de captura del analito que es específico para un tipo de analito puede tener acoplado a este un primer dominio de código de barras del agente de captura (por ejemplo, que incluye un primer código de barras de la porción de unión al analito), mientras que un agente de captura del analito que es específico para un analito diferente puede tener un dominio de código de barras del agente de captura diferente (por ejemplo, que incluye un segundo código de barras de la porción de unión al analito) que se acopla a este. En algunos aspectos, tal dominio de código de barras del agente de captura puede incluir un código de barras de la porción de unión al analito que permite la identificación de la porción de unión al analito al que se acopla el dominio de código de barras del agente de captura. La selección del dominio de código de barras del agente de captura puede permitir una diversidad significativa en términos de secuencia, mientras que también puede unirse fácilmente a la mayoría de las porciones de unión al analito (por ejemplo, anticuerpos) así como también detectarse fácilmente (por ejemplo, mediante el uso de tecnologías de

secuenciación o matriz). En algunas modalidades, los agentes de captura del analito pueden incluir porciones de unión al analito con dominios de código de barras de agente de captura que se une a ellos. Por ejemplo, un agente de captura del analito puede incluir una primera porción de unión al analito (por ejemplo, un anticuerpo que se une a un analito, por ejemplo, un primer elemento de la superficie celular) que tiene asociada con este un dominio de código de barras del agente de captura que incluye un primer código de barras de la porción de unión al analito.

En algunas modalidades, el dominio de código de barras del agente de captura de un agente de captura del analito incluye una secuencia de captura del analito. Como se usa en la presente descripción, el término "secuencia de captura del analito" se refiere a una región o porción que se configura para hibridarse, unirse, acoplarse o interactuar de cualquier otra manera con un dominio de captura de una sonda de captura. En algunas modalidades, una secuencia de captura del analito incluye una secuencia de ácido nucleico que es complementaria o sustancialmente complementaria al dominio de captura de una sonda de captura de manera que la secuencia de captura del analito se hibrida con el dominio de captura de la sonda de captura. En algunas modalidades, una secuencia de captura del analito comprende una secuencia de ácido nucleico poli(A) que se hibrida con un dominio de captura que comprende una secuencia de ácido nucleico poli(T). En algunas modalidades, una secuencia de captura del analito comprende una secuencia de ácido nucleico poli(T) que se hibrida con un dominio de captura que comprende una secuencia de ácido nucleico poli(A). En algunas modalidades, una secuencia de captura del analito comprende una secuencia de ácido nucleico que no es homopolimérica que se hibrida con un dominio de captura que comprende una secuencia de ácido nucleico que no es homopolimérica que es complementaria (o sustancialmente complementaria) a la secuencia de ácido nucleico que no es homopolimérica de la región de captura del analito.

En algunas modalidades de cualquiera de los métodos de análisis espacial descritos en la presente descripción que emplean un agente de captura del analito, el dominio de código de barras del agente de captura puede acoplarse directamente a la porción de unión al analito, o pueden unirse a una perla, red molecular, por ejemplo, una linear, globular, reticulado, u otro polímero, u otro marco que se une o se asocia de cualquier otra manera con la porción de unión al analito, lo que permite la unión de múltiples dominios de código de barras de agente de captura a una única porción de unión al analito. La unión (acoplamiento) de los dominios de código de barras del agente de captura a las porciones de unión al analito puede lograrse a través de cualquiera de una variedad de asociaciones o uniones directas o indirectas, covalentes o no covalentes. Por ejemplo, en el caso de un dominio de código de barras del agente de captura que se acopla a una porción de unión al analito que incluye un anticuerpo o fragmento de unión a antígeno, tales dominios de código de barras del agente de captura pueden unirse covalentemente a una porción del anticuerpo o fragmento de unión a antígeno mediante el uso de técnicas de conjugación química (por ejemplo, kits de marcaje de anticuerpos Lightning-Link® disponibles en Innova Biosciences). En algunas modalidades, un dominio de código de barras del agente de captura puede acoplarse a un anticuerpo o fragmento de unión a antígeno mediante el uso de mecanismos de unión no covalentes (por ejemplo, mediante el uso de anticuerpos biotinilados y oligonucleótidos o perlas que incluyen uno o más enlazadores biotinilados, que se acoplan a oligonucleótidos con un enlazador de avidina o estreptavidina). Pueden usarse técnicas de biotinilación de anticuerpos y oligonucleótidos, y se describen por ejemplo en Fang y otros, *Nucleic Acids Res.* (2003), 31(2): 708-715. Igualmente, se han desarrollado técnicas de biotinilación de proteínas y péptidos y pueden usarse, y se describen por ejemplo en la patente de Estados Unidos núm. 6,265,552. Además, la química de la reacción clic tal como una reacción de éster de metiltetrazina-PEG5-NHS, una reacción de éster de TCO-PEG4-NHS, o similares, puede usarse para acoplar dominios de código de barras del agente de captura a porciones de unión al analito. La porción reactiva en la porción de unión al analito también puede incluir amina para dirigirse a aldehídos, amina para dirigirse a maleimida (por ejemplo, tioles libres), azida para dirigirse a compuestos químicos clic (por ejemplo, alquinos), biotina para dirigirse a estreptavidina, fosfatos para dirigirse a EDC, que a su vez se dirige al éster activo (por ejemplo, NH₂). La porción reactiva en la porción de unión al analito puede ser un compuesto químico o grupo que se une a la porción reactiva sobre la porción de unión al analito. Las estrategias ilustrativas para conjugar la porción de unión al analito al dominio de código de barras del agente de captura incluyen el uso de kits comerciales (por ejemplo, Solulink, Thunder link), conjugación de reducción leve de la región de bisagra y marcaje de maleimida, reacción química clic promovida por tinción a amidas marcadas (por ejemplo, sin cobre) y conjugación de oxidación peryodada de cadena de azúcar y conjugación de amina. En los casos en los que la porción de unión al analito es un anticuerpo, el anticuerpo puede modificarse previo a, o simultáneamente con la conjugación del oligonucleótido. Por ejemplo, el anticuerpo puede glicosilarse con un mutante que permite el sustrato de β -1,4-galactosiltransferasa, GalT (Y289L) y uridina difosfato de N-acetilgalactosamina análogo de uridina difosfato de -GalNAz que porta azida. El anticuerpo modificado puede conjugarse a un oligonucleótido con un grupo dibenzociclooctina-PEG4-NHS. En algunas modalidades, pueden evitarse determinadas etapas (por ejemplo, activación de COOH (por ejemplo, EDC) y reticuladores homobifuncionales) para evitar que las porciones de unión al analito se conjuguen a sí mismos. En algunas modalidades de cualquiera de los métodos de caracterización espacial descritos en la presente descripción, el agente de captura del analito (por ejemplo, una porción de unión al analito que se acopla a un oligonucleótido) puede suministrarse en la célula, por ejemplo, mediante transfección (por ejemplo, mediante el uso de transfectamina, polímeros catiónicos, fosfato cálcico o electroporación), mediante transducción (por ejemplo, mediante el uso de un bacteriófago o vector viral recombinante), mediante suministro mecánico (por ejemplo, perlas magnéticas), mediante lípidos (por ejemplo, 1,2-Dioleoil-sn-glicero-3-fosfolina (DOPC)), o mediante proteínas transportadoras. Un agente de captura del analito puede suministrarse a una célula mediante el uso de exosomas. Por ejemplo, puede generarse una primera célula que libera exosomas que comprenden un agente de captura del analito. Un agente de captura del analito puede unirse a una membrana exosómica. Un agente de captura del analito puede estar contenido dentro del citosol de un exosoma. Los exosomas que se liberan pueden cosecharse

y proporcionarse a una segunda célula, de esta manera se suministra el agente de captura del analito en la segunda célula. Un agente de captura del analito puede liberarse de una membrana del exosoma antes, durante o después del suministro a una célula. En algunas modalidades, la célula se permeabiliza para permitir que el agente de captura del analito se acople con constituyentes celulares intracelulares (tales como, sin limitación, proteínas intracelulares, metabolitos y proteínas de la membrana nuclear). Después del suministro intracelular, pueden usarse agentes de captura del analito para analizar los constituyentes intracelulares como se describe en la presente descripción.

En algunas modalidades de cualquiera de los métodos de caracterización espacial descritos en la presente descripción, el dominio de código de barras del agente de captura que se acopla a un agente de captura del analito puede incluir modificaciones que lo hacen no extensible por una polimerasa. En algunas modalidades, cuando se une a un dominio de captura de una sonda de captura o ácido nucleico en una muestra para una reacción de extensión del cebador, el dominio de código de barras del agente de captura puede servir como un molde, no como un cebador. Cuando el dominio de código de barras del agente de captura también incluye un código de barras (por ejemplo, un código de barras de la porción de unión al analito), tal diseño puede aumentar la eficiencia de la codificación molecular mediante código de barras por el aumento de la afinidad entre el dominio de código de barras del agente de captura y los ácidos nucleicos de la muestra sin codificar, y elimina la formación potencial de artefactos de adaptadores. En algunas modalidades, el dominio de código de barras del agente de captura puede incluir una secuencia de N-mer aleatoria que se protege con modificaciones que la hacen no extensible por una polimerasa. En algunos casos, la composición de la secuencia de N-mer aleatoria puede diseñarse para maximizar la eficiencia de unión a moléculas de ADNmc libres, no codificadas mediante código de barras. El diseño puede incluir una composición de secuencia aleatoria con un mayor contenido de GC, una secuencia aleatoria parcial con G o C fija en posiciones específicas, el uso de guanosinas, el uso de ácidos nucleicos bloqueados o cualquiera de sus combinaciones.

Una modificación para bloquear la extensión del cebador por una polimerasa puede ser un grupo separador de carbono de diferentes longitudes o un dideoxinucleótido. En algunas modalidades, la modificación puede ser un sitio abásico que tiene una estructura de apurina o apirimidina, un análogo de base o un análogo de una cadena principal de fosfato, tal como una cadena principal de N-(2-aminoetil)-glicina que se une mediante enlaces amida, tetrahidrofurano o 1',2'-Dideoxirribosa. La modificación puede ser, además, una base de uracilo, ARN modificado 2'OMe, separadores C3-18 (por ejemplo, estructuras con átomos de carbono consecutivos 3-18, tales como espaciador C3), separadores de multímeros de etilenglicol (por ejemplo, espaciador 18 (espaciador de hexa-etilenglicol), biotina, trifosfato de dideoxinucleótido, etilenglicol, amina o fosfato).

En algunas modalidades de cualquiera de los métodos de caracterización espacial descritos en la presente descripción, el dominio de código de barras del agente de captura que se acopla a la porción de unión al analito incluye un dominio escindible. Por ejemplo, después de que el agente de captura del analito se une a un analito (por ejemplo, un analito de la superficie celular), el dominio de código de barras del agente de captura puede escindirse y recolectarse para el análisis aguas abajo de acuerdo con los métodos como se describe en la presente descripción. En algunas modalidades, el dominio escindible del dominio de código de barras del agente de captura incluye un elemento de escisión en U que permite que las especies se liberen de la perla. En algunas modalidades, el elemento de escisión en U puede incluir una secuencia de ADN monocatenario (ADNmc) que contiene al menos un uracilo. Las especies pueden unirse a una perla mediante la secuencia de ADNmc. Las especies pueden liberarse mediante una combinación de uracilo-ADN glucosilasa (por ejemplo, para eliminar el uracilo) y una endonucleasa (por ejemplo, para inducir una ruptura del ADNmc). Si la endonucleasa genera un grupo fosfato 5' a partir de la escisión, después puede incluirse un tratamiento enzimático adicional en el procesamiento aguas abajo para eliminar el grupo fosfato, por ejemplo, previo a la ligación de elementos adicionales de manejo de secuenciación, por ejemplo, secuencia de P5 completa de Illumina, secuencia parcial de P5, secuencia completa de R1 y/o secuencia parcial de R1.

En algunas modalidades, una porción de unión al analito de un agente de captura de un analito incluye uno o más anticuerpos o fragmentos de unión a antígeno de estos. Los anticuerpos o fragmentos de unión a antígeno que incluyen la porción de unión al analito pueden unirse específicamente a un analito diana. En algunas modalidades, el analito es una proteína (por ejemplo, una proteína sobre una superficie de la muestra biológica (por ejemplo, una célula) o una proteína intracelular). En algunas modalidades, una pluralidad de agentes de captura del analito que comprenden una pluralidad de porciones de unión al analito se une a una pluralidad de analitos presentes en una muestra biológica. En algunas modalidades, la pluralidad de analitos incluye una única especie del analito (por ejemplo, una única especie de polipéptido). En algunas modalidades en las que la pluralidad de analitos incluye una única especie del analito, las porciones de unión al analito de la pluralidad de agentes de captura del analito son los mismos. En algunas modalidades en las que la pluralidad de analitos incluye una única especie del analito, las porciones de unión al analito de la pluralidad de agentes de captura del analito son diferentes (por ejemplo, los miembros de la pluralidad de agentes de captura del analito pueden tener dos o más especies de porciones de unión al analito, en donde cada una de las dos o más especies de porciones de unión al analito se unen a una única especie del analito, por ejemplo, en sitios de unión diferentes). En algunas modalidades, la pluralidad de analitos incluye múltiples especies diferentes de analitos (por ejemplo, múltiples especies diferentes de polipéptidos).

En algunas modalidades, múltiples especies diferentes de analitos (por ejemplo, polipéptidos) de la muestra biológica pueden asociarse subsecuentemente con la una o más propiedades físicas de la muestra biológica. Por ejemplo, las múltiples especies diferentes de analitos pueden asociarse con ubicaciones de los analitos en la muestra biológica.

Tal información (por ejemplo, información proteómica cuando el o la(s) porción(es) de unión al analito reconocen un(os) polipéptido(s)) puede usarse en asociación con otra información espacial (por ejemplo, información genética de la muestra biológica, tal como información de la secuencia de ADN, información del transcriptoma (es decir, secuencias de transcritos), o ambas). Por ejemplo, una proteína de la superficie celular de una célula puede asociarse con una o más propiedades físicas de la célula (por ejemplo, una forma, tamaño, actividad, o un tipo de célula). La una o más propiedades físicas pueden caracterizarse mediante la obtención de imágenes de la célula. La célula puede unirse mediante un agente de captura del analito que comprende una porción de unión al analito que se une a la proteína de la superficie celular y un código de barras de la porción de unión al analito que identifica esa porción de unión al analito, y la célula puede someterse a análisis espacial (por ejemplo, cualquiera de la variedad de métodos de análisis espacial descrito en la presente descripción). Por ejemplo, el agente de captura del analito que se une a la proteína de la superficie celular puede unirse a una sonda de captura (por ejemplo, una sonda de captura sobre una matriz), dicha sonda de captura incluye un dominio de captura que interactúa con una secuencia de captura del analito presente sobre el dominio de código de barras del agente de captura del analito. Todo o parte del dominio de código de barras del agente de captura (que incluye el código de barras de la porción de unión al analito) puede copiarse con una polimerasa mediante el uso de un extremo 3' del dominio de captura como sitio de imprimación, lo que genera una sonda de captura extendida que incluye toda o parte de la sonda de captura (que incluye un código de barras espacial presente sobre la sonda de captura) y una copia del código de barras de la porción de unión al analito. En algunas modalidades, la matriz espacial con la(s) sonda(s) de captura extendida puede ponerse en contacto con una muestra, donde el/los agente(s) de captura del analito asociados con la matriz espacial capturan el/los analito(s) diana. El/los agente(s) de captura del analito que contienen la(s) sonda(s) de captura extendida, que incluye el/los código(s) de barras espacial(es) de la(s) sonda(s) de captura y el/los código(s) de barras de la porción de unión al analito, pueden después desnaturalizarse de la(s) sonda(s) de captura de la matriz espacial. Esto permite que la matriz espacial se reutilice. La muestra puede disociarse en células no agregadas (por ejemplo, células únicas) y analizarse mediante los métodos de célula única/gota descritos en la presente descripción. La sonda de captura extendida puede secuenciarse para obtener una secuencia de ácido nucleico, en la que el código de barras espacial de la sonda de captura se asocia con el código de barras de la porción de unión al analito del agente de captura del analito. Por lo tanto, la secuencia de ácido nucleico de la sonda de captura extendida puede asociarse con el analito (por ejemplo, proteína de la superficie celular) y, a su vez, con la una o más propiedades físicas de la célula (por ejemplo, una forma o tipo de célula). En algunas modalidades, la secuencia de ácido nucleico de la sonda de captura extendida puede asociarse con un analito intracelular de una célula cercana, donde el analito intracelular se libera mediante el uso de cualquiera de las técnicas de permeabilización celular o migración del analito descritas en la presente descripción.

En algunas modalidades de cualquiera de los métodos de caracterización espacial descritos en la presente descripción, los dominios de código de barras del agente de captura que se liberan de los agentes de captura del analito pueden someterse después al análisis de secuencia para identificar qué agentes de captura del analito se unieron a los analitos. En base a los dominios de código de barras del agente de captura que se asocian con un elemento (por ejemplo, un elemento en una ubicación particular) sobre una matriz espacial y la presencia de la secuencia de código de barras de la porción de unión al analito, puede crearse un perfil del analito para una muestra biológica. Los perfiles de células individuales o poblaciones de células pueden compararse con perfiles de otras células, por ejemplo, células "normales", para identificar variaciones en los analitos, que pueden proporcionar información diagnóstica relevante. En algunas modalidades, estos perfiles pueden ser útiles en el diagnóstico de una variedad de trastornos que se caracterizan por variaciones en los receptores de la superficie celular, tales como el cáncer y otros trastornos.

VII. Sustratos y matrices

En algunas modalidades, un sustrato en la presente descripción (por ejemplo, el primer sustrato y/o el segundo sustrato en la presente descripción) puede ser cualquier soporte que es insoluble en un líquido acuoso y que permite el posicionamiento de muestras biológicas, analitos, elementos, y/o reactivos (por ejemplo, sondas tal como sondas de captura) sobre el soporte. En algunas modalidades, una muestra biológica puede unirse a un sustrato. La unión de la muestra biológica puede ser irreversible o reversible, en dependencia de la naturaleza de la muestra y de las etapas subsecuentes en el método analítico. En determinadas modalidades, la muestra puede unirse al sustrato de manera reversible mediante la aplicación de un recubrimiento de polímero adecuado al sustrato, y poniendo en contacto la muestra con el recubrimiento de polímero. Después, la muestra puede desprenderse del sustrato, por ejemplo, mediante el uso de un solvente orgánico que disuelve al menos parcialmente el recubrimiento de polímero. Los hidrogeles son ejemplos de polímeros que son adecuados para este propósito.

En algunas modalidades, el sustrato puede recubrirse o funcionalizarse con una o más sustancias para facilitar la unión de la muestra al sustrato. Las sustancias adecuadas que pueden usarse para recubrir o funcionalizar el sustrato incluyen, pero no se limitan a, lectinas, polilisina, anticuerpos y polisacáridos.

En general, un sustrato funciona como un soporte para la unión directa o indirecta de las sondas de captura a elementos de la matriz. Además, en algunas modalidades, un sustrato (por ejemplo, el mismo sustrato o un sustrato diferente) puede usarse para proporcionar soporte a una muestra biológica, particularmente, por ejemplo, una sección

de tejido delgado. En consecuencia, un "sustrato" es un soporte que es insoluble en líquido acuoso y que permite el posicionamiento de muestras biológicas, analitos, elementos y/o sondas de captura sobre el sustrato.

Una amplia variedad de diferentes sustratos puede usarse para los propósitos anteriores. En general, un sustrato puede ser cualquier material de soporte adecuado. Los sustratos ilustrativos incluyen, pero no se limitan a, vidrio y vidrio modificado y/o funcionalizado, hidrogeles, películas, membranas, plásticos (que incluyen, por ejemplo, acrílicos, poliestireno, copolímeros de estireno y otros materiales, polipropileno, polietileno, polibutileno, poliuretanos, Teflon™, olefinas cíclicas, poliimidas, etc.), nailon, cerámica, resinas, Zeonor, sílice o materiales basados en sílice, que incluyen silicio y silicio modificado, carbono, metales, vidrios inorgánicos, haces de fibras ópticas y polímeros, tales como poliestireno, copolímeros de olefinas cíclicas (COC), polímeros de olefinas cíclicas (COP), polipropileno, polietileno y policarbonato.

El sustrato puede, además, corresponder a una celda de flujo. Las celdas de flujo pueden formarse de cualquiera de los materiales anteriores, y pueden incluir canales que permiten que los reactivos, solventes, elementos y moléculas pasen a través de la celda.

Entre los ejemplos de materiales de sustrato mencionados anteriormente, el poliestireno es un material hidrófobo adecuado para unir macromoléculas cargadas negativamente porque normalmente contiene pocos grupos hidrófilos. Para los ácidos nucleicos que se inmovilizan sobre portaobjetos de vidrio, al aumentar la hidrofobicidad de la superficie vítrea puede aumentarse la inmovilización del ácido nucleico. Tal mejora puede permitir una formación relativamente más densamente empaquetada (por ejemplo, proporcionar una especificidad y resolución mejoradas).

En algunas modalidades, un sustrato se recubre con un tratamiento de superficie tal como poli(L)-lisina. Adicional o alternativamente, el sustrato puede tratarse mediante silanación, por ejemplo, con epoxi-silano, amino-silano, y/o mediante un tratamiento con poliacrilamida.

El sustrato puede tener generalmente cualquier forma o formato adecuado. Por ejemplo, el sustrato puede ser plano, curvado, por ejemplo, curvado de forma convexa o cóncava hacia el área donde tiene lugar la interacción entre una muestra biológica, por ejemplo, muestra de tejido y el sustrato. En algunas modalidades, el sustrato es plano, por ejemplo, planar, chip o portaobjetos. El sustrato puede contener una o más superficies con patrones dentro del sustrato (por ejemplo, canales, pocillos, proyecciones, crestas, hendiduras, etc.).

Un sustrato puede tener cualquier forma deseada. Por ejemplo, un sustrato puede tener típicamente una forma delgada y plana (por ejemplo, un cuadrado o un rectángulo). En algunas modalidades, una estructura de sustrato tiene esquinas redondas (por ejemplo, para una mayor seguridad o robustez). En algunas modalidades, una estructura de sustrato tiene una o más esquinas cortadas (por ejemplo, para su uso con una abrazadera deslizante o mesa transversal). En algunas modalidades, donde una estructura de sustrato es plana, la estructura de sustrato puede ser cualquier tipo apropiado de soporte que tenga una superficie plana (por ejemplo, un chip o un portaobjetos tal como un portaobjetos de microscopio).

Los sustratos pueden incluir opcionalmente varias estructuras tales como, pero no se limitan a, proyecciones, crestas, y canales. Un sustrato puede tener micropatrones para limitar la difusión lateral (por ejemplo, para evitar el solapamiento de códigos de barras espaciales). Un sustrato modificado con tales estructuras puede modificarse para permitir la asociación de analitos, elementos o sondas en sitios individuales. Por ejemplo, los sitios donde un sustrato se modifica con varias estructuras pueden ser contiguos o no contiguos con otros sitios.

En algunas modalidades, la superficie de un sustrato puede modificarse de manera que se forman sitios discretos que solo pueden tener o acomodar un único elemento. En algunas modalidades, la superficie de un sustrato puede modificarse de manera que los elementos se adhieran a sitios aleatorios.

En algunas modalidades, la superficie de un sustrato se modifica para contener uno o más pocillos, mediante el uso de técnicas tales como (pero no limitadas a) técnicas de estampado, técnicas de micrograbado y técnicas de moldeo. En algunas modalidades en las que un sustrato incluye uno o más pocillos, el sustrato puede ser un portaobjetos de concavidad o portaobjetos de cavidad. Por ejemplo, los pocillos pueden formarse por una o más depresiones poco profundas en la superficie del sustrato. En algunas modalidades, donde un sustrato incluye uno o más pocillos, los pocillos pueden formarse mediante la unión de un casete (por ejemplo, un casete que contiene una o más cámaras) a una superficie de la estructura de sustrato.

En algunas modalidades, las estructuras de un sustrato, (por ejemplo, pocillos) pueden portar cada una sonda de captura diferente. Pueden identificarse diferentes sondas de captura que se unen a cada estructura de acuerdo con las ubicaciones de las estructuras en o sobre la superficie del sustrato. Los sustratos ilustrativos incluyen matrices en las que las estructuras separadas se localizan sobre el sustrato que incluye, por ejemplo, aquellos que tienen pocillos que acomodan los elementos.

En algunas modalidades, un sustrato incluye una o más marcas sobre una superficie del sustrato, por ejemplo, para proporcionar orientación para correlacionar la información espacial con la caracterización del analito de interés. Por

ejemplo, un sustrato puede marcarse con una rejilla de líneas (por ejemplo, para permitir que el tamaño de los objetos vistos con aumento se calcule fácilmente y/o para proporcionar áreas de referencia para contar objetos). En algunas modalidades, los marcadores fiduciales pueden incluirse sobre el sustrato. Tales marcas pueden hacerse mediante el uso de técnicas que incluyen, pero no se limitan a, impresión, pulimento y deposición sobre la superficie.

En algunas modalidades donde el sustrato se modifica para contener una o más estructuras, que incluyen pero no se limitan a, pocillos, proyecciones, crestas, o marcas, las estructuras pueden incluir sitios físicamente alterados. Por ejemplo, un sustrato modificado con varias estructuras puede incluir propiedades físicas, que incluyen, pero no se limitan a, configuraciones físicas, fuerzas magnéticas o de compresión, sitios químicamente funcionalizados, sitios químicamente alterados y/o sitios alterados electrostáticamente.

En algunas modalidades donde el sustrato se modifica para contener varias estructuras, que incluyen pero no se limitan a pocillos, proyecciones, crestas, o marcas, las estructuras se aplican en un patrón. Alternativamente, las estructuras pueden distribuirse aleatoriamente.

En algunas modalidades, un sustrato se trata para minimizar o reducir la hibridación inespecífica del analito dentro de o entre los elementos. Por ejemplo, el tratamiento puede incluir recubrir el sustrato con un hidrogel, película y/o membrana que crea una barrera física a la hibridación inespecífica. Puede usarse cualquier hidrogel adecuado. Por ejemplo, pueden usarse matrices de hidrogel preparadas de acuerdo con los métodos establecidos en las patentes de Estados Unidos núms. 6,391,937, 9,512,422 y 9,889,422, y en las publicaciones de solicitud de patente de Estados Unidos núms. U.S. 2017/0253918 y U.S. 2018/0052081.

El tratamiento puede incluir añadir un grupo funcional que es reactivo o capaz de activarse de manera que se vuelva reactivo después de recibir un estímulo (por ejemplo, fotorreactivo). El tratamiento puede incluir tratar con polímeros que tienen una o más propiedades físicas (por ejemplo, mecánicas, eléctricas, magnéticas y/o térmicas) que minimizan la unión inespecífica (por ejemplo, que activan un sustrato en determinadas ubicaciones para permitir la hibridación del analito en esas ubicaciones).

El sustrato (por ejemplo, un elemento sobre una matriz) puede incluir decenas a cientos de miles o millones de moléculas de oligonucleótidos individuales (por ejemplo, al menos aproximadamente 10 000, 50 000, 100 000, 500 000, 1 000 000, 10 000 000, 100 000 000, 1 000 000 000 o 10 000 000 000 moléculas de oligonucleótidos).

En algunas modalidades, la superficie de un sustrato se recubre con un recubrimiento permisivo para las células para permitir la adherencia de células vivas. Un "recubrimiento permisivo para las células" es un recubrimiento que permite o ayuda a las células a mantener la viabilidad celular (por ejemplo, permanecer viables) sobre el sustrato. Por ejemplo, un recubrimiento permisivo para las células puede mejorar la unión celular, el crecimiento celular y/o la diferenciación celular, por ejemplo, un recubrimiento permisivo para las células puede proporcionar nutrientes a las células vivas. Un recubrimiento permisivo para las células puede incluir un material biológico y/o un material sintético. Los ejemplos no limitantes de un recubrimiento permisivo para las células incluyen recubrimientos que presentan uno o más componentes de matriz extracelular (ECM) (por ejemplo, proteoglicanos y proteínas fibrosas tales como colágeno, elastina, fibronectina y laminina), polilisina, poli(L)-ornitina y/o una silicona biocompatible (por ejemplo, CYTOSOFT®). Por ejemplo, un recubrimiento permisivo para las células que incluye uno o más componentes de la matriz extracelular puede incluir colágeno tipo I, colágeno tipo II, colágeno tipo IV, elastina, fibronectina, laminina y/o vitronectina. En algunas modalidades, el recubrimiento permisivo para las células incluye una preparación de membrana basal solubilizada que se extrae del sarcoma de ratón Engelbreth-Holm-Swarm (EHS) (por ejemplo, MATRIGEL®). En algunas modalidades, el recubrimiento permisivo para las células incluye colágeno. Puede usarse un recubrimiento permisivo para las células para cultivar células adherentes sobre una matriz de código de barras espacial, o para mantener la viabilidad celular de una muestra o sección de tejido mientras está en contacto con una matriz de código de barras espacial.

A. Sustratos para módulos de ensayo *in situ*

En algunas modalidades, se proporciona una muestra biológica sobre un primer sustrato para uno o más módulos de ensayo *in situ* del ensayo integrado descrito en la presente descripción. En algunas modalidades, la muestra biológica sobre el primer sustrato se pone en contacto con una o más sondas de ácido nucleico para uno o más módulos de ensayo *in situ*. La una o más sondas de ácido nucleico pueden hibridarse directa o indirectamente con un primer ácido nucleico diana o un complemento o un producto de amplificación de este en la muestra biológica. En algunas modalidades, el primer sustrato comprende una pluralidad de agentes de captura que se inmovilizan sobre el mismo, y los agentes de captura son capaces de capturar directa o indirectamente un segundo ácido nucleico diana o un complemento de este o un producto de amplificación de este.

Puede usarse una amplia variedad de diferentes sustratos para el módulo de ensayo *in situ*, siempre que el sustrato sea compatible con la muestra y con el procesamiento de muestras, los reactivos y reacciones *in situ*, y la detección de señales *in situ* (por ejemplo, imágenes ópticas tales como microscopía de fluorescencia). Un sustrato puede ser cualquier material de soporte adecuado y generalmente es transparente. Por ejemplo, puede usarse un portaobjeto de vidrio tal como un cubreobjetos. El primer sustrato puede incluir, pero no se limita a, vidrio, vidrio modificado y/o

funcionalizado, hidrogeles, películas, membranas, plásticos, nailon, cerámica, resinas, Zeonor, sílice o materiales basados en sílice, que incluyen silicio y silicio modificado, carbono, metales, vidrios inorgánicos, haces de fibras ópticas y polímeros, tales como poliestireno, copolímeros de olefinas cíclicas (COC), polímeros de olefinas cíclicas (COP), polipropileno, polietileno y policarbonato. El primer sustrato puede, además, corresponder a una celda de flujo.

En algunas modalidades, el primer sustrato está entre aproximadamente 0,01 mm y aproximadamente 5 mm, por ejemplo, entre aproximadamente 0,05 mm y aproximadamente 3 mm, entre aproximadamente 0,1 mm y aproximadamente 2,5 mm, entre aproximadamente 0,2 mm y aproximadamente 2 mm, entre aproximadamente 0,5 mm y aproximadamente 1,5 mm, o aproximadamente 1 mm de grosor. En algunas modalidades, el primer sustrato es o es aproximadamente 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9 o 2,0 mm de grosor, o de un grosor entre cualquiera de los valores mencionados anteriormente.

B. Sustratos para módulos de ensayo espacial

En algunas modalidades, el primer sustrato y el segundo sustrato son los mismos. En algunas modalidades, el mismo sustrato puede usarse para uno o más módulos de ensayo *in situ* así como también uno o más módulos de ensayo espacial descritos en la presente descripción. Por ejemplo, una muestra biológica se une a un sustrato previo a un módulo de ensayo *in situ* y permanece unida al mismo sustrato durante la obtención de imágenes *in situ* y un análisis espacial, que incluye, por ejemplo, la captura de los analitos mediante agentes de captura que se proporcionan por el mismo sustrato. En algunas modalidades, el mismo sustrato comprende una pluralidad de agentes de captura que se inmovilizan sobre el mismo. En algunas modalidades, la pluralidad de agentes de captura permanece inmovilizada sobre el sustrato durante un ensayo *in situ* y un ensayo espacial, y las moléculas en la muestra biológica se liberan, suministran y/o se dirigen hacia el sustrato para que los agentes de captura capturen las moléculas. En algunas modalidades, la pluralidad de agentes de captura permanece inmovilizada sobre el sustrato durante un ensayo *in situ*, pero se libera, suministra y/o se dirige hacia la muestra biológica y/o moléculas en ella y/o sobre ella para un ensayo espacial.

En algunas modalidades, el primer sustrato y el segundo sustrato son diferentes. En algunas modalidades, el primer sustrato y el segundo sustrato son sustratos separados. Por ejemplo, el primer sustrato que tiene una muestra que se une al mismo para un ensayo *in situ* puede no comprender una pluralidad de agentes de captura que se inmovilizan sobre el primer sustrato. En su lugar, los agentes de captura se proporcionan sobre uno o más segundos sustratos, que se proporcionan a la muestra biológica durante o después de un módulo de ensayo *in situ*. Por ejemplo, un primer sustrato, una muestra biológica y un segundo sustrato pueden formar un sándwich para facilitar la interacción molecular y/o la transferencia de materiales entre la muestra y los sustratos. En algunas modalidades, la pluralidad de agentes de captura permanece inmovilizada sobre el segundo sustrato durante un ensayo espacial, y las moléculas en la muestra biológica se liberan, suministran y/o se dirigen hacia el segundo sustrato para que los agentes de captura capturen las moléculas. En algunas modalidades, la pluralidad de agentes de captura se libera del segundo sustrato. En algunas modalidades, la pluralidad de agentes de captura del segundo sustrato se suministra y/o se conduce hacia la muestra biológica y/o moléculas en ella y/o sobre ella.

En algunas modalidades, un segundo sustrato que comprende una pluralidad de agentes de captura se proporciona a una muestra después de la obtención de imágenes *in situ* de la muestra sobre un primer sustrato. La pluralidad de agentes de captura puede liberarse del segundo sustrato y suministrarse y/o conducirse hacia la muestra biológica y/o moléculas, en ella y/o sobre ella, sobre el primer sustrato, y las interacciones de moléculas (por ejemplo, captura del analito) y las etapas posteriores de un módulo de ensayo espacial (por ejemplo, transcripción inversa de moléculas de ARNm capturadas) se realiza sobre el primer sustrato.

Aún en otras modalidades, después de un módulo de ensayo *in situ*, la muestra biológica se acerca con un segundo sustrato para un módulo de ensayo espacial. En algunas modalidades, el primer sustrato no se elimina de la muestra biológica, y el sándwich que se forma por la muestra y los sustratos se usan en un módulo de ensayo espacial. En otras modalidades, el primer sustrato puede eliminarse. En algunas modalidades, después de un módulo de ensayo *in situ* sobre un primer sustrato, la muestra biológica se transfiere sobre un segundo sustrato que comprende una pluralidad de agentes de captura capaces de capturar analitos en o sobre la muestra para codificar, mediante código de barras, especialmente los analitos capturados. Un dispositivo puede usarse para facilitar la transferencia de una muestra entre sustratos, por ejemplo, desde un portaobjetos de microscopio ordinario a un sustrato que comprende una pluralidad de agentes de captura (por ejemplo, un césped de sondas de captura) que se inmovilizan sobre el mismo.

En cualquiera de las modalidades anteriores, los agentes de captura no necesitan proporcionarse sobre el mismo sustrato en el que está la muestra biológica. En otras palabras, la muestra biológica puede estar sobre un primer sustrato para el análisis *in situ*, y las moléculas en la muestra que han estado a través del análisis *in situ* pueden ponerse en contacto con y/o transferirse a uno o más segundos sustratos. En el caso de múltiples segundos sustratos, la muestra puede ponerse en contacto con cada segundo sustrato secuencialmente, o en paralelo si dos o más segundos sustratos cubren solo una subregión de la muestra. Uno o más de los segundos sustratos pueden ponerse en contacto con uno o más terceros sustratos para producir réplicas del/de los segundo(s) sustrato(s).

Puede usarse una amplia variedad de diferentes sustratos para el módulo de ensayo espacial, siempre que el sustrato sea compatible con la muestra y el procesamiento de la muestra, los reactivos y reacciones del ensayo espacial, y la preparación de bibliotecas de secuenciación para la lectura espacial. Un sustrato puede ser cualquier material de soporte adecuado. El sustrato para un módulo de ensayo espacial puede, pero no necesita ser transparente, si no se usa también para un módulo de ensayo *in situ*. Por ejemplo, puede usarse un portaobjeto de vidrio tal como un cubreobjetos. El sustrato para un módulo de ensayo espacial (por ejemplo, el primer sustrato o el segundo sustrato) puede incluir, pero no se limitan a, vidrio, vidrio modificado y/o funcionalizado, hidrogeles, películas, membranas, plásticos, nailon, cerámica, resinas, Zeonor, sílice o materiales a base de sílice que incluyen silicio y silicio modificado, carbono, metales, vidrios inorgánicos, conjuntos de fibra óptica, y polímeros, tales como el poliestireno, copolímeros de olefina cíclica (COC), polímeros de olefina cíclica (COP), polipropileno, polietileno y policarbonato. El sustrato para un módulo de ensayo espacial (por ejemplo, el primer o segundo sustrato) puede corresponder a, además, una celda de flujo.

En algunas modalidades, el sustrato para un módulo de ensayo espacial (por ejemplo, el primer o segundo sustrato) está entre aproximadamente 0,01 mm y aproximadamente 5 mm, por ejemplo, entre aproximadamente 0,05 mm y aproximadamente 3 mm, entre aproximadamente 0,1 mm y aproximadamente 2,5 mm, entre aproximadamente 0,2 mm y aproximadamente 2 mm, entre aproximadamente 0,5 mm y aproximadamente 1,5 mm, o aproximadamente 1 mm de grosor. En algunas modalidades, el sustrato para un módulo de ensayo espacial es o es aproximadamente 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9 o 2,0 mm de grosor, o de un grosor entre cualquiera de los valores mencionados anteriormente.

En muchos de los métodos descritos en la presente descripción, los elementos (como se describe posteriormente más abajo) se posicionan colectivamente sobre un sustrato. Una "matriz" es una disposición específica de una pluralidad de elementos que es irregular o forma un patrón regular. Los elementos individuales en la matriz difieren entre sí sobre la base de sus ubicaciones espaciales relativas. En general, al menos dos de la pluralidad de elementos en la matriz incluyen una sonda de captura distinta (por ejemplo, cualquiera de los ejemplos de sondas de captura descritas en la presente descripción).

Las matrices pueden usarse para medir un gran número de analitos simultáneamente. En algunas modalidades, los oligonucleótidos se usan, al menos en parte, para crear una matriz. Por ejemplo, una o más copias de una única especie de oligonucleótido (por ejemplo, sonda de captura) pueden corresponder o unirse directa o indirectamente a un elemento dado en la matriz. En algunas modalidades, un elemento dado en la matriz incluye dos o más especies de oligonucleótidos (por ejemplo, sondas de captura). En algunas modalidades, las dos o más especies de oligonucleótidos (por ejemplo, sondas de captura) que se unen directa o indirectamente a un elemento dado sobre la matriz incluyen un código de barras espacial común (por ejemplo, idéntico).

Un "elemento" es una entidad que actúa como un soporte o depósito para varias entidades moleculares que se usan en el análisis de muestras. Los ejemplos de elementos incluyen, pero no se limitan a, una perla, un punto de cualquier geometría bi- o tridimensional (por ejemplo, un punto de chorro de tinta, un punto enmascarado, un cuadrado sobre una rejilla), un pocillo, y una almohadilla de hidrogel. En algunas modalidades, los elementos se unen o se fijan directa o indirectamente a un sustrato. En algunas modalidades, los elementos no se unen o se fijan directa o indirectamente a un sustrato, sino que, por ejemplo, se disponen dentro de un espacio tridimensional cerrado o parcialmente cerrado (por ejemplo, pocillos o hendiduras).

Además de los anteriores, puede usarse una amplia variedad de otros elementos para formar las matrices descritas en la presente descripción. Por ejemplo, en algunas modalidades, pueden usarse elementos que se forman a partir de polímeros y/o biopolímeros que se imprimen por chorro, se serigrafian o se depositan electrostáticamente sobre un sustrato para formar matrices. La impresión por chorro de biopolímeros se describe, por ejemplo, en la publicación de solicitud de patente PCT núm. WO 2014/085725. La impresión por chorro de polímeros se describe, por ejemplo, en de Gans y otros, *Adv Mater.* 16(3): 203-213 (2004). Los métodos para la deposición electrostática de polímeros y biopolímeros se describen, por ejemplo, en Hoyer y otros, *Anal. Chem.* 68(21): 3840-3844 (1996).

Como otro ejemplo, en algunas modalidades, los elementos se forman por micropartículas o nanopartículas metálicas. Los métodos adecuados para depositar tales partículas para formar matrices se describen, por ejemplo, en Lee y otros, *Beilstein J. Nanotechnol.* 8: 1049-1055 (2017).

Como un ejemplo adicional, en algunas modalidades, los elementos se forman mediante partículas magnéticas que se ensamblan sobre un sustrato. Ejemplos de tales partículas y métodos para ensamblar matrices se describen en Ye y otros, *Scientific Reports* 6: 23145 (2016).

Como otro ejemplo, en algunas modalidades, los elementos corresponden a regiones de un sustrato en las que se han incorporado una o más marcas ópticas, y/o que se han alterado mediante un proceso tal como fotoblanqueo permanente. Los sustratos adecuados para implementar elementos de esta manera incluyen una amplia variedad de polímeros, por ejemplo. Los métodos para formar tales elementos se describen, por ejemplo, en Moshrefzadeh y otros, *Appl. Phys. Lett.* 62: 16 (1993).

Como otro ejemplo más, en algunas modalidades, los elementos pueden corresponder a partículas coloidales ensambladas (por ejemplo, mediante autoensamblaje) para formar una matriz. Las partículas coloidales adecuadas se describen, por ejemplo, en Sharma, *Resonance* 23(3): 263-275 (2018).

Como un ejemplo adicional, en algunas modalidades, los elementos pueden formarse mediante fotopolimerización de matriz de puntos de una solución de monómero sobre un sustrato. En particular, la polimerización de dos y tres fotones puede usarse para fabricar elementos de dimensiones relativamente pequeñas (por ejemplo, submicrónicas). Los métodos adecuados para preparar elementos sobre un sustrato de esta manera se describen, por ejemplo, en Nguyen y otros, *Materials Today* 20(6): 314-322 (2017).

En algunas modalidades, los elementos se unen o se fijan directa o indirectamente a un sustrato que es permeable a los líquidos. En algunas modalidades, los elementos se unen o se fijan directa o indirectamente a un sustrato que es biocompatible. En algunas modalidades, los elementos se unen o se fijan directa o indirectamente a un sustrato que es un hidrogel.

La figura 17 representa un arreglo ilustrativo de elementos con código de barras dentro de una matriz. De izquierda a derecha, la Figura 17 muestra (L) un portaobjetos que incluye seis matrices con código de barras espacial, (C) un esquema ampliado de una de las seis matrices con código de barras espacial, que muestra una cuadrícula de elementos con código de barras en relación con una muestra biológica, y (R) un esquema ampliado de una sección de una matriz, que muestra la identificación específica de múltiples elementos dentro de la matriz (marcados como ID578, ID579, ID560, etc.).

En algunas modalidades, los elementos pueden formarse sobre perlas de una matriz de perlas. Como se usa en la presente descripción, el término “matriz de perlas” se refiere a una matriz que incluye una pluralidad de perlas como los elementos en la matriz. En algunas modalidades, las perlas se unen a un sustrato. Por ejemplo, las perlas pueden unirse opcionalmente a un sustrato tal como un portaobjetos de microscopio y cerca de una muestra biológica (por ejemplo, una sección de tejido que incluye células). Las perlas pueden suspenderse, además, en una solución y depositarse sobre una superficie (por ejemplo, una membrana, una sección de tejido, o un sustrato (por ejemplo, un portaobjetos de microscopio)). Los ejemplos de matrices de perlas sobre o dentro de un sustrato incluyen las perlas que se ubican en pocillos tales como la matriz BeadChip (disponible en Illumina Inc., San Diego CA), matrices que se usan en plataformas de secuenciación de 454 LifeSciences (una subsidiaria de Roche, Basilea Suiza) y matrices que se usan en plataformas de secuenciación de Ion Torrent (una subsidiaria de Life Technologies, Carlsbad California). Los ejemplos de matrices de perlas se describen en, por ejemplo, las patentes de Estados Unidos núms. 6,266,459; 6,355,431; 6,770,441; 6,859,570; 6,210,891; 6,258,568; y 6,274,320; las publicaciones de solicitud de patente de Estados Unidos núms. 2009/0026082; 2009/0127589; 2010/0137143; y 2010/0282617; y Publicación de solicitud de patente PCT núms. WO 00/063437 y WO 2016/162309.

Una “matriz flexible” incluye una pluralidad de elementos con código de barras espacial que se unen a, o que se incluyen en, un sustrato flexible (por ejemplo, una membrana o cinta) colocados sobre una muestra biológica. En algunas modalidades, una matriz flexible incluye una pluralidad de elementos con código de barras espacial que se incluyen dentro de una matriz de hidrogel. Para formar tal matriz, los elementos de una micromatriz se copian en un hidrogel, y el tamaño del hidrogel se reduce mediante la eliminación del agua. Estas etapas pueden realizarse varias veces. Por ejemplo, en algunas modalidades, un método para preparar una matriz de códigos de barras espaciales de alta densidad puede incluir copiar una pluralidad de elementos de una micromatriz en un primer hidrogel, donde el primer hidrogel está en contacto con la micromatriz; reducir el tamaño del primer hidrogel, que incluye los elementos copiados mediante la eliminación del agua, formar un primer hidrogel encogido que incluye los elementos copiados; copiar los elementos en el primer hidrogel encogido en un segundo hidrogel, donde el segundo hidrogel está en contacto con el primer hidrogel; y reducir el tamaño del segundo hidrogel, que incluye los elementos copiados al eliminar el agua, formar un segundo hidrogel encogido que incluye los elementos copiados, de esta manera se genera una matriz de códigos de barras espaciales de alta densidad. El resultado es una matriz flexible de alta densidad que incluye elementos con código de barras espacial.

En algunas modalidades, las perlas con código de barras espacial pueden cargarse sobre un sustrato (por ejemplo, un hidrogel) para producir una matriz de perlas autoensambladas de alta densidad.

Las matrices flexibles pueden preequilibrarse, combinarse con tampones de reacción y enzimas a concentraciones funcionales (por ejemplo, una mezcla de transcripción inversa). En algunas modalidades, las matrices de perlas flexibles pueden almacenarse durante períodos prolongados (por ejemplo, días) o congelarse hasta que estén listas para usar. En algunas modalidades, la permeabilización de muestras biológicas (por ejemplo, una sección de tejido) puede realizarse con la adición de enzimas/detergentes previo a entrar en contacto con la matriz flexible. La matriz flexible puede colocarse directamente sobre la muestra, o colocarse en contacto indirecto con la muestra biológica (por ejemplo, con una capa o sustancia intermedia entre la muestra biológica y la matriz de perlas flexible). En algunas modalidades, una vez que se aplica una matriz flexible a la muestra, la transcripción inversa y la captura dirigida de analitos pueden realizarse sobre microesferas sólidas, o perlas circulares de un primer tamaño y perlas circulares de un segundo tamaño.

Una “matriz microcapilar” es una serie de elementos dispuestos que se particionan por microcapilares. Un “canal microcapilar” es una partición individual que se crea por los microcapilares. Por ejemplo, los canales microcapilares pueden aislarse de manera fluida de otros canales microcapilares, de manera que el fluido u otro contenido en un canal microcapilar en la matriz se separan del fluido u otro contenido en un canal microcapilar vecino en la matriz. La densidad y el orden de los microcapilares pueden ser cualquier densidad adecuada u orden de sitios discretos.

En algunas modalidades, las matrices microcapilares se tratan para generar condiciones que facilitan la carga. Un ejemplo es el uso de un tubo coronal (BD-20AC, Electro Technic Products) para generar una superficie hidrófila. En algunas modalidades, un elemento (por ejemplo, una perla con la sonda de captura unida) se carga en una matriz microcapilar de manera que se conoce la posición exacta del elemento dentro de la matriz. Por ejemplo, una sonda de captura que contiene un código de barras espacial puede colocarse en un canal microcapilar de manera que el código de barras espacial pueda permitir la identificación de la ubicación desde la que se derivó la secuencia de código de barras de la molécula de ácido nucleico con código de barras.

En algunas modalidades, cuando se usa una distribución aleatoria para distribuir elementos, pueden realizarse pruebas empíricas para generar condiciones de carga/distribución que faciliten un único elemento por microcapilar. En algunas modalidades, puede ser conveniente lograr condiciones de distribución que faciliten solo un único elemento (por ejemplo, perla) por canal microcapilar. En algunas modalidades, puede ser conveniente lograr condiciones de distribución que faciliten más de un elemento (por ejemplo, perla) por canal microcapilar, mediante el flujo de los elementos a través del canal microcapilar.

En algunas modalidades, algunos o todos los elementos en una matriz incluyen una sonda de captura. En algunas modalidades, una matriz puede incluir una sonda de captura que se une directa o indirectamente al sustrato.

La sonda de captura incluye un dominio de captura (por ejemplo, una secuencia de nucleótidos) que puede unirse específicamente (por ejemplo, hibridar) a un analito diana (por ejemplo, ARNm, ADN o proteína) dentro de una muestra. En algunas modalidades, la unión de la sonda de captura a la diana (por ejemplo, hibridación) puede detectarse y cuantificarse mediante la detección de una señal visual, por ejemplo, un fluoróforo, un metal pesado (por ejemplo, ion de plata) o una marca quimioluminiscente, que se ha incorporado en la diana. En algunas modalidades, la intensidad de la señal visual se correlaciona con la abundancia relativa de cada analito en la muestra biológica. Dado que una matriz puede contener miles o millones de sondas de captura (o más), una matriz de elementos con sondas de captura puede interrogar muchos analitos en paralelo.

En algunas modalidades, un sustrato incluye una o más sondas de captura que se diseñan para capturar analitos de uno o más organismos. En un ejemplo no limitante, un sustrato puede contener una o más sondas de captura diseñadas para capturar ARNm de un organismo (por ejemplo, un humano) y una o más sondas de captura diseñadas para capturar ADN de un segundo organismo (por ejemplo, una bacteria).

Las sondas de captura pueden unirse a un sustrato o elemento mediante el uso de una variedad de técnicas. En algunas modalidades, la sonda de captura se une directamente a un elemento que se fija sobre una matriz. En algunas modalidades, las sondas de captura se inmovilizan a un sustrato mediante inmovilización química. Por ejemplo, puede tener lugar una inmovilización química entre los grupos funcionales sobre el sustrato y los elementos funcionales correspondientes sobre las sondas de captura. Los elementos funcionales correspondientes ilustrativos en las sondas de captura pueden ser un grupo químico inherente de la sonda de captura, por ejemplo, un grupo hidroxilo, o puede introducirse un elemento funcional sobre la sonda de captura. Un ejemplo de un grupo funcional sobre el sustrato es un grupo amina. En algunas modalidades, la sonda de captura a inmovilizar incluye un grupo amina funcional o se modifica químicamente para incluir un grupo amina funcional. Los medios y métodos para tal una modificación química se conocen bien en la técnica.

En algunas modalidades, la sonda de captura es un ácido nucleico. En algunas modalidades, la sonda de captura se inmoviliza sobre el elemento o el sustrato mediante su extremo 5'. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 5' e incluye desde el extremo 5' hasta 3': uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI) y uno o más dominios de captura. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento mediante su extremo 5' e incluye desde el extremo 5' hasta 3': un código de barras (por ejemplo, un código de barras espacial o un UMI) y un dominio de captura. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 5' e incluye desde el extremo 5' hasta 3': un dominio de escisión, un dominio funcional, uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI), y un dominio de captura.

En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 5' e incluye desde el extremo 5' hasta 3': un dominio de escisión, un dominio funcional, uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI), un segundo dominio funcional y un dominio de captura. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 5' e incluye desde el extremo 5' hasta 3': un dominio de escisión, un dominio funcional, un código de barras espacial, un UMI y un dominio de captura. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 5' y no incluye un código de barras espacial. En algunas modalidades, la sonda de

captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 5' y no incluye un UMI. En algunas modalidades, la sonda de captura incluye una secuencia para iniciar una reacción de secuenciación.

En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 3'. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 3' e incluye desde el extremo 3' hasta 5': uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI) y uno o más dominios de captura. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 3' e incluye desde el extremo 3' hasta 5': un código de barras (por ejemplo, un código de barras espacial o un UMI) y un dominio de captura. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 3' e incluye desde el extremo 3' hasta 5': un dominio de escisión, un dominio funcional, uno o más códigos de barras (por ejemplo, un código de barras espacial y/o un UMI), y un dominio de captura. En algunas modalidades, la sonda de captura se inmoviliza sobre un elemento o un sustrato mediante su extremo 3' e incluye desde el extremo 3' hasta 5': un dominio de escisión, un dominio funcional, un código de barras espacial, un UMI y un dominio de captura.

La localización del grupo funcional dentro de la sonda de captura a inmovilizar puede usarse para controlar y dar forma al comportamiento de unión y/u orientación de la sonda de captura, por ejemplo, el grupo funcional puede colocarse en el extremo 5' o 3' de la sonda de captura o dentro de la secuencia de la sonda de captura. En algunas modalidades, una sonda de captura puede incluir además un sustrato (por ejemplo, un soporte que se une a la sonda de captura, un soporte que se une al elemento, o un soporte que se une al sustrato). Un sustrato típico para que una sonda de captura se inmovilice incluye porciones que son capaces de unirse a tales sondas de captura, por ejemplo, a ácidos nucleicos funcionalizados con amina. Ejemplos de tales sustratos son soportes de carboxi, aldehído o epoxi.

En algunas modalidades, los sustratos sobre los cuales las sondas de captura pueden inmovilizarse pueden activarse químicamente, por ejemplo mediante la activación de grupos funcionales, disponibles sobre el sustrato. El término "sustrato activado" se refiere a un material en el que se establecieron o habilitaron grupos funcionales químicos reactivos o interactivos mediante procedimientos de modificación química. Por ejemplo, un sustrato que incluye grupos carboxilo puede activarse antes de usar. Además, determinados sustratos contienen grupos funcionales que pueden reaccionar con porciones específicas ya presentes en las sondas de captura.

En algunas modalidades, se usa un enlace covalente para acoplar directamente una sonda de captura a un sustrato. En algunas modalidades una sonda de captura se acopla indirectamente a un sustrato a través de un enlazador que separa el "primer" nucleótido de la sonda de captura del sustrato, es decir, un enlazador químico. En algunas modalidades, una sonda de captura no se une directamente a la matriz, sino que interactúa indirectamente, por ejemplo, mediante la unión a una molécula que se une directa o indirectamente a la matriz. En algunas modalidades, la sonda de captura se une indirectamente a un sustrato (por ejemplo, mediante una solución que incluye un polímero). En algunas modalidades donde la sonda de captura se inmoviliza sobre el elemento de la matriz indirectamente, por ejemplo, mediante hibridación con una sonda de superficie capaz de unirse a la sonda de captura, la sonda de captura puede incluir además una secuencia aguas arriba (5' con respecto a la secuencia que se hibrida con el ácido nucleico, por ejemplo, ARN de la muestra de tejido) que es capaz de hibridarse con el extremo 5' de la sonda de superficie. Solo, el dominio de captura de la sonda de captura puede verse como un oligonucleótido del dominio de captura, que puede usarse en la síntesis de la sonda de captura en modalidades donde la sonda de captura se inmoviliza sobre la matriz indirectamente.

En algunas modalidades, un sustrato comprende un material o matriz inerte (por ejemplo, portaobjetos de vidrio) que se ha funcionalizado, por ejemplo, mediante tratamiento con un material que comprende grupos reactivos que facilitan la inmovilización de las sondas de captura. Véase, por ejemplo, el documento WO 2017/019456. Los ejemplos no limitantes incluyen hidrogeles de poliacrilamida soportados sobre un sustrato inerte (por ejemplo, portaobjetos de vidrio; véase el documento WO 2005/065814 y la solicitud de patente de Estados Unidos núm. 2008/0280773).

En algunas modalidades, las biomoléculas funcionalizadas (por ejemplo, sondas de captura) se inmovilizan sobre un sustrato funcionalizado mediante el uso de métodos covalentes. Los métodos para la unión covalente incluyen, por ejemplo, condensación de aminas y ésteres carboxílicos activados (por ejemplo, ésteres de N-hidroxisuccinimida); condensación de amina y aldehídos en condiciones de aminación reductiva; y reacciones de cicloadición tales como la reacción Diels-Alder [4+2], reacciones de cicloadición 1,3-dipolar, y reacciones de cicloadición [2+2]. Los métodos para la unión covalente incluyen además, por ejemplo, reacciones químicas clic, que incluyen reacciones de cicloadición [3+2] (por ejemplo, reacción de cicloadición de 1,3-dipolar de Huisgen y cicloadición de azida-alquino catalizada por cobre (I) (CuAAC)); reacciones de tiol-eno; la reacción Diels-Alder y la reacción de demanda de electrones inversa Diels-Alder; cicloadición [4+1] de isonitrilos y tetrazinas; y apertura de anillo nucleófilo de carbociclos pequeños (por ejemplo, apertura de epóxido con oligonucleótidos amino). Los métodos para la unión covalente incluyen además, por ejemplo, maleimidas y tioles; y oligonucleótidos funcionalizados con ésteres de *para*-nitrofenilo y sustrato funcionalizado con polilisina. Los métodos para la unión covalente incluyen además, por ejemplo, reacciones disulfuro; reacciones de radicales (véase, por ejemplo, la patente de Estados Unidos núm. 5,919,626); y sustrato funcionalizado con hidrazida (por ejemplo, en donde el grupo funcional con hidrazida se une directa o indirectamente al sustrato) y oligonucleótidos funcionalizados con aldehído (véase, por ejemplo, Yershov y otros (1996) Proc. Natl. Acad. Sci. USA 93, 4913–4918).

En algunas modalidades, las biomoléculas funcionalizadas (por ejemplo, sondas de captura) se inmovilizan sobre un sustrato funcionalizado mediante el uso de métodos covalentes fotoquímicos. Los métodos para la unión covalente fotoquímica incluyen, por ejemplo, la inmovilización de oligonucleótidos conjugados con antraquinona (véase, por ejemplo, Koch y *otros*, (2000) *Bioconjugate Chem.* 11, 474–483).

En algunas modalidades, las biomoléculas funcionalizadas (por ejemplo, sondas de captura se inmovilizan sobre un sustrato funcionalizado mediante el uso de métodos no covalentes. Los métodos para la unión no covalente incluyen, por ejemplo, oligonucleótidos funcionalizados con biotina y sustratos tratados con estreptavidina (véase, por ejemplo, Holmstrøm y *otros*, (1993) *Analytical Biochemistry* 209, 278–283 y Gilles y *otros* (1999) *Nature Biotechnology* 17, 365–370).

En algunas modalidades, un oligonucleótido (por ejemplo, una sonda de captura) puede unirse a un sustrato o elemento de acuerdo con los métodos establecidos en las patentes de Estados Unidos núms. 6,737,236, 7,259,258, 7,375,234, 7,427,678, 5,610,287, 5,807,522, 5,837,860, y 5,472,881; las publicaciones de solicitud de patente de Estados Unidos núms. 2008/0280773 y 2011/0059865; Shalon y *otros*, (1996) *Genome Research*, 639–645; Rogers y *otros*, (1999) *Analytical Biochemistry* 266, 23–30; Stimpson y *otros*, (1995) *Proc. Natl. Acad. Sci. USA* 92, 6379–6383; Beattie y *otros*, (1995) *Clin. Chem.* 45, 700–706; Lamture y *otros*, (1994) *Nucleic Acids Research* 22, 2121–2125; Beier y *otros*, (1999) *Nucleic Acids Research* 27, 1970–1977; Joos y *otros*, (1997) *Analytical Biochemistry* 247, 96–101; Nikiforov y *otros*, (1995) *Analytical Biochemistry* 227, 201–209; Timofeev y *otros*, (1996) *Nucleic Acids Research* 24, 3142–3148; Chrisey y *otros*, (1996) *Nucleic Acids Research* 24, 3031–3039; Guo y *otros*, (1994) *Nucleic Acids Research* 22, 5456–5465; Running y Urdea (1990) *BioTechniques* 8, 276–279; Fahy y *otros*, (1993) *Nucleic Acids Research* 21, 1819–1826; Zhang y *otros* (1991) 19, 3929–3933; y Rogers y *otros*, (1997) *Gene Therapy* 4, 1387–1392. Las matrices pueden prepararse mediante una variedad de métodos. En algunas modalidades, las matrices se preparan a través de la síntesis (por ejemplo, síntesis *in situ*) de oligonucleótidos sobre la matriz, o mediante impresión por chorro o litografía. Por ejemplo, la síntesis dirigida por luz de oligonucleótidos de ADN de alta densidad puede lograrse mediante fotolitografía o síntesis de ADN en fase sólida. Para implementar la síntesis fotolitográfica, los enlazadores sintéticos modificados con grupos protectores fotoquímicos pueden unirse a un sustrato y los grupos protectores fotoquímicos pueden modificarse mediante el uso de una máscara fotolitográfica (aplicada a áreas específicas del sustrato) y luz, de esta manera se produce una matriz que tiene fotodesprotección localizada. Muchos de estos métodos se conocen en la técnica, y se describen, por ejemplo, en Miller y *otros*, “Basic concepts of microarrays and potential applications in clinical microbiology.” *Clinical microbiology reviews* 22.4 (2009): 611–633; US201314111482A; US9593365B2; US2019203275; y WO2018091676.

En algunas modalidades, las matrices se “cubren de puntos” o se “imprimen” con oligonucleótidos y estos oligonucleótidos (por ejemplo, sondas de captura) se unen después al sustrato. Los oligonucleótidos pueden aplicarse ya sea mediante impresión por contacto o sin contacto. Una impresora sin contacto puede usar el mismo método que las impresoras de computadoras (por ejemplo, chorro de burbujas o chorro de tinta) para expulsar pequeñas gotitas de la solución de la sonda sobre el sustrato. La impresora especializada tipo chorro de tinta puede expulsar un volumen de gotas de la solución de oligonucleótidos, en lugar de tinta, de nanolitros a picolitros sobre el sustrato. En la impresión por contacto, cada pasador de impresión aplica directamente la solución de oligonucleótidos sobre una ubicación específica sobre la superficie. Los oligonucleótidos pueden unirse a la superficie del sustrato mediante la interacción electrostática de la carga negativa de la cadena principal de fosfato del ADN con un recubrimiento cargado positivamente de la superficie del sustrato o mediante enlaces covalentes reticulados por UV entre las bases de timidina en el ADN y los grupos de amina sobre la superficie del sustrato tratado. En algunas modalidades, el sustrato es un portaobjetos de vidrio. En algunas modalidades, los oligonucleótidos (por ejemplo, sondas de captura) se unen al sustrato mediante un enlace covalente a una matriz química, por ejemplo, epoxi-silano, amino-silano, lisina, poliacrilamida, etc.

Las matrices pueden prepararse, además, mediante síntesis *in situ*. En algunas modalidades, estas matrices pueden prepararse mediante el uso de fotolitografía. El método típicamente se basa en el enmascaramiento UV y la síntesis química combinatoria dirigida por la luz sobre un sustrato para sintetizar selectivamente sondas directamente sobre la superficie de la matriz, un nucleótido a la vez por punto, para muchos puntos simultáneamente. En algunas modalidades, un sustrato contiene moléculas enlazadoras covalentes que tienen un grupo protector sobre el extremo libre que puede eliminarse mediante la luz. La luz UV se dirige a través de una máscara fotolitográfica para desproteger y activar sitios seleccionados con grupos hidroxilo que inician el acoplamiento con nucleótidos protegidos entrantes que se unen a los sitios activados. La máscara se diseña de tal manera que los sitios expuestos pueden seleccionarse, y por lo tanto se especifican las coordenadas sobre la matriz donde puede unirse cada nucleótido. El proceso puede repetirse, se aplica una nueva máscara que activa diferentes conjuntos de sitios y acopla diferentes bases, lo que permite construir oligonucleótidos arbitrarios en cada sitio. Este proceso puede usarse para sintetizar cientos de miles de diferentes oligonucleótidos. En algunas modalidades, puede usarse la tecnología de síntesis de matriz sin máscara. Esta usa una matriz de microespejos programables para crear máscaras digitales que reflejan el patrón deseado de luz UV para desproteger los elementos.

En algunas modalidades, el proceso de formar puntos por chorro de tinta puede usarse, además, para la síntesis de oligonucleótidos *in situ*. Los diferentes precursores de nucleótidos más el catalizador pueden imprimirse sobre el

sustrato, y después se combinan con las etapas de acoplamiento y desprotección. Este método se basa en imprimir volúmenes de picolitros de nucleótidos sobre la superficie de la matriz en rondas repetidas de impresión base por base que extienden la longitud de las sondas de oligonucleótidos sobre la matriz.

Las matrices pueden prepararse, además, por hibridación activa mediante campos eléctricos para controlar el transporte del ácido nucleico. Los ácidos nucleicos cargados negativamente pueden transportarse a sitios o elementos específicos, cuando se aplica una corriente positiva a uno o más sitios de prueba sobre la matriz. La superficie de la matriz puede contener una molécula de unión, por ejemplo, estreptavidina, que permite la formación de enlaces (por ejemplo, enlaces estreptavidina-biotina) una vez que las sondas biotiniladas dirigidas electrónicamente alcanzan su ubicación diana. La corriente positiva se elimina después de los elementos activos, y los nuevos sitios de prueba pueden activarse mediante la aplicación dirigida de una corriente positiva. El proceso se repite hasta que se cubren todos los sitios de la matriz.

Una matriz para el análisis espacial puede generarse mediante varios métodos como se describe en la presente descripción. En algunas modalidades, la matriz tiene una pluralidad de sondas de captura que comprenden códigos de barras espacial. Pueden determinarse estos códigos de barras espaciales y su relación con las ubicaciones sobre la matriz. En algunos casos, tal información está fácilmente disponible, debido a que los oligonucleótidos se disponen como puntos, se imprimen o se sintetizan sobre la matriz con un patrón predeterminado. En algunos casos, el código de barras espacial puede decodificarse mediante los métodos descritos en la presente descripción, por ejemplo, mediante secuenciación in situ, mediante varias marcas asociadas con los códigos de barras espaciales, etc. En algunas modalidades, puede usarse una matriz como un molde para generar una matriz hija. Por lo tanto, el código de barras espacial puede transferirse a la matriz hija con un patrón conocido.

En algunas modalidades, puede generarse una matriz que comprende sondas con código de barras mediante la ligación de una pluralidad de oligonucleótidos. En algunos casos, un oligonucleótido de la pluralidad contiene una porción de un código de barras, y el código de barras completo se genera tras la ligación de la pluralidad de oligonucleótidos. Por ejemplo, un primer oligonucleótido que contiene una primera porción de un código de barras puede unirse a un sustrato (por ejemplo, mediante el uso de cualquiera de los métodos para unir un oligonucleótido a un sustrato descrito en la presente descripción), y un segundo oligonucleótido que contiene una segunda porción del código de barras puede ligarse después al primer oligonucleótido para generar un código de barras completo. Pueden usarse diferentes combinaciones de la primera, segunda y cualquier porción adicional de un código de barras para aumentar la diversidad de los códigos de barras. En los casos donde el segundo oligonucleótido también se une al sustrato previo a la ligación, el primer y/o el segundo oligonucleótido pueden unirse al sustrato mediante un enlazador de superficie que contiene un sitio de escisión. Después de la ligación, el oligonucleótido ligado se linealiza mediante la escisión en el sitio de escisión.

Para aumentar la diversidad de los códigos de barras, una pluralidad de segundos oligonucleótidos que comprenden dos o más secuencias de código de barras diferentes puede ligarse a una pluralidad de primeros oligonucleótidos que comprenden la misma secuencia de código de barras, de esta manera se generan dos o más especies diferentes de códigos de barras. Para lograr la ligación selectiva, un primer oligonucleótido que se une a un sustrato que contiene una primera porción de un código de barras puede protegerse inicialmente con un grupo protector (por ejemplo, un grupo protector fotoescindible), y el grupo protector puede eliminarse previo a la ligación entre el primer y el segundo oligonucleótido. En los casos donde las sondas con código de barras sobre una matriz se generan mediante la ligación de dos o más oligonucleótidos, puede aplicarse un gradiente de concentración de los oligonucleótidos a un sustrato de manera que diferentes combinaciones de los oligonucleótidos se incorporan en una sonda con código de barras en dependencia de su ubicación sobre el sustrato.

Las sondas con código de barras sobre una matriz pueden generarse, además, mediante la adición de nucleótidos únicos a oligonucleótidos existentes sobre una matriz, por ejemplo, mediante el uso de polimerasas que funcionan de una manera independiente del molde. Pueden añadirse nucleótidos únicos a los oligonucleótidos existentes en un gradiente de concentración, de esta manera se generan las sondas con longitud variable, en dependencia de la ubicación de las sondas sobre la matriz.

Las matrices pueden prepararse, además, mediante modificación de matrices existentes, por ejemplo, mediante modificación de los oligonucleótidos que se unen a las matrices. Por ejemplo, las sondas pueden generarse sobre una matriz que comprende oligonucleótidos que se unen a la matriz en el extremo 3' y tienen un extremo 5' libre. Los oligonucleótidos pueden ser oligonucleótidos sintetizados in situ, y pueden incluir un código de barras. La longitud de los oligonucleótidos puede ser menos de 50 nucleótidos (nt) (por ejemplo, menos de 45, 40, 35, 30, 25, 20, 15 o 10 nt). Para generar sondas mediante el uso de estos oligonucleótidos, puede usarse un cebador complementario a una porción de un oligonucleótido (por ejemplo, una secuencia constante compartida por los oligonucleótidos) para hibridar con el oligonucleótido y extenderse (mediante el uso del oligonucleótido como un molde) para formar un dúplex y para crear un saliente en 3'. El saliente 3' permite por lo tanto añadir nucleótidos u oligonucleótidos adicionales en el dúplex. Una sonda de captura puede generarse, por ejemplo, mediante la adición de uno o más oligonucleótidos al extremo del 3' saliente (por ejemplo, mediante ligación mediada por oligonucleótidos de soporte), donde los oligonucleótidos que se añaden pueden incluir la secuencia o una porción de la secuencia de un dominio de captura.

En los casos en los que los oligonucleótidos sobre una matriz existente incluyan una secuencia de reconocimiento que puede hibridarse con un oligonucleótido de soporte, las sondas pueden generarse, además, mediante la ligación directa de oligonucleótidos adicionales sobre los oligonucleótidos existentes mediante el oligonucleótido de soporte. La secuencia de reconocimiento puede estar en el extremo 5' libre o en el extremo 3' libre de un oligonucleótido sobre la matriz existente. Las secuencias de reconocimiento útiles para los métodos de la presente descripción pueden no contener sitios de reconocimiento de enzimas de restricción o estructuras secundarias (por ejemplo, horquillas), y pueden incluir altos contenidos de nucleótidos de guanina y citosina y, por lo tanto, tener alta estabilidad.

Las matrices de perlas pueden generarse mediante la unión de perlas (por ejemplo, perlas con código de barras) a un sustrato en un patrón regular, o una disposición irregular. Las perlas pueden unirse a regiones selectivas sobre un sustrato, por ejemplo, mediante la activación selectiva de regiones sobre el sustrato para permitir la unión de las perlas. La activación de regiones selectivas sobre el sustrato puede incluir la activación de un recubrimiento (por ejemplo, un recubrimiento fotoescindible) o un polímero que se aplica sobre el sustrato. Las perlas pueden unirse iterativamente, por ejemplo, un subconjunto de las perlas puede unirse a la vez, y el mismo proceso puede repetirse para unir las perlas restantes. Alternativamente, todas las perlas pueden unirse al sustrato en una etapa.

Las perlas con código de barras, o perlas que comprenden una pluralidad de sondas con código de barras, pueden generarse, primero, mediante la preparación de una pluralidad de sondas con código de barras sobre un sustrato, depositando una pluralidad de perlas sobre el sustrato y generando sondas que se unen a las perlas mediante el uso de las sondas sobre el sustrato como un molde.

Los métodos de fabricación comercial a gran escala permiten que millones de oligonucleótidos se unan a una matriz. Las matrices disponibles comercialmente incluyen las de Roche NimbleGen, Inc., (Wisconsin) y Affymetrix (ThermoFisher Scientific).

En algunas modalidades, las matrices pueden prepararse de acuerdo con los métodos establecidos en los documentos WO 2012/140224, WO 2014/060483, WO 2016/162309, WO 2017/019456, WO 2018/091676, y WO 2012/140224, y en la Solicitud de Patente de Estados Unidos núm. 2018/0245142.

C. Captura de analitos para el análisis espacial

En esta sección, se describen aspectos generales de los métodos y sistemas para capturar analitos. Las etapas individuales del método y los elementos del sistema pueden presentarse en combinación en muchas modalidades diferentes; las combinaciones específicas descritas en la presente descripción no limitan de ninguna manera otras combinaciones de etapas y elementos.

Generalmente, pueden capturarse analitos cuando se pone en contacto una muestra biológica con, por ejemplo, un sustrato que comprende sondas o agentes de captura (por ejemplo, un sustrato con sondas de captura incluidas, dispuestas en forma de puntos, impresas sobre el sustrato o un sustrato con elementos (por ejemplo, perlas, pocillos) que comprenden sondas de captura).

Como se usa en la presente descripción, "contacto", "en contacto" y/o "poner en contacto", una muestra biológica con un sustrato que comprende elementos puede referirse a cualquier contacto (por ejemplo, directo o indirecto) de manera que las sondas de captura pueden interactuar (por ejemplo, capturar) con analitos de la muestra biológica. Por ejemplo, el sustrato puede estar cerca o adyacente a la muestra biológica sin contacto físico directo, pero capaz de capturar analitos de la muestra biológica. En algunas modalidades la muestra biológica está en contacto físico directo con el sustrato. En algunos aspectos, los analitos se liberan y están cerca o en contacto con los agentes de captura. En algunas modalidades, la muestra biológica está en contacto físico indirecto con el sustrato. Por ejemplo, una capa líquida puede estar entre la muestra biológica y el sustrato. En algunas modalidades, los analitos se difunden a través de la capa líquida. En algunas modalidades, las sondas de captura se difunden a través de la capa líquida. En algunas modalidades, los reactivos pueden suministrarse mediante la capa líquida entre la muestra biológica y el sustrato. En algunas modalidades, el contacto físico indirecto puede ser la presencia de un segundo sustrato (por ejemplo, un hidrogel, una película, una membrana porosa) entre la muestra biológica y el primer sustrato que comprende elementos con sondas de captura. En algunas modalidades, los reactivos pueden suministrarse por el segundo sustrato a la muestra biológica.

i. Medios/tapas resistentes a la difusión

Para aumentar la eficiencia al fomentar la difusión del analito hacia las sondas de captura marcadas espacialmente, puede usarse un medio resistente a la difusión. En general, la difusión molecular de analitos biológicos ocurre en todas las direcciones, que incluye hacia las sondas de captura (es decir, hacia la matriz con código de barras espacial) y lejos de las sondas de captura (es decir, hacia la solución a granel). El aumento de la difusión hacia la matriz con código de barras espacial reduce la difusión del analito lejos de la matriz con código de barras espacial y aumenta la eficiencia de captura de las sondas de captura.

En algunas modalidades, una muestra biológica se coloca sobre la parte superior de un sustrato con código de barras espacial y un medio resistente a la difusión se coloca sobre la parte superior de la muestra biológica. Por ejemplo, el medio resistente a la difusión puede colocarse sobre una matriz que se ha colocado en contacto con una muestra biológica. En algunas modalidades, el medio resistente a la difusión y la matriz marcada espacialmente son el mismo componente. Por ejemplo, el medio resistente a la difusión puede contener sondas de captura marcadas espacialmente dentro de o sobre el medio resistente a la difusión (por ejemplo, cubreobjetos, portaobjetos, hidrogel o membrana). En algunas modalidades, una muestra se coloca sobre un sustrato y un medio resistente a la difusión se coloca en la parte superior de la muestra biológica. Adicionalmente, una matriz de sondas de captura con código de barras espacial puede colocarse muy cerca sobre el medio resistente a la difusión. Por ejemplo, un medio resistente a la difusión puede intercalarse entre una matriz marcada espacialmente y una muestra sobre un sustrato. En algunas modalidades, el medio resistente a la difusión se dispone o se aplica como puntos sobre la muestra. En otras modalidades, el medio resistente a la difusión se coloca cerca de la muestra.

En general, el medio resistente a la difusión puede ser cualquier material conocido por limitar la difusividad de los analitos biológicos. Por ejemplo, un medio resistente a la difusión puede ser una tapa sólida (por ejemplo, un cubreobjetos o un portaobjetos de vidrio). En algunas modalidades, el medio resistente a la difusión puede fabricarse de vidrio, silicio, papel, monolitos de polímero de hidrogel u otro material. En algunas modalidades, el portaobjetos de vidrio puede ser un portaobjetos de vidrio acrilado. En algunas modalidades, el medio resistente a la difusión es una membrana porosa. En algunas modalidades, el material puede ser naturalmente poroso. En algunas modalidades, el material puede tener poros o pocillos grabados en material sólido. En algunas modalidades, el tamaño de los poros puede manipularse para minimizar la pérdida de analitos diana. En algunas modalidades, la composición química de la membrana puede manipularse para minimizar la pérdida de analitos diana. En algunas modalidades, el medio resistente a la difusión (es decir, hidrogel) se une de manera covalente a un sustrato (es decir, portaobjetos de vidrio). En algunas modalidades, el medio resistente a la difusión puede ser cualquier material conocido por limitar la difusividad de los transcritos de poli(A). En algunas modalidades, el medio resistente a la difusión puede ser cualquier material conocido por limitar la difusividad de las proteínas. En algunas modalidades, el medio resistente a la difusión puede ser cualquier material conocido por limitar la difusividad de los constituyentes macromoleculares.

En algunas modalidades, un medio resistente a la difusión incluye uno o más medios resistentes a la difusión. Por ejemplo, uno o más medios resistentes a la difusión pueden combinarse de diversas maneras previo a colocar el medio en contacto con una muestra biológica que incluye, sin limitación, recubrimiento, estratificación o aplicación de puntos. Como otro ejemplo, un hidrogel puede colocarse sobre una muestra biológica seguido de la colocación de una cubierta (por ejemplo, un portaobjetos de vidrio) sobre el hidrogel.

En algunas modalidades, se aplica una fuerza (por ejemplo, presión hidrodinámica, vibración ultrasónica, contrastes de solutos, radiación con microondas, circulación vascular u otras fuerzas eléctricas, mecánicas, magnéticas, centrífugas y/o térmicas) para controlar la difusión y mejorar la captura del analito. En algunas modalidades, una o más fuerzas y uno o más medios resistentes a la difusión se usan para controlar la difusión y mejorar la captura. Por ejemplo, una fuerza centrífuga y un portaobjetos de vidrio pueden usarse simultáneamente. Cualquiera de una variedad de combinaciones de una fuerza y un medio resistente a la difusión puede usarse para controlar o mitigar la difusión y mejorar la captura de los analitos.

En algunas modalidades, el medio resistente a la difusión, junto con la matriz con código de barras espacial y la muestra, se sumerge en una solución a granel. En algunas modalidades, la solución a granel incluye reactivos de permeabilización. En algunas modalidades, el medio resistente a la difusión incluye al menos un reactivo de permeabilización. En algunas modalidades, el medio resistente a la difusión (es decir, hidrogel) se empapa en reactivos de permeabilización antes de poner en contacto el medio resistente a la difusión con la muestra. En algunas modalidades, el medio resistente a la difusión puede incluir pocillos (por ejemplo, micro, nano o picopocillos) que contienen un tampón o reactivos de permeabilización. En algunas modalidades, el medio resistente a la difusión puede incluir reactivos de permeabilización. En algunas modalidades, el medio resistente a la difusión puede contener reactivos o monómeros secos para suministrar reactivos de permeabilización cuando el medio resistente a la difusión se aplica a una muestra biológica. En algunas modalidades, el medio resistente a la difusión se añade a la matriz con código de barras espacial y al ensamble de la muestra antes de que el conjunto se sumerja en una solución a granel. En algunas modalidades, el medio resistente a la difusión se añade a la matriz con código de barras espacial y al ensamble de la muestra después de que la muestra se ha expuesto a reactivos de permeabilización. En algunas modalidades, los reactivos de permeabilización se hacen fluir a través de una cámara o canal microfluídico sobre el medio resistente a la difusión. En algunas modalidades, el flujo controla el acceso de la muestra a los reactivos de permeabilización. En algunas modalidades, los analitos diana difunden fuera de la muestra y hacia una solución a granel y consiguen incluirse en un medio resistente a la difusión que incluye la sonda de captura marcada espacialmente. En algunas modalidades, se intercala una solución libre entre la muestra biológica y un medio resistente a la difusión.

La Figura 18 es una ilustración de un uso ilustrativo de un medio resistente a la difusión. Un medio resistente a la difusión 1802 puede ponerse en contacto con una muestra 1803. En la figura 18, un portaobjetos de vidrio 1804 se rellena con sondas de captura con código de barras espacial 1806, y la muestra 1803, 1805 se pone en contacto con la matriz 1804, 1806. Un medio resistente a la difusión 1802 puede aplicarse a la muestra 1803, en donde la muestra

1803 se intercala entre un medio resistente a la difusión 1802 y un portaobjetos recubierto de la sonda de captura 1804. Cuando se aplica una solución de permeabilización 1801 a la muestra, el uso del medio/tapa resistente a la difusión 1802 dirige la migración de los analitos 1805 hacia las sondas de captura 1806 al reducir la difusión de los analitos fuera del medio. Alternativamente, la tapa puede contener reactivos de permeabilización.

5 ii. Condiciones de captura

Las sondas de captura sobre el sustrato (o en un elemento sobre el sustrato) interactúan con los analitos que se liberan a través de un dominio de captura, descrito en otra parte, para capturar los analitos. En algunas modalidades, se realizan determinadas etapas para mejorar la transferencia o captura de los analitos mediante las sondas de captura de la matriz. Ejemplos de tales modificaciones incluyen, pero no se limitan a, condiciones de ajuste para poner en contacto el sustrato con una muestra biológica (por ejemplo, tiempo, temperatura, orientación, niveles de pH, tratamiento previo de muestras biológicas, etc.), usar la fuerza para transportar analitos (por ejemplo, electroforética, centrífuga, mecánico, etc.), realizar reacciones de amplificación para aumentar la cantidad de analitos biológicos (por ejemplo, amplificación por PCR, amplificación in situ, amplificación clonal), y/o usar sondas marcadas para detectar amplicones y códigos de barras.

En algunas modalidades, la captura de los analitos se facilita mediante el tratamiento de la muestra biológica con reactivos de permeabilización. Si una muestra biológica no se permeabiliza lo suficiente, la cantidad del analito capturado sobre el sustrato puede ser demasiado baja para permitir un análisis adecuado. Por el contrario, si la muestra biológica es demasiado permeable, el analito puede difundirse lejos de su origen en la muestra biológica, de manera que se pierde la relación espacial relativa de los analitos dentro de la muestra biológica. Por tanto, es deseado un equilibrio entre permeabilizar la muestra biológica lo suficiente como para obtener una buena intensidad de señal, al tiempo que se mantiene aún la resolución espacial de la distribución del analito en la muestra. Los métodos para preparar muestras biológicas fácilmente se conocen en la técnica y pueden modificarse en dependencia de la muestra biológica y de cómo se prepara la muestra biológica (por ejemplo, congelada fresca, FFPE, etc.).

20 iii. Métodos de captura pasiva

En algunas modalidades, los analitos pueden migrar desde una muestra a un sustrato. Los métodos para facilitar la migración pueden ser pasivos (por ejemplo, difusión) y/o activos (por ejemplo, migración electroforética de ácidos nucleicos). Los ejemplos no limitantes de migración pasiva pueden incluir difusión simple y presión osmótica creada por la rehidratación de objetos deshidratados.

La migración pasiva mediante difusión usa gradientes de concentración. La difusión es el movimiento de objetos no atados hacia el equilibrio. Por lo tanto, cuando hay una región de alta concentración del objeto y una región de baja concentración del objeto, el objeto (sonda de captura, el analito, etc.) se mueve a un área de menor concentración. En algunas modalidades, los analitos no atados se mueven hacia abajo en un gradiente de concentración.

En algunas modalidades, pueden añadirse diferentes reactivos a la muestra biológica, de manera que la muestra biológica se rehidrate mientras mejora la captura de los analitos. En algunas modalidades, la muestra biológica puede rehidratarse con reactivos de permeabilización. En algunas modalidades, la muestra biológica puede rehidratarse con una solución de tinción (por ejemplo, tinción con hematoxilina y eosina).

45 iv. Métodos de captura activa

En algunos ejemplos de cualquiera de los métodos descritos en la presente descripción, un analito en una célula o una muestra biológica puede transportarse (por ejemplo, de manera pasiva o activa) a una sonda de captura (por ejemplo, una sonda de captura fija a una superficie sólida).

Por ejemplo, los analitos en una célula o una muestra biológica pueden transportarse a una sonda de captura (por ejemplo, una sonda de captura inmovilizada) mediante el uso de un campo eléctrico (por ejemplo, mediante el uso de electroforesis), un gradiente de presión, flujo de fluido, un gradiente de concentración química, un gradiente de temperatura y/o un campo magnético. Por ejemplo, los analitos pueden transportarse a través de, por ejemplo, un gel (por ejemplo, matriz de hidrogel), un fluido, o una célula permeabilizada, a una sonda de captura (por ejemplo, una sonda de captura inmovilizada).

En algunos ejemplos, puede aplicarse un campo electroforético a los analitos para facilitar la migración de los analitos hacia una sonda de captura. En algunos ejemplos, una muestra entra en contacto con un sustrato y sondas de captura fijadas sobre un sustrato (por ejemplo, un portaobjetos, un cubreobjetos o una perla), y se aplica una corriente eléctrica para promover la migración direccional de analitos cargados hacia las sondas de captura fijadas sobre el sustrato. Puede usarse un conjunto de electroforesis, donde una célula o una muestra biológica está en contacto con un cátodo y las sondas de captura (por ejemplo, sondas de captura fijadas sobre un sustrato), y donde las sondas de captura (por ejemplo, sondas de captura fijadas sobre un sustrato) están en contacto con la célula o muestra biológica y un ánodo, para aplicar la corriente.

La transferencia electroforética de analitos puede realizarse mientras se mantiene la alineación espacial relativa de los analitos en la muestra. Como tal, un analito capturado por las sondas de captura (por ejemplo, sondas de captura fijadas sobre un sustrato) retiene la información espacial de la célula o de la muestra biológica. La aplicación de un campo electroforético a analitos también puede resultar en un aumento de la temperatura (por ejemplo, calor). En algunas modalidades, el aumento de temperatura (por ejemplo, calor) puede facilitar la migración de los analitos hacia una sonda de captura.

En algunos ejemplos, se usa una matriz de microelectrodos direccionables espacialmente para la captura con limitaciones espaciales de al menos un analito de interés cargado por una sonda de captura. La matriz de microelectrodos puede configurarse para incluir una alta densidad de sitios discretos que tienen un área pequeña para aplicar un campo eléctrico para promover la migración de analito(s) cargado(s) de interés. Por ejemplo, la captura electroforética puede realizarse en una región de interés mediante el uso de una matriz de microelectrodos direccionables espacialmente.

Puede usarse una alta densidad de sitios discretos sobre una matriz de microelectrodos para dispositivos pequeños. La superficie puede incluir cualquier densidad adecuada de sitios discretos (por ejemplo, una densidad adecuada para procesar la muestra sobre el sustrato conductor en una cantidad de tiempo dada). En una modalidad, la superficie tiene una densidad de sitios discretos mayor que o igual a aproximadamente 500 sitios por 1 mm². En algunas modalidades, la superficie tiene una densidad de sitios discretos de aproximadamente 100, aproximadamente 200, aproximadamente 300, aproximadamente 400, aproximadamente 500, aproximadamente 600, aproximadamente 700, aproximadamente 800, aproximadamente 900, aproximadamente 1000, aproximadamente 2000, aproximadamente 3000, aproximadamente 4000, aproximadamente 5000, aproximadamente 6000, aproximadamente 7000, aproximadamente 8000, aproximadamente 9000, aproximadamente 10 000, aproximadamente 20 000, aproximadamente 40 000, aproximadamente 60 000, aproximadamente 80 000, aproximadamente 100 000 o aproximadamente 500 000 sitios por 1 mm². En algunas modalidades, la superficie tiene una densidad de sitios discretos de al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 400, al menos aproximadamente 500, al menos aproximadamente 600, al menos aproximadamente 700, al menos aproximadamente 800, al menos aproximadamente 900, al menos aproximadamente 1000, al menos aproximadamente 2000, al menos aproximadamente 3000, al menos aproximadamente 4000, al menos aproximadamente 5000, al menos aproximadamente 6000, al menos aproximadamente 7000, al menos aproximadamente 8000, al menos aproximadamente 9000, al menos aproximadamente 10 000, al menos aproximadamente 20 000, al menos aproximadamente 40 000, al menos aproximadamente 60 000, al menos aproximadamente 80 000, al menos aproximadamente 100 000, o al menos aproximadamente 500 000 sitios por 1 mm².

Los esquemas que ilustran un sistema electroforético de transferencia que se configura para dirigir los analitos de transcritos hacia una matriz de sonda de captura con código de barras espacial se muestran en la figura 19A y la figura 19B. En esta configuración ilustrativa de un sistema electroforético, una muestra 1902 se intercala entre el cátodo 1901 y la matriz de sonda de captura con código de barras espacial 1904, 1905, y la matriz de sonda de captura con código de barras espacial 1904, 1905, se intercala entre la muestra 1902 y el ánodo 1903, de manera que la muestra 1902, 1906 está en contacto con las sondas de captura con código de barras espacial 1907. Cuando se aplica un campo eléctrico al sistema electroforético de transferencia, los analitos de ARNm cargados negativamente 1906 se tirarán hacia el ánodo cargado positivamente 1903 y hacia la matriz con códigos de barras espacial 1904, 1905 que contiene las sondas de captura con códigos de barras espacial 1907. Las sondas de captura con código de barras espacial 1907 después interactúan/hibridan con/inmovilizan el analito de ARNm diana 1906, lo que hace que la captura del analito sea más eficiente. La configuración del sistema electroforético puede cambiar en dependencia del analito diana. Por ejemplo, las proteínas pueden ser positivas, negativas, neutras o polares en dependencia de la proteína, así como también de otros factores (por ejemplo, punto isoeléctrico, solubilidad, etc.). El profesional experto tiene el conocimiento y la experiencia para disponer el sistema electroforético de transferencia para facilitar la captura de un analito diana particular.

La figura 20 es una ilustración que muestra un protocolo de flujo de trabajo ilustrativo que utiliza un sistema electroforético de transferencia. En el ejemplo, el Panel A representa una matriz de elementos flexible con códigos de barras espacial que se pone en contacto con una muestra. La muestra puede ser una matriz flexible, en donde la matriz se inmoviliza sobre un hidrogel, membrana u otro sustrato flexible. El Panel B representa el contacto de la matriz con la muestra y la obtención de imágenes del conjunto de la matriz y la muestra. La imagen del conjunto de la matriz y la muestra puede usarse para verificar la colocación de la muestra, elegir una región de interés o cualquier otra razón para obtener imágenes de una muestra sobre una matriz como se describe en la presente descripción. El Panel C representa la aplicación de un campo eléctrico mediante el uso de un sistema electroforético de transferencia para ayudar en la captura eficiente de un analito diana. Aquí, los analitos diana de ARNm cargado negativamente migran hacia el ánodo cargado positivamente. El Panel D representa la aplicación de reactivos de transcripción inversa y la síntesis de ADNc de la primera hebra de los analitos diana capturados. El Panel E representa la eliminación y preparación de la matriz para la construcción de bibliotecas (Panel F) y secuenciación de nueva generación (Panel G).

D. Región de interés

Una muestra biológica puede tener regiones que muestran elementos(s) morfológicos que pueden indicar la presencia de enfermedad o el desarrollo de un fenotipo de enfermedad. Por ejemplo, los elementos morfológicos en un sitio específico dentro de una muestra de biopsia tumoral pueden indicar la agresividad, resistencia terapéutica, potencial metastásico, migración, etapa, diagnóstico y/o pronóstico de cáncer en un sujeto. Un cambio en los elementos morfológicos en un sitio específico dentro de una muestra de biopsia tumoral a menudo se correlaciona con un cambio en el nivel o la expresión de un analito en una célula dentro del sitio específico, que puede, a su vez, usarse para proporcionar información con respecto a la agresividad, resistencia terapéutica, potencial metastásico, migración, etapa, diagnóstico y/o pronóstico de cáncer en un sujeto. Una región o área dentro de una muestra biológica que se selecciona para un análisis específico (por ejemplo, una región en una muestra biológica que tiene elementos morfológicos de interés) se describe a menudo como "una región de interés". Una región de interés puede identificarse, detectarse, observarse, analizarse previo a, simultáneamente con y/o después del ensayo *in situ* y/o ensayos espaciales descritos en la presente descripción. En algunos casos, pueden realizarse una o más etapas de microscopía para visualizar o tomar imágenes de la muestra biológica previo a, durante y/o después del ensayo *in situ* y/o ensayos espaciales descritos en la presente descripción. Por ejemplo, cualquier imagen adecuada tal como imágenes de campo claro y/o fluorescencia puede usarse para visualizar muestras biológicas teñidas o no teñidas (por ejemplo, hematoxilina, eosina, DAPI, etc.).

Una región de interés en una muestra biológica puede usarse para analizar un área específica de interés dentro de una muestra biológica y, de esta manera, enfocar la experimentación y la recogida de datos en una región específica de una muestra biológica (en lugar de una muestra biológica completa). Esto resulta en una mayor eficiencia en el tiempo del análisis de una muestra biológica.

Una región de interés puede identificarse en una muestra biológica mediante el uso de una variedad de diferentes técnicas, por ejemplo, microscopía de expansión, microscopía de campo claro, microscopía de campo oscuro, microscopía de contraste de fase, microscopía electrónica, microscopía de fluorescencia, microscopía de reflexión, microscopía de interferencia, microscopía confocal e identificación visual (por ejemplo, a simple vista) y sus combinaciones. Por ejemplo, la tinción y la obtención de imágenes de una muestra biológica pueden realizarse para identificar una región de interés. En algunos ejemplos, la región de interés puede corresponder a una estructura específica de la citoarquitectura. En algunas modalidades, una muestra biológica puede teñirse previo a la visualización para proporcionar contraste entre las diferentes regiones de la muestra biológica. El tipo de tinción puede elegirse en dependencia del tipo de muestra biológica y de la región de las células a teñir. En algunas modalidades puede usarse más de una tinción para visualizar diferentes aspectos de la muestra biológica, por ejemplo, diferentes regiones de la muestra, estructuras celulares específicas (por ejemplo, organelos) o diferentes tipos de células. En otras modalidades, la muestra biológica puede visualizarse u obtenerse imágenes sin teñir la muestra biológica.

En algunas modalidades, las imágenes pueden realizarse mediante el uso de uno o más marcadores fiduciales, es decir, objetos colocados en el campo visual de un sistema de obtención de imágenes que aparecen en la imagen producida. Los marcadores fiduciales se usan típicamente como un punto de referencia o escala de medición. Los marcadores fiduciales pueden incluir, pero no se limitan a, marcas detectables tales como marcas fluorescentes, radiactivas, quimioluminiscentes y colorimétricas. El uso de marcadores fiduciales para estabilizar y orientar muestras biológicas se describe, por ejemplo, en Carter y otros, *Applied Optics* 46:421-427, 2007). En algunas modalidades, un marcador fiducial puede ser una partícula física (por ejemplo, una nanopartícula, una microesfera, una nanoesfera, una perla, o cualquiera de las otras partículas físicas ilustrativas descritas en la presente descripción o conocidas en la técnica).

En algunas modalidades, un marcador fiducial puede estar presente sobre un sustrato para proporcionar la orientación de la muestra biológica. En algunas modalidades, una microesfera puede acoplarse a un sustrato para ayudar en la orientación de la muestra biológica. En algunos ejemplos, una microesfera que se acopla a un sustrato puede producir una señal óptica (por ejemplo, fluorescencia). En otro ejemplo, una microesfera puede unirse a una porción (por ejemplo, esquina) de una matriz en un patrón o diseño específico (por ejemplo, diseño hexagonal) para ayudar en la orientación de una muestra biológica sobre una matriz de elementos sobre el sustrato. En algunas modalidades, un punto de cuantificación puede acoplarse al sustrato para ayudar en la orientación de la muestra biológica. En algunos ejemplos, un punto de cuantificación que se acopla a un sustrato puede producir una señal óptica.

En algunas modalidades, un marcador fiducial puede ser una molécula inmovilizada con la que una molécula señal detectable puede interactuar para generar una señal. Por ejemplo, un ácido nucleico marcador puede enlazarse o acoplarse a una porción química capaz de fluorescer cuando se somete a la luz de una longitud de onda específica (o intervalo de longitudes de onda). Tal molécula de ácido nucleico marcadora puede ponerse en contacto con una matriz antes, al mismo tiempo con, o después de que la muestra de tejido se tiñe para visualizar o tomar imágenes de la sección de tejido. Aunque no se requiere, puede ser ventajoso usar un marcador que pueda detectarse mediante el uso de las mismas condiciones (por ejemplo, condiciones de obtención de imágenes) usadas para detectar un ADNc marcado.

En algunas modalidades, los marcadores fiduciales se incluyen para facilitar la orientación de una muestra de tejido o una imagen de esta en relación con una sonda de captura que se inmoviliza sobre un sustrato. Puede usarse cualquier número de métodos para marcar una matriz de manera que un marcador sea detectable solo cuando se obtienen

imágenes de una sección de tejido. Por ejemplo, una molécula, por ejemplo una molécula fluorescente, que genera una señal, puede inmovilizarse directa o indirectamente sobre la superficie de un sustrato. Los marcadores pueden proporcionarse sobre un sustrato en un patrón (por ejemplo, un borde, una o más hileras, una o más líneas, etc.).

En algunas modalidades, un marcador fiducial puede colocarse aleatoriamente en el campo visual. Por ejemplo, un oligonucleótido que contiene un fluoróforo puede imprimirse, estamparse, sintetizarse o unirse aleatoriamente a un sustrato (por ejemplo, un portaobjetos de vidrio) en una posición aleatoria sobre el sustrato. Una sección de tejido puede ponerse en contacto con el sustrato de manera que el oligonucleótido que contiene el fluoróforo entre en contacto, o esté cerca de, una célula de la sección de tejido o un componente de la célula (por ejemplo, una molécula de ARNm o de ADN). Puede obtenerse una imagen del sustrato y de la sección de tejido, y puede determinarse la posición del fluoróforo dentro de la imagen de la sección de tejido (por ejemplo, al revisar una imagen óptica de la sección de tejido superpuesta con la detección del fluoróforo). En algunas modalidades, los marcadores fiduciales pueden colocarse con precisión en el campo visual (por ejemplo, en ubicaciones conocidas sobre un sustrato). En este caso, un marcador fiducial puede estamparse, unirse o sintetizarse sobre el sustrato y ponerse en contacto con una muestra biológica. Típicamente, se toma una imagen de la muestra y del marcador fiducial, y la posición del marcador fiducial sobre el sustrato puede confirmarse al ver la imagen.

En algunas modalidades, un marcador fiducial puede ser una molécula inmovilizada (por ejemplo, una partícula física) que se une al sustrato. Por ejemplo, un marcador fiducial puede ser una nanopartícula, por ejemplo, una nanovarilla, un nanoalambre, un nanocubo, una nanopirámide, o una nanopartícula esférica. En algunos ejemplos, la nanopartícula puede fabricarse de un metal pesado (por ejemplo, oro). En algunas modalidades, la nanopartícula puede fabricarse de diamante. En algunas modalidades, el marcador fiducial puede ser visible a simple vista. Los marcadores fiduciales ilustrativos y los usos de estos (por ejemplo, para el análisis espacial) se describen, por ejemplo, en los documentos WO 2020/047002, WO 2020/047004, WO 2020/047005, WO 2020/047007, WO 2020/047010, y WO2020/123320.

En algunas modalidades, un módulo de ensayo *in situ* puede proporcionar uno o más marcadores fiduciales para un módulo de ensayo espacial de un método integrado descrito en la presente descripción, por ejemplo, como se describe en la Sección VIII. En algunas modalidades, un primer analito seleccionado por una sonda en el análisis *in situ* puede proporcionar un marcador fiducial molecular para uno o más analitos en un módulo de ensayo espacial subsecuente. Por ejemplo, un primer ácido nucleico, un complemento, una sonda que se hibrida directa o indirectamente con el primer ácido nucleico en un módulo de ensayo *in situ*, un producto de ligación de la sonda y/o un producto de amplificación (por ejemplo, producto de RCA) de la sonda pueden analizarse primero en un módulo de ensayo *in situ*, y después capturarse sobre un sustrato para el análisis espacial. Una molécula capturada se asocia con una señal *in situ* y puede servir como referencia espacial para proporcionar información con respecto a uno o más de otros analitos (por ejemplo, moléculas de ácido nucleico) para analizarse en un módulo de ensayo espacial, pero no se ha seleccionado por la una o más sondas y/o analizado en el módulo de ensayo *in situ*.

En algunas modalidades, la tinción y la obtención de imágenes de una muestra biológica previo a poner en contacto la muestra biológica con una matriz espacial se realiza para seleccionar muestras para el análisis espacial. En algunas modalidades, la tinción incluye aplicar un marcador fiducial como se describió anteriormente, que incluye marcadores detectables fluorescentes, radiactivos, quimioluminiscentes o colorimétricos. En algunas modalidades, la tinción y la obtención de imágenes de muestras biológicas permite al usuario identificar la muestra específica (o región de interés) que el usuario desea evaluar.

En algunas modalidades, un sustrato se trata para minimizar o reducir la hibridación inespecífica del analito dentro de o entre los elementos. Por ejemplo, el tratamiento puede incluir recubrir el sustrato con un hidrogel, película y/o membrana que crea una barrera física a la hibridación inespecífica. Puede usarse cualquier hidrogel adecuado. Por ejemplo, pueden usarse matrices de hidrogel preparadas de acuerdo con los métodos establecidos en las patentes de Estados Unidos núms. 6,391,937, 9,512,422 y 9,889,422, y en las publicaciones de solicitud de patente de Estados Unidos núms. U.S. 2017/0253918 y U.S. 2018/0052081.

El tratamiento puede incluir añadir un grupo funcional que es reactivo o capaz de activarse de manera que se vuelva reactivo después de recibir un estímulo (por ejemplo, fotorreactivo). El tratamiento puede incluir tratar con polímeros que tienen una o más propiedades físicas (por ejemplo, mecánicas, eléctricas, magnéticas y/o térmicas) que minimizan la unión inespecífica (por ejemplo, que activan un sustrato en determinadas ubicaciones para permitir la hibridación del analito en esas ubicaciones).

En algunos ejemplos, una matriz (por ejemplo, cualquiera de las matrices ilustrativas descritas en la presente descripción) puede ponerse en contacto con solo una porción de una muestra biológica (por ejemplo, una célula, un elemento o una región de interés). En algunos ejemplos, una muestra biológica se pone en contacto con solo una porción de una matriz (por ejemplo, cualquiera de las matrices ilustrativas descritas en la presente descripción). En algunos ejemplos, una porción de la matriz puede desactivarse de manera que no interactúe con los analitos en la muestra biológica (por ejemplo, desactivación óptica, desactivación química, desactivación térmica o bloqueo de las sondas de captura en la matriz (por ejemplo, mediante el uso de sondas de bloqueo)). En algunos ejemplos, una región de interés puede eliminarse de una muestra biológica y después la región de interés puede ponerse en contacto con la matriz (por ejemplo, cualquiera de las matrices descritas en la presente descripción). Una región de interés

puede eliminarse de una muestra biológica mediante el uso de microcirugía, microdissección de captura láser, corte, un micrótopo, troceado, tripsinización, marcaje y/o clasificación celular asistida por fluorescencia. En algunas modalidades, los analitos o derivados de estos (por ejemplo, una molécula con código de barras asociada con un analito) de una región de interés pueden eliminarse de la muestra y analizarse.

VIII. Análisis de los analitos etiquetados y/o capturados

Después de que los analitos de la muestra se han hibridado o asociado de cualquier otra manera con sondas de captura, agentes de captura del analito, u otra secuencia de oligonucleótidos con código de barras de acuerdo con cualquiera de los métodos descritos anteriormente en relación con la metodología analítica general basada en células en el espacio, los constructos de código de barras que resultan de la hibridación/asociación se analizan mediante secuenciación para identificar los analitos.

En algunas modalidades, donde una muestra se codifica con código de barras espacial directamente mediante hibridación con sondas de captura o agentes que capturan analitos hibridados, unidos o asociados con ya sea la superficie celular, o introducidos en la célula, como se describió anteriormente, la secuenciación puede realizarse sobre la muestra intacta.

Pueden usarse una amplia variedad de diferentes métodos de secuenciación para analizar constructos de analitos con código de barras. En general, los polinucleótidos secuenciados pueden ser, por ejemplo, moléculas de ácido nucleico tales como ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN), que incluyen variantes o derivados de estos (por ejemplo, ADN de una única hebra o híbridos ADN/ARN, y moléculas de ácido nucleico con un análogo de nucleótido).

La secuenciación de polinucleótidos puede realizarse mediante varios sistemas comerciales. Más generalmente, la secuenciación puede realizarse mediante el uso de amplificación de ácido nucleico, reacción en cadena de la polimerasa (PCR) (por ejemplo, PCR digital y PCR digital de gotas (ddPCR), PCR cuantitativa, PCR en tiempo real, PCR multiplex, métodos de PCR basados en plexo único, PCR de emulsión) y/o amplificación isotérmica.

Otros ejemplos de métodos para la secuenciación de material genético incluyen, pero no se limitan a, métodos de hibridación de ADN (por ejemplo, transferencia Southern), métodos de digestión mediante enzimas de restricción, métodos de secuenciación Sanger, métodos de secuenciación de nueva generación (por ejemplo, secuenciación en tiempo real de molécula única, secuenciación de nanoporos y secuenciación de Polony), métodos de ligación y métodos de micromatriz. Ejemplos adicionales de métodos de secuenciación que pueden usarse incluyen secuenciación dirigida, secuenciación en tiempo real de molécula única, secuenciación de exones, secuenciación basada en microscopía electrónica, secuenciación de paneles, secuenciación mediada por transistor, secuenciación directa, secuenciación de disparo aleatorio, secuenciación de terminación de didesoxi de Sanger, secuenciación del genoma completo, secuenciación mediante hibridación, pirosecuenciación, electroforesis capilar, electroforesis en gel, secuenciación dúplex, secuenciación de ciclo, secuenciación de extensión de base única, secuenciación de fase sólida, secuenciación de alto rendimiento, secuenciación de firma masivamente paralela, coamplificación a una temperatura de desnaturalización más baja-PCR (COLD-PCR), secuenciación por terminador de colorante reversible, secuenciación de extremo pareado, secuenciación a corto plazo, secuenciación de exonucleasas, secuenciación mediante ligación, secuenciación de lectura breve, secuenciación de molécula única, secuenciación mediante síntesis, secuenciación en tiempo real, secuenciación de terminación inversa, secuenciación de nanoporos, secuenciación 454, secuenciación del analizador de genoma Solexa, secuenciación SOLiD™, secuenciación MS-PET, y cualquiera de sus combinaciones.

El análisis de secuencia de las moléculas de ácido nucleico (que incluyen moléculas de ácido nucleico con código de barras o derivados de estas) puede ser directo o indirecto. Por lo tanto, el sustrato de análisis de secuencias (que puede verse como la molécula que se somete a la etapa o proceso de análisis de secuencias) puede ser directamente la molécula de ácido nucleico con código de barras o puede ser una molécula que se deriva de esta (por ejemplo, un complemento de esta). Por lo tanto, por ejemplo, en la etapa de análisis de secuencias de una reacción de secuenciación, el molde de secuenciación puede ser la molécula de ácido nucleico con código de barras o puede ser una molécula derivada de esta. Por ejemplo, una molécula de ADN de la primera y/o segunda hebra puede someterse directamente al análisis de secuencias (por ejemplo, secuenciación), es decir, puede participar directamente en la reacción o proceso de análisis de secuencias (por ejemplo, la reacción de secuenciación o proceso de secuenciación, o ser la molécula que se secuencia o se identifica de cualquier otra manera). Alternativamente, la molécula de ácido nucleico con código de barras puede someterse a una etapa de síntesis o amplificación de la segunda hebra antes del análisis de secuencias (por ejemplo, secuenciación o identificación por otra técnica). El sustrato de análisis de secuencias (por ejemplo, molde) puede ser por lo tanto un amplicón o una segunda hebra de una molécula de ácido nucleico con código de barras.

En algunas modalidades, ambas hebras de una molécula bicatenaria pueden someterse al análisis de secuencias (por ejemplo, secuenciadas). En algunas modalidades, pueden analizarse moléculas monocatenarias (por ejemplo, moléculas de ácido nucleico con código de barras) (por ejemplo, secuenciadas). Para realizar la secuenciación de molécula única, la hebra de ácido nucleico puede modificarse en el extremo 3'.

Pueden usarse técnicas de secuenciación masivamente paralelas para la secuenciación de ácidos nucleicos, como se describió anteriormente. En una modalidad, una técnica de secuenciación masivamente paralela puede basarse en terminadores de colorante reversibles. Como un ejemplo, las moléculas de ADN se unen primero a los cebadores sobre, por ejemplo, un sustrato de vidrio o silicio, y se amplifican de manera que se formen colonias clonales locales (amplificación puente). Se añaden cuatro tipos de ddNTP y los nucleótidos no incorporados se eliminan mediante lavado. A diferencia de la pirosecuenciación, el ADN solo se extiende un nucleótido a la vez debido a un grupo de bloqueo (por ejemplo, el grupo de bloqueo 3' presente en la porción de azúcar del ddNTP). Un detector adquiere imágenes de los nucleótidos marcados con fluorescencia, y después el colorante junto con el grupo de bloqueo terminal 3' se elimina químicamente del ADN, como un precursor de un ciclo subsecuente. Este proceso puede repetirse hasta que se obtengan los datos de secuencia requeridos.

Como otro ejemplo, además pueden usarse técnicas de pirosecuenciación masivamente paralelas para la secuenciación de ácidos nucleicos. En la pirosecuenciación, el ácido nucleico se amplifica dentro de gotitas de agua en una solución oleosa (PCR en emulsión), con cada gotita que contiene un único molde de ADN que se une a una única perla recubierta con cebador que después forma una colonia clonal. El sistema de secuenciación contiene muchos pocillos de volumen de picolitros, cada uno de los cuales contiene una única perla y enzimas de secuenciación. La pirosecuenciación usa luciferasa para generar luz para la detección de los nucleótidos individuales que se añaden al ácido nucleico naciente y los datos combinados se usan para generar lecturas de secuencias.

Como otro ejemplo de la aplicación de la pirosecuenciación, el PPi que se libera puede detectarse al convertirse inmediatamente en trifosfato de adenosina (ATP) por la ATP sulfurilasa, y el nivel de ATP que se genera puede detectarse mediante fotones producidos por luciferasa, tales como se describe en Ronaghi, y otros, Anal. Biochem. 242(1), 84-9 (1996); Ronaghi, Genome Res. 11(1), 3-11 (2001); Ronaghi y otros, Science 281 (5375), 363 (1998); y patente de Estados Unidos núms. 6,210,891, 6,258,568 y 6,274,320

En algunas modalidades, la secuenciación se realiza mediante la detección de iones de hidrógeno que se liberan durante la polimerización del ADN. Un micropocillo que contiene una hebra de ADN molde a secuenciar puede inundarse con un único tipo de nucleótido. Si el nucleótido introducido es complementario al nucleótido molde delantero, este se incorpora en la hebra complementaria en crecimiento. Esto provoca la liberación de un ion de hidrógeno que activa un sensor de iones hipersensible, lo que indica que ha ocurrido una reacción. Si hay repeticiones de homopolímeros presentes en la secuencia molde, se incorporarán múltiples nucleótidos en un único ciclo. Esto conduce a un número correspondiente de iones de hidrógeno que se liberan y una señal electrónica proporcionalmente mayor.

En algunas modalidades, donde las sondas de captura no contienen un código de barras espacial, el código de barras espacial puede añadirse después de que la sonda de captura, captura los analitos de una muestra biológica y antes del análisis de los analitos. Cuando se añade un código de barras espacial después de capturar un analito, el código de barras puede añadirse después de la amplificación del analito. En algunas modalidades, se añade un código de barras espacial después de la amplificación de una secuencia de ácido nucleico (por ejemplo, ADN o ARN) en un analito, y el analito puede incluir una molécula de ácido nucleico endógeno (por ejemplo, ADN o ARN) en la muestra (por ejemplo, una molécula de ADN o ARN presente en la muestra previo a un módulo de ensayo *in situ*), una molécula de ADN o ARN que se añade a la muestra durante un módulo de ensayo *in situ* (por ejemplo, una sonda o agente de marcaje que se hibrida/ une directa o indirectamente a un analito), una molécula de ADN o ARN que se genera en un módulo de ensayo *in situ* (por ejemplo, un producto de transcripción inversa, un producto de extensión de la polimerasa, un producto de ligación tal como un producto de ligación con molde, y/o un producto de amplificación tal como un producto de RCA). En algunas modalidades, se añade un código de barras espacial después de la transcripción inversa de un ARN y de la amplificación mediante polimerasa de un ADNc). En algunas modalidades, el análisis del analito usa secuenciación directa de uno o más analitos capturados, tal como secuenciación directa de una molécula de ácido nucleico endógeno (por ejemplo, ADN o ARN) en la muestra (por ejemplo, una molécula de ADN o ARN presente en la muestra previo a un módulo de ensayo *in situ*), una molécula de ADN o ARN que se añade a la muestra durante un módulo de ensayo *in situ* (por ejemplo, una sonda o agente de marcaje que se hibrida/ une directa o indirectamente a un analito), una molécula de ADN o ARN que se genera en un módulo de ensayo *in situ* (por ejemplo, un producto de transcripción inversa, un producto de extensión de la polimerasa, un producto de ligación tal como un producto de ligación con molde, y/o un producto de amplificación tal como un producto de RCA). En algunas modalidades, la secuenciación directa se realiza después de la transcripción inversa de un ARN capturado. En algunas modalidades, la secuenciación directa se realiza después de la amplificación de la transcripción inversa de un ARN capturado.

En algunas modalidades, la secuenciación directa de uno o más analitos capturados se realiza por secuenciación mediante síntesis (SBS). En algunas modalidades, un cebador de secuenciación es complementario a una secuencia en uno o más de los dominios de una sonda de captura (por ejemplo, dominio funcional). En tales modalidades, la secuenciación mediante síntesis puede incluir la transcripción inversa y/o amplificación para generar una secuencia molde (por ejemplo, un dominio funcional) a partir de la cual puede unirse una secuencia del cebador.

La SBS puede involucrar la hibridación de un cebador apropiado, a veces denominado como un cebador de secuenciación, con el molde de ácido nucleico a secuenciar, la extensión del cebador y la detección de los nucleótidos que se usan para extender el cebador. Preferentemente, el ácido nucleico que se usa para extender el cebador se detecta antes de que se añada un nucleótido adicional a la cadena de ácido nucleico en crecimiento, lo que permite la secuenciación del ácido nucleico in situ base por base. La detección de nucleótidos incorporados se facilita mediante la inclusión de uno o más nucleótidos marcados en la reacción de extensión del cebador. Para permitir la hibridación de un cebador de secuenciación apropiado con el molde de ácido nucleico a secuenciar, el molde de ácido nucleico debe estar normalmente en una forma monocatenaria. Si los moldes de ácido nucleico que componen los puntos de ácido nucleico están presentes en una forma bicatenaria, estas pueden procesarse para proporcionar moldes de ácido nucleico monocatenarios mediante el uso de métodos bien conocidos en la técnica, por ejemplo mediante desnaturalización, escisión, etc. Los cebadores de secuenciación que se hibridan con el molde de ácido nucleico y se usan para la extensión del cebador son, preferentemente, oligonucleótidos cortos, por ejemplo, de 15 a 25 nucleótidos de longitud. Los cebadores de secuenciación pueden proporcionarse en solución o en una forma inmovilizada. Una vez que el cebador de secuenciación se ha hibridado con el molde de ácido nucleico a secuenciar al someter el molde de ácido nucleico y el cebador de secuenciación a condiciones apropiadas, la extensión del cebador se realiza, por ejemplo mediante el uso de una polimerasa de ácido nucleico y un suministro de nucleótidos, al menos algunos de los cuales se proporcionan en una forma marcada, y condiciones adecuadas para la extensión del cebador si se proporciona un nucleótido adecuado.

Preferentemente, después de cada etapa de extensión del cebador, se incluye una etapa de lavado para eliminar los nucleótidos no incorporados que pueden interferir con las etapas subsecuente. Una vez que se ha realizado la etapa de extensión del cebador, la colonia de ácido nucleico se monitorea para determinar si un nucleótido marcado se ha incorporado en un cebador extendido. La etapa de extensión del cebador puede repetirse después para determinar los nucleótidos siguientes y subsecuentes incorporados en un cebador extendido. Si se desconoce la secuencia a determinarse, los nucleótidos aplicados a una colonia dada se aplican usualmente en un orden elegido que después se repite a lo largo del análisis, por ejemplo dATP, dTTP, dCTP, dGTP.

Las técnicas de SBS que pueden usarse se describen, por ejemplo, pero no limitado a, las de la solicitud de patente de Estados Unidos. Pub. Núm. 2007/0166705, solicitud de patente de Estados Unidos Pub. Núm. 2006/0188901, patente de Estados Unidos núm. 7,057,026, solicitud de patente de Estados Unidos. Pub. Núm. 2006/0240439, solicitud de patente de Estados Unidos Pub. Núm. 2006/0281109, solicitud de Patente PCT. Pub. Núm. WO 05/065814, solicitud de Patente de Estados Unidos Pub. Núm. 2005/0100900, solicitud de Patente PCT. Pub. Núm. WO 06/064199, solicitud de Patente PCT. Pub. Núm. WO07/010,251, solicitud de Patente de Estados Unidos Pub. Núm. 2012/0270305, solicitud de Patente de Estados Unidos Pub. Núm. 2013/0260372, y solicitud de Patente de Estados Unidos Pub. Núm. 2013/0079232.

En algunas modalidades, la secuenciación directa del ARN capturado se realiza mediante hibridación fluorescente secuencial (por ejemplo, secuenciación mediante hibridación). En algunas modalidades, una reacción de hibridación, donde el ARN se hibrida con una sonda de captura, se realiza in situ. En algunas modalidades, el ARN capturado no se amplifica, previo a la hibridación, con una sonda de secuenciación. En algunas modalidades, el ARN se amplifica, previo a la hibridación, con sondas de secuenciación (por ejemplo, transcripción inversa a ADNc y amplificación del ADNc). En algunas modalidades, la amplificación se realiza mediante el uso de la reacción en cadena de hibridación de molécula única. En algunas modalidades, la amplificación se realiza mediante el uso de la amplificación por círculo rodante.

La hibridación fluorescente secuencial puede implicar la hibridación secuencial de sondas que incluyen secuencias de cebadores degenerados y una marca detectable. Una secuencia de cebador degenerado es una secuencia de oligonucleótidos corta que es capaz de hibridarse con cualquier fragmento de ácido nucleico independientemente de la secuencia de dicho fragmento de ácido nucleico. Por ejemplo, tal método podría incluir las etapas de: (a) proporcionar una mezcla que incluye cuatro sondas, cada una de las cuales incluye A, C, G, o T en el extremo 5', incluir además la secuencia de nucleótidos degenerados de 5 a 11 nucleótidos de longitud, y que incluye además un dominio funcional (por ejemplo, molécula fluorescente) que es distinta para sondas con A, C, G, o T en el extremo 5'; (b) asociar las sondas de la etapa (a) a las secuencias de polinucleótidos diana, cuyas secuencias necesitan determinarse mediante este método; (c) medir las actividades de los cuatro dominios funcionales y registrar la ubicación espacial relativa de las actividades; (d) eliminar los reactivos de las etapas (a)-(b) de las secuencias de polinucleótidos diana; y repetir las etapas (a)-(d) para n ciclos, hasta que se determina la secuencia de nucleótidos del dominio espacial para cada perla, con la modificación de que los oligonucleótidos que se usan en la etapa (a) son complementarios a una parte de las secuencias de polinucleótidos diana y las posiciones 1 hasta n que flanquean la parte de las secuencias. Debido a que las secuencias de código de barras son diferentes, en algunas modalidades, estas secuencias flanqueantes adicionales son secuencias degeneradas. La señal fluorescente de cada punto sobre la matriz para los ciclos 1 hasta n puede usarse para determinar la secuencia de las secuencias de polinucleótidos diana.

En algunas modalidades, la secuenciación directa del ARN capturado mediante el uso de hibridación fluorescente secuencial se realiza in vitro. En algunas modalidades, el ARN capturado se amplifica previo a la hibridación con una sonda de secuenciación (por ejemplo, transcripción inversa a ADNc y amplificación del ADNc). En algunas

modalidades, una sonda de captura que contiene ARN capturado se expone a la sonda de secuenciación que se dirige a las regiones codificantes de ARN. En algunas modalidades, una o más sondas de secuenciación se dirigen a cada región codificante. En algunas modalidades, la sonda de secuenciación se diseña para hibridarse con reactivos de secuenciación (por ejemplo, oligonucleótidos de lectura marcados con colorante). Una sonda de secuenciación puede entonces hibridarse con reactivos de secuenciación. En algunas modalidades, se obtienen imágenes de la salida de la reacción de secuenciación. En algunas modalidades, una secuencia específica de ADNc se resuelve a partir de una imagen de una reacción de secuenciación. En algunas modalidades, la transcripción inversa del ARN capturado se realiza previo a la hibridación con la sonda de secuenciación. En algunas modalidades, la sonda de secuenciación se diseña para dirigirse a secuencias complementarias de las regiones codificantes de ARN (por ejemplo, que se dirige al ADNc).

En algunas modalidades, un ARN capturado se secuencia directamente mediante el uso de un método basado en nanoporos. En algunas modalidades, la secuenciación directa se realiza mediante el uso de secuenciación de ARN directa de nanoporos en la que el ARN capturado se transloca a través de un nanoporo. Una corriente de nanoporo puede registrarse y convertirse en una secuencia base. En algunas modalidades, el ARN capturado permanece unido a un sustrato durante la secuenciación de nanoporos. En algunas modalidades, el ARN capturado se libera del sustrato previo a la secuenciación de nanoporos. En algunas modalidades, donde el analito de interés es una proteína, la secuenciación directa de la proteína puede realizarse mediante el uso de métodos basados en nanoporos. Los ejemplos de métodos de secuenciación basados en nanoporos que pueden usarse se describen en Deamer y otros, Trends Biotechnol. 18, 14 7-151 (2000); Deamer y otros, Acc. Chem. Res. 35:817-825 (2002); Li y otros, Nat. Mater. 2:611-615 (2003); Soni y otros, Clin. Chem. 53, 1996-2001 (2007); Healy y otros, Nanomed. 2, 459-481 (2007); Cockcroft y otros, J. Am. Chem. Soc. 130, 818-820 (2008); y en la Patente de Estados Unidos 7,001,792.

En algunas modalidades, la secuenciación directa del ARN capturado se realiza mediante el uso de secuenciación de molécula única mediante ligación. Tales técnicas utilizan una ADN ligasa para incorporar oligonucleótidos e identificar la incorporación de tales oligonucleótidos. Los oligonucleótidos tienen típicamente diferentes marcadores que se correlacionan con la identidad de un nucleótido en particular en una secuencia con la que se hibridan los oligonucleótidos. Los aspectos y elementos implicados en la secuenciación mediante ligación se describen, por ejemplo, en Shendure y otros, Science (2005), 309: 1728-1732, y en las patentes de Estados Unidos núms. 5,599,675; 5,750,341; 6,969,488; 6,172,218; y 6,306,597.

En algunas modalidades, puede usarse la hibridación de ácido nucleico para la secuenciación. Estos métodos utilizan sondas descodificadoras de ácido nucleico marcadas que son complementarias a al menos una porción de una secuencia de código de barras. La descodificación multiplexada puede realizarse con mezclas de muchas sondas diferentes con marcadores distinguibles. Ejemplos no limitantes de secuenciación por hibridación de ácido nucleico se describen, por ejemplo, en la patente de Estados Unidos núm. 8,460,865, y en Gunderson y otros, Genome Research 14:870-877 (2004).

En algunas modalidades, pueden usarse técnicas comerciales de secuenciación digital de alto rendimiento para analizar secuencias de código de barras, en las que los moldes de ADN se preparan para la secuenciación no una a la vez, sino en un proceso masivo, y donde muchas secuencias se leen preferentemente en paralelo, o alternativamente mediante el uso de un proceso en serie de ultra alto rendimiento que puede paralelizarse a sí misma. Ejemplos de tales técnicas incluyen la secuenciación de Illumina® (por ejemplo, técnicas de secuenciación basadas en celdas de flujo), secuenciación por síntesis mediante el uso de nucleótidos modificados (tales como los comercializados en la tecnología TruSeq™ y HiSeq™ de Illumina, Inc., San Diego, CA), HeliScope™ por Helicos Biosciences Corporation, Cambridge, MA, y PacBio RS de Pacific Biosciences de California, Inc., Menlo Park, CA), secuenciación mediante tecnologías de detección de iones (Ion Torrent, Inc., del Sur de San Francisco, CA), y secuenciación de nanobolas de ADN (Complete Genomics, Inc., Mountain View, CA).

En algunas modalidades, la detección de un protón que se libera después de la incorporación de un nucleótido en un producto de extensión puede usarse en los métodos descritos en la presente descripción. Por ejemplo, los métodos y sistemas de secuenciación descritos en las publicaciones de solicitud de patente de Estados Unidos núms. 2009/0026082, 2009/0127589, 2010/0137143 y 2010/0282617, pueden usarse para secuenciar directamente códigos de barras.

En algunas modalidades, puede usarse el monitoreo en tiempo real de la actividad de la ADN polimerasa durante la secuenciación. Por ejemplo, las incorporaciones de nucleótidos pueden detectarse a través de transferencia de energía por resonancia de fluorescencia (FRET), como se describió, por ejemplo, en Levene y otros, Science (2003), 299, 682-686, Lundquist y otros, Opt. Lett. (2008), 33, 1026-1028, y Korlach y otros, Proc. Natl. Acad. Sci. USA (2008), 105, 1176-1181.

En algunas modalidades, los métodos descritos en la presente descripción pueden usarse para evaluar los niveles y/o la expresión del analito en una célula o una muestra biológica a lo largo del tiempo (por ejemplo, antes o después del tratamiento con un agente o diferentes etapas de diferenciación). En algunos ejemplos, los métodos descritos en la presente descripción pueden realizarse en múltiples muestras biológicas similares o células que se obtienen del sujeto en diferentes puntos de tiempo (por ejemplo, antes o después del tratamiento con un agente, diferentes etapas de

diferenciación, diferentes etapas de progresión de la enfermedad, diferentes edades del sujeto, o antes o después del desarrollo de resistencia a un agente).

5 En algunas modalidades, puede usarse una tabla de búsqueda (LUT) para asociar una propiedad con otra propiedad de un elemento. Estas propiedades incluyen, por ejemplo, ubicaciones, códigos de barras (por ejemplo, moléculas de ácido nucleico de código de barras), códigos de barras espaciales, marcas ópticas, marcas moleculares y otras propiedades.

10 En algunas modalidades, una tabla de búsqueda puede asociar la pluralidad de moléculas de ácido nucleico de código de barras con los elementos. En algunas modalidades, la marca óptica de un elemento puede permitir asociar el elemento con la partícula biológica (por ejemplo, célula o núcleos). La asociación del elemento con la partícula biológica puede permitir además asociar una secuencia de ácido nucleico de una molécula de ácido nucleico de la partícula biológica a una o más propiedades físicas de la partícula biológica (por ejemplo, un tipo de célula o una ubicación de la célula). Por ejemplo, en base a la relación entre el código de barras y la marca óptica, la marca óptica
15 puede usarse para determinar la ubicación de un elemento, de esta manera se asocia la ubicación del elemento con la secuencia de código de barras del elemento. El análisis subsecuente (por ejemplo, secuenciación) puede asociar la secuencia de código de barras y el analito de la muestra. En consecuencia, en base a la relación entre la ubicación y la secuencia de código de barras, puede determinarse la ubicación del analito biológico (por ejemplo, en un tipo específico de célula, en una célula en una ubicación específica de la muestra biológica).

20 En algunas modalidades, el elemento puede tener una pluralidad de moléculas de código de barras de ácido nucleico que se une a este. La pluralidad de moléculas de código de barras de ácido nucleico puede incluir secuencias de código de barras. La pluralidad de moléculas de ácido nucleico que se une a un elemento dado puede tener las mismas secuencias de código de barras, o dos o más secuencias de código de barras diferentes. Pueden usarse diferentes
25 secuencias de código de barras para proporcionar una precisión de ubicación espacial mejorada.

Como se discutió anteriormente, los analitos que se obtienen a partir de una muestra, tales como ARN, ADN, péptidos, lípidos y proteínas, pueden procesarse adicionalmente. En particular, el contenido de células individuales de la muestra puede proporcionarse con secuencias de código de barras espaciales únicas de manera que, después de la
30 caracterización de los analitos, los analitos pueden atribuirse como derivados de la misma célula. Más generalmente, los códigos de barras espaciales pueden usarse para atribuir analitos a ubicaciones espaciales correspondientes en la muestra. Por ejemplo, el posicionamiento espacial jerárquico de múltiples pluralidades de códigos de barras espaciales puede usarse para identificar y caracterizar analitos sobre una región espacial particular de la muestra. En algunas modalidades, la región espacial corresponde a una región espacial particular de interés identificada
35 previamente, por ejemplo, una estructura particular de la citoarquitectura identificada previamente. En algunas modalidades, la región espacial corresponde a una pequeña estructura o grupo de células que no pueden verse a simple vista. En algunas modalidades, puede usarse un identificador molecular único para identificar y caracterizar los analitos a un nivel de célula única.

40 El analito puede incluir una molécula de ácido nucleico, que puede codificarse con una secuencia de código de barras de una molécula de código de barras de ácido nucleico. En algunas modalidades, el analito con código de barras puede secuenciarse para obtener una secuencia de ácido nucleico. En algunas modalidades, la secuencia de ácido nucleico puede incluir información genética asociada con la muestra. La secuencia de ácido nucleico puede incluir la
45 secuencia de código de barras, o un complemento de esta. La secuencia de código de barras, o un complemento de esta, de la secuencia de ácido nucleico puede asociarse electrónicamente con la propiedad (por ejemplo, color y/o intensidad) del analito mediante el uso de la LUT para identificar el elemento asociado en una matriz.

En algunas modalidades, la caracterización espacial bidimensional o tridimensional de uno o más analitos presentes en una muestra biológica puede realizarse mediante el uso de una reacción de captura por proximidad, que es una
50 reacción que detecta dos analitos que están espacialmente cerca entre sí y/o que interactúan entre sí. Por ejemplo, puede usarse una reacción de captura por proximidad para detectar secuencias de ADN que están cerca en el espacio entre sí, por ejemplo, las secuencias de ADN pueden estar dentro del mismo cromosoma, pero separadas por aproximadamente 700 pb o menos. Como otro ejemplo, puede usarse una reacción de captura por proximidad para detectar asociaciones de proteínas, por ejemplo, dos proteínas que interactúan entre sí. Una reacción de captura por
55 proximidad puede realizarse in situ para detectar dos analitos que están espacialmente cerca entre sí y/o que interactúan entre sí dentro de una célula. Los ejemplos no limitantes de reacciones de captura por proximidad incluyen nanoscopia de ADN, microscopia de ADN y métodos de captura de conformación cromosómica. La captura de conformación cromosómica (3C) y los procedimientos experimentales derivados pueden usarse para estimar la proximidad espacial entre diferentes elementos genómicos. Los ejemplos no limitantes de métodos de captura de
60 cromatina incluyen captura de conformación cromosómica (3-C), captura de conformación en chip (4-C), 5-C, ChIA-PET, Hi-C, captura de cromatina seleccionada (T2C). Los ejemplos de tales métodos se describen, por ejemplo, en Miele y otros, *Methods Mol Biol.* (2009), 464, Simonis y otros, *Nat. Genet.* (2006), 38(11): 1348-54, Raab y otros, *Embo. J.* (2012), 31(2): 330-350, y Eagen y otros, *Trends Biochem. Sci.* (2018) 43(6): 469-478.

65 En algunas modalidades, la reacción de captura por proximidad incluye ligación por proximidad. En algunas modalidades, la ligación por proximidad puede incluir el uso de anticuerpos con cadenas de ADN unidas que pueden

participar en reacciones de ligación, replicación y decodificación de secuencias. Por ejemplo, una reacción de ligación por proximidad puede incluir oligonucleótidos que se unen a pares de anticuerpos que pueden unirse por ligación si los anticuerpos se han traído cerca de cada oligonucleótido, por ejemplo, mediante la unión de la misma proteína diana (complejo), y los productos de ligación de ADN que se forman se usan después como molde de la amplificación por PCR, como se describe por ejemplo en Söderberg y otros, *Methods*. (2008), 45(3): 227-32. En algunas modalidades, la ligación por proximidad puede incluir métodos de captura de la conformación cromosómica.

En algunas modalidades, la reacción de captura por proximidad se realiza sobre analitos dentro de aproximadamente una distancia de 400 nm (por ejemplo, aproximadamente 300 nm, aproximadamente 200 nm, aproximadamente 150 nm, aproximadamente 100 nm, aproximadamente 50 nm, aproximadamente 25 nm, aproximadamente 10 nm o aproximadamente 5 nm) entre sí. En general, las reacciones de captura por proximidad pueden ser reversibles o irreversibles.

En algunas modalidades, un método descrito en la presente descripción comprende correlacionar, comparar y/o integrar un resultado del ensayo *in situ* con un resultado del ensayo espacial. En algunas modalidades, un método descrito en la presente descripción comprende correlacionar, comparar y/o integrar la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad de un primer analito (por ejemplo, un primer ácido nucleico o analito de proteína) del ensayo *in situ* con la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad de un segundo analito (por ejemplo, un segundo ácido nucleico o analito de proteína) del ensayo espacial.

En algunas modalidades, en la presente descripción se describe un método para analizar una muestra biológica, que comprende poner en contacto una muestra biológica sobre un primer sustrato con una o más sondas que comprenden moléculas de ácido nucleico que se hibridan directa o indirectamente con una primera diana que es un ácido nucleico (o un complemento o un producto de amplificación de este) en la muestra biológica. La primera diana puede ser una molécula de ARN. En algunas modalidades, el método comprende además detectar la una o más sondas en una ubicación espacial de la muestra biológica, y proporcionar condiciones para permitir que una pluralidad de agentes de captura capturen una segunda diana, que puede ser una proteína diana. En algunas modalidades, la proteína diana se une por un aglutinante que comprende una marca del ácido nucleico que corresponde al aglutinante y/o a la proteína diana, y un agente de captura puede capturar la marca del ácido nucleico, de esta manera se captura la proteína diana. La marca del ácido nucleico capturada (correspondiente al aglutinante y/o a la proteína diana), o un complemento de esta o un producto de amplificación de esta, puede analizarse en un ensayo espacial. Por ejemplo, la pluralidad de agentes de captura puede unirse directa o indirectamente al primer sustrato o a un segundo sustrato, y un agente de captura de la pluralidad de agentes de captura comprende un dominio de captura capaz de capturar un ácido nucleico (por ejemplo, la marca del ácido nucleico) y un código de barras espacial correspondiente a la posición del agente de captura sobre el primer sustrato o sobre el segundo sustrato. En algunas modalidades, el método comprende además generar un polinucleótido marcado espacialmente que comprende (i) una secuencia de la marca del ácido nucleico (correspondiente al aglutinante y/o a la proteína diana) o complemento de esta y (ii) una secuencia del código de barras espacial o complemento de este. En cualquiera de las modalidades en la presente descripción, la muestra biológica puede ponerse en contacto con el aglutinante para el analito de proteína antes, durante o después de detectar la una o más sondas en una ubicación espacial de la muestra. En cualquiera de las modalidades en la presente descripción, la muestra biológica puede ponerse en contacto con el aglutinante para el analito de proteína antes, durante o después de un módulo de secuenciación *in situ* que se realiza sobre la muestra para la primera diana que es un ácido nucleico. En algunas modalidades, el analito de proteína o una subunidad o secuencia polipeptídica de este puede codificarse por una secuencia del analito de ácido nucleico.

En algunas modalidades, en la presente descripción se describe un método para analizar una muestra biológica, que comprende poner en contacto una muestra biológica sobre un primer sustrato con una o más sondas que comprenden moléculas de ácido nucleico que se hibridan directa o indirectamente con una marca de ácido nucleico (o un complemento o un producto de amplificación de este) de un aglutinante que se une a una primera diana, que es una diana que no es ácido nucleico, en la muestra biológica. La primera diana puede ser una proteína. La marca del ácido nucleico puede corresponder al aglutinante y/o a la proteína diana. En algunas modalidades, el método comprende además detectar la una o más sondas en una ubicación espacial de la muestra biológica, y proporcionar condiciones para permitir que una pluralidad de agentes de captura capturen una segunda diana que es una diana de ácido nucleico tal como un ARNm. La diana de ácido nucleico capturado, o un complemento de este o un producto de amplificación de este, puede analizarse en un ensayo espacial. Por ejemplo, la pluralidad de agentes de captura puede unirse directa o indirectamente al primer sustrato o a un segundo sustrato, y un agente de captura de la pluralidad de agentes de captura comprende un dominio de captura capaz de capturar un ácido nucleico (por ejemplo, la diana de ácido nucleico tal como un ARNm), y un código de barras espacial correspondiente a la posición del agente de captura sobre el primer sustrato o sobre el segundo sustrato. En algunas modalidades, el método comprende además generar un polinucleótido marcado espacialmente que comprende (i) una secuencia de la diana de ácido nucleico tal como un ARNm o complemento de este y (ii) una secuencia del código de barras espacial o complemento de este. En cualquiera de las modalidades en la presente descripción, el analito de proteína o una subunidad o secuencia polipeptídica de este puede codificarse por una secuencia del analito de ácido nucleico.

En algunas modalidades, un método descrito en la presente descripción integra elementos de tejidos intactos de una primera pluralidad de analitos en una muestra *in situ* con etapas de ensayo capaces de un análisis de resolución de

nucleótidos, transcriptoma completo (por ejemplo, secuencias de ARN completa) de una segunda pluralidad de analitos en la misma muestra. En algunas modalidades, la primera y la segunda pluralidad de analitos comprenden secuencias de ácido nucleico de interés. En algunas modalidades, la primera y la segunda pluralidad de analitos son transcritos de ARNm. En algunas modalidades, la primera pluralidad de analitos es un subconjunto de la segunda pluralidad de analitos, por ejemplo, la primera pluralidad es un panel de transcritos de ARNm para el análisis dirigido y la segunda pluralidad es el transcriptoma completo o un subconjunto de este para un análisis no dirigido. En algunas modalidades, la primera pluralidad de analitos comprende analitos de proteína y la segunda pluralidad de analitos comprende moléculas de ácido nucleico (por ejemplo, transcritos de ARNm) que corresponden a al menos algunos de los analitos de proteína.

En algunas modalidades, un método descrito en la presente descripción comprende el análisis *in situ* de una primera pluralidad de analitos en un contexto de tejido espacialmente intacto y el análisis espacial de una segunda pluralidad de analitos, donde el análisis espacial puede ser confirmatorio o complementario al análisis *in situ*. En algunas modalidades, el análisis *in situ* comprende un análisis 2D de una muestra biológica, por ejemplo, una sección de tejido que se aísla de un organismo o un cultivo de tejido sobre un sustrato. En algunas modalidades, el análisis *in situ* comprende un análisis 3D de una muestra biológica, por ejemplo, una sección de tejido que se aísla de un organismo o un cultivo de tejido tal como un cultivo de organoides en forma 3D.

En algunas modalidades, un método descrito en la presente descripción comprende el análisis *in situ* de una primera pluralidad de analitos de ácido nucleico o de proteínas, por ejemplo, para el fenotipado celular en una muestra de tejido mediante el uso de analitos de ácido nucleico o de proteínas como biomarcadores, y análisis espacial de una segunda pluralidad de analitos de ácido nucleico, por ejemplo, para una secuenciación más profunda de muchas otras moléculas de ácido nucleico (por ejemplo, ARNm) en un modo de descubrimiento, por ejemplo, para identificar moléculas de ácido nucleico asociadas con uno o más fenotipos celulares particulares.

En algunas modalidades, un método descrito en la presente descripción comprende el uso de un resultado del análisis *in situ* de una muestra para validar un resultado del ensayo espacial de la misma muestra. Por ejemplo, los resultados de la secuenciación *in situ* de una pluralidad de analitos de ácido nucleico o de proteínas pueden usarse para validar el análisis espacial de los mismos analitos de ácido nucleico o relacionados o las moléculas de ácido nucleico (secuencias de ADN de un análisis genómico espacial o secuencias de transcritos de ARN de un análisis transcriptómico espacial) que corresponden a los analitos de proteínas analizados *in situ*. En otro ejemplo, los resultados del análisis espacial de una pluralidad de analitos de ácido nucleico pueden usarse para validar el análisis *in situ* de los mismos analitos de ácido nucleico o relacionados, por ejemplo, al proporcionar información de la morfología del tejido y/o relación espacial de un analito de ácido nucleico con respecto a la morfología del tejido y/u otras moléculas en el tejido.

En algunas modalidades, un método descrito en la presente descripción comprende el análisis *in situ* de uno o más analitos de proteína en una muestra y el análisis espacial de uno o más analitos de ácido nucleico, por ejemplo, ARNm, en la misma muestra. En algunas modalidades, el análisis *in situ* comprende poner en contacto la muestra con una o más sondas, donde una sonda comprende una porción de unión al analito (por ejemplo, un anticuerpo) que se une a un analito de proteína o una porción (por ejemplo, un epítipo) de este y una secuencia de código de barras de ácido nucleico que corresponde a la porción de unión al analito y/o al analito de proteína o porción de este. En algunas modalidades, el análisis *in situ* comprende además analizar la una o más sondas, por ejemplo, mediante imágenes ópticas. Por ejemplo, la una o más sondas pueden ser sondas con código de barras que comprenden una o más secuencias de código de barras de ácido nucleico, que pueden unirse directa o indirectamente mediante sondas de detección marcadas de manera detectable. Una señal detectable o una serie de señales tales como fluorescencia que comprende un patrón espacial y/o un patrón temporal puede analizarse para revelar la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión o actividad del uno o más analitos de proteína en la muestra. En algunas modalidades, el uno o más analitos de proteína se analizan (por ejemplo, mediante imágenes) *in situ* en una muestra de tejido sin migrar fuera de una célula de la muestra de tejido. En algunas modalidades, el uno o más analitos de proteína se analizan (por ejemplo, mediante imágenes) *in situ* en una muestra de tejido sin migrar fuera de la muestra de tejido, por ejemplo, sobre un sustrato. En algunas modalidades, la sonda comprende la porción de unión al analito (por ejemplo, anticuerpo) y la secuencia de código de barras de ácido nucleico no se escinde durante el análisis *in situ*. Por ejemplo, para el análisis *in situ*, la secuencia de código de barras de ácido nucleico no se libera de la porción de unión al analito (por ejemplo, anticuerpo) de la sonda que se une al analito de proteína o capturada por un agente de captura sobre un sustrato; sin embargo, después del análisis *in situ*, la secuencia de código de barras de ácido nucleico puede liberarse y capturarse por un agente de captura para el análisis espacial junto con otras moléculas de ácido nucleico (por ejemplo, transcritos de ARNm) que se liberan de la muestra.

En algunas modalidades, un método descrito en la presente descripción comprende el análisis *in situ* de uno o más analitos no poliadenilados (por ejemplo, transcritos de ARNm no poliadenilados) en una muestra, y el análisis espacial de uno o más analitos poliadenilados (por ejemplo, transcritos de ARNm con colas poli-A) en la misma muestra.

En algunas modalidades, un método descrito en la presente descripción comprende el análisis *in situ* de una primera región de una muestra de tejido y el análisis espacial de una segunda región en la misma muestra de tejido. En algunas modalidades, la primera y la segunda regiones no se solapan. En algunas modalidades, la primera y la segunda

regiones se solapan. Las regiones pueden ser idénticas o una región puede estar completamente dentro de la otra región. En un ejemplo, una porción de una célula en una muestra se analiza *in situ* para una primera pluralidad de analitos (por ejemplo, un panel de transcritos de ARNm de interés), por ejemplo, con un microscopio de superresolución, y una región (por ejemplo, un segmento de tejido de 1 cm × 1 cm) que comprende la célula se somete a un ensayo espacial descrito en la presente descripción para una segunda pluralidad de analitos, por ejemplo, todos los transcritos de ARNm para el análisis transcriptómico no dirigido.

En algunas modalidades, un método descrito en la presente descripción comprende el análisis *in situ* de una primera pluralidad de analitos (por ejemplo, analitos de ácido nucleico de interés) mediante el uso de una pluralidad de sondas. La pluralidad de sondas puede comprender sondas primarias, sondas secundarias y/o incluso sondas de orden aún superior, cualquiera de las cuales puede comprender secuencias de código de barras de ácido nucleico. La unión de una sonda a un analito u otra sonda puede ser directa (por ejemplo, hibridación directa) o indirecta (por ejemplo, mediante un soporte o sonda puente). En algunas modalidades, un método descrito en la presente descripción comprende el análisis *in situ* de un analito de ácido nucleico (por ejemplo, ADN o ARNm), mediante el uso de una o más sondas que se unen directa o indirectamente al analito de ácido nucleico o complemento o producto (por ejemplo, un producto de hibridación, un producto de ligación, un producto de extensión (por ejemplo, mediante una ADN o ARN polimerasa), un producto de replicación, un producto de transcripción/transcripción inversa y/o un producto de amplificación) de este. En algunas modalidades, el método comprende además un análisis espacial descrito en la presente descripción, donde se proporcionan condiciones para permitir que los agentes de captura capturen directa o indirectamente no solo el analito de ácido nucleico (por ejemplo, ADN o ARNm), sino además al menos una de la una o más sondas. En algunas modalidades, el método comprende además generar un primer polinucleótido marcado espacialmente que comprende (i) una secuencia del analito de ácido nucleico o complemento de este y (ii) una secuencia de código de barras espacial o complemento de esta de un agente de captura descrito en la presente descripción, donde el código de barras espacial corresponde a la posición del agente de captura sobre un sustrato (por ejemplo, el primer sustrato o el segundo sustrato descrito en la presente descripción). En algunas modalidades, el método comprende además generar un segundo polinucleótido marcado espacialmente que comprende (i) una secuencia de una de la una o más sondas o complemento de esta y (ii) una secuencia del código de barras espacial o complemento de este de un agente de captura descrito en la presente descripción. En algunas modalidades, el método comprende analizar tanto el primer polinucleótido marcado espacialmente (para analizar el analito de ácido nucleico) como el segundo polinucleótido marcado espacialmente (para analizar una sonda que se une directa o indirectamente al analito de ácido nucleico), y el análisis de uno puede usarse para validar o complementar el otro.

En algunas modalidades, el método comprende capturar el analito de ácido nucleico (por ejemplo, ADN o ARNm) y al menos una de la una o más sondas sobre diferentes moléculas del agente de captura en la misma ubicación sobre un sustrato. En algunas modalidades, el método comprende capturar el analito de ácido nucleico (por ejemplo, ADN o ARNm) y al menos una de la una o más sondas sobre diferentes moléculas de agente de captura que tienen la(s) misma(s) secuencia(s) de código de barras espacial. En algunas modalidades, el método comprende capturar el analito de ácido nucleico (por ejemplo, ADN o ARNm) y al menos una de la una o más sondas sobre el mismo agente de captura. En algunas modalidades, el método comprende generar un polinucleótido marcado espacialmente que comprende (i) una secuencia del analito de ácido nucleico o complemento de este, (ii) una secuencia de la al menos una de la una o más sondas o complemento de esta, y (iii) una secuencia del código de barras espacial o complemento de este de un agente de captura descrito en la presente descripción.

En algunas modalidades, la(s) sonda(s) capturadas sobre el sustrato pueden servir como una referencia espacial para proporcionar información con respecto a uno o más de otros analitos (por ejemplo, moléculas de ácido nucleico endógenas) que no se seleccionan por la una o más sondas en el análisis *in situ*.

En algunas modalidades, se usa un módulo de ensayo *in situ* como un marcador fiducial para el módulo de ensayo espacial. Por ejemplo, un panel de sonda que comprende una sonda P1 que se dirige a un primer analito de mRNA1 del gen Núm. 1 puede usarse para analizar una sección de tejido cerebral *in situ*. Se conoce que el mRNA1 se expresa en el cerebro y este transcrito se detecta en la posición X en la muestra de tejido durante la obtención de imágenes *in situ*. La sonda P1 y los transcritos que incluyen mRNA1 del gen Núm. 1 se capturan por agentes de captura sobre un sustrato, se etiquetan con códigos de barras espaciales (que incluyen código(s) de barras espaciales correspondientes a la posición X) y se someten a secuenciación. Las lecturas de secuenciación de la posición X incluyen no solo aquellas que comprenden secuencias correspondientes a P1 y aquellas que comprenden secuencias correspondientes al mRNA1 (como una validación de la lectura *in situ*), sino además lecturas de secuenciación que comprenden una secuencia correspondiente al mRNA2. El mRNA2 puede ser un transcrito del gen Núm. 2 que es diferente del gen Núm. 1, o una variante (por ejemplo, variante de corte y empalme) del mRNA1 del gen Núm. 1. El mRNA2 puede seleccionarse o no por una sonda (por ejemplo, sonda P1) en el panel de sonda *in situ*. Independientemente, las lecturas de secuenciación que comprenden una secuencia correspondiente al mRNA2 y el/los código(s) de barras espacial o complemento(s) de este correspondientes a la posición X indican que el mRNA2 además está presente y/o se expresa en la posición X, aunque el mRNA2 no se representa por una sonda en el análisis *in situ*. En este ejemplo, una sonda (por ejemplo, P1) capturada sobre el sustrato sirve como una referencia espacial en una posición (por ejemplo, posición X) sobre un sustrato, y el análisis de polinucleótidos marcados espacialmente que comprenden una secuencia del(de los) código(s) de barras espacial(s) o complemento(s) de este correspondientes a la posición pueden

proporcionar información de la presencia/ausencia, distribución, ubicación, cantidad, nivel, expresión, o actividad de un analito (por ejemplo, ARNm2) que no se representa o se selecciona por una sonda en el análisis *in situ*.

IX. Composiciones y kits

Además, en la presente descripción se proporcionan kits, que no son parte de la invención, por ejemplo, que comprenden uno o más polinucleótidos, descritos en la presente descripción y reactivos para realizar los métodos que se proporcionan en la presente descripción, por ejemplo, reactivos requeridos para una o más etapas que comprenden hibridación, ligación, amplificación, detección, secuenciación, preparación de matriz, captura del analito y/o preparación de muestra, como se describe en la presente descripción. En algunas modalidades, el kit comprende uno o más sustratos (por ejemplo, un primer sustrato y/o un segundo sustrato). En algunos ejemplos, un sustrato puede comprender una pluralidad de agentes de captura (por ejemplo, sondas de captura) que se inmovilizan directa o indirectamente sobre el mismo. En algunas modalidades, el kit comprende además un ácido nucleico diana. En algunas modalidades, cualquiera o todos los polinucleótidos son moléculas de ADN. En algunas modalidades, el ácido nucleico diana es una molécula de ARN mensajero.

Los diversos componentes del kit pueden estar presentes en contenedores separados o determinados componentes compatibles pueden combinarse previamente en un único contenedor. En algunas modalidades, los kits contienen además instrucciones para usar los componentes del kit para llevar a la práctica los métodos que se proporcionan.

En algunas modalidades, los kits pueden contener reactivos y/o consumibles necesarios para realizar una o más etapas de los métodos que se proporcionan. En algunas modalidades, los kits contienen reactivos para fijar, incluir, y/o permeabilizar la muestra biológica. En algunas modalidades, los kits contienen reactivos, tales como enzimas y tampones para la ligación y/o la amplificación, tales como ligasas y/o polimerasas. En algunos aspectos, el kit puede comprender además cualquiera de los reactivos descritos en la presente descripción, por ejemplo, tampón de lavado y tampón de ligación. En algunas modalidades, los kits contienen reactivos para la detección y/o la secuenciación, tales como sondas de detección de códigos de barras o marcadores detectables. En algunas modalidades, los kits contienen opcionalmente otros componentes, por ejemplo, cebadores de ácido nucleico, enzimas y reactivos, tampones, nucleótidos, nucleótidos modificados, reactivos para ensayos adicionales.

X. Terminología

Se usa terminología específica a lo largo de esta descripción para explicar diversos aspectos de los aparatos, sistemas, métodos y composiciones que se describen.

Habiendo descrito algunas modalidades ilustrativas de la invención, debería ser evidente para los expertos en la técnica que lo anterior es solamente ilustrativo y no limitante, y que se ha presentado solo a manera de ejemplo. Numerosas modificaciones y otras modalidades ilustrativas están dentro del alcance de un experto en la técnica y se contemplan como que caen dentro del alcance de la invención. En particular, aunque muchos de los ejemplos presentados en la presente descripción implican combinaciones específicas de actos del método o elementos del sistema, debe entenderse que esos actos y esos elementos pueden combinarse de otras maneras para lograr los mismos objetivos.

Como se usa en la presente, las formas singulares "un", "una" y "el/la" incluyen los referentes plurales a menos que el contexto lo indique claramente de cualquier otra manera. Por ejemplo, "un" o "una" significa "al menos uno/una" o "uno/una o más".

El término "aproximadamente", como se usa en la presente, se refiere al intervalo de error habitual para el valor respectivo fácilmente conocido por el experto en este campo técnico. La referencia a "aproximadamente" un valor o parámetro en la presente descripción incluye (y describe) modalidades que se dirigen a ese valor o parámetro per se. A lo largo de esta descripción, diversos aspectos del objeto reivindicado se presentan en un formato de intervalo. Debe entenderse que la descripción en formato de intervalo es solamente para conveniencia y brevedad y no debe interpretarse como una limitación inflexible en el alcance del objeto reivindicado. En consecuencia, la descripción de un intervalo debe considerarse que ha descrito específicamente todos los posibles subintervalos, así como también valores numéricos individuales dentro de ese intervalo. Por ejemplo, cuando se proporciona un intervalo de valores, se entiende que cada valor intermedio, entre el límite superior e inferior de ese intervalo, y cualquier otro valor indicado o intermedio en ese intervalo indicado, está abarcado dentro del objeto reivindicado. Los límites superior e inferior de estos intervalos más pequeños pueden incluirse independientemente en los intervalos más pequeños, y también están abarcados dentro del objeto reivindicado, sujeto a cualquier límite específicamente excluido en el intervalo indicado. Cuando el intervalo indicado incluye uno o ambos de los límites, los intervalos que excluyen uno o ambos de esos límites incluidos también se incluyen en el objeto reivindicado. Esto se aplica independientemente de la amplitud del intervalo.

El uso de términos ordinales tales como "primero", "segundo", "tercero", etc., en las reivindicaciones para modificar un elemento reivindicado no connota por sí mismo ninguna prioridad, precedencia, u orden de un elemento reivindicado sobre otro o el orden temporal en el que se realizan los actos de un método, sino que se usan solamente como

marcadores para distinguir un elemento reivindicado que tiene un determinado nombre de otro elemento que tenga el mismo nombre (excepto por el uso del término ordinal) para distinguir los elementos reivindicados. De manera similar, el uso de a), b), etc., o i), ii), etc. no connota por sí mismo ninguna prioridad, precedencia, u orden de las etapas en las reivindicaciones. De manera similar, el uso de estos términos en la descripción no connota por sí mismo ninguna

(i) Código de barras

Un “código de barras” es una marca, o identificador, que transmite o que 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 en particular puede ser único con relación a otros códigos de barras.

Los códigos de barras pueden tener una variedad de formatos diferentes. Por ejemplo, los códigos de barras pueden incluir códigos de barras de polinucleótidos, secuencias de aminoácidos y/o ácido nucleico aleatorias, y secuencias de aminoácidos y/o ácido nucleico sintéticas. Un código de barras puede unirse a un analito o a otra porción o estructura de manera reversible o irreversible. Un código de barras puede añadirse, por ejemplo, a un fragmento de una muestra de ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN) antes o durante la secuenciación de la muestra. Los códigos de barras pueden permitir la identificación y/o cuantificación de lecturas de secuenciación individuales (por ejemplo, un código de barras puede ser o puede incluir un identificador molecular único o “UMI”).

Los códigos de barras pueden resolver espacialmente los componentes moleculares encontrados en muestras biológicas, por ejemplo, en la resolución de una célula única (por ejemplo, un código de barras puede ser o puede incluir un “código de barras espacial”). En algunas modalidades, un código de barras incluye tanto un UMI como un código de barras espacial. En algunas modalidades, un código de barras incluye dos o más subcódigos de barras que funcionan juntos como un único código de barras. Por ejemplo, un código de barras de polinucleótidos puede incluir dos o más secuencias polinucleotídicas (por ejemplo, subcódigos de barras) que están separadas por una o más secuencias que no son de código de barras.

(ii) Ácido nucleico y nucleótido

Los términos “ácido nucleico” y “nucleótido” pretenden ser consistentes con su uso en la técnica e incluir especies de origen natural o análogos funcionales de estas. Los análogos funcionales particularmente útiles de ácidos nucleicos son capaces de hibridarse con un ácido nucleico de una manera específica de secuencia (por ejemplo, capaces de hibridarse con dos ácidos nucleicos de manera que puede producirse una ligación entre los dos ácidos nucleicos hibridados) o son capaces de usarse como molde para la replicación de una secuencia de nucleótidos en particular. Los ácidos nucleicos de origen natural tienen generalmente una cadena principal que contiene enlaces fosfodiéster. Una estructura análoga puede tener un enlace de cadena principal alternativo que incluye cualquiera de una variedad de los conocidos en la técnica. Los ácidos nucleicos de origen natural tienen generalmente un azúcar desoxirribosa (por ejemplo, que se encuentra en el ácido desoxirribonucleico (ADN)) o un azúcar ribosa (por ejemplo, que se encuentra en el ácido ribonucleico (ARN)).

Un ácido nucleico puede contener nucleótidos que tienen cualquiera de una variedad de análogos de estas porciones de azúcar que se conocen en la técnica. Un ácido nucleico puede incluir nucleótidos naturales o no naturales. Con respecto a esto, un ácido desoxirribonucleico natural puede tener una o más bases seleccionadas del grupo que consiste en adenina (A), timina (T), citosina (C), o guanina (G), y un ácido ribonucleico puede tener una o más bases seleccionadas del grupo que consiste en uracilo (U), adenina (A), citosina (C), o guanina (G). En la técnica se conocen bases no naturales útiles que pueden incluirse en un ácido nucleico o nucleótido.

(iii) Sonda y diana

Una “sonda” o una “diana”, cuando se usa en referencia a un ácido nucleico o secuencia de un ácido nucleico, pretende ser un identificador semántico para el ácido nucleico o secuencia en el contexto de un método o composición, y no limita la estructura o función del ácido nucleico o secuencia más allá de lo que se indica expresamente.

(iv) Oligonucleótido y polinucleótido

Los términos “oligonucleótido” y “polinucleótido” se usan indistintamente para referirse a un multímero monocatenario de nucleótidos de aproximadamente 2 a aproximadamente 500 nucleótidos de longitud. Los oligonucleótidos pueden ser sintéticos, generados enzimáticamente (por ejemplo, mediante polimerización), o mediante el uso de un método de “separar-mezclar”. Los oligonucleótidos pueden incluir monómeros de ribonucleótidos (es decir, pueden ser oligorribonucleótidos) y/o monómeros de desoxirribonucleótidos (es decir, oligodesoxirribonucleótidos). En algunos ejemplos, los oligonucleótidos pueden incluir una combinación tanto de monómeros de desoxirribonucleótidos como de monómeros de ribonucleótidos en el oligonucleótido (por ejemplo, combinación aleatoria u ordenada de monómeros de desoxirribonucleótidos y monómeros de ribonucleótidos). Por ejemplo, un oligonucleótido puede tener de 4 a 10, 10 a 20, 21 a 30, 31 a 40, 41 a 50, 51 a 60, 61 a 70, 71 a 80, 80 a 100, 100 a 150, 150 a 200, 200 a 250, 250 a 300,

300 a 350, 350 a 400, o 400-500 nucleótidos de longitud. Los oligonucleótidos pueden incluir una o más porciones funcionales que se unen (por ejemplo, covalente o no covalentemente) a la estructura del multímero. Por ejemplo, un oligonucleótido puede incluir uno o más marcadores detectables (por ejemplo, un radioisótopo o fluoróforo).

5 (v) Sujeto

Un "sujeto" es un animal, tal como un mamífero (por ejemplo, humano o un simio no humano), o ave (por ejemplo, pájaro), u otro organismo, tal como una planta. Los ejemplos de sujetos incluyen, pero no se limitan a, un mamífero tal como un roedor, ratón, rata, conejo, cobaya, ungulado, caballo, oveja, cerdo, cabra, vaca, gato, perro, primate (es decir, primate humano o no humano); una planta tal como *Arabidopsis thaliana*, maíz, sorgo, copos de avena, trigo, arroz, canola, o soja; una alga tal como *Chlamydomonas reinhardtii*; un nematodo tal como *Caenorhabditis elegans*; un insecto tal como *Drosophila melanogaster*, mosquito, mosca de fruta, o abeja melífera; una arácnida tal como una araña; un pescado tal como un pez cebra; un reptil; un anfibio tal como una rana o *Xenopus laevis*; un *Dictyostelium discoideum*; un hongo tal como *Pneumocystis carinii*, *Takifugu rubripes*, levadura, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*; o un *Plasmodium falciparum*.

(vi) Oligonucleótido de soporte

Un "oligonucleótido de soporte" es un oligonucleótido que, cuando se hibrida con otros polinucleótidos, actúa como un "soporte" para posicionar los polinucleótidos próximos entre sí de manera que puedan ligarse juntos. En algunas modalidades, el oligonucleótido de soporte es ADN o ARN. El oligonucleótido de soporte 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 soporte participa en la ligación de un oligonucleótido "donante" y un oligonucleótido "aceptor". Generalmente, se usa una ARN ligasa, una ADN ligasa, u otra variedad de ligasa para ligar dos secuencias de nucleótidos juntas.

En algunas modalidades, el oligonucleótido de soporte tiene entre 10 y 50 nucleótidos de longitud, por ejemplo, entre 10 y 45, 10 y 40, 10 y 35, 10 y 30, 10 y 25, o 10 y 20 nucleótidos de longitud. En algunas modalidades, el oligonucleótido de soporte tiene entre 15 y 50, 15 y 45, 15 y 40, 15 y 35, 15 y 30, 15 y 25, o 15 y 20 nucleótidos de longitud.

(vii) Adaptador, adaptador y etiqueta

Un "adaptador", un "adaptador" y una "etiqueta" son términos que se usan indistintamente en esta descripción, y se refieren a especies que pueden acoplarse a una secuencia de polinucleótidos (en un proceso denominado "etiquetado") mediante el uso de una cualquiera de muchas técnicas diferentes que incluyen (pero no se limitan a) ligación, hibridación y etiquetado. Los adaptadores también pueden ser secuencias de ácido nucleico que añaden una función, por ejemplo, secuencias espaciadoras, secuencias/sitios cebadores, secuencias de código de barras, secuencias de identificadores moleculares únicos.

40 (viii) Hibridación, hibridar, apareamiento y aparear

Los términos "hibridación" e "hibridar", "apareamiento" y "aparear", se usan indistintamente en esta descripción, y se refieren al apareamiento de secuencias de ácido nucleico sustancialmente complementarias o complementarias dentro de dos moléculas diferentes. El apareamiento puede lograrse mediante cualquier proceso en el que una secuencia de ácido nucleico se une a una secuencia sustancial o completamente complementaria a través del apareamiento de bases para formar un complejo de hibridación. Para los propósitos de hibridación, dos secuencias de ácido nucleico son "sustancialmente complementarias" si al menos el 60 % (por ejemplo, al menos el 70 %, al menos el 80 %, o al menos el 90 %) de sus bases individuales son complementarias entre sí.

50 (ix) Cebador

Un "cebador" es una secuencia de ácido nucleico monocatenario que tiene un extremo 3' que puede usarse como un sustrato para una polimerasa de ácido nucleico en una reacción de extensión de ácido nucleico. Los cebadores de ARN están formados de nucleótidos de ARN, y se usan en la síntesis de ARN, mientras que los cebadores de ADN están formados de nucleótidos de ADN y se usan en la síntesis de ADN. Los cebadores pueden incluir además tanto nucleótidos de ARN como nucleótidos de ADN (por ejemplo, en un patrón aleatorio o diseñado). Los cebadores pueden incluir además otros nucleótidos naturales o sintéticos descritos en la presente descripción que pueden tener una funcionalidad adicional. En algunos ejemplos, los cebadores de ADN pueden usarse para el cebado de la síntesis de ARN y viceversa (por ejemplo, los cebadores de ARN pueden usarse para el cebado de la síntesis de ADN). Los cebadores pueden variar en longitud. Por ejemplo, los cebadores pueden tener de aproximadamente 6 bases a aproximadamente 120 bases. Por ejemplo, los cebadores pueden incluir hasta aproximadamente 25 bases. Un cebador puede, en algunos casos, referirse a una secuencia de unión al cebador.

(x) Extensión del cebador

Una "extensión del cebador" se refiere a cualquier método donde dos secuencias de ácido nucleico (por ejemplo, una región constante de cada una de dos sondas de captura diferentes) se unen (por ejemplo, se hibridan) mediante un solapamiento de sus respectivas secuencias de ácido nucleico complementarias terminales (es decir, por ejemplo, los extremos 3' terminales). Tal unión puede ir seguida de la extensión del ácido nucleico (por ejemplo, una extensión enzimática) de uno, o ambos extremos terminales mediante el uso de la otra secuencia de ácido nucleico como un molde para la extensión. La extensión enzimática puede realizarse mediante una enzima, lo que incluye, pero no se limita a, una polimerasa y/o una transcriptasa inversa.

(xi) Ligación por proximidad

Una "ligación por proximidad" es un método de ligación de dos (o más) secuencias de ácido nucleico que están próximas entre sí a través de medios enzimáticos (por ejemplo, una ligasa). En algunas modalidades, la ligación por proximidad puede incluir una etapa de "relleno de espacios" que implica la incorporación de uno o más ácidos nucleicos mediante una polimerasa, basado en la secuencia de ácido nucleico de una molécula de ácido nucleico molde, que abarca una distancia entre las dos moléculas de ácido nucleico de interés (véase, por ejemplo, la patente de Estados Unidos núm. 7,264,929).

Puede usarse una amplia variedad de métodos diferentes para la ligación por proximidad de moléculas de ácido nucleico, lo que incluye (pero no se limita a) ligaciones de "extremo adhesivo" y "extremo romo". Adicionalmente, la ligación monocatenaria puede usarse para realizar la ligación por proximidad en una molécula de ácido nucleico monocatenario. Las ligaciones por proximidad, de extremo adhesivo, implican la hibridación de secuencias complementarias monocatenarias entre las dos moléculas de ácido nucleico para unirse, previo al evento de ligación en sí mismo. Las ligaciones por proximidad, de extremo romo, no incluyen generalmente la hibridación de regiones complementarias de cada molécula de ácido nucleico, porque ambas moléculas de ácido nucleico carecen de un saliente monocatenario en el sitio de ligación.

(xii) Extensión del ácido nucleico

Una "extensión del ácido nucleico" implica generalmente la incorporación de uno o más ácidos nucleicos (por ejemplo, A, G, C, T, U, análogos de nucleótidos, o derivados de estos) en una molécula (tal como, pero no limitada a, una secuencia de ácido nucleico) de una manera dependiente de molde, de manera que los ácidos nucleicos consecutivos se incorporan mediante una enzima (tal como una polimerasa o transcriptasa inversa), de esta manera se genera una molécula de ácido nucleico recién sintetizada. Por ejemplo, un cebador que se hibrida con una secuencia de ácido nucleico complementaria puede usarse para sintetizar una nueva molécula de ácido nucleico mediante el uso de la secuencia de ácido nucleico complementaria como un molde para la síntesis del ácido nucleico. De manera similar, una cola poliadenilada 3' de un transcrito de ARNm que se hibrida con una secuencia poli (dT) (por ejemplo, dominio de captura) puede usarse como un molde para la síntesis de una única hebra de una molécula de ADNc correspondiente.

(xiii) Amplificación por PCR

Una "amplificación por PCR" se refiere al uso de una reacción en cadena de la polimerasa (PCR) para generar copias de material genético, lo que incluye secuencias de ADN y ARN. Los reactivos y condiciones adecuadas para implementar la PCR se describen, por ejemplo, en las patentes de Estados Unidos 4,683,202, 4,683,195, 4,800,159, 4,965,188, y 5,512,462. En una amplificación por PCR típica, la mezcla de reacción incluye el material genético a amplificar, una enzima, uno o más cebadores que se emplean en una reacción de extensión del cebador, y reactivos para la reacción. Los cebadores de oligonucleótidos son de una longitud suficiente para proporcionar la hibridación al material genético complementario en condiciones de apareamiento. La longitud de los cebadores depende generalmente de la longitud de los dominios de amplificación, pero será típicamente de al menos 4 bases, al menos 5 bases, al menos 6 bases, al menos 8 bases, al menos 9 bases, al menos 10 pares de bases (pb), al menos 11 pb, al menos 12 pb, al menos 13 pb, al menos 14 pb, al menos 15 pb, al menos 16 pb, al menos 17 pb, al menos 18 pb, al menos 19 pb, al menos 20 pb, al menos 25 pb, al menos 30 pb, al menos 35 pb, y puede ser de hasta 40 pb o más, donde la longitud de los cebadores variará generalmente de 18 a 50 pb. El material genético puede ponerse en contacto con un único cebador o un conjunto de dos cebadores (cebadores directos e inversos), en dependencia de si se desea la extensión del cebador, la amplificación lineal o exponencial del material genético.

En algunas modalidades, el proceso de amplificación por PCR usa una enzima ADN polimerasa. La actividad ADN polimerasa puede proporcionarse por una o más enzimas ADN polimerasas distintas. En determinadas modalidades, la enzima ADN polimerasa es de una bacteria, por ejemplo, la enzima ADN polimerasa es una enzima ADN polimerasa bacteriana. Por ejemplo, la ADN polimerasa puede ser de una bacteria del género *Escherichia*, *Bacillus*, *Thermophilus*, o *Pyrococcus*.

Ejemplos adecuados de ADN polimerasas que pueden usarse incluyen, pero no se limitan a: las enzimas ADN polimerasa I de *E.coli*, ADN polimerasa Bsu, ADN polimerasa Bst, ADN polimerasa Taq, ADN polimerasa VENT™, ADN polimerasa DEEPVENT™, ADN polimerasa LongAmp® Taq, ADN polimerasa LongAmp® Hot Start Taq, ADN polimerasa Crimson LongAmp® Taq, ADN polimerasa Crimson Taq, ADN polimerasa OneTaq®, ADN polimerasa

OneTaq® Quick-Load®, ADN polimerasa Hemo KlenTaq®, ADN polimerasa REDTaq®, ADN polimerasa Phusion®, ADN polimerasa Phusion® de alta fidelidad, ADN polimerasa Pfx Platinum, ADN polimerasa Pfx AccuPrime, ADN polimerasa Phi29, fragmento de Klenow, ADN polimerasa Pwo, ADN polimerasa Pfu, ADN polimerasa T4 y ADN polimerasa T7.

El término “ADN polimerasa” incluye no solo enzimas de origen natural, sino también todos los derivados modificados de estas, lo que incluye además derivados de enzimas ADN polimerasas de origen natural. Por ejemplo, en algunas modalidades, la ADN polimerasa puede haberse modificado para eliminar la actividad exonucleasa 5'-3'. Los derivados de secuencia modificada o mutantes de enzimas ADN polimerasa que pueden usarse incluyen, pero no se limitan a, mutantes que retienen al menos parte de la actividad funcional, por ejemplo, la actividad ADN polimerasa de la secuencia de tipo silvestre. Las mutaciones pueden afectar el perfil de actividad de las enzimas, por ejemplo, potenciar o reducir la velocidad de polimerización, en diferentes condiciones de reacción, por ejemplo, temperatura, concentración del molde, concentración del cebador, etc. Las mutaciones o modificaciones de secuencias también pueden afectar la actividad exonucleasa y/o la termoestabilidad de la enzima.

En algunas modalidades, la amplificación por PCR puede incluir reacciones tales como, pero no limitadas a, una reacción de amplificación por desplazamiento de hebra, una reacción de amplificación por círculo rodante, una reacción en cadena con ligasa, una reacción de amplificación mediada por transcripción, una reacción de amplificación isotérmica, y/o una reacción de amplificación mediada por lazo.

En algunas modalidades, la amplificación por PCR usa un único cebador que es complementario a la etiqueta 3' de los fragmentos del ADN diana. En algunas modalidades, la amplificación por PCR usa un primer y un segundo cebador, donde al menos una porción del extremo 3' del primer cebador es complementaria a al menos una porción de la etiqueta 3' de los fragmentos del ácido nucleico diana, y donde al menos una porción del extremo 3' del segundo cebador exhibe la secuencia de al menos una porción de la etiqueta 5' de los fragmentos del ácido nucleico diana. En algunas modalidades, una porción del extremo 5' del primer cebador no es complementaria a la etiqueta 3' de los fragmentos del ácido nucleico diana, y una porción del extremo 5' del segundo cebador no exhibe la secuencia de al menos una porción de la etiqueta 5' de los fragmentos del ácido nucleico diana. En algunas modalidades, el primer cebador incluye una primera secuencia universal y/o el segundo cebador incluye una segunda secuencia universal.

En algunas modalidades (por ejemplo, cuando la amplificación por PCR amplifica ADN capturado), los productos de la amplificación por PCR pueden ligarse a secuencias adicionales mediante el uso de una enzima ADN ligasa. La actividad ADN ligasa puede proporcionarse por una o más enzimas ADN ligasa distintas. En algunas modalidades, la enzima ADN ligasa es de una bacteria, por ejemplo, la enzima ADN ligasa es una enzima ADN ligasa bacteriana. En algunas modalidades, la enzima ADN ligasa es de un virus (por ejemplo, un bacteriófago). Por ejemplo, la ADN ligasa puede ser ADN ligasa T4. Otras enzimas apropiadas para la etapa de ligación incluyen, pero no se limitan a, ADN ligasa Tth, ADN ligasa Taq, ADN ligasa de *Thermococcus* sp. (cepa 9°N) (ADN ligasa 9°N™, disponible en New England Biolabs, Ipswich, MA), y Ampligase™ (disponible en Epicentre Biotechnologies, Madison, WI). Además, pueden usarse derivados, por ejemplo, derivados de secuencia modificada, y/o mutantes de estos.

En algunas modalidades, el material genético se amplifica mediante reacción en cadena de la polimerasa con transcripción inversa (RT-PCR). La actividad transcriptasa inversa deseada puede proporcionarse por una o más enzimas transcriptasas inversas distintas, cuyos ejemplos adecuados incluyen, pero no se limitan a: las enzimas M-MLV, MuLV, AMV, HIV, ArrayScript™, MultiScribe™, ThermoScript™, y SuperScript® I, II, III, y IV. “Transcriptasa inversa” incluye no solo enzimas de origen natural sino también todos los derivados modificados de estas, lo que incluye además derivados de enzimas transcriptasas inversas de origen natural.

Además, la transcripción inversa puede realizarse mediante el uso de derivados de secuencia modificada o mutantes de las enzimas transcriptasas inversas M-MLV, MuLV, AMV, y HIV, lo que incluye mutantes que retienen al menos parte de la actividad funcional, por ejemplo, transcriptasa inversa, de la secuencia de tipo silvestre. La enzima transcriptasa inversa puede proporcionarse como parte de una composición que incluye otros componentes, por ejemplo, componentes estabilizantes, que potencian o mejoran la actividad de la enzima transcriptasa inversa, tal como inhibidor(es) de ARNasa, inhibidores de la síntesis de ADN dependiente de ADN, por ejemplo, actinomicina D. Muchos derivados de secuencia modificada o mutantes de enzimas transcriptasas inversas, por ejemplo, M-MLV, y composiciones, lo que incluye enzimas no modificadas y modificadas, están disponibles comercialmente, por ejemplo, enzimas ArrayScript™, MultiScribe™, ThermoScript™, y SuperScript® I, II, III, y IV.

Determinadas enzimas transcriptasas inversas (por ejemplo, la transcriptasa inversa del virus de la mieloblastosis aviar (AMV) y la transcriptasa inversa del virus de la leucemia murina de Moloney (M-MuLV, MMLV)) pueden sintetizar una hebra de ADN complementaria mediante el uso tanto de ARN (síntesis de ADNc) como de ADN monocatenario (ADNmc) como un molde. De esta forma, en algunas modalidades, la reacción de transcripción inversa puede usar una enzima (transcriptasa inversa) que es capaz de usar tanto ARN como ADNmc como el molde para una reacción de extensión, por ejemplo, una transcriptasa inversa AMV o MMLV.

En algunas modalidades, la cuantificación de ARN y/o ADN se realiza mediante PCR en tiempo real (también conocida como PCR cuantitativa o qPCR), mediante el uso de técnicas bien conocidas en la técnica, tales como, pero no

limitadas a "TAQMANTM" o "SYBR®", o en capilares ("capilares LightCycler®"). En algunas modalidades, la cuantificación del material genético se determina por absorbancia óptica y con PCR en tiempo real. En algunas modalidades, la cuantificación del material genético se determina mediante PCR digital. En algunas modalidades, los genes analizados pueden compararse con un extracto de ácido nucleico (ADN y ARN) de referencia que corresponde a la expresión (ARNm) y cantidad (ADN) para comparar los niveles de expresión de los ácidos nucleicos diana.

(xiv) Anticuerpo

Un "anticuerpo" es una molécula polipeptídica que reconoce y se une a un antígeno diana complementario. Los anticuerpos tienen típicamente una forma de estructura molecular que se asemeja a una forma de Y. Los anticuerpos de origen natural, denominados inmunoglobulinas, pertenecen a una de las clases de inmunoglobulinas IgG, IgM, IgA, IgD, e IgE. Los anticuerpos también pueden producirse sintéticamente. Por ejemplo, los anticuerpos recombinantes, que son anticuerpos monoclonales, pueden sintetizarse mediante el uso de genes sintéticos al recuperar los genes del anticuerpo de las células fuente, amplificar en un vector apropiado, e introducir el vector en un huésped para provocar que el huésped exprese el anticuerpo recombinante. En general, los anticuerpos recombinantes pueden clonarse a partir de cualquier especie de animal productor de anticuerpos, mediante el uso de cebadores de oligonucleótidos y/o sondas de hibridación adecuados. Pueden usarse técnicas recombinantes para generar anticuerpos y fragmentos de anticuerpo, lo que incluye especies no endógenas.

Los anticuerpos sintéticos pueden derivarse de fuentes no inmunoglobulínicas. Por ejemplo, los anticuerpos pueden generarse a partir de ácidos nucleicos (por ejemplo, aptámeros) y a partir de andamios de proteínas no inmunoglobulínicas (tales como aptámeros peptídicos) en los que se insertan lazos hipervariables para formar sitios de unión a antígeno. Los anticuerpos sintéticos basados en ácidos nucleicos o estructuras peptídicas pueden ser más pequeños que los anticuerpos derivados de inmunoglobulina, lo que conduce a una mayor penetración en el tejido.

Los anticuerpos también pueden incluir proteínas afímeras, que son reactivos de afinidad que tienen típicamente un peso molecular de aproximadamente 12-14 kDa. Las proteínas afímeras se unen generalmente a una diana (por ejemplo, una proteína diana) tanto con alta afinidad como con alta especificidad. Los ejemplos de dichas dianas incluyen, pero no se limitan a, cadenas de ubiquitina, inmunoglobulinas, y proteína C reactiva. En algunas modalidades, las proteínas afímeras se derivan de inhibidores de cisteína proteasa, e incluyen lazos peptídicos y una secuencia N-terminal variable que proporciona el sitio de unión.

Anticuerpos también puede referirse a un "fragmento de unión a epítipo" o "fragmento de anticuerpo", que como se usa en la presente descripción, se refiere generalmente a una porción de un anticuerpo completo capaz de unirse al mismo epítipo que el anticuerpo completo, aunque no necesariamente en la misma medida. Aunque son posibles múltiples tipos de fragmentos de unión a epítipo, un fragmento de unión a epítipo comprende típicamente al menos un par de regiones variables de las cadenas pesada y ligera (VH y VL, respectivamente) que se unen (*por ejemplo*, mediante enlaces disulfuro) para preservar el sitio de unión a antígeno, y no contiene toda o una porción de la región Fc. Los fragmentos de unión a epítipo de un anticuerpo pueden obtenerse a partir de un anticuerpo dado mediante cualquier técnica adecuada (*por ejemplo*, tecnología de ADN recombinante o escisión enzimática o química de un anticuerpo completo), y pueden tamizarse típicamente para determinar la especificidad de la misma manera en que se tamizan los anticuerpos completos. En algunas modalidades, un fragmento de unión a epítipo comprende un fragmento F(ab')₂, fragmento Fab', fragmento Fab, fragmento Fd, o fragmento Fv. En algunas modalidades, el término "anticuerpo" incluye polipéptidos derivados de anticuerpos, tales como fragmentos variables de cadena única (scFv), diacuerpos u otros scFv multiméricos, anticuerpos de cadena pesada, anticuerpos de dominio único, u otros polipéptidos que comprenden una porción suficiente de un anticuerpo (*por ejemplo*, una o más regiones determinantes de la complementariedad (CDR)) para conferir capacidad de unión a antígeno específica al polipéptido.

(xv) Grupo de afinidad

Un "grupo de afinidad" es una molécula o porción molecular que tiene una alta afinidad o preferencia por asociarse o unirse con otra molécula o porción específica o particular. La asociación o unión con otra molécula o porción específica o particular puede ser mediante una interacción no covalente, tal como puente de hidrógeno, fuerzas iónicas, e interacciones de van der Waals. Un grupo de afinidad puede ser, por ejemplo, biotina, que tiene una alta afinidad o preferencia para asociar o unirse a la proteína avidina o estreptavidina. Un grupo de afinidad, por ejemplo, puede referirse, además, a avidina o estreptavidina que tiene afinidad por la biotina. Otros ejemplos de un grupo de afinidad y molécula o porción específica o particular con el que se une o se asocia incluyen, pero no se limitan a, anticuerpos o fragmentos de anticuerpos y sus respectivos antígenos, tales como anticuerpos de digoxigenina y anti-digoxigenina, lectina, y carbohidratos (por ejemplo, un azúcar, un monosacárido, un disacárido, o un polisacárido), y receptores y ligandos de receptores.

Cualquier par de grupos de afinidad y su molécula o porción específica o particular con el que se une o se asocia puede tener sus funciones inversas, por ejemplo, de manera que entre una primera molécula y una segunda molécula, en un primer caso la primera molécula se caracteriza como un grupo de afinidad por la segunda molécula, y en un segundo caso la segunda molécula se caracteriza como un grupo de afinidad por la primera molécula.

(xvi) Marcador, marcador detectable y marcador óptico

Los términos “marcador detectable”, “marcador óptico” y “marcador” se usan indistintamente en la presente descripción para referirse a una porción detectable directa o indirectamente que se asocia con (por ejemplo, conjugado a) una molécula a detectar, por ejemplo, una sonda para ensayo *in situ*, una sonda de captura o analito. El marcador detectable puede detectarse directamente por sí mismo (por ejemplo, marcadores de radioisótopos o marcadores fluorescentes) o, en el caso de un marcador enzimático, puede detectarse indirectamente, por ejemplo, mediante la catálisis de alteraciones químicas de un compuesto o composición de sustrato, cuyo compuesto o composición de sustrato es detectable directamente. Los marcadores detectables pueden ser adecuados para la detección a pequeña escala y/o adecuados para el tamizaje de alto rendimiento. Como tal, los marcadores detectables adecuados incluyen, pero no se limitan a, radioisótopos, fluoróforos, compuestos quimioluminiscentes, compuestos bioluminiscentes, y colorantes.

El marcador detectable puede detectarse cualitativamente (por ejemplo, ópticamente o espectralmente), o puede cuantificarse. La detección cualitativa incluye generalmente un método de detección en el que se confirma la existencia o presencia del marcador detectable, mientras que la detección cuantificable incluye generalmente un método de detección que tiene un valor cuantificable (por ejemplo, que puede informarse numéricamente) tal como una intensidad, duración, polarización, y/u otras propiedades. En algunas modalidades, el marcador detectable se une a un elemento o a una sonda de captura asociada con un elemento. Por ejemplo, los elementos marcados de manera detectable pueden incluir un marcador fluorescente, colorimétrico o quimioluminiscente que se unen a una perla (véase, por ejemplo, Rajeswari y otros, *J. Microbiol Methods* 139:22-28, 2017, y Forcucci y otros, *J. Biomed Opt.* 10:105010, 2015).

En algunas modalidades, una pluralidad de marcadores detectables pueden unirse a un elemento, sonda de captura, o composición a detectar. Por ejemplo, pueden incorporarse marcadores detectables durante la polimerización o amplificación de ácido nucleico (por ejemplo, nucleótidos marcados con Cy5®, tales como Cy5®-dCTP). Puede usarse cualquier marcador detectable adecuado. En algunas modalidades, el marcador detectable es un fluoróforo. Por ejemplo, el fluoróforo puede ser de un grupo que incluye: 7-AAD (7-aminoactinomicina D), naranja de acridina (+ADN), naranja de acridina (+ARN), Alexa Fluor® 350, Alexa Fluor® 430, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 555, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 633, Alexa Fluor® 647, Alexa Fluor® 660, Alexa Fluor® 680, Alexa Fluor® 700, Alexa Fluor® 750, alofococianina (APC), AMCA/AMCA-X, 7-aminoactinomicina D (7-AAD), 7-amino-4-metilcoumarina, 6-aminoquinolina, azul de anilina, ANS, APC-Cy7, ATTO-TAG™ CBQCA, ATTO-TAG™ FQ, auramina O-Feulgen, BCECF (pH alto), BFP (proteína fluorescente azul), BFP/GFP FRET, BOBO™-1/BO-PRO™-1, BOBO™-3/BO-PRO™-3, BODIPY® FL, BODIPY® TMR, BODIPY® TR-X, BODIPY® 530/550, BODIPY® 558/568, BODIPY® 564/570, BODIPY® 581/591, BODIPY® 630/650-X, BODIPY® 650-665-X, BTC, calceína, azul calceína, Calcium Crimson™, Calcium Green-1™, Calcium Orange™, Calcofluor® White, 5-carboxifluoresceína (5-FAM), 5-carboxinaftofluoresceína, 6-carboxirrodamina 6G, 5-carboxitetrametilrodamina (5-TAMRA), carboxi-X-rodamina (5-ROX), Cascade Blue®, Cascade Yellow™, CCF2 (GeneBLAzer™), CFP (proteína fluorescente cian), CFP/YFP FRET, cromomicina A3, Cl-NERF (pH bajo), CPM, 6-CR 6G, CTC Formazan, Cy2®, Cy3®, Cy3.5®, Cy5®, Cy5.5®, Cy7®, cicromo (PE-Cy5), dansilamina, dansil cadaverina, cloruro de dansilo, DAPI, dapoxilo, DCFH, DHR, DiA (4-Di-16-ASP), DiD (DiIC18(5)), DIDS, DiI (DiIC18(3)), DiO (DiOC18(3)), DiR (DiIC18(7)), Di-4 ANEPPS, Di-8 ANEPPS, DM-NERF (pH 4,5-6,5), DsRed (proteína fluorescente roja), EBFP, ECFP, EGFP, alcohol ELF® -97, eosina, eritrosina, bromuro de etidio, homodímero de etidio-1 (EthD-1), cloruro de europio (III), 5-FAM (5-carboxifluoresceína), Fast Blue, fosforamidita de fluoresceína-dT, FITC, Fluo-3, Fluo-4, FluorX®, Fluoro-Gold™ (pH alto), Fluoro-Gold™ (pH bajo), Fluoro-Jade, FM® 1-43, Fura-2 (alto contenido de calcio), Fura-2/BCECF, Fura Red™ (alto contenido de calcio), Fura Red™/Fluo-3, GeneBLAzer™ (CCF2), GFP desplazada al rojo (rsGFP), GFP de tipo silvestre, GFP/BFP FRET, GFP/DsRed FRET, Hoechst 33342 y 33258, 7-hidroxi-4-metilcoumarina (pH 9), 1,5 IAEDANS, indo-1 (alto contenido de calcio), Indo-1 (bajo contenido de calcio), indodicarbocianina, indotricarbocianina, JC-1, 6-JOE, JOJO™-1/JO-PRO™-1, LDS 751 (+ADN), LDS 751 (+ARN), LOLO™-1/LO-PRO™-1, amarillo Lucifer, azul LysoSensor™ (pH 5), verde LysoSensor™ (pH 5), amarillo/azul LysoSensor™ (pH 4,2), verde LysoTracker®, rojo LysoTracker®, amarillo MitoTracker®, naranja MitoTracker®, rojo MitoTracker®, NBD (amina), Rojo Nilo, Oregon Green® 488, Oregon Green® 500, Oregon Green® 514, Azul pacífico, PBF1, PE (R-ficoeritrina), PE-Cy5, PE-Cy7, PE-rojo Texas, PerCP (proteína peridina clorofila), PerCP-Cy5.5 (TruRed), PharRed (APC-Cy7), C-ficocianina, R-ficocianina, R-ficoeritrina (PE), IP (yoduro de propidio), PKH26, PKH67, POPO™-1/PO-PRO™-1, POPO™-3/PO-PRO™-3, yoduro de propidio (PI), PyMPO, pireno, pironina Y, Quantam Red (PE-Cy5), mostaza de quinacrina, R670 (PE-Cy5), rojo 613 (PE-Texas Red), proteína fluorescente roja (DsRed), resorufina, RH 414, Rhod-2, rodamina B, Rhodamine Green™, Rhodamine Red™, rodamina faloidina, rodamina 110, rodamina 123, 5-ROX (carboxi-X-rodamina), S65A, S65C, S65L, S65T, SBFI, SITS, SNAFL®-1 (pH alto), SNAFL®-2, SNARF®-1 (pH alto), SNARF®-1 (pH bajo), Sodium Green™, SpectrumAqua®, SpectrumGreen® #1, SpectrumGreen® #2, SpectrumOrange®, SpectrumRed®, SYTO® 11, SYTO® 13, SYTO® 17, SYTO® 45, azul SYTOX®, verde SYTOX®, naranja SYTOX®, 5-TAMRA (5-carboxitetrametilrodamina), tetrametilrodamina (TRITC), Texas Red®/Texas Red®-X, Texas Red®-X (éster NHS), tiadicarbocianina, naranja Tiazol, TOTO®-1/TO-PRO®-1, TOTO®-3/TO-PRO®-3, TO-PRO®-5, Tri-color (PE-Cy5), TRITC (tetrametilrodamina), TruRed (PerCP-Cy5.5), WW 781, X-rodamina (XRITC), Y66F, Y66H, Y66W, YFP (proteína fluorescente amarilla), YOYO®-1/YO-PRO®-1, YOYO®-3/YO-PRO®-3, 6-FAM (fluoresceína), 6-FAM (éster NHS), 6-FAM (azida), HEX, TAMRA (éster NHS), amarillo Yakima, MAX, TET, TEX615, ATTO 488, ATTO 532, ATTO 550, ATTO 565, ATTO Rho101, ATTO 590, ATTO 633, ATTO 647N, TYE 563, TYE 665, TYE 705, 5' IRDye® 700, 5'

IRDye® 800, 5' IRDye® 800CW (éster NHS), colorante WellRED D4, colorante WellRED D3, colorante WellRED D2, Lightcycler® 640 (éster NHS), y Dy 750 (éster NHS).

Como se mencionó anteriormente, en algunas modalidades, un marcador detectable es o incluye una porción luminiscente o quimioluminiscente. Las porciones luminiscentes/quimioluminiscentes comunes incluyen, pero no se limitan a, peroxidasa tales como peroxidasa de rábano picante (HRP), peroxidasa de soja (SP), fosfatasa alcalina, y luciferasa. Estas porciones de proteínas pueden catalizar reacciones quimioluminiscentes dados los sustratos apropiados (por ejemplo, un reactivo oxidante más un compuesto quimioluminiscente). Se conoce que una serie de familias de compuestos proporcionan quimioluminiscencia bajo una variedad de condiciones. Los ejemplos no limitantes de familias de compuestos quimioluminiscentes incluyen 2,3-dihidro-1,4-ftalazindiona luminol, 5-amino-6,7,8-trimetoxi- y el análogo dimetilamino[ca]benz. Estos compuestos pueden emitir luminiscencia en presencia de peróxido de hidrógeno alcalino o hipoclorito de calcio y base. Otros ejemplos de familias de compuestos quimioluminiscentes incluyen, por ejemplo, 2,4,5-trifenilimidazoles, sustituyentes para-dimetilamino y -metoxi, oxalatos tales como ésteres activos de oxalilo, p-nitrofenilo, ésteres de N-alquilacridinio, luciferinas, lucigeninas, o ésteres de acridinio. En algunas modalidades, un marcador detectable es o incluye un marcador basado en metales o basado en masa. Por ejemplo, los conglomerados pequeños de iones metálicos, metales, o semiconductores pueden actuar como un código de masa. En algunos ejemplos, los metales pueden seleccionarse de los grupos 3-15 de la tabla periódica, por ejemplo, Y, La, Ag, Au, Pt, Ni, Pd, Rh, Ir, Co, Cu, Bi, o una de sus combinaciones.

(xvii) Oligonucleótido que cambian de molde

Un "oligonucleótido que cambia de molde" es un oligonucleótido que se hibrida con nucleótidos no templados añadidos por una transcriptasa inversa (por ejemplo, enzima con actividad transferasa terminal) durante la transcripción inversa. En algunas modalidades, un oligonucleótido que cambia de molde hibrida con nucleótidos poli(C) no templados añadidos por una transcriptasa inversa. En algunas modalidades, el oligonucleótido que cambia de molde añade una secuencia 5' común al ADNc de longitud completa que se usa para la amplificación del ADNc.

En algunas modalidades, el oligonucleótido que cambia de molde añade una secuencia común sobre el extremo 5' del ARN que se transcribe de manera inversa. Por ejemplo, un oligonucleótido que cambia de molde puede hibridarse con nucleótidos poli(C) no templados que se añaden al extremo de una molécula de ADNc y proporcionar un molde para que la transcriptasa inversa continúe la replicación al extremo 5' del oligonucleótido que cambia de molde, de esta manera se genera un ADNc de longitud completa listo para amplificación adicional. En algunas modalidades, una vez que se genera una molécula de ADNc de longitud completa, el oligonucleótido que cambia de molde puede servir como un cebador en una reacción de amplificación de ADNc.

En algunas modalidades, se añade un oligonucleótido que cambia de molde antes, simultáneamente con, o después de una transcripción inversa, u otra reacción terminal basada en transferasa. Por ejemplo, en algunas modalidades, un oligonucleótido que cambia de molde se incluye en la sonda de captura. En determinadas modalidades, los métodos de análisis de muestras mediante el uso de oligonucleótidos que cambian de molde pueden involucrar la generación de productos de ácido nucleico a partir de analitos de la muestra de tejido, seguido del procesamiento posterior de los productos de ácido nucleico con el oligonucleótido que cambia de molde.

Los oligonucleótidos que cambian de molde pueden incluir una región de hibridación y una región de molde. La región de hibridación puede incluir cualquier secuencia capaz de hibridarse con la diana. En algunas modalidades, la región de hibridación puede incluir, por ejemplo, una serie de bases G para complementar las bases C salientes en el extremo 3' de una molécula de ADNc. La serie de bases G puede incluir 1 base G, 2 bases G, 3 bases G, 4 bases G, 5 bases G o más de 5 bases G. La secuencia molde puede incluir cualquier secuencia que se incorpore en el ADNc. En otras modalidades, la región de hibridación puede incluir al menos una base además de al menos una base G. En otras modalidades, la hibridación puede incluir bases que no son una base G. En algunas modalidades, la región del molde incluye al menos 1 (por ejemplo, al menos 2, 3, 4, 5 o más) secuencias de etiquetas y/o secuencias funcionales. En algunas modalidades, la región del molde y la región de hibridación se separan por un espaciador.

En algunas modalidades, las regiones del molde incluyen una secuencia de código de barras. La secuencia de código de barras puede actuar como un código de barras espacial y/o como un identificador molecular único. Los oligonucleótidos que cambian de molde pueden incluir ácidos desoxirribonucleicos; ácidos ribonucleicos; ácidos nucleicos modificados que incluyen 2-aminopurina, 2,6-diaminopurina (2-amino-dA), dT invertido, 5-metil dC, 2'-desoxiinosina, Super T (5-hidroxibutini-2'-desoxiuridina), Super G (8-aza-7-deazaguanosina), ácidos nucleicos bloqueados (LNA), ácidos nucleicos desbloqueados (UNA, por ejemplo, UNA-A, UNA-U, UNA-C, UNA-G), Iso-dG, Iso-dC, Bases de flúor 2' (por ejemplo, Fluoro C, Fluoro U, Fluoro A, y Fluoro G), o cualquier combinación de lo anterior.

En algunas modalidades, la longitud de un oligonucleótido que cambia de molde puede ser al menos aproximadamente 1, 2, 10, 20, 50, 75, 100, 150, 200 o 250 nucleótidos o más larga. En algunas modalidades, la longitud de un oligonucleótido que cambia de molde puede ser a lo máximo de aproximadamente 2, 10, 20, 50, 100, 150, 200 o 250 nucleótidos o más larga.

Ejemplos

Los siguientes ejemplos se incluyen solo con fines ilustrativos y no pretenden limitar el alcance de la presente descripción.

Ejemplo 1: Análisis de una muestra mediante el uso de sondas SNAIL para la secuenciación *in situ* seguido de un análisis basado en matriz espacial.

Este ejemplo ilustra un método para analizar una muestra biológica mediante la generación de secuencia e información espacial de moléculas de ácido nucleico diana mediante el uso de secuenciación *in situ* (por ejemplo, mediante el uso de microscopía de fluorescencia como lectura), seguido de un análisis basado en matriz espacial (por ejemplo, mediante el uso de secuenciación NGS de moléculas capturadas sobre la matriz) de la misma muestra (figura 1).

La muestra biológica (por ejemplo, muestra de tejido) se coloca sobre un portaobjetos de matriz fina, tal como un cubreobjetos con suficiente resistencia. El tejido puede fijarse mediante el uso de varias técnicas conocidas en la técnica. Por ejemplo, el tejido puede fijarse mediante el uso de paraformaldehído (PFA, formaldehído polimérico) y un agente de reticulación reversible. En algunos casos, las moléculas de ácido nucleico (por ejemplo, ARN) en la muestra pueden bloquearse en su lugar mediante la inclusión de la muestra en un hidrogel y reticular las moléculas de ácido nucleico al hidrogel. La muestra de tejido primero puede permeabilizarse con pepsina y HCl, y subsecuentemente fijarse. El tejido puede fijarse antes y/o después de la etapa de permeabilización y/o el tratamiento con el inductor de reticulaciones reversibles.

Después de la fijación del tejido, los ARNm se seleccionan por sondas y se analizan *in situ*, siguiendo enfoques típicos de obtención de imágenes *in situ* altamente multiplexados, por ejemplo, hibridación de sonda, ligación, amplificación por círculo rodante, seguido de secuenciación mediante ligación o secuenciación mediante hibridación, con lecturas de microscopía. Por ejemplo, sondas primarias *in situ* (por ejemplo, sondas SNAIL), una transcriptasa inversa termoestable (RT) (por ejemplo, cualquier enzima RT que funcione bien a temperaturas de hibridación de 40 °C o más alta), dNTP, oligonucleótidos que cambian de molde y otros componentes, pueden añadirse al tejido, junto con un tampón adecuado. Por lo tanto, en este ejemplo, el ADNc se genera temprano durante la transcripción inversa en el flujo de trabajo ilustrativo (por ejemplo, RT durante la hibridación de la sonda *in situ*).

En algunos casos, puede añadirse un catalizador que elimina las reticulaciones para capturar ARNm en portaobjetos, o para revertir las reticulaciones si se usa un formador de reticulaciones reversibles. El proceso de RT genera ADNc estable que se conserva durante el resto del protocolo *in situ*. El OH 3' de las sondas SNAIL está disponible para la amplificación por círculo rodante (RCA) pero no está disponible para RT (por ejemplo, porque el extremo 3' del cebador se hibrida con el candado y no con el ARN diana), de manera que no se produce una RT no deseada a partir del ARN de sondas candado no ligadas.

En algunos casos, puede realizarse una etapa de posfijación mediante el uso de PFA u otros agentes de reticulación. Para preservar el transcriptoma endógeno durante el protocolo *in situ*, pueden usarse varios inhibidores, que incluyen inhibidores de ribonucleasa (ARNasa) y cebadores aleatorios para proteger de la degradación del ARN mediada por la ARNasa A y/o ARNasa H en los híbridos ADN-ARN.

Una vez que se completa el protocolo *in situ*, los tejidos se permeabilizan opcionalmente con Proteinasa K (ProK) y/o una etapa de eliminación de las reticulaciones puede realizarse opcionalmente de manera que moléculas tales como ARNm, ADNc, sondas, productos de ligación y/o productos de amplificación ya no se bloquean en su lugar (por ejemplo, a un hidrogel). En algunos casos, las sondas (por ejemplo, sondas SNAIL) migran a una matriz que comprende sondas de captura con código de barras espacial. En algunos casos, la sonda candado del conjunto de sondas SNAIL comprende una secuencia que puede capturarse mediante la matriz portaobjeto. El tejido se lisa y el transcriptoma (por ejemplo, el ADNc que se genera) se captura en la matriz portaobjeto. Después se realiza la preparación y secuenciación de la biblioteca.

En algunos casos para analizar los ARNm directamente, pueden usarse varios enfoques para bloquear los ARNm en su lugar sobre la matriz portaobjeto. Por ejemplo, pueden realizarse sondas modificadas (por ejemplo, sondas oligo(dT)) captura de ARNm ((por ejemplo, oligo-dT FISH), seguido de reticulación y del uso de sondas candado de código de barras que se dirigen aleatoriamente lo que permiten la RCA con nucleótidos modificados, bloquear el amplicón en su lugar. En este enfoque, pueden analizarse *in situ* algunos cientos de ARN, mientras que los ARN restantes se analizan a través de la captura sobre una matriz espacial y por secuenciación.

Ejemplo 2: Análisis de una muestra mediante el uso de sondas candado y sondas puente para la secuenciación *in situ* seguido de un análisis basado en matriz espacial.

Este ejemplo ilustra un método alternativo para analizar un ácido nucleico diana en una muestra, que genera tanto información espacial como de secuenciación (figura 1).

La muestra biológica (por ejemplo, muestra de tejido) se fija sobre un portaobjetos y opcionalmente se permeabiliza como se describe en el Ejemplo 1. Después de la fijación de tejidos, los ARNm se seleccionan por sondas primarias *in situ* que comprenden sondas candado, cada una que comprende una secuencia de anclaje y una secuencia de

código de barras, por ejemplo, como se describe en Gyllborg y otros, "Hybridization-based *In Situ* Sequencing (HybISS): spatial transcriptomic detection in human and mouse brain tissue," bioRxiv 2020.02.03.931618. Las sondas primarias (sondas candado) se hibridan con la muestra de tejido, y las sondas no unidas se lavan de la muestra. Las sondas candado hibridan con el ARN correspondiente a los genes de interés y se ligan mediante el uso de la ligación con molde de ARN, de manera que el 3' OH de las sondas candado ya no está disponible para la RT. La RT se realiza después de la ligación de la sonda candado en el flujo de trabajo ilustrativo.

Después, el círculo cerrado se amplifica mediante una ADN polimerasa en una reacción de RCA. El proceso de RT genera ADNc estable que se conserva durante el resto del protocolo *in situ*.

En algunos casos, puede realizarse una etapa de posfijación mediante el uso de PFA u otros agentes de reticulación. Para preservar el transcriptoma endógeno durante el protocolo *in situ*, pueden usarse varios inhibidores, que incluyen inhibidores de ribonucleasa (ARNasa) y cebadores aleatorios para proteger de la degradación del ARN mediada por la ARNasa A y/o ARNasa H en los híbridos ADN-ARN.

Una vez que se completa el protocolo *in situ*, los tejidos se permeabilizan opcionalmente con Proteinasa K (ProK) y/o una etapa de eliminación de las reticulaciones puede realizarse opcionalmente de manera que moléculas tales como ARNm, ADNc, sondas, productos de ligación y/o productos de amplificación ya no se bloquean en su lugar (por ejemplo, a un hidrogel). En algunos casos, las sondas migran hacia una matriz que comprende sondas de captura con código de barras espacial. En algunos casos, la sonda candado comprende una secuencia que puede capturarse por la matriz. El tejido se lisa y el transcriptoma (por ejemplo, el ADNc que se genera) se captura en la matriz portaobjeto. Después se realiza la preparación de la biblioteca y la secuenciación del transcriptoma. Este método puede capturar la transcripción completa sobre una matriz con resolución de nucleótidos, en tándem con una lectura de secuenciación completa.

La presente invención no pretende limitarse en alcance a las modalidades particulares descritas, que se proporcionan, por ejemplo, para ilustrar diversos aspectos de la invención. Diversas modificaciones a las composiciones y métodos descritos se harán evidentes a partir de la descripción y las enseñanzas en la presente descripción.

REIVINDICACIONES

1. Un método para analizar una muestra biológica, que comprende:
 - 5 (a) poner en contacto la muestra biológica sobre un primer sustrato con una o más primeras sondas de ácido nucleico que se hibridan con un primer ácido nucleico diana en la muestra biológica, en donde el primer ácido nucleico diana es un ARN;
 - (b) obtener imágenes de la muestra biológica para detectar la una o más primeras sondas de ácido nucleico en una ubicación espacial de la muestra biológica, en donde la una o más primeras sondas de ácido nucleico comprenden una o más secuencias de código de barras que corresponden a una secuencia del primer ácido nucleico diana;
 - 10 (c) proporcionar condiciones para permitir que una pluralidad de agentes de captura interactúen con un producto de ligación que se ha liberado de la muestra biológica, en donde el producto de ligación se genera a partir de la ligación con molde de un par de sondas que se hibridan con un segundo ácido nucleico diana, en donde la pluralidad de agentes de captura se unen directa o indirectamente a un segundo sustrato, y en donde un agente de captura de la pluralidad de agentes de captura comprende:
 - (i) un dominio de captura capaz de capturar el producto de la ligación, y
 - (ii) un código de barras espacial; y
 - 15 (d) generar un polinucleótido marcado espacialmente que comprende (i) una secuencia del producto de ligación o complemento de esta y (ii) una secuencia del código de barras espacial o complemento de esta.
2. El método de la reivindicación 1, en donde el primer ácido nucleico diana y/o el segundo ácido nucleico diana son moléculas de ARNm.
- 25 3. El método de la reivindicación 1 o la reivindicación 2, en donde la muestra biológica es una muestra de tejido, tal como una sección de tejido.
4. El método de cualquiera de las reivindicaciones 1-3, en donde el par de sondas se pone en contacto con la muestra biológica en la etapa (a).
- 30 5. El método de cualquiera de las reivindicaciones 1-3, en donde el par de sondas se pone en contacto con la muestra biológica después de la etapa (b).
6. El método de cualquiera de las reivindicaciones 1-5, en donde la una o más primeras sondas de ácido nucleico comprenden una sonda primaria que se hibrida con el primer ácido nucleico diana, en donde la sonda primaria es una sonda circularizable.
- 35 7. El método de la reivindicación 6, que comprende generar un producto de amplificación por círculo rodante (RCA) de la sonda circularizable.
- 40 8. El método de la reivindicación 7, en donde la etapa (b) comprende poner en contacto la muestra biológica con una o más sondas marcadas de manera detectable capaces de hibridarse con el producto de RCA.
9. El método de la reivindicación 7, en donde la etapa (b) comprende poner en contacto la muestra biológica con una o más sondas secundarias capaces de hibridarse con el producto de RCA, y con una o más sondas marcadas de manera detectable capaces de hibridarse con la una o más sondas secundarias.
- 45 10. El método de cualquiera de las reivindicaciones 1-9, en donde el agente de captura comprende además un dominio universal que es 5' con respecto al código de barras espacial, en donde el dominio universal comprende:
 - (i) un dominio de amplificación; y/o
 - 50 (ii) un dominio de escisión para liberar el polinucleótido que se genera marcado espacialmente desde la superficie del segundo sustrato.
11. El método de cualquiera de las reivindicaciones 1-10, en donde el dominio de captura no captura el primer ácido nucleico diana, un complemento de este, o un producto de amplificación de este previo a la etapa (c).
- 55 12. El método de cualquiera de las reivindicaciones 1-11, en donde el polinucleótido marcado espacialmente o una porción de este se libera del segundo sustrato para el análisis, y el polinucleótido marcado espacialmente liberado o porción de este se analiza mediante secuenciación.
- 60 13. El método de cualquiera de las reivindicaciones 1-12, en donde el método comprende además correlacionar el código de barras espacial del polinucleótido marcado espacialmente con la ubicación espacial detectada de la una o más primeras sondas de ácido nucleico.

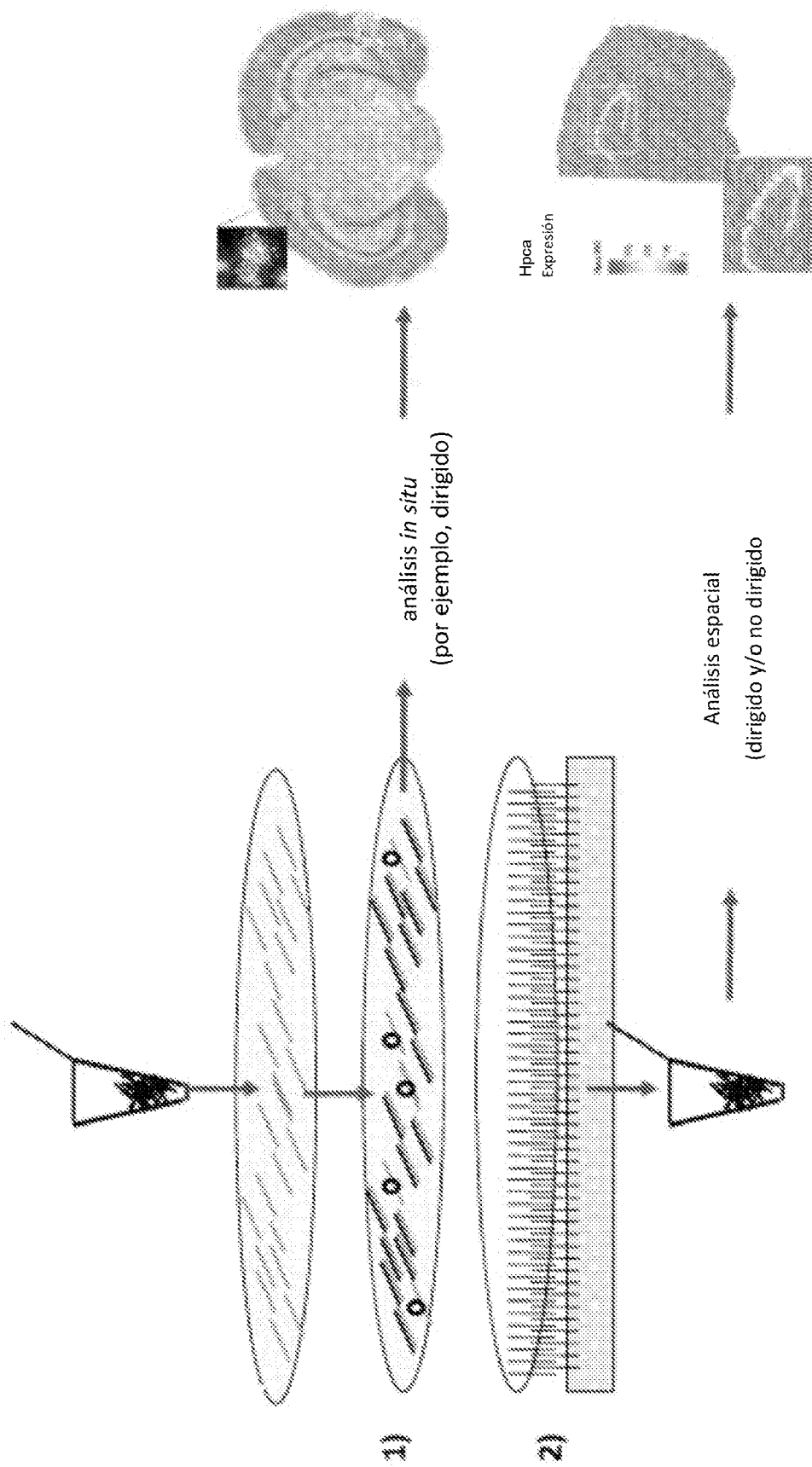


FIGURA 1

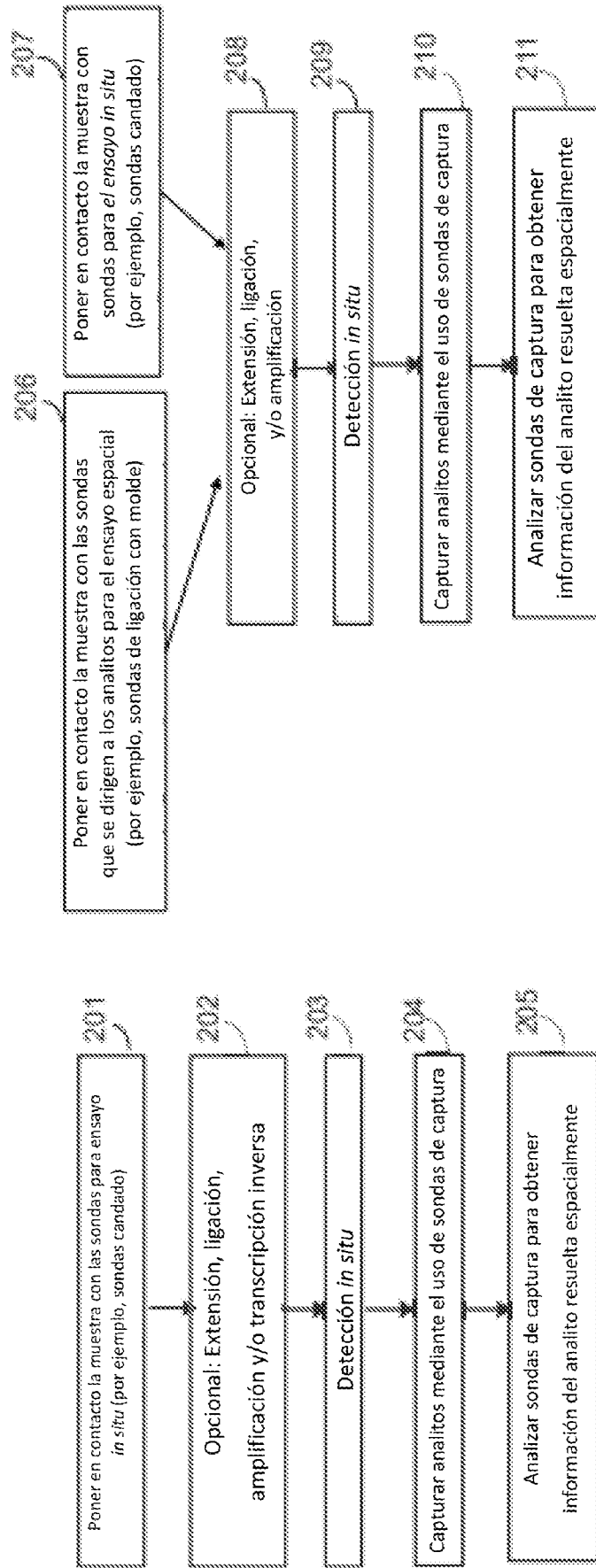


FIGURA 2A

FIGURA 2B

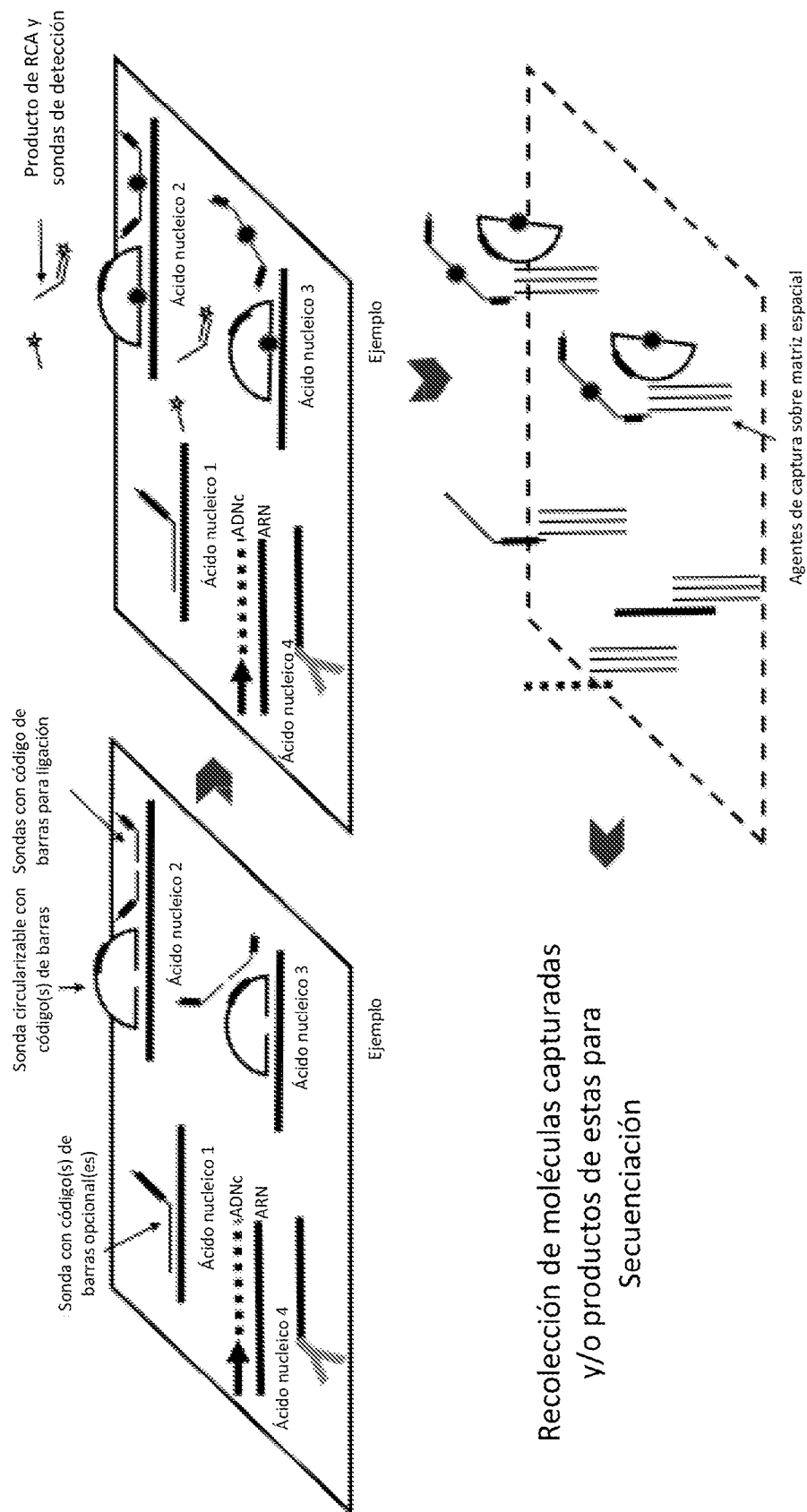


FIGURA 3

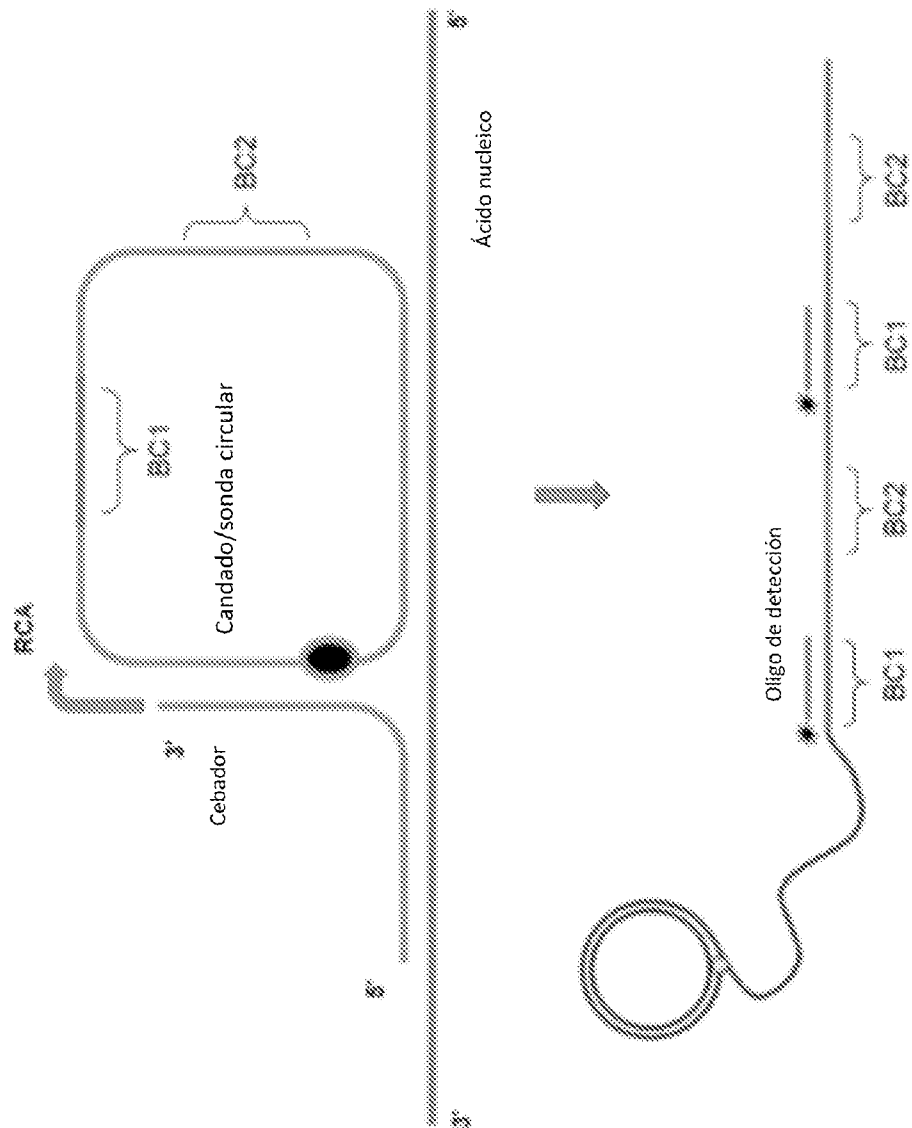


FIGURA 4A

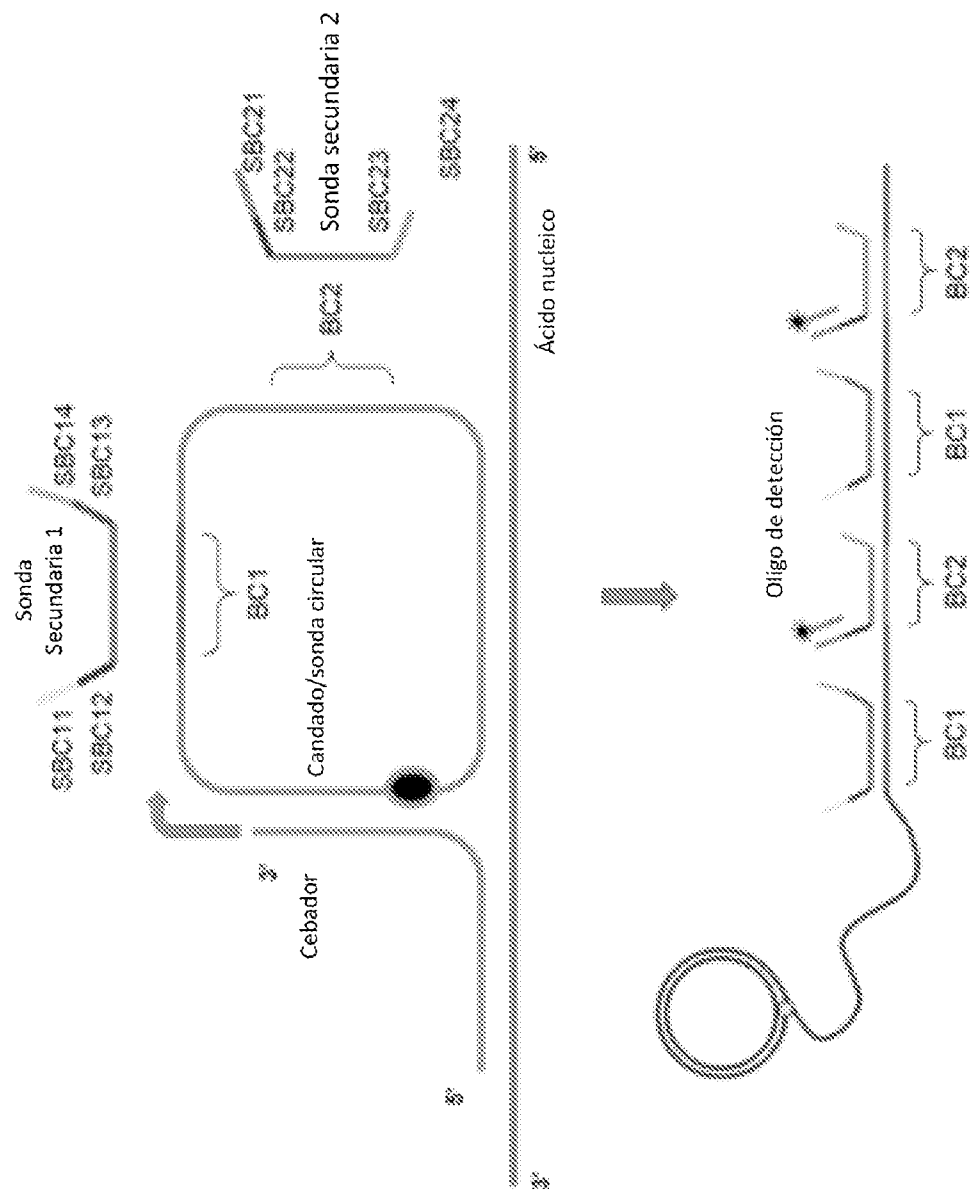


FIGURA 4B

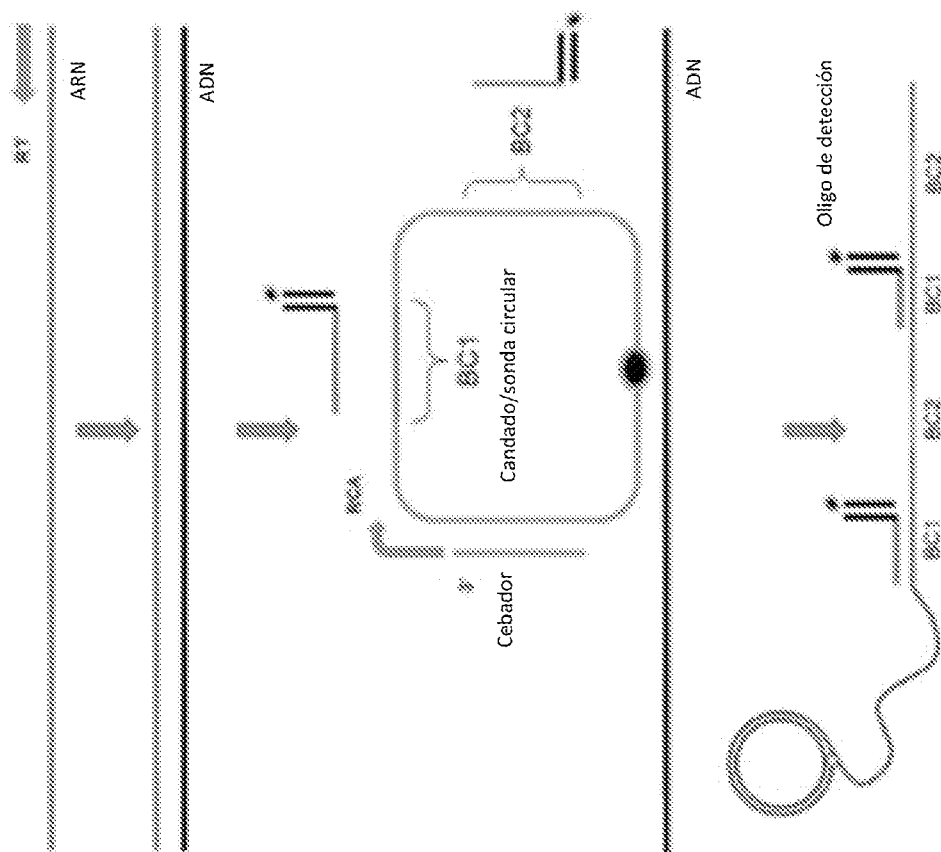


FIGURA 4C

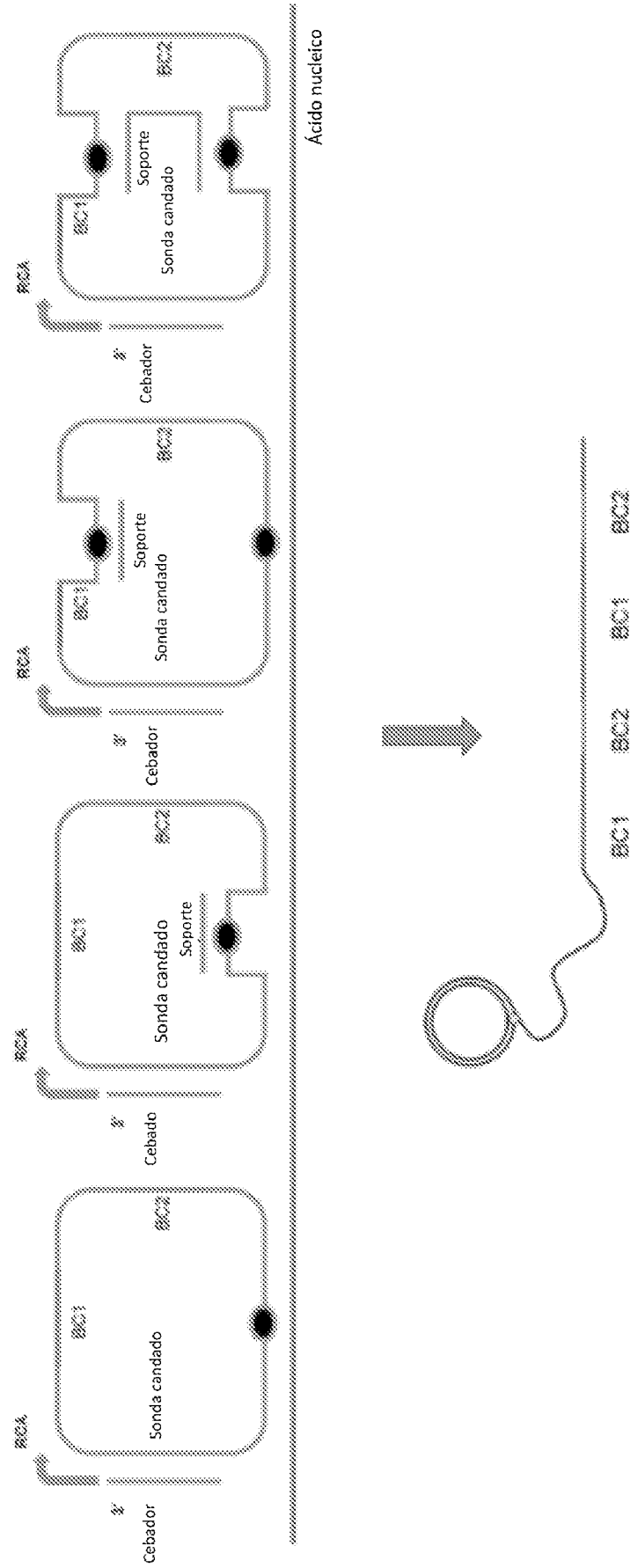


FIGURA 5

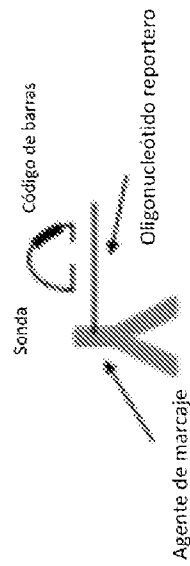
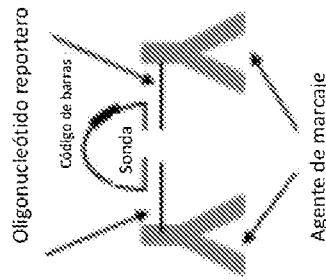
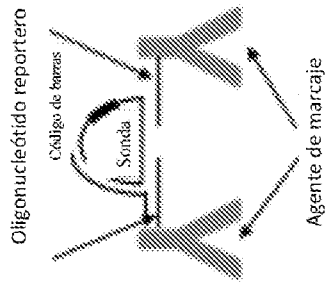


FIGURA 6A

FIGURA 6B

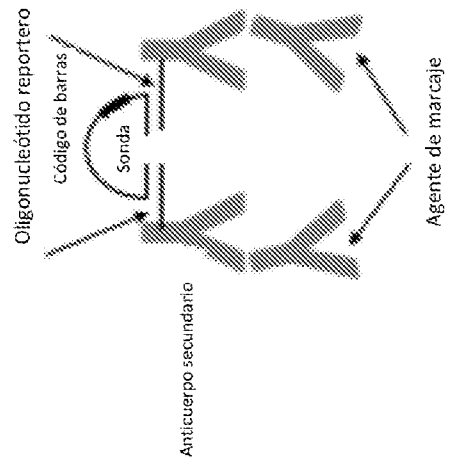


FIGURA 6F

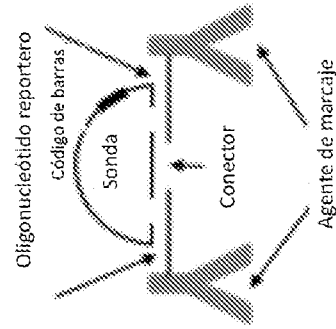


FIGURA 6E

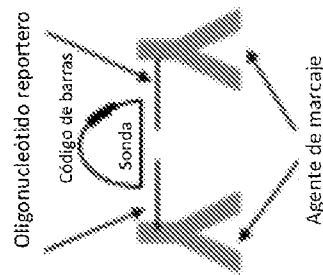


FIGURA 6D

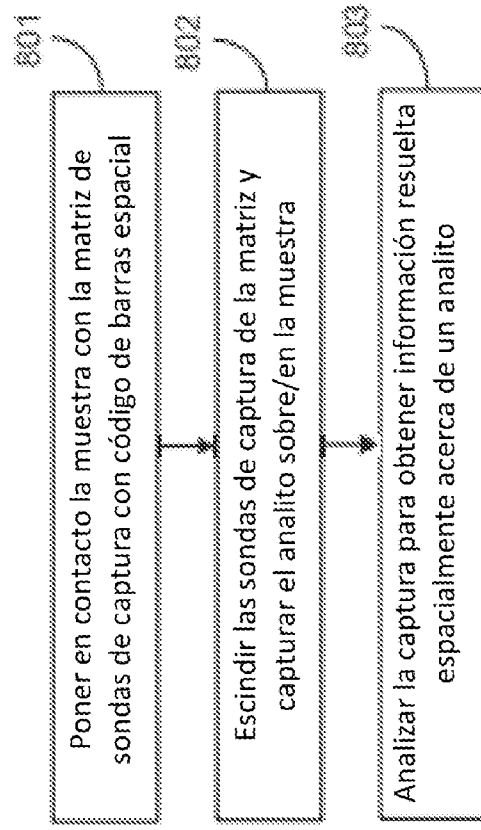


FIGURA 8

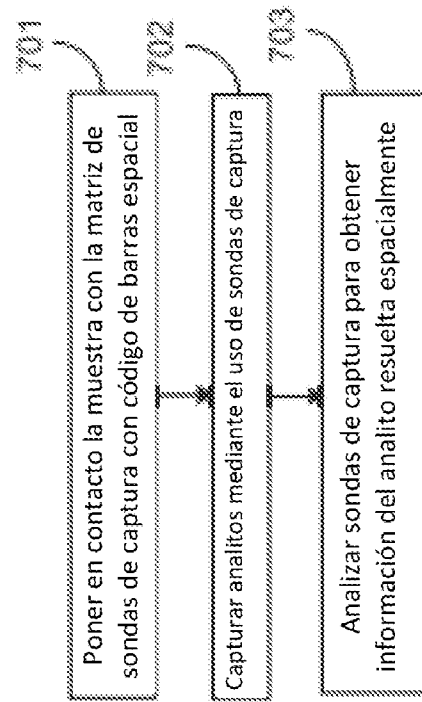


FIGURA 7

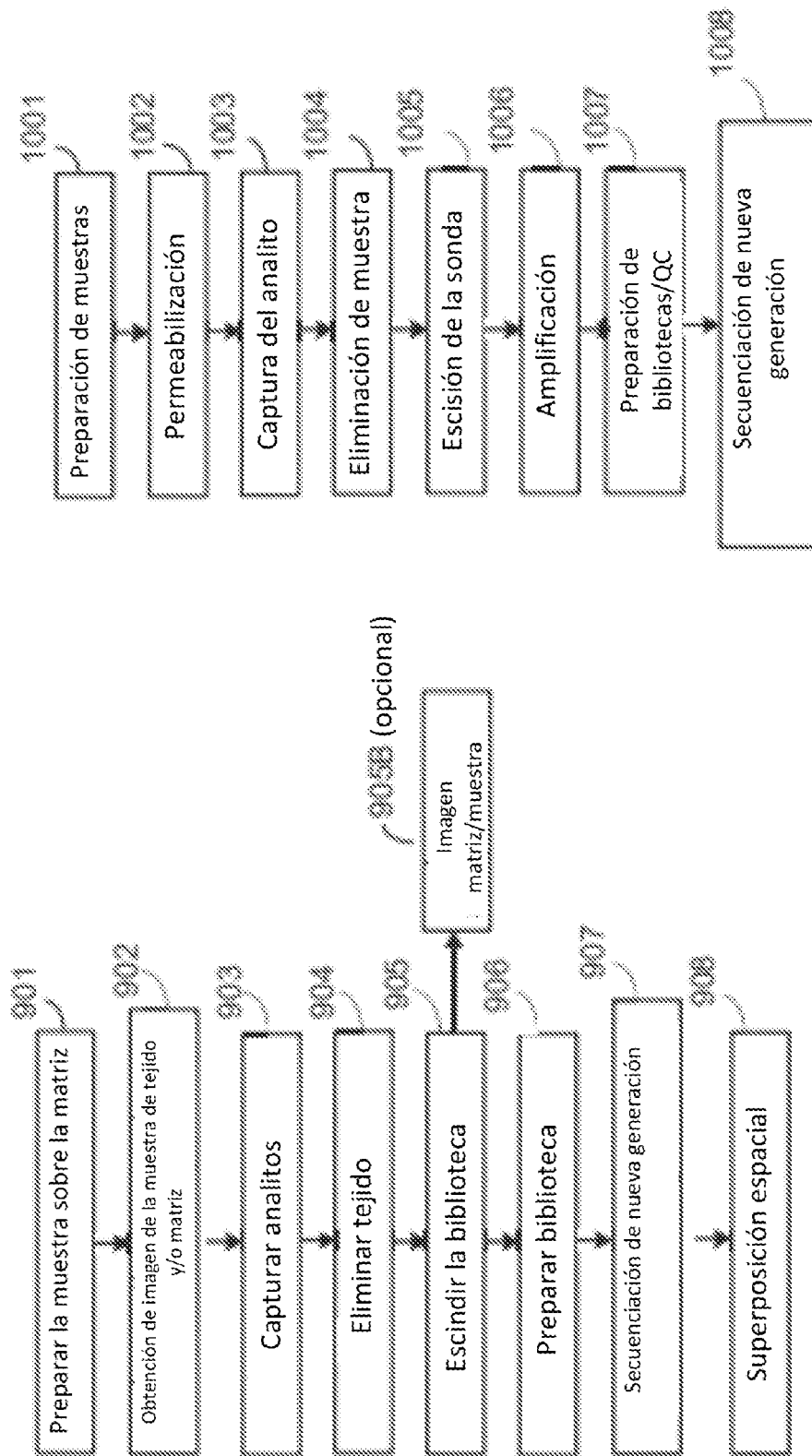


FIGURA 9

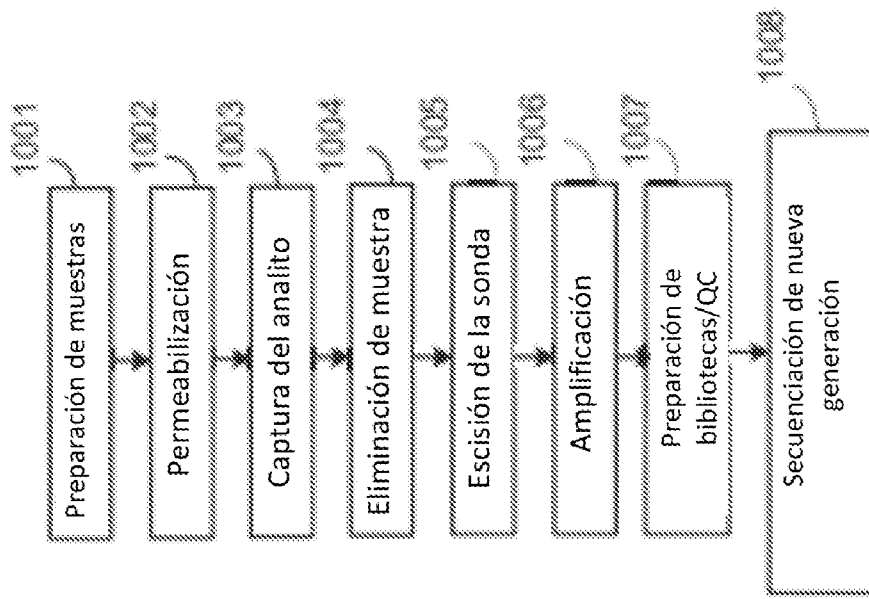


FIGURA 10

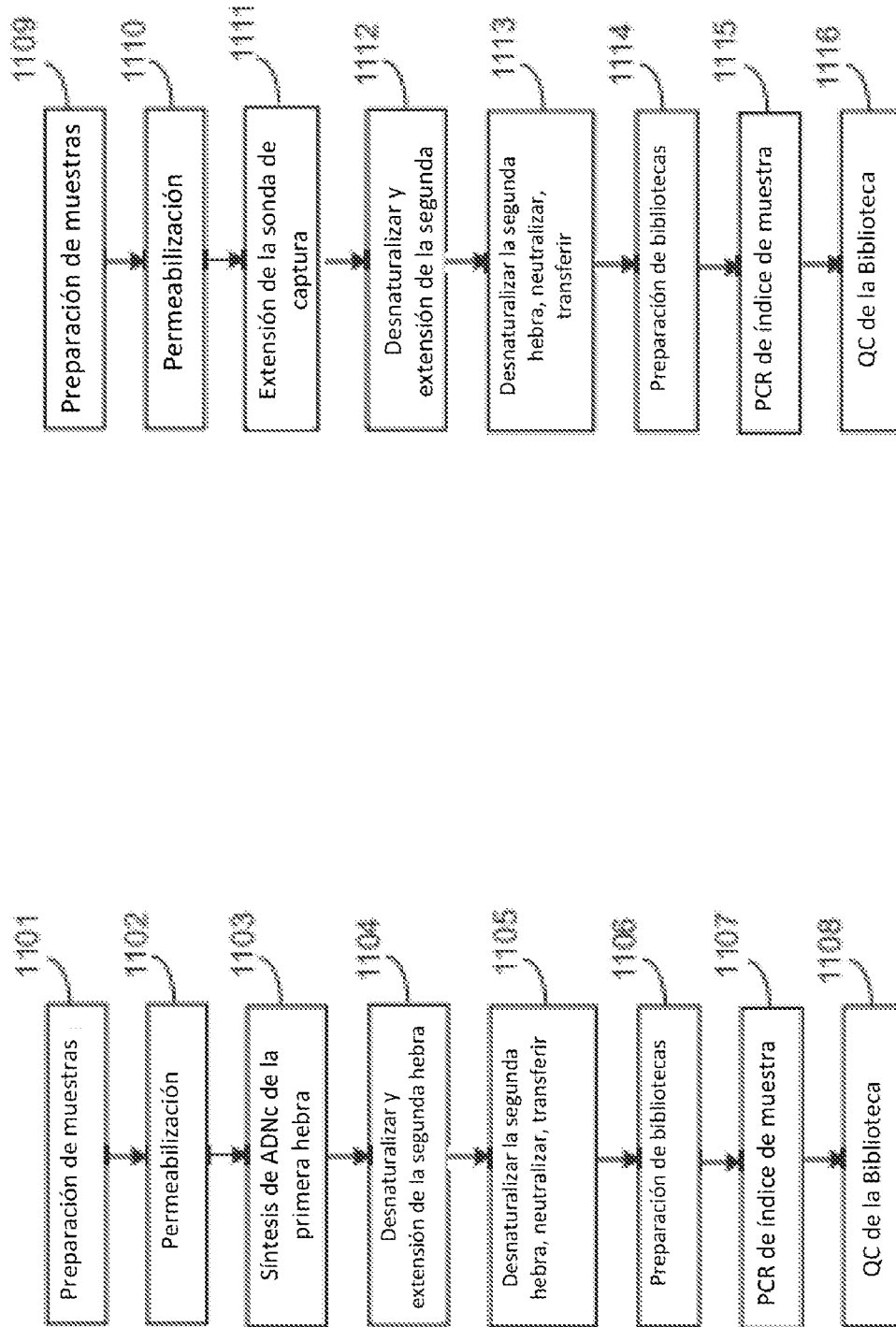


FIGURA 11B

FIGURA 11A

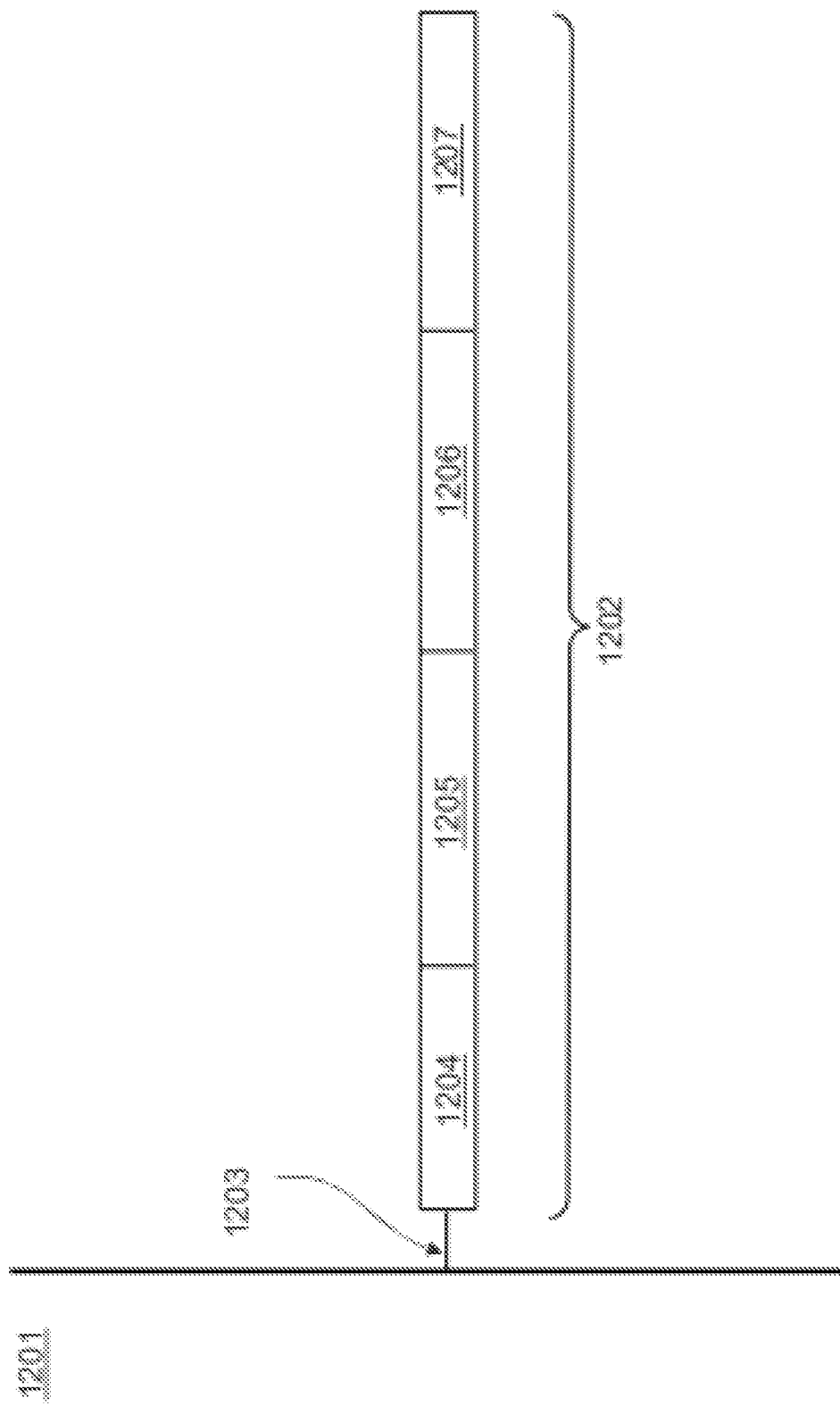


FIGURA 12

FIGURA 13A

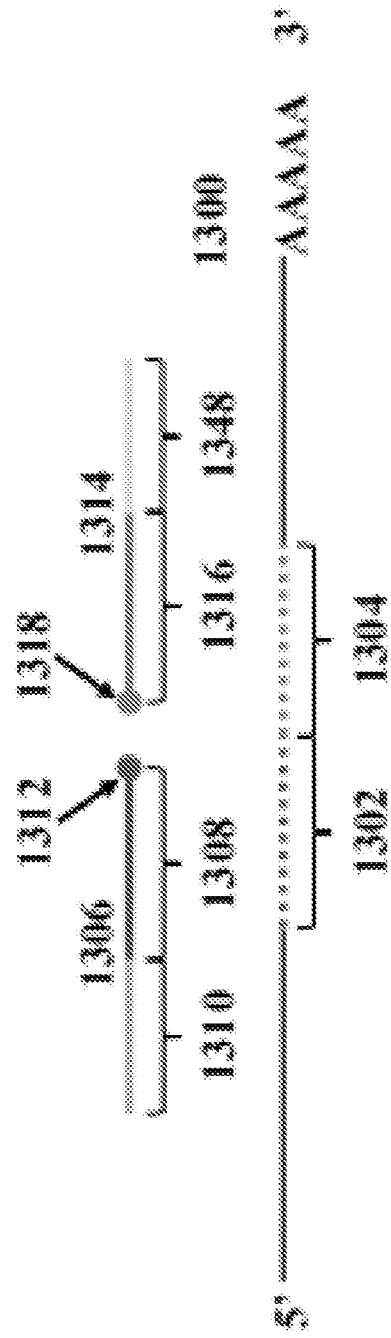


FIGURA 13B

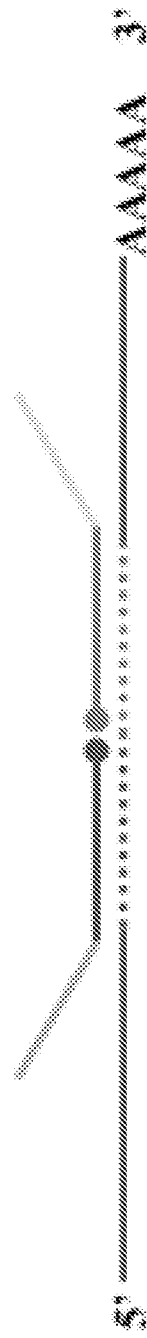
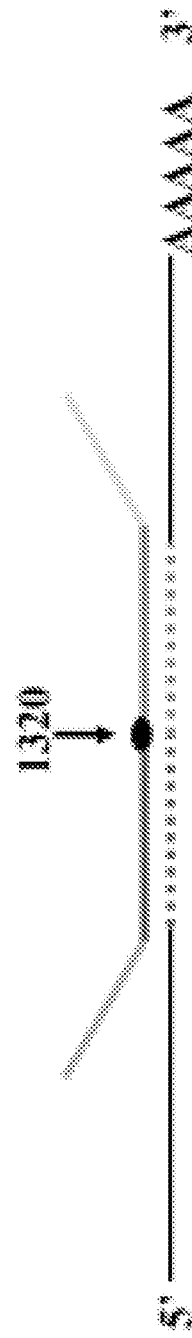


FIGURA 13C



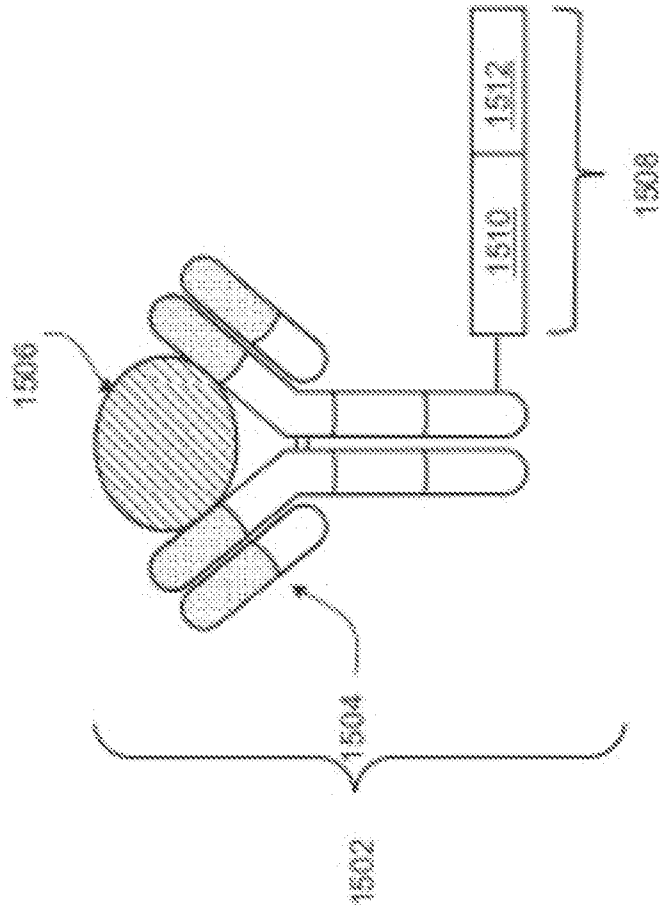


FIGURE 15

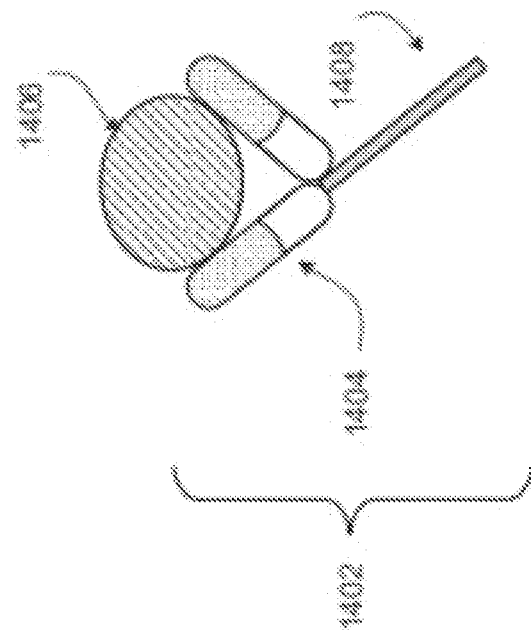


FIGURE 14

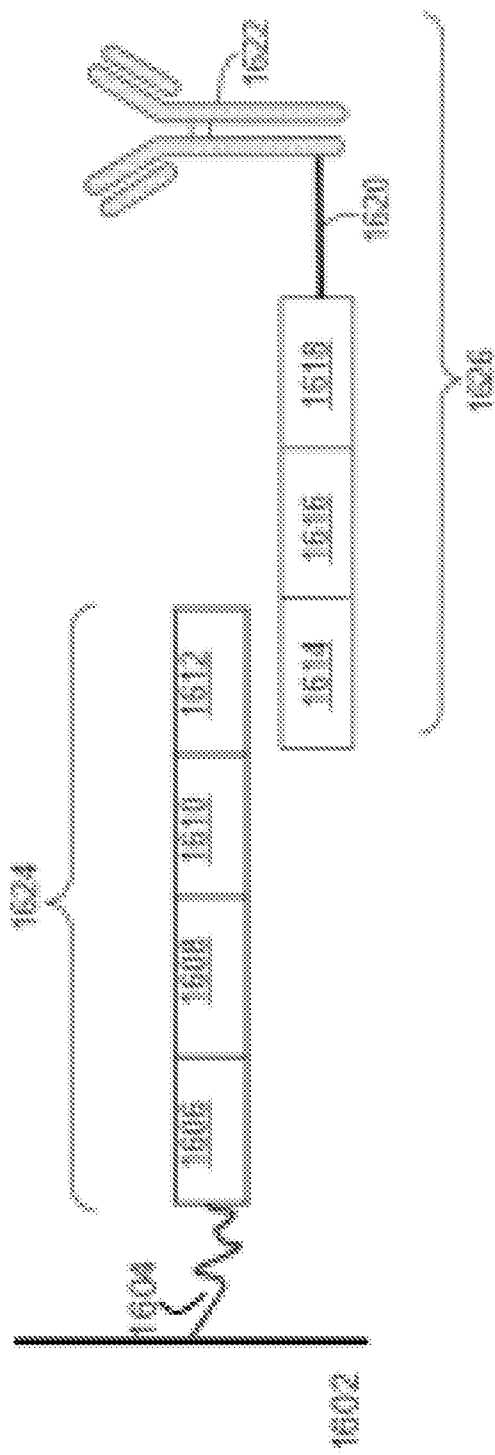


FIGURA 16

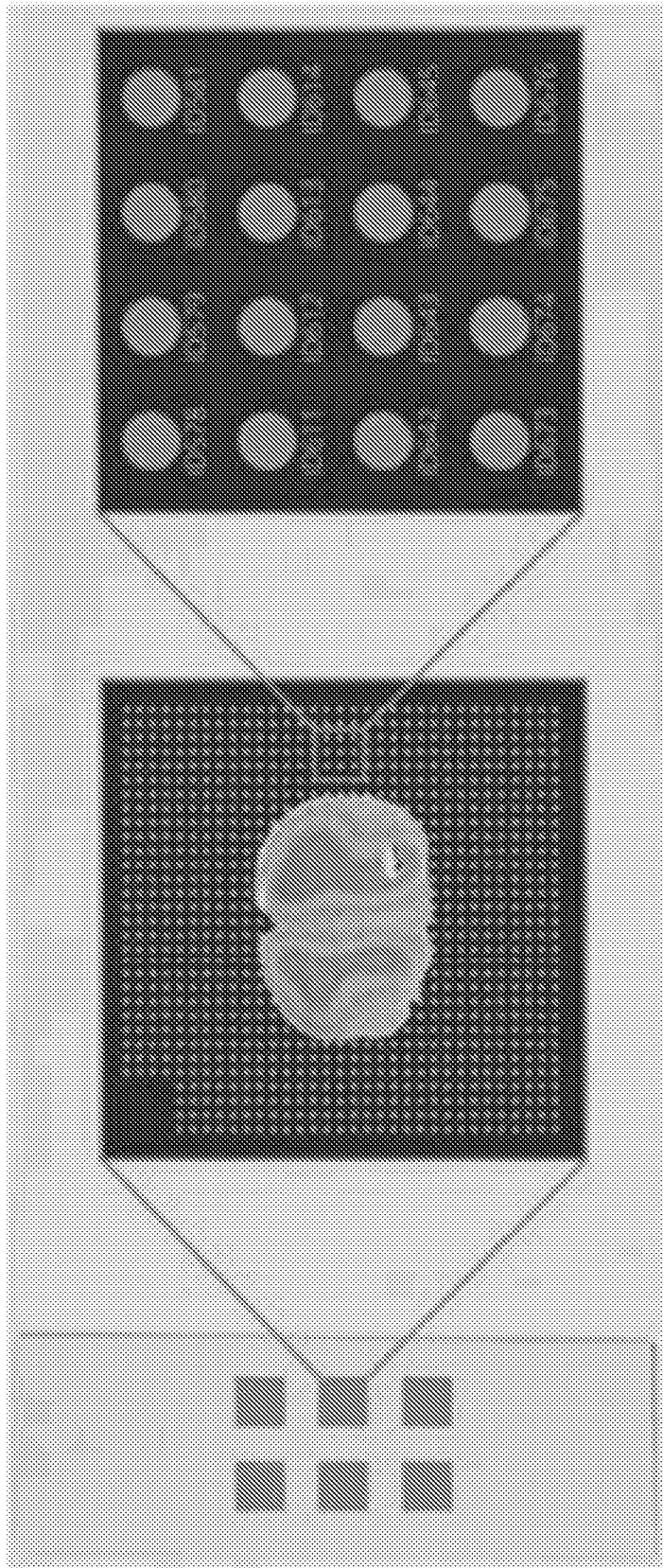


FIGURA 17

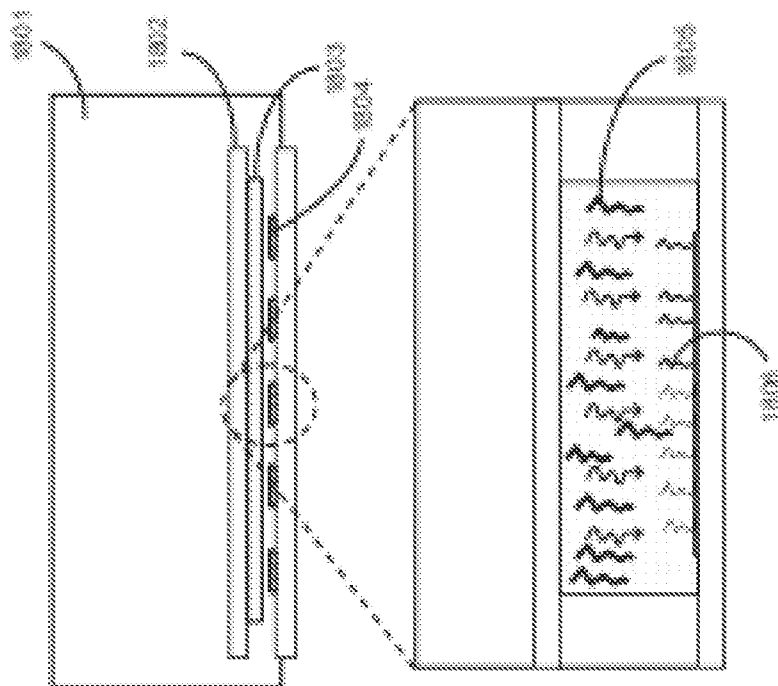


FIGURE 18

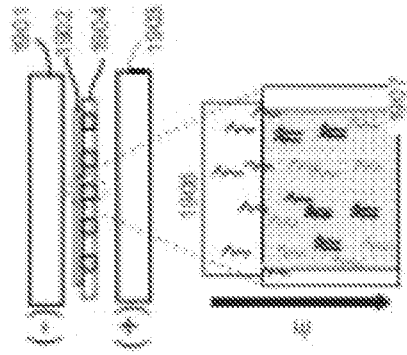


FIGURE 19B

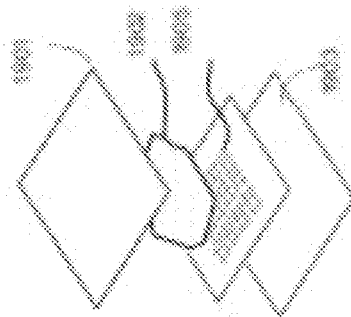
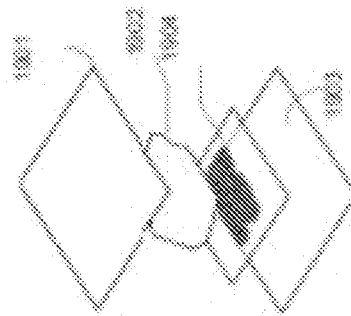


FIGURE 19A



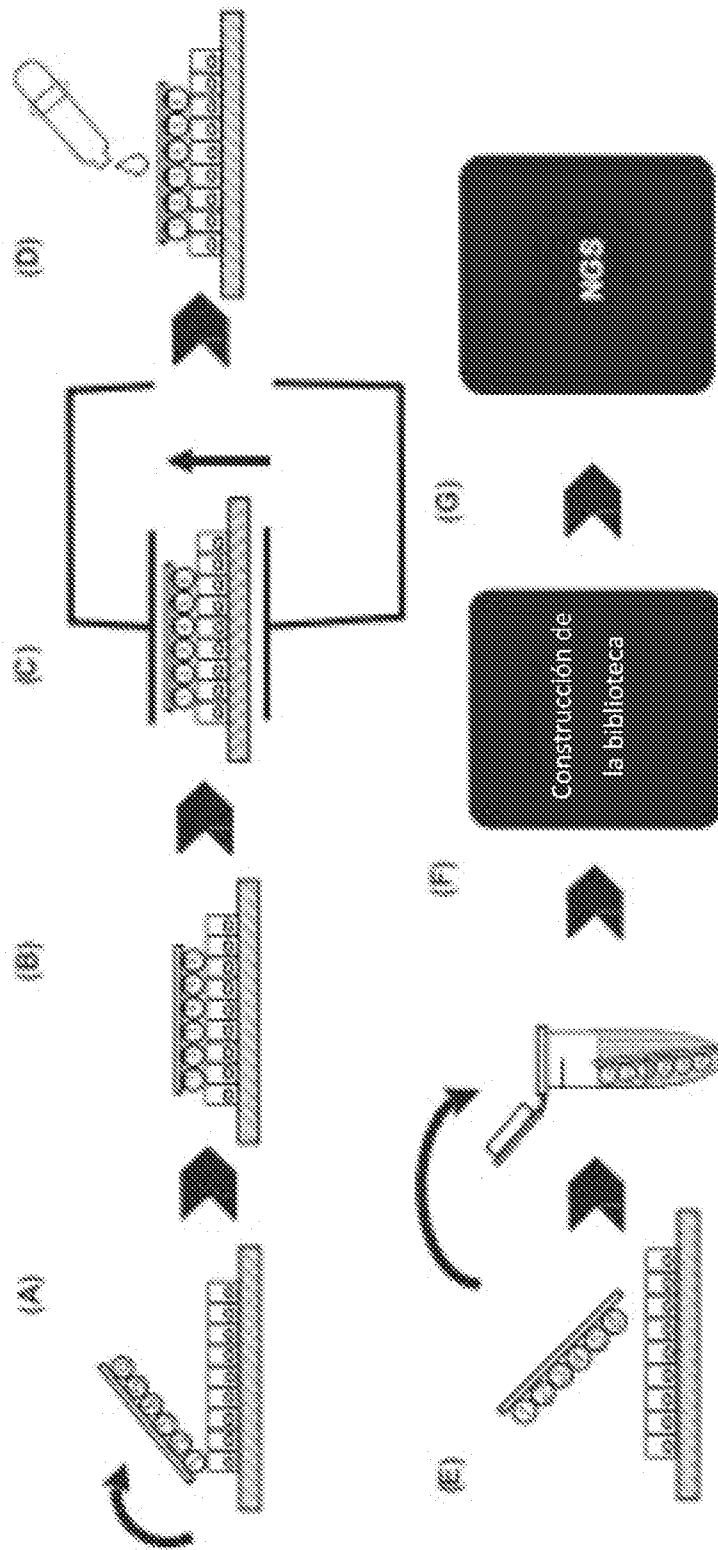


FIGURA 20