

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 945 191**

51 Int. Cl.:

C12Q 1/6806 (2008.01)

C12Q 1/6869 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.05.2019** **PCT/US2019/030245**

87 Fecha y número de publicación internacional: **07.11.2019** **WO19213294**

96 Fecha de presentación y número de la solicitud europea: **01.05.2019** **E 19724003 (9)**

97 Fecha y número de publicación de la concesión europea: **05.04.2023** **EP 3788171**

54 Título: **Análisis de muestras multiómicas de alto rendimiento**

30 Prioridad:

03.05.2018 US 201862666483 P

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

29.06.2023

73 Titular/es:

BECTON, DICKINSON AND COMPANY (100.0%)

1 Becton Drive

Franklin Lakes, NJ 07417, US

72 Inventor/es:

FAN, CHRISTINA y

WALCZAK, ELISABETH, MARIE

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

ES 2 945 191 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

Análisis de muestras multiómicas de alto rendimiento

5 SOLICITUDES RELACIONADAS

La presente solicitud reivindica el beneficio de la Solicitud Provisional de los Estados Unidos Nº 62/666,483, presentada el 3 de mayo de 2018.

10 Campo

La presente divulgación se refiere de manera general al campo de la biología molecular y, en particular, a análisis multiómico de células que usa códigos de barras moleculares.

15 Descripción de la técnica relacionada

Los métodos y técnicas, como los códigos de barras moleculares, son útiles para el análisis transcriptómico de células individuales, incluyendo el descifrado de perfiles de expresión génica para determinar los estados de las células usando, por ejemplo, transcripción inversa, amplificación por reacción en cadena de la polimerasa (PCR) y secuenciación de próxima generación (NGS). El código de barras molecular también es útil para el análisis proteómico de células individuales. Hay una necesidad de métodos y técnicas para el análisis multiómico de células individuales. La WO2016/061517 se refiere a transposición que conserva la contigüidad.

SUMARIO

25

La invención reivindicada se define en las reivindicaciones adjuntas.

En la presente se divulgan, pero no forman parte de la invención reivindicada, realizaciones de un método de análisis de muestras. Por ejemplo, el análisis de muestras puede comprender, consistir esencialmente en, o consistir en análisis de células individuales. En algunas realizaciones, el método incluye: poner en contacto ácido desoxirribonucleico de cadena doble (ADNdc) (por ejemplo, ADN genómico (ADNg)) de una célula, ya sea que el ADNg esté en la célula, un orgánulo de la célula como el núcleo o una mitocondria, o una fracción celular o extracto durante el contacto) con un transposoma. El transposoma puede comprender una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc (por ejemplo, una transposasa), y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5'. El método puede comprender codificar con códigos de barras la pluralidad de fragmentos de ADNdc salientes usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADN codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden secuencias de marcadores moleculares diferentes, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcador celular idéntica. El método puede comprender detectar secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras. El método puede comprender determinar la información que relaciona las secuencias de ADNdc con la estructura que comprende el ADNdc sobre la base de las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación. El método puede comprender además poner en contacto la pluralidad de fragmentos de ADNdc salientes con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, comprendiendo cada uno una secuencia complementaria a por lo menos una parte del saliente 5'; y desnaturalizar la pluralidad de fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc), en los que los fragmentos de ADNmc están codificados con códigos de barras, codificando con barras de este modo los fragmentos de ADN. En algunas realizaciones, el ADNdc comprende, consiste esencialmente en, o consiste en ADNg. En cualquier método de análisis de muestras como se describe en la presente, el transposoma puede dirigirse a una estructura específica que comprende ADNdc, por ejemplo, cromatina, un estado de metilación de ADN particular, un ADN en un orgánulo específico o similar. Se contempla que el método de análisis de muestras pueda identificar secuencias de ADN particulares asociadas con estructuras dirigidas por el transposoma, por ejemplo, ADN accesible por cromatina, ADN de constructos, ADN de orgánulos, o similares.

Se divulga, pero no forma parte de la invención reivindicada, un método de análisis de muestras que incluye: generar una pluralidad de fragmentos de ácidos nucleicos a partir de ADNdc (por ejemplo, ADNg de una célula, ya sea que el ADNg esté en la célula o en el núcleo de la célula, durante el contacto), en donde cada uno de la pluralidad de fragmentos de ácidos nucleicos comprende una secuencia de captura, un complemento de la secuencia de captura, un complemento inverso de la secuencia de captura, o una combinación de los mismos; codificar con códigos de barras la pluralidad de fragmentos de ácidos nucleicos usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADN codificados con códigos de barras, en donde cada uno de

la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden diferentes secuencias de marcadores moleculares, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcador celular idéntica; y detectar secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras. El método puede comprender además determinar la información que relaciona las secuencias de ADNdc con una estructura que comprende el ADNdc sobre la base de las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, generar la pluralidad de fragmentos de ácidos nucleicos puede comprender: poner en contacto el ADNdc con un transposoma, en el que el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y dos copias de un adaptador que comprende la secuencia de captura, para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a la secuencia de captura. La nucleasa de cadena doble puede cargarse con las dos copias del adaptador. El método puede comprender además desnaturalizar los fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc). El método puede comprender codificar con códigos de barras de la pluralidad de fragmentos de ADNmc, generando así la pluralidad de fragmentos de ADN codificados con códigos de barras. El método puede comprender además desnaturalizar el fragmento de ADN codificado con códigos de barras para generar fragmentos de ADN de cadena sencilla (ADNmc) codificados con códigos de barras.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, generar la pluralidad de fragmentos de ácidos nucleicos puede comprender: poner en contacto el ADNdc con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura, para generar una pluralidad de fragmentos de ADNdc salientes, cada uno con dos copias de los salientes 5'; y poner en contacto la pluralidad de fragmentos de ADNdc salientes que tienen los salientes 5' con una polimerasa para generar la pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria con por lo menos una parte de los salientes 5'. La nucleasa de cadena doble puede cargarse con las dos copias del adaptador. El método puede comprender además la desnaturalización de los fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc). El método puede comprender codificar con códigos de barras los fragmentos de ADNmc, generando de este modo el ADN codificado con códigos de barras. El método puede comprender además desnaturalizar el fragmento de ADN codificado con códigos de barras para generar fragmentos de ADN de cadena sencilla (ADNmc) codificados con códigos de barras. En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, los fragmentos de ADN codificados con códigos de barras pueden ser fragmentos de ADNmc.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, ninguno de la pluralidad de fragmentos de ADNdc complementarios comprende un saliente (por ejemplo, un saliente 3' o un saliente 5'). En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, el adaptador puede comprender una secuencia final de ADN del transposón. A modo de ejemplo, la nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc puede comprender una transposasa, tal como una transposasa Tn5. En la presente se describen ejemplos de otras transposasas adecuadas. En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, la pluralidad de fragmentos de ADNdc complementarios comprende cada uno extremos romos.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, generar la pluralidad de fragmentos de ácidos nucleicos comprende: fragmentar el ADNdc para generar una pluralidad de fragmentos de ADNdc. Fragmentar el ADNdc puede comprender poner en contacto el ADNdc con una enzima de restricción para generar la pluralidad de fragmentos de ADNdc, cada uno con uno o dos extremos romos. En algunas realizaciones, por lo menos uno de la pluralidad de fragmentos de ADNdc puede comprender un extremo romo. En algunas realizaciones, por lo menos uno de la pluralidad de fragmentos de ADNdc puede comprender un saliente 5' y/o un saliente 3'. En algunas realizaciones, ninguno de la pluralidad de fragmentos de ADNdc comprende un extremo romo.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, fragmentar el ADNdc puede comprender poner en contacto el ADNdc con una proteína asociada a CRISPR (por ejemplo, Cas9 o Cas12a) para generar la pluralidad de fragmentos de ADNdc. A modo de ejemplo, puede usarse un ARN guía complementario con un motivo o secuencia de ADN objetivo para dirigir la proteína asociada a CRISPR para generar roturas de ADN de cadena doble en el motivo o secuencia de ADN objetivo.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, generar la pluralidad de fragmentos de ácidos nucleicos comprende: agregar dos copias de un adaptador que comprende una

secuencia complementaria con una secuencia de captura a por lo menos uno de la pluralidad de fragmentos de ADNdc para generar una pluralidad de fragmentos de ADNdc. Por ejemplo, los adaptadores pueden agregarse mediante una transposasa como se describe en la presente. Por ejemplo, agregar las dos copias del adaptador puede comprender ligar las dos copias del adaptador a por lo menos uno de la pluralidad de fragmentos de ADNdc para generar la pluralidad de fragmentos de ADNdc que comprenden el adaptador.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, la secuencia de captura comprende una región poli(dT). La secuencia complementaria con la secuencia de captura puede comprender una región poli(dA).

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, fragmentar el ADNdc puede comprender poner en contacto el ADNdc con una enzima de restricción para generar la pluralidad de fragmentos de ADNdc, en donde por lo menos uno de la pluralidad de fragmentos de ADNdc comprende la secuencia de captura. La secuencia de captura puede ser complementaria a las secuencias de los salientes 5'. La secuencia complementaria a la secuencia de captura puede comprender la secuencia del saliente 5'. En algunas realizaciones, la secuencia de captura comprende una secuencia que no comprende tres, cuatro, cinco, seis o más T consecutivas. Por ejemplo, la secuencia de captura puede comprender una secuencia característica de una o ambas cadenas del ADNdc objetivo.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, el ADNdc está dentro de un orgánulo de la célula, por ejemplo, un núcleo. El método puede incluir permeabilizar un núcleo para generar un núcleo permeabilizado, por ejemplo usando un detergente como Triton X-100. El método puede incluir la fijación de una célula que comprende el núcleo antes de permeabilizar el núcleo. En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, el ADNdc está dentro de por lo menos uno de un núcleo, un nucléolo, una mitocondria o un cloroplasto. En algunas realizaciones, el ADNdc se selecciona del grupo que consiste en: ADN nuclear (por ejemplo, como parte de la cromatina), ADN nucleolar, ADN genómico, ADN mitocondrial, ADN de cloroplasto, ADN de constructo, ADN viral o una combinación de dos o más de los elementos enumerados. Los ejemplos de constructos de ADN pueden incluir plásmidos, vectores de clonación, vectores de expresión, vectores híbridos, minicírculos, cósmidos, vectores virales, BAC, YAC y HAC. A modo de ejemplo, el ADN viral puede insertarse en un genoma huésped, o estar presente en un ADN extragenómico. Por ejemplo, un método de análisis de muestras como se describe en la presente puede cuantificar el ADN o una clase de ADN en uno o más orgánulos de una célula. Por ejemplo, un método de análisis de muestras como se describe en la presente puede cuantificar el ADN viral o una carga viral de ADN en una célula. Por ejemplo, un método de análisis de muestras como se describe en la presente puede cuantificar el ADN del constructo en una célula (por ejemplo, plásmidos, vectores de clonación, vectores de expresión, vectores híbridos, minicírculos, cósmidos, vectores virales, BAC, YAC y/o HAC). Por tanto, se contempla que el método pueda proporcionar información acerca de las estructuras accesibles al transposoma que comprenden el ADNdc.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, el método comprende desnaturalizar la pluralidad de fragmentos de ácidos nucleicos para generar una pluralidad de fragmentos de ADNmc, en donde la codificación con códigos de barras de la pluralidad de fragmentos de ácidos nucleicos comprende codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando la pluralidad de códigos de barras para generar la pluralidad de fragmentos de ADNmc codificados con códigos de barras. En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, el adaptador comprende una secuencia promotora. La generación de la pluralidad de fragmentos de ácidos nucleicos puede comprender transcribir la pluralidad de fragmentos de ADNdc usando transcripción in vitro para generar una pluralidad de moléculas de ácido ribonucleico (ARN), y en donde la codificación con códigos de barras de la pluralidad de fragmentos de ácidos nucleicos comprende codificar con códigos de barras la pluralidad de moléculas de ARN. La secuencia promotora puede comprender una secuencia promotora T7.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, determinar la información referente al ADNdc (por ejemplo, ADNg) comprende determinar la accesibilidad a la cromatina del ADNdc (por ejemplo, ADNg) en función de las secuencias y/o la abundancia de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación obtenidos. Determinar la accesibilidad a la cromatina del ADNdc puede comprender: alinear las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras con una secuencia de referencia del ADNdc (por ejemplo, ADNg); identificar regiones del ADNdc correspondientes a los extremos de los fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) de la pluralidad de fragmentos de ADNmc con accesibilidad por encima de un umbral. Determinar la accesibilidad a la cromatina del ADNdc (por ejemplo, ADNg) puede comprender: alinear las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc) con una secuencia de referencia del ADNdc (por ejemplo, ADNg); y determinar la accesibilidad de las regiones del ADNdc (por ejemplo, ADNg) correspondientes a los extremos de los fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) de la pluralidad de fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) basándose en los números de fragmentos de ADN codificados con

códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) de la pluralidad de fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) en los datos de secuenciación.

5 En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, determinar la información relacionada con el ADNdc (por ejemplo, ADNg) comprende determinar la información del genoma del ADNdc en función de las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) en los datos de secuenciación obtenidos. El método de análisis de muestras puede comprender digerir nucleosomas asociados con el ADNdc.
10 Determinar la información del genoma del ADNdc puede comprender: determinar por lo menos una secuencia parcial del ADNdc alineando las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) con una secuencia de referencia del ADNdc.

15 En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, determinar la información que relaciona el ADNdc (por ejemplo, ADNg) con la estructura que comprende ADNdc comprende determinar la información del metiloma del ADNdc (por ejemplo, ADNg) en función de las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación obtenidos. El método de análisis de muestras puede comprender digerir nucleosomas asociados con el ADNdc. El método de
20 análisis de muestras puede comprender realizar la conversión con bisulfito de las bases de citosina de una pluralidad de fragmentos de ADN de cadena sencilla de la pluralidad de fragmentos de ADN salientes o de la pluralidad de fragmentos de ácidos nucleicos (por ejemplo, obtenidos mediante la desnaturalización de fragmentos de ADN salientes o de la pluralidad de fragmentos de fragmentos) para generar una pluralidad de ADNmc convertido con bisulfito con bases de uracilo. La codificación con códigos de barras de la pluralidad de fragmentos de ADN salientes o la codificación con códigos de barras de la pluralidad de fragmentos de ácidos nucleicos puede comprender
25 codificar con códigos de barras la pluralidad de ADNmc convertidos con bisulfito usando la pluralidad de códigos de barras para generar la pluralidad de fragmentos de ADNmc codificados con códigos de barras. Determinar la información del metiloma puede comprender: determinar una posición de la pluralidad de fragmentos de ADN codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras) en los datos de secuenciación que tiene una base de timina y la posición correspondiente en una secuencia de referencia del ADNdc que tiene una base de citosina para determinar la posición correspondiente en el ADNdc que tiene una base de metilcitosina.
30

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, la
35 codificación con códigos de barras comprende: codificar estocásticamente la pluralidad de fragmentos de ADN (por ejemplo, fragmentos de ADNmc) o la pluralidad de ácidos nucleicos usando la pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADN codificados con códigos de barras estocásticamente. La codificación con códigos de barras puede comprender: codificar con códigos de barras la pluralidad de fragmentos de ADN (por ejemplo, fragmentos de ADNmc) o la pluralidad de fragmentos de ácidos nucleicos usando la pluralidad de códigos
40 de barras asociados con una partícula para generar la pluralidad de fragmentos de ADNmc codificados con códigos de barras, en donde los códigos de barras asociados con la partícula comprenden un secuencia de marcado celular idéntica y por lo menos 100 secuencias de marcado molecular diferentes.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, puede
45 inmovilizarse sobre la partícula por lo menos un código de barras de la pluralidad de códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede inmovilizarse parcialmente en la partícula. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar encerrado en la partícula. Puede estar parcialmente encerrado en la partícula por lo menos un código de barras de la pluralidad de códigos de barras. La partícula puede ser disruptiva. La partícula puede comprender una partícula de hidrogel fragmentable. La
50 partícula puede comprender una perla de Sepharose, una perla de estreptavidina, una perla de agarosa, una perla magnética, una perla conjugada, una perla conjugada con proteína A, una perla conjugada con proteína G, una perla conjugada con proteína A/G, una perla conjugada con proteína L, una perla conjugada con oligo(dT), una perla de sílice, una perla similar a sílice, una microperla anti-biotina, una microesfera antifluorocromo, o cualquier combinación de las mismas. La partícula puede comprender un material seleccionado del grupo que consiste en
55 polidimetilsiloxano (PDMS), poliestireno, vidrio, polipropileno, agarosa, gelatina, hidrogel, paramagnético, cerámica, plástico, vidrio, metilestireno, polímero acrílico, titanio, látex, sefarsa, celulosa, nailon, silicona y cualquier combinación de los mismos. En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, por lo menos un código de barras de la pluralidad de códigos de barras puede dividirse de los otros códigos de barras. Se contempla que la división puede comprender, por ejemplo, disponer el código de barras en un
60 soporte sólido como una partícula como se describe en la presente, disponer el código de barras en una gotita (por ejemplo, una microgotita) como una gotita de hidrogel, o en un pocillo de un sustrato, como un micropocillo, o cámara de un dispositivo fluidoico (por ejemplo, un dispositivo microfluídico).

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, los
65 códigos de barras de la partícula pueden comprender marcadores moleculares con por lo menos 1000 secuencias

de marcadores moleculares diferentes. Los códigos de barras de la partícula pueden comprender marcadores moleculares con por lo menos 10000 secuencias de marcadores moleculares diferentes. Los marcadores moleculares de los códigos de barras pueden comprender secuencias aleatorias. La partícula puede comprender por lo menos 10000 códigos de barras.

En cualquiera de los métodos de análisis de células individuales descritos en la presente, la codificación con códigos de barras de la pluralidad de fragmentos de ADN salientes o la pluralidad de fragmentos de ácidos nucleicos puede comprender: poner en contacto una pluralidad de ADNmc (de los fragmentos de ADN o fragmentos de ácidos nucleicos) con la secuencia de captura de la pluralidad de códigos de barras; y transcribir la pluralidad de ADNmc usando la pluralidad de códigos de barras para generar la pluralidad de fragmentos de ADNmc codificados con códigos de barras. El método de análisis de muestras puede incluir: antes de obtener los datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras, amplificar la pluralidad de fragmentos de ADNmc codificados con códigos de barras para generar una pluralidad de fragmentos de ADN codificados con códigos de barras amplificados. La amplificación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras puede comprender: amplificar los fragmentos de ADNmc codificados con códigos de barras mediante la reacción en cadena de la polimerasa (PCR).

En algunas realizaciones, cualquier método de análisis de muestras descrito en la presente puede incluir: codificar con códigos de barras una pluralidad de objetivos del núcleo usando la pluralidad de códigos de barras para generar una pluralidad de objetivos codificados con códigos de barras; y obtener datos de secuenciación de los objetivos codificados con códigos de barras.

En algunas realizaciones, para cualquiera de los métodos de análisis de muestras descritos en la presente, el ADNdc de la célula se selecciona del grupo que consiste en: ADN nuclear, ADN nucleolar, ADN genómico, ADN mitocondrial, ADN de cloroplasto, ADN de constructo, ADN viral o una combinación de dos o más de los elementos enumerados. En algunas realizaciones, para cualquiera de los métodos de análisis de muestras descritos en la presente, los salientes 5' comprenden secuencias poli dT. En algunas realizaciones, para cualquiera de los métodos de análisis de muestras descritos en la presente, el método comprende además capturar un fragmento de ADNmc de la pluralidad de fragmentos de ADNs codificados con códigos de barras en una partícula que comprende un oligonucleótido que comprende la secuencia de captura, la secuencia del marcador celular y la secuencia del marcador molecular, en donde la secuencia de captura comprende una secuencia poli dT que se une a una cola poli A en el fragmento de ADNmc, dicho fragmento de ADNmc capturado comprendiendo una citidina metilada, realizando una reacción de conversión de bisulfuro en el fragmento de ADNmc para convertir la citidina metilada en una timidina, extendiendo el fragmento de ADNmc en la dirección de 5' a 3' para producir el fragmento de ADNmc codificado con códigos de barras que comprende la timidina, el ADNmc codificado con códigos de barras comprendiendo la secuencia de captura, la secuencia del marcador molecular y la secuencia del marcador celular, extendiendo el oligonucleótido en la dirección de 5' a 3' usando una transcriptasa inversa o polimerasa o una combinación de las mismas para producir una cadena de ADN complementaria al ADNmc codificado con códigos de barras que comprende la timidina, desnaturalizando el ADNmc codificado con códigos de barras y la cadena de ADN complementaria para producir secuencias de cadena sencilla y amplificando las secuencias de cadena sencilla. El método puede comprender además determinar si una posición de la pluralidad de fragmentos de ADNmc en los datos de secuenciación tiene una base de timina y la posición correspondiente en una secuencia de referencia del ADNdc tiene una base de citosina que comprende, después de la reacción de conversión de bisulfuro, determinar la posición correspondiente de la base de timina en la secuencia de referencia para ser una base de citosina.

En algunas realizaciones, para cualquiera de los métodos de análisis de muestras descritos en la presente, la nucleasa de cadena doble del transposoma se selecciona del grupo que consiste en una transposasa, una endonucleasa de restricción, una proteína asociada a CRISPR, una nucleasa específica de dúplex o una combinación de estas. En algunas realizaciones, para cualquiera de los métodos de análisis de muestras descritos en la presente, el transposoma comprende además un anticuerpo o fragmento del mismo, un aptámero o un dominio de unión al ADN que se une a la estructura que comprende el ADNdc. En algunas realizaciones, para cualquiera de los métodos de análisis de muestras descritos en la presente, el transposoma comprende además una ligasa.

En algunas realizaciones, se describe un reactivo de ácido nucleico. El reactivo de ácido nucleico puede comprender una secuencia de captura, un código de barras, un sitio de unión del cebador y un agente de unión al ADN de cadena doble. La secuencia de captura puede comprender una región poli(A). El sitio de unión al cebador puede comprender un sitio de unión al cebador universal. El reactivo de ácido nucleico puede ser impermeable a la membrana plasmática. En algunas realizaciones, el reactivo de ácido nucleico está configurado para unirse específicamente a las células muertas. En algunas realizaciones, el reactivo de ácido nucleico no se une a las células vivas.

En algunas realizaciones, para cualquiera de los métodos de análisis de muestras descritos en la presente, el método comprende además poner en contacto una célula con un reactivo de ácido nucleico. El reactivo de ácido nucleico puede ser como se describe en la presente. El reactivo de ácido nucleico puede comprender una secuencia de captura, un código de barras, un sitio de unión del cebador; y un agente de unión a ADN de cadena doble. La

célula puede ser una célula muerta y el reactivo de unión de ácido nucleico puede unirse al ADN de cadena doble en la célula muerta. El método puede comprender lavar la célula muerta para eliminar el exceso del reactivo de unión al ácido nucleico. El método puede comprender lisar la célula muerta. La lisis puede liberar el reactivo de unión al ácido nucleico. El método puede comprender codificar con códigos de barras el reactivo de unión al ácido nucleico. En el método de algunas realizaciones, la célula está asociada con un soporte sólido que comprende un oligonucleótido que comprende una secuencia de marcador celular, codificar con códigos de barras comprende codificar con códigos de barras el reactivo de unión al ácido nucleico con la secuencia de marcador celular. El soporte sólido puede comprender una pluralidad de oligonucleótidos, comprendiendo cada uno la secuencia de marcador celular y una secuencia de marcador molecular diferente. En algunas realizaciones, el método comprende además secuenciar los reactivos de unión de ácido nucleico codificados con códigos de barras y determinar la presencia de una célula muerta basándose en la presencia del código de barras del reactivo de ácido nucleico. En algunas realizaciones, el método comprende además asociar dos o más células, cada una con diferentes soportes sólidos que comprenden diferentes marcadores celulares, por lo que cada una de las dos o más células se asocia uno a uno con un marcador celular diferente. En algunas realizaciones, el método comprende además determinar un número de células muertas en la muestra basándose en el número de marcadores celulares únicos asociados con un código de barras de un reactivo de ácido nucleico. Determinar el número de secuencias de marcadores moleculares con secuencias distintas asociadas con el marcador celular y la secuencia de código de barras de control puede comprender determinar el número de secuencias de marcadores moleculares con el mayor número de secuencias distintas asociadas con el marcador celular y la secuencia de código de barras de control para cada marcador celular en los datos de secuenciación. En el método de algunas realizaciones, el reactivo de unión de ácido nucleico no se introduce en una célula viva y, por lo tanto, no se une al ADN de cadena doble en la célula viva. En algunas realizaciones, el método comprende además poner en contacto una célula muerta con un reactivo de unión a proteínas asociado con un oligonucleótido identificador único, en el que el reactivo de unión a proteínas se une a una proteína de la célula muerta; y codificar con códigos de barras el oligonucleótido identificador único. En el método de algunas realizaciones, el reactivo de unión a proteínas comprende un anticuerpo, un tetrámero, un aptámero, un andamiaje proteico, una invasina o una combinación de los mismos. En el método de algunas realizaciones, un objetivo proteico del reactivo de unión a proteínas se selecciona de un grupo que comprende 10-100 objetivos de proteínas diferentes, o se selecciona un objetivo de componente celular del reactivo de unión a componentes celulares de un grupo que comprende 10-100 objetivos de componentes de proteínas diferentes. En el método de algunas realizaciones, un objetivo de proteínas del reactivo de unión a proteínas comprende un carbohidrato, un lípido, una proteína, una proteína extracelular, una proteína de superficie celular, un marcador celular, un receptor de células B, un receptor de células T, un complejo mayor de histocompatibilidad, un antígeno tumoral, un receptor, una integrina, una proteína intracelular o cualquier combinación de los mismos. En el método de algunas realizaciones, el reactivo de unión a proteínas comprende un anticuerpo o fragmento del mismo que se une a una proteína de la superficie celular. En el método de algunas realizaciones, la codificación con códigos de barras es con un código de barras que comprende una secuencia de marcadores moleculares.

Se describe, pero no forma parte de la invención reivindicada, un método de análisis de muestras. El método puede comprender poner en contacto una célula muerta de una muestra con un reactivo de unión de ácido nucleico, un reactivo de unión de ácido nucleico que comprende una secuencia de captura, un código de barras, un sitio de unión de cebador y un agente de unión de ADN de cadena doble. El reactivo de unión nucleico puede unirse al ADN de cadena doble en la célula muerta. El método puede comprender lavar el exceso de reactivo de unión de ácido nucleico de la célula muerta. El método puede comprender lisar la célula muerta, liberando de este modo el reactivo de unión de ácido nucleico de la célula muerta. El método puede comprender codificar con códigos de barras el reactivo de unión de ácido nucleico. En el método de algunas realizaciones, el código de barras comprende capturar la célula muerta en un soporte sólido, como una perla, el soporte sólido comprendiendo una secuencia de marcador celular y una secuencia de marcador molecular. En algunas realizaciones, el método comprende además determinar una cantidad de secuencias de marcadores moleculares distintas asociadas con cada secuencia de marcador celular, y determinar una cantidad de células muertas en la muestra en función de la cantidad de secuencias de marcadores celulares distintas asociadas con secuencias de marcadores moleculares. En el método de algunas realizaciones, determinar el número de secuencias de marcadores moleculares con secuencias distintas asociadas con el marcador celular y la secuencia de código de barras de control comprende determinar el número de secuencias de marcadores moleculares con el mayor número de secuencias distintas asociadas con el marcador celular para cada marcador celular en los datos de secuenciación. En algunas realizaciones, el método comprende además poner en contacto una célula muerta con un reactivo de unión a proteínas asociado con un oligonucleótido identificador único. El reactivo de unión a proteínas puede unirse a una proteína de la célula muerta. El método puede comprender además codificar con códigos de barras el oligonucleótido identificador único. En el método de algunas realizaciones, el reactivo de unión a proteínas se asocia con dos o más oligonucleótidos de indexación de muestras con una secuencia idéntica. En el método de algunas realizaciones, el reactivo de unión a proteínas está asociado con dos o más oligonucleótidos de indexación de muestras con diferentes secuencias de indexación de muestras. En el método de algunas realizaciones, el reactivo de unión a proteínas comprende un anticuerpo, un tetrámero, un aptámero, un andamiaje proteico, una invasina o una combinación de los mismos. En el método de algunas realizaciones, un objetivo proteico del reactivo de unión a proteínas se selecciona de un grupo que comprende 10-100 objetivos proteicos diferentes, o en donde un objetivo de componente celular del reactivo de unión a componente celular se selecciona de un grupo que comprende 10-100 objetivos de componentes celulares

diferentes. En el método de algunas realizaciones, un objetivo proteico del reactivo de unión a proteínas comprende un carbohidrato, un lípido, una proteína, una proteína extracelular, una proteína de superficie celular, un marcador celular, un receptor de células B, un receptor de células T, un complejo mayor de histocompatibilidad, un antígeno tumoral, un receptor, una integrina, una proteína intracelular o cualquier combinación de los mismos. En el método de algunas realizaciones, el reactivo de unión a proteínas comprende un anticuerpo o fragmento del mismo que se une a una proteína de la superficie celular.

En el método de algunas realizaciones, la secuencia de captura y la secuencia complementaria a la secuencia de captura son un par específico de ácidos nucleicos complementarios de por lo menos 5 nucleótidos a aproximadamente 25 nucleótidos de longitud.

Se describe, pero no forma parte de la invención reivindicada, un método de análisis de muestras. El método puede comprender poner en contacto ácido desoxirribonucleico de cadena doble (ADNdc) de una célula con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5'. El método puede comprender poner en contacto la pluralidad de fragmentos de ADNdc salientes con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a por lo menos una parte de cada uno de los salientes 5'. El método puede comprender desnaturalizar la pluralidad de fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc). El método puede comprender codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de etiqueta molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden secuencias de marcadores moleculares diferentes, y en donde la pluralidad de códigos de barras comprende una secuencia de marcadores celulares idénticos. El método puede comprender obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras. El método puede comprender cuantificar una cantidad de ADNdc en la célula sobre la base de una cantidad de secuencias de marcadores moleculares únicas asociadas con la misma secuencia de marcador celular. En algunas realizaciones, el método comprende además capturar un fragmento de ADNmc de la pluralidad de fragmentos de ADNmc en un soporte sólido que comprende un oligonucleótido que comprende la secuencia de captura, la secuencia de marcador celular y la secuencia de marcador molecular, en donde la secuencia de captura comprende una secuencia poli dT que se une a una cola poli A en el fragmento de ADNmc; extender el fragmento de ADNmc en la dirección de 5' a 3' para producir el fragmento de ADNmc codificado con códigos de barras, el ADNmc codificado con códigos de barras comprendiendo la secuencia de captura, la secuencia de marcador molecular y la secuencia de marcador celular; extender el oligonucleótido en la dirección de 5' a 3' usando una transcriptasa inversa o polimerasa o una combinación de las mismas para producir una cadena de ADN complementaria que es complementaria al ADNmc codificado con códigos de barras; desnaturalizar el ADNmc codificado con códigos de barras y la cadena de ADN complementaria para producir secuencias de cadena sencilla; y amplificar las secuencias de cadena sencilla. En algunas realizaciones, el método comprende además una conversión con bisulfito de bases de citosina de la pluralidad de fragmentos de ADNmc para generar una pluralidad de fragmentos de ADNmc convertidos con bisulfito que comprenden bases de uracilo.

En cualquiera de los métodos descritos en la presente, el ADNdc puede comprender un ADN constructo. El ADN constructo puede seleccionarse del grupo que consiste en plásmidos, vectores de clonación, vectores de expresión, vectores híbridos, minicúculos, cósmidos, vectores virales, BAC, YAC y HAC. En algunas realizaciones, el número de constructos de ADN varía entre 1 y aproximadamente 1×10^6 .

En cualquiera de los métodos descritos en la presente, el ADNdc puede comprender ADN viral. La carga de ADN viral en la célula puede variar entre aproximadamente 1×10^2 - 1×10^6 .

Se describe, pero no forma parte de la invención reivindicada, un kit para el análisis de muestras. El kit puede comprender un transposoma como se describe en la presente y una pluralidad de códigos de barras como se describe en la presente. Cada transposoma puede comprender una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc (por ejemplo, una transposasa como se describe en la presente) y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura. Opcionalmente, el transposoma comprende además una ligasa. Cada código de barras puede comprender una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, por ejemplo, una secuencia poli T. Por lo menos dos de la pluralidad de códigos de barras comprenden secuencias de marcadores moleculares diferentes, y por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcadores celulares idénticos. Por ejemplo, los códigos de barras pueden comprender por lo menos 10, 50, 100, 500, 1000, 5000, 10000, 50000 o 100000 marcadores moleculares diferentes. Los códigos de barras pueden inmovilizarse en partículas como se describe en la presente. Todos los códigos de barras en la misma partícula pueden comprender el mismo marcador celular. En el kit de algunas realizaciones, los códigos de barras se

reparten en pocillos de un sustrato. Todos los códigos de barras repartidos en cada pocillo pueden comprender la misma secuencia de marcadores celulares, y en donde diferentes pocillos comprenden diferentes secuencias de marcadores celulares.

5 BREVE DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1 ilustra un código de barras ejemplar no limitativo.

La FIG. 2 muestra un flujo de trabajo ejemplar no limitativo de codificación con códigos de barras y recuento digital.

10 La FIG. 3 es una ilustración esquemática que muestra un proceso ejemplar no limitativo para generar una biblioteca indexada de los objetivos codificados con códigos de barras a partir de una pluralidad de objetivos.

Las FIGS. 4A-4B muestran una ilustración esquemática de métodos ejemplares no limitativos de captura de alto rendimiento de información multiómica de células individuales.

15 Las FIGS. 5A-5B ilustran esquemáticamente un método ejemplar no limitativo de captura de información de accesibilidad genómica y cromática de células individuales con intensidad de señal mejorada.

La FIG. 6 ilustra esquemáticamente un reactivo de ácido nucleico ejemplar no limitativo de algunas realizaciones.

DESCRIPCIÓN DETALLADA

20 En la siguiente descripción detallada, se hace referencia a los dibujos acompañantes, que forman parte de la misma. En los dibujos, los símbolos similares típicamente identifican componentes similares, a menos que el contexto indique lo contrario. No se pretende que las realizaciones ilustrativas descritas en la descripción detallada, los dibujos y las reivindicaciones sean limitativas. Se entenderá fácilmente que los aspectos de la presente divulgación, como se describen en general en la presente y se ilustran en las Figuras, pueden disponerse, sustituirse, combinarse, separarse y diseñarse en una amplia variedad de configuraciones diferentes, todas las
25 cuales se contemplan explícitamente en la presente y forman parte de la divulgación de la presente.

Los códigos de barras, como los códigos de barras estocásticos, con marcadores moleculares (también referidas como índices moleculares (MI)) que tienen diferencias de diferentes marcadores moleculares pueden
30 usarse para determinar la abundancia de objetivos de ácidos nucleicos, como la abundancia relativa o absoluta de los objetivos de ácidos nucleicos. La codificación con códigos de barras estocásticos puede realizarse usando el ensayo Precise™ (Cellular Research, Inc. (Palo Alto, CA)) y el ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ)). El ensayo Precise™ o el ensayo Rhapsody™, pueden utilizar un grupo no agotador de códigos de barras estocásticos con un gran número, por ejemplo de 6561 a 65536, secuencias de marcadores
35 moleculares únicos en oligonucleótidos poli(T) para hibridar con todos los ARNm poli(A) en una muestra durante el paso de transcripción inversa (RT). Un código de barras estocástico puede comprender un sitio de cebado de PCR universal. Durante la RT, las moléculas del gen objetivo reaccionan aleatoriamente con códigos de barras estocásticos. Cada molécula objetivo puede hibridarse con un código de barras estocástico que da como resultado generar moléculas de ácido ribonucleótido complementario (ADNc) codificadas con barras estocásticas). Después
40 del marcado, las moléculas de ADNc codificadas con barras estocásticas de los micropocillos de una placa de micropocillos pueden agruparse en un único tubo para la amplificación y secuenciación por PCR. Pueden analizarse los datos de secuenciación sin procesar para producir el número de lecturas, el número de códigos de barras estocásticos con secuencias de marcadores moleculares únicas, y el número de moléculas de ARNm.

45 En la presente se divulgan realizaciones de un método de análisis de muestras. Por ejemplo, cualquiera de los métodos de análisis de muestras descritos en la presente puede comprender, consistir en, o consistir esencialmente en análisis de células individuales. El método de análisis de muestras puede usarse para el análisis multiómico usando codificación con códigos de barras moleculares (como el ensayo Precise™ y el ensayo Rhapsody™. En algunas realizaciones, el método de análisis de muestras incluye: poner en contacto ácido
50 desoxirribonucleico de cadena doble (ADNdc) con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc, y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura para generar una pluralidad de fragmentos salientes de ADN de cadena doble (ADNdc), cada uno con dos copias de los salientes 5'. La nucleasa de cadena doble (por ejemplo, una transposasa) puede cargarse con las dos copias del
55 adaptador. El método puede comprender poner en contacto la pluralidad de fragmentos de ADNdc salientes (que comprenden los salientes 5') con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a por lo menos una parte del saliente 5'; desnaturalizar la pluralidad de fragmentos de ADNdc complementarios (cada uno de los cuales comprende la secuencia complementaria de por lo menos una parte del saliente 5') para generar una pluralidad de
60 fragmentos de ADN de cadena sencilla (ADNmc); codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden secuencias de marcadores moleculares diferentes, y
65 en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcadores

celulares idénticos; obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras; y determinar la información referente al ADNdc (por ejemplo, ADNg) sobre la base de las secuencias de la pluralidad de fragmentos de ADNmc en los datos de secuenciación obtenidos.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, un ADN de cadena doble puede comprender, consistir esencialmente en o consistir en cualquier ADN de cadena doble, por ejemplo, ADN genómico (ADNg), ADN de orgánulos (por ejemplo, ADN nuclear, ADN nucleolar, ADN genómico, ADN mitocondrial y ADN de cloroplasto), ADN viral y/o ADN de constructo (por ejemplo, plásmidos, vectores de clonación, vectores de expresión, vectores híbridos, minicírculos, cósmidos, vectores virales y/o cromosomas artificiales como BAC, YAC y HAC).

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, el ADN constructo se selecciona del grupo que consiste en plásmidos, vectores de clonación, vectores de expresión, vectores híbridos, minicírculos, cósmidos, vectores virales, BAC, YAC y HAC, o una combinación de dos o más de cualquiera de los elementos enumerados.

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, el número de ADN constructos varía de 1 a aproximadamente 1×10^6 .

En algunas realizaciones, para cualquier método de análisis de muestras descrito en la presente, una carga de ADN viral varía de aproximadamente 1×10^2 - 1×10^6 .

En reactivos de ácido nucleico y métodos de análisis de muestras como se describe en la presente pueden usarse varios reactivos de unión a ADN de cadena doble adecuados. En algunas realizaciones, para cualquier reactivo de ácido nucleico y/o método de análisis de muestras descrito en la presente, se selecciona un reactivo de unión a ácido de ADN de cadena doble, sin limitaciones, del grupo que consiste en antraciclinas (por ejemplo, aclarubicina, aldorrubicina, amrubicina, annamicina, ácido bohemico, carubicina, cosmomicina B, daunorrubicina, doxorrubicina, epirubicina, idarrubicina, menogaril, nogalamicina, pirarrubicina, sabarrubicina, valrubicina, zoptarelina doxorrubicina y zorrubicina), amikhelline, 9-aminoacridina, 7-aminoactinomicina D, amsacrina, dactinomicina, daunorrubicina, doxorrubicina, elipticina, bromuro de etidio, mitoxantrona, pirarrubicina, pixantrona, proflavina y psoraleno, o una combinación de dos o más de los elementos enumerados.

En algunas realizaciones, cualquiera de los métodos de análisis de muestras descritos en la presente incluye: generar una pluralidad de fragmentos de ácidos nucleicos a partir de ácido desoxirribonucleico de cadena doble (ADNdc) de una célula, en donde cada uno de la pluralidad de fragmentos de ácidos nucleicos comprende una secuencia de captura, un complemento de la secuencia de captura, un complemento inverso de la secuencia de captura o una combinación de los mismos; codificar con códigos de barras la pluralidad de fragmentos de ácidos nucleicos usando la pluralidad de códigos de barras para generar una pluralidad de fragmentos de ácido desoxirribonucleico de cadena sencilla (ADNmc) codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden diferentes secuencias de marcadores moleculares, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcador celular idéntica; obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras; y determinar información referente al ADNdc basándose en las secuencias de la pluralidad de fragmentos de ADNmc en los datos de secuenciación obtenidos.

A menos que se defina de otro modo, los términos técnicos y científicos usados en la presente tienen el mismo significado que el entendido comúnmente por los expertos en la técnica a la que pertenece la presente divulgación. Consultar, por ejemplo, Singleton et al., Dictionary of Microbiology and Molecular Biology 2^a ed., J. Wiley & Sons (Nueva York, NY 1994); Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY 1989). A los efectos de la presente divulgación, a continuación se proporciona información sobre los siguientes términos.

Como se usa en la presente, el término "adaptador" tiene su significado habitual y ordinario en la técnica a la vista de esta memoria descriptiva. Se refiere a una secuencia para facilitar la amplificación, secuenciación y/o captura de ácidos nucleicos asociados. Los ácidos nucleicos asociados pueden comprender ácidos nucleicos objetivo. Los ácidos nucleicos asociados pueden comprender uno o más marcadores espaciales, marcadores objetivo, marcadores de muestra, marcadores de indexación o secuencias de código de barras (por ejemplo, marcadores moleculares). Los adaptadores pueden ser lineales. Los adaptadores pueden ser adaptadores preadenilados. Los adaptadores pueden ser de cadena doble o sencilla. Uno o más adaptadores pueden localizarse en el extremo 5' o 3' de un ácido nucleico. Cuando los adaptadores comprenden secuencias conocidas en los extremos 5' y 3', las secuencias conocidas pueden ser secuencias iguales o diferentes. Un adaptador localizado en el extremo 5' y/o 3' de un polinucleótido puede ser capaces de hibridar con uno o más oligonucleótidos inmovilizados en una superficie. Un adaptador puede, en algunas realizaciones, comprender una secuencia universal. Una secuencia universal puede ser una región de secuencia de nucleótidos que es común para dos o más moléculas de

ácidos nucleicos. Las dos o más moléculas de ácidos nucleicos también pueden tener regiones de diferente secuencia. Por tanto, por ejemplo, los adaptadores 5' pueden comprender secuencias de ácidos nucleicos idénticas y/o universales y los adaptadores 3' pueden comprender secuencias idénticas y/o universales. Una secuencia universal que puede estar presente en diferentes miembros de una pluralidad de moléculas de ácidos nucleicos puede permitir la replicación o amplificación de múltiples secuencias diferentes usando un único cebador universal que es complementario a la secuencia universal. De manera similar, por lo menos uno, dos (por ejemplo, un par) o más secuencias universales que pueden estar presentes en diferentes miembros de una colección de moléculas de ácidos nucleicos pueden permitir la replicación o amplificación de múltiples secuencias diferentes usando por lo menos uno, dos (por ejemplo, un par) o más cebadores universales únicos que son complementarios a las secuencias universales. Por tanto, un cebador universal incluye una secuencia que puede hibridar con dicha secuencia universal. Las moléculas portadoras de las secuencias de ácidos nucleicos objetivo pueden modificarse para unir adaptadores universales (por ejemplo, secuencias de ácidos nucleicos no objetivo) a uno o ambos extremos de las diferentes secuencias de ácidos nucleicos objetivo. El uno o más cebadores universales unidos al ácido nucleico objetivo pueden proporcionar sitios para la hibridación de cebadores universales. El uno o más cebadores universales unidos al ácido nucleico objetivo pueden ser iguales o diferentes entre sí.

Como se usa en la presente, el término "asociado" o "asociado con" tiene su significado habitual y ordinario en la técnica a la vista de esta memoria descriptiva. Puede referirse a dos o más especies que son identificables como colocadas en un punto temporal. Una asociación puede referirse a dos o más especies que están o estuvieron dentro de un recipiente similar. Una asociación puede referirse a una asociación informática. Por ejemplo, puede almacenarse la información digital referente a dos o más especies y puede usarse para determinar que una o más de las especies estaban colocadas en un punto temporal. Una asociación también puede referirse a una asociación física. En algunas realizaciones, dos o más especies asociadas están "atadas", "unidas" o "inmovilizadas" entre sí o a una superficie sólida o semisólida común. Una asociación puede referirse a medios covalentes o no covalentes para unir marcadores a soportes sólidos o semisólidos, como perlas. Una asociación puede referirse a un enlace covalente entre un objetivo y un marcador. Una asociación puede comprender la hibridación entre dos moléculas (como una molécula objetivo y un marcador).

Como se usa en la presente, el término "complementario" tiene su significado habitual y ordinario en la técnica a la vista de esta memoria descriptiva. Puede referirse a la capacidad de emparejamiento preciso entre dos nucleótidos. Por ejemplo, si un nucleótido en una posición dada de un ácido nucleico es capaz de formar enlaces de hidrógeno con un nucleótido de otro ácido nucleico, entonces los dos ácidos nucleicos se consideran complementarios entre sí en esa posición. La complementariedad entre dos moléculas de ácidos nucleicos de cadena sencilla puede ser "parcial", en la que sólo se unen algunos de los nucleótidos, o puede ser completa cuando existe una complementariedad total entre las moléculas de cadena sencilla. Puede decirse que una primera secuencia de nucleótidos es el "complemento" de una segunda secuencia si la primera secuencia de nucleótidos es complementaria a la segunda secuencia de nucleótidos. Puede decirse que una primera secuencia de nucleótidos es el "complemento inverso" de una segunda secuencia, si la primera secuencia de nucleótidos es complementaria a una secuencia que es inversa (es decir, el orden de los nucleótidos se invierte) de la segunda secuencia. Como se usa en la presente, una secuencia "complementaria" puede referirse a un "complemento" o un "complemento inverso" de una secuencia. Se entiende por la divulgación que, si una molécula puede hibridarse con otra molécula, puede ser complementaria, o parcialmente complementaria, a la molécula que está hibridando.

Como se usa en la presente, el término "recuento digital" puede referirse a un método para estimar un número de moléculas objetivo en una muestra. El recuento digital puede incluir el paso de determinar una cantidad de marcadores únicos que se han asociado con objetivos en una muestra. Esta metodología, que puede ser de naturaleza estocástica, transforma el problema de contar moléculas de localizar e identificar moléculas idénticas a una serie de preguntas digitales de sí/no relacionadas con la detección de un conjunto de marcadores predefinidos.

Como se usa en la presente, el término "marcador" o "marcadores" tienen sus significados habituales y ordinarios en la técnica a la vista de esta memoria descriptiva. Pueden referirse a códigos de ácidos nucleicos asociados con un objetivo dentro de una muestra. Un marcador puede comprender, consistir esencialmente en, o consistir en, por ejemplo, un marcador de ácido nucleico. Un marcador puede ser un marcador total o parcialmente amplificable. Un marcador puede ser un marcador total o parcialmente secuenciable. Un marcador puede ser una porción de un ácido nucleico nativo que es identificable como distinto. Un marcador puede comprender, consistir esencialmente en, o consistir en una secuencia conocida. Un marcador puede comprender una unión de secuencias de ácidos nucleicos, por ejemplo, una unión de una secuencia nativa y no nativa. Como se usa en la presente, el término "marcador" puede usarse indistintamente con los términos "índice", "etiqueta" o "marcador-etiqueta". Los marcadores pueden transportar información. Por ejemplo, en varias realizaciones, los marcadores pueden usarse para determinar la identidad de una muestra, una fuente de una muestra, una identidad de una célula y/o un objetivo.

Como se usa en la presente, el término "depósitos que no se agotan" puede referirse a un grupo de códigos de barras (por ejemplo, códigos de barras estocásticos) compuesto por muchos marcadores diferentes. Un depósito que no se agota puede comprender un gran número de códigos de barras diferentes, de tal manera que cuando el depósito que no se agota está asociado con un grupo de objetivos, es probable que cada objetivo esté asociado con

un código de barras único. La singularidad de cada molécula objetivo marcada puede determinarse mediante las estadísticas de elección aleatoria y depende del número de copias de moléculas objetivo idénticas en la colección en comparación con la diversidad de marcadores. El tamaño del conjunto resultante de moléculas objetivo marcadas puede determinarse por la naturaleza estocástica del proceso de codificación con códigos de barras, y el análisis del número de códigos de barras detectados permite luego calcular el número de moléculas objetivo presentes en la colección o muestra original. Cuando la relación entre el número de copias de una molécula objetivo presente y el número de códigos de barras únicos es baja, las moléculas objetivo marcadas son altamente únicas (es decir, hay una probabilidad muy baja de que más de una molécula objetivo haya sido marcada con un marcador dado).

Como se usa en la presente, el término "ácido nucleico" tiene su significado habitual y ordinario en la técnica a la vista de esta memoria descriptiva. Se refiere a una secuencia de polinucleótidos, o a un fragmento de la misma. Un ácido nucleico puede comprender, consistir esencialmente en, o consistir en nucleótidos. Un ácido nucleico puede ser exógeno o endógeno a una célula. Un ácido nucleico puede existir en un entorno libre de células. Un ácido nucleico puede comprender, consistir esencialmente en o consistir en un gen o un fragmento del mismo. Un ácido nucleico puede comprender, consistir esencialmente en o consistir en ADN. Un ácido nucleico puede comprender, consistir esencialmente en o consistir en ARN. Un ácido nucleico puede comprender, consistir esencialmente en, o consistir en uno o más análogos (por ejemplo, estructura principal alterada, azúcar o nucleobase). Algunos ejemplos no limitativos de análogos incluyen: 5-bromouracilo, ácido nucleico peptídico, ácido xenonucleico, morfolinós, ácidos nucleicos bloqueados, ácidos nucleicos de glicol, ácidos nucleicos de treosa, dideoxinucleótidos, cordicepina, 7-deaza-GTP, fluoróforos (por ejemplo, rodamina o fluoresceína enlazada a azúcar), nucleótidos que contienen tiol, nucleótidos enlazados a biotina, análogos de bases fluorescentes, islas CpG, metil-7-guanosina, nucleótidos metilados, inosina, tiouridina, pseudouridina, dihidrouridina, queuosina y wyosina. "Ácido nucleico", "polinucleótido", "polinucleótido objetivo" y "ácido nucleico objetivo" pueden usarse indistintamente.

Un ácido nucleico puede comprender una o más modificaciones (por ejemplo, una modificación de la base, una modificación de la estructura principal), para proporcionar al ácido nucleico una característica nueva o mejorada (por ejemplo, una estabilidad mejorada). Un ácido nucleico puede comprender una etiqueta de afinidad de ácidos nucleicos. Un nucleósido puede ser una combinación de base-azúcar. La porción de base del nucleósido puede ser una base heterocíclica. Las dos clases más comunes de tales bases heterocíclicas son las purinas y las pirimidinas. Los nucleótidos pueden ser nucleósidos que además incluyen un grupo fosfato enlazado covalentemente a la porción de azúcar del nucleósido. Para aquellos nucleósidos que incluyen un azúcar de pentofuranosilo, el grupo fosfato puede enlazarse a la fracción hidroxilo 2', 3' o 5' del azúcar. En la formación de ácidos nucleicos, los grupos fosfato pueden enlazar covalentemente nucleósidos adyacentes entre sí para formar un compuesto polimérico lineal. A su vez, los extremos respectivos de este compuesto polimérico lineal pueden unirse adicionalmente para formar un compuesto circular; sin embargo, los compuestos lineales son generalmente adecuados. Además, los compuestos lineales pueden tener complementariedad interna de bases de nucleótidos y, por lo tanto, pueden plegarse de manera que produzcan un compuesto total o parcialmente de cadena doble. Dentro de los ácidos nucleicos, los grupos fosfato pueden denominarse comúnmente como formadores de la estructura principal internucleosídica del ácido nucleico. El enlace o estructura principal puede ser un enlace fosfodiéster de 3' a 5'.

Un ácido nucleico puede comprender una estructura principal modificada y/o enlaces internucleosídicos modificados. Las estructuras principales modificadas pueden incluir aquellas que retienen un átomo de fósforo en la estructura principal y aquellos que no tienen un átomo de fósforo en la estructura principal. Las estructuras principales de ácidos nucleicos modificadas adecuadas que contienen un átomo de fósforo pueden incluir, por ejemplo, fosforotioatos, fosforotioatos quirales, fosforoditioatos, fosfotriésteres, fosfotriésteres de aminoalquilo, metilo y otros fosfonatos de alquilo como fosfonatos de 3'-alquileo, fosfonatos de 5'-alquileo, fosfonatos quirales, fosfinatos, fosforamidatos que incluyen 3'-amino fosforamidato y fosforamidatos de aminoalquilo, fosforodiamidatos, tionofosforamidatos, tionoalquilfosfonatos, tionoalquilfosfotriésteres, selenofosfatos y boranofosfatos que tienen enlaces 3' 5' normales, análogos enlazados 2' 5', y los que tienen polaridad invertida en donde uno o más enlaces internucleotídicos son un enlace de 3' a 3', de 5' a 5' o de 2' a 2'.

Un ácido nucleico puede comprender estructuras principales de polinucleótidos que están formadas por enlaces internucleosídicos de alquilo o cicloalquilo de cadena corta, enlaces internucleosídicos mixtos de heteroátomos y alquilo o cicloalquilo, o uno o más enlaces internucleosídicos heteroatómicos o heterocíclicos de cadena corta. Estos pueden incluir aquellos que tienen enlaces morfolino (formados en parte a partir de la porción de azúcar de un nucleósido); estructuras principales de siloxano; estructuras principales de sulfuro, sulfóxido y sulfona; estructuras principales de formacetilo y tioformacetilo; estructuras principales de metileno formacetilo y tioformacetilo; estructuras principales de riboacetilo; estructuras principales que contienen alquenos; estructuras principales de sulfamato; estructuras principales de metilenoimino y metilenoimidazino; estructuras principales de sulfonato y sulfonamida; estructuras principales de amida; y otros que tienen partes componentes mixtas de N, O, S y CH₂.

Un ácido nucleico puede comprender, consistir esencialmente en, o consistir en un mimético de ácido nucleico. Se pretende que el término "mimético" incluya polinucleótidos en los que solo el anillo de furanosa o tanto el anillo de furanosa como el enlace internucleotídico se reemplazan con grupos que no son furanosa, al reemplazo

del anillo de furanosa también puede hacerse referencia como un sustituto de azúcar. La fracción de base heterocíclica o una fracción de base heterocíclica modificada puede mantenerse para la hibridación con un ácido nucleico objetivo apropiado. Uno de tales ácidos nucleicos puede ser un ácido nucleico peptídico (PNA). En un PNA, la estructura principal de azúcar de un polinucleótido puede reemplazarse con una estructura principal que contiene amida, en particular, una estructura principal de aminoetilglicina. Los nucleótidos pueden retenerse y se unen directa o indirectamente a los átomos de nitrógeno aza de la porción amida de la estructura principal. La estructura principal en los compuestos de PNA puede comprender dos o más unidades de aminoetilglicina enlazadas que dan al PNA una estructura principal que contiene amida. Las fracciones de base heterocíclicas pueden unirse directa o indirectamente a átomos de nitrógeno aza de la porción de amida de la estructura principal.

Un ácido nucleico puede comprender, consistir esencialmente en, o consistir en una estructura de estructura principal de morfolino. Por ejemplo, un ácido nucleico puede comprender un anillo de morfolino de 6 miembros en lugar de un anillo de ribosa. En algunas de estas realizaciones, un enlace internucleosídico fosfordiamidato u otro no fosfodiéster puede reemplazar un enlace fosfodiéster.

Un ácido nucleico puede comprender, consistir esencialmente en, o consistir en unidades de morfolino enlazadas (por ejemplo, ácido nucleico de morfolino) que tienen bases heterocíclicas unidas al anillo de morfolino. Los grupos de enlace pueden enlazar las unidades monoméricas de morfolino en un ácido nucleico de morfolino. Los compuestos oligoméricos basados en morfolino no iónicos pueden tener menos interacciones no deseadas con las proteínas celulares. Los polinucleótidos basados en morfolino pueden ser miméticos no iónicos de ácidos nucleicos. Pueden unirse una variedad de compuestos dentro de la clase de morfolino usando diferentes grupos de enlace. A una clase adicional de miméticos de polinucleótidos puede hacerse referencia como ácidos nucleicos de ciclohexenilo (CeNA). El anillo de furanosa normalmente presente en una molécula de ácido nucleico puede reemplazarse con un anillo de ciclohexenilo. Pueden prepararse monómeros de fosforamidita protegidos con CeNA DMT y usarse para la síntesis de compuestos oligoméricos usando química de fosforamidita. La incorporación de monómeros de CeNA en una cadena de ácidos nucleicos puede aumentar la estabilidad de un híbrido de ADN/ARN. Los oligoadenilatos de CeNA pueden formar complejos con complementos de ácidos nucleicos con una estabilidad similar a la de los complejos nativos. Una modificación adicional puede incluir ácidos nucleicos bloqueados (LNA) en los que el grupo 2'-hidroxilo está enlazado al átomo de carbono 4' del anillo de azúcar formando de este modo un enlace 2'-C, 4'-C-oximetileno formando de este modo una fracción de azúcar bicíclica. El enlace puede ser un metileno (-CH₂), grupo que forma puentes entre el átomo de oxígeno 2' y el átomo de carbono 4' en donde n es 1 o 2. Los LNA y los análogos de LNA pueden mostrar estabilidades térmicas dúplex muy altas con ácido nucleico complementario (Tm=+3 a +10° C), estabilidad frente a la degradación 3'-exonucleolítica y buenas propiedades de solubilidad.

Un ácido nucleico también puede incluir modificaciones o sustituciones de nucleobases (a menudo denominadas simplemente "bases"). Como se usa en la presente, las nucleobases "no modificadas" o "naturales" pueden incluir las bases de purina (por ejemplo, adenina (A) y guanina (G)), y las bases de pirimidina (por ejemplo, timina (T), citosina (C) y uracilo (U)). Las nucleobases modificadas pueden incluir otras nucleobases sintéticas y naturales como 5-metilcitosina (5-me-C), 5-hidroximetilcitosina, xantina, hipoxantina, 2-aminoadenina, 6-metilo y otros derivados alquílicos de adenina y guanina, 2-propilo y otros derivados alquílicos de adenina y guanina, 2-tiouracilo, 2-tiotimina y 2-tiocitosina, 5-halouracilo y citosina, 5-propinilo (-C≡C-CH₃) uracilo y citosina y otros derivados alquílicos de bases pirimidínicas, 6-azo uracilo, citosina y timina, 5-uracilo (pseudouracilo), 4-tiouracilo, 8-halo, 8-amino, 8-tiol, 8-tioalquilo, 8-hidroxilo y otras adeninas y guaninas 8-sustituidas, 5-halo, particularmente 5-bromo, 5-trifluorometilo y otros uracilos y citosinas 5-sustituidos, 7-metilguanina y 7-metiladenina, 2-F-adenina, 2-aminoadenina, 8-azaguanina y 8-azaadenina, 7-deazaguanina y 7-deazaadenina y 3-deazaguanina y 3-deazaadenina. Las nucleobases modificadas pueden incluir pirimidinas tricíclicas como fenoxazina citidina (1H-pirimido(5,4-b)(1,4)benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido(5,4-b)(1,4)benzotiazina-2(3H)-ona), abrazaderas G como una citidina de fenoxazina sustituida (por ejemplo, 9-(2-aminoetoxi)-H-pirimido(5,4-b)(1,4)benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido(5,4-b)(1,4)benzotiazin-2(3H)-ona), abrazaderas G como una citidina de fenotiazina sustituida (por ejemplo, 9-(2-aminoetoxi)-H-pirimido(5,4-b)(1,4)benzoxazin-2(3H)-ona), carbazol citidina (2H-pirimido(4,5-b)indol-2-ona), piridoindol citidina (H-pirido(3',2':4,5)pirrolo[2,3-d]pirimidin-2-ona).

Como se usa en la presente, el término "muestra" puede referirse a una composición que comprende objetivos. Las muestras adecuadas para el análisis mediante los métodos, dispositivos y sistemas divulgados incluyen células, tejidos, órganos u organismos. En algunas realizaciones, la muestra comprende, consiste esencialmente en o consiste en una única célula. En algunas realizaciones, la muestra comprende, consiste esencialmente en, o consiste en por lo menos 100.000, 200.000, 300.000, 500.000, 800.000 o 1.000.000 de células individuales.

Como se usa en la presente, el término "dispositivo de muestreo" o "dispositivo" puede referirse a un dispositivo que puede tomar una sección de una muestra y/o colocar la sección sobre un sustrato. Un dispositivo de muestra puede referirse, por ejemplo, a una máquina de clasificación de células activadas por fluorescencia (FACS), una máquina de clasificación celular, una aguja de biopsia, un dispositivo de biopsia, un dispositivo de corte de tejido, un dispositivo microfluidico, una rejilla de láminas y/o un micrótopo.

Como se usa en la presente, el término "soporte sólido" tiene su significado habitual y ordinario en la técnica a la vista de esta memoria descriptiva. Puede referirse a superficies discretas sólidas o semisólidas a las que pueden unirse una pluralidad de códigos de barras (por ejemplo, códigos de barras estocásticos). Un soporte sólido puede abarcar cualquier tipo de esfera, bola, cojinete, cilindro u otra configuración similar sólida, porosa o hueca compuesta de material plástico, cerámico, metálico o polimérico (por ejemplo, hidrogel) sobre el cual puede inmovilizarse un ácido nucleico (por ejemplo, covalentemente o no covalentemente). Un soporte sólido puede comprender una partícula discreta que puede ser esférica (por ejemplo, microesferas) o tener una forma no esférica o irregular, como cúbica, cuboide, piramidal, cilíndrica, cónica, oblonga o con forma de disco, y similares. Una perla puede tener forma no esférica. Una pluralidad de soportes sólidos espaciados en una matriz puede no comprender un sustrato. Un soporte sólido puede usarse indistintamente con el término "perla". Se contempla que para cualquier realización de la presente en la que el código de barras esté inmovilizado en un soporte sólido, partícula, perla o similar, el código de barras también puede dividirse, por ejemplo, en una gotita (por ejemplo, una microgotita) como una gotita de hidrogel, o en un pocillo de un sustrato, como un micropocillo, o cámara de un dispositivo fluídico (por ejemplo, un dispositivo microfluídico). Por consiguiente, siempre que se divulgue en la presente la agrupación, clasificación o división de ácidos nucleicos por medio de un "soporte sólido" (por ejemplo, una perla), también se contempla expresamente la división en un fluido (por ejemplo, una gotita, como una microgotita) o espacio físico, por ejemplo un micropocillo (por ejemplo, en una placa de múltiples pocillos) o una cámara (por ejemplo, en un dispositivo fluídico).

Como se usa en la presente, el término "código de barras estocástico" puede referirse a una secuencia de polinucleótidos que comprende marcadores de la presente divulgación. Un código de barras estocástico puede ser una secuencia de polinucleótidos que puede usarse para la codificación con códigos de barras estocásticos. Los códigos de barras estocásticos pueden usarse para cuantificar objetivos dentro de una muestra. Los códigos de barras estocásticos pueden usarse para controlar los errores que pueden producirse después de asociar un marcador con un objetivo. Por ejemplo, un código de barras estocástico puede usarse para evaluar errores de amplificación o de secuenciación. Un código de barras estocástico asociado con un objetivo puede denominarse objetivo de código de barras estocástico u objetivo código de barras-etiqueta estocástico.

Como se usa en la presente, el término "código de barras estocástico específico de gen" puede referirse a una secuencia de polinucleótidos que comprende marcadores y una región de unión a objetivo que es específica del gen. Un código de barras estocástico puede ser una secuencia de polinucleótidos que puede usarse para la codificación con códigos de barras estocásticos. Los códigos de barras estocásticos pueden usarse para cuantificar objetivos dentro de una muestra. Los códigos de barras estocásticos pueden usarse para controlar los errores que pueden producirse después de asociar un marcador con un objetivo. Por ejemplo, un código de barras estocástico puede usarse para evaluar los errores de amplificación o de secuenciación. Un código de barras estocástico asociado con un objetivo puede denominarse objetivo de código de barras estocástico u objetivo código de barras-etiqueta estocástico.

Como se usa en la presente, el término "codificación con códigos de barras estocásticos" puede referirse al marcado aleatorio (por ejemplo, codificación con códigos de barras) de ácidos nucleicos. La codificación con códigos de barras estocásticos puede utilizar una estrategia de Poisson recursiva para asociar y cuantificar los marcadores asociados con los objetivos. Como se usa en la presente, el término "codificación con códigos de barras estocásticos" puede usarse indistintamente con "marcado estocástico".

Como se usa en la presente, el término "objetivo" tiene su significado habitual y ordinario en la técnica a la vista de esta memoria descriptiva. Puede referirse a una composición que puede asociarse con un código de barras (por ejemplo, un código de barras estocástico). Ejemplos de objetivos adecuados para el análisis mediante los métodos, dispositivos y sistemas divulgados incluyen oligonucleótidos, ADN, ARN, ARNm, microARN, ARNt y similares. Los objetivos pueden ser de cadena sencilla o de cadena doble. En algunas realizaciones, los objetivos pueden ser proteínas, péptidos o polipéptidos. En algunas realizaciones, los objetivos son lípidos. Como se usa en la presente, "objetivo" puede usarse indistintamente con "especie".

Como se usa en la presente, el término "transcriptasas inversas" tiene su significado habitual y ordinario en la técnica a la vista de esta memoria descriptiva. Puede referirse a un grupo de enzimas que tienen actividad de transcriptasa inversa (es decir, que catalizan la síntesis de ADN a partir de una plantilla de ARN). En general, tales enzimas incluyen, pero no se limitan a, transcriptasa inversa retroviral, transcriptasa inversa de retrotransposones, transcriptasas inversas de retroplásmidos, transcriptasas inversas de retrones, transcriptasas inversas bacterianas, transcriptasa inversa derivada de intrones del grupo II y mutantes, variantes o derivados de las mismas. Las transcriptasas inversas no retrovirales incluyen transcriptasas inversas de retrotransposones no LTR, transcriptasas inversas de retroplásmidos, transcriptasas inversas de retrones y transcriptasas inversas de intrones del grupo II. Los ejemplos de transcriptasas inversas de intrones del grupo II incluyen la transcriptasa inversa del intrón LI.LtrB de *Lactococcus lactis*, la transcriptasa inversa del intrón Tel4c de *Thermosynechococcus* o la transcriptasa inversa del intrón Gsl-IIC de *Geobacillus stearothermophilus*. Otras clases de transcriptasas inversas pueden incluir muchas clases de transcriptasas inversas no retrovirales (es decir, retrones, intrones del grupo II y retroelementos

generadores de diversidad, entre otros).

Los términos "cebador de adaptador universal", "adaptador cebador universal" o "secuencia adaptadora universal" se usan indistintamente para referirse a una secuencia de nucleótidos que puede usarse para hibridar con códigos de barras (por ejemplo, códigos de barras estocásticos) para generar códigos de barras específicos de genes. Una secuencia adaptadora universal puede, por ejemplo, ser una secuencia conocida que sea universal en todos los códigos de barras usados en los métodos de la divulgación. Por ejemplo, cuando se marcan múltiples objetivos usando los métodos divulgados en la presente, cada una de las secuencias específicas del objetivo puede estar enlazada a la misma secuencia adaptadora universal. En algunas realizaciones, en los métodos divulgados en la presente puede usarse más de una secuencia adaptadora universal. Por ejemplo, cuando se están marcando múltiples objetivos usando los métodos divulgados en la presente, por lo menos dos de las secuencias específicas del objetivo están enlazadas a diferentes secuencias adaptadoras universales. Un cebador de adaptador universal y su complemento pueden incluirse en dos oligonucleótidos, uno de los cuales comprende una secuencia específica del objetivo y el otro comprende un código de barras. Por ejemplo, una secuencia adaptadora universal puede ser parte de un oligonucleótido que comprende una secuencia específica del objetivo para generar una secuencia de nucleótidos que es complementaria a un ácido nucleico objetivo. Un segundo oligonucleótido que comprende un código de barras y una secuencia complementaria de la secuencia adaptadora universal puede hibridar con la secuencia de nucleótidos y generar un código de barras específico del objetivo (por ejemplo, un código de barras estocástico específico del objetivo). En algunas realizaciones, un cebador de adaptador universal tiene una secuencia que es diferente de un cebador de PCR universal usado en los métodos de esta divulgación.

Códigos de barras

La codificación con códigos de barras, como la codificación con códigos de barras estocásticos, se ha descrito, por ejemplo, en la US 2015/0299784, la WO 2015/031691, y Fu et al, Proc Natl Acad Sci U.S.A 31 de mayo de 2011; 108 (22): 9026-31. En algunas realizaciones, el código de barras divulgado en la presente puede ser un código de barras estocástico que puede ser una secuencia de polinucleótidos que puede usarse para marcar estocásticamente (por ejemplo, codificar con códigos de barras, etiquetar) un objetivo. A los códigos de barras puede hacerse referencia como códigos de barras estocásticos si la proporción del número de secuencias de códigos de barras diferentes de los códigos de barras estocásticos y el número de ocurrencias de cualquiera de los objetivos a marcar puede ser, o ser aproximadamente, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, o un número o un intervalo entre dos cualquiera de estos valores. Un objetivo puede ser una especie de ARNm que comprende moléculas de ARNm con secuencias idénticas o casi idénticas. A los códigos de barras puede hacerse referencia como códigos de barras estocásticos si la proporción del número de secuencias de códigos de barras diferentes de los códigos de barras estocásticos y el número de ocurrencias de cualquiera de los objetivos a marcar es por lo menos, o como máximo, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, o 100:1. Las secuencias de códigos de barras de códigos de barras estocásticos puede hacerse referencia como marcadores moleculares.

Un código de barras, por ejemplo un código de barras estocástico, puede comprender uno o más marcadores. Los marcadores ejemplares pueden incluir un marcador universal, un marcador celular, una secuencia de código de barras (por ejemplo, un marcador molecular), un marcador de muestra, un marcador de placa, un marcador espacial y/o un marcador preespacial. La FIG. 1 ilustra un código de barras ejemplar 104 con un marcador espacial. El código de barras 104 puede comprender una amina 5' que puede enlazar el código de barras a un soporte sólido 105. El código de barras puede comprender un marcador universal, un marcador de dimensión, un marcador espacial, un marcador celular y/o un marcador molecular. El código de barras puede comprender un marcador universal, un marcador celular y un marcador molecular. El código de barras puede comprender un marcador universal, un marcador espacial, un marcador celular y un marcador molecular. El código de barras puede comprender un marcador universal, un marcador espacial, un marcador celular y un marcador molecular. El código de barras puede comprender un marcador universal, un marcador espacial, un marcador celular y un marcador molecular. El orden de los diferentes marcadores (incluyendo, pero no limitados a, el marcador universal, el marcador de dimensión, el marcador espacial, el marcador celular y/o el marcador molecular) en el código de barras puede variar. Por ejemplo, como se muestra en la FIG. 1, el marcador universal puede ser el marcador más 5' y el marcador molecular puede ser el marcador más 3'. El marcador espacial, el marcador de dimensión y el marcador celular pueden estar en cualquier orden. En algunas realizaciones, el marcador universal, el marcador espacial, el marcador de dimensión, el marcador celular y el marcador molecular están en cualquier orden. El código de barras puede comprender una región de unión al objetivo. La región de unión al objetivo puede interactuar con un objetivo (por ejemplo, ácido nucleico, ARN, ARNm, ADN objetivo) en una muestra. Por ejemplo, una región de unión al objetivo puede comprender una secuencia oligo(dT) que puede interactuar con las colas poli(A) de los ARNm. En algunos casos, los marcadores del código de barras (por ejemplo, marcador universal, marcador de dimensión, marcador espacial, marcador celular y secuencia de código de barras) pueden estar separados por 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 o más nucleótidos.

Un marcador, por ejemplo el marcador celular, puede comprender un conjunto único de subsecuencias de ácidos nucleicos de longitud definida, por ejemplo, de siete nucleótidos cada una (equivalente a la cantidad de bits

usados en algunos códigos de corrección de errores de Hamming), que puede estar diseñado para proporcionar la capacidad de corrección de errores. El conjunto de subsecuencias de corrección de errores comprende siete secuencias de nucleótidos que pueden diseñarse de tal manera que cualquier combinación de secuencias por pares en el conjunto muestre una "distancia genética" definida (o número de bases malapareadas), por ejemplo, un conjunto de subsecuencias de corrección de errores puede diseñarse para que muestren una distancia genética de tres nucleótidos. En este caso, la revisión de las secuencias de corrección de errores en el conjunto de datos de secuencias para moléculas de ácidos nucleicos objetivo marcadas (descritas con más detalle a continuación) puede permitir detectar o corregir errores de amplificación o secuenciación. En algunas realizaciones, la longitud de las subsecuencias de ácido nucleico usadas para crear códigos de corrección de errores puede variar, por ejemplo, pueden tener o tener aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 31, 40, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. En algunas realizaciones, pueden usarse subsecuencias de ácidos nucleicos de otras longitudes para crear códigos de corrección de errores.

El código de barras puede comprender una región de unión al objetivo. La región de unión al objetivo puede interactuar con un objetivo en una muestra. El objetivo puede ser, o comprender, ácidos ribonucleicos (ARN), ARN mensajeros (ARNm), microARN, ARN interferente pequeño (ARNip), productos de degradación de ARN, ARN que comprenden cada uno una cola de poli(A), o cualquier combinación de los mismos. En algunas realizaciones, la pluralidad de objetivos puede incluir ácidos desoxirribonucleicos (ADN).

En algunas realizaciones, una región de unión al objetivo puede comprender una secuencia oligo(dT) que puede interactuar con las colas poli(A) de los ARNm. Uno o más de los marcadores del código de barras (por ejemplo, el marcador universal, el marcador de dimensión, el marcador espacial, el marcador celular y las secuencias del código de barras (por ejemplo, marcador molecular)) pueden estar separados por un espaciador de otro o dos de los marcadores restantes del código de barras. El espaciador puede tener, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20, o más nucleótidos. En algunas realizaciones, ninguno de los marcadores del código de barras está separado por un espaciador.

Marcadores Universales

Un código de barras puede comprender uno o más marcadores universales. En algunas realizaciones, uno o más marcadores universales pueden ser los mismos para todos los códigos de barras en el conjunto de códigos de barras unidos a un soporte sólido dado. En algunas realizaciones, el uno o más marcadores universales pueden ser los mismos para todos los códigos de barras unidos a una pluralidad de perlas. En algunas realizaciones, un marcador universal puede comprender una secuencia de ácidos nucleicos que es capaz de hibridar con un cebador de secuenciación. Los cebadores de secuenciación pueden usarse para secuenciar códigos de barras que comprenden un marcador universal. Los cebadores de secuenciación (por ejemplo, cebadores de secuenciación universales) pueden comprender cebadores de secuenciación asociados con plataformas de secuenciación de alto rendimiento. En algunas realizaciones, un marcador universal puede comprender una secuencia de ácidos nucleicos que es capaz de hibridar con un cebador de PCR. En algunas realizaciones, el marcador universal puede comprender una secuencia de ácidos nucleicos que es capaz de hibridar con un cebador de secuenciación y un cebador de PCR. La secuencia de ácidos nucleicos del marcador universal que es capaz de hibridar con un cebador de secuenciación o de PCR puede denominarse sitio de unión del cebador. Un marcador universal puede comprender una secuencia que puede usarse para iniciar la transcripción del código de barras. Un marcador universal puede comprender una secuencia que puede usarse para la extensión del código de barras o una región dentro del código de barras. Un marcador universal puede tener 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. Por ejemplo, un marcador universal puede comprender por lo menos aproximadamente 10 nucleótidos. Un marcador universal puede tener como mínimo, o como máximo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, o 300 nucleótidos de longitud. En algunas realizaciones, puede formar parte de la secuencia marcadora universal un conector escindible o un nucleótido modificado para permitir que el código de barras se escinda del soporte.

Marcadores de dimensión

Un código de barras puede comprender uno o más marcadores de dimensión. En algunas realizaciones, un marcador de dimensión puede comprender una secuencia de ácidos nucleicos que proporciona información sobre una dimensión en la que se produjo el marcado (por ejemplo, marcado estocástico). Por ejemplo, un marcador de dimensión puede proporcionar información sobre el momento en que se codificó con códigos de barras un objetivo. Un marcador de dimensión puede asociarse con un momento de codificación con códigos de barras (por ejemplo, codificación con códigos de barras estocásticos) en una muestra. Un marcador de dimensión puede activarse en el momento del marcado. Pueden activarse diferentes marcadores de dimensión en diferentes momentos. El marcador de dimensión proporciona información sobre el orden en que se codificaron con códigos de barras los objetivos, grupos de objetivos y/o muestras. Por ejemplo, puede codificarse con códigos de barras una población de células en la fase G0 del ciclo celular. Las células pueden pulsarse de nuevo con códigos de barras (por ejemplo, códigos de barras estocásticos) en la fase G1 del ciclo celular. Las células pueden pulsarse de nuevo con códigos de barras en la fase S del ciclo celular, y así sucesivamente. Los códigos de barras en cada pulso (por ejemplo, cada fase del

ciclo celular) pueden comprender diferentes marcadores de dimensión. De esta manera, el marcador de dimensión proporciona información sobre qué objetivos se marcaron en qué fase del ciclo celular. Los marcadores de dimensión pueden interrogar muchos momentos biológicos diferentes. Los momentos biológicos ejemplares pueden incluir, pero no se limitan a, el ciclo celular, la transcripción (por ejemplo, el inicio de la transcripción) y la degradación del transcrito. En otro ejemplo, puede marcarse una muestra (por ejemplo, una célula, una población de células) antes y/o después del tratamiento con un fármaco y/o terapia. Los cambios en el número de copias de distintos objetivos pueden ser indicativos de la respuesta de la muestra al fármaco y/o la terapia.

Un marcador de dimensión puede ser activable. Un marcador de dimensión activable puede activarse en un punto temporal específico. El marcador activable puede ser, por ejemplo, activado constitutivamente (por ejemplo, no apagado). El marcador de dimensión activable puede activarse, por ejemplo, reversiblemente (por ejemplo, el marcador de dimensión activable puede activarse y desactivarse). El marcador de dimensión puede ser, por ejemplo, reversiblemente activable por lo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más veces. El marcador de dimensión puede activarse reversiblemente, por ejemplo, por lo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más veces. En algunas realizaciones, el marcador de dimensión puede activarse con fluorescencia, luz, un evento químico (por ejemplo, escisión, ligamiento de otra molécula, adición de modificaciones (por ejemplo, pegilarse, sumoilarse, acetilarse, metilarse, desacetilarse, desmetilarse), un evento fotoquímico (por ejemplo, fotoenjaulado), y la introducción de un nucleótido no natural.

El marcador de dimensión puede, en algunas realizaciones, ser idéntico para todos los códigos de barras (por ejemplo, códigos de barras estocásticos) unidos a un soporte sólido dado (por ejemplo, una perla), pero diferente para diferentes soportes sólidos (por ejemplo, perlas). En algunas realizaciones, por lo menos el 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99% o 100% de los códigos de barras en el mismo soporte sólido pueden comprender el mismo marcador de dimensión. En algunas realizaciones, por lo menos el 60% de los códigos de barras en el mismo soporte sólido pueden comprender el mismo marcador de dimensión. En algunas realizaciones, por lo menos el 95% de los códigos de barras en el mismo soporte sólido pueden comprender el mismo marcador de dimensión.

Puede haber hasta 10^6 o más secuencias de marcadores de dimensiones únicas representadas en una pluralidad de soportes sólidos (por ejemplo, perlas). Un marcador de dimensión puede tener, o tener aproximadamente 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. Un marcador de dimensión puede tener por lo menos, o como máximo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200 o 300 nucleótidos de longitud. Un marcador de dimensión puede comprender entre aproximadamente 5 y aproximadamente 200 nucleótidos. Un marcador de dimensión puede comprender entre aproximadamente 10 y aproximadamente 150 nucleótidos. Un marcador de dimensión puede comprender entre aproximadamente 20 y aproximadamente 125 nucleótidos de longitud.

Marcadores espaciales

Un código de barras puede comprender uno o más marcadores espaciales. En algunas realizaciones, un marcador espacial puede comprender una secuencia de ácidos nucleicos que proporciona información sobre la orientación espacial de una molécula objetivo que está asociada con el código de barras. Un marcador espacial puede estar asociado con una coordenada en una muestra. La coordenada puede ser una coordenada fija. Por ejemplo, una coordenada puede ser fija con respecto a un sustrato. Un marcador espacial puede hacer referencia a una cuadrícula de dos o tres dimensiones. Una coordenada puede ser fija con respecto a un punto de referencia. El punto de referencia puede ser identificable en el espacio. Un punto de referencia puede ser una estructura de la que puede obtenerse una imagen. Un punto de referencia puede ser una estructura biológica, por ejemplo, un punto de referencia anatómico. Un punto de referencia puede ser un punto de referencia celular, por ejemplo, un orgánulo. Un punto de referencia puede ser un punto de referencia no natural, como una estructura con un identificador identificable, como un código de color, un código de barras, una propiedad magnética, fluorescentes, radiactividad o un tamaño o forma únicos. Un marcador espacial puede estar asociado con una partición física (por ejemplo, un pocillo, un recipiente o una gotita). En algunas realizaciones, se usan juntos múltiples marcadores espaciales para codificar una o más posiciones en el espacio.

El marcador espacial puede ser idéntico para todos los códigos de barras unidos a un soporte sólido dado (por ejemplo, una perla), pero diferente para diferentes soportes sólidos (por ejemplo, perlas). En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprende el mismo marcador espacial puede ser, o ser de aproximadamente, un 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, o un número o un intervalo entre dos cualquiera de estos valores. En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprende el mismo marcador espacial puede ser por lo menos, o como máximo, del 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, o 100%. En algunas realizaciones, por lo menos el 60% de los códigos de barras en el mismo soporte sólido pueden comprender el mismo marcador espacial. En algunas realizaciones, por lo menos el 95% de los códigos de barras en el mismo soporte sólido pueden comprender el mismo marcador espacial.

Puede haber hasta 10^6 o más secuencias de marcadores espaciales únicos representadas en una

pluralidad de soportes sólidos (por ejemplo, perlas). Un marcador espacial puede tener, o tener aproximadamente, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. Un marcador espacial puede tener por lo menos o como máximo 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200 o 300 nucleótidos de longitud. Un marcador espacial puede comprender entre aproximadamente 5 y aproximadamente 200 nucleótidos. Un marcador espacial puede comprender entre aproximadamente 10 y aproximadamente 150 nucleótidos. Un marcador espacial puede comprender entre aproximadamente 20 y aproximadamente 125 nucleótidos de longitud.

Marcadores celulares

Un código de barras (por ejemplo, un código de barras estocástico) puede comprender uno o más marcadores celulares. En algunas realizaciones, un marcador celular puede comprender una secuencia de ácidos nucleicos que proporciona información para determinar qué ácido nucleico objetivo se originó a partir de qué célula. En algunas realizaciones, el marcador celular es idéntico para todos los códigos de barras unidos a un soporte sólido dado (por ejemplo, una perla), pero diferente para diferentes soportes sólidos (por ejemplo, perlas). En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprende el mismo marcador celular puede ser de aproximadamente el 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, o un número o un intervalo entre dos cualquiera de estos valores. En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprende el mismo marcador celular puede ser de aproximadamente el 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99% o 100%. Por ejemplo, por lo menos el 60% de los códigos de barras en el mismo soporte sólido pueden comprender el mismo marcador celular. Como otro ejemplo, por lo menos el 95% de los códigos de barras en el mismo soporte sólido pueden comprender el mismo marcador celular.

Puede haber hasta 10^6 o más secuencias de marcadores celulares únicos representadas en una pluralidad de soportes sólidos (por ejemplo, perlas). Un marcador celular puede tener, o tener aproximadamente, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. Un marcador celular puede tener por lo menos, o como máximo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200 o 300 nucleótidos de longitud. Por ejemplo, un marcador celular puede comprender entre aproximadamente 5 y aproximadamente 200 nucleótidos. Como otro ejemplo, un marcador celular puede comprender entre aproximadamente 10 y aproximadamente 150 nucleótidos. Como otro ejemplo más, un marcador celular puede comprender entre aproximadamente 20 y aproximadamente 125 nucleótidos de longitud.

Secuencias de código de barras

Un código de barras puede comprender una o más secuencias de códigos de barras. En algunas realizaciones, una secuencia de código de barras puede comprender una secuencia de ácidos nucleicos que proporciona información de identificación para el tipo específico de especie de ácido nucleico objetivo hibridada con el código de barras. Una secuencia de código de barras puede comprender una secuencia de ácidos nucleicos que proporciona un contador (por ejemplo, que proporciona una aproximación aproximada) para la aparición específica de la especie de ácido nucleico objetivo hibridada con el código de barras (por ejemplo, región de unión al objetivo).

En algunas realizaciones, se une un conjunto diverso de secuencias de códigos de barras a un soporte sólido dado (por ejemplo, una perla). En algunas realizaciones, puede haber, o haber aproximadamente, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , o un número o un intervalo entre dos cualquiera de estos valores, secuencias de marcadores moleculares únicas. Por ejemplo, una pluralidad de códigos de barras puede comprender aproximadamente 6561 secuencias de códigos de barras con secuencias distintas. Como otro ejemplo, una pluralidad de códigos de barras puede comprender aproximadamente 65536 secuencias de códigos de barras con secuencias distintas. En algunas realizaciones, puede haber por lo menos, o como máximo, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 o 10^9 , secuencias de códigos de barras únicas. Las secuencias de marcadores moleculares únicas pueden unirse a un soporte sólido dado (por ejemplo, una perla). En algunas realizaciones, la secuencia de marcador molecular única está parcial o totalmente abarcada por una partícula (por ejemplo, una perla de hidrogel).

La longitud de un código de barras puede ser diferente en diferentes implementaciones. Por ejemplo, un código de barras puede tener, o tener aproximadamente, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, o un número o un intervalo entre dos cualesquiera de estos valores, nucleótidos de longitud. Como otro ejemplo, un código de barras puede tener por lo menos, o como máximo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200 o 300 nucleótidos de longitud.

Marcadores moleculares

Un código de barras (por ejemplo, un código de barras estocástico) puede comprender uno o más marcadores moleculares. Los marcadores moleculares pueden incluir secuencias de códigos de barras. En algunas realizaciones, un marcador molecular puede comprender una secuencia de ácidos nucleicos que proporciona información de identificación para el tipo específico de especie de ácido nucleico objetivo hibridada con el código de barras. Un marcador molecular puede comprender una secuencia de ácidos nucleicos que proporciona un contador

para la aparición específica de la especie de ácido nucleico objetivo hibridada con el código de barras (por ejemplo, región de unión al objetivo).

En algunas realizaciones, se une un conjunto diverso de marcadores moleculares a un soporte sólido dado (por ejemplo, una perla). En algunas realizaciones, puede haber, o haber aproximadamente, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , o un número o un intervalo entre dos cualquiera de estos valores, de secuencias de marcadores moleculares únicas. Por ejemplo, una pluralidad de códigos de barras puede comprender aproximadamente 6561 marcadores moleculares con secuencias distintas. Como otro ejemplo, una pluralidad de códigos de barras puede comprender aproximadamente 65536 marcadores moleculares con secuencias distintas. En algunas realizaciones, puede haber por lo menos, o como máximo, 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 o 10^9 , secuencias de marcadores moleculares únicas. Los códigos de barras con secuencias de marcadores moleculares únicas pueden unirse a un soporte sólido determinado (por ejemplo, una perla).

Para codificación con códigos de barras (por ejemplo, codificación con códigos de barras estocásticos) que usan una pluralidad de códigos de barras estocásticos, la proporción del número de secuencias de marcadores moleculares diferentes y el número de apariciones de cualquiera de los objetivos puede ser, o ser aproximadamente, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, o un número o un intervalo entre dos cualquiera de estos valores. Un objetivo puede ser una especie de ARNm que comprende moléculas de ARNm con secuencias idénticas o casi idénticas. En algunas realizaciones, la proporción del número de secuencias de marcadores moleculares diferentes y el número de apariciones de cualquiera de los objetivos es por lo menos, o como máximo, 1:1, 2:1, 3:1, 4:1, 5:1, 6:1, 7:1, 8:1, 9:1, 10:1, 11:1, 12:1, 13:1, 14:1, 15:1, 16:1, 17:1, 18:1, 19:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, o 100:1.

Un marcador molecular puede tener, o tener aproximadamente, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. Un marcador molecular puede tener por lo menos, o como máximo, 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200 o 300 nucleótidos de longitud.

Región de unión al objetivo

Un código de barras puede comprender una o más regiones de unión al objetivo, como sondas de captura. En algunas realizaciones, una región de unión al objetivo puede hibridar con un objetivo de interés. En algunas realizaciones, las regiones de unión al objetivo pueden comprender una secuencia de ácidos nucleicos que hibrida específicamente con un objetivo (por ejemplo, ácido nucleico objetivo, molécula objetivo, por ejemplo, un ácido nucleico celular a analizar), por ejemplo, con una secuencia génica específica. En algunas realizaciones, una región de unión al objetivo puede comprender una secuencia de ácidos nucleicos que puede unirse (por ejemplo, hibridar) a una localización específica de un ácido nucleico objetivo específico. En algunas realizaciones, la región de unión al objetivo puede comprender una secuencia de ácidos nucleicos que es capaz de una hibridación específica con un saliente de sitio de enzima de restricción (por ejemplo, un extremo cohesivo de EcoRI). Luego, el código de barras puede ligarse a cualquier molécula de ácido nucleico que comprenda una secuencia complementaria con el saliente del sitio de restricción.

En algunas realizaciones, una región de unión al objetivo puede comprender una secuencia de ácidos nucleicos objetivo no específica. Una secuencia de ácidos nucleicos objetivo no específica puede referirse a una secuencia que puede unirse a múltiples ácidos nucleicos objetivo, independientemente de la secuencia específica del ácido nucleico objetivo. Por ejemplo, la región de unión al objetivo puede comprender una secuencia multimérica aleatoria o una secuencia oligo(dT) que hibrida con la cola poli(A) en moléculas de ARNm. Una secuencia multimérica aleatoria puede ser, por ejemplo, una secuencia multimérica aleatoria de dímero, trímero, cuádamero, pentámero, hexámero, septámero, octámero, nonámero, decámero o superior de cualquier longitud. En algunas realizaciones, la región de unión al objetivo es la misma para todos los códigos de barras unidos a una perla dada. En algunas realizaciones, las regiones de unión al objetivo para la pluralidad de códigos de barras unidos a una perla dada pueden comprender dos o más secuencias de unión al objetivo diferentes. Una región de unión al objetivo puede tener, o tener aproximadamente, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. Una región de unión al objetivo puede tener como máximo aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más nucleótidos de longitud.

En algunas realizaciones, una región de unión al objetivo puede comprender un oligo(dT) que puede hibridar con ARNm que comprenden extremos poliadenilados. Una región de unión al objetivo puede ser específica de un gen. Por ejemplo, una región de unión al objetivo puede configurarse para hibridar con una región específica de un objetivo. En algunas realizaciones, una región de unión al objetivo no comprende un oligo(dT). Una región de unión al objetivo puede tener, o tener aproximadamente, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos de longitud. Una región de unión al objetivo puede tener por lo menos, o como máximo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 nucleótidos de longitud. Una región de unión al objetivo puede tener una longitud de aproximadamente 5-30 nucleótidos.

Propiedad de orientación

Un código de barras estocástico (por ejemplo, un código de barras estocástico) puede comprender una o más propiedades de orientación que pueden usarse para orientar (por ejemplo, alinear) los códigos de barras. Un código de barras puede comprender una fracción para enfoque isoelectrico. Diferentes códigos de barras pueden comprender diferentes puntos de enfoque isoelectrico. Cuando estos códigos de barras se introducen en una muestra, la muestra puede someterse a enfoque isoelectrico para orientar los códigos de barras de una manera conocida. De esta manera, puede usarse la propiedad de orientación para desarrollar un mapa conocido de códigos de barras en una muestra. Las propiedades de orientación ejemplares pueden incluir movilidad electroforética (por ejemplo, sobre la base del tamaño del código de barras), punto isoelectrico, espín, conductividad y/o autoensamblaje. Por ejemplo, los códigos de barras con una propiedad de orientación de autoensamblaje pueden autoensamblarse en una orientación específica (por ejemplo, nanoestructura de ácido nucleico) tras la activación.

Propiedad de afinidad

Un código de barras (por ejemplo, un código de barras estocástico) puede comprender una o más propiedades de afinidad. Por ejemplo, un marcador espacial puede comprender una propiedad de afinidad. Una propiedad de afinidad puede incluir una fracción química y/o biológica que puede facilitar la unión del código de barras a otra entidad (por ejemplo, receptor celular). Por ejemplo, una propiedad de afinidad puede comprender un anticuerpo, por ejemplo, un anticuerpo específico para una fracción específica (por ejemplo, receptor) en una muestra. En algunas realizaciones, el anticuerpo puede guiar el código de barras a un tipo de célula o molécula específica. Los objetivos en y/o cerca del tipo de célula o molécula específicos pueden estar marcados (por ejemplo, marcados estocásticamente). La propiedad de afinidad puede, en algunas realizaciones, proporcionar información espacial además de la secuencia de nucleótidos del marcador espacial porque el anticuerpo puede guiar al código de barras a una localización específica. El anticuerpo puede ser un anticuerpo terapéutico, por ejemplo, un anticuerpo monoclonal o un anticuerpo policlonal. El anticuerpo puede ser humanizado o quimérico. El anticuerpo puede ser un anticuerpo desnudo o un anticuerpo de fusión.

El anticuerpo puede ser una molécula de inmunoglobulina de longitud completa (es decir, de origen natural o formada por procesos de recombinación de fragmentos de genes de inmunoglobulina normales) (por ejemplo, un anticuerpo de IgG) o una porción inmunológicamente activa (es decir, que se une específicamente) de una molécula de inmunoglobulina, como un fragmento de anticuerpo

El fragmento de anticuerpo puede ser, por ejemplo, una porción de un anticuerpo como F(ab')₂, Fab', Fab, Fv, sFv y similares. En algunas realizaciones, el fragmento de anticuerpo puede unirse con el mismo antígeno que reconoce el anticuerpo de longitud completa. El fragmento de anticuerpo puede incluir fragmentos aislados que consisten en las regiones variables de los anticuerpos, como los fragmentos "Fv" que consisten en las regiones variables de las cadenas pesada y ligera y moléculas polipeptídicas de cadena sencilla recombinantes en las que las regiones variables ligeras y pesadas están conectadas por un conector peptídico ("proteínas scFv"). Los anticuerpos ejemplares pueden incluir, pero no se limitan a, anticuerpos para células cancerosas, anticuerpos para virus, anticuerpos que se unen a receptores de la superficie celular (CD8, CD34, CD45) y anticuerpos terapéuticos.

Cebador de adaptador universal

Un código de barras puede comprender uno o más cebadores adaptadores universales. Por ejemplo, un código de barras específico de un gen, como un código de barras estocástico específico de un gen, puede comprender un cebador de adaptador universal. Un cebador de adaptador universal puede hacer referencia a una secuencia de nucleótidos que es universal en todos los códigos de barras. Un cebador de adaptador universal puede usarse para crear códigos de barras específicos de genes. Un cebador de adaptador universal puede tener, o tener aproximadamente, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, o un número o un intervalo entre dos cualquiera de estos nucleótidos de longitud. Un cebador de adaptador universal puede tener como mínimo, o como máximo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 nucleótidos de longitud. Un cebador de adaptador universal puede tener una longitud de 5 a 30 nucleótidos.

Conector

Cuando un código de barras comprende más de un tipo de marcador (por ejemplo, más de un marcador celular o más de una secuencia de código de barras, como un marcador molecular), los marcadores pueden estar intercalados con una secuencia marcadora de conector. Una secuencia marcadora de conector puede tener por lo menos aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más nucleótidos de longitud. Una secuencia marcadora de conector puede tener como máximo aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más nucleótidos de longitud. En algunos casos, una secuencia marcadora de conector tiene una longitud de 12 nucleótidos. Una secuencia marcadora de conector puede usarse para facilitar la síntesis del código de barras. El

marcador del conector puede comprender un código de corrección de errores (por ejemplo, Hamming).

Soportes Sólidos

Los códigos de barras, como los códigos de barras estocásticos, divulgados en la presente pueden, en algunas realizaciones, estar asociados con un soporte sólido. El soporte sólido puede ser, por ejemplo, una partícula sintética. En algunas realizaciones, algunas o todas las secuencias de códigos de barras, como marcadores moleculares para códigos de barras estocásticos (por ejemplo, las primeras secuencias de códigos de barras) de una pluralidad de códigos de barras (por ejemplo, la primera pluralidad de códigos de barras) en un soporte sólido difieren en por lo menos un nucleótido. Los marcadores celulares de los códigos de barras sobre un mismo soporte sólido pueden ser los mismos. Los marcadores celulares de los códigos de barras en diferentes soportes sólidos pueden diferir en por lo menos un nucleótido. Por ejemplo, los primeros marcadores celulares de una primera pluralidad de códigos de barras sobre un primer soporte sólido pueden tener la misma secuencia, y los segundos marcadores celulares de una segunda pluralidad de códigos de barras sobre un segundo soporte sólido pueden tener la misma secuencia. Los primeros marcadores celulares de la primera pluralidad de códigos de barras en el primer soporte sólido y los segundos marcadores celulares de la segunda pluralidad de códigos de barras en el segundo soporte sólido pueden diferir en por lo menos un nucleótido. Un marcador celular puede tener, por ejemplo, una longitud de aproximadamente 5-20 nucleótidos. Una secuencia de código de barras puede tener, por ejemplo, una longitud de aproximadamente 5-20 nucleótidos. La partícula sintética puede ser, por ejemplo, una perla.

La perla puede ser, por ejemplo, una perla de gel de sílice, una perla de vidrio de poro controlado, una perla magnética, una perla Dynabead, una perla de Sephadex/Sepharose, una perla de celulosa, una perla de poliestireno o cualquier combinación de las mismas. La perla puede comprender un material como polidimetilsiloxano (PDMS), poliestireno, vidrio, polipropileno, agarosa, gelatina, hidrogel, paramagnético, cerámico, plástico, vidrio, metilestireno, polímero acrílico, titanio, látex, sefarsa, celulosa, nailon, silicona, o cualquier combinación de los mismos.

En algunas realizaciones, la perla puede ser una perla polimérica, por ejemplo, una perla deformable o una perla de gel, funcionalizada con códigos de barras o códigos de barras estocásticos (como perlas de gel de 10X Genomics (San Francisco, CA). En alguna implementación, una perla de gel puede comprender geles a base de polímeros. Pueden generarse perlas de gel, por ejemplo, encapsulando uno o más precursores poliméricos en gotitas. Tras la exposición de los precursores poliméricos a un acelerador (por ejemplo, tetrametilendiamina (TEMED)), puede generarse una perla de gel.

En algunas realizaciones, la partícula puede ser alterable (por ejemplo, soluble o degradable). Por ejemplo, la perla polimérica puede disolverse, fundirse o degradarse, por ejemplo, en las condiciones deseadas. La condición deseada puede incluir una condición ambiental. La condición deseada puede dar como resultado que la perla polimérica se disuelva, funda o degrade de manera controlada. Una perla de gel puede disolverse, fundirse o degradarse debido a un estímulo químico, un estímulo físico, un estímulo biológico, un estímulo térmico, un estímulo magnético, un estímulo eléctrico, un estímulo de luz o cualquier combinación de los mismos.

Los analitos y/o reactivos, como códigos de barras de oligonucleótidos, por ejemplo, pueden acoplarse/inmovilizarse a la superficie interior de una perla de gel (por ejemplo, el interior accesible a través de la difusión de un código de barras de oligonucleótidos y/o materiales usados para generar un código de barras de oligonucleótidos) y/o la superficie exterior de una perla de gel o cualquier otra microcápsula descrita en la presente. El acoplamiento/inmovilización puede ser a través de cualquier forma de enlace químico (por ejemplo, enlace covalente, enlace iónico) o fenómenos físicos (por ejemplo, fuerzas de Van der Waals, interacciones dipolo-dipolo, etc.). En algunas realizaciones, el acoplamiento/inmovilización de un reactivo a una perla de gel o cualquier otra microcápsula descrita en la presente puede ser reversible como, por ejemplo, a través de una fracción lábil (por ejemplo, a través de un reticulante químico, incluyendo los reticulantes químicos descritos en la presente). Tras aplicar un estímulo, la fracción lábil puede escindirse y puede liberarse el reactivo inmovilizado. En algunas realizaciones, la fracción lábil es un enlace disulfuro. Por ejemplo, en el caso de que un código de barras de oligonucleótidos se inmovilice en una perla de gel a través de un enlace disulfuro, la exposición del enlace disulfuro a un agente reductor puede escindir el enlace disulfuro y liberar el código de barras de oligonucleótidos de la perla. La fracción lábil puede incluirse como parte de una perla de gel o microcápsula, como parte de un conector químico que enlaza un reactivo o analito a una perla de gel o microcápsula y/o como parte de un reactivo o analito. En algunas realizaciones, por lo menos un código de barras de la pluralidad de códigos de barras puede inmovilizarse en la partícula, inmovilizarse parcialmente en la partícula, encerrarse en la partícula, encerrarse parcialmente en la partícula o cualquier combinación de los mismos.

En algunas realizaciones, una perla de gel puede comprender una amplia variedad de polímeros diferentes, incluyendo, pero no limitados a: polímeros, polímeros sensibles al calor, polímeros fotosensibles, polímeros magnéticos, polímeros sensibles al pH, polímeros sensibles a las sales, polímeros químicamente sensibles, polielectrolitos, polisacáridos, péptidos, proteínas y/o plásticos. Los polímeros pueden incluir, pero no se limitan a, materiales como poli(N-isopropilacrilamida) (PNIPAAm), poli(sulfonato de estireno) (PSS), poli(alilamina) (PAAm), poli(ácido acrílico) (PAA), poli(etilenimina) (PEI), poli(cloruro de dialildimetilamonio) (PDADMAC), poli(pirrola) (PPy),

poli(vinilpirrolidona) (PVPON), poli(vinilpiridina) (PVP), poli(ácido metacrílico) (PMAA)), poli(metacrilato de metilo) (PMMA), poliestireno (PS), poli(tetrahidrofurano) (PTHF), poli(ftaladehído) (PTHF), poli(hexilviológeno) (PHV), poli(L-lisina) (PLL), poli(L-arginina) (PARG), poli(ácido láctico-coglicólico) (PLGA).

Pueden usarse numerosos estímulos químicos para desencadenar la alteración, disolución o degradación de las perlas. Los ejemplos de estos cambios químicos pueden incluir, pero no se limitan a, cambios mediados por el pH en la pared de la perla, disgregación de la pared de la perla a través de la escisión química de los enlaces de reticulación, despolimerización desencadenada de la pared de la perla y reacciones de cambio de pared de la perla. También pueden usarse cambios de volumen para desencadenar la alteración de las perlas.

Los cambios de volumen o físicos en la microcápsula a través de varios estímulos también ofrecen muchas ventajas en el diseño de cápsulas para liberar reactivos. Los cambios de volumen o masivos se producen en una escala macroscópica, en donde la ruptura de las perlas es el resultado de fuerzas mecanofísicas inducidas por un estímulo. Estos procesos pueden incluir, pero no se limitan a, ruptura inducida por presión, fusión de la pared de la perla o cambios en la porosidad de la pared de la perla.

También pueden usarse estímulos biológicos para desencadenar la alteración, disolución o degradación de las perlas. En general, los desencadenantes biológicos se asemejan a los desencadenantes químicos, pero muchos ejemplos usan biomoléculas o moléculas que se encuentran comúnmente en los sistemas vivos como enzimas, péptidos, sacáridos, ácidos grasos, ácidos nucleicos y similares. Por ejemplo, las perlas pueden comprender polímeros con enlaces cruzados de péptidos que son sensibles a la escisión por proteasas específicas. Más específicamente, un ejemplo puede comprender una microcápsula que comprende enlaces cruzados de péptido GFLGK. Tras la adición de un activador biológico como la proteasa catépsina B, los enlaces cruzados peptídicos de la cubierta se escinden y se libera el contenido de las perlas. En otros casos, las proteasas pueden activarse por calor. En otro ejemplo, las perlas comprenden una pared de cubierta que comprende celulosa. La adición de la enzima hidrolítica quitosano sirve como desencadenante biológico para la escisión de los enlaces celulósicos, la despolimerización de la pared de la cubierta y la liberación de su contenido interno.

También puede inducirse a las perlas para que liberen su contenido tras la aplicación de un estímulo térmico. Un cambio de temperatura puede provocar una variedad de cambios en las perlas. Un cambio en el calor puede provocar la fusión de una perla de tal manera que la pared de la perla se disgregue. En otros casos, el calor puede aumentar la presión interna de los componentes internos de la perla de tal manera que la perla se rompa o explote. En otros casos más, el calor puede transformar la perla en un estado deshidratado y encogido. El calor también puede actuar sobre polímeros sensibles al calor dentro de la pared de una perla para provocar la ruptura de la perla.

La inclusión de nanopartículas magnéticas en la pared de la perla de las microcápsulas puede permitir la ruptura desencadenada de las perlas, así como guiar las perlas en una matriz. Un dispositivo de esta divulgación puede comprender perlas magnéticas para cualquier propósito. En un ejemplo, la incorporación de nanopartículas de Fe_3O_4 en perlas que contienen polielectrolitos desencadena la ruptura en presencia de un estímulo de campo magnético oscilante.

Una perla también puede alterarse, disolverse o degradarse como resultado de la estimulación eléctrica. De manera similar a las partículas magnéticas descritas en la sección anterior, las perlas eléctricamente sensibles pueden permitir tanto la ruptura desencadenada de las perlas como otras funciones, como la alineación en un campo eléctrico, la conductividad eléctrica o las reacciones redox. En un ejemplo, las perlas que contienen material eléctricamente sensible se alinean en un campo eléctrico de tal manera que puede controlarse la liberación de los reactivos internos. En otros ejemplos, los campos eléctricos pueden inducir reacciones redox dentro de la propia pared de la perla que puede aumentar la porosidad.

También puede usarse un estímulo de luz para alterar las perlas. Son posibles numerosos activadores de luz y pueden incluir sistemas que usan varias moléculas, como nanopartículas y cromóforos, capaces de absorber fotones de intervalos específicos de longitudes de onda. Por ejemplo, pueden usarse recubrimientos de óxido de metal como disparadores de cápsulas. La irradiación UV de cápsulas de polielectrolito recubiertas con SiO_2 puede dar como resultado la disgregación de la pared de la perla. En otro ejemplo más, pueden incorporarse en la pared de la perla materiales fotoconmutables como grupos azobenceno. Tras la aplicación de luz UV o visible, los productos químicos como estos experimentan una isomerización reversible de cis a trans tras la absorción de fotones. En este aspecto, la incorporación de interruptores de fotones da como resultado una pared de perlas que puede disgregarse o volverse más porosa con la aplicación de un desencadenante de luz.

Por ejemplo, en un ejemplo no limitativo de codificación con códigos de barras (por ejemplo, codificación con códigos de barras estocástico) ilustrado en la FIG. 2, después de introducir células como células individuales en una pluralidad de micropocillos de una matriz de micropocillos en el bloque 208, pueden introducirse perlas en la pluralidad de micropocillos de la matriz de micropocillos en el bloque 212. Cada micropocillo puede comprender una perla. Las perlas pueden comprender una pluralidad de códigos de barras. Un código de barras puede comprender

una región de amina 5' unida a una perla. El código de barras puede comprender un marcador universal, una secuencia de código de barras (por ejemplo, un marcador molecular), una región de unión al objetivo o cualquier combinación de los mismos.

Los códigos de barras divulgados en la presente pueden asociarse con (por ejemplo, unirse a) un soporte sólido (por ejemplo, una perla). Los códigos de barras asociados con un soporte sólido pueden comprender cada uno una secuencia de códigos de barras seleccionada de un grupo que comprende por lo menos 100 o 1000 secuencias de códigos de barras con secuencias únicas. En algunas realizaciones, diferentes códigos de barras asociados con un soporte sólido pueden comprender códigos de barras con diferentes secuencias. En algunas realizaciones, un porcentaje de los códigos de barras asociados con un soporte sólido comprende el mismo marcador celular. Por ejemplo, el porcentaje puede ser, o ser de aproximadamente el 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, o un número o un intervalo entre dos cualquiera de estos valores. Como otro ejemplo, el porcentaje puede ser como mínimo o como máximo el 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99% o 100%. En algunas realizaciones, los códigos de barras asociados a un soporte sólido pueden tener el mismo marcador celular. Los códigos de barras asociados con diferentes soportes sólidos pueden tener diferentes marcadores celulares seleccionados de un grupo que comprende por lo menos 100 o 1000 marcadores celulares con secuencias únicas.

Los códigos de barras divulgados en la presente pueden asociarse a (por ejemplo, unirse a) un soporte sólido (por ejemplo, una perla). En algunas realizaciones, la codificación con códigos de barras de la pluralidad de objetivos en la muestra puede realizarse con un soporte sólido que incluye una pluralidad de partículas sintéticas asociadas con la pluralidad de códigos de barras. En algunas realizaciones, el soporte sólido puede incluir una pluralidad de partículas sintéticas asociadas con la pluralidad de códigos de barras. Los marcadores espaciales de la pluralidad de códigos de barras sobre diferentes soportes sólidos pueden diferir en por lo menos un nucleótido. El soporte sólido puede, por ejemplo, incluir la pluralidad de códigos de barras en dos dimensiones o en tres dimensiones. Las partículas sintéticas pueden ser perlas. Las perlas pueden ser perlas de gel de sílice, perlas de vidrio de poro controlado, perlas magnéticas, Dynabeads, Sephadex/Sepharose, perlas de celulosa, perlas de poliestireno, o cualquier combinación de las mismas. El soporte sólido puede incluir un polímero, una matriz, un hidrogel, un dispositivo de matriz de agujas, un anticuerpo o cualquier combinación de los mismos. En algunas realizaciones, los soportes sólidos pueden flotar libremente. En algunas realizaciones, los soportes sólidos pueden incorporarse en una matriz semisólida o sólida. Los códigos de barras pueden no estar asociados a soportes sólidos. Los códigos de barras pueden ser nucleótidos individuales. Los códigos de barras pueden asociarse con un sustrato. En algunas realizaciones, los códigos de barras pueden asociarse con células individuales en particiones, por ejemplo, gotitas como microgotitas, o pocillos de un sustrato como micropocillos (por ejemplo, en una placa de múltiples pocillos) o cámaras (por ejemplo, en un dispositivo fluídico). Las gotitas ejemplares pueden incluir gotitas de hidrogel. Los códigos de barras en las particiones pueden estar inmovilizados sobre un soporte sólido, o pueden estar libres en solución.

Como se usa en la presente, los términos "atado", "unido" e "inmovilizado" se usan indistintamente y pueden referirse a medios covalentes o no covalentes para unir códigos de barras a un soporte sólido. Puede usarse cualquiera de una variedad de diferentes soportes sólidos como soporte sólido para unir códigos de barras presintetizados o para la síntesis de código de barras en fase sólida *in situ*.

En algunas realizaciones, el soporte sólido es una perla. La perla puede comprender uno o más tipos de esfera sólida, porosa o hueca, bola, cojinete, cilindro u otra configuración similar en la que pueda inmovilizarse un ácido nucleico (por ejemplo, covalente o no covalentemente). La perla puede estar compuesta, por ejemplo, de plástico, cerámica, metal, material polimérico o cualquier combinación de los mismos. Una perla puede ser, o comprender, una partícula discreta que es esférica (por ejemplo, microesferas) o que tiene una forma no esférica o irregular, como cúbica, cuboide, piramidal, cilíndrica, cónica, oblonga o en forma de disco, y similiares. En algunas realizaciones, una perla puede tener una forma no esférica.

Las perlas pueden comprender una variedad de materiales que incluyen, pero no se limitan a, materiales paramagnéticos (por ejemplo, magnesio, molibdeno, litio y tántalo), materiales superparamagnéticos (por ejemplo, nanopartículas de ferrita (Fe_3O_4 ; magnetita)), materiales ferromagnéticos (por ejemplo, hierro, níquel, cobalto, algunas de sus aleaciones y algunos compuestos de metales de tierras raras), cerámica, plástico, vidrio, poliestireno, sílice, metilestireno, polímeros acrílicos, titanio, látex, sefarsa, agarosa, hidrogel, polímero, celulosa, nailon, o cualquier combinación de los mismos.

En algunas realizaciones, la perla (por ejemplo, la perla a la que se unen los marcadores) es una perla de hidrogel. En algunas realizaciones, la perla comprende hidrogel.

Algunas realizaciones divulgadas en la presente incluyen una o más partículas (por ejemplo, perlas). Cada una de las partículas puede comprender una pluralidad de oligonucleótidos (por ejemplo, códigos de barras). Cada uno de la pluralidad de oligonucleótidos puede comprender una secuencia de código de barras (por ejemplo, una secuencia de marcador molecular), un marcador celular y una región de unión al objetivo (por ejemplo, una secuencia de oligo(dT), una secuencia específica de gen, un multímero aleatorio, o una combinación de los mismos).

La secuencia del marcador celular de cada uno de la pluralidad de oligonucleótidos puede ser la misma. Las secuencias de marcadores celulares de oligonucleótidos en diferentes partículas pueden ser diferentes de manera que puedan identificarse los oligonucleótidos en diferentes partículas. El número de secuencias de marcadores celulares diferentes puede ser diferente en diferentes implementaciones. En algunas realizaciones, el número de secuencias de marcadores celulares puede ser, o ser aproximadamente 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , 10^9 , un número o un intervalo entre dos cualquiera de estos valores, o más. En algunas realizaciones, el número de secuencias de marcadores celulares puede ser por lo menos, o como máximo 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , o 10^9 . En algunas realizaciones, no más de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 o más de la pluralidad de partículas incluyen oligonucleótidos con la misma secuencia celular. En alguna realización, la pluralidad de partículas que incluyen oligonucleótidos con la misma secuencia celular puede ser como máximo del 0,1%, 0,2%, 0,3%, 0,4%, 0,5%, 0,6%, 0,7%, 0,8%, 0,9%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10% o más. En algunas realizaciones, ninguna de la pluralidad de partículas tiene la misma secuencia de marcador celular.

La pluralidad de oligonucleótidos en cada partícula puede comprender diferentes secuencias de códigos de barras (por ejemplo, marcadores moleculares). En algunas realizaciones, el número de secuencias de código de barras puede ser, o ser de aproximadamente 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 , 10^9 , o un número o un intervalo entre dos cualquiera de estos valores. En algunas realizaciones, el número de secuencias de código de barras puede ser por lo menos o como máximo de 10, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 10^6 , 10^7 , 10^8 o 10^9 . Por ejemplo, por lo menos 100 de la pluralidad de oligonucleótidos comprenden diferentes secuencias de código de barras. Como otro ejemplo, en una única partícula, por lo menos 100, 500, 1000, 5000, 10000, 15000, 20000, 50000, un número o un intervalo entre dos cualquiera de estos valores, o más de la pluralidad de oligonucleótidos comprenden diferentes secuencias de código de barras. Algunas realizaciones proporcionan una pluralidad de partículas que comprenden códigos de barras. En algunas realizaciones, la proporción de una aparición (o una copia o un número) de un objetivo a marcar y las diferentes secuencias de código de barras puede ser por lo menos 1:1, 1:2, 1:3, 1:4, 1: 5, 1:6, 1:7, 1:8, 1:9, 1:10, 1:11, 1:12, 1:13, 1:14, 1:15, 1:16, 1:17, 1:18, 1:19, 1:20, 1:30, 1:40, 1:50, 1:60, 1:70, 1:80, 1:90 o más. En algunas realizaciones, cada una de la pluralidad de oligonucleótidos comprende además un marcador de muestra, un marcador universal o ambos. La partícula puede ser, por ejemplo, una nanopartícula o micropartícula.

El tamaño de las perlas puede variar. Por ejemplo, el diámetro de la perla puede variar entre 0,1 micrómetro y 50 micrómetro. En algunas realizaciones, el diámetro de la perla puede ser, o ser aproximadamente, 0,1, 0,5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50 micrómetros, o un número o un intervalo entre dos cualquiera de estos valores.

El diámetro de la perla puede relacionarse con el diámetro de los pocillos del sustrato. En algunas realizaciones, el diámetro de la perla puede ser, o ser aproximadamente, un 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, o un número o un intervalo entre dos cualquiera de estos valores, más largo o más corto que el diámetro del pocillo. El diámetro de las perlas puede estar relacionado con el diámetro de una célula (por ejemplo, una única célula atrapada en un pocillo del sustrato). En algunas realizaciones, el diámetro de la perla puede ser por lo menos, o como máximo, un 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% o 100% más largo o más corto que el diámetro del pocillo. El diámetro de las perlas puede estar relacionado con el diámetro de una célula (por ejemplo, una única célula atrapada en un pocillo del sustrato). En algunas realizaciones, el diámetro de la perla puede ser, o ser aproximadamente, un 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 250%, 300%, o un número o un intervalo entre dos cualquiera de estos valores, más largo o más corto que el diámetro de la célula. En algunas realizaciones, el diámetro de las perlas puede ser por lo menos, o como máximo, un 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200%, 250% o 300% más largo o más corto que el diámetro de la célula.

Una perla puede unirse y/o incorporarse en un sustrato. Una perla puede unirse y/o incorporarse en un gel, hidrogel, polímero y/o matriz. La posición espacial de una perla dentro de un sustrato (por ejemplo, gel, matriz, andamiaje o polímero) puede identificarse usando el marcador espacial presente en el código de barras de la perla que puede servir como dirección de localización.

Los ejemplos de perlas pueden incluir, pero no se limitan a, perlas de estreptavidina, perlas de agarosa, perlas magnéticas, Dynabeads®, microperlas MACS®, perlas conjugadas con anticuerpos (por ejemplo, microperlas antiinmunoglobulina), perlas conjugadas con proteína A, perlas conjugadas con proteína G, perlas conjugadas con proteína A/G, perlas conjugadas con proteína L, perlas conjugadas con oligo(dT), perlas de sílice, perlas similares a sílice, microperlas antibiotina, microperlas anti fluorocromo y perlas magnéticas terminadas en carboxilo BcMag™.

Una perla puede asociarse (por ejemplo, impregnarse con) puntos cuánticos o colorantes fluorescentes para hacerla fluorescente en un canal óptico o en múltiples canales ópticos de fluorescencia. Una perla puede

asociarse con óxido de hierro u óxido de cromo para hacerla paramagnética o ferromagnética. Las perlas pueden ser identificables. Por ejemplo, puede obtenerse una imagen de una perla usando una cámara. Una perla puede tener un código detectable asociado con la perla. Por ejemplo, una perla puede comprender un código de barras. Una perla puede cambiar de tamaño, por ejemplo, debido al hinchamiento en una solución orgánica o inorgánica. Una perla puede ser hidrófoba. Una perla puede ser hidrófila. Una perla puede ser biocompatible.

Puede visualizarse un soporte sólido (por ejemplo, una perla). El soporte sólido puede comprender una etiqueta de visualización (por ejemplo, un colorante fluorescente). Un soporte sólido (por ejemplo, una perla) puede grabarse con un identificador (por ejemplo, un número). El identificador puede visualizarse a través de imagenología de las perlas.

Un soporte sólido puede comprender un material insoluble, semisoluble o soluble. Puede hacerse referencia a un soporte sólido como "funcionalizado" cuando incluye un conector, un andamiaje, un bloque funcional u otra fracción reactiva unida al mismo, mientras que un soporte sólido puede estar "no funcionalizado" cuando carece de dicha fracción reactiva unida al mismo. El soporte sólido puede emplearse libre en solución, como en un formato de pocillo de microtitulación; en un formato continuo, como en una columna; o en una tira reactiva.

El soporte sólido puede comprender una membrana, papel, plástico, superficie recubierta, superficie plana, vidrio, portaobjetos, chip o cualquier combinación de los mismos. Un soporte sólido puede tomar la forma de resinas, geles, microesferas u otras configuraciones geométricas. Un soporte sólido puede comprender chips de sílice, micropartículas, nanopartículas, placas, matrices, capilares, soportes planos como filtros de fibra de vidrio, superficies de vidrio, superficies metálicas (acero, oro, plata, aluminio, silicio y cobre), soportes de vidrio, soportes de plástico, soportes de silicona, chips, filtros, membranas, placas de micropocillos, portaobjetos, materiales plásticos, incluyendo placas o membranas de múltiples pocillos (por ejemplo, formados por polietileno, polipropileno, poliamida, difluoruro de polivinilideno) y/u obleas, peines, clavijas o agujas (por ejemplo, matrices de clavijas adecuados para síntesis o análisis combinatorios) o perlas en una matriz de hoyos o pocillos de nanolitros de superficies planas como obleas (por ejemplo, obleas de silicio), obleas con hoyos con o sin fondo de filtro.

El soporte sólido puede comprender una matriz polimérica (por ejemplo, gel, hidrogel). La matriz polimérica puede penetrar en el espacio intracelular (por ejemplo, alrededor de los orgánulos). La matriz polimérica puede bombearse por todo el sistema circulatorio.

Sustratos y matriz de micropocillos

Como se usa en la presente, un sustrato puede referirse a un tipo de soporte sólido. Un sustrato puede referirse a un soporte sólido que puede comprender códigos de barras o códigos de barras estocásticos de la divulgación. Un sustrato puede, por ejemplo, comprender una pluralidad de micropocillos. Por ejemplo, un sustrato puede ser una matriz de pocillos que comprende dos o más micropocillos. En algunas realizaciones, un micropocillo puede comprender una cámara de reacción pequeña de volumen definido. En algunas realizaciones, un micropocillo puede atrapar una o más células. En algunas realizaciones, un micropocillo puede atrapar solo una célula. En algunas realizaciones, un micropocillo puede atrapar uno o más soportes sólidos. En algunas realizaciones, un micropocillo puede atrapar solo un soporte sólido. En algunas realizaciones, un micropocillo atrapa una única célula y un único soporte sólido (por ejemplo, una perla). Un micropocillo puede comprender reactivos de código de barras de la divulgación.

Métodos de codificación con códigos de barras

La divulgación proporciona métodos para estimar el número de objetivos distintos en localizaciones distintas en una muestra física (por ejemplo, tejido, órgano, tumor, célula). Los métodos pueden comprender colocar códigos de barras (por ejemplo, códigos de barras estocásticos) muy cerca de la muestra, lisar la muestra, asociar distintos objetivos con los códigos de barras, amplificar los objetivos y/o contar digitalmente los objetivos. El método puede comprender además analizar y/o visualizar la información obtenida de los marcadores espaciales en los códigos de barras. En algunas realizaciones, un método comprende visualizar la pluralidad de objetivos en la muestra. Mapear la pluralidad de objetivos sobre el mapa de la muestra puede incluir generar un mapa bidimensional o un mapa tridimensional de la muestra. El mapa bidimensional y el mapa tridimensional pueden generarse antes o después de codificar con códigos de barras (por ejemplo, codificar estocásticamente con códigos de barras) la pluralidad de objetivos en la muestra. Visualizar la pluralidad de objetivos en la muestra puede incluir mapear la pluralidad de objetivos en un mapa de la muestra. Mapear la pluralidad de objetivos sobre el mapa de la muestra puede incluir generar un mapa bidimensional o un mapa tridimensional de la muestra. El mapa bidimensional y el mapa tridimensional pueden generarse antes o después de codificar con códigos de barras la pluralidad de objetivos en la muestra. En algunas realizaciones, el mapa bidimensional y el mapa tridimensional pueden generarse antes o después de lisar la muestra. Lisar la muestra antes o después de generar el mapa bidimensional o el mapa tridimensional puede incluir calentar la muestra, poner en contacto la muestra con un detergente, cambiar el pH de la muestra o cualquier combinación de los mismos.

En algunas realizaciones, codificar con códigos de barras la pluralidad de objetivos comprende hibridar una pluralidad de códigos de barras con una pluralidad de objetivos para crear objetivos codificados con códigos de barras (por ejemplo, objetivos codificados con códigos de barras estocástico). Codificar con códigos de barras la pluralidad de objetivos puede comprender generar una biblioteca indexada de los objetivos codificados con códigos de barras. Generar una biblioteca indexada de los objetivos codificados con códigos de barras puede realizarse con un soporte sólido que comprende la pluralidad de códigos de barras (por ejemplo, códigos de barras estocásticos).

Poner en contacto una muestra y un código de barras

La divulgación proporciona métodos para poner en contacto una muestra (por ejemplo, células) con un sustrato de la divulgación. Una muestra que comprende, por ejemplo, una sección delgada de célula, órgano o tejido, puede ponerse en contacto con códigos de barras (por ejemplo, códigos de barras estocásticos). Las células pueden ponerse en contacto, por ejemplo, mediante flujo por gravedad, en donde las células pueden asentarse y crear una monocapa. La muestra puede ser una sección delgada de tejido. La sección delgada puede colocarse sobre el sustrato. La muestra puede ser unidimensional (por ejemplo, forma una superficie plana). La muestra (por ejemplo, células) puede esparcirse por el sustrato, por ejemplo, haciendo crecer/cultivando las células sobre el sustrato.

Cuando los códigos de barras están muy cerca de los objetivos, los objetivos pueden hibridar con el código de barras. Los códigos de barras pueden ponerse en contacto en una proporción no agotable de tal manera que cada objetivo distinto pueda asociarse con un código de barras distinto de la divulgación. Para asegurar una asociación eficiente entre el objetivo y el código de barras, los objetivos pueden reticularse con el código de barras.

Lisis celular

Después de la distribución de las células y los códigos de barras, las células pueden lisarse para liberar las moléculas objetivo. La lisis celular puede lograrse mediante cualquiera de una variedad de medios, por ejemplo, mediante medios químicos o bioquímicos, por choque osmótico o por medio de lisis térmica, lisis mecánica o lisis óptica. Las células pueden lisarse mediante la adición de un tampón de lisis celular que comprende un detergente (por ejemplo, SDS, dodecil sulfato de Li, Triton X-100, Tween-20 o NP-40), un solvente orgánico (por ejemplo, metanol o acetona) o enzimas digestivas (por ejemplo, proteinasa K, pepsina o tripsina), o cualquier combinación de los mismos. Para aumentar la asociación de un objetivo y un código de barras, la tasa de difusión de las moléculas objetivo puede alterarse, por ejemplo, reduciendo la temperatura y/o aumentando la viscosidad del lisado.

En algunas realizaciones, la muestra puede lisarse usando un papel de filtro. El papel de filtro puede empaparse con un tampón de lisis encima del papel de filtro. El papel de filtro puede aplicarse a la muestra con presión, lo que puede facilitar la lisis de la muestra y la hibridación de los objetivos de la muestra con el sustrato.

En algunas realizaciones, la lisis puede realizarse mediante lisis mecánica, lisis por calor, lisis óptica y/o lisis química. La lisis química puede incluir el uso de enzimas digestivas como proteinasa K, pepsina y tripsina. La lisis puede realizarse mediante la adición de un tampón de lisis al sustrato. Un tampón de lisis puede comprender Tris HCl. Un tampón de lisis puede comprender por lo menos aproximadamente 0,01, 0,05, 0,1, 0,5 o 1 M o más de Tris HCl. Un tampón de lisis puede comprender como máximo aproximadamente 0,01, 0,05, 0,1, 0,5 o 1 M o más de Tris HCl. Un tampón de lisis puede comprender aproximadamente Tris HCl 0,1 M. El pH del tampón de lisis puede ser de por lo menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más. El pH del tampón de lisis puede ser como máximo de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más. En algunas realizaciones, el pH del tampón de lisis es de aproximadamente 7,5. El tampón de lisis puede comprender una sal (por ejemplo, LiCl). La concentración de sal en el tampón de lisis puede ser de por lo menos aproximadamente 0,1, 0,5 o 1 M o más. La concentración de sal en el tampón de lisis puede ser de como máximo aproximadamente 0,1, 0,5 o 1 M o más. En algunas realizaciones, la concentración de sal en el tampón de lisis es de aproximadamente 0,5 M. El tampón de lisis puede comprender un detergente (por ejemplo, SDS, dodecil sulfato de Li, triton X, tween, NP-40). La concentración del detergente en el tampón de lisis puede ser de por lo menos aproximadamente un 0,0001%, 0,0005%, 0,001%, 0,005%, 0,01%, 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 6% o 7%, o más. La concentración del detergente en el tampón de lisis puede ser como máximo de aproximadamente el 0,0001%, 0,0005%, 0,001%, 0,005%, 0,01%, 0,05%, 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 6% o 7%, o más. En algunas realizaciones, la concentración del detergente en el tampón de lisis es de aproximadamente el 1% de dodecil sulfato de Li. El tiempo usado en el método de lisis puede depender de la cantidad de detergente usada. En algunas realizaciones, cuanto más detergente se usa, menos tiempo se necesita para la lisis. El tampón de lisis puede comprender un agente quelante (por ejemplo, EDTA, EGTA). La concentración de un agente quelante en el tampón de lisis puede ser de por lo menos aproximadamente 1, 5, 10, 15, 20, 25 o 30 mM o más. La concentración de un agente quelante en el tampón de lisis puede ser como máximo de aproximadamente 1, 5, 10, 15, 20, 25 o 30 mM o más. En algunas realizaciones, la concentración de agente quelante en el tampón de lisis es de aproximadamente 10 mM. El tampón de lisis puede comprender un reactivo reductor (por ejemplo, beta-mercaptoetanol, DTT). La concentración del reactivo reductor en el tampón de lisis puede ser de por lo menos aproximadamente 1, 5, 10, 15 o 20 mM o más. La concentración del reactivo reductor en el tampón de lisis puede ser como máximo de aproximadamente 1, 5, 10, 15 o 20 mM o más. En algunas

realizaciones, la concentración de reactivo reductor en el tampón de lisis es de aproximadamente 5 mM. En algunas realizaciones, un tampón de lisis puede comprender TrisHCl aproximadamente 0,1 M, pH aproximadamente 7,5, LiCl aproximadamente 0,5 M, dodecilsulfato de litio aproximadamente al 1%, EDTA aproximadamente 10 mM y DTT aproximadamente 5 mM.

La lisis puede realizarse a una temperatura de aproximadamente 4, 10, 15, 20, 25 o 30° C. La lisis puede realizarse durante aproximadamente 1, 5, 10, 15 o 20 o más minutos. Una célula lisada puede comprender por lo menos aproximadamente 100000, 200000, 300000, 400000, 500000, 600000 o 700000 o más moléculas de ácidos nucleicos objetivo. Una célula lisada puede comprender como máximo aproximadamente 100000, 200000, 300000, 400000, 500000, 600000 o 700000 o más moléculas de ácidos nucleicos objetivo.

Unión de códigos de barras a moléculas de ácidos nucleicos objetivo

Después de la lisis de las células y la liberación de las moléculas de ácidos nucleicos de las mismas, las moléculas de ácidos nucleicos pueden asociarse aleatoriamente con los códigos de barras del soporte sólido colocalizado. La asociación puede comprender la hibridación de la región de reconocimiento de objetivo de un código de barras con una porción complementaria de la molécula de ácido nucleico objetivo (por ejemplo, el oligo(dT) del código de barras puede interactuar con una cola poli(A) de un objetivo). Las condiciones de ensayo usadas para la hibridación (por ejemplo, pH del tampón, fuerza iónica, temperatura, etc.) pueden elegirse para promover la formación de híbridos específicos y estables. En algunas realizaciones, las moléculas de ácidos nucleicos liberadas de las células lisadas pueden asociarse con la pluralidad de sondas en el sustrato (por ejemplo, hibridar con las sondas en el sustrato). Cuando las sondas comprenden oligo(dT), las moléculas de ARNm pueden hibridar con las sondas y transcribirse inversamente. La porción de oligo(dT) del oligonucleótido puede actuar como cebador para la síntesis de la primera cadena de la molécula de ADNc. Por ejemplo, en un ejemplo no limitativo de codificación con códigos de barras ilustrado en la FIG. 2, en el bloque 216, las moléculas de ARNm pueden hibridar con códigos de barras en perlas. Por ejemplo, los fragmentos de nucleótidos de cadena sencilla pueden hibridar con las regiones de unión al objetivo de los códigos de barras.

La unión puede comprender además la ligadura de la región de reconocimiento objetivo de un código de barras y una porción de la molécula de ácido nucleico objetivo. Por ejemplo, la región de unión al objetivo puede comprender una secuencia de ácidos nucleicos que puede ser capaz de una hibridación específica con un saliente del sitio de restricción (por ejemplo, un saliente del extremo cohesivo de EcoRI). El procedimiento de ensayo puede comprender además el tratamiento de los ácidos nucleicos objetivo con una enzima de restricción (por ejemplo, EcoRI) para crear un saliente de sitio de restricción. Luego, el código de barras puede ligarse a cualquier molécula de ácido nucleico que comprenda una secuencia complementaria al saliente del sitio de restricción. Para unir los dos fragmentos puede usarse una ligasa (por ejemplo, ADN ligasa T4).

Por ejemplo, en un ejemplo no limitativo de codificación con códigos de barras ilustrado en la FIG. 2, en el bloque 220, los objetivos marcados de una pluralidad de células (o una pluralidad de muestras) (por ejemplo, moléculas de código de barras objetivo) pueden agruparse posteriormente, por ejemplo, en un tubo. Los objetivos marcados pueden agruparse, por ejemplo, recuperando los códigos de barras y/o las perlas a las que se unen las moléculas de código de barras objetivo.

La recuperación de colecciones basadas en soportes sólidos de moléculas de código de barras objetivo unidas puede implementarse mediante el uso de perlas magnéticas y un campo magnético aplicado externamente. Una vez que se han agrupado las moléculas del código de barras objetivo, todo el procesamiento posterior puede continuar en un solo recipiente de reacción. El procesamiento adicional puede incluir, por ejemplo, reacciones de transcripción inversa, reacciones de amplificación, reacciones de escisión, reacciones de disociación y/o reacciones de extensión de ácidos nucleicos. Pueden realizarse más reacciones de procesamiento dentro de los micropocillos, es decir, sin agrupar primero las moléculas de ácidos nucleicos objetivo marcadas de una pluralidad de células.

Transcripción inversa

La divulgación proporciona un método para crear un conjugado de objetivo-código de barras usando transcripción inversa (por ejemplo, en el bloque 224 de la FIG. 2). El conjugado de objetivo-código de barras puede comprender el código de barras y una secuencia complementaria de todo o una parte del ácido nucleico objetivo (es decir, una molécula de ADNc codificada con códigos de barras, como una molécula de ADNc codificada con códigos de barras estocástico). La transcripción inversa de la molécula de ARN asociada puede producirse mediante la adición de un cebador de transcripción inversa junto con la transcriptasa inversa. El cebador de transcripción inversa puede ser un cebador oligo(dT), un cebador de hexanucleótidos aleatorio o un cebador oligonucleotídico específico del objetivo. Los cebadores oligo(dT) pueden tener, o pueden tener aproximadamente, 12-18 nucleótidos de longitud y se unen a la cola poli(A) endógena en el extremo 3' del ARNm de mamífero. Los cebadores de hexanucleótidos aleatorios pueden unirse al ARNm en una variedad de sitios complementarios. Los cebadores de oligonucleótidos específicos del objetivo típicamente ceban selectivamente el ARNm de interés.

En algunas realizaciones, la transcripción inversa de la molécula de ARN marcada puede producirse mediante la adición de un cebador de transcripción inversa. En algunas realizaciones, el cebador de transcripción inversa es un cebador oligo(dT), un cebador de hexanucleótidos aleatorio o un cebador oligonucleotídico específico del objetivo. En general, los cebadores oligo(dT) tienen una longitud de 12 a 18 nucleótidos y se unen a la cola poli(A) endógena en el extremo 3' del ARNm de mamífero. Los cebadores de hexanucleótidos aleatorios pueden unirse al ARNm en una variedad de sitios complementarios. Los cebadores de oligonucleótidos específicos del objetivo típicamente ceban selectivamente el ARNm de interés.

La transcripción inversa puede producirse repetidamente para producir múltiples moléculas de ADNc marcadas. Los métodos divulgados en la presente pueden comprender realizar por lo menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 reacciones de transcripción inversa. El método puede comprender realizar por lo menos aproximadamente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 reacciones de transcripción inversa.

Amplificación

Pueden realizarse una o más reacciones de amplificación de ácidos nucleicos (por ejemplo, en el bloque 228 de la FIG. 2) para crear múltiples copias de las moléculas de ácidos nucleicos objetivo marcadas. La amplificación puede realizarse de manera multiplexada, en la que múltiples secuencias de ácidos nucleicos objetivo se amplifican simultáneamente. La reacción de amplificación puede usarse para añadir adaptadores de secuenciación a las moléculas de ácidos nucleicos. Las reacciones de amplificación pueden comprender la amplificación de por lo menos una parte de un marcador de muestra, si lo hay. Las reacciones de amplificación pueden comprender la amplificación de por lo menos una parte del marcador celular y/o la secuencia del código de barras (por ejemplo, un marcador molecular). Las reacciones de amplificación pueden comprender amplificar por lo menos una parte de un marcador de muestra, un marcador celular, un marcador espacial, una secuencia de código de barras (por ejemplo, un marcador molecular), un ácido nucleico objetivo o una combinación de los mismos. Las reacciones de amplificación pueden comprender amplificar un 0,5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 100%, o un intervalo o un número entre dos cualquiera de estos valores, de la pluralidad de ácidos nucleicos. El método puede comprender además la realización de una o más reacciones de síntesis de ADNc para producir una o más copias de ADNc de moléculas de código de barras objetivo que comprenden un marcador de muestra, un marcador celular, un marcador espacial y/o una secuencia de código de barras (por ejemplo, un marcador molecular).

En algunas realizaciones, la amplificación puede realizarse mediante una reacción en cadena de la polimerasa (PCR). Como se usa en la presente, PCR puede referirse a una reacción para la amplificación in vitro de secuencias de ADN específicas mediante la extensión simultánea de cebadores de hebras complementarias de ADN. Como se usa en la presente, la PCR puede abarcar formas derivadas de la reacción, incluyendo pero no limitadas a, RT-PCR, PCR en tiempo real, PCR anidada, PCR cuantitativa, PCR multiplexada, PCR digital y PCR ensamblada.

La amplificación de los ácidos nucleicos marcados puede comprender métodos no basados en PCR. Los ejemplos de métodos no basados en PCR incluyen, pero no se limitan a, amplificación por desplazamiento múltiple (MDA), amplificación mediada por transcripción (TMA), amplificación basada en secuencias de ácidos nucleicos (NASBA), amplificación por desplazamiento de cadena (SDA), SDA en tiempo real, amplificación de círculo rodante o amplificación de círculo a círculo. Otros métodos de amplificación no basados en PCR incluyen ciclos múltiples de amplificación por transcripción de ARN impulsada por ARN polimerasa dependiente de ADN o síntesis y transcripción de ADN dirigida por ARN para amplificar objetivos de ADN o ARN, una reacción en cadena de ligasa (LCR) y un método de replicasa Q β (Q β), uso de sondas palindrómicas, amplificación por desplazamiento de cadena, amplificación dirigida por oligonucleótidos usando una endonucleasa de restricción, un método de amplificación en el que un cebador se hibrida con una secuencia de ácidos nucleicos y el dúplex resultante se escinde antes de la reacción de extensión y la amplificación, amplificación por desplazamiento de cadena usando una polimerasa de ácido nucleico que carece de actividad de exonucleasa 5', amplificación por círculo rodante y amplificación por extensión de ramificación (RAM). En algunas realizaciones, la amplificación no produce transcritos circulares.

En algunas realizaciones, los métodos divulgados en la presente comprenden además realizar una reacción en cadena de la polimerasa en el ácido nucleico marcado (por ejemplo, ARN marcado, ADN marcado, ADNc marcado) para producir un amplicón marcado (por ejemplo, un amplicón marcado estocásticamente). El amplicón marcado puede ser una molécula de cadena doble. La molécula de cadena doble puede comprender una molécula de ARN de cadena doble, una molécula de ADN de cadena doble o una molécula de ARN hibridada con una molécula de ADN. Una o ambas cadenas de la molécula de cadena doble pueden comprender un marcador de muestra, un marcador espacial, un marcador celular y/o una secuencia de código de barras (por ejemplo, un marcador molecular). El amplicón marcado puede ser una molécula de cadena sencilla. La molécula de cadena sencilla puede comprender ADN, ARN o una combinación de los mismos. Los ácidos nucleicos de la divulgación pueden comprender ácidos nucleicos sintéticos o alterados.

La amplificación puede comprender el uso de uno o más nucleótidos no naturales. Los nucleótidos no naturales pueden comprender nucleótidos fotolábiles o activables. Los ejemplos de nucleótidos no naturales pueden incluir, pero no se limitan a, ácido nucleico peptídico (PNA), morfolino y ácido nucleico bloqueado (LNA), así como ácido nucleico de glicol (GNA) y ácido nucleico de treosa (TNA). Pueden añadirse nucleótidos no naturales a uno o más ciclos de una reacción de amplificación. La adición de nucleótidos no naturales puede usarse para identificar productos como ciclos específicos o puntos temporales en la reacción de amplificación.

Realizar una o más reacciones de amplificación puede comprender el uso de uno o más cebadores. El uno o más cebadores pueden comprender, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 o más nucleótidos. El uno o más cebadores pueden comprender por lo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 o más nucleótidos. El uno o más cebadores pueden comprender menos de 12-15 nucleótidos. El uno o más cebadores pueden aparearse con por lo menos una parte de la pluralidad de objetivos marcados (por ejemplo, objetivos marcados estocásticamente). El uno o más cebadores pueden aparearse con el extremo 3' o el extremo 5' de la pluralidad de objetivos marcados. El uno o más cebadores pueden aparearse con una región interna de la pluralidad de objetivos marcados. La región interna puede tener por lo menos aproximadamente 50, 100, 150, 200, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540 550 560 570 580, 590, 600, 650, 700, 750, 800, 850, 900 o 1000 nucleótidos desde los extremos 3' de la pluralidad de objetivos marcados. El uno o más cebadores pueden comprender un panel fijo de cebadores. El uno o más cebadores pueden comprender por lo menos uno o más cebadores personalizados. El uno o más cebadores pueden comprender por lo menos uno o más cebadores de control. El uno o más cebadores pueden comprender por lo menos uno o más cebadores específicos de genes.

El uno o más cebadores pueden comprender un cebador universal. El cebador universal puede aparearse con un sitio de unión del cebador universal. El uno o más cebadores personalizados pueden aparearse con un primer marcador de muestra, un segundo marcador de muestra, un marcador espacial, un marcador celular, una secuencia de código de barras (por ejemplo, un marcador molecular), un objetivo o cualquier combinación de los mismos. El uno o más cebadores pueden comprender un cebador universal y un cebador personalizado. El cebador personalizado puede diseñarse para amplificar uno o más objetivos. Los objetivos pueden comprender un subconjunto de los ácidos nucleicos totales en una o más muestras. Los objetivos pueden comprender un subconjunto del total de objetivos marcados en una o más muestras. El uno o más cebadores pueden comprender por lo menos 96 o más cebadores personalizados. El uno o más cebadores pueden comprender por lo menos 960 o más cebadores personalizados. El uno o más cebadores pueden comprender por lo menos 9600 o más cebadores personalizados. El uno o más cebadores personalizados pueden aparearse con dos o más ácidos nucleicos marcados diferentes. Los dos o más ácidos nucleicos marcados diferentes pueden corresponder a uno o más genes.

En los métodos de la presente divulgación puede usarse cualquier esquema de amplificación. Por ejemplo, en un esquema, la primera ronda de PCR puede amplificar moléculas unidas a la perla usando un cebador específico de gen y un cebador contra la secuencia del cebador 1 de secuenciación universal de Illumina. La segunda ronda de PCR puede amplificar los primeros productos de PCR usando un cebador específico de gen anidado flanqueado por la secuencia del cebador 2 de secuenciación de Illumina y un cebador contra la secuencia del cebador 1 de secuenciación universal de Illumina. La tercera ronda de PCR añade P5 y P7 y el índice de muestra para convertir los productos de PCR en una biblioteca de secuenciación de Illumina. La secuenciación usando secuenciación de 150 pb x 2 puede revelar el marcador celular y la secuencia del código de barras (por ejemplo, marcador molecular) en la lectura 1, el gen en la lectura 2 y el índice de muestra en la lectura del índice 1.

En algunas realizaciones, los ácidos nucleicos pueden eliminarse del sustrato usando escisión química. Por ejemplo, puede usarse un grupo químico o una base modificada presente en un ácido nucleico para facilitar su eliminación de un soporte sólido. Por ejemplo, puede usarse una enzima para eliminar un ácido nucleico de un sustrato. Por ejemplo, un ácido nucleico puede eliminarse de un sustrato a través de una digestión con endonucleasa de restricción (a la que también puede hacerse referencia en la presente como "enzima de restricción"). Por ejemplo, puede usarse el tratamiento de un ácido nucleico que contiene un dUTP o ddUTP con uracil-d-glicosilasa (UDG) para eliminar un ácido nucleico de un sustrato. Por ejemplo, un ácido nucleico puede eliminarse de un sustrato usando una enzima que realiza la escisión de nucleótidos, como una enzima reparadora de escisión de bases, como una endonucleasa apurínica/apirimidínica (AP). En algunas realizaciones, un ácido nucleico puede eliminarse de un sustrato usando un grupo fotoescindible y luz. En algunas realizaciones, para eliminar un ácido nucleico del sustrato puede usarse un conector escindible. Por ejemplo, el conector escindible puede comprender por lo menos uno de biotina/avidina, biotina/estreptavidina, biotina/neutravidina, Ig-proteína A, un conector fotolábil, un grupo conector lábil a ácidos o bases, o un aptámero.

Cuando las sondas son específicas de genes, las moléculas pueden hibridar con las sondas y ser transcritas y/o amplificadas de manera inversa. En algunas realizaciones, después de que se haya sintetizado el ácido nucleico (por ejemplo, transcrito inverso), puede amplificarse. La amplificación puede realizarse de manera multiplex, en la que múltiples secuencias de ácido nucleico objetivo se amplifican simultáneamente. La amplificación puede añadir adaptadores de secuenciación al ácido nucleico.

En algunas realizaciones, la amplificación puede realizarse en el sustrato, por ejemplo, con amplificación puente. Los ADNc pueden tener una cola de homopolímero para generar un extremo compatible para la amplificación puente usando sondas de oligo(dT) en el sustrato. En la amplificación puente, el cebador que es complementario al extremo 3' del ácido nucleico plantilla puede ser el primer cebador de cada par que se une covalentemente a la partícula sólida. Cuando una muestra que contiene el ácido nucleico plantilla se pone en contacto con la partícula y se realiza un único ciclo térmico, la molécula plantilla puede aparearse con el primer cebador y el primer cebador se alarga en la dirección directa mediante la adición de nucleótidos para formar una molécula dúplex que consiste de la molécula plantilla y una cadena de ADN recién formada que es complementaria a la plantilla. En el paso de calentamiento del siguiente ciclo, puede desnaturalizarse la molécula dúplex, liberando la molécula plantilla de la partícula y dejando la cadena de ADN complementaria unida a la partícula a través del primer cebador. En la etapa de apareamiento del paso de apareamiento y elongación que sigue, la cadena complementaria puede hibridar con el segundo cebador, que es complementario a un segmento de la cadena complementaria en una localización eliminada del primer cebador. Esta hibridación puede hacer que la cadena complementaria forme un puente entre el primer y el segundo cebadores asegurado al primer cebador por un enlace covalente y al segundo cebador por hibridación. En la etapa de elongación, el segundo cebador puede alargarse en dirección inversa mediante la adición de nucleótidos en la misma mezcla de reacción, convirtiendo de este modo el puente en un puente de cadena doble. Entonces comienza el siguiente ciclo, y el puente de cadena doble puede desnaturalizarse para producir dos moléculas de ácidos nucleicos de cadena sencilla, teniendo cada una un extremo unido a la superficie de la partícula a través del primer y el segundo cebadores, respectivamente, con el otro extremo de cada uno sin unir. En el paso de apareamiento y elongación de este segundo ciclo, cada cadena puede hibridar con un cebador complementario adicional, no usado con anterioridad, en la misma partícula, para formar nuevos puentes de cadena sencilla. Los dos cebadores no usados con anterioridad que ahora están hibridados se alargan para convertir los dos nuevos puentes en puentes de cadena doble.

Las reacciones de amplificación pueden comprender amplificar por lo menos un 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97% o 100% de la pluralidad de ácidos nucleicos ácidos.

La amplificación de los ácidos nucleicos marcados puede comprender métodos basados en PCR o métodos no basados en PCR. La amplificación de los ácidos nucleicos marcados puede comprender la amplificación exponencial de los ácidos nucleicos marcados. La amplificación de los ácidos nucleicos marcados puede comprender la amplificación lineal de los ácidos nucleicos marcados. La amplificación puede realizarse por reacción en cadena de la polimerasa (PCR). La PCR puede referirse a una reacción para la amplificación in vitro de secuencias de ADN específicas mediante la extensión simultánea de cebadores de cadenas complementarias de ADN. La PCR puede abarcar formas derivadas de la reacción, incluyendo pero no limitadas a, RT-PCR, PCR en tiempo real, PCR anidada, PCR cuantitativa, PCR multiplexada, PCR digital, PCR de supresión, PCR semisupresora y PCR de ensamblaje.

En algunas realizaciones, la amplificación de los ácidos nucleicos marcados comprende métodos no basados en PCR. Los ejemplos de métodos no basados en PCR incluyen, pero no se limitan a, amplificación por desplazamiento múltiple (MDA), amplificación mediada por transcripción (TMA), amplificación basada en secuencias de ácidos nucleicos (NASBA), amplificación por desplazamiento de cadena (SDA), SDA en tiempo real, amplificación de círculo rodante o amplificación de círculo a círculo. Otros métodos de amplificación no basados en PCR incluyen ciclos múltiples de amplificación de transcripción de ARN impulsada por ARN polimerasa dependiente de ADN o síntesis y transcripción de ADN dirigida por ARN para amplificar objetivos de ADN o ARN, una reacción en cadena de ligasa (LCR), una replicasa Q β (Q β), uso de sondas palindrómicas, amplificación por desplazamiento de cadena, amplificación dirigida por oligonucleótidos usando una endonucleasa de restricción, un método de amplificación en el que un cebador se hibrida con una secuencia de ácidos nucleicos y el dúplex resultante se escinde antes de la reacción de extensión y amplificación, amplificación por desplazamiento de cadena usando una polimerasa de ácido nucleico que carece de actividad de exonucleasa 5', amplificación por círculo rodante y/o amplificación por extensión de ramificación (RAM).

En algunas realizaciones, los métodos divulgados en la presente comprenden además la realización de una reacción en cadena de la polimerasa anidada en el amplicón amplificado (por ejemplo, el objetivo). El amplicón puede ser una molécula de cadena doble. La molécula de cadena doble puede comprender una molécula de ARN de cadena doble, una molécula de ADN de cadena doble o una molécula de ARN hibridada con una molécula de ADN. Una o ambas cadenas de la molécula de cadena doble pueden comprender un marcador de muestra o un marcador de identificador molecular. Alternativamente, el amplicón puede ser una molécula de cadena sencilla. La molécula de cadena sencilla puede comprender ADN, ARN o una combinación de los mismos. Los ácidos nucleicos descritos en la presente pueden comprender ácidos nucleicos sintéticos o alterados.

En algunas realizaciones, el método comprende amplificar repetidamente el ácido nucleico marcado para producir múltiples amplicones. Los métodos divulgados en la presente pueden comprender realizar por lo menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 reacciones de amplificación.

Alternativamente, el método comprende realizar por lo menos aproximadamente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 reacciones de amplificación.

La amplificación puede comprender además añadir uno o más ácidos nucleicos de control a una o más muestras que comprenden una pluralidad de ácidos nucleicos. La amplificación puede comprender además añadir uno o más ácidos nucleicos de control a una pluralidad de ácidos nucleicos. Los ácidos nucleicos de control pueden comprender un marcador de control.

La amplificación puede comprender usar uno o más nucleótidos no naturales. Los nucleótidos no naturales pueden comprender nucleótidos fotolábiles y/o activables. Los ejemplos de nucleótidos no naturales incluyen, pero no se limitan a, ácido nucleico peptídico (PNA), morfolino y ácido nucleico bloqueado (LNA), así como ácido nucleico de glicol (GNA) y ácido nucleico de treosa (TNA). A uno o más ciclos de una reacción de amplificación pueden añadirse nucleótidos no naturales. La adición de nucleótidos no naturales puede usarse para identificar productos como ciclos específicos o puntos temporales en la reacción de amplificación.

La realización de una o más reacciones de amplificación puede comprender el uso de uno o más cebadores. El uno o más cebadores pueden comprender uno o más oligonucleótidos. El uno o más oligonucleótidos pueden comprender por lo menos aproximadamente 7-9 nucleótidos. El uno o más oligonucleótidos pueden comprender menos de 12-15 nucleótidos. El uno o más cebadores pueden aparearse con por lo menos una parte de la pluralidad de ácidos nucleicos marcados. El uno o más cebadores pueden aparearse con el extremo 3' y/o el extremo 5' de la pluralidad de ácidos nucleicos marcados. El uno o más cebadores pueden aparearse con una región interna de la pluralidad de ácidos nucleicos marcados. La región interna puede tener por lo menos 50, 100, 150, 200, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 650, 700, 750, 800, 850, 900 o 1000 nucleótidos desde los extremos 3' de la pluralidad de ácidos nucleicos marcados. El uno o más cebadores pueden comprender un panel fijo de cebadores. El uno o más cebadores pueden comprender por lo menos uno o más cebadores personalizados. El uno o más cebadores pueden comprender por lo menos uno o más cebadores de control. El uno o más cebadores pueden comprender por lo menos uno o más cebadores de genes constitutivos. El uno o más cebadores pueden comprender un cebador universal. El cebador universal puede aparearse con un sitio de unión del cebador universal. El uno o más cebadores personalizados pueden aparearse con la primera etiqueta de muestra, la segunda etiqueta de muestra, el marcador identificador molecular, el ácido nucleico o un producto de los mismos. El uno o más cebadores pueden comprender un cebador universal y un cebador personalizado. El cebador personalizado puede diseñarse para amplificar uno o más ácidos nucleicos objetivo. Los ácidos nucleicos objetivo pueden comprender un subconjunto de los ácidos nucleicos totales en una o más muestras. En algunas realizaciones, los cebadores son las sondas unidas a la matriz de la divulgación.

En algunas realizaciones, la codificación con códigos de barras (por ejemplo, codificación con códigos de barras estocásticos) de la pluralidad de objetivos en la muestra comprende además generar una biblioteca indexada de los objetivos codificados con códigos de barras (por ejemplo, objetivos codificados con códigos de barras estocásticos) o fragmentos codificados con códigos de barras de los objetivos. Las secuencias de códigos de barras de diferentes códigos de barras (por ejemplo, los marcadores moleculares de diferentes códigos de barras estocásticos) pueden ser diferentes entre sí. La generación de una biblioteca indexada de los objetivos codificados con códigos de barras incluye la generación de una pluralidad de polinucleótidos indexados a partir de la pluralidad de objetivos en la muestra. Por ejemplo, para una biblioteca indexada de objetivos codificados con códigos de barras que comprende un primer objetivo indexado y un segundo objetivo indexado, la región marcada del primer polinucleótido indexado puede diferir de la región marcada del segundo polinucleótido indexado en, aproximadamente, por lo menos, o como máximo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, o un número o un intervalo entre dos cualquiera de estos valores, nucleótidos. En algunas realizaciones, generar una biblioteca indexada de los objetivos codificados con códigos de barras incluye poner en contacto una pluralidad de objetivos, por ejemplo, moléculas de ARNm, con una pluralidad de oligonucleótidos que incluyen una región poli(T) y una región marcadora; y llevar a cabo una síntesis de la primera cadena usando una transcriptasa inversa para producir moléculas de ADNc marcadas de cadena sencilla, cada una de las cuales comprende una región de ADNc y una región marcadora, en donde la pluralidad de objetivos incluye por lo menos dos moléculas de ARNm de secuencias diferentes y la pluralidad de oligonucleótidos incluye por lo menos dos oligonucleótidos de secuencias diferentes. Generar una biblioteca indexada de los objetivos codificados con códigos de barras puede comprender además amplificar las moléculas de ADNc marcadas de cadena sencilla para producir moléculas de ADNc marcadas de cadena doble; y llevar a cabo una PCR anidada en las moléculas de ADNc marcadas de cadena doble para producir amplicones marcados. En algunas realizaciones, el método puede incluir generar un amplicón marcado con adaptador.

La codificación con códigos de barras (por ejemplo, codificación con códigos de barras estocásticos) puede incluir el uso de etiquetas o códigos de barras de ácido nucleico para marcar moléculas individuales de ácido nucleico (por ejemplo, ADN o ARN). En algunas realizaciones, implica añadir etiquetas o códigos de barras de ADN a las moléculas de ADNc a medida que se generan a partir del ARNm. Puede realizarse una PCR anidada para minimizar el sesgo de amplificación de la PCR. Pueden añadirse adaptadores para secuenciar usando, por ejemplo,

secuenciación de próxima generación (NGS). Los resultados de la secuenciación pueden usarse para determinar marcadores celulares, marcadores moleculares y secuencias de fragmentos de nucleótidos de una o más copias de los objetivos, por ejemplo, en el bloque 232 de la FIG. 2.

La FIG. 3 es una ilustración esquemática que muestra un proceso ejemplar no limitativo para generar una biblioteca indexada de los objetivos codificados con códigos de barras (por ejemplo, objetivos codificados con códigos de barras estocásticos), como ARNm codificados con códigos de barras o fragmentos de los mismos. Como se muestra en el paso 1, el proceso de transcripción inversa puede codificar cada molécula de ARNm con una secuencia de marcador molecular única, una secuencia de marcador celular y un sitio de PCR universal. En particular, las moléculas de ARN 302 se pueden transcribir inversamente para producir moléculas de ADNc 304 marcadas, incluyendo una región de ADNc 306, mediante hibridación (por ejemplo, hibridación estocástica) de un conjunto de códigos de barras (por ejemplo, códigos de barras estocásticos) 310 con la región de cola poli(A) 308 de las moléculas de ARN 302. Cada uno de los códigos de barras 310 puede comprender una región de unión al objetivo, por ejemplo, una región poli(dT) 312, una región marcadora 314 (por ejemplo, una secuencia de código de barras o una molécula), y una región de PCR universal 316.

En algunas realizaciones, la secuencia del marcador celular puede incluir de 3 a 20 nucleótidos. En algunas realizaciones, la secuencia del marcador molecular puede incluir de 3 a 20 nucleótidos. En algunas realizaciones, cada uno de la pluralidad de códigos de barras estocásticos comprende además uno o más de un marcador universal y un marcador celular, en donde los marcadores universales son los mismos para la pluralidad de códigos de barras estocásticos en el soporte sólido y los marcadores celulares son los mismos para la pluralidad de códigos de barras estocásticos sobre el soporte sólido. En algunas realizaciones, el marcador universal puede incluir de 3 a 20 nucleótidos. En algunas realizaciones, el marcador celular comprende de 3 a 20 nucleótidos.

En algunas realizaciones, la región del marcador 314 puede incluir una secuencia de código de barras o un marcador molecular 318 y un marcador celular 320. En algunas realizaciones, la región del marcador 314 puede incluir uno o más de un marcador universal, un marcador de dimensión y un marcador celular. La secuencia de código de barras o marcador molecular 318 puede tener, puede tener aproximadamente, puede tener por lo menos o puede tener como máximo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, o un número o un intervalo entre cualquiera de estos valores, nucleótidos de longitud. El marcador celular 320 puede tener, puede tener aproximadamente, puede tener por lo menos o puede tener como máximo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, o un número o un intervalo entre cualquiera de estos valores, nucleótidos de longitud. El marcador universal puede tener, puede tener aproximadamente, puede tener por lo menos, o puede tener como máximo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, o un número o un intervalo entre cualquiera de estos valores, nucleótidos de longitud. Los marcadores universales pueden ser los mismos para la pluralidad de códigos de barras estocásticos sobre el soporte sólido y los marcadores celulares son los mismos para la pluralidad de códigos de barras estocásticos sobre el soporte sólido. El marcador de dimensión puede tener, puede tener aproximadamente, puede tener por lo menos o puede tener como máximo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, o un número o un intervalo entre cualquiera de estos valores, nucleótidos de longitud.

En algunas realizaciones, la región del marcador 314 puede comprender, comprender aproximadamente, comprender por lo menos o comprender como máximo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o un intervalo entre cualquiera de estos valores, marcadores diferentes, como una secuencia de código de barras o un marcador molecular 318 y un marcador celular 320. Cada marcador puede tener, puede tener aproximadamente, puede tener por lo menos o puede tener como máximo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, o un número o un intervalo entre cualquiera de estos valores, nucleótidos de longitud. Un conjunto de códigos de barras o códigos de barras estocásticos 310 puede contener, contener aproximadamente, contener por lo menos, o puede tener como máximo, 10, 20, 40, 50, 70, 80, 90, 10², 10³, 10⁴, 10⁵, 10⁶, 10⁷, 10⁸, 10⁹, 10¹⁰, 10¹¹, 10¹², 10¹³, 10¹⁴, 10¹⁵, 10²⁰, o un número o intervalo entre cualquiera de estos valores, códigos de barras o códigos de barras estocásticos 310. Y el conjunto de códigos de barras o códigos de barras estocásticos 310 puede, por ejemplo, contener cada uno una región de marcador único 314. Las moléculas de ADNc marcadas 304 pueden purificarse para eliminar el exceso de códigos de barras o códigos de barras estocásticos 310. La purificación puede comprender la purificación con perlas Ampure.

Como se muestra en el paso 2, los productos del proceso de transcripción inversa en el paso 1 pueden agruparse en 1 tubo y amplificarse por PCR con un primer grupo de cebadores de PCR y un primer cebador de PCR universal. El agrupamiento es posible debido a la región de marcador única 314. En particular, las moléculas de ADNc marcadas 304 pueden amplificarse para producir amplicones marcados con PCR anidados 322. La amplificación puede comprender amplificación por PCR multiplex. La amplificación puede comprender una amplificación por PCR multiplex con 96 cebadores multiplex en un único volumen de reacción. En algunas realizaciones, la amplificación por PCR multiplex puede utilizar, utilizar aproximadamente, utilizar por lo menos o utilizar como máximo 10, 20, 40, 50, 70, 80, 90, 10², 10³, 10⁴, 10⁵, 10⁶, 10⁷, 10⁸, 10⁹, 10¹⁰, 10¹¹, 10¹², 10¹³, 10¹⁴, 10¹⁵, 10²⁰, o un número o un intervalo entre dos cualquiera de estos valores, cebadores multiplexados en un único volumen de reacción. La amplificación puede comprender usar un primer grupo de cebadores de PCR 324 que comprende cebadores personalizados 326A-C dirigidos a genes específicos y un cebador universal 328. Los

cebadores personalizados 326 pueden hibridar con una región dentro de la porción de ADNc 306' de la molécula de ADNc marcada 304. El cebador universal 328 puede hibridar con la región PCR universal 316 de la molécula de ADNc marcada 304.

Como se muestra en el paso 3 de la FIG. 3, los productos de la amplificación por PCR en el paso 2 pueden amplificarse con un grupo de cebadores de PCR anidados y un segundo cebador de PCR universal. La PCR anidada puede minimizar el sesgo de amplificación de la PCR. En particular, los amplicones marcados con PCR anidada 322 pueden amplificarse adicionalmente mediante PCR anidada. La PCR anidada puede comprender PCR multiplex con un grupo de cebadores de PCR anidados 330 de cebadores de PCR anidados 332a-c y un segundo cebador de PCR universal 328' en un único volumen de reacción. El grupo de cebadores de PCR anidados 328 puede contener, contener aproximadamente, contener por lo menos o contener como máximo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o un intervalo entre cualquiera de estos valores, diferentes cebadores de PCR anidados 330. Los cebadores de PCR anidados 332 pueden contener un adaptador 334 e hibridar con una región dentro de la porción de ADNc 306" del amplicón marcado 322. El cebador universal 328' puede contener un adaptador 336 e hibridar con la región de PCR universal 316 del amplicón marcado 322. Por tanto, el paso 3 produce el amplicón marcado con adaptador 338. En algunas realizaciones, los cebadores de PCR anidados 332 y el segundo cebador de PCR universal 328' pueden no contener los adaptadores 334 y 336. En su lugar, los adaptadores 334 y 336 pueden ligarse a los productos de la PCR anidada para producir el amplicón marcado con adaptador 338.

Como se muestra en el paso 4, los productos de PCR del paso 3 pueden amplificarse mediante PCR para la secuenciación usando cebadores de amplificación de bibliotecas. En particular, los adaptadores 334 y 336 pueden usarse para realizar uno o más ensayos adicionales en el amplicón marcado con adaptador 338. Los adaptadores 334 y 336 pueden hibridar con los cebadores 340 y 342. El uno o más cebadores 340 y 342 pueden ser cebadores de amplificación por PCR. Los uno o más cebadores 340 y 342 pueden ser cebadores de secuenciación. El uno o más adaptadores 334 y 336 pueden usarse para amplificación adicional de los amplicones marcados con adaptador 338. El uno o más adaptadores 334 y 336 pueden usarse para secuenciar el amplicón marcado con adaptador 338. El cebador 342 puede contener un índice de placa 344 de tal manera que los amplicones generados usando el mismo conjunto de códigos de barras o códigos de barras estocásticos 310 pueden secuenciarse en una reacción de secuenciación usando secuenciación de próxima generación (NGS).

Análisis multiómico

En la presente se divulgan realizaciones de un método para el análisis de muestras de alto rendimiento. El método puede usarse con cualquier plataforma o sistema de análisis de muestras para dividir células individuales con partículas individuales, como plataformas y sistemas basados en gotitas (por ejemplo, Solución 3' de células individuales Chromium™ (10X Genomics (San Francisco, CA))), micropocillos (por ejemplo, ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ))), cámaras microfluídicas y sustratos con patrones. El método puede capturar información multiómica, incluyendo el genoma, la accesibilidad genómica (por ejemplo, accesibilidad a la cromatina) y el metiloma. El método puede usarse con métodos para análisis transcriptómicos, análisis proteómicos y/o seguimiento de muestras. El uso de codificación con códigos de barras para el análisis proteómico se ha descrito en Solicitud de Estados Unidos Nº 15/715028, publicada como US 2018/0088112. El uso de codificación con códigos de barras para el seguimiento de muestras se ha descrito en Solicitud de Estados Unidos Nº 15/937,713, publicada como US 2018/0346970. En algunas realizaciones, la información multiómica, como la genómica, la accesibilidad de la cromatina, la metilómica, la transcriptómica y la proteómica, de células individuales puede obtenerse usando codificación con códigos de barras.

En algunas realizaciones, el método incluye agregar una secuencia complementaria a la de las sondas de captura con marcadores o índices celulares y moleculares al final de los fragmentos de ADN genómico. Por ejemplo, puede añadirse una cola de poli(dA) (o cualquier secuencia) a los fragmentos genómicos de tal manera que puedan ser capturados por sondas de oligo(dT) (o una secuencia complementaria a la secuencia añadida) flanqueada con códigos de barras celulares y moleculares. El método puede usarse para capturar todo o parte de lo siguiente a partir de células individuales con un alto rendimiento, incluyendo el genoma, el metiloma, la accesibilidad a la cromatina, el transcriptoma y el proteoma.

El método puede incluir la preparación de muestras antes de cargar materiales celulares en cualquiera de estos sistemas de análisis de muestras. Por ejemplo, utilizando cortadores enzimáticos, por ejemplo, nucleasas de cadena doble como se describe en la presente (como transposasa, enzimas de restricción y proteínas asociadas a CRISPR), el ADNc (por ejemplo, ADNg) puede fragmentarse en fragmentos genómicos dentro de células o núcleos fijados. Puede usarse una enzima de restricción para el análisis de muestras multiómicas de alto rendimiento. Por ejemplo, el método puede incluir incubar células con enzimas de restricción (seguido de la eliminación de la enzima de restricción, por ejemplo). Como otro ejemplo, el método puede incluir incubar células con una ligasa y adaptadores con poli(dT)/poli(dA) o poli(dT)/poli(dA) con secuencias del promotor T7 flanqueadas con la secuencia de restricción. Como otro ejemplo más, la sonda de captura puede tener una secuencia del sitio de restricción. En esta realización, puede no ser necesaria la adición de adaptadores dTs/dAs. En algunas realizaciones, puede usarse

Cas9/CRISPR para cortar en localizaciones específicas del genoma.

Las células o núcleos pueden ser frescos o fijados (por ejemplo, células fijadas con fijadores como aldehídos, agentes oxidantes, fijadores con efecto de protección de solvente orgánico mediado por tampón de hepes-ácido glutámico (HOPE)). En algunas realizaciones, el método comprende poner en contacto las células con un reactivo de ácido nucleico como se describe en la presente. Luego, las células pueden lavarse para eliminar el exceso de reactivo de ácido nucleico. Como se describe en la presente, el reactivo de ácido nucleico puede unirse a ADNdc en células muertas, pero no en células vivas, de tal manera que solo las células muertas permanecerán marcadas con el reactivo de ácido nucleico después del lavado. En algunas realizaciones, se agrega luego una secuencia complementaria a las sondas de captura (por ejemplo, códigos de barras como códigos de barras estocásticos) a cada extremo de los fragmentos genómicos. Las sondas de captura pueden anclarse en un soporte sólido o en solución. La sonda de captura de un sistema de análisis transcriptómico de células individuales puede ser una secuencia poli(dT). Por tanto, a cada extremo de los fragmentos genómicos se le puede añadir una secuencia poli(dA). Luego, las células o los núcleos pueden calentarse o exponerse a productos químicos para desnaturalizar los fragmentos genómicos de cadena doble con una secuencia poli(dA) agregada en cada extremo y luego cargarse en un sistema de análisis de muestras. Tras la lisis celular y/o del núcleo, los fragmentos genómicos con la secuencia agregada pueden ser capturados por las sondas de captura presentes, al igual que las moléculas de ARNm con colas poli(A) pueden ser capturadas por secuencias poli(dT) de sondas de captura. Puede añadirse transcriptasa inversa y/o ADN polimerasa para copiar (por ejemplo, transcribir de manera inversa) los fragmentos genómicos y agregar los marcadores o índices celulares y moleculares a los fragmentos genómicos.

En la presente se divulgan realizaciones de un método de análisis de muestras. Las FIGS. 4A-4B muestran una ilustración esquemática de realizaciones ejemplares no limitativas de un método 400 de captura de alto rendimiento de información multiómica de células individuales. En algunas realizaciones, el método 400 incluye usar un transposoma para generar fragmentos de ADN de cadena doble con salientes 5' (o salientes 3') que comprenden una secuencia de captura. El método 400 puede incluir: poner en contacto 410 ácido desoxirribonucleico de cadena doble (ADNdc), por ejemplo, un ADN genómico (ADNg), con un transposoma 428. El transposoma 428 puede comprender una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc 430 y dos copias 432a, 432b de un adaptador que tiene un saliente 5' que comprende una secuencia de captura (por ejemplo, una secuencia poli(dT) 434a, 434b). La nucleasa de cadena doble 430 puede cargarse con las dos copias 432a y 432b del adaptador 434a, 434b. Cada copia 436a, 436b del adaptador puede comprender una secuencia final de ADN del transposón (por ejemplo, una secuencia Tn5 436a, 436b o una subsecuencia de la misma). La nucleasa de cadena doble puede ser, o comprender, una transposasa como la transposasa Tn5. Poner en contacto 410 ADNdc (por ejemplo, ADNg) con un transposoma 428 puede generar una pluralidad de fragmentos de ADNdc salientes 438 cada uno con dos copias 432a, 432b de los salientes 5' 434a, 434b.

En algunas realizaciones, el método 400 incluye poner en contacto (por ejemplo, en el bloque 412) la pluralidad de fragmentos de ADNdc salientes (con los salientes 5') 438 con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria 434a', 434b' a por lo menos una parte del saliente 5' 434a, 434b. El método 400 puede incluir desnaturalizar (por ejemplo, en el bloque 414) la pluralidad de fragmentos de ADNdc complementarios 440, cada uno de los cuales comprende la secuencia complementaria de por lo menos una parte del saliente 5' para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc) 442, y codificar con códigos de barras (por ejemplo, en el bloque 424) la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras 444 para generar una pluralidad de fragmentos de ADNmc codificados con códigos de barras (por ejemplo, fragmentos de ADNmc codificados con códigos de barras 446 o una secuencia complementaria de los mismos). Por lo menos algunos (por ejemplo, por lo menos 1, 2, 3, 4, 5, 6, 7, 8, 10, 100, 1000, 10000, 100000, 1000000, 10000000 o más) de la pluralidad de códigos de barras 444 comprenden un marcador celular 448, un marcador molecular 450, y la secuencia de captura 434. Los marcadores moleculares 448 de por lo menos dos códigos de barras de la pluralidad de códigos de barras 444 pueden comprender diferentes secuencias de marcadores moleculares. Por lo menos dos códigos de barras de la pluralidad de códigos de barras 444 pueden comprender marcadores celulares 450 con una secuencia de marcadores celulares idéntica. El método 400 puede incluir obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras 446 (o una secuencia complementaria de los mismos), y determinar la información relacionada con el ADNdc (por ejemplo, ADNg) sobre la base de las secuencias de la pluralidad de fragmentos de ADNmc 446 (o una secuencia complementaria de los mismos) en los datos de secuenciación obtenidos.

El método 400 puede incluir usar un transposoma 428 (que puede comprender, por ejemplo, una transposasa, una endonucleasa de restricción y/o una proteína asociada a CRISPR como Cas9 o Cas12a) para generar fragmentos de ADN a partir del ADN genómico de una célula. En algunas realizaciones, el método 400 puede incluir: generar una pluralidad de fragmentos de ácidos nucleicos a partir de ácido desoxirribonucleico de cadena doble (ADNdc), por ejemplo, ADNg, de una célula. Por ejemplo, la pluralidad de fragmentos de ácidos nucleicos puede no generarse a partir de la amplificación. Como otro ejemplo, la pluralidad de fragmentos de ácidos nucleicos puede ser, o incluir, moléculas de ARN producidas por transcripción in vitro.

En algunas realizaciones, cada uno de la pluralidad de fragmentos de ácidos nucleicos puede comprender una secuencia de captura 434a, 434b, un complemento de la secuencia de captura, un complemento inverso de la secuencia de captura o una combinación de los mismos. El método 400 puede incluir codificar con códigos de barras 424 la pluralidad de fragmentos de ácidos nucleicos usando la pluralidad de códigos de barras 444 para generar una pluralidad de fragmentos 446 de ácido desoxirribonucleico de cadena sencilla (ADNmc) codificados con códigos de barras (o una secuencia complementaria de los mismos, como un complemento o un complemento inverso 446). Por lo menos algunos (por ejemplo, por lo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 100, 1000, 10000, 100000, 1000000, 10000000 o más) de la pluralidad de códigos de barras 444 pueden comprender un marcador celular 450, un marcador molecular 448 y la secuencia de captura 434 (o un complemento de la secuencia de captura, un complemento inverso de la secuencia de captura, o una combinación de los mismos). Los marcadores moleculares 448 de por lo menos dos códigos de barras de la pluralidad de códigos de barras 444 comprenden diferentes secuencias de marcadores moleculares. Por lo menos dos códigos de barras de la pluralidad de códigos de barras 444 pueden comprender marcadores celulares 450 con una secuencia de marcadores celulares idéntica. El método 400 puede incluir obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras 446 (o una secuencia complementaria de los mismos); y determinar la información relativa al ADNdc (por ejemplo, ADNg) basándose en las secuencias de la pluralidad de fragmentos de ADNmc 445 en los datos de secuenciación obtenidos.

En algunas realizaciones, el ADNdc (por ejemplo, ADNg) está dentro de un núcleo 452. El método 400 puede incluir opcionalmente permeabilizar (por ejemplo, en el bloque 402) un núcleo 452 para generar un núcleo permeabilizado. El método 400 puede incluir opcionalmente fijar una célula (por ejemplo, en el bloque 402) que comprende el núcleo 452 antes de permeabilizar el núcleo.

En algunas realizaciones, el método 400 comprende desnaturalizar 414 la pluralidad de fragmentos de ácidos nucleicos 440 para generar una pluralidad de fragmentos de ADNmc 442. Codificar con códigos de barras 424 la pluralidad de fragmentos de ácidos nucleicos puede comprender codificar con códigos de barras 424 la pluralidad de fragmentos de ADNmc 442 usando la pluralidad de códigos de barras 444 para generar la pluralidad de fragmentos 446 de ADNmc codificados con códigos de barras y/o secuencias complementarias de los mismos.

En algunas realizaciones, para cualquier método de análisis de muestras como se describe en la presente, el método comprende además poner en contacto una célula con un reactivo de ácido nucleico como se describe en la presente. El reactivo de ácido nucleico puede comprender una secuencia de captura, un código de barras, un sitio de unión del cebador y un agente de unión al ADN de cadena doble. La célula puede ser una célula muerta. El reactivo de ácido nucleico puede unirse al ADN de cadena doble en la célula muerta. El método puede comprender además lavar la célula para eliminar el exceso del reactivo de ácido nucleico. El método puede comprender además lisar la célula, liberando de este modo el reactivo de ácido nucleico. El método puede comprender además codificar con código de barras el reactivo de ácido nucleico. Se contempla que las células muertas sean permeables al reactivo de ácido nucleico, mientras que las células vivas no lo sean, o sean permeables a no más que cantidades traza del reactivo de ácido nucleico. Por consiguiente, se contempla que el método descrito en la presente pueda identificar células muertas y vivas identificando si el reactivo de ácido nucleico se ha unido al ADN de una célula (por ejemplo, determinando si un código de barras asociado con el reactivo de ácido nucleico está asociado con la célula) y/o determinando si por lo menos un número umbral de reactivos de ácido nucleico se ha unido a la célula (por ejemplo, determinando si por lo menos un el recuento umbral de códigos de barras está asociado con el reactivo de ácido nucleico que está asociado con la célula, por ejemplo, por lo menos 10, 50, 100, 500, 1000, 5000 o 10000 códigos de barras diferentes).

Uso de un transposoma para generar fragmentos de ADN

En métodos y kits de algunas realizaciones, pueden generarse fragmentos de ADN con un transposoma. Como se usa en la presente, un "transposoma" comprende (i) una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y (ii) por lo menos dos copias de un adaptador que comprende una secuencia de captura. El adaptador puede configurarse para añadirlo a un extremo de un ADNdc. Por tanto, el adaptador puede configurarse para añadir la secuencia de captura a los extremos del ADNdc después de que la fracción haya inducido la ruptura de cadena doble en el ADNdc. La nucleasa de cadena doble puede comprender una enzima como una transposasa (por ejemplo, Tn5, Tn7, Tn10, Tc3 o una transposasa mariner como Mos1), una endonucleasa de restricción (por ejemplo, EcoRI, NotI, HindIII, HhaI, BamHI o Sal I), una proteína asociada a CRISPR (por ejemplo, Cas9 o Cas12a), nucleasa específica de dúplex (DSN), o una combinación de estas. Se contempla que mientras algunas nucleasas de cadena doble como la transposasa pueden facilitar la adición de un adaptador a un extremo de un fragmento de ADNdc, otras, por ejemplo, las endonucleasas de restricción, no lo hacen. Como tal, un transposoma puede comprender opcionalmente una ligasa (por ejemplo, ADN ligasa T4, T7 o Taq). Se contempla además que un transposoma puede dirigirse a una estructura particular que comprende ADNdc, por ejemplo, cromatina, ADNdc metilado, un complejo de iniciación de la transcripción o similares. Por consiguiente, dirigiendo los adaptadores a la estructura que comprende ADNdc, fragmentando el ADNdc y codificando con códigos de barras el ADNdc para obtener información de secuencia en el

ADNdc, el transposoma puede proporcionar información sobre las secuencias de ADN asociadas con la estructura que comprende el ADNdc. Como tal, el transposoma puede comprender además una fracción que dirige el transposoma a la estructura que comprende el ADNdc, por ejemplo, un anticuerpo (por ejemplo, el anticuerpo HTA28 que se une específicamente a la histona S28 fosforilada de la histona H3, o) o un fragmento del mismo, un aptámero (ácido nucleico o péptido) o un dominio de unión a ADN (por ejemplo, un dominio de unión a dedos de zinc). En cualquier método de análisis de muestras como se describe en la presente, el transposoma puede dirigirse a una estructura específica que comprende ADNdc, por ejemplo, cromatina, un estado de metilación de ADN particular, un ADN en un orgánulo específico o similar. Se contempla que el método de análisis de muestras pueda identificar secuencias de ADN particulares asociadas con estructuras objetivo del transposoma, por ejemplo, ADN accesible a la cromatina, ADN de constructo, ADN de orgánulos o similares. En algunas realizaciones, se describe un kit para el análisis de muestras. El kit puede comprender un transposoma como se describe en la presente y una pluralidad de códigos de barras como se describe en la presente. Los códigos de barras pueden inmovilizarse en partículas como se describe en la presente.

A modo de ejemplo, generar la pluralidad de fragmentos de ácidos nucleicos puede comprender: poner en contacto el ADNdc (por ejemplo, ADNg) con un transposoma 428, en donde el transposoma 428 comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc (por ejemplo, una transposasa) 430 y dos copias 434a, 434b de un adaptador que comprende la secuencia de captura (por ejemplo, una secuencia poli(dT)), para generar una pluralidad de fragmentos de ADN de cadena doble (ADNdc) 440, cada uno de los cuales comprende una secuencia complementaria 434a', 434b' a la secuencia de captura 434a, 434b. Por ejemplo, el adaptador puede no incluir un saliente 5', como un saliente de poli(dT) 434a, 434b. La nucleasa de cadena doble (por ejemplo, transposasa) 430 puede cargarse con las dos copias 434a, 434b del adaptador. En algunas realizaciones, la secuencia de captura 434a, 434b comprende una región poli(dT). La secuencia complementaria 434a' 434b' a la secuencia de captura puede comprender una región poli(dA).

La generación de la pluralidad de fragmentos de ácidos nucleicos puede comprender: poner en contacto el ADNdc (por ejemplo, ADNg) con un transposoma 428, en donde el transposoma 428 comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc (por ejemplo, transposasa) 430 y dos copias 432a, 432b de un adaptador que tiene un saliente 5' 434a, 434b que comprende una secuencia de captura, para generar una pluralidad de fragmentos de ADN de cadena doble (ADNdc) 438 cada uno de los cuales comprende dos copias de los salientes 5' 434a, 434b. La nucleasa de cadena doble 430 puede cargarse con las dos copias 432a, 432b del adaptador. El método 400 puede, en algunas realizaciones, incluir poner en contacto 412 la pluralidad de fragmentos ADNdc 438 que tienen los salientes 5' 434a, 434b con una polimerasa para generar la pluralidad de fragmentos de ácidos nucleicos 440 que comprenden una pluralidad de fragmentos de ADNdc, cada uno de los cuales comprende una secuencia complementaria 434a', 434b' (por ejemplo, un complemento o un complemento inverso) a por lo menos una parte del saliente 5'. En algunas realizaciones, ninguno de la pluralidad de fragmentos de ADNdc 442 comprende un saliente (por ejemplo, un saliente 3' o un saliente 5' como los salientes 5' 434a', 434b').

Intensidad de señal más alta

Para capturar información genómica y de accesibilidad de la cromatina, la señal (por ejemplo, el número de fragmentos de ADNdc de interés, como los fragmentos de ADNdc para el análisis de accesibilidad de la cromatina, puede amplificarse aún más incorporando un promotor (por ejemplo, un promotor T7) delante del cola poli(dA) en el transposoma 428. Por ejemplo, el ADNdc (por ejemplo, ADNg) puede amplificarse aún más (por ejemplo, 1000 veces) incorporando la transcripción in vitro dentro de los núcleos 452 o la célula antes de cargarlo en un sistema o plataforma de células individuales 416. Por ejemplo, un promotor T7 502 en la secuencia puede unirse a los extremos de los fragmentos de ADNdc (por ejemplo, ADNg).

Después de la transposición y la adición de la secuencia poli(dT) y el promotor, incubar las células o los núcleos fijados con la mezcla de reacción de transcripción in vitro (IVT). Se producirían miles de copias del ARN que lleva la secuencia de ADNdc (por ejemplo, ADNg) y se contendrían dentro de la célula o los núcleos fijados. La captura y lisis de una célula individual (por ejemplo, en el bloque 418 en las FIGS. 4A-4B) puede producirse como se describe en la presente.

Las FIGS. 5A-5B ilustran esquemáticamente un método ejemplar no limitativo para capturar información de accesibilidad genómica y cromática de células individuales con intensidad de señal mejorada. En algunas realizaciones, el adaptador 432a, 432b comprende opcionalmente una secuencia promotora. La secuencia promotora puede comprender una secuencia promotora T7 502. La generación de la pluralidad de fragmentos de ácidos nucleicos puede comprender la transcripción de la pluralidad de fragmentos de ADNdc usando la transcripción in vitro para generar una pluralidad de moléculas de ácido ribonucleico (ARN) 504. La codificación con códigos de barras 424 de la pluralidad de fragmentos de ácidos nucleicos comprende codificar con códigos de barras la pluralidad de moléculas de ARN 504.

Uso de una enzima de restricción para generar fragmentos de ADNdc con extremos romos

En algunas realizaciones, generar la pluralidad de fragmentos de ácidos nucleicos comprende: fragmentar el ADNdc (por ejemplo, ADNg) para generar una pluralidad de fragmentos de ADNdc con extremos romos usando una enzima de restricción. Fragmentar el ADNdc (por ejemplo, ADNg) puede comprender poner en contacto el ADNdc (por ejemplo, ADNg) con una enzima de restricción para generar la pluralidad de fragmentos de ADNdc, cada uno con extremos romos. Por lo menos uno de la pluralidad de fragmentos de ADNdc puede comprender un extremo romo. Por lo menos uno de la pluralidad de fragmentos de ADNdc puede comprender un saliente 5' o un saliente 3'. Ninguno de la pluralidad de fragmentos de ADNdc puede comprender un extremo romo. Fragmentar el ADNdc (por ejemplo, ADNg) puede comprender poner en contacto el ADNg de cadena doble con una enzima de restricción para generar la pluralidad de fragmentos de ADNdc con extremos romos. Por lo menos uno, algunos, o todos de los fragmentos de ADNdc puede incluir extremos romos.

Uso de proteína asociada a CRISPR para generar fragmentos de ADNdc

En algunas realizaciones, generar la pluralidad de fragmentos de ácidos nucleicos comprende: fragmentar el ADNdc (por ejemplo, ADNg) para generar una pluralidad de fragmentos de ácido desoxirribonucleico de cadena doble (ADNdc) usando una proteína asociada a CRISPR como Cas9 o Cas12a. Fragmentar el ADNdc (por ejemplo, ADNg) puede comprender poner en contacto el ADNg de cadena doble con la proteína asociada a CRISPR para generar la pluralidad de fragmentos de ADNdc. Por lo menos uno, algunos o todos los fragmentos de ADNdc pueden incluir extremos romos. Se contempla que en algunas realizaciones, las roturas en el ADNdc pueden dirigirse a secuencias o motivos particulares usando un ARN guía (gRNA) dirigido a la secuencia o motivo particular, de manera que la proteína asociada a CRISPR induzca roturas de cadena doble en la secuencia o motivo particular.

Generación de fragmentos de ácidos nucleicos

En algunas realizaciones, generar la pluralidad de fragmentos de ácidos nucleicos (por ejemplo, usando una enzima de restricción o una proteína asociada a CRISPR) comprende: agregar (por ejemplo, en el bloque 410 analizado con referencia a las FIGS. 4A-4B) dos copias de un adaptador que comprende una secuencia complementaria a una secuencia de captura de por lo menos algunos (por ejemplo, por lo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 100, 1000, 10000, 100000, 1000000, 10000000 o más) de la pluralidad de fragmentos de ADNdc para generar una pluralidad de fragmentos de ADNdc (por ejemplo, una pluralidad de fragmentos de ADNdc con extremos romos). Agregar las dos copias del adaptador puede comprender ligar las dos copias del adaptador a por lo menos algunos (por ejemplo, por lo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 100, 1000, 10000, 100000, 1000000, 10000000 o más) de la pluralidad de fragmentos de ADNdc para generar la pluralidad de fragmentos de ADNdc que comprenden el adaptador.

Uso de una enzima de restricción para generar fragmentos de ADNdc con salientes

En algunas realizaciones, generar la pluralidad de fragmentos de ácidos nucleicos comprende: fragmentar el ADNdc (por ejemplo, ADNg) para generar una pluralidad de fragmentos de ADNdc con salientes usando una enzima de restricción para que no sea necesario añadir adaptadores. Fragmentar el ADNdc (por ejemplo, ADNg) puede comprender poner en contacto el ADNdc (por ejemplo, ADNg) con una enzima de restricción para generar la pluralidad de fragmentos de ADNdc, en donde por lo menos uno de la pluralidad de fragmentos de ADNdc comprende la secuencia de captura. La secuencia de captura puede ser complementaria a las secuencias de los salientes 5'. La secuencia complementaria a la secuencia de captura puede comprender la secuencia del saliente 5'.

Accesibilidad a la cromatina

En referencia a las FIGS. 4A-4B, para capturar la información de accesibilidad a la cromatina 406a, los núcleos 452 pueden incubarse con cortadores enzimáticos (por ejemplo, una transposasa, una enzima de restricción y Cas9) y los fragmentos de ADNdc (por ejemplo, ADNg) pueden unirse con adaptadores 432a, 432b. El corte puede producirse en lugares donde las cromatinas están expuestas (por ejemplo, más expuestas, más expuestas que la media y expuestas en la medida deseable). Por ejemplo, una transposasa 432 puede insertar los adaptadores 432a, 432b en el ADNdc (por ejemplo, ADNg).

En algunas realizaciones, determinar la información relacionada con el ADNdc (por ejemplo, ADNg) comprende determinar la accesibilidad a la cromatina 406a del ADNdc (por ejemplo, ADNg) sobre la base de las secuencias y/o la abundancia de la pluralidad de fragmentos de ADNmc 442 en los datos de secuenciación obtenidos. Determinar la accesibilidad a la cromatina 442 del ADNdc (por ejemplo, ADNg) puede comprender: alinear las secuencias de la pluralidad de fragmentos ADNmc 442 con una secuencia de referencia del ADNdc (por ejemplo, ADNg); determinar regiones del ADNdc (por ejemplo, ADNg) correspondientes a los extremos de los fragmentos de ADNmc de la pluralidad de fragmentos de ADNmc 442 para que sean accesibles o tengan cierta accesibilidad (por ejemplo, altamente accesible, accesibilidad superior a la media y accesibilidad superior a un umbral o medida deseada). Determinar la accesibilidad a la cromatina del ADNdc (por ejemplo, ADNg) puede

comprender: alinear las secuencias de la pluralidad de fragmentos de ADNmc con una secuencia de referencia del ADNdc (por ejemplo, ADNg); y determinar la accesibilidad de las regiones del ADNdc (por ejemplo, ADNg) correspondientes a los extremos de los fragmentos de ADNmc de la pluralidad de fragmentos de ADNmc basándose en los números de fragmentos de ADNmc de la pluralidad de fragmentos de ADNmc en los datos de secuenciación.

Por ejemplo, el corte puede producirse en localizaciones donde las cromatinas tienen una accesibilidad superior a la media. Las regiones del ADNdc (por ejemplo, ADNg) que corresponden a los extremos de los fragmentos de ADNmc pueden tener una accesibilidad superior a la media. Tales regiones del ADNdc (por ejemplo, ADNg) pueden tener una abundancia superior a la media en los datos de secuenciación obtenidos. Como otro ejemplo, el ADNdc (por ejemplo, ADNg) comprende la región A-región B1-región B2-región C. Si la región B1 y la región B2 tienen una accesibilidad superior a la media mientras que la región A y la región C tienen una accesibilidad inferior a la media, la región B1 y la región B2 pueden cortarse (por ejemplo, entre la región B1 y la región B2), mientras que la región A y la región C no se cortan. Los datos de secuenciación pueden incluir una abundancia superior a la media de secuencias de la región B1 y la región B2 donde se produce el corte (y alrededor de donde se produce el corte). Las secuencias de la región A y la región C pueden no estar presentes (o tener poca abundancia) en los datos de secuenciación. Por tanto, la accesibilidad cromática del ADNdc (por ejemplo, ADNg) puede determinarse sobre la base de la secuencia y el número de cada uno de la pluralidad de fragmentos de ADNmc.

Información del genoma

Para capturar información del genoma 406b, los núcleos 452 pueden exponerse primero a reactivos para digerir 408 la estructura del nucleosoma (por ejemplo, para eliminar las proteínas del nucleosoma/histona), antes de someterlos a cortadores enzimáticos y la adición de adaptadores. En algunas realizaciones, determinar la información relacionada con el ADNdc (por ejemplo, ADNg) comprende determinar la información del genoma 406b del ADNdc (por ejemplo, ADNg) basándose en las secuencias de la pluralidad de fragmentos de ADNmc 442 en los datos de secuenciación obtenidos. El método puede comprender digerir 408 nucleosomas asociados con el ADNdc de cadena doble (por ejemplo, ADNg). Determinar la información del genoma del ADNdc (por ejemplo, ADNg) puede comprender: determinar por lo menos una secuencia parcial del ADNdc (por ejemplo, ADNg) alineando las secuencias de la pluralidad de fragmentos de ADNmc 442 con una secuencia de referencia del ADNdc (por ejemplo, ADNg). En algunas realizaciones, puede determinarse un genoma completo o parcial de una célula. En algunas realizaciones, el ADNdc es el ADN genómico (ADNg) de una célula. En algunas realizaciones, el ADNdc es el ADN genómico de un orgánulo de la célula, por ejemplo, una mitocondria o un cloroplasto.

Información de metiloma

Para capturar información de metiloma 406c, después de que los fragmentos de ADNdc (por ejemplo, ADNg) sean capturados por la sonda de captura 444 y permanezcan de cadena sencillas 442, se usa tratamiento con bisulfito 422 para convertir las bases de metilcitosina 454mc en bases de timina. Posteriormente, el ADNdc (por ejemplo, ADNg) puede copiarse mediante RT 424 o ADN polimerasa.

En algunas realizaciones, determinar la información relacionada con el ADNdc (por ejemplo, ADNg) comprende determinar la información del metiloma 406c del ADNdc (por ejemplo, ADNg) basándose en las secuencias de la pluralidad de fragmentos 442 de ADNmc en los datos de secuenciación obtenidos. El método puede comprender: digerir 408 nucleosomas asociados con el ADNdc (por ejemplo, ADNg). El método 400 puede comprender: realizar la conversión de bisulfito 422 de bases de citosina de la pluralidad de ADN de cadena sencilla 442 para generar una pluralidad de ADNmc convertido con bisulfito 442b con bases de uracilo 454u. La codificación con códigos de barras 424 de la pluralidad de fragmentos 442 de ADNmc puede comprender codificar con códigos de barras 424 de la pluralidad de ADNmc convertidos con bisulfito 452b usando la pluralidad de códigos de barras 444 para generar la pluralidad de fragmentos de ADNmc codificados con códigos de barras 446 y/o secuencias complementarias de los mismos. Determinar la información del metiloma 406c puede comprender: determinar una posición de la pluralidad de fragmentos de ADNmc 442 en los datos de secuenciación que tienen una base de timina (o base de uracilo 454u) y la posición correspondiente en una secuencia de referencia del ADNdc (por ejemplo, ADNg) que tiene una base de citosina para determinar la posición correspondiente en el ADNdc (por ejemplo, ADNg) que tiene una base de metilcitosina 454mc.

En algunas realizaciones, determinar la información del metiloma comprende un método de análisis de muestras que comprende poner en contacto ácido desoxirribonucleico de cadena doble (ADNdc) de una célula con un transposoma, en el que el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc cargado con dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5'. El método puede comprender además poner en contacto la pluralidad de fragmentos de ADNdc salientes con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a por lo menos una parte de cada uno de los salientes 5', desnaturalizar la pluralidad de fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc),

codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados con códigos de barras, en el que cada uno de la pluralidad de códigos de barras comprende una secuencia de marcadores celulares, una secuencia de marcadores moleculares y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden secuencias de marcadores moleculares diferentes, y en el que por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcadores celulares idéntica, obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras, y determinar la información relacionada con el ADNdc basándose en secuencias de la pluralidad de fragmentos de ADNmc codificados con códigos de barras en los datos de secuenciación. En algunas realizaciones, el método comprende además capturar un fragmento de ADNmc de la pluralidad de fragmentos de ADNmc codificados con códigos de barras en una partícula que comprende un oligonucleótido que comprende la secuencia de captura, la secuencia del marcador celular y la secuencia del marcador molecular. A modo de ejemplo, la secuencia de captura puede comprender una secuencia poli dT que se une a una cola poli A en el fragmento de ADNmc. El fragmento de ADNmc capturado puede comprender una citidina metilada, realizar una reacción de conversión de bisulfuro en el fragmento de ADNmc para convertir la citidina metilada en una timidina, extender el fragmento de ADNmc en la dirección 5' a 3' para producir el fragmento de ADNmc codificado con códigos de barras que comprende la timidina, el ADNmc codificado con códigos de barras que comprende la secuencia de captura, la secuencia del marcador molecular y la secuencia del marcador celular, extender el oligonucleótido en la dirección 5' a 3' usando una transcriptasa inversa o polimerasa o una combinación de las mismas para producir una cadena de ADN complementaria que sea complementaria al ADNmc codificado con código de barras que comprende la timidina, desnaturalizar el ADNmc codificado con código de barras y la cadena de ADN complementaria para producir secuencias de cadena sencilla y amplificar las secuencias de cadena sencilla.

En algunas realizaciones, obtener la información del metiloma comprende determinar una posición de la pluralidad de fragmentos de ADNmc en los datos de secuenciación que tienen una base de timina y la posición correspondiente en una secuencia de referencia del ADNdc que tiene una base de citosina que comprende una conversión de bisulfuro de una citosina metilada de un fragmento de ADNmc de la pluralidad, convirtiendo así la base de citosina metilada en la base de timina, y determinar la posición correspondiente de la base de timina en la secuencia de referencia para que sea una base de citosina.

Multiómica

En algunas realizaciones, el método puede incluir: codificar con códigos de barras una pluralidad de objetivos (por ejemplo, objetivos en el núcleo 452) usando la pluralidad de códigos de barras 444 para generar una pluralidad de objetivos codificados con códigos de barras; y obtener datos de secuenciación de los objetivos codificados con códigos de barras. Los objetivos pueden ser objetivos de ácido nucleico, como objetivos de ARNm, oligonucleótidos de indexación de muestras (por ejemplo, descritos en la Solicitud de Estados Unidos Nº 15/937,713, publicada como US 2018/0346970.), y oligonucleótidos para determinar la expresión de proteínas (por ejemplo, descritos en la Solicitud de Estados Unidos Nº 15/715028, publicada como US2018/0088112). En algunas realizaciones, pueden determinarse dos o más de la información del genoma, la accesibilidad a la cromatina, el metiloma, el transcriptoma y el proteoma en células individuales.

En algunas realizaciones, un método de análisis de muestras comprende poner en contacto ácido desoxirribonucleico de cadena doble (ADNdc) de una célula con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc cargado con dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5', poner en contacto la pluralidad de fragmentos de ADNdc salientes con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios que comprenden cada uno una secuencia complementaria a por lo menos una porción de cada uno de los salientes 5', desnaturalizar la pluralidad de fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc), codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados con código de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden diferentes secuencias de marcadores moleculares, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcador celular idéntica, obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras y determinar la información relacionada con el ADNdc basándose en las secuencias de la pluralidad de fragmentos de ADNmc codificados con códigos de barras en los datos de secuenciación. En algunas realizaciones, el método comprende además poner en contacto una célula con un reactivo de ácido nucleico, el reactivo de ácido nucleico comprendiendo una secuencia de captura, un código de barras, un sitio de unión del cebador, un agente de unión a ADN de cadena doble, en donde la célula es una célula muerta, y en donde el reactivo de unión nucleico se une al ADN de cadena doble en la célula muerta, lavar la célula muerta para eliminar el exceso del reactivo de ácido nucleico, lisar la célula muerta, liberando de este modo el reactivo de ácido nucleico y codificar con códigos de barras el reactivo de ácido nucleico.

En algunas realizaciones del método de análisis de muestras, la célula está asociada con un soporte sólido que comprende un oligonucleótido que comprende una secuencia de marcador celular, y en donde la codificación con código de barras comprende codificar con código de barras el reactivo de ácido nucleico con la secuencia de marcador celular.

En algunas realizaciones del método de análisis de muestras, el soporte sólido comprende una pluralidad de oligonucleótidos, cada uno comprendiendo la secuencia de marcador celular y una secuencia de marcador molecular diferente.

En algunas realizaciones, el método de análisis de muestras comprende además secuenciar los reactivos de ácido nucleico codificados con códigos de barras y determinar la presencia de una célula muerta basándose en la presencia del código de barras del reactivo de ácido nucleico.

En algunas realizaciones, el método de análisis de muestras comprende además asociar dos o más células, cada una con diferentes soportes sólidos que comprenden diferentes marcadores celulares, por lo que cada una de las dos o más células se asocia una a una con un marcador de célula diferente.

En algunas realizaciones, el método de análisis de muestras comprende además determinar una cantidad de células muertas en la muestra en función de la cantidad de marcadores celulares únicos asociados con un código de barras de un reactivo de ácido nucleico.

En algunas realizaciones, el método de análisis de muestras comprende determinar el número de secuencias de marcadores moleculares con secuencias distintas asociadas con el marcador celular y la secuencia de código de barras de control comprende determinar el número de secuencias de marcadores moleculares con el mayor número de secuencias distintas asociadas con el marcador celular y la secuencia de código de barras de control para cada marcador celular en los datos de secuenciación.

En algunas realizaciones del método de análisis de muestras, la célula es una célula viva y en donde el reactivo de ácido nucleico no se introduce en la célula viva y, por lo tanto, no se une al ADN de cadena doble en la célula viva.

En algunas realizaciones, el método de análisis de muestras comprende además poner en contacto una célula muerta con un reactivo de unión a proteínas asociado con un oligonucleótido identificador único, por lo que el reactivo de unión a proteínas se une a una proteína de la célula muerta y codificar con códigos de barras el oligonucleótido identificador único.

En algunas realizaciones del método de análisis de muestras, el reactivo de unión a proteínas comprende un anticuerpo, un tetrámero, un aptámero, un andamiaje proteico, una invasina o una combinación de los mismos. En algunas realizaciones, el reactivo de unión a proteínas comprende un anticuerpo o fragmentos del mismo, aptámero, molécula pequeña, ligando, péptido, oligonucleótido o cualquier combinación de los mismos. A modo de ejemplo, el reactivo de unión a proteínas puede comprender, consistir esencialmente en o consistir en un anticuerpo policlonal, un anticuerpo monoclonal, un anticuerpo recombinante, un anticuerpo de cadena sencilla (scAb) o fragmentos de los mismos, como Fab, Fv, scFv, o similares. A modo de ejemplo, el anticuerpo puede comprender, consistir esencialmente en o consistir en un anticuerpo Abseq (consultar Shahi et al. (2017), Sci Rep. 7:44447). El identificador único del reactivo de unión a proteínas puede comprender una secuencia de nucleótidos. En algunas realizaciones, el identificador único comprende una secuencia de nucleótidos de 25 a 45 nucleótidos de longitud. En algunas realizaciones, el identificador único no es homólogo a las secuencias genómicas de la muestra o célula. En algunas realizaciones, el reactivo de unión a proteínas puede asociarse covalentemente con el oligonucleótido identificador único. En algunas realizaciones, el reactivo de unión a proteínas puede asociarse covalentemente con el oligonucleótido identificador único. Por ejemplo, el reactivo de unión a proteínas puede asociarse con el oligonucleótido identificador único a través de un conector. En algunas realizaciones, el conector puede comprender un grupo químico que une reversiblemente el oligonucleótido a los reactivos de unión a proteínas. El grupo químico puede conjugarse con el conector, por ejemplo, a través de un grupo amina. En algunas realizaciones, el conector puede comprender un grupo químico que forma un enlace estable con otro grupo químico conjugado con el reactivo de unión a proteínas. Por ejemplo, el grupo químico puede ser un grupo fotoescindible UV, estreptavidina, biotina, amina, etc. En algunas realizaciones, el grupo químico puede conjugarse con el reactivo de unión a proteínas a través de una amina primaria en un aminoácido, como lisina, o el extremo N-terminal. El oligonucleótido puede conjugarse con cualquier sitio adecuado del reactivo de unión a proteínas, siempre que no interfiera con la unión específica entre el reactivo de unión a proteínas y su proteína objetivo. En las realizaciones en las que el reactivo de unión a proteínas es un anticuerpo, el oligonucleótido puede conjugarse con el anticuerpo en cualquier lugar que no sea el sitio de unión al antígeno, por ejemplo, la región Fc, el dominio C_H1, el dominio C_H2, el dominio C_H3, el dominio C_L, etc. En algunas realizaciones, cada reactivo de unión a proteínas puede conjugarse con una sola molécula de oligonucleótido. En algunas realizaciones, cada reactivo de unión a proteínas puede conjugarse con más de una molécula de oligonucleótido, por ejemplo, por lo menos 2, por lo menos 3, por lo menos 4, por lo menos

5, por lo menos 10, por lo menos 20, por lo menos 30, por lo menos 40, por lo menos 50, por lo menos 100, por lo menos 1.000 o más moléculas de oligonucleótidos, en donde cada una de las moléculas de oligonucleótidos comprende el mismo identificador único.

5 En algunas realizaciones del método de análisis de muestras, un objetivo proteico del reactivo de unión a proteínas se selecciona de un grupo que comprende 10-100 objetivos de proteínas diferentes, o un objetivo de componente celular del reactivo de unión a componentes celulares se selecciona de un grupo que comprende 10-100 objetivos de componentes celulares diferentes.

10 En algunas realizaciones del método de análisis de muestras, un objetivo de proteínas del reactivo de unión a proteínas comprende un carbohidrato, un lípido, una proteína, una proteína extracelular, una proteína de la superficie celular, un marcador celular, un receptor de células B, un receptor de células T, un complejo mayor de histocompatibilidad, un antígeno tumoral, un receptor, una integrina, una proteína intracelular o cualquier combinación de los mismos.

15 En algunas realizaciones del método de análisis de muestras, el reactivo de unión a proteínas comprende un anticuerpo o fragmento del mismo que se une a una proteína de la superficie celular.

20 En algunas realizaciones del método de análisis de muestras, el código de barras es un código de barras que comprende una secuencia de marcadores moleculares.

25 En algunas realizaciones, un método de análisis de muestras comprende poner en contacto una célula muerta de una muestra con un reactivo de ácido nucleico. El reactivo de ácido nucleico puede comprender, consistir esencialmente en o consistir en cualquier agente de ácido nucleico como se describe en la presente. Por ejemplo, el agente de unión a ácidos nucleicos puede comprender una secuencia de captura, un código de barras, un sitio de unión al cebador y un agente de unión a ADN de cadena doble. A modo de ejemplo, el código de barras puede comprender un marcador celular, un marcador molecular y una región de unión al objetivo como se describe en la presente. El reactivo de ácido nucleico puede unirse al ADN de cadena doble en la célula muerta. El método puede comprender lavar el exceso de reactivo de ácido nucleico de la célula muerta, por ejemplo, centrifugando la muestra, aspirando el fluido de la muestra y aplicando un nuevo fluido, como un tampón, a la muestra. El lavado puede eliminar el reactivo de ácido nucleico no unido, mientras que el reactivo de unión de ácido nucleico unido a ADN de cadena doble puede permanecer unido al ADN de cadena doble de la célula muerta. Se contempla que para las células vivas, el lavado eliminará todo (o eliminará todo excepto cantidades traza del reactivo de ácido nucleico). El método puede comprender lisar la célula muerta. La lisis puede liberar el reactivo de ácido nucleico de la célula muerta. A modo de ejemplo, la célula muerta puede lisarse mediante la adición de un tampón de lisis celular que comprende un detergente (por ejemplo, SDS, dodecil sulfato de Li, Triton X-100, Tween-20 o NP-40), un solvente orgánico (por ejemplo, metanol o acetona), una enzima digestiva (por ejemplo, proteinasa K, pepsina o tripsina), o cualquier combinación de los mismos. El método puede comprender codificar con códigos de barras del reactivo de ácido nucleico como se describe en la presente. La codificación con códigos de barras puede producir un ácido nucleico que comprende el código de barras del reactivo de ácido nucleico (o un complemento del mismo) marcado con un marcador celular. Opcionalmente, el ácido nucleico puede comprender además un marcador molecular. Se contempla que el marcador celular pueda asociar el reactivo de ácido nucleico uno a uno con una célula (por ejemplo, la célula muerta), y el marcador molecular pueda usarse para cuantificar el número de reactivos de ácido nucleico asociados con una sola célula (por ejemplo, la célula muerta).

45 En algunas realizaciones de un método de análisis de muestras, codificar con códigos de barras comprende capturar la célula muerta en un soporte sólido, como una perla, el soporte sólido comprendiendo una secuencia de marcador celular y una secuencia de marcador molecular.

50 En algunas realizaciones, un método de análisis de muestras comprende además determinar una cantidad de secuencias de marcadores moleculares distintas asociadas con cada secuencia de marcadores celulares, y determinar una cantidad de células muertas en la muestra sobre la base de la cantidad de secuencias de marcadores celulares distintas asociadas con secuencias de marcadores moleculares. Por ejemplo, en algunas realizaciones, la presencia de un código de barras de un reactivo de ácido nucleico como se describe en la presente puede indicar que una célula es una célula muerta. Por ejemplo, en algunas realizaciones, una cantidad de códigos de barras de reactivos de ácidos nucleicos que superan un umbral pueden indicar que una célula es una célula muerta. El umbral puede comprender, por ejemplo, un límite de detección o una cantidad de códigos de barras de reactivos de ácido nucleico que supera un control negativo, por ejemplo, una célula viva conocida. En algunas realizaciones, una cantidad de por lo menos 10, 50, 100, 500, 1000, 5000, o 10000 códigos de barras de reactivos de ácido nucleico asociados con la célula puede indicar que la célula es una célula muerta.

65 En algunas realizaciones de un método de análisis de muestras, determinar el número de secuencias de marcadores moleculares con secuencias distintas asociadas con el marcador celular y la secuencia de código de barras de control comprende determinar el número de secuencias de marcadores moleculares con el mayor número de secuencias distintas asociadas con el marcador celular para cada marcador celular en los datos de

secuenciación.

En algunas realizaciones, un método de análisis de muestras comprende además poner en contacto una célula muerta con un reactivo de unión a proteínas asociado con un oligonucleótido identificador único, por lo que el reactivo de unión a proteínas se une a una proteína de la célula muerta. El método puede comprender además codificar con códigos de barras el oligonucleótido identificador único. Opcionalmente, el reactivo de unión a proteínas puede ponerse en contacto con la célula muerta antes de lavar la célula muerta. En algunas realizaciones, la célula muerta se pone en contacto con dos o más reactivos de unión a proteínas diferentes, cada uno asociado con un oligonucleótido identificador único. Por tanto, por lo menos dos proteínas diferentes de la célula muerta, si están presentes, pueden unirse con los diferentes reactivos de unión a proteínas.

En algunas realizaciones de un método de análisis de muestras, el reactivo de unión a proteínas se asocia con dos o más oligonucleótidos de indexación de muestras con una secuencia idéntica.

En algunas realizaciones de un método de análisis de muestras, el reactivo de unión a proteínas se asocia con dos o más oligonucleótidos de indexación de muestras con diferentes secuencias de indexación de muestras.

En algunas realizaciones de un método de análisis de muestras, el reactivo de unión a proteínas comprende un anticuerpo, un tetrámero, un aptámero, un andamiaje proteico, una invasina o una combinación de los mismos.

En algunas realizaciones de un método de análisis de muestras, un objetivo proteico del reactivo de unión a proteínas se selecciona de un grupo que comprende 10-100 objetivos de proteínas diferentes, o en donde un componente celular objetivo del reactivo de unión a componentes celulares se selecciona de un grupo que comprende 10-100 objetivos de componentes celulares diferentes.

En algunas realizaciones de un método de análisis de muestras, un objetivo proteico del reactivo de unión a proteínas comprende un carbohidrato, un lípido, una proteína, una proteína extracelular, una proteína de superficie celular, un marcador celular, un receptor de células B, un receptor de células T, un complejo mayor de histocompatibilidad, un antígeno tumoral, un receptor, una integrina, una proteína intracelular o cualquier combinación de los mismos.

En algunas realizaciones de un método de análisis de muestras, el reactivo de unión a proteínas comprende un anticuerpo o fragmento del mismo que se une a una proteína de la superficie celular.

En algunas realizaciones de un método de análisis de muestras, la secuencia de captura y la secuencia complementaria a la secuencia de captura son un par específico de ácidos nucleicos complementarios de por lo menos 5 nucleótidos a aproximadamente 25 nucleótidos de longitud.

En algunas realizaciones, un método de análisis de muestras comprende poner en contacto ácido desoxirribonucleico de cadena doble (ADNdc) de una célula con un transposoma. El transposoma puede comprender una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc cargada con dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5'. El método puede comprender poner en contacto la pluralidad de fragmentos de ADNdc salientes con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a por lo menos una parte de cada uno de los salientes 5'. El método puede comprender desnaturalizar la pluralidad de fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc). El método puede comprender codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura. Todas las secuencias de marcadores celulares asociadas con una única célula pueden ser las mismas, para asociar cada célula única, una a una, con una secuencia de marcador celular. Por lo menos dos de la pluralidad de códigos de barras pueden comprender diferentes secuencias de marcadores moleculares. El método puede comprender obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras. El método puede comprender cuantificar una cantidad del ADNdc en la célula sobre la base de una cantidad de secuencias de marcadores moleculares únicas asociadas con la misma secuencia de marcador celular.

En algunas realizaciones, un método de análisis de muestras comprende además capturar un fragmento de ADNmc de la pluralidad de fragmentos de ADNmc en un soporte sólido que comprende un oligonucleótido que comprende la secuencia de captura, la secuencia de marcador celular y la secuencia de marcador molecular. La secuencia de captura puede comprender una secuencia de unión al objetivo que hibrida con una secuencia del fragmento de ADNmc que es complementaria a la secuencia de unión al objetivo. Por ejemplo, la secuencia de

captura puede comprender una secuencia poli dT que se une a una cola poli A en el fragmento de ADNmc. El método puede comprender extender el fragmento de ADNmc en la dirección de 5' a 3' para producir el fragmento de ADNmc codificado con códigos de barras. Por ejemplo, la extensión puede realizarse con una ADN polimerasa. El ADNmc codificado con códigos de barras puede comprender la secuencia de captura, la secuencia del marcador molecular y la secuencia del marcador celular. El método puede comprender extender el oligonucleótido en la dirección de 5' a 3' usando una transcriptasa o polimerasa inversa o una combinación de las mismas para producir una cadena de ADN complementaria que es complementaria al ADNmc codificado con códigos de barras. El método puede comprender desnaturalizar el ADNmc codificado con códigos de barras y la cadena de ADN complementaria para producir secuencias de cadena sencilla. El método puede comprender amplificar las secuencias de cadena sencilla.

En algunas realizaciones, el método de análisis de muestras comprende además la conversión con bisulfito de bases de citosina de la pluralidad de fragmentos de ADNmc para generar una pluralidad de fragmentos de ADNmc convertidos con bisulfito que comprenden bases de uracilo. Por consiguiente, se contempla que cuando se produzcan cadenas de ADN complementarias a los ADNmc codificados con códigos de barras, las posiciones complementarias a las bases de uracilo comprenderán adenina (en lugar de guanina, como sería de esperar si la base de citosina no se hubiera metilado y, por lo tanto, permaneciera como citosina después del proceso de conversión de bisulfito). Por consiguiente, se contempla que la presencia de adenina (en lugar de guanina) en las posiciones que se espera que contengan guanina en las cadenas de ADN complementarias pueda indicar la metilación de una citosina en esa posición. La presencia de la adenina puede determinarse secuenciando directamente la cadena de ADN complementaria o secuenciando su complemento. Opcionalmente, la secuencia puede compararse con una secuencia de referencia, como una secuencia de referencia genómica. La secuencia de referencia puede ser una referencia almacenada electrónicamente.

Codificación con códigos de barras

En algunas realizaciones, la codificación con códigos de barras 424 comprende cargar células 416 en una plataforma de una célula individual. Los fragmentos de ADNmc 442 o los ácidos nucleicos pueden hibridar 420 con la secuencia de captura 434 para la codificación con códigos de barras. Los fragmentos de ADNmc codificados con códigos de barras 446, un complemento, un complemento inverso 446rc, o una combinación de los mismos, pueden amplificarse 426 antes de y/o para la secuenciación como se describe con referencia a la FIG. 3.

En algunas realizaciones, la codificación con códigos de barras 424 puede incluir: codificar con códigos de barras estocásticos la pluralidad de fragmentos 442 de ADNmc o la pluralidad de ácidos nucleicos usando la pluralidad de códigos de barras 444 para generar una pluralidad de fragmentos 446 de ADNmc codificados con códigos de barras estocásticamente. La codificación con códigos de barras 424 puede comprender: codificar con códigos de barras la pluralidad de fragmentos 442 de ADNmc usando la pluralidad de códigos de barras 444 asociados con una partícula 456 para generar la pluralidad de fragmentos 446 de ADNmc codificados con códigos de barras, en donde los códigos de barras 444 asociados con la partícula 456 comprenden una secuencia de marcadores celulares idéntica y por lo menos 100 secuencias de marcadores moleculares diferentes.

En algunas realizaciones, puede inmovilizarse en la partícula por lo menos un código de barras de la pluralidad de códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede inmovilizarse parcialmente en la partícula. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar encerrado en la partícula. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente encerrado en la partícula. La partícula puede ser alterable (por ejemplo, soluble o degradable). La partícula puede comprender una partícula de hidrogel alterable. La partícula puede comprender una perla de Sefarosa, una perla de estreptavidina, una perla de agarosa, una perla magnética, una perla conjugada, una perla conjugada con proteína A, una perla conjugada con proteína G, una perla conjugada con proteína A/G, una perla conjugada con proteína L, una perla conjugada con oligo(dT), una perla de sílice, una perla similar a la sílice, una microperla antibiotina, una microperla antilfluorocromo, o cualquier combinación de las mismas. La partícula puede comprender un material seleccionado del grupo que consiste en polidimetilsiloxano (PDMS), poliestireno, vidrio, polipropileno, agarosa, gelatina, hidrogel, paramagnético, cerámica, plástico, vidrio, metilestireno, polímero acrílico, titanio, látex, sefarosa, celulosa, nailon, silicona y cualquier combinación de los mismos.

En algunas realizaciones, los códigos de barras de la partícula pueden comprender marcadores moleculares con por lo menos 1000 secuencias de marcadores moleculares diferentes. Los códigos de barras de la partícula pueden comprender marcadores moleculares con por lo menos 10000 secuencias de marcadores moleculares diferentes. Los marcadores moleculares de los códigos de barras pueden comprender secuencias aleatorias. La partícula puede comprender por lo menos 10000 códigos de barras.

La codificación con códigos de barras de la pluralidad de fragmentos de ADNmc puede comprender: poner en contacto la pluralidad de fragmentos de ADNmc con la secuencia de captura de la pluralidad de códigos de barras; y transcribir la pluralidad de ADNmc usando la pluralidad de códigos de barras para generar la pluralidad de fragmentos de ADNmc codificados con códigos de barras. El método puede incluir: antes de obtener los datos de

secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras, amplificar la pluralidad de fragmentos de ADNmc codificados con códigos de barras para generar una pluralidad de fragmentos de ADN codificados con códigos de barras amplificados. La amplificación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras puede comprender: amplificar los fragmentos de ADNmc codificados con códigos de barras mediante la reacción en cadena de la polimerasa (PCR).

Reactivos de ácidos nucleicos

En algunas realizaciones, un reactivo de ácido nucleico comprende, consiste esencialmente en o consiste en una secuencia de captura, un código de barras, un sitio de unión al cebador y un agente de unión al ADN de cadena doble. El código de barras del reactivo de ácido nucleico puede comprender una secuencia identificadora, que indica que el código de barras está asociado con el reactivo de ácido nucleico. Opcionalmente, de acuerdo con los métodos y kits como se describe en la presente, diferentes reactivos de ácidos nucleicos de moléculas pueden comprender diferentes secuencias de código de barras. El ácido nucleico puede usarse en cualquiera de los métodos de análisis de muestras que se describen en la presente. En algunas realizaciones, un kit comprende, consiste esencialmente en o consiste en un reactivo de ácido nucleico como se describe en la presente. Opcionalmente, el kit comprende además un soporte sólido (por ejemplo, una partícula) como se describe en la presente.

En la FIG. 6 se ilustra un ejemplo de reactivo de ácido nucleico 600 de algunas realizaciones. El reactivo de ácido nucleico 600 puede comprender un agente de unión a ADN de cadena doble 610. El reactivo de ácido nucleico 600 puede comprender un sitio de unión de cebador 620, por ejemplo, un asidero de PCR. El reactivo de ácido nucleico 600 puede comprender un código de barras 630. El código de barras puede comprender una secuencia de identificación única. El reactivo de ácido nucleico 600 puede comprender una secuencia de captura 640, por ejemplo, una cola poli(A).

En algunas realizaciones, el reactivo de ácido nucleico es impermeable a la membrana plasmática. Sin querer estar limitados por la teoría, se contempla que tal reactivo de ácido nucleico no pueda atravesar una membrana plasmática intacta (o puede atravesar una membrana plasmática intacta en no más que pequeñas cantidades), y por lo tanto, no se introducirá en los núcleos de las células vivas (o no entrará en los núcleos de las células vivas más que en cantidades traza). Por el contrario, el reactivo de ácido nucleico puede introducirse en los núcleos de las células muertas porque la membrana plasmática de las células muertas no está intacta. En algunas realizaciones, el reactivo de ácido nucleico está configurado para unirse específicamente a las células muertas y el reactivo de ácido nucleico no se une a las células vivas.

En algunas realizaciones del reactivo de ácido nucleico, la secuencia de captura comprende una región poli(A).

En algunas realizaciones del reactivo de ácido nucleico, el sitio de unión al cebador comprende un sitio de unión al cebador universal.

En algunas realizaciones, se describe un método para unir un reactivo de ácido nucleico a una célula. El método puede comprender marcar células de una muestra con reactivos de ácido nucleico. El exceso de reactivos de ácido nucleico puede lavarse. Opcionalmente, las células también se marcan con uno o más códigos de barras como se describe en la presente, por ejemplo, reactivos de unión a proteínas asociados con una secuencia de identificación única, por ejemplo, un anticuerpo Abseq. Luego, las células pueden asociarse con una partícula que comprende códigos de barras inmovilizados en la misma. Los ácidos nucleicos de la célula (por ejemplo, ARNm) y/o las secuencias identificadoras únicas (de reactivos de unión a proteínas, como los anticuerpos Abseq), y los reactivos de unión a ácidos nucleicos de la célula pueden asociarse con un marcador celular único, por ejemplo, inmovilizados en un soporte sólido, o en una división. Los ácidos nucleicos pueden codificarse con códigos de barras con el marcador de una célula individual y un marcador molecular como se describe en la presente. Puede prepararse una biblioteca de ácidos nucleicos codificados con códigos de barras. La biblioteca puede secuenciarse. Se observa que además de proporcionar información sobre recuentos de proteínas y/o ácidos nucleicos de las células, la secuenciación puede proporcionar información sobre si el reactivo de ácido nucleico (o una cantidad umbral del reactivo de ácido nucleico) se asoció con la célula. La asociación del reactivo de ácido nucleico con la célula o la cantidad umbral de reactivo de ácido nucleico (por ejemplo, por lo menos 10, 50, 100, 500, 1000, 5000 o 10000 moléculas de reactivo de ácido nucleico) con la célula puede indicar que la célula es una célula muerta.

Aunque en la presente se han divulgado varios aspectos y realizaciones, para los expertos en la técnica resultarán evidentes otros aspectos y realizaciones. Los varios aspectos y realizaciones divulgados en la presente tienen propósitos ilustrativos y no se pretende que sean limitativos.

Un experto en la técnica apreciará que, para este y otros procesos y métodos divulgados en la presente, las funciones realizadas en los procesos y métodos pueden implementarse en diferente orden. Además, los pasos y operaciones indicados solo se proporcionan como ejemplos, y algunos de los pasos y operaciones pueden ser

opcionales, combinarse en menos pasos y operaciones, o ampliarse en pasos y operaciones adicionales sin restar valor a la esencia de las realizaciones divulgadas.

5 Con respecto al uso de sustancialmente cualquier término plural y/o singular en la presente, los expertos en la técnica pueden traducir del plural al singular y/o del singular al plural según sea apropiado para el contexto y/o la aplicación. Las varias permutaciones de singular/plural pueden exponerse expresamente en la presente en aras de la claridad.

10 Los expertos en la técnica entenderán que, en general, los términos usados en la presente, y especialmente en las reivindicaciones adjuntas (por ejemplo, los cuerpos de las reivindicaciones adjuntas), generalmente se entienden como términos "abiertos" (por ejemplo, el término "que incluye" debe interpretarse como "que incluye pero no se limita a", el término "que tiene" debe interpretarse como "que tiene por lo menos", el término "que incluye" debe interpretarse como "que incluye pero no se limita a", etc.). Los especialistas en la técnica comprenderán además que si se pretende un número específico de una mención de reivindicación introducida, dicha
15 intención se enumerará explícitamente en la reivindicación y, en ausencia de dicha mención, dicha intención no estará presente. Por ejemplo, como ayuda para la comprensión, las siguientes reivindicaciones adjuntas pueden contener el uso de las frases introductorias "por lo menos uno" y "uno o más" para introducir enumeraciones de reivindicaciones. Sin embargo, no debe interpretarse el uso de tales frases en el sentido de que la introducción de una mención de reivindicación mediante los artículos indefinidos "un" o "uno" limita cualquier reivindicación particular que contenga dicha mención de reivindicación introducida a las realizaciones que contengan solo una mención de este tipo, incluso cuando la misma reivindicación incluye las frases introductorias "uno o más" o "por lo menos uno" y artículos indefinidos como "un" o "uno" (por ejemplo, "un" y/o "uno" debe interpretarse como "por lo menos uno" o "uno o más"); lo mismo se aplica al uso de artículos definidos para introducir menciones de reivindicaciones. Además, incluso si se menciona explícitamente un número específico de una mención de reivindicación introducida,
20 los expertos en la técnica reconocerán que debe interpretarse que dicha mención como por lo menos el número mencionado (por ejemplo, la simple mención de "dos menciones", sin otros modificadores, significa por lo menos dos menciones, o dos o más menciones). Además, en aquellos casos en los que se usa una convención análoga a "por lo menos uno de A, B y C, etc.", en general, dicha construcción se pretende en el sentido de que un experto en la técnica entendería la convención (por ejemplo, "un sistema que tiene por lo menos uno de A, B y C" incluiría, pero no está limitado a, sistemas que tienen A solo, B solo, C solo, A y B juntos, A y C juntos, B y C juntos, y/o A, B y C juntos, etc.). En aquellos casos en los que se usa una convención análoga a "por lo menos uno de A, B o C, etc.", en general, dicha construcción se pretende en el sentido de que un experto en la técnica entendería la convención (por ejemplo, "un sistema que tiene por lo menos uno de A, B o C" incluiría, pero no estaría limitado a, sistemas que tienen A solo, B solo, C solo, A y B juntos, A y C juntos, B y C juntos, y/o A, B y C juntos, etc.). Los expertos en la técnica comprenderán además que debe entenderse que prácticamente cualquier palabra y/o frase disyuntiva que presente dos o más términos alternativos, ya sea en la descripción, las reivindicaciones o los dibujos, contempla las posibilidades de incluir uno de los términos, cualquiera de los términos, o ambos términos. Por ejemplo, se entenderá que la frase "A o B" incluye las posibilidades de "A" o "B" o "A y B".

40 Además, cuando las características o aspectos de la divulgación se describen en términos de grupos Markush, los expertos en la técnica reconocerán que la divulgación también se describe en términos de cualquier miembro individual o subgrupo de miembros del grupo Markush.

45 Como entenderá un experto en la técnica, para todos y cada uno de los propósitos, como en términos de proporcionar una descripción escrita, todos los intervalos divulgados en la presente también abarcan todos y cada uno de los posibles subintervalos y combinaciones de los mismos. Cualquier intervalo enumerado puede reconocerse fácilmente como una descripción suficiente y permitir que el mismo intervalo se divida en mitades, tercios, cuartos, quintos, décimos, etc., por lo menos iguales. Como ejemplo no limitativo, cada intervalo analizado en la presente puede dividirse fácilmente en un tercio inferior, un tercio medio y un tercio superior, etc. Como
50 también entenderá un experto en la técnica, todos los lenguajes como "hasta", "por lo menos" y similares incluyen el número mencionado y se refieren a intervalos que puede dividirse posteriormente en subintervalos como se ha analizado con anterioridad. Finalmente, como comprenderán los expertos en la técnica, un intervalo incluye cada miembro individual. Así, por ejemplo, un grupo que tiene de 1 a 3 células se refiere a grupos que tienen 1, 2 o 3 células. De manera similar, un grupo que tiene de 1 a 5 células se refiere a grupos que tienen 1, 2, 3, 4 o 5 células, y así sucesivamente.

A partir de lo anterior, se apreciará que en la presente se han divulgado varias realizaciones de la presente divulgación con propósitos ilustrativos. Por consiguiente, no se pretende que las varias realizaciones divulgadas en la presente sean limitativas.

60

REIVINDICACIONES

1. Un método de análisis de muestras, que comprende:

poner en contacto ácido desoxirribonucleico de cadena doble (ADNdc) de una célula con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura, en donde la secuencia de captura comprende una región poli(dT), para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5';

poner en contacto la pluralidad de fragmentos de ADNdc salientes con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a por lo menos una porción del saliente 5', en donde la secuencia complementaria a la secuencia de captura comprende una región poli(dA);

desnaturalizar la pluralidad de fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc); codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden secuencias de marcadores moleculares diferentes, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcadores celulares idéntica;

detectar secuencias de la pluralidad de fragmentos de ADNmc codificados con códigos de barras; y

determinar la información que relaciona las secuencias de ADNdc con la estructura que comprende ADNdc, basándose en secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación.

2. Un método de análisis de muestras, que comprende:

generar una pluralidad de fragmentos de ácidos nucleicos a partir de ácido desoxirribonucleico de cadena doble (ADNdc) de una célula, en donde generar la pluralidad de fragmentos de ácidos nucleicos comprende poner en contacto el ADNdc con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y dos copias de un adaptador que comprende la secuencia de captura, en donde la secuencia de captura comprende una región poli(dT), para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a la secuencia de captura, en donde la secuencia complementaria a la secuencia de captura comprende una región poli(dA), y en donde cada uno de la pluralidad de fragmentos de ácidos nucleicos comprende una secuencia de captura, un complemento de la secuencia de captura, un complemento inverso de la secuencia de captura, o una combinación de los mismos;

codificar con códigos de barras la pluralidad de fragmentos de ácidos nucleicos usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADN codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprende diferentes secuencias de marcadores moleculares, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcadores celulares idéntica; y

detectar secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras.

3. Un método de análisis de muestras, que comprende:

generar una pluralidad de fragmentos de ácidos nucleicos a partir de ácido desoxirribonucleico de cadena doble (ADNdc) de una célula, en donde generar la pluralidad de fragmentos de ácidos nucleicos comprende:

poner en contacto el ADNdc con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura, en donde la secuencia de captura comprende una región poli(dT), para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5'; y

poner en contacto la pluralidad de fragmentos de ADNdc salientes que comprenden los salientes 5' con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a por lo menos una parte de los salientes 5', en donde la secuencia complementaria a la secuencia de captura comprende una región poli (dA);

en donde cada uno de la pluralidad de fragmentos de ácidos nucleicos comprende una secuencia de captura, un complemento de la secuencia de captura, un complemento inverso de la secuencia de captura o una combinación de los mismos;

codificar con códigos de barras la pluralidad de fragmentos de ácidos nucleicos usando una pluralidad de

códigos de barras para generar una pluralidad de fragmentos de ADN codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprende diferentes secuencias de marcadores moleculares, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcadores celulares idéntica; y detectar secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras.

4. El método de la reivindicación 2 o 3, que comprende además determinar la información que relaciona las secuencias de ADNdc con una estructura que comprende el ADNdc basándose en secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras.

5. El método de una cualquiera de las reivindicaciones 2-4, en donde generar la pluralidad de fragmentos de ácidos nucleicos comprende fragmentar el ADNdc para generar una pluralidad de fragmentos de ADNdc y, opcionalmente, agregar dos copias de un adaptador que comprende una secuencia complementaria a una secuencia de captura a la pluralidad de fragmentos de ADNdc para generar una pluralidad de fragmentos de ácidos nucleicos, opcionalmente en donde agregar las dos copias del adaptador comprende ligar las dos copias del adaptador a la pluralidad de fragmentos de ADNdc para generar la pluralidad de fragmentos de ácidos nucleicos que comprenden el adaptador.

6. El método de cualquiera de las reivindicaciones 1 o 4-5, en donde determinar la información que relaciona la secuencia de ADNdc con la estructura comprende:

(a) determinar la accesibilidad a la cromatina del ADNdc basándose en las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación obtenidos, opcionalmente en donde determinar la accesibilidad a la cromatina del ADNdc comprende:

(i) alinear las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras con una secuencia de referencia del ADNdc; e

identificar regiones del ADNdc correspondientes a los extremos de los fragmentos de ADN codificados con códigos de barras de la pluralidad de fragmentos de ADN codificados con códigos de barras para tener una accesibilidad por encima de un umbral; o

(ii) alinear las secuencias de la pluralidad de fragmentos de ADNmc con una secuencia de referencia del ADNdc; y

determinar la accesibilidad de las regiones del ADNdc correspondientes a los extremos de los fragmentos de ADNmc de la pluralidad de fragmentos de ADNmc basándose en los números de los fragmentos de ADNmc de la pluralidad de fragmentos de ADNmc en los datos de secuenciación; y/o

(b) determinar la información del metiloma del ADNdc basándose en las secuencias de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación obtenidos, opcionalmente:

(i) que comprende además:

(A) digerir nucleosomas asociados con el ADNdc de la célula; y/o

(B) conversión de bisulfito de bases de citosina de una pluralidad de ADN de cadena sencilla de la pluralidad de fragmentos de ADN salientes o pluralidad de fragmentos de ácidos nucleicos para generar una pluralidad de ADNmc convertidos con bisulfito que comprenden bases de uracilo; y/o

(ii) en donde determinar la información del metiloma comprende:

determinar si una posición de la pluralidad de fragmentos de ADN codificados con códigos de barras en los datos de secuenciación tiene una base de timina y la posición correspondiente en una secuencia de referencia del ADNdc tiene una base de citosina para determinar la posición correspondiente en el ADNdc que tiene una base de metilcitosina.

7. El método de cualquiera de las reivindicaciones 1-6, en donde codificar con códigos de barras comprende:

(a) codificar con códigos de barras estocásticamente la pluralidad de fragmentos de ADN o la pluralidad de fragmentos de ácidos nucleicos usando la pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados estocásticamente; o

(b) codificar con códigos de barras la pluralidad de fragmentos de ADN o la pluralidad de fragmentos de ácidos nucleicos usando la pluralidad de códigos de barras asociados con una partícula para generar la pluralidad de fragmentos de ADNmc codificados con códigos de barras, en donde los códigos de barras asociados con la partícula comprenden una secuencia de marcador celular idéntica y por lo menos 100 secuencias de marcadores moleculares diferentes, opcionalmente en donde:

(i) por lo menos un código de barras de la pluralidad de códigos de barras:

- (A) se inmoviliza sobre la partícula;
- (B) se inmoviliza parcialmente sobre la partícula;
- (C) se encierra en la partícula; o
- (D) se encierra parcialmente en la partícula;

5

y opcionalmente

(ii) en donde la partícula:

- (A) es alterable; o
- (B) comprende una partícula de hidrogel alterable.

10

8. El método de la reivindicación 7(b), en donde:

- (a) la partícula comprende un material seleccionado del grupo que consiste en polidimetilsiloxano (PDMS), poliestireno, vidrio, polipropileno, agarosa, gelatina, hidrogel, paramagnético, cerámica, plástico, vidrio, metilestireno, polímero acrílico, titanio, látex, sefarosa, celulosa, nailon, silicona y cualquier combinación de los mismos; y/o
- (b) los códigos de barras de la partícula comprenden marcadores moleculares con por lo menos 1000 secuencias de marcadores moleculares diferentes, opcionalmente en donde los códigos de barras de la partícula comprenden marcadores moleculares con por lo menos 10000 secuencias de marcadores moleculares diferentes.

15

20

9. El método de cualquiera de las reivindicaciones 1-8, en donde el ADNdc de la célula se selecciona del grupo que consiste en: ADN nuclear, ADN nucleolar, ADN genómico, ADN mitocondrial, ADN de cloroplasto, ADN de constructo, ADN viral o una combinación de dos o más de los elementos enumerados.

25

10. Un método de análisis de muestras, que comprende:

- poner en contacto ácido desoxirribonucleico de cadena doble (ADNdc) de una célula con un transposoma, en donde el transposoma comprende una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc y dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura, en donde la secuencia de captura comprende una región poli(dT), para generar una pluralidad de fragmentos de ADNdc salientes, cada uno de los cuales comprende dos copias de los salientes 5';
- poner en contacto la pluralidad de fragmentos de ADNdc salientes con una polimerasa para generar una pluralidad de fragmentos de ADNdc complementarios, cada uno de los cuales comprende una secuencia complementaria a por lo menos una porción de cada uno de los salientes 5', en donde la secuencia complementaria a la secuencia de captura comprende una región poli(dA);
- desnaturalizar la pluralidad de fragmentos de ADNdc complementarios para generar una pluralidad de fragmentos de ADN de cadena sencilla (ADNmc);
- codificar con códigos de barras la pluralidad de fragmentos de ADNmc usando una pluralidad de códigos de barras para generar una pluralidad de fragmentos de ADNmc codificados con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden diferentes secuencias de marcadores moleculares, y en donde si la pluralidad de códigos de barras comprende una secuencia de marcadores celulares idéntica;
- obtener datos de secuenciación de la pluralidad de fragmentos de ADNmc codificados con códigos de barras; y
- cuantificar una cantidad de ADNdc en la célula sobre la base de una cantidad de secuencias de marcadores moleculares únicas asociadas con la misma secuencia de marcador celular.

30

35

40

45

50

11. El método de la reivindicación 10, que comprende además:

- capturar un fragmento de ADNmc de la pluralidad de fragmentos de ADNmc en un soporte sólido que comprende un oligonucleótido que comprende la secuencia de captura, la secuencia del marcador celular y la secuencia del marcador molecular, en donde la secuencia de captura comprende una secuencia poli dT que se une a una cola poli A en el fragmento de ADNmc;
- extender el fragmento de ADNmc en la dirección de 5' a 3' para producir el fragmento de ADNmc codificado con códigos de barras, el ADNmc codificado con códigos de barras comprendiendo la secuencia de captura, la secuencia de marcador molecular y la secuencia de marcador celular;
- extender el oligonucleótido en la dirección de 5' a 3' usando una transcriptasa o polimerasa inversa o una combinación de las mismas para producir una cadena de ADN complementaria al ADNmc codificado con códigos de barras;
- desnaturalizar el ADNmc codificado con códigos de barras y la cadena de ADN complementaria para producir secuencias de cadena sencilla; y
- amplificar las secuencias de cadena sencilla; opcionalmente que comprende además una conversión de bisulfito

55

60

65

de bases de citosina de la pluralidad de fragmentos de ADNmc para generar una pluralidad de fragmentos de ADNmc convertidos con bisulfito que comprenden bases de uracilo.

12. Un kit para el análisis de muestras, que comprende:

5 un transposoma que comprende:

una nucleasa de cadena doble configurada para inducir una rotura de ADN de cadena doble en una estructura que comprende ADNdc; y

10 dos copias de un adaptador que tiene un saliente 5' que comprende una secuencia de captura, en donde la secuencia de captura comprende una región poli(dT); y

una pluralidad de códigos de barras, cada código de barras puede comprender una secuencia de marcador celular, una secuencia de marcador molecular y la secuencia de captura, en donde por lo menos dos de la pluralidad de códigos de barras comprenden diferentes secuencias de marcadores moleculares, y en donde por lo menos dos de la pluralidad de códigos de barras comprenden una secuencia de marcador celular idéntica;

15 opcionalmente en donde:

(a) la nucleasa de cadena doble comprende una transposasa, una endonucleasa de restricción, una proteína asociada a CRISPR, una nucleasa específica de dúplex (DSN) o una combinación de estas;

20 (b) el transposoma comprende además una ligasa;

(c) la pluralidad de códigos de barras comprende por lo menos 10, 50, 100, 500, 1000, 5000, 10000, 50000 o 100000 marcadores moleculares diferentes; y/o

(d) los códigos de barras:

25 (i) se inmovilizan en partículas, opcionalmente en donde todos los códigos de barras inmovilizados en cada partícula comprenden la misma secuencia de marcador celular, y en donde diferentes partículas comprenden diferentes secuencias de marcadores celulares; o

(ii) se dividen en pocillos de un sustrato, opcionalmente en donde todos los códigos de barras divididos en cada pocillo comprenden la misma secuencia de marcador celular, y en donde diferentes pocillos comprenden secuencias de marcadores celulares diferentes.

30

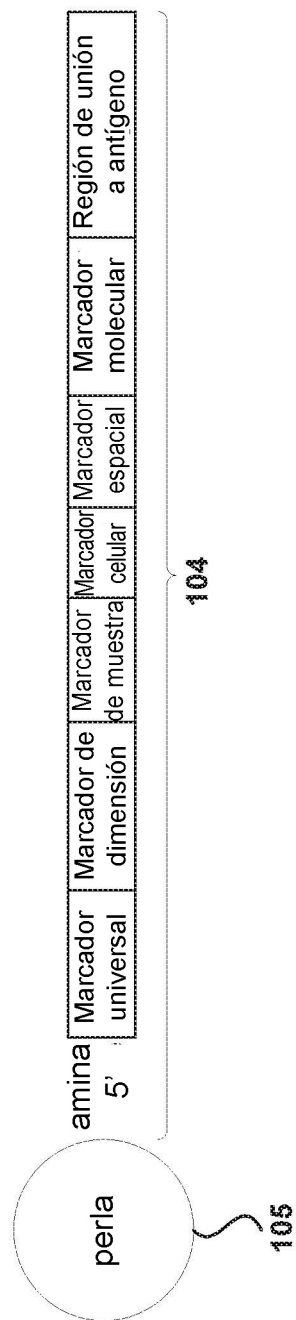


FIG. 1

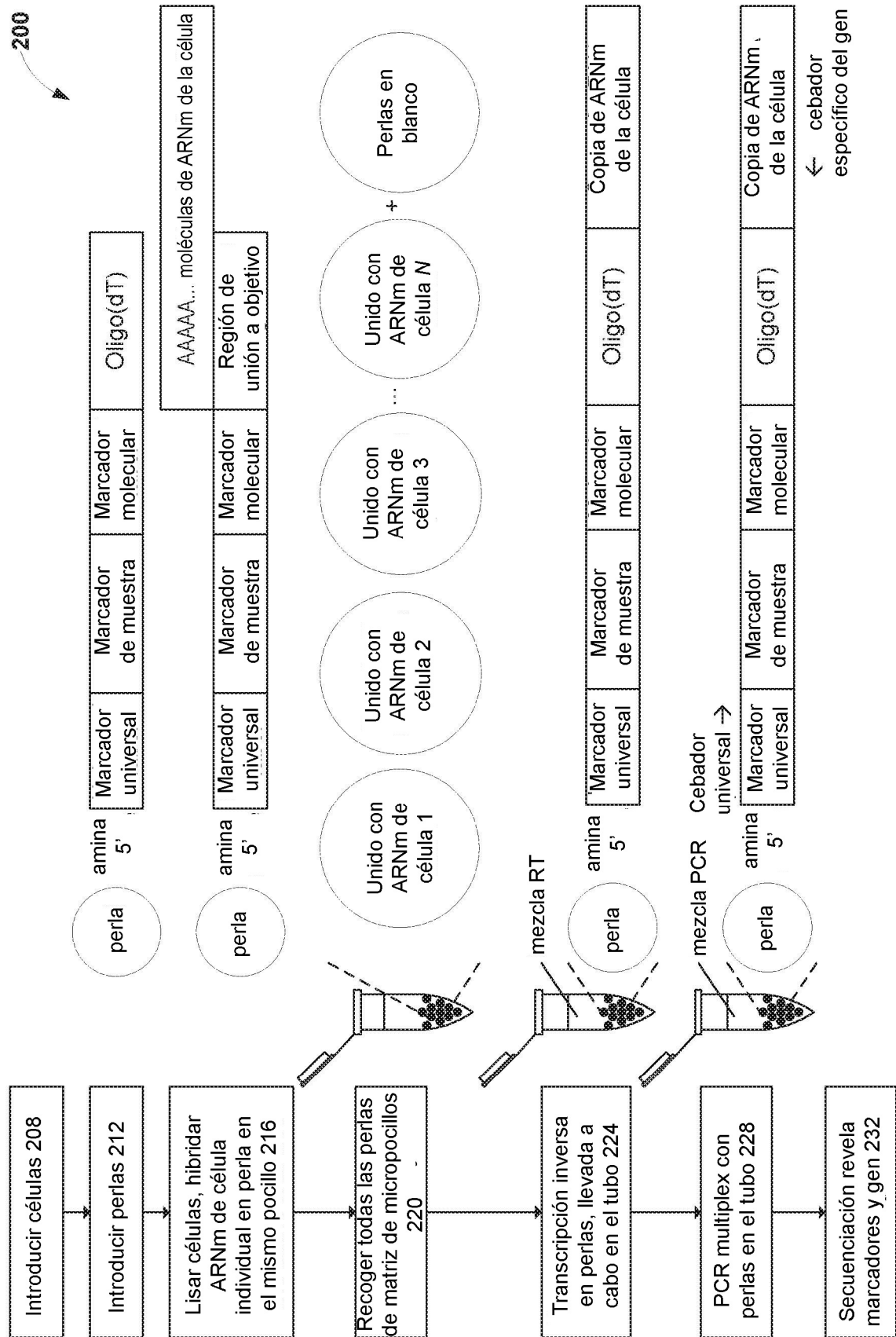


FIG. 2

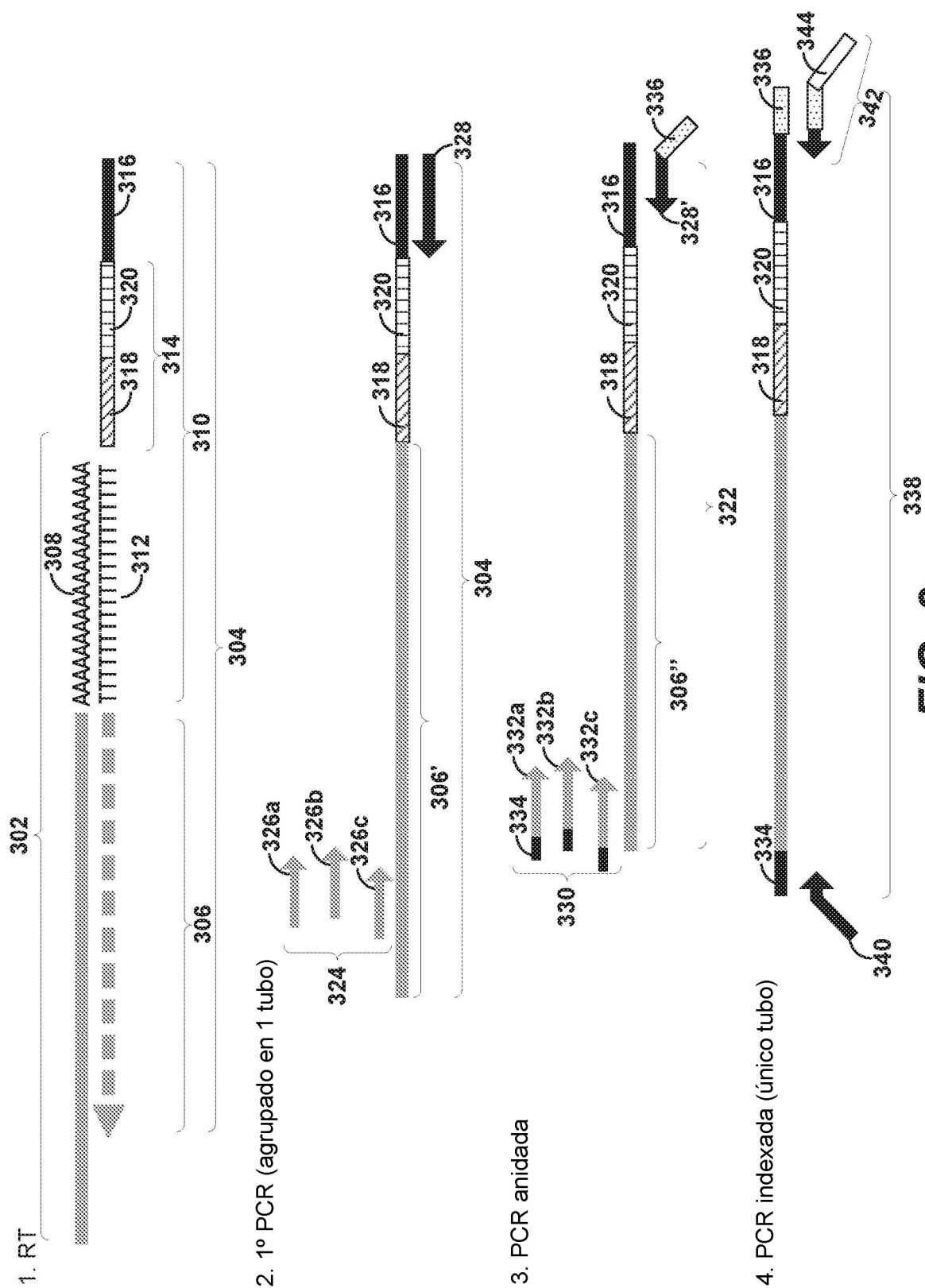
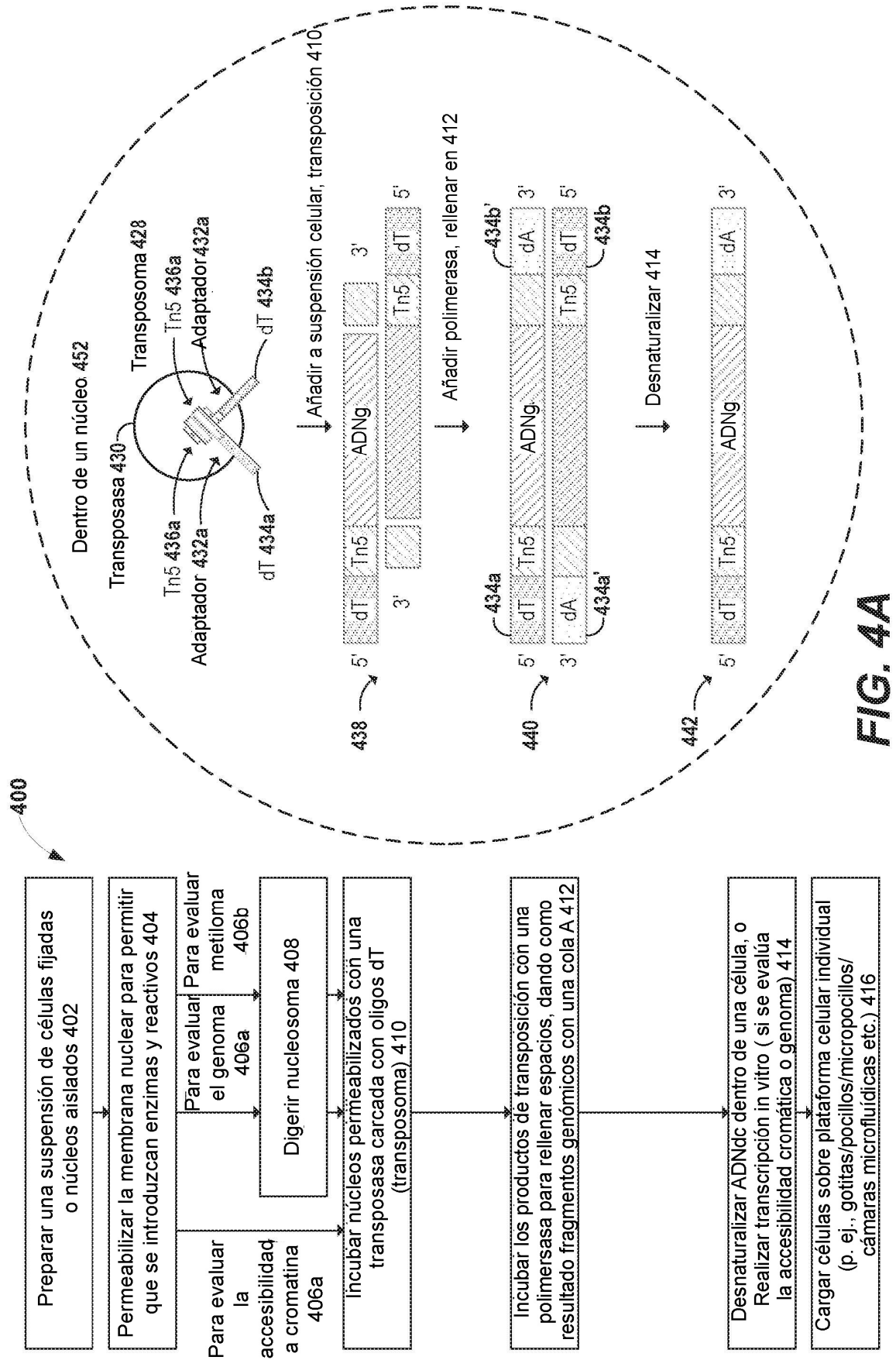


FIG. 3



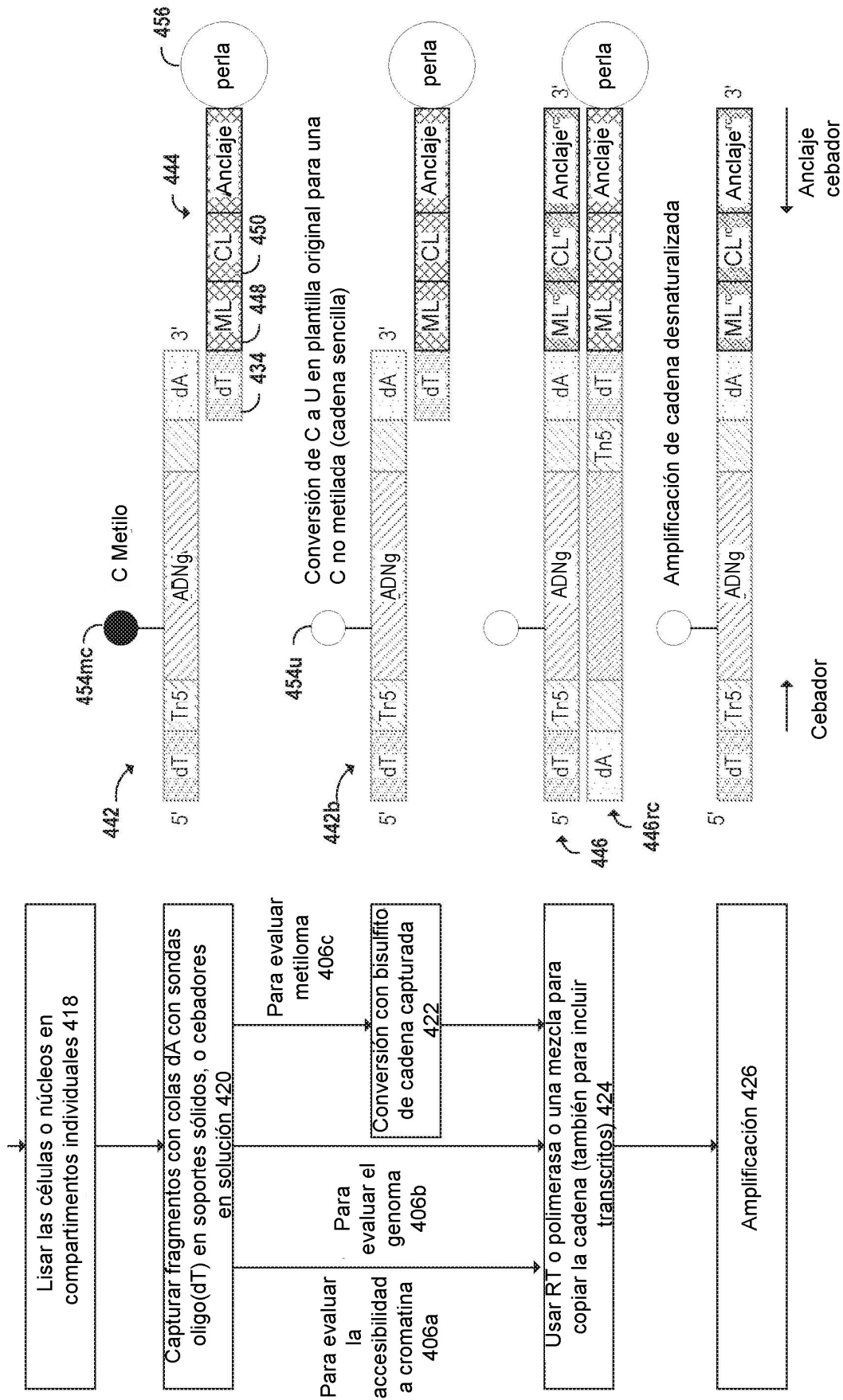


FIG. 4B

FIG. 5A

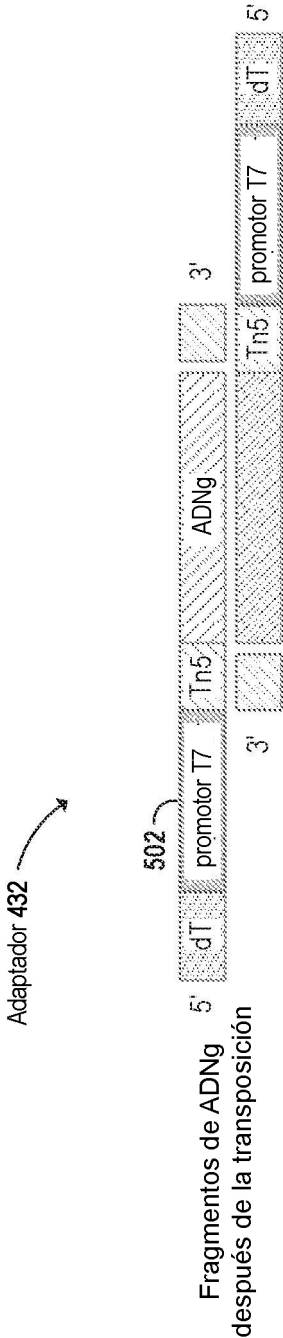
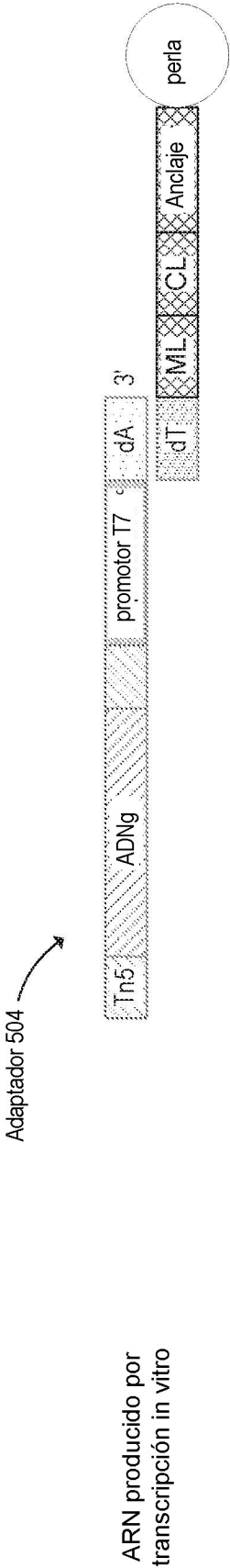


FIG. 5B



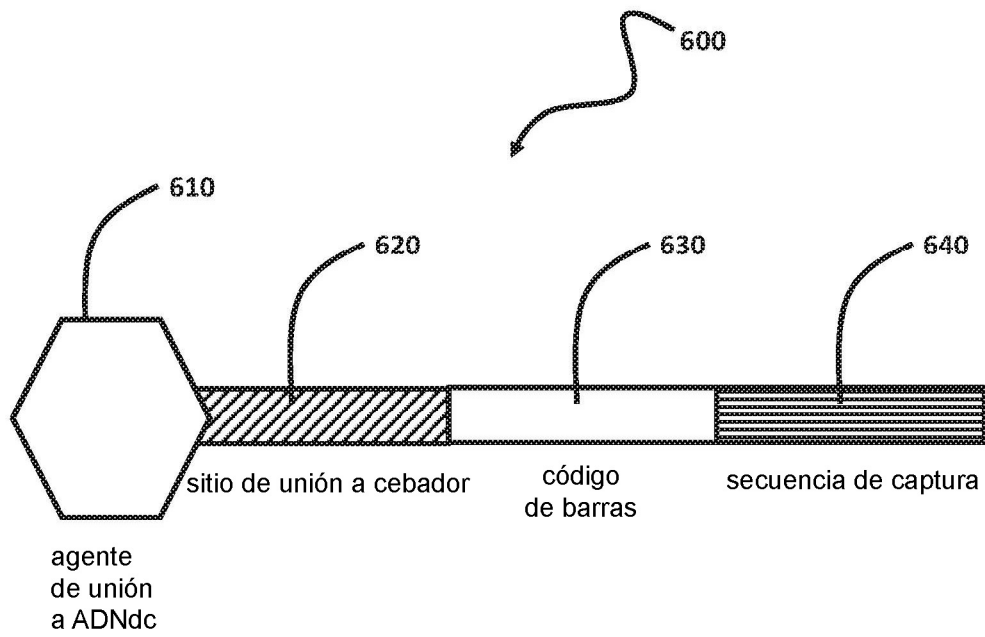


FIG. 6