

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 989 072**

51 Int. Cl.:

**C12N 15/10** (2006.01)

**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: **29.07.2019** **PCT/US2019/043949**

87 Fecha y número de publicación internacional: **06.02.2020** **WO20028266**

96 Fecha de presentación y número de la solicitud europea: **29.07.2019** **E 19752792 (2)**

97 Fecha y número de publicación de la concesión europea: **28.08.2024** **EP 3830262**

54 Título: **Identificación con códigos de barras de núcleos y captura en células individuales**

30 Prioridad:

**03.08.2018 US 201862714222 P**

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

**25.11.2024**

73 Titular/es:

**BECTON DICKINSON AND COMPANY (100.0%)**  
**1 Becton Drive**  
**Franklin Lakes, NJ 07417-1880, US**

72 Inventor/es:

**WALCZAK, ELISABETH, MARIE;**  
**SHUM, ELEEN;**  
**CHANG, CHRISTINA y**  
**FAN, CHRISTINA**

74 Agente/Representante:

**IZQUIERDO BLANCO, María Alicia**

ES 2 989 072 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

Identificación con códigos de barras de núcleos y captura en células individuales

## 5 ANTECEDENTES

**Campo**

10 La presente divulgación se refiere de manera general al campo de la identificación con códigos de barras moleculares de dianas y más particularmente a la identificación con códigos de barras de núcleos.

**Descripción de la técnica relacionada**

15 Los métodos y técnicas como la identificación con códigos de barras usando perlas acopladas a códigos de barras de ácidos nucleicos son útiles para el análisis de células individuales, como descifrar perfiles de expresión génica para determinar los estados de células individuales usando, por ejemplo, la transcripción inversa, la amplificación por reacción en cadena de la polimerasa (PCR) y la secuenciación de próxima generación (NGS). Sin embargo, la preparación de suspensiones de células individuales para el análisis de células individuales puede ser desafiante.

20 Preissl, S. et al. (2018) Nat Neurosci 21(3):432-9 describe el análisis de un solo núcleo de cromatina accesible en el cerebro anterior de ratón en desarrollo, que revela la regulación transcripcional específica de cada tipo celular. Cao, J. et al. (2017) Science 357(6352):661-7 describe la realización de perfiles transcripcional de células individuales completa de un organismo multicelular. Hu, P. et al. (2017) Molecular Cell 68(5):1006-15 describe la disección de la composición del tipo de célula y el estado transcripcional dependiente de la actividad en cerebros de mamíferos mediante Sec. de ARN de núcleos individuales masivamente paralela. Kausch, A. et al. (1999) Biotechniques Rapid Dispatches 26(2):336-43 describe el aislamiento de orgánulos mediante inmunoabsorción magnética.

## SUMARIO

30 La invención se expone en el conjunto de reivindicaciones adjunto. Los términos "realización" y "realizaciones" deben interpretarse como realización o realizaciones de la invención sólo en la medida en que entren dentro del alcance de las presentes reivindicaciones. De lo contrario, se refieren únicamente a realizaciones de la divulgación.

35 Lo divulgado en la presente incluye realizaciones de un método para determinar el número de dianas en una pluralidad de células. En algunas realizaciones, el método comprende: aislar una pluralidad de núcleos de una pluralidad de células usando una composición de aislamiento de núcleos, en donde la composición de aislamiento de núcleos comprende un reactivo de unión a núcleos, y en donde el reactivo de unión a núcleos es capaz de unirse específicamente a uno o más componentes de un núcleo; identificar con códigos de barras una pluralidad de dianas en la pluralidad de núcleos usando una pluralidad de códigos de barras para generar una pluralidad de dianas identificadas con códigos de barras, en donde cada uno de los códigos de barras comprende una secuencia de marcador molecular y una región de unión a la diana, y en donde las secuencias de marcadores moleculares de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden secuencias diferentes; obtener datos de secuenciación de la pluralidad de dianas identificadas con códigos de barras; y estimar el número de cada una de la pluralidad de dianas en la pluralidad de células usando las secuencias de marcadores moleculares de la pluralidad de códigos de barras en los datos de secuenciación.

50 En algunas realizaciones, aislar la pluralidad de núcleos comprende: poner en contacto la pluralidad de núcleos de la pluralidad de células con la composición de aislamiento de núcleos para generar núcleos unidos al reactivo de unión a núcleos. El aislamiento de la pluralidad de núcleos puede comprender: aislar los núcleos unidos al reactivo de unión a núcleos usando un reactivo capaz de unirse específicamente al reactivo de unión a núcleos.

55 En algunas realizaciones, el reactivo de unión a núcleos está asociado a un primer epítipo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un reactivo de unión al primer epítipo. El primer epítipo puede comprender biotina, hapteno o una combinación de los mismos. El hapteno puede comprender digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender un anticuerpo antihapteno. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender avidina, estreptavidina, neutravidina o una combinación de los mismos. En algunas realizaciones, el reactivo de unión a núcleos comprende un grupo funcional seleccionado del grupo que consiste en biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, o una combinación de los mismos. En algunas realizaciones, el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un grupo funcional seleccionado del grupo que consiste en biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias,

carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, o una combinación de los mismos.

En algunas realizaciones, el reactivo de unión a núcleos comprende un anticuerpo primario capaz de unirse específicamente al uno o más componentes del núcleo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario.

En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a carbohidratos. El reactivo de unión a carbohidratos puede comprender una proteína de unión a carbohidratos. La proteína de unión a carbohidratos puede comprender una lectina. La lectina puede comprender una lectina de unión a manosa, una lectina de unión a galactosa, una lectina de unión a N-acetilgalactosamina, una lectina de unión a N-acetilglucosamina, una lectina de unión a ácido N-acetilneuramínico, una lectina de unión a fucosa o una combinación de las mismas. La lectina puede comprender Concanavalina A (ConA), lectina de lenteja (LCH), lectina de campanilla de invierno (GNA), aglutinina de *Ricinus communis* (RCA), aglutinina de cacahuete (PNA), jacalina (AEL), lectina de veza peluda (VVL), aglutinina de germen de trigo (WGA), lectina del saúco (SNA), leucoaglutinina de *Maackia amurensis* (MAL), hemoaglutinina de *Maackia amurensis* (MAH), aglutinina de *Ulex europaeus* (UEA), lectina de *Aleuria aurantia* (AAL), o una combinación de las mismas. La lectina puede ser, o comprender, una aglutinina. La aglutinina puede ser, o comprender, aglutinina de germen de trigo (WGA). La proteína de unión a carbohidratos puede proceder o derivarse de un animal, una bacteria, un virus, un hongo o una combinación de los mismos. La proteína de unión a carbohidratos puede proceder o derivarse de una planta. La planta puede ser *Canavalia ensiformis*, *Lens culinaris*, *Galanthus nivalis*, *Ricinus communis*, *Arachis hypogaea*, *Artocarpus integrifolia*, *Vicia villosa*, *Triticum vulgare*, *Sambucus nigra*, *Maackia amurensis*, *Ulex europaeus*, *Aleuria aurantia*, o una combinación de las mismas.

En algunas realizaciones, los uno o más componentes del núcleo comprenden un azúcar, un oligosacárido, un polisacárido, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender un monosacárido, un disacárido, un poliol, un malto-oligosacárido, un no-malto-oligosacárido, un almidón, un polisacárido no de almidón, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender glucosa, galactosa, fructosa, xilosa, sacarosa, lactosa, maltosa, trehalosa, sorbitol, manitol, maltodextrina, rafinosa, estaquiosa, fructooligosacárido, amilosa, amilopectina, almidón modificado, glucógeno, celulosa, hemicelulosa, pectina, hidrocoloide, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender un residuo de  $\alpha$ -D-manosilo, un residuo de  $\alpha$ -D-glucosilo, una estructura  $\alpha$ -manosídica ramificada de tipo  $\alpha$ -manosa alta, una estructura  $\alpha$ -manosídica ramificada de tipo híbrido y N-Glicano de tipo complejo biantenar, una región central fucosilada de N-Glicano de tipo complejo bi- y triantenar, una estructura de manosa alta enlazada  $\alpha$  1-3 y  $\alpha$  1-6, Gal $\beta$ 1-4GalNAc $\beta$ 1-R, Gal $\beta$ 1-3GalNAc $\alpha$ 1-Ser/Thr, (Sia)Gal $\beta$ 1-3GalNAc $\alpha$ 1-Ser/Thr, GalNAc $\alpha$ -Ser/Thr, GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc, Neu5Ac (ácido siálico), Neu5Ac $\alpha$ 2-6Gal(NAc)-R, Neu5Ac/Gc $\alpha$ 2,3Gal $\beta$ 1,4Glc(NAc), NeuSAc/Gc $\alpha$ 2,3Gal $\beta$ 1,3(Neu5Ac $\alpha$ 2,6)GalNAc, Fuca1-2Gal-R, Fuca1-2Gal $\beta$ 1-4(Fuca1-3/4)Gal $\beta$ 1-4GlcNAc, R2-GlcNAc $\beta$ 1-4(Fuca1-6)GlcNAc-R1, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender una glicoproteína, un glicolípido o una combinación de los mismos.

En algunas realizaciones, los uno o más componentes del núcleo comprenden Lamin, Emerin, Nesprin, Nurim, UNC-83, Klar, ZYG-12, Kms1p, UNC-84, Klaroid, SUN-1, Sad1p, LBR, MAN1, LAP1, LAP2, LINK, un complejo de poro nuclear, una porción del mismo, o una combinación de los mismos.

En algunas realizaciones, el reactivo de unión a núcleos está asociado a una partícula de aislamiento del núcleo. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede estar asociado con la partícula de aislamiento del núcleo. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede estar inmovilizado, o parcialmente inmovilizado, en la partícula de aislamiento del núcleo. Por ejemplo, el reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede asociarse de manera reversible, no reversible, covalente, no covalente, o una combinación de las mismas, a la partícula de aislamiento del núcleo. Como otro ejemplo, el reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede estar insertado, parcialmente insertado, no insertado, encerrado, parcialmente encerrado, no encerrado, o una combinación de los mismos, en la partícula de aislamiento del núcleo. En algunas realizaciones, el reactivo capaz de unirse específicamente al reactivo de unión a núcleos se asocia con la partícula de aislamiento del núcleo a través de un conector escindible. El conector escindible puede, por ejemplo, comprender un conector escindible químicamente, un enlace fotoescindible, un conector lábil al ácido, un enlace sensible al calor, un enlace escindible enzimáticamente, o una combinación de los mismos.

En algunas realizaciones, la partícula de aislamiento del núcleo comprende una perla de aislamiento del núcleo. La partícula de aislamiento del núcleo 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 antifluorocromo, o cualquier combinación de las mismas. La partícula de aislamiento del núcleo puede comprender un material seleccionado del grupo que consiste en polidimetilsiloxano (PDMS), poliestireno, vidrio, polipropileno, agarosa, gelatina, hidrogel, paramagnético, cerámica, plástico, vidrio, metileno, polímero acrílico, titanio, látex,

sefarosa, celulosa, nailon, silicona y cualquier combinación de los mismos. La partícula de aislamiento del núcleo puede ser alterable. La partícula de aislamiento del núcleo puede comprender una partícula de hidrogel alterable de aislamiento del núcleo. En algunas realizaciones, el aislamiento de los núcleos unidos al reactivo de unión a núcleos comprende poner en contacto los núcleos unidos al reactivo de unión a núcleos con una pluralidad de partículas de aislamiento de núcleos. En algunas realizaciones, el método comprende, antes de poner en contacto los núcleos unidos al reactivo de unión a núcleos con una pluralidad de partículas de aislamiento de núcleos, poner en contacto los núcleos unidos al reactivo de unión a núcleos con una pluralidad de partículas nulas. En algunas realizaciones, la proporción de partículas nulas con respecto a las partículas de aislamiento de núcleos es de por lo menos 10:1. La partícula nula, en algunas realizaciones, no comprende una propiedad magnética. En algunas realizaciones, la partícula nula no comprende un reactivo capaz de unirse específicamente al reactivo de unión a núcleos. En algunas realizaciones, la partícula nula 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, nylon, silicona, o una combinación de los mismos. En algunas realizaciones, la pluralidad de partículas nulas reduce la aglutinación de las partículas de aislamiento del núcleo, o evita la aglutinación de las partículas de aislamiento del núcleo, o ambas cosas. En algunas realizaciones, poner en contacto los núcleos unidos al reactivo de unión a núcleos con una pluralidad de partículas de aislamiento de núcleos genera una pluralidad de núcleos unidos a partículas de aislamiento de núcleos a través del reactivo capaz de unirse específicamente al reactivo de unión a núcleos. En algunas realizaciones, el porcentaje de núcleos unidos a una única partícula de aislamiento del núcleo después de poner en contacto los núcleos unidos al reactivo de unión a núcleos con una pluralidad de partículas de aislamiento de núcleos es de por lo menos el 90%, por lo menos el 95% o por lo menos el 99%. En algunas realizaciones, el porcentaje de partículas de aislamiento de núcleos unidas a un único núcleo o a ningún núcleo después de poner en contacto los núcleos unidos al reactivo de unión a núcleos con una pluralidad de partículas de aislamiento de núcleos es de por lo menos el 90%, por lo menos el 95% o por lo menos el 99%.

En algunas realizaciones, aislar los núcleos unidos al reactivo de unión a núcleos puede comprender: aislar la partícula de aislamiento del núcleo mediante extracción magnética, centrifugación, o cualquier combinación de los mismos. En algunas realizaciones, el método comprende, antes de aislar la pluralidad de partículas de aislamiento de núcleos, poner en contacto la pluralidad de núcleos unidos a partículas de aislamiento de núcleos con el primer epítipo libre, en donde el primer epítipo libre comprende la totalidad o una parte del primer epítipo. En algunas realizaciones, la pluralidad de códigos de barras está asociada con la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede inmovilizarse en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente inmovilizado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar encerrado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente encerrado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede no estar encerrado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar insertado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente insertado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede no estar insertado en la partícula de aislamiento del núcleo.

En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a componentes de la superficie de la envoltura nuclear, y el reactivo de unión a núcleos puede ser capaz de unirse específicamente a uno o más componentes de la superficie de la envoltura nuclear.

En algunas realizaciones, el método comprende: antes de identificar con códigos de barras la pluralidad de dianas en la pluralidad de núcleos usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras, lisar la pluralidad de núcleos. En algunas realizaciones, el método comprende: antes de aislar la pluralidad de núcleos de la pluralidad de células usando la composición de aislamiento de núcleos, lisar la pluralidad de células sin lisar núcleos de la pluralidad de células. En algunas realizaciones, el método comprende antes de lisar la pluralidad de núcleos, dividir la pluralidad de núcleos. En algunas realizaciones, la división de la pluralidad de núcleos comprende dividir la pluralidad de núcleos en una pluralidad de divisiones, en las que una división de la pluralidad de divisiones comprende un único núcleo de la pluralidad de núcleos. En algunas realizaciones, la pluralidad de divisiones comprende micropocillos de una matriz de micropocillos. En algunas realizaciones, la pluralidad de divisiones comprende una pluralidad de gotitas. En algunas realizaciones, dos o más divisiones de la pluralidad de divisiones comprenden un único núcleo. En algunas realizaciones, una división de la pluralidad de divisiones comprende un único núcleo y una única partícula de aislamiento de núcleo. En algunas realizaciones, una división de la pluralidad de divisiones comprende un único núcleo y una única partícula de identificación con códigos de barras. En algunas realizaciones, una división de la pluralidad de divisiones comprende un único núcleo, una única partícula de aislamiento del núcleo y una única partícula de identificación con códigos de barras.

En algunas realizaciones, el método comprende: antes de aislar la pluralidad de núcleos de la pluralidad de células usando la composición de aislamiento de núcleos, permeabilizar (por ejemplo, extendiendo la membrana plasmática de) la pluralidad de células; y agotar uno o más orgánulos de la pluralidad de células usando una

composición de captura de orgánulos que comprende un reactivo de unión a orgánulos, en donde el reactivo de unión a orgánulos es capaz de unirse específicamente a uno o más componentes de uno o más orgánulos de la pluralidad de células. El agotamiento de los uno o más orgánulos puede comprender: poner en contacto los uno o más orgánulos de la pluralidad de células con la composición de captura de orgánulos para generar uno o más orgánulos unidos al reactivo de unión a componentes de orgánulos. El agotamiento de los uno o más orgánulos puede comprender: el agotamiento de los uno o más orgánulos unidos al reactivo de unión a orgánulos usando un reactivo capaz de unirse específicamente al reactivo de unión a orgánulos. El reactivo de unión a orgánulos puede estar asociado a un segundo epítipo, y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un segundo reactivo de unión a epítipo.

En algunas realizaciones, el segundo epítipo puede comprender biotina, hapteno o una combinación de los mismos. El hapteno puede comprender digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos. El reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un anticuerpo antihapteno. El reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender avidina, estreptavidina, neutravidina o una combinación de los mismos. El reactivo de unión a orgánulos puede comprender un anticuerpo primario capaz de unirse específicamente a uno o más componentes de uno o más orgánulos de la pluralidad de células, y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario.

En algunas realizaciones, el reactivo de unión a orgánulos se asocia con una partícula de captura de orgánulos. El reactivo de unión a orgánulos puede estar inmovilizado, o parcialmente inmovilizado, en la partícula de captura de orgánulos. Por ejemplo, el reactivo de unión a orgánulos puede estar asociado de manera reversible, no reversible, covalente, no covalente, o una combinación de las mismas, con la partícula de captura de orgánulos. Como otro ejemplo, el reactivo de unión a orgánulos puede estar insertado, parcialmente insertado, no insertado, encerrado, parcialmente encerrado, no encerrado, o una combinación de los mismos, en la partícula de captura de orgánulos.

En algunas realizaciones, la partícula de captura de orgánulos comprende una perla de captura de orgánulos. La partícula de captura de orgánulos 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 antifluorocromo, o cualquier combinación de las mismas. La partícula de captura de orgánulos 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, nylon, silicona, y cualquier combinación de los mismos.

En algunas realizaciones, el agotamiento de los orgánulos de la pluralidad de células usando la composición de captura de orgánulos puede comprender: el agotamiento de una o más partículas de captura de orgánulos mediante extracción magnética, centrifugación o cualquier combinación de las mismas. Los orgánulos pueden comprender mitocondrias de la pluralidad de células. Los uno o más componentes de los uno o más orgánulos de la pluralidad de células puede comprender: ABCD3, ESR2, NOS3, ALB, HIF1A, NR3C1, ATP5A1, HK1, PGR, CASQ1, HSPA1A, PHB, CLTC, HSPD1, PLN, COX4I1, IFM1, SOD1, CPS1, LGALS3, TP53, Citocromo C Oxidasa, MAPT, TP5B, ERN1, MT-CO1, VDAC1, o una combinación de los mismos.

En algunas realizaciones, el reactivo de unión a orgánulos puede comprender un reactivo de unión a componentes de superficie de orgánulos, uno o más componentes de los uno o más orgánulos pueden comprender uno o más componentes de superficie de orgánulos, y el reactivo de unión a orgánulos puede ser capaz de unirse específicamente al uno o más componentes de superficie de orgánulos. El reactivo de unión a núcleos puede comprender un oligonucleótido de indexación de núcleos, y en donde el oligonucleótido de indexación de núcleos comprende una secuencia de indexación de núcleos.

En algunas realizaciones, el método comprende: identificar con códigos de barras los oligonucleótidos de indexación de núcleos usando la pluralidad de códigos de barras para generar una pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras; y obtener datos de secuenciación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. La identificación con códigos de barras de los oligonucleótidos de indexación de núcleos puede comprender: identificar estocásticamente con códigos de barras usando la pluralidad de códigos de barras para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. Identificar con códigos de barras los oligonucleótidos de indexación de núcleos usando la pluralidad de códigos de barras puede comprender: poner en contacto la pluralidad de códigos de barras con los oligonucleótidos de indexación de núcleos para generar códigos de barras hibridados con los oligonucleótidos de indexación de núcleos; y extender los códigos de barras hibridados con los oligonucleótidos de indexación de núcleos para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. Extender los códigos de barras puede comprender: extender los códigos de barras usando una ADN polimerasa para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. Extender los códigos de barras puede comprender: extender los códigos de barras usando una transcriptasa

inversa para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras.

En algunas realizaciones, el método puede comprender: amplificar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras para producir una pluralidad de amplicones de indexación de núcleos identificados con códigos de barras. La amplificación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras puede comprender: amplificar, mediante reacción en cadena de la polimerasa (PCR), por lo menos una parte de la secuencia del marcador molecular y por lo menos una parte del oligonucleótido de indexación de núcleos. Obtener los datos de secuenciación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras puede comprender: obtener los datos de secuenciación de la pluralidad de amplicones de indexación de núcleos identificados con códigos de barras. Obtener los datos de secuenciación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras puede comprender: secuenciar la por lo menos una parte de la secuencia del marcador molecular y la por lo menos una parte del oligonucleótido de indexación de núcleos.

En algunas realizaciones, la pluralidad de códigos de barras está asociada a una partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede inmovilizarse en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente inmovilizado en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar encerrado en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente encerrado en la partícula de identificación con códigos de barras. La partícula de identificación con códigos de barras puede ser alterable. La partícula de identificación con códigos de barras puede comprender una perla de identificación con código de barras. La partícula de identificación con códigos de barras 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 antifluorocromo, o cualquier combinación de las mismas. La partícula de identificación con códigos de barras 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, Sepharose, celulosa, nylon, silicona, y cualquier combinación de los mismos. La partícula de identificación con códigos de barras puede comprender una partícula de hidrogel alterable.

En algunas realizaciones, identificar con códigos de barras la pluralidad de dianas usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras comprende: poner en contacto copias de las dianas con regiones de unión a diana de los códigos de barras; y transcribir inversamente la pluralidad de dianas usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras. El método puede comprender: antes de obtener los datos de secuenciación de la pluralidad de dianas identificadas con códigos de barras, amplificar la pluralidad de dianas identificadas con códigos de barras para generar una pluralidad de dianas identificadas con códigos de barras amplificadas. La amplificación de las dianas identificadas con código de barras para generar la pluralidad de dianas identificadas con códigos de barras amplificadas puede comprender: amplificar, usando una reacción en cadena de la polimerasa (PCR), las dianas identificadas con códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras amplificadas. El método puede comprender: amplificar la pluralidad de dianas identificadas con códigos de barras amplificadas para generar una pluralidad de amplicones de dianas identificadas con códigos de barras. La amplificación de la pluralidad de dianas amplificadas con código de barras puede comprender: amplificar la secuencia del marcador molecular y la secuencia de una diana de la pluralidad de dianas, o una parte de la misma, para generar la pluralidad de amplicones de dianas identificadas con códigos de barras. La amplificación de la pluralidad de dianas identificadas con códigos de barras amplificadas puede comprender: amplificar, usando reacción en cadena de la polimerasa (PCR), la pluralidad de dianas identificadas con códigos de barras amplificadas para generar la pluralidad de amplicones de dianas identificadas con códigos de barras. La identificación con códigos de barras de la pluralidad de dianas de la célula usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras puede comprender: identificar estocásticamente con códigos de barras la pluralidad de dianas de la célula usando una pluralidad de códigos de barras estocásticos para generar una pluralidad de dianas identificadas estocásticamente con códigos de barras.

En algunas realizaciones, cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, un sitio de unión para un cebador universal, o cualquier combinación de los mismos, y las secuencias de marcador celular de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden una secuencia idéntica. La región de unión a la diana puede comprender una región poli(dT). Por lo menos 100 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Por lo menos 1000 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Por lo menos 10000 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Las secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias aleatorias. La región de unión a la diana puede comprender una secuencia específica de gen, una secuencia oligo(dT), un multímero aleatorio o cualquier combinación de los mismos,

la pluralidad de células comprende una muestra de tejido. La pluralidad de células puede, por ejemplo, comprender una muestra de tejido epitelial, células congeladas, células fijadas, células insertadas en parafina fijadas en formalina, células tumorales, células tumorales fijadas, células tumorales congeladas, células tumorales insertadas en parafina fijadas en formalina, o una combinación de las mismas. En algunas realizaciones, la pluralidad de células comprende uno o más componentes celulares extranucleares. Ejemplos no limitativos de componentes celulares extranucleares comprenden mitocondrias, peroxisomas, citosol, vesículas, lisosomas, membranas plasmáticas, cloroplastos, matrices mitocondriales internas, membranas mitocondriales internas, espacios intermembrana, membranas mitocondriales externas, vesículas secretoras, retículos endoplásmicos lisos, retículos endoplásmicos rugosos, cuerpos de golgi, fagosomas, endosomas, exosomas, membranas plasmáticas, microtúbulos, microfilamentos, filamentos intermedios, filopodios, remolinos, lamelipodios, sarcómeros, contactos focales, podosomas, ribosomas, microsomas, balsas lipídicas, paredes celulares, o cualquier combinación de los mismos. En algunas realizaciones, uno o más componentes celulares extranucleares comprenden una o más especies de ácidos nucleicos indeseables. En algunas realizaciones, la abundancia del por lo menos uno de los uno o más componentes celulares extranucleares se reduce aislando la pluralidad de núcleos, se reduce agotando uno o más orgánulos, o ambos. En algunas realizaciones, la pluralidad de células comprende una pluralidad de dianas y una o más especies de ácidos nucleicos indeseables. En algunas realizaciones, las especies de ácidos nucleicos indeseables se derivan de orgánulos no nucleares. En algunas realizaciones, las especies de ácidos nucleicos indeseables comprenden ARN ribosómico, ARN mitocondrial, ADN mitocondrial. En algunas realizaciones, la abundancia de la por lo menos una de las una o más especies de ácidos nucleicos indeseables se reduce aislando la pluralidad de núcleos, agotando los uno o más orgánulos, o ambos. En algunas realizaciones, la una o más especies de ácidos nucleicos indeseables asciende a aproximadamente el 50%, aproximadamente el 60%, aproximadamente el 70%, aproximadamente el 80%, o más del contenido de ácido nucleico de la pluralidad de células. En algunas realizaciones, la especie de ácido nucleico indeseable representa menos del 40%, menos del 20%, menos del 10%, menos del 5% de la pluralidad de amplicones de dianas identificadas con códigos de barras. En algunas realizaciones, la obtención de datos de secuenciación de la pluralidad de dianas identificadas con códigos de barras comprende la generación de una pluralidad de lecturas de secuenciación. Las lecturas de secuenciación para la especie de ácido nucleico indeseable pueden ser, por ejemplo, menos del 40%, menos del 20%, menos del 10%, menos del 5% del total de lecturas de secuenciación.

Lo divulgado en la presente incluye realizaciones de una composición de identificación con códigos de barras. En algunas realizaciones, la composición de identificación con códigos de barras comprende: una composición de aislamiento de núcleos que comprende un reactivo de unión a núcleos, en donde el reactivo de unión a núcleos es capaz de unirse específicamente a uno o más componentes de un núcleo; y una pluralidad de códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador molecular y una región de unión a diana, y en donde las secuencias de marcadores moleculares de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden secuencias diferentes.

En algunas realizaciones, la composición comprende: un reactivo capaz de unirse específicamente al reactivo de unión a núcleos. El reactivo de unión a núcleos puede estar asociado con un primer epítipo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender un primer reactivo de unión a epítipo. El primer epítipo puede comprender biotina, hapteno o una combinación de los mismos. El hapteno puede comprender digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender un anticuerpo antihapteno. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender avidina, estreptavidina, neutravidina o una combinación de las mismas.

En algunas realizaciones, el reactivo de unión a núcleos comprende un anticuerpo primario capaz de unirse específicamente a uno o más componentes del núcleo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario.

En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a carbohidratos. El reactivo de unión a carbohidratos puede comprender una proteína de unión a carbohidratos. La proteína de unión a carbohidratos puede comprender una lectina. La lectina puede comprender una lectina de unión a manosa, una lectina de unión a galactosa, una lectina de unión a N-acetilgalactosamina, una lectina de unión a N-acetilglucosamina, una lectina de unión a ácido N-acetilneuramínico, una lectina de unión a fucosa o una combinación de las mismas. La lectina puede comprender Concanavalina A (ConA), lectina de lenteja (LCH), lectina de campanilla de invierno (GNA), aglutinina de *Ricinus communis* (RCA), aglutinina de cacahuete (PNA), jacalina (AEL), lectina de veza peluda (VVL), aglutinina de germen de trigo (WGA), lectina del saúco (SNA), leucoaglutinina de *Maackia amurensis* (MAL), hemoaglutinina de *Maackia amurensis* (MAH), aglutinina de *Ulex europaeus* (UEA), lectina de *Aleuria aurantia* (AAL), o una combinación de las mismas. La lectina puede ser, o comprender, una aglutinina. La aglutinina puede ser, o comprender, aglutinina de germen de trigo (WGA). La proteína de unión a carbohidratos puede proceder o derivarse de un animal, una bacteria, un virus o un hongo. La proteína de unión a carbohidratos puede proceder o derivarse de una planta. La planta puede ser *Canavalia ensiformis*, *Lens culinaris*, *Galanthus nivalis*, *Ricinus communis*, *Arachis hypogaea*, *Artocarpus integrifolia*, *Vicia villosa*, *Triticum vulgare*, *Sambucus nigra*, *Maackia amurensis*, *Ulex europaeus*, *Aleuria aurantia*, o una combinación de las mismas.

En algunas realizaciones, los uno o más componentes del núcleo comprenden un azúcar, un oligosacárido, un polisacárido, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender un monosacárido, un disacárido, un poliol, un malto-oligosacárido, un no-malto-oligosacárido, un almidón, un polisacárido no de almidón, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender glucosa, galactosa, fructosa, xilosa, sacarosa, lactosa, maltosa, trehalosa, sorbitol, manitol, maltodextrina, rafinosa, estaquiosa, fructooligosacárido, amilosa, amilopectina, almidón modificado, glucógeno, celulosa, hemicelulosa, pectina, hidrocoloide, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender un residuo de  $\alpha$ -D-manosilo, un residuo de  $\alpha$ -D-glucosilo, una estructura  $\alpha$ -manosídica ramificada de tipo  $\alpha$ -manosa alta, una estructura  $\alpha$ -manosídica ramificada de tipo híbrido y N-Glicano de tipo complejo biantenar, una región central fucosilada de N-Glicano de tipo complejo bi- y triantenar, una estructura de manosa alta enlazada  $\alpha$  1-3 y  $\alpha$  1-6, Gal $\beta$ 1-4GalNAc $\beta$ 1-R, Gal $\beta$ 1-3GalNAc $\alpha$ 1-Ser/Thr, (Sia)Gal $\beta$ 1-3GalNAc $\alpha$ 1-Ser/Thr, GalNAc $\alpha$ -Ser/Thr, GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc, Neu5Ac (ácido siálico), Neu5Ac $\alpha$ 2-6Gal(NAc)-R, Neu5Ac/Gc $\alpha$ 2,3Gal $\beta$ 1,3(Neu5Ac $\alpha$ 2,6)GalNAc, Fuca $\alpha$ 1-2Gal-R, Fuca $\alpha$ 1-2Gal $\beta$ 1-4(Fuca $\alpha$ 1-3/4)Gal $\beta$ 1-4GlcNAc, R2-GlcNAc $\beta$ 1-4(Fuca $\alpha$ 1-6)GlcNAc-R1, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender una glicoproteína, un glicolípido o una combinación de los mismos.

En algunas realizaciones, el reactivo de unión a núcleos está asociado a una partícula de aislamiento del núcleo. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede asociarse a la partícula de aislamiento del núcleo. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede estar inmovilizado, o parcialmente inmovilizado, en la partícula de aislamiento del núcleo. La partícula de aislamiento del núcleo puede comprender una perla de aislamiento del núcleo. La partícula de aislamiento del núcleo 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 micropelícula antibiotina, una micropelícula antiluminocromo, o cualquier combinación de las mismas. La partícula de aislamiento del núcleo puede comprender un material seleccionado del grupo que consiste en polidimetilsiloxano (PDMS), poliestireno, vidrio, polipropileno, agarosa, gelatina, hidrogel, paramagnético, cerámica, plástico, vidrio, metilmetacrilato, polímero acrílico, titanio, látex, sefarosa, celulosa, nailon, silicona y cualquier combinación de los mismos. La partícula de aislamiento del núcleo puede ser alterable. La partícula de aislamiento del núcleo puede comprender una partícula de hidrogel alterable de aislamiento del núcleo.

En algunas realizaciones, la pluralidad de códigos de barras está asociada a la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar inmovilizado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente inmovilizado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar encerrado en la partícula de aislamiento del núcleo. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente encerrado en la partícula de aislamiento del núcleo.

En algunas realizaciones, los uno o más componentes del núcleo pueden comprender Lamin, Emerin, Nesprin, Nurim, UNC-83, Klar, ZYG-12, Kms1p, UNC-84, Klaroid, SUN-1, Sad1p, LBR, MAN1, LAP1, LAP2, LINK, un complejo de poro nuclear, una parte del mismo, o una combinación de los mismos. En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a componentes de la superficie de la envoltura nuclear, y el reactivo de unión a núcleos puede ser capaz de unirse específicamente a uno o más componentes de la superficie de la envoltura nuclear.

En algunas realizaciones, la composición comprende una composición de captura de orgánulos que comprende un reactivo de unión a orgánulos, en donde el reactivo de unión a orgánulos es capaz de unirse específicamente a uno o más componentes de uno o más orgánulos. La composición puede comprender un reactivo capaz de unirse específicamente al reactivo de unión a orgánulos. El reactivo de unión a orgánulos puede estar asociado a un segundo epítipo, y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un segundo reactivo de unión a epítipo. El segundo epítipo puede comprender biotina, hapteno o una combinación de los mismos. El hapteno puede comprender digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos. El reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un anticuerpo antihapteno. El reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender avidina, estreptavidina, neutravidina o una combinación de los mismos. El reactivo de unión a orgánulos puede comprender un anticuerpo primario capaz de unirse específicamente al uno o más componentes de uno o más orgánulos de la pluralidad de células, y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario.

En algunas realizaciones, el reactivo de unión a orgánulos está asociado con una partícula de captura de orgánulos. El reactivo de unión a orgánulos puede estar inmovilizado, o parcialmente inmovilizado, en la partícula de captura de orgánulos. La partícula de captura de orgánulos puede comprender una perla de captura de orgánulos. La partícula de captura de orgánulos 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 antifluorocromo, o cualquier combinación de las mismas. La partícula de captura de orgánulos 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, Sepharose, celulosa, nylon, silicona, y cualquier combinación de los mismos.

En algunas realizaciones, los orgánulos comprenden mitocondrias de la pluralidad de células. Los uno o más componentes de los uno o más orgánulos de la pluralidad de células puede comprender ABCD3, ESR2, NOS3, ALB, HIF1A, NR3C1, ATP5A1, HK1, PGR, CASQ1, HSPA1A, PHB, CLTC, HSPD1, PLN, COX4I1, IFM1, SOD1, CPS1, LGALS3, TP53, Citocromo C Oxidasa, MAPT, TP5B, ERN1, MT-CO1, VDAC1, o una combinación de los mismos.

En algunas realizaciones, el reactivo de unión a orgánulos comprende un reactivo de unión a componentes de superficie de orgánulos, los uno o más componentes de los uno o más orgánulos pueden comprender uno o más componentes de superficie de orgánulos, y el reactivo de unión a orgánulos puede ser capaz de unirse específicamente a los uno o más componentes de superficie de orgánulos.

En algunas realizaciones, el reactivo de unión a núcleos puede comprender un oligonucleótido de indexación de núcleos, y en donde el oligonucleótido de indexación de núcleos comprende una secuencia de indexación de núcleos. La pluralidad de códigos de barras puede estar asociada a una partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede inmovilizarse en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente inmovilizado en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar encerrado en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente encerrado en la partícula de identificación con códigos de barras. La partícula de identificación con códigos de barras puede ser alterable. La partícula de identificación con códigos de barras puede comprender una perla de identificación con código de barras. La partícula de identificación con códigos de barras 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 antifluorocromo o cualquier combinación de las mismas. La partícula de identificación con códigos de barras 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, Sepharose, celulosa, nylon, silicona, y cualquier combinación de los mismos. La partícula de identificación con códigos de barras puede comprender una partícula de hidrogel alterable.

En algunas realizaciones, cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, un sitio de unión para un cebador universal, o cualquier combinación de los mismos, y las secuencias de marcador celular de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden una secuencia idéntica.

La región de unión a la diana puede comprender una región poli(dT). Por lo menos 100 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Por lo menos 1000 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Por lo menos 10000 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Las secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias aleatorias. La región de unión a la diana puede comprender una secuencia específica de gen, una secuencia oligo(dT), un multímero aleatorio o cualquier combinación de los mismos.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1 ilustra un código de barras ejemplar no limitativo (por ejemplo, un código de barras estocástico).

La FIG. 2 muestra un flujo de trabajo ejemplar no limitativo de identificación con códigos de barras y recuento digital (por ejemplo, identificación con códigos de barras estocásticos y recuento digital).

La FIG. 3 es una ilustración esquemática que muestra un proceso ejemplar no limitativo para generar una biblioteca indexada de objetivos identificados con códigos de barras (por ejemplo, objetivos identificados con códigos de barras estocásticos) a partir de una pluralidad de objetivos.

Las FIGS. 4A-4B muestran una ilustración esquemática ejemplar no limitativa de un flujo de trabajo de identificación con códigos de barras de núcleos (por ejemplo, identificación con códigos de barras estocásticos).

Las FIGS. 5A-5B muestran una ilustración esquemática ejemplar no limitativa de un flujo de trabajo de identificación con códigos de barras (por ejemplo, identificación con códigos de barras estocásticos) con eliminación de orgánulos (por ejemplo, mitocondrias).

Las FIGS. 6A-6F muestran una ilustración esquemática ejemplar no limitativa de un flujo de trabajo de captura de

núcleos e identificación con códigos de barras.

Las FIGS. 7A-7C muestran una ilustración esquemática ejemplar no limitativa de un flujo de trabajo de captura de núcleos e identificación con códigos de barras.

## 5 DESCRIPCIÓN DETALLADA

En la siguiente descripción detallada, se hace referencia a los dibujos acompañantes, que forman parte de la misma. En los dibujos, símbolos similares identifican típicamente 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. Pueden utilizarse otras realizaciones, y pueden realizarse otros cambios, sin apartarse del alcance de la materia presentada en la presente. Se entenderá fácilmente que los aspectos de la presente divulgación, tal 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 cuales se contemplan explícitamente en la presente y forman parte de la divulgación de la presente.

Determinar el número de ácidos nucleicos o dianas, por ejemplo moléculas de ácido ribonucleico mensajero (ARNm), es clínicamente importante para identificar, por ejemplo, los genes que se expresan en una célula en diferentes etapas de desarrollo o bajo diferentes condiciones ambientales. Sin embargo, puede ser muy difícil determinar el número absoluto de moléculas de ácido nucleico (por ejemplo, moléculas de ARNm), especialmente cuando el número de moléculas es muy pequeño. Un método para determinar el número absoluto de moléculas en una muestra es la reacción en cadena de la polimerasa (PCR) digital. Idealmente, la PCR produce una copia idéntica de una molécula en cada ciclo. Sin embargo, la PCR puede tener desventajas tales como que cada molécula se replica con una probabilidad estocástica, y esta probabilidad varía según el ciclo de PCR y la secuencia del gen, lo que da como resultado un sesgo de amplificación y mediciones inexactas de la expresión génica.

Los códigos de barras (por ejemplo, códigos de barras estocásticos) con marcadores moleculares únicos (ML, también denominadas índices moleculares (MI)) pueden usarse para contar el número de moléculas. Los códigos de barras con marcadores moleculares que son únicos para cada marcador celular pueden usarse para contar el número de moléculas en cada célula. Algunos ensayos ejemplares no limitativos para identificación con códigos de barras incluyen el ensayo Precise™ (Cellular Research, Inc. (Palo Alto, CA)), el ensayo Resolve™ (Cellular Research, Inc. (Palo Alto, CA)), o el ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ)). Sin embargo, estos métodos y técnicas pueden introducir errores que, si no se corrigen, pueden dar lugar a recuentos celulares sobreestimados.

El ensayo Rhapsody™ puede utilizar un grupo no agotable de códigos de barras (por ejemplo, códigos de barras estocásticos) con un gran número, por ejemplo de 6561 a 65536, marcadores moleculares únicos en oligonucleótidos poli(T) para hibridar con todos los ARNm poli(A) de una muestra durante el paso de transcripción inversa (RT). Además de los marcadores moleculares, los marcadores celulares de los códigos de barras pueden usarse para identificar cada célula individual en cada pocillo de una placa de micropocillos. Un código de barras puede incluir un sitio de cebado de PCR universal. Durante la RT, las moléculas del gen diana reaccionan aleatoriamente con los códigos de barras. Cada molécula diana puede hibridar con un código de barras (por ejemplo, un código de barras estocástico) dando lugar a moléculas de ácido ribonucleótido complementario (ADNc) identificadas con códigos de barras (por ejemplo, moléculas de ADNc identificadas con códigos de barras estocásticos). Después del marcado, las moléculas de ADNc identificadas con códigos de barras de los micropocillos de una placa de micropocillos pueden agruparse en un único tubo para la amplificación por PCR y la secuenciación. Los datos brutos de secuenciación pueden analizarse para obtener el número de códigos de barras con marcadores moleculares únicos.

Lo divulgado en la presente incluye realizaciones de un método para determinar el número de dianas en una pluralidad de células. En algunas realizaciones, el método comprende: aislar una pluralidad de núcleos de una pluralidad de células usando una composición de aislamiento de núcleos, en donde la composición de aislamiento de núcleos comprende un reactivo de unión a núcleos, y en donde el reactivo de unión a núcleos es capaz de unirse específicamente a uno o más componentes de un núcleo; identificar con códigos de barras una pluralidad de dianas en la pluralidad de núcleos usando una pluralidad de códigos de barras para generar una pluralidad de dianas identificadas con códigos de barras, en donde cada uno de los códigos de barras comprende una secuencia de marcador molecular y una región de unión a la diana, y en donde las secuencias de marcadores moleculares de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden secuencias diferentes; obtener datos de secuenciación de la pluralidad de dianas identificadas con códigos de barras; y estimar el número de cada una de las dianas en la pluralidad de células usando las secuencias de marcadores moleculares de la pluralidad de códigos de barras en los datos de secuenciación.

Lo divulgado en la presente incluye realizaciones de una composición de identificación con códigos de barras. En algunas realizaciones, la composición de identificación con códigos de barras comprende: una composición de aislamiento de núcleos que comprende un reactivo de unión a núcleos, en donde el reactivo de unión a núcleos es capaz de unirse específicamente a uno o más componentes de un núcleo; y una pluralidad de códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador molecular y una región

de unión a diana, y en donde las secuencias de marcadores moleculares de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden secuencias diferentes.

### Definiciones

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

Como se usa en la presente, el término "adaptador" puede significar una secuencia para facilitar la amplificación o secuenciación de ácidos nucleicos asociados. Los ácidos nucleicos asociados pueden comprender ácidos nucleicos diana. Los ácidos nucleicos asociados pueden comprender uno o más de marcadores espaciales, marcadores diana, marcadores de muestra, marcadores de indexación, códigos de barras, códigos de barras estocásticos o 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 estar situados 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 iguales o diferentes. Un adaptador situado en los extremos 5' y/o 3' de un polinucleótido puede ser capaz de hibridar con uno o más oligonucleótidos inmovilizados en una superficie. En algunas realizaciones, un adaptador puede comprender una secuencia universal. Una secuencia universal puede ser una región de secuencia de nucleótidos que es común a dos o más moléculas de ácido nucleico. Las dos o más moléculas de ácido nucleico también pueden tener regiones de secuencia diferente. 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 ácido nucleico puede permitir la replicación o amplificación de múltiples secuencias diferentes usando un único cebador universal que sea complementario a la secuencia universal. De manera similar, por lo menos una, dos (por ejemplo, un par) o más secuencias universales que pueden estar presentes en diferentes miembros de una colección de moléculas de ácido nucleico pueden permitir la replicación o amplificación de múltiples secuencias diferentes usando por lo menos uno, dos (por ejemplo, un par) o más cebadores universales individuales 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 secuencias de ácido nucleico diana pueden modificarse para unir adaptadores universales (por ejemplo, secuencias de ácido nucleico no diana) a uno o ambos extremos de las diferentes secuencias de ácido nucleico diana. Los uno o más cebadores universales unidos al ácido nucleico diana pueden proporcionar sitios para la hibridación de cebadores universales. Los uno o más cebadores universales unidos al ácido nucleico diana pueden ser iguales o diferentes entre sí.

Como se usa en la presente, el término "asociado" o "asociado con" puede significar que dos o más especies son identificables como estando colocalizadas en un punto en el tiempo. Una asociación puede significar que dos o más especies están o estaban dentro de un recipiente similar. Una asociación puede ser una asociación informática, donde por ejemplo se almacena información digital referente a dos o más especies y puede usarse para determinar que una o más de las especies estaban conlocalizadas en un punto temporal. Una asociación puede ser una asociación física, directa o indirecta. 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 ser una unión covalente entre una diana y un marcador.

Como se usa en la presente, el término "complementario" 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 un enlace de hidrógeno con un nucleótido de otro ácido nucleico, entonces se considera que los dos ácidos nucleicos son complementarios entre sí en esa posición. La complementariedad entre dos moléculas de ácido nucleico de cadena sencilla puede ser "parcial", en la que sólo se unen algunos de los nucleótidos, o puede ser completa cuando existe 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 de 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 de una secuencia que es inversa (es decir, el orden de los nucleótidos está invertido) de la segunda secuencia. Como se usa en la presente, los términos "complemento", "complementario" o "complemento inverso" pueden usarse indistintamente. Se entiende de la divulgación que si una molécula puede hibridar con otra molécula puede ser el complemento de 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 diana en una muestra. El recuento digital puede incluir el paso de determinar un número de marcadores únicos que se han asociado con dianas en una muestra. Esta metodología estocástica transforma el problema del

recuento de moléculas de uno de localización e identificación de moléculas idénticas a una serie de preguntas digitales de sí/no referentes a la detección de un conjunto de marcadores predefinidos.

Como se usa en la presente, el término "marcador" o "marcadores" puede referirse a códigos de ácidos nucleicos asociados con una diana dentro de una muestra. Un marcador puede ser, por ejemplo, un marcador de ácido nucleico. Un marcador puede ser total o parcialmente amplificable. Un marcador puede ser total o parcialmente secuenciable. Un marcador puede ser una parte de un ácido nucleico nativo que es identificable como distinto. Un marcador puede ser una secuencia conocida. Un marcador puede comprender una unión de secuencias de ácido nucleico, 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 transmitir información. Por ejemplo, en varias realizaciones, los marcadores pueden usarse para determinar una identidad de una muestra, una fuente de una muestra, una identidad de una célula y/o una diana.

Como se usa en la presente, el término "depósitos no agotables" puede referirse a un grupo de códigos de barras estocásticos compuesto por muchos marcadores diferentes. Un depósito no agotable puede comprender un gran número de códigos de barras diferentes, de tal manera que cuando el depósito no agotable se asocia con un grupo de dianas, es probable que cada diana se asocie con un código de barras estocástico único. La unicidad de cada molécula diana marcada puede determinarse mediante la estadística de elección aleatoria y depende del número de copias de moléculas diana idénticas en la colección en comparación con la diversidad de marcadores. El tamaño del conjunto resultante de moléculas diana marcadas puede determinarse por la naturaleza estocástica del proceso de identificación con códigos de barras, y el análisis del número de códigos de barras estocásticos detectados permite entonces calcular el número de moléculas diana presentes en la colección o muestra original. Cuando la proporción entre el número de copias de una molécula diana presente y el número de códigos de barras estocásticos únicos es baja, las moléculas diana marcadas son altamente únicas (es decir, hay una probabilidad muy baja de que más de una molécula diana haya sido marcada con un marcador dado).

Como se usa en la presente, el término "ácido nucleico" se refiere a una secuencia de polinucleótidos, o fragmento de la misma. Un ácido nucleico puede comprender 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 ser un gen o un fragmento del mismo. Un ácido nucleico puede ser ADN. Un ácido nucleico puede ser ARN. Un ácido nucleico puede comprender uno o más análogos (por ejemplo, una estructura principal, azúcar o nucleobase alteradas). Algunos ejemplos no limitativos de análogos incluyen: 5-bromouracilo, ácido nucleico peptídico, ácido xeno nucleico, morfolinos, ácidos nucleicos bloqueados, ácidos nucleicos glicólicos, ácidos nucleicos de treosa, dideoxinucleótidos, cordicepina, 7-deaza-GTP, fluoróforos (por ejemplo, rodamina o fluoresceína enlazada al azúcar), nucleótidos que contienen tioles, nucleótidos enlazados a biotina, análogos de bases fluorescentes, islas de CpG, metil-7-guanosina, nucleótidos metilados, inosina, tiouridina, pseudouridina, dihidrouridina, queuosina y wiosina. Los términos "ácido nucleico", "polinucleótido", "polinucleótido diana" y "ácido nucleico diana" pueden usarse indistintamente.

Un ácido nucleico puede comprender una o más modificaciones (por ejemplo, una modificación de bases, una modificación de la estructura principal), para proporcionar al ácido nucleico una característica nueva o mejorada (por ejemplo, estabilidad mejorada). Un ácido nucleico puede comprender una etiqueta de afinidad de ácido nucleico. Un nucleósido puede ser una combinación de base y 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 incluyen además un grupo fosfato enlazado covalentemente a la porción de azúcar del nucleósido. En el caso de los nucleósidos que incluyen un azúcar pentofuranosílico, el grupo fosfato puede estar enlazado a la fracción hidroxilo 2', 3' o 5' del azúcar. Al formar 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 aquellas que no tienen un átomo de fósforo en la estructura principal. Las estructuras principales de ácidos nucleicos modificados adecuados que contienen un átomo de fósforo en las mismas pueden incluir, por ejemplo, fosforotioatos, fosforotioatos quirales, fosforoditioatos, fosfotriesteres, fosfotriesteres de aminoalquilo, fosfonatos de metilo y otros alquilo, como fosfonatos de 3'-alquileo, fosfonatos de 5'-alquileo, fosfonatos quirales, fosfinatos, fosforamidatos, incluyendo 3'-aminofosforamidato y fosforamidatos de aminoalquilo, fosforodiamidatos, tionofosforamidatos, tionalquilfosfonatos, tionalquilfosfotriesteres, selenofosfatos y boranofosfatos que tienen enlaces 3'-5' normales, análogos 2'-5' enlazados y aquellos que tienen polaridad invertida en los que uno o más enlaces internucleotídicos es un enlace de 3' a 3', de 5' a 5' o de 2' a 2'.

Un ácido nucleico puede comprender estructuras principales de polinucleótidos formadas por enlaces internucleosídicos de alquilo o cicloalquilo de cadena corta, enlaces internucleosídicos mixtos de heteroátomo y alquilo o cicloalquilo, o uno o más enlaces internucleosídicos heteroatómicos o heterocíclicos de cadena corta. Estos pueden incluir los que tienen enlaces morfolinos (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 alqueno; estructuras principales de sulfamato; estructuras principales de metilenimino y metilenedihidrazino; estructuras principales de sulfonato y sulfonamida; estructuras principales de amida; y otras que tienen partes componentes de N, O, S y CH<sub>2</sub> mixtas.

Un ácido nucleico puede comprender un mimético de ácido nucleico. Puede pretenderse que el término "mimético" incluya polinucleótidos en los que sólo el anillo de furanosa o tanto el anillo de furanosa como el enlace internucleotídico se sustituyen por grupos no furanosa; la sustitución de sólo el anillo de furanosa también puede denominarse sustituto de azúcar. La fracción de base heterocíclica o una fracción de base heterocíclica modificada puede mantenerse para la hibridación con un ácido nucleico diana 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 sustituirse por una estructura principal que contenga amida, en particular una estructura principal de aminoetilglicina. Los nucleótidos pueden conservarse y unirse directa o indirectamente a los átomos de nitrógeno aza de la porción amida de la estructura principal. La estructura principal de los compuestos de PNA puede comprender dos o más unidades de aminoetilglicina enlazadas, lo que confiere al PNA una estructura principal que contiene amida. Las fracciones de base heterocíclica pueden unirse directa o indirectamente a los átomos de nitrógeno aza de la porción amida de la estructura principal.

Un ácido nucleico puede comprender 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 fosfordiamidato u otro enlace internucleosídico no fosfodiéster puede sustituir a un enlace fosfodiéster.

Un ácido nucleico puede comprender 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 no iónicos basados en morfolino pueden tener menos interacciones no deseadas con las proteínas celulares. Los polinucleótidos basados en morfolino pueden ser imitadores no iónicos de ácidos nucleicos. Una variedad de compuestos dentro de la clase morfolino pueden unirse usando diferentes grupos de enlace. Una clase adicional de miméticos de polinucleótidos puede denominarse ácidos nucleicos de ciclohexenilo (CeNA). El anillo de furanosa normalmente presente en una molécula de ácido nucleico puede sustituirse por un anillo de ciclohexenilo. Para la síntesis de compuestos oligoméricos mediante química de fosforamidita pueden prepararse y usarse los monómeros de fosforamidita protegidos con CeNA DMT. La incorporación de monómeros de CeNA en una cadena de ácido nucleico puede aumentar la estabilidad de un híbrido 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 los 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 bicíclico. El enlace puede ser un grupo metileno (-CH<sub>2</sub>), grupo que forma un puente entre el átomo de oxígeno 2' y el átomo de carbono 4' en donde n es 1 o 2. El LNA y los análogos del LNA pueden mostrar estabilidades térmicas dúplex muy altas con el ácido nucleico complementario (T<sub>m</sub>=+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 la 5-metilcitosina (5-me-C), la 5-hidroximetilcitosina, la xantina, la hipoxantina, la 2-aminoadenina, la 6-metil y otros derivados alquílicos de la adenina y la guanina, 2-propil y otros derivados alquílicos de la adenina y la guanina, 2-tiouracilo, 2-tiotimina y 2-tiocitosina, 5-halouracilo y citosina, 5-propinilo (-C≡C-CH<sub>3</sub>) uracilo y citosina y otros derivados alquílicos de las bases de pirimidina, 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 la fenoxazina citidina (1H-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 fenoxazina citidina 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 fenoxazina citidina sustituida (por ejemplo, 9-(2-aminoetoxi)-H-pirimido(5,4-b)(1,4)benzoxazin-2(3H)-ona), citidina carbazol (2H-pirimido(4,5-b)indol-2-ona), citidina piridindol (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 dianas. Las muestras adecuadas para el análisis mediante los métodos, dispositivos y sistemas divulgados incluyen células, tejidos, órganos u organismos.

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 en un sustrato. Un dispositivo de muestreo puede referirse, por ejemplo, a una máquina de clasificación celular activada por fluorescencia (FACS), una máquina clasificadora de células, una aguja de biopsia, un dispositivo de biopsia, un dispositivo de seccionamiento de tejidos, un dispositivo microfluídico, una rejilla de cuchillas y/o un micrótopo.

Como se usa en la presente, el término "soporte sólido" puede referirse a superficies sólidas o semisólidas discretas a las que pueden unirse una pluralidad de códigos de barras estocásticos. Un soporte sólido puede abarcar cualquier tipo de esfera sólida, porosa o hueca, bola, cojinete, cilindro u otra configuración similar compuesta de plástico, cerámica, metal o material polimérico (por ejemplo, hidrogel) sobre la que puede inmovilizarse un ácido nucleico (por ejemplo, de manera covalente o no covalente). 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 pluralidad de soportes sólidos espaciados en una matriz pueden no comprender un sustrato. Un soporte sólido puede ser una partícula, como una partícula sintética o una perla.

Un soporte sólido puede referirse a un "sustrato". Un sustrato puede ser un tipo de soporte sólido. Un sustrato puede referirse a una superficie sólida o semisólida continua sobre la que pueden realizarse los métodos de la divulgación. Un sustrato puede referirse a una matriz, un cartucho, un chip, un dispositivo y un portaobjetos, por ejemplo.

Como se usa en la presente, el término "marcador espacial" puede referirse a un marcador que puede asociarse con una posición en el espacio.

Como se usa en la presente, el término "código de barras estocástico" puede referirse a una secuencia de polinucleótidos que comprende marcadores. Un código de barras estocástico puede ser una secuencia de polinucleótidos que puede usarse para identificar 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 que se haya asociado un marcador con una diana. Por ejemplo, un código de barras estocástico puede usarse para evaluar errores de amplificación o secuenciación. Un código de barras estocástico asociado con una diana puede denominarse código de barras estocástico-diana o código de barras estocástico-etiqueta-diana.

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 diana que es específica de gen. Un código de barras estocástico puede ser una secuencia de polinucleótidos que puede usarse para identificar con códigos de barras estocásticos. Los códigos de barras estocásticos pueden usarse para cuantificar dianas dentro de una muestra. Los códigos de barras estocásticos pueden usarse para controlar los errores que pueden producirse después de que un marcador se haya asociado a una diana. Por ejemplo, un código de barras estocástico puede usarse para evaluar errores de amplificación o secuenciación. Un código de barras estocástico asociado a una diana puede denominarse código de barras estocástico-diana o código de barras estocástico-etiqueta-diana.

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

Como se usa en la presente, el término "diana" puede referirse a una composición que puede asociarse con un código de barras estocástico. Las dianas ejemplares adecuadas para el análisis mediante los métodos, dispositivos y sistemas divulgados incluyen oligonucleótidos, ADN, ARN, ARNm, microARN, ARNt y similares. Las dianas pueden ser de cadena simple o doble. En algunas realizaciones, las dianas pueden ser proteínas. En algunas realizaciones, las dianas son lípidos.

Como se usa en la presente, el término "transcriptasas inversas" 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, entre otras, la transcriptasa inversa retroviral, la transcriptasa inversa retrotransposónica, las transcriptasas inversas retroplásmidas, las transcriptasas inversas retrónicas, las transcriptasas inversas bacterianas, las transcriptasas inversas derivadas 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 de intrones LI.LtrB de *Lactococcus lactis*, la transcriptasa inversa de intrones Tel4c de *Thermosynechococcus elongatus* o la transcriptasa inversa de intrones 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).

En la presente se divulgan sistemas y métodos para identificar un marcador de célula señal. En algunas realizaciones, el método comprende: (a) identificar estocásticamente con códigos de barras una pluralidad de dianas en una muestra de células usando una pluralidad de códigos de barras estocásticos para crear una pluralidad de objetivos identificados estocásticamente con códigos de barras, en donde cada uno de la pluralidad de códigos de barras estocásticos comprende un marcador celular y un marcador molecular; (b) obtener datos de secuenciación de la pluralidad de objetivos identificados estocásticamente con códigos de barras; (c) determinar el número de marcadores moleculares con secuencias distintas asociadas a cada uno de los marcadores celulares de la pluralidad de códigos de barras estocásticos; (d) determinar un rango de cada uno de los marcadores celulares de la pluralidad de códigos de barras estocásticos sobre la base del número de marcadores moleculares con secuencias distintas asociadas a cada uno de los marcadores celulares; (e) generar un gráfico de suma acumulativa sobre la base del número de marcadores moleculares con secuencias distintas asociadas con cada uno de los marcadores celulares determinados en (c) y el rango de cada uno de los marcadores celulares determinados en (d); (f) generar un segundo gráfico derivado del gráfico de suma acumulativa; (g) determinar un mínimo del segundo gráfico derivado del diagrama de suma acumulativa, en donde el mínimo del segundo gráfico derivado corresponde a un umbral de marcador celular; y (h) identificar cada uno de los marcadores celulares como un marcador celular de señal o un marcador celular de ruido basándose en el número de marcadores moleculares con secuencias distintas asociadas con cada uno de los marcadores celulares determinados en (c) y el umbral de marcador celular determinado en (g).

#### Códigos de barras

La identificación con códigos de barras, como la codificación con códigos de barras estocásticos, se ha descrito, por ejemplo, en la US20150299784, la WO2015031691, y Fu et al, Proc Natl Acad Sci U.S.A. 31 de mayo 2011;108(22):9026-31 y Fan et al., Science (2015) 347(6222):1258367. 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, identificar con códigos de barras, etiquetar) una diana. Los códigos de barras pueden denominarse códigos de barras estocásticos si la proporción entre el número de secuencias de códigos de barras diferentes de los códigos de barras estocásticos y el número de apariciones de cualquiera de las dianas que se van a marcar puede ser, o ser de 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 intervalo entre dos cualesquiera de estos valores. Una diana puede ser, por ejemplo, una especie de ARNm que comprende moléculas de ARNm con secuencias idénticas o casi idénticas. Los códigos de barras pueden denominarse códigos de barras estocásticos si la proporción entre el número de secuencias de códigos de barras diferentes de los códigos de barras estocásticos y el número de apariciones de cualquiera de las dianas que deben marcarse es por lo menos, o como máximo, de 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 estocásticos pueden denominarse 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 104 ejemplar con un marcador espacial. El código de barras 104 puede comprender una amina 5' que puede unir el código de barras a un soporte sólido 108. 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 orden de los diferentes marcadores (incluyendo pero no limitados al marcador universal, marcador de dimensión, marcador espacial, marcador celular y 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 a la diana. La región de unión a la diana puede interactuar con una diana (por ejemplo, ácido nucleico, ARN, ARNm, ADN diana) en una muestra. Por ejemplo, una región de unión a la diana puede comprender una secuencia oligo(dT) que puede interactuar con colas poli(A) de los ARNm. En algunos casos, los marcadores del código de barras (por ejemplo, el marcador universal, el marcador de dimensión, el marcador espacial, el marcador celular y la secuencia del código de barras) pueden estar separadas por 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleótidos.

Un marcador, por ejemplo el marcador celular, puede comprender un conjunto único de subsecuencias de ácido nucleico de longitud definida, por ejemplo de siete nucleótidos cada una (equivalente al número de bits usados en algunos códigos de corrección de errores de Hamming), que pueden diseñarse para proporcionar capacidad de



corrección de errores. El conjunto de subsecuencias de corrección de errores comprende siete secuencias de nucleótidos y puede diseñarse de tal manera que cualquier combinación de secuencias por pares del conjunto muestre una "distancia genética" definida (o número de bases malapareadas); por ejemplo, un conjunto de subsecuencias de corrección de errores puede diseñarse de manera que muestre una distancia genética de tres nucleótidos. En este caso, la revisión de las secuencias de corrección de errores en el conjunto de datos de secuencia para moléculas de ácido nucleico diana marcadas (descritas más detalladamente 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 cualesquiera de estos valores, nucleótidos de longitud. En algunas realizaciones, para crear códigos de corrección de errores pueden usarse subsecuencias de ácido nucleico de otras longitudes.

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

En algunas realizaciones, una región de unión a diana puede comprender una secuencia oligo(dT) que puede interactuar con colas poli(A) de 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 la secuencia del código de barras (por ejemplo, un marcador molecular)) pueden estar separados por uno o más espaciadores de otro u otros 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, 20, o más nucleótidos de longitud. 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, los 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, los uno o más marcadores universales pueden ser los mismos para todos los códigos de barras unidos a una pluralidad de partículas (por ejemplo, partículas sintéticas como perlas). En algunas realizaciones, un marcador universal puede comprender una secuencia de ácido nucleico capaz de hibridarse 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 ácido nucleico capaz de hibridar con un cebador de PCR. En algunas realizaciones, el marcador universal puede comprender una secuencia de ácido nucleico capaz de hibridar con un cebador de secuenciación y un cebador de PCR. La secuencia de ácido nucleico del marcador universal capaz de hibridar con un cebador de secuenciación o de PCR puede denominarse sitio de unión a 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 de una región dentro del código de barras. Un marcador universal 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. Por ejemplo, un marcador universal puede comprender por lo menos aproximadamente 10 nucleótidos. Un marcador universal 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. En algunas realizaciones, un conector escindible o nucleótido modificado puede formar parte de la secuencia del marcador universal 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 ácido nucleico 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 el que se identificó estocásticamente con código de barras un objetivo. Un marcador de dimensión puede asociarse con un momento de identificación con código de barras (por ejemplo, identificación con código de barras estocástico) en una muestra. Un marcador de dimensión puede activarse en el momento del marcado. En diferentes momentos pueden activarse diferentes marcadores de dimensión. El marcador de dimensión proporciona información sobre el orden en el que se identificaron con códigos de barras estocásticamente los objetivos, grupos de objetivos y/o muestras. Por ejemplo, puede identificarse con códigos de barras estocásticamente una población de células en la fase G0 del ciclo celular. Las células pueden ser pulsadas de nuevo con códigos de barras (por ejemplo, códigos de barras estocásticos) en la fase G1 del ciclo celular. Las células pueden ser pulsadas 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 y en qué fase del ciclo celular. Los marcadores de dimensión pueden interrogar muchos momentos biológicos diferentes. Los momentos biológicos ejemplares pueden incluir, entre otros, el ciclo celular, la transcripción (por ejemplo, el inicio de la transcripción) y la degradación de la transcripción. En otro ejemplo, una muestra (por ejemplo, una célula, una población de células) puede marcarse estocásticamente antes y/o después del tratamiento con un fármaco y/o terapia. Los cambios en el número de copias de distintas dianas pueden ser indicativos de la respuesta de la muestra al fármaco y/o terapia.

Un marcador de dimensión puede ser activable. Un marcador de dimensión activable puede activarse en un punto temporal específico. El marcador activable puede, por ejemplo, activarse constitutivamente (por ejemplo, no apagarse). El marcador de dimensión activable puede, por ejemplo, activarse reversiblemente (por ejemplo, el marcador de dimensión activable puede encenderse y apagarse). 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 ser reversiblemente activable, 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, ligadura de otra molécula, adición de modificaciones (por ejemplo, pegilarse, sumoilarse, acetilarse, metilarse, desacetilarse, desmetilarse), un evento fotoquímico (por ejemplo, fotocaptura), 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 partícula sintética, como una perla), pero diferente para diferentes soportes sólidos (por ejemplo, partículas sintéticas). 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 incluir el mismo marcador de dimensión. En algunas realizaciones, por lo menos el 95% de los códigos de barras en el mismo soporte sólido pueden incluir el mismo marcador de dimensión.

En una pluralidad de soportes sólidos (por ejemplo, partículas sintéticas) puede haber representadas tantas como  $10^6$  o más secuencias de marcadores de dimensión únicas. 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 cualesquiera 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 tener entre 5 y 200 nucleótidos. Un marcador de dimensión puede tener entre 10 y 150 nucleótidos. Un marcador de dimensión puede tener entre 20 y 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 ácido nucleico que proporciona información sobre la orientación espacial de una molécula diana que está asociada con el código de barras. Un marcador espacial puede asociarse a una coordenada en una muestra. La coordenada puede ser una coordenada fija. Por ejemplo, una coordenada puede fijarse en referencia a un sustrato. Un marcador espacial puede referirse a una cuadrícula bi- o tridimensional. Una coordenada puede fijarse en referencia 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 puedan obtenerse imágenes. 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 asociarse a una división física (por ejemplo, un pocillo, un recipiente, una microesfera, un tubo, una microcápsula o una gotita). En algunas realizaciones, se usan varios marcadores espaciales juntos 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 partícula sintética como una perla), pero diferente para diferentes soportes sólidos (por ejemplo, partículas sintéticas). En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprenden el mismo marcador espacial puede ser, o ser aproximadamente, del 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, o un número o un intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprenden el mismo marcador espacial puede ser como mínimo, 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 del mismo soporte sólido pueden incluir el mismo marcador espacial. En algunas realizaciones, por lo menos el 95% de los códigos de barras del mismo soporte sólido pueden incluir el mismo marcador espacial.

En una pluralidad de soportes sólidos (por ejemplo, partículas sintéticas como una perla) puede haber representadas hasta  $10^6$  o más secuencias de marcadores espaciales únicas. 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 cualesquiera 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 tener entre 10 y 150 nucleótidos. Un marcador espacial puede tener entre 20 y 125 nucleótidos de longitud.

#### Marcadores celulares

Un código de barras puede comprender uno o más marcadores celulares. En algunas realizaciones, un marcador celular puede comprender una secuencia de ácido nucleico que proporciona información para determinar qué ácido nucleico diana 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 partícula sintética como una perla), pero diferente para diferentes soportes sólidos (por ejemplo, partículas sintéticas). En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprende el mismo marcador celular 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. En algunas realizaciones, el porcentaje de códigos de barras en el mismo soporte sólido que comprende el mismo marcador celular puede ser, o ser aproximadamente del 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.

En una pluralidad de soportes sólidos (por ejemplo, partículas sintéticas) puede haber representadas hasta  $10^6$  o más secuencias de marcadores celulares únicas. 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 cualesquiera 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, un marcador celular puede comprender entre aproximadamente 20 y aproximadamente 125 nucleótidos de longitud.

#### Secuencias de códigos 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 ácido nucleico que proporciona información de identificación para el tipo específico de especie de ácido nucleico diana hibridada con el código de barras. Una secuencia de código de barras puede comprender una secuencia de ácido nucleico que proporciona un contador (por ejemplo, que proporciona una aproximación) para la aparición específica de la especie de ácido nucleico diana hibridada con el código de barras (por ejemplo, región de unión a diana).

En algunas realizaciones, un conjunto diverso de secuencias de código de barras se unen a un soporte sólido dado (por ejemplo, una partícula sintética, como 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 cualesquiera de estos valores, secuencias de marcadores moleculares únicos. 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 código de barras únicas. Las secuencias de marcadores moleculares únicos pueden unirse a un soporte sólido dado (por ejemplo, una partícula sintética).

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

En algunas realizaciones, un conjunto diverso de marcadores moleculares se unen a un soporte sólido dado (por ejemplo, una partícula sintética como 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 de secuencias de marcadores moleculares únicos. Por ejemplo, una pluralidad de códigos de barras estocásticos puede comprender aproximadamente 6561 marcadores moleculares con secuencias distintas. Como otro ejemplo, una pluralidad de códigos de barras estocásticos 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 únicos. Los códigos de barras estocásticos con las secuencias de marcadores moleculares únicos pueden unirse a un soporte sólido dado (por ejemplo, una partícula sintética).

Para la identificación con códigos de barras estocásticos usando una pluralidad de códigos de barras estocásticos, la proporción entre el número de secuencias de marcadores moleculares diferentes y el número de apariciones de cualquiera de las dianas puede ser, o ser aproximadamente, de 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 intervalo entre dos cualesquiera de estos valores. Una diana puede ser una especie de ARNm que comprende moléculas de ARNm con secuencias idénticas o casi idénticas. En algunas realizaciones, la proporción entre el número de secuencias de marcadores moleculares diferentes y el número de apariciones de cualquiera de las dianas es como mínimo, o como máximo, de 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 cualesquiera 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 a la diana

Un código de barras puede comprender una o más regiones de unión a diana, como sondas de captura. En algunas realizaciones, una región de unión a diana puede hibridar con una diana de interés. En algunas realizaciones, una región de unión a diana puede comprender una secuencia de ácido nucleico que hibrida específicamente con una diana (por ejemplo, ácido nucleico diana, como un ácido nucleico celular a analizar). Por ejemplo, una región de unión a diana puede hibridar con un ácido nucleico diana en una secuencia génica específica. En algunas realizaciones, una región de unión a diana puede comprender una secuencia de ácido nucleico que puede unirse (por ejemplo, hibridar) con una ubicación específica de un ácido nucleico diana específico. En algunas realizaciones, la región de unión a la diana puede comprender una secuencia de ácido nucleico que es capaz de hibridar específicamente con un saliente de sitio de enzima de restricción (por ejemplo, un saliente de extremo pegajoso EcoRI). El código de barras puede entonces ligarse a cualquier molécula de ácido nucleico que comprenda una secuencia complementaria al saliente del sitio de restricción.

En algunas realizaciones, una región de unión a diana puede comprender una secuencia de ácido nucleico diana no específica. Una secuencia de ácido nucleico diana no específica puede referirse a una secuencia que puede unirse a múltiples ácidos nucleicos diana, independientemente de la secuencia específica del ácido nucleico diana. Por ejemplo, la región de unión a la diana puede comprender una secuencia multimérica aleatoria, o una secuencia oligo(dT) que hibrida con las colas poli(A) de las moléculas de ARNm. Una secuencia multimérica aleatoria puede ser, por ejemplo, un dímero, trímero, cuádrmero, pentámero, hexámero, septámero, octámero, nonámero, decámero o una secuencia multimérica superior aleatoria de cualquier longitud. En algunas realizaciones, la región de unión a la diana es la misma para todos los códigos de barras unidos a una partícula sintética determinada (por ejemplo, una perla). En algunas realizaciones, las regiones de unión a diana para la pluralidad de códigos de barras unidos a una partícula sintética dada pueden comprender dos o más secuencias de unión a diana diferentes. Una región de unión a diana puede tener una longitud de 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 nucleótidos, o un número o intervalo entre dos cualesquiera de estos valores. Una región de unión a diana 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 a diana puede comprender un oligo(dT) que puede hibridar con ARNm que comprenden extremos poliadenilados. Una región de unión a diana puede ser específica de un gen. Por ejemplo, una región de unión a diana puede configurarse para que hibride con una región específica de una diana. Una región de unión a diana 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 intervalo entre dos cualesquiera de estos valores, nucleótidos de longitud. Una región de unión a diana 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 a la diana puede tener entre 5 y 30 nucleótidos de longitud. Cuando un código de barras comprende una región de unión a diana específica de un gen, en la presente el código de barras puede denominarse código de barras específico de un gen.

En algunas realizaciones, un código de barras no incluye una región de unión a diana. En algunas realizaciones, un código de barras incluye una región que corresponde a una región de unión a diana con una secuencia que tiene una afinidad de unión baja (por ejemplo, no se une) a algunas, la mayoría, sustancialmente todas

o todas las moléculas de ARNm en una célula o células de una muestra. Por ejemplo, un código de barras puede comprender una región que corresponde a una región de unión a diana que puede tener una secuencia que no se une a moléculas de ARNm de interés. Si la región de unión a la diana de un primer código de barras comprende una secuencia oligo(dT) que es capaz de hibridar con la cola poli(A) de una molécula de ARNm, la región correspondiente de un segundo código de barras puede incluir, por ejemplo, una secuencia que no sea similar, o sustancialmente similar, a una secuencia poli(dT). Si la región de unión a la diana de un primer código de barras comprende una región de unión a la diana con una secuencia capaz de hibridar específicamente con una secuencia génica particular, la región correspondiente de un segundo código de barras puede incluir, por ejemplo, una secuencia que no sea similar, o sustancialmente similar, a la región de unión a la diana.

#### Propiedad de orientación

Un código de barras 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 experimentar un enfoque isoelectrico para orientar los códigos de barras de una manera conocida. De esta manera, la propiedad de orientación puede usarse para desarrollar un mapa conocido de códigos de barras en una muestra. Las propiedades de orientación ejemplares pueden incluir, movilidad electroforética (por ejemplo, basada en el 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 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. Las dianas en y/o cerca del tipo de célula o molécula específica pueden marcarse 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 el 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 recombinatorios de fragmentos de genes de inmunoglobulina normales) (por ejemplo, un anticuerpo IgG) o una parte inmunológicamente activa (es decir, de unión específica) de una molécula de inmunoglobulina, como un fragmento de anticuerpo.

El fragmento de anticuerpo puede ser, por ejemplo, una parte de un anticuerpo como F(ab')<sub>2</sub>, Fab', Fab, Fv, sFv y similares. En algunas realizaciones, el fragmento de anticuerpo puede unirse al mismo antígeno reconocido por el anticuerpo de longitud completa. El fragmento de anticuerpo puede incluir fragmentos aislados consistentes 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 recombinantes de cadena sencilla en las que las regiones variables ligera y pesada están conectadas por un conector peptídico ("proteínas scFv"). Los anticuerpos ejemplares pueden incluir, entre otros, anticuerpos contra células cancerígenas, anticuerpos contra virus, anticuerpos que se unen a receptores de la superficie celular (CD8, CD34, CD45) y anticuerpos terapéuticos.

#### Cebador 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 adaptador universal. Un cebador adaptador universal puede referirse a una secuencia de nucleótidos que es universal para todos los códigos de barras. Un cebador adaptador universal puede usarse para construir códigos de barras específicos de genes. Un cebador 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 intervalo entre dos cualesquiera de estos nucleótidos de longitud. Un cebador adaptador universal 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. Un cebador adaptador universal puede tener una longitud de 5-30 nucleótidos.

Conector

Quando 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 intercalarse con una secuencia de marcador conector. Una secuencia de marcador conector puede tener por lo menos aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 o más nucleótidos de longitud. Una secuencia de marcador 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 de marcador conector tiene 12 nucleótidos de longitud. Una secuencia de marcador conector puede usarse para facilitar la síntesis del código de barras. El marcador conector puede incluir 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 o una partícula sintética. En algunas realizaciones, algunas o todas las secuencias de códigos de barras, como los marcadores moleculares de los 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) sobre un soporte sólido difieren en por lo menos un nucleótido. Los marcadores celulares de los códigos de barras en el mismo soporte sólido pueden ser iguales. 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 en un primer soporte sólido pueden tener la misma secuencia, y los segundos marcadores celulares de una segunda pluralidad de códigos de barras en 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, aproximadamente entre 5 y 20 nucleótidos de longitud. Una secuencia de código de barras puede tener, por ejemplo, aproximadamente entre 5 y 20 nucleótidos de longitud. La partícula sintética puede ser, por ejemplo, una perla.

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

En algunas realizaciones, una partícula sintética puede ser una partícula polimérica, por ejemplo una partícula deformable o una partícula de gel, funcionalizada con códigos de barras o códigos de barras estocásticos (como las perlas de gel de 10X Genomics (San Francisco, CA)). En alguna realización, una partícula de gel puede comprender uno o más geles basados en polímeros. Las partículas de gel pueden generarse, por ejemplo, encapsulando uno o más precursores poliméricos en gotitas. Tras la exposición de los precursores poliméricos a un acelerador (por ejemplo, tetrametiletilendiamina (TEMED)), puede generarse una partícula de gel.

En algunas realizaciones, la partícula puede ser degradable. Por ejemplo, la partícula o perla polimérica puede disolverse, fundirse o degradarse, por ejemplo, bajo una condición deseada. La condición deseada puede incluir una condición ambiental. La condición deseada puede dar como resultado que la partícula polimérica se disuelva, funda o degrade de manera controlada. Una partícula 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 luminoso o cualquier combinación de los mismos.

Los reactivos, como los códigos de barras de oligonucleótidos, pueden acoplarse/inmovilizarse en la superficie interior de una partícula sintética (por ejemplo, el interior accesible por difusión de un código de barras de oligonucleótidos y/o los materiales usados para generar un código de barras de oligonucleótidos) y/o la superficie exterior de una partícula de gel o cualquier otra microcápsula descrita en la presente. La partícula puede ser, por ejemplo, una perla de gel. La asociación (por ejemplo, acoplamiento o inmovilización) puede realizarse mediante cualquier forma de enlace químico (por ejemplo, enlace covalente, enlace iónico) o fenómeno físico (por ejemplo, fuerzas de Van der Waals, interacciones dipolo-dipolo, etc.). En algunas realizaciones, la asociación (por ejemplo, acoplamiento o inmovilización) de un reactivo a una partícula o a cualquier otro soporte sólido (por ejemplo, microcápsula) descrito 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 la aplicación de 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 esté inmovilizado en una partícula de gel mediante 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 partícula. La fracción lábil puede incluirse como parte de una partícula, perla o microcápsula de gel, como parte de un

conector químico que enlaza un reactivo a una perla o microcápsula de gel, y/o como parte de un reactivo. En algunas realizaciones, por lo menos un código de barras de la pluralidad de códigos de barras puede estar inmovilizado en la partícula, parcialmente inmovilizado en la partícula, encerrado en la partícula, parcialmente encerrado en la partícula, o cualquier combinación de los mismos.

En algunas realizaciones, una partícula (por ejemplo una perla de gel) puede comprender una amplia variedad de diferentes polímeros, incluyendo pero no limitados a: polímeros, polímeros sensibles al calor, polímeros fotosensibles, polímeros magnéticos, polímeros sensibles al pH, polímeros sensibles a la sal, polímeros químicamente sensibles, polielectrolitos, polisacáridos, péptidos, proteínas y/o plásticos. Los polímeros pueden incluir, entre otros, materiales como poli(N-isopropilacrilamida) (PNIPAAm), poli(sulfonato de estireno) (PSS), poli(alil amina) (PAAm), poli(ácido acrílico) (PAA), polietileno imina (PEI), poli(cloruro de dialildimetil-amonio) (PDADMAC), poli(pirrol) (PPy), poli(vinilpirrolidona) (PVPON), poli(vinilpiridina) (PVP), poli(ácido metacrílico) (PMAA), poli(metacrilato de metilo) (PMMA), poliestireno (PS), poli(tetrahidrofurano) (PTHF), poli(ftalaldehído) (PTHF), poli(hexil viologeno) (PHV), poli(L-lisina) (PLL), poli(L-arginina) (PARG), poli(ácido láctico-co-glicólico) (PLGA).

Para desencadenar la alteración, disolución o degradación de las partículas pueden usarse numerosos estímulos químicos. Ejemplos de estos cambios químicos pueden incluir, pero no se limitan a, cambios en la pared de la partícula mediados por el pH, disgregación de la pared de la partícula a través de la escisión química de los enlaces cruzados, despolimerización desencadenada de la pared de la partícula y reacciones de cambio de la pared de la partícula. Para desencadenar la disgregación de las partículas también pueden usarse cambios de volumen.

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

También pueden usarse estímulos biológicos para desencadenar la alteración, disolución o degradación de partículas. Generalmente, los desencadenantes biológicos se parecen 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 partículas pueden comprender polímeros con enlaces cruzados peptídicos que son sensibles a la escisión por proteasas específicas. Más específicamente, un ejemplo puede comprender una microcápsula que comprende enlaces cruzados peptídicos GFLGK. Tras la adición de un desencadenante biológico, como la proteasa catepsina B, los enlaces cruzados peptídicos de la microcápsula se escinden y se libera el contenido de las partículas. En otros casos, las proteasas pueden activarse por calor. En otro ejemplo, las partículas 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 pared de cubierta y la liberación de su contenido interno.

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

La inclusión de nanopartículas magnéticas en la pared de partículas de las microcápsulas puede permitir la ruptura desencadenada de las partículas, así como guiar las partículas en una matriz. Un dispositivo de la presente divulgación puede comprender partículas magnéticas para ambos propósitos. En un ejemplo, la incorporación de nanopartículas de  $\text{Fe}_3\text{O}_4$  en partículas que contienen polielectrolitos desencadena la ruptura en presencia de un estímulo de campo magnético oscilante.

Una partícula 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 partículas eléctricamente sensibles pueden permitir tanto la ruptura desencadenada de las partículas como otras funciones tales como la alineación en un campo eléctrico, la conductividad eléctrica o las reacciones redox. En un ejemplo, las partículas que contienen material eléctricamente sensible se alinean en un campo eléctrico de tal manera que pueda controlarse la liberación de reactivos internos. En otros ejemplos, los campos eléctricos pueden inducir reacciones redox dentro de la propia pared de la partícula que pueden aumentar la porosidad.

Para alterar las partículas también puede usarse un estímulo luminoso. Son posibles numerosos activadores lumínicos 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 metálico como activadores de cápsulas. La irradiación UV de cápsulas de polielectrolito recubiertas con  $\text{SiO}_2$  puede

provocar la disgregación de la pared de la partícula. En otro ejemplo más, pueden incorporarse a la pared de la partícula materiales fotoconmutables, como grupos azobenceno. Tras la aplicación de luz UV o visible, estos productos químicos experimentan una isomerización reversible de cis a trans por absorción de fotones. En este aspecto, la incorporación de interruptores de fotones da como resultado una pared de partículas que puede disgregarse o volverse más porosa tras la aplicación de un disparador lumínico.

Por ejemplo, en un ejemplo no limitativo de identificación con códigos de barras (por ejemplo, identificación con códigos de barras estocásticos) 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, en el bloque 212 pueden introducirse partículas en la pluralidad de micropocillos de la matriz de micropocillos. Cada micropocillo puede comprender una partícula. Las partículas pueden comprender una pluralidad de códigos de barras. Un código de barras puede comprender una región amina 5' unida a una partícula. El código de barras puede comprender un marcador universal, una secuencia de código de barras (por ejemplo, un marcador molecular), una región de unión a diana o cualquier combinación de las mismas.

Los códigos de barras divulgados en la presente pueden asociarse con (por ejemplo, unirse a) un soporte sólido (por ejemplo, una partícula o una perla). Los códigos de barras asociados con un soporte sólido pueden comprender cada uno una secuencia de código 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 secuencias de códigos de barras de secuencias diferentes. En algunas realizaciones, un porcentaje de los códigos de barras asociados a un soporte sólido comprende el mismo marcador celular. Por ejemplo, el porcentaje puede ser, o ser aproximadamente del 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, 100%, o un número o un intervalo entre dos cualesquiera de estos valores. Como otro ejemplo, el porcentaje puede ser como mínimo, o como máximo del 60%, 70%, 80%, 85%, 90%, 95%, 97%, 99%, o 100%. En algunas realizaciones, los códigos de barras asociados con 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 partícula, como una perla). En algunas realizaciones, identificar estocásticamente con códigos de barras la pluralidad de dianas 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 en 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 partículas pueden ser perlas de gel de sílice, perlas de vidrio de poro controlado, perlas magnéticas, Dynabeads, perlas 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 estar insertados en una matriz semisólida o sólida. Los códigos de barras pueden no estar asociados a los soportes sólidos. Los códigos de barras pueden ser nucleótidos individuales. Los códigos de barras pueden estar asociados a un sustrato.

Como se usan 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. Para unir códigos de barras presintetizados o para la síntesis *in situ* de códigos de barras en fase sólida puede usarse cualquiera de una variedad de diferentes soportes sólidos.

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

Las partículas 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 ( $\text{Fe}_3\text{O}_4$ ; magnetita)), materiales ferromagnéticos (por ejemplo hierro, níquel, cobalto, algunas aleaciones de los mismos y algunos compuestos metálicos de tierras raras), cerámica, plástico, vidrio, poliestireno, sílice, metilpoliestireno, polímeros acrílicos, titanio, látex, sefrosa, agarosa, hidrogel, polímero, celulosa, nailon o cualquier combinación de los mismos.

En algunas realizaciones, la partícula (por ejemplo, la partícula a la que se unen los marcadores) es una perla

de hidrogel. En algunas realizaciones, la partícula 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, un marcador molecular), un marcador celular y una región de unión a diana (por ejemplo, una secuencia oligo(dT), una secuencia específica de un 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 igual. Las secuencias de el marcador celular de los oligonucleótidos en partículas diferentes pueden ser diferentes, de tal manera que puedan identificarse los oligonucleótidos en partículas diferentes. En diferentes implementaciones el número de secuencias de marcadores celulares diferentes puede ser diferente. En algunas realizaciones, el número de secuencias de marcado celular puede ser, o ser aproximadamente 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,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ , un número o un intervalo entre dos cualesquiera de estos valores, o más. En algunas realizaciones, el número de secuencias de marcadores celulares puede ser como mínimo 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  $10^9$ . 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 las partículas incluyen oligonucleótidos con la misma secuencia celular. En algunas realizaciones, 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 código de barras (por ejemplo, marcadores moleculares). En algunas realizaciones, el número de secuencias de código de barras puede ser, o ser aproximadamente 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,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ , o un número o un intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de secuencias de códigos de barras puede ser como mínimo 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 cualesquiera de estos valores, o más de la pluralidad de oligonucleótidos comprenden secuencias de código de barras diferentes. Algunas realizaciones proporcionan una pluralidad de las partículas que comprenden códigos de barras. En algunas realizaciones, la proporción de una aparición (o una copia o un número) de una diana a marcar y las diferentes secuencias de código de barras puede ser de 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 uno 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 una micropartícula.

El tamaño de una partícula puede variar. Por ejemplo, el diámetro de la partícula puede variar entre 0,1 micrómetros y 50 micrómetros. En algunas realizaciones, los diámetros de las partículas pueden ser, o ser de 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 partícula puede estar relacionado con el diámetro de los pocillos del sustrato. En algunas realizaciones, el diámetro de la partícula puede ser, o ser aproximadamente, un 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, o un número o un intervalo entre dos cualesquiera de estos valores, más largo o más corto que el diámetro del pocillo. El diámetro de la partícula puede estar relacionado con el diámetro de una célula (por ejemplo, una célula individual atrapada por un pocillo del sustrato). En algunas realizaciones, el diámetro de la partícula 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 la partícula puede estar relacionado con el diámetro de una célula (por ejemplo, una única célula atrapada por un pocillo del sustrato). En algunas realizaciones, el diámetro de la partícula 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 cualesquiera de estos valores, más largo o más corto que el diámetro de la célula. En algunas realizaciones, el diámetro de la partícula 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 partícula puede estar unida y/o insertada en un sustrato. Una partícula puede estar unida y/o insertada en un gel, hidrogel, polímero y/o matriz. La posición espacial de una partícula 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 partícula que puede servir como dirección de localización.



Ejemplos de partículas pueden incluir, pero no se limitan a, perlas de estreptavidina, perlas de agarosa, perlas magnéticas, perlas Dynabeads®, perlas MACS®, perlas conjugadas con anticuerpos (por ejemplo, perlas 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 la sílice, microperlas antibiotina, microperlas antifluorocromo, y perlas magnéticas terminadas en carboxilo BcMag™.

Una partícula puede asociarse con (por ejemplo, impregnarse con) puntos cuánticos o tintes fluorescentes para hacerla fluorescente en un canal óptico de fluorescencia o en múltiples canales ópticos. Una partícula puede asociarse con óxido de hierro u óxido de cromo para hacerla paramagnética o ferromagnética. Las partículas pueden ser identificables. Por ejemplo, pueden obtenerse imágenes de una partícula mediante una cámara. Una partícula puede tener asociado un código detectable. Por ejemplo, una partícula puede incluir un código de barras. Una partícula puede cambiar de tamaño, por ejemplo debido al hinchamiento en una solución orgánica o inorgánica. Una partícula puede ser hidrófoba. Una partícula puede ser hidrófila. Una partícula puede ser biocompatible.

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

Un soporte sólido puede comprender un material insoluble, semisoluble o soluble. Un soporte sólido puede denominarse "funcionalizado" cuando incluye un conector, un andamiaje, un bloque de construcción u otra fracción reactiva unida al mismo, mientras que un soporte sólido puede ser "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 de flujo 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 adoptar 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, plata dorada, aluminio, silicio y cobre), soportes de vidrio, soportes de plástico, soportes de silicio, chips, filtros, membranas, placas de micropocillos, portaobjetos, materiales plásticos, incluyendo placas o membranas multipocillo (por ejemplo, formadas de polietileno, polipropileno, poliamida, polivinilidendifluoruro), y/u obleas, peines, alfileres o agujas (por ejemplo, matrices de alfileres adecuadas para síntesis o análisis combinatorios) o partículas (por ejemplo, partículas sintéticas como perlas) en una matriz de fosas o pocillos de nanolitros de superficies planas como obleas (por ejemplo, obleas de silicio), obleas con fosas con o sin fondos filtrantes.

El soporte sólido puede comprender una matriz polimérica (por ejemplo, gel, hidrogel). La matriz polimérica puede ser capaz de permeabilizar el espacio intracelular (por ejemplo, alrededor de orgánulos). La matriz polimérica puede ser capaz de ser bombeada a través del sistema circulatorio.

Un soporte sólido puede ser una molécula biológica. Por ejemplo, un soporte sólido puede ser, o puede comprender, un ácido nucleico, una proteína, un anticuerpo, una histona, un compartimento celular, un lípido, un carbohidrato, y similares. Los soportes sólidos que son moléculas biológicas pueden amplificarse, traducirse, transcribirse, degradarse y/o modificarse (por ejemplo, pegarse, sumoarse, acetilarse, metilarse). Un soporte sólido que es una molécula biológica puede proporcionar información espacial y temporal además del marcador espacial que está unida a la molécula biológica. Por ejemplo, una molécula biológica puede comprender una primera confirmación cuando está sin modificar, pero puede cambiar a una segunda confirmación cuando se modifica. Las diferentes conformaciones pueden exponer los códigos de barras (por ejemplo, códigos de barras estocásticos) de la divulgación a los objetivos. Por ejemplo, una molécula biológica puede comprender códigos de barras que son inaccesibles debido al plegamiento de la molécula biológica. Después de la modificación de la molécula biológica (por ejemplo, acetilación), la molécula biológica puede cambiar su conformación para exponer los códigos de barras. La cadencia de la modificación puede proporcionar otra dimensión temporal al método de identificación con códigos de barras de la divulgación.

En algunas realizaciones, la molécula biológica que comprende reactivos de código de barras de la divulgación puede estar localizada en el citoplasma de una célula. Tras la activación, la molécula biológica puede desplazarse al núcleo, donde puede tener lugar la identificación con códigos de barras. De esta manera, la modificación de la molécula biológica puede codificar información espacio-temporal adicional para las dianas identificadas por los códigos de barras.

#### 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 y 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 pequeña cámara de reacción de volumen definido. En algunas realizaciones, un micropocillo puede atrapar una o más células. En algunas realizaciones, un micropocillo puede atrapar sólo una célula. En algunas realizaciones, un micropocillo puede atrapar uno o más soportes sólidos. En algunas realizaciones, un micropocillo puede atrapar sólo un soporte sólido. En algunas realizaciones, un micropocillo atrapa una única célula y un único soporte sólido (por ejemplo, una partícula). Un micropocillo puede comprender reactivos combinatorios de códigos de barras de la divulgación.

#### Síntesis de códigos de barras sobre soportes sólidos

Un código de barras (por ejemplo, un código de barras estocástico) puede sintetizarse sobre un soporte sólido (por ejemplo, una partícula, como una partícula sintética o una perla). Los códigos de barras presintetizados (por ejemplo, que comprenden la amina 5' que puede unirse al soporte sólido) pueden unirse a soportes sólidos (por ejemplo, partículas) a través de cualquiera de una variedad de técnicas de inmovilización que implican pares de grupos funcionales en el soporte sólido y el código de barras. El código de barras puede comprender un grupo funcional. El soporte sólido (por ejemplo, una partícula) puede incluir un grupo funcional. El grupo funcional del código de barras y el grupo funcional del soporte sólido pueden comprender, por ejemplo, biotina, estreptavidina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, y cualquier combinación de los mismos. Un código de barras (por ejemplo, un código de barras estocástico) puede fijarse a un soporte sólido, por ejemplo, acoplado (por ejemplo, usando 1-etil-3-(3-dimetilaminopropil)carbodiimida) un grupo amino 5' del código de barras al grupo carboxilo del soporte sólido funcionalizado. Los códigos de barras residuales no acoplados pueden eliminarse de la mezcla de reacción realizando múltiples pasos de enjuagado. En algunas realizaciones, el código de barras y el soporte sólido se unen indirectamente a través de moléculas conectoras (por ejemplo, moléculas cortas de hidrocarburo funcionalizado o moléculas de óxido de polietileno) usando técnicas químicas de unión similares. Los conectores pueden ser conectores escindibles, por ejemplo, conectores lábiles al ácido o fotoescindibles.

Los códigos de barras (por ejemplo, códigos de barras estocásticos) pueden sintetizarse sobre soportes sólidos (por ejemplo, partículas) usando cualquiera de las técnicas de síntesis de oligonucleótidos en fase sólida, como la síntesis de fosfodiéster, la síntesis de fosfortriéster, la síntesis de triéster fosfito y la síntesis de fosforamidita. Los nucleótidos individuales pueden acoplarse paso a paso al código de barras anclado en crecimiento. Al código de barras anclado en crecimiento puede acoplarse una secuencia corta presintetizada (o bloque) de varios oligonucleótidos.

Los códigos de barras (por ejemplo, los códigos de barras estocásticos) pueden sintetizarse intercalando reacciones de acoplamiento por pasos o en bloque con una o más rondas de síntesis de grupos divididos, en la que el conjunto total de partículas de síntesis se divide en una serie de grupos individuales más pequeños, cada uno de los cuales se somete después a una reacción de acoplamiento diferente, seguido de la recombinación y el mezclado de los grupos individuales para aleatorizar la secuencia del código de barras en crecimiento en el grupo total de partículas. La síntesis de grupos divididos es un ejemplo de proceso de síntesis combinatoria en el que se sintetiza un número máximo de compuestos químicos usando un número mínimo de pasos de acoplamiento químico. La diversidad potencial de la biblioteca de compuestos creada de este modo viene determinada por el número de bloques de construcción únicos (por ejemplo, nucleótidos) disponibles para cada paso de acoplamiento y el número de pasos de acoplamiento usados para crear la biblioteca. Por ejemplo, una síntesis de grupos divididos que comprende 10 rondas de acoplamiento usando 4 nucleótidos diferentes en cada paso producirá  $4^{10} = 1.048.576$  secuencias de nucleótidos únicas. En algunas realizaciones, la síntesis de grupos divididos puede llevarse a cabo usando métodos enzimáticos tales como reacciones de extensión o ligación de polimerasas en lugar de acoplamiento químico. Por ejemplo, en cada ronda de una reacción de extensión de polimerasa de grupos divididos, los extremos 3' de los códigos de barras unidos a partículas en un grupo dado pueden hibridar con los extremos 5' de un conjunto de cebadores semialeatorios, por ejemplo cebadores que tengan una estructura de 5'-(M)<sub>k</sub>-(X)<sub>i</sub>-(N)<sub>j</sub>-3', donde (X)<sub>i</sub> es una secuencia aleatoria de nucleótidos que tiene una longitud de i nucleótidos (el conjunto de cebadores que comprende todas las combinaciones posibles de (X)<sub>i</sub>), (N)<sub>j</sub> es un nucleótido específico (o una serie de j nucleótidos) y (M)<sub>k</sub> es un nucleótido específico (o serie de k nucleótidos), en donde se añade un desoxirribonucleótido trifosfato (dNTP) diferente a cada grupo y es incorporado en los oligonucleótidos ligados por la polimerasa.

#### Métodos de identificación con códigos de barras

La divulgación proporciona métodos para estimar el número de dianas distintas en una muestra. En algunas realizaciones, la identificación con códigos de barras de la pluralidad de dianas comprende hibridar una pluralidad de códigos de barras con una pluralidad de dianas para crear dianas identificadas con códigos de barras (por ejemplo, dianas identificadas estocásticamente con códigos de barras). La identificación con códigos de barras de la pluralidad de dianas puede comprender generar una biblioteca indexada de las dianas identificadas con códigos de barras. La generación de una biblioteca indexada de las dianas identificadas 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, por flujo de 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 extenderse por el sustrato, por ejemplo, haciendo crecer/cultivando las células en el sustrato.

Cuando los códigos de barras están en proximidad cercana de las dianas, las dianas 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 diana distinta pueda asociarse con un código de barras distinto de la divulgación. Para garantizar una asociación eficaz entre la diana y el código de barras, las dianas pueden reticularse con el código de barras.

Lisis celular

Después de la distribución de células y códigos de barras, las células pueden lisarse para liberar las moléculas diana. La lisis celular puede realizarse mediante cualquiera de una variedad de medios, por ejemplo, por 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 comprenda 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 una diana y un código de barras, puede alterarse la velocidad de difusión de las moléculas diana, 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 sobre el 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 las dianas de la muestra con el sustrato.

En algunas realizaciones, la lisis puede llevarse a cabo 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, o como máximo aproximadamente, 0,01 M, 0,05 M, 0,1 M, 0,5 M, o 1 M o más o menos Tris HCl. Un tampón de lisis puede contener aproximadamente 0,01 M, 0,05 M, 0,1 M, 0,5 M o 1 M de Tris HCl. El pH del tampón de lisis puede ser por lo menos de 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 contener 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 como máximo de 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 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. La concentración del detergente en el tampón de lisis puede ser como máximo del 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 un 1% de dodecil sulfato de Li. El tiempo usado en el método para la lisis puede depender de la cantidad de detergente usado. En algunas realizaciones, cuanto más detergente se usa, menos tiempo será necesario para la lisis. El tampón de lisis puede incluir 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 mM, 5 mM, 10 mM, 15 mM, 20 mM, 25 mM o 30 mM, o más. La concentración de un agente quelante en el tampón de lisis puede ser como máximo de aproximadamente 1 mM, 5 mM, 10 mM, 15 mM, 20 mM, 25 mM 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 por lo menos de aproximadamente 1 mM, 5 mM, 10 mM, 15 mM o 20 mM o más. La concentración del reactivo reductor en el tampón de lisis puede ser como máximo de aproximadamente 1 mM, 5 mM, 10 mM, 15 mM 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 0,1M, aproximadamente pH 7,5, LiCl aproximadamente 0,5M, aproximadamente un 1% de dodecil sulfato de litio, EDTA aproximadamente 10mM, y DTT aproximadamente 5mM.

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, 20 o más minutos. Una célula lisada puede comprender por lo menos aproximadamente 100000, 200000, 300000, 400000, 500000, 600000, 700000 o más moléculas de ácido nucleico

diana. Una célula lisada puede comprender como máximo aproximadamente 100000, 200000, 300000, 400000, 500000, 600000, 700000 o más moléculas de ácido nucleico diana.

#### Unión de códigos de barras a moléculas de ácido nucleico diana

Después de la lisis de las células y la liberación de moléculas de ácido nucleico de las mismas, las moléculas de ácido nucleico pueden asociarse aleatoriamente con los códigos de barras del soporte sólido colocalizado. La asociación puede comprender la hibridación de la región de reconocimiento de la diana de un código de barras con una parte complementaria de la molécula de ácido nucleico diana (por ejemplo, el oligo(dT) del código de barras puede interactuar con una cola de poli(A) de una diana). Las condiciones de ensayo usadas para la hibridación (por ejemplo, pH del tampón, fuerza iónica, temperatura, etc.) pueden elegirse para que promuevan la formación de híbridos específicos y estables. En algunas realizaciones, las moléculas de ácido nucleico liberadas de las células lisadas pueden asociarse con la pluralidad de sondas del sustrato (por ejemplo, hibridar con las sondas del sustrato). Cuando las sondas comprenden oligo(dT), las moléculas de ARNm pueden hibridar con las sondas y transcribirse inversamente. La parte 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 identificación con códigos de barras ilustrado en la FIG. 2, en el bloque 216, las moléculas de ARNm pueden hibridar con los códigos de barras en las partículas. Por ejemplo, los fragmentos de nucleótidos de cadena sencilla pueden hibridar con las regiones de unión a diana de los códigos de barras.

La unión puede comprender además la ligación de una región de reconocimiento de diana de un código de barras y una parte de la molécula de ácido nucleico diana. Por ejemplo, la región de reconocimiento de diana puede comprender una secuencia de ácido nucleico que puede ser capaz de hibridación específica con un saliente de sitio de restricción (por ejemplo, un saliente de extremo pegajoso de EcoRI). El procedimiento de ensayo puede comprender además el tratamiento de los ácidos nucleicos diana con una enzima de restricción (por ejemplo, EcoRI) para crear un saliente de sitio de restricción. A continuación, 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 identificación con códigos de barras ilustrado en la FIG. 2, en el bloque 220, los dianas marcadas de una pluralidad de células (o una pluralidad de muestras) (por ejemplo, moléculas de diana-código de barras) pueden agruparse posteriormente, por ejemplo, en un tubo. Las dianas marcadas pueden agruparse, por ejemplo, recuperando los códigos de barras y/o las partículas a las que están unidas las moléculas de diana-código de barras.

La recuperación de colecciones basadas en soportes sólidos de moléculas de diana-código de barras unidas puede implementarse mediante el uso de partículas magnéticas y un campo magnético aplicado externamente. Una vez que se hayan agrupado las moléculas de diana-código de barras, todo el procesamiento posterior puede realizarse en un único recipiente de reacción. El procesamiento posterior 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. Las reacciones de procesamiento posterior pueden realizarse dentro de los micropocillos, es decir, sin agrupar primero las moléculas de ácido nucleico diana marcadas de una pluralidad de células.

#### Reacción de extensión de ácidos nucleicos (por ejemplo, transcripción inversa)

La divulgación proporciona un método para crear un conjugado diana-código de barras usando una reacción de extensión de ácidos nucleicos, como la transcripción inversa (por ejemplo, en el bloque 224 de la FIG. 2). El conjugado diana-código de barras puede comprender el código de barras y una secuencia complementaria de todo o parte del ácido nucleico diana (es decir, una molécula de ADNc identificada con código de barras, como una molécula de ADNc identificada estocásticamente con código de barras). 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 hexanucleotídico aleatorio o un cebador oligonucleotídico específico de diana. Los cebadores oligo(dT) pueden tener una longitud de entre 12 y 18 nucleótidos y unirse a la cola de poli(A) endógena en el extremo 3' del ARNm de mamífero. Los cebadores hexanucleotídicos aleatorios pueden unirse al ARNm en una variedad de sitios complementarios. Los cebadores oligonucleotídicos específicos de diana típicamente ceban selectivamente el ARNm de interés.

En algunas realizaciones, la transcripción inversa de la molécula de ARN marcado 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 hexanucleotídico aleatorio o un cebador oligonucleotídico específico de diana. Generalmente, los cebadores oligo(dT) tienen una longitud de 12-18 nucleótidos y se unen a la cola endógena poli(A)+ en el extremo 3' del ARNm de mamíferos. Los cebadores hexanucleotídicos aleatorios pueden unirse al ARNm en una variedad de sitios complementarios. Los cebadores oligonucleotídicos específicos de diana 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 la realización de 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 diana marcadas. La amplificación puede realizarse de manera multiplexada, en donde múltiples secuencias de ácido nucleico diana se amplifican simultáneamente. La reacción de amplificación puede usarse para añadir adaptadores de secuenciación a las moléculas de ácido nucleico. 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 de la secuencia del marcador celular y/o del código de barras (por ejemplo, el marcador molecular). Las reacciones de amplificación pueden comprender la amplificación de por lo menos una parte de una etiqueta de muestra, un marcador celular, un marcador espacial, un código de barras (por ejemplo, un marcador molecular), un ácido nucleico diana, 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 cualesquiera de estos valores, de la pluralidad de ácidos nucleicos. El método puede comprender además llevar a cabo una o más reacciones de síntesis de ADNc para producir una o más copias de ADNc de moléculas de diana-código de barras 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 llevarse a cabo usando 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 específicas de ADN mediante la extensión simultánea de cebadores de cadenas complementarias de ADN. Como se usa en la presente, 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 de ensamblaje.

La amplificación de los ácidos nucleicos marcados puede comprender métodos no basados en PCR. 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 secuencia de ácido nucleico (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 dianas de ADN o ARN, una reacción en cadena de ligasa (LCR), y un método de Q $\beta$  replicasa (Q $\beta$ ), uso de sondas palindrómicas, amplificación por desplazamiento de cadena, amplificación impulsada por oligonucleótidos usando una endonucleasa de restricción, un método de amplificación en el que un cebador hibrida con una secuencia de ácido nucleico 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 amplificación por extensión de ramificaciones (RAM). En algunas realizaciones, la amplificación no produce transcritos circularizados.

En algunas realizaciones, los métodos divulgados en la presente comprenden además llevar a cabo 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 cadena doble de ARN, una molécula de cadena doble de ADN o una molécula de ARN hibridada con una molécula de ADN. Una o ambas cadenas de la molécula de cadena doble pueden incluir 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. 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 glicólico (GNA) y ácido nucleico de treosa (TNA). Los nucleótidos no naturales pueden añadirse a uno o más ciclos de una reacción de amplificación. La adición de los 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. Los 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. Los 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. Los uno o más cebadores pueden comprender menos de 12-15 nucleótidos. Los uno o más cebadores pueden aparearse con por lo menos una parte de la pluralidad de dianas marcadas (por ejemplo, dianas marcadas estocásticamente). Los uno o más cebadores pueden unirse al extremo 3' o al extremo 5' de la pluralidad de dianas marcadas. Los uno o más cebadores pueden unirse a una región interna de la pluralidad de dianas marcadas. La región interna puede estar por lo menos a 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 de los extremos 3' de la pluralidad de dianas marcadas. Los uno o más cebadores puede comprender un panel fijo de cebadores. Los uno o más cebadores pueden comprender por lo menos uno o más cebadores personalizados. Los uno o más cebadores puede comprender por lo menos uno o más cebadores de control. Los uno o más cebadores puede comprender por lo menos uno o más cebadores específicos de gen.

Los uno o más cebadores puede comprender un cebador universal. El cebador universal puede aparearse con un sitio de unión de cebador universal. Los 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), una diana o cualquier combinación de las mismas. Los uno o más cebadores puede comprender un cebador universal y un cebador personalizado. El cebador personalizado puede diseñarse para que amplifique una o más dianas. Las dianas pueden comprender un subconjunto de los ácidos nucleicos totales en una o más muestras. Las dianas pueden comprender un subconjunto del total de dianas marcadas en una o más muestras. Los uno o más cebadores pueden comprender por lo menos 96 o más cebadores personalizados. Los uno o más cebadores puede comprender por lo menos 960 o más cebadores personalizados. Los uno o más cebadores pueden comprender por lo menos 9600 o más cebadores personalizados. Los 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 partícula (por ejemplo, una 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 anidado específico del gen 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 e 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, el marcador molecular) en la lectura 1, el gen en la lectura 2 y el índice de la muestra en la lectura del índice 1.

En algunas realizaciones, los ácidos nucleicos pueden eliminarse del sustrato mediante escisión química. Por ejemplo, para facilitar su eliminación de un soporte sólido puede usarse un grupo químico o una base modificada presente en un ácido nucleico. Por ejemplo, puede usarse una enzima para eliminar un ácido nucleico de un sustrato. Por ejemplo, un ácido nucleico puede eliminarse de un sustrato mediante una digestión con endonucleasas de restricción. Por ejemplo, para eliminar un ácido nucleico de un sustrato puede usarse el tratamiento de un ácido nucleico que contiene un dUTP o ddUTP con uracilo-d-glicosilasa (UDG). Por ejemplo, un ácido nucleico puede eliminarse de un sustrato usando una enzima que realiza la escisión de nucleótidos, como una enzima de reparación por 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 foto-lábil, un grupo conector ácido- o base-lábil, o un aptámero.

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

En algunas realizaciones, la amplificación puede llevarse a cabo en el sustrato, por ejemplo, con amplificación en puente. A los ADNc se les pueden añadir colas de homopolímero para generar un extremo compatible para la amplificación en puente usando sondas oligo(dT) en el sustrato. En la amplificación en puente, el cebador complementario al extremo 3' del ácido nucleico plantilla puede ser el primer cebador de cada par unido 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 de avance mediante la adición de nucleótidos para formar una molécula dúplex que consiste en la molécula plantilla y una cadena de ADN recién formada que es complementaria a la plantilla. En el paso de calentamiento del siguiente ciclo, la molécula dúplex puede desnaturalizarse, 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 alejada del primer cebador. Esta hibridación puede hacer que la cadena complementaria forme un puente entre el primer y el segundo cebadores fijado al primer cebador por un enlace covalente y al segundo cebador por hibridación. En la etapa de elongación, el segundo cebador puede alargarse en sentido inverso mediante la adición de nucleótidos en la misma mezcla de reacción, convirtiendo de este modo el puente en un puente de cadena doble. A continuación, comienza el siguiente ciclo, y el puente de cadena doble puede desnaturalizarse para dar proporcionar os moléculas de ácido nucleico de cadena sencilla, cada una de las cuales tiene un extremo unido a la superficie de la partícula a través del primer y el segundo cebadores, respectivamente, con el otro extremo de cada una de ellas sin unir. En el paso de apareamiento y elongación de este segundo ciclo, cada cadena puede hibridar con otro cebador complementario, 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 elongan para convertir los dos nuevos puentes en puentes de cadena doble.

Las reacciones de amplificación pueden comprender la amplificación de por lo menos el 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.

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 mediante reacción en cadena de la polimerasa (PCR). La PCR puede referirse a una reacción para la amplificación in vitro de secuencias específicas de ADN mediante la extensión simultánea de cebadores de cadenas complementarias de ADN. La PCR puede abarcar formas derivadas de la reacción, incluyendo, entre otras, RT-PCR, PCR en tiempo real, PCR anidada, PCR cuantitativa, PCR multiplexada, PCR digital, PCR de supresión, PCR semisupresiva y PCR de ensamblaje.

En algunas realizaciones, la amplificación de los ácidos nucleicos marcados comprende métodos no basados en PCR. 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 secuencia de ácido nucleico (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 la 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 dianas de ADN o ARN, una reacción en cadena de ligasa (LCR), una replicasa Q $\beta$  (Q $\beta$ ), uso de sondas palindrómicas, amplificación por desplazamiento de cadena, amplificación impulsada por oligonucleótidos usando una endonucleasa de restricción, un método de amplificación en el que un cebador se hibrida con una secuencia de ácido nucleico 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 ramificada (RAM).

En algunas realizaciones, los métodos divulgados en la presente comprenden además llevar a cabo una reacción en cadena de la polimerasa anidada en el amplicón amplificado (por ejemplo, diana). 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 contener una etiqueta de muestra o un marcador de identificación 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 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 llevar a cabo 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 el uso de uno o más nucleótidos no naturales. Los nucleótidos no naturales pueden comprender nucleótidos fotolábiles y/o activables. 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 glicólico (GNA) y ácido nucleico de treosa (TNA). Los nucleótidos no naturales pueden añadirse a uno o más ciclos de una reacción de amplificación. La adición de los 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. Los uno o más cebadores pueden comprender uno o más oligonucleótidos. Los uno o más oligonucleótidos pueden comprender por lo menos aproximadamente 7-9 nucleótidos. El o los oligonucleótidos pueden comprender menos de 12-15 nucleótidos. Los uno o más cebadores pueden aparearse con por lo menos una parte de la pluralidad de ácidos nucleicos marcados. Los uno o más cebadores pueden aparearse al extremo 3' y/o al extremo 5' de la pluralidad de ácidos nucleicos marcados. Los uno o más cebadores pueden aparearse en una región interna de la pluralidad de ácidos nucleicos marcados. La región interna puede ser de 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 de los extremos 3' de la pluralidad de ácidos nucleicos marcados. Los uno o más cebadores puede comprender un panel fijo de cebadores. Los uno o más cebadores pueden comprender por lo menos uno o más cebadores personalizados. Los uno o más cebadores puede comprender por lo menos uno o más cebadores de control. Los uno o más cebadores puede comprender por lo menos uno o más cebadores de genes constitutivos. Los uno o más cebadores puede comprender un cebador universal. El cebador universal puede aparearse con un sitio de unión del cebador universal. Los uno o más cebadores personalizados pueden aparearse con la primera etiqueta de muestra, la segunda etiqueta de muestra, la etiqueta identificadora molecular, al ácido nucleico o a un producto de los mismos. Los uno o más cebadores pueden comprender un cebador universal y un cebador personalizado. El cebador personalizado puede diseñarse para que amplifique uno o más ácidos nucleicos diana. Los ácidos nucleicos diana 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 identificación con códigos de barras (por ejemplo, identificación estocástica con códigos de barras) de la pluralidad de dianas en la muestra comprende además la generación de una biblioteca indexada de los fragmentos identificados con códigos de barras. 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 las dianas con códigos de barras (por ejemplo, dianas identificadas estocásticamente con códigos de barras) incluye la generación de una pluralidad de polinucleótidos indexados a partir de la pluralidad de dianas de la muestra. Por ejemplo, para una biblioteca indexada de dianas identificadas con códigos de barras que comprende una primera diana indexada y una segunda diana indexada, la región del marcador del primer polinucleótido indexado puede diferir de la región del marcador 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 intervalo entre dos cualesquiera de estos valores, nucleótidos. En algunas realizaciones, la generación de una biblioteca indexada de las dianas identificadas con códigos de barras incluye poner en contacto una pluralidad de dianas, por ejemplo moléculas de ARNm, con una pluralidad de oligonucleótidos que incluyen una región poli(T) y una región de marcador; y llevar a cabo una síntesis de primera cadena usando una transcriptasa inversa para producir moléculas de ADNc marcadas de cadena sencilla que cada una comprenden una región de ADNc y una región de marcador, en donde la pluralidad de dianas 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. La generación de una biblioteca indexada de las dianas identificadas con códigos de barras puede comprender además la amplificación de las moléculas de ADNc marcadas de cadena sencilla para producir moléculas de ADNc marcadas de cadena doble; y la realización de PCR anidada en las moléculas de ADNc marcadas de cadena doble para producir amplicones marcados. En algunas realizaciones, el método puede incluir la generación de un amplicón marcado con adaptador.

La identificación con códigos de barras estocásticos puede usar códigos de barras o etiquetas de ácido nucleico para marcar moléculas individuales de ácido nucleico (por ejemplo, ADN o ARN). En algunas realizaciones, implica añadir códigos de barras o etiquetas de ADN a moléculas de ADNc a medida que se generan a partir de ARNm. Puede realizarse una PCR anidada para minimizar el sesgo de amplificación de la PCR. Pueden añadirse adaptadores para la secuenciación usando, por ejemplo, la secuenciación de próxima generación (NGS). Los resultados de la secuenciación pueden usarse para determinar los marcadores celulares, las secuencias de códigos de barras (por ejemplo, marcadores moleculares) y las secuencias de fragmentos de nucleótidos de la una o más copias de las dianas, por ejemplo en el bloque 232 de la FIG. 2.

La FIG. 3 es una ilustración esquemática que muestra un proceso ejemplar no limitativo de generación de una biblioteca indexada de dianas identificadas con códigos de barras (por ejemplo, dianas identificadas estocásticamente con códigos de barras), por ejemplo ARNm. Como se muestra en el paso 1, el proceso de transcripción inversa puede codificar cada molécula de ARNm con una secuencia de código de barras única (por ejemplo, marcador molecular), un marcador celular y un sitio de PCR universal. Por ejemplo, las moléculas de ARN 302 pueden transcribirse inversamente para producir moléculas de ADNc marcadas 304, incluyendo una región de ADNc 306, mediante la hibridación (por ejemplo, hibridación estocástica) de un conjunto de códigos de barras (por ejemplo, códigos de barras estocásticos) 310 a 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 a diana, por ejemplo una región poli(dT) 312, una secuencia de código de barras o un marcador molecular 314, y una región de PCR universal 316.



En algunas realizaciones, el marcador celular puede incluir de 3 a 20 nucleótidos. En algunas realizaciones, la secuencia de código de barras (por ejemplo, marcador molecular) puede incluir de 3 a 20 nucleótidos. En algunas realizaciones, cada uno de la pluralidad de códigos de barras estocásticos comprende además una 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 en 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 de 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 de 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, de 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, de 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, de nucleótidos de longitud. Los marcadores universales pueden ser 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 en 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, de nucleótidos de longitud.

En algunas realizaciones, la región de 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, 7, 8, 9, 10, 20, 30, 40, 50, 60, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000. 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 ser como máximo, 10, 20, 40, 50, 70, 80, 90, 10<sup>2</sup>, 10<sup>3</sup>, 10<sup>4</sup>, 10<sup>5</sup>, 10<sup>6</sup>, 10<sup>7</sup>, 10<sup>8</sup>, 10<sup>9</sup>, 10<sup>10</sup>, 10<sup>11</sup>, 10<sup>12</sup>, 10<sup>13</sup>, 10<sup>14</sup>, 10<sup>15</sup>, 10<sup>20</sup>, o un número o un 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 1º grupo de cebadores de PCR y un 1º cebador de PCR universal. La agrupación es posible gracias a la región de marcador único 314. En particular, las moléculas de ADNc marcadas 304 pueden amplificarse para producir amplicones marcados por PCR anidada 322. La amplificación puede comprender una amplificación PCR multiplex. La amplificación puede comprender una amplificación PCR multiplex con 96 cebadores multiplex en un único volumen de reacción. En algunas realizaciones, la amplificación PCR multiplex puede utilizar, utilizar aproximadamente, utilizar por lo menos, o utilizar como máximo, 10, 20, 40, 50, 70, 80, 90, 10<sup>2</sup>, 10<sup>3</sup>, 10<sup>4</sup>, 10<sup>5</sup>, 10<sup>6</sup>, 10<sup>7</sup>, 10<sup>8</sup>, 10<sup>9</sup>, 10<sup>10</sup>, 10<sup>11</sup>, 10<sup>12</sup>, 10<sup>13</sup>, 10<sup>14</sup>, 10<sup>15</sup>, 10<sup>20</sup>, o un número o un intervalo entre cualquiera de estos valores, cebadores multiplex en un único volumen de reacción. La amplificación puede comprender un 1º grupo de cebadores de PCR 324 de 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 parte de ADNc 306' de la molécula de ADNc marcada 304. El cebador universal 328 puede hibridar con la región de 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 PCR en el paso 2 pueden amplificarse con un grupo de cebadores de PCR anidada y un 2º cebador de PCR universal. La PCR anidada puede minimizar el sesgo de amplificación por PCR. Por ejemplo, los amplicones marcados por PCR anidada 322 pueden amplificarse adicionalmente por PCR anidada. La PCR anidada puede comprender PCR multiplex con el conjunto de cebadores de PCR anidada 330 de cebadores de PCR anidada 332a-c y un 2º cebador de PCR universal 328' en un único volumen de reacción. El grupo de cebadores de PCR anidada 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, cebadores de PCR anidada 330 diferentes. Los cebadores de PCR anidada 332 pueden contener un adaptador 334 e hibridar con una región dentro de la parte 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 anidada 332 y el 2º cebador de PCR universal 328' pueden no contener los adaptadores 334 y 336. Los adaptadores 334 y 336 hibridan en la región de PCR universal 316 del amplicón marcado 322. 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 por PCR para secuenciación usando cebadores de amplificación de biblioteca. 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. Los uno o más cebadores 340 y 342 pueden ser cebadores de amplificación de PCR. Los uno o más cebadores 340 y 342 pueden ser cebadores de secuenciación. Los uno o más adaptadores 334 y 336 pueden usarse para la amplificación adicional de los amplicones marcados con adaptadores 338. Los 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 puedan secuenciarse en una reacción de secuenciación usando secuenciación de próxima generación (NGS).

#### Método y composición de identificación con códigos de barras de núcleos

En la presente se divulgan métodos, composiciones y kits para realizar tanto la captura de núcleos individuales como la identificación con códigos de barras (por ejemplo, en un solo paso), en donde solo el método forma parte de la invención. La secuenciación de núcleos individuales puede usarse para el análisis del transcriptoma de células individuales (u otros análisis ómicos o multiómicos de células individuales, como el análisis proteómico para determinar los perfiles de expresión proteica de células individuales) a la vez que se minimiza (por ejemplo, se elimina) la necesidad de realizar o preparar suspensiones de células individuales. La preparación de suspensiones de células individuales puede ser técnicamente difícil, por ejemplo, para ciertas preparaciones de muestras, como tejidos congelados. El método incluye un paso de aislamiento de núcleos. Por ejemplo, para separar los núcleos de otros componentes subcelulares puede usarse un anticuerpo específico de la envoltura nuclear, como el anti-LAMIN.

En algunas realizaciones, estos anticuerpos pueden estar asociados con (por ejemplo, conjugados con) un epítipo (como un epítipo fuerte, incluyendo biotina, digoxigenina o fluoresceína) o un código de barras molecular que representa ese anticuerpo particular y/o ambos. El epítipo asociado (por ejemplo, un epítipo conjugado) puede permitir un paso rápido de aislamiento de núcleos sin usar un paso de ultracentrifugación. Para realizar la indexación de muestras puede usarse el código de barras asociado (por ejemplo, un código de barras conjugado), por ejemplo, cuando se agrupan varias muestras para realizar un único procedimiento de preparación de bibliotecas de secuenciación de próxima generación (NGS). En algunas realizaciones, la indexación de muestras puede incluir la indexación de muestras en núcleos individuales. En algunas realizaciones, la indexación de muestras puede incluir la indexación de muestras en núcleos individuales y células individuales.

En algunas realizaciones, el código de barras del anticuerpo puede ser, o puede incluir, un código de barras de nucleótidos. Por ejemplo, un código de barras de nucleótidos puede tener por lo menos 10 pares de bases de longitud. El código de barras de nucleótidos puede ser, o puede incluir, ácido desoxirribonucleico (ADN), ácido ribonucleico (ARN), ácido nucleico bloqueado (LNA), ácido nucleico peptídico (PNA), u otro nucleótido sintético compatible con la transcripción inversa u otro paso de polimerización del ADN. El código de barras de nucleótidos puede incluir opcionalmente una cola 3' de poli(dA) para imitar los ARNm celulares endógenos y hacer que el código de barras sea compatible con la preparación de bibliotecas posterior. La información de la secuencia del código de barras se puede demultiplexar después de la secuenciación junto con cada núcleo para determinar de qué muestra se originó el núcleo. Los anticuerpos identificados con código de barras y sus usos, como la indexación de muestras de células usando anticuerpos identificados con códigos de barras, se han descrito en la Publicación de Solicitud de Patente de Estados Unidos N° 2018/0088112 y la Solicitud de Patente de Estados Unidos N° 15/937.713.

#### Método de identificación con códigos de barras de núcleos

Las FIGS. 4A-4B muestran una ilustración esquemática ejemplar no limitativa de un flujo de trabajo de identificación con códigos de barras de núcleos (por ejemplo, identificación con códigos de barras estocásticos). La lisis celular puede realizarse usando un homogeneizador (por ejemplo, un homogeneizador Dounce), un detergente o un método enzimático. Los núcleos pueden capturarse mediante, por ejemplo, un anticuerpo específico para una proteína de la envoltura nuclear (como LAMIN) conjugada con un epítipo (por ejemplo, un epítipo fuerte como biotina, digoxigenina (DIG) y fluoresceína) y, opcionalmente, un oligonucleótido de código de barras para marcar los núcleos. Puede usarse un anticuerpo secundario contra el epítipo para separar los núcleos para su purificación (por ejemplo, de los fragmentos citoplasmáticos), seguido de la captura de ARN/ADN mediante un ensayo de análisis de células individuales (como el ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ)).

En la presente se divulgan realizaciones de un método 400 para determinar el número de dianas en una pluralidad de células (como la célula 404 ilustrada en las FIGS. 4A-4B). El método 400 puede incluir lisis celular, separación y aislamiento de núcleos, lisis de núcleos e identificación con códigos de barras para análisis de células

individuales (por ejemplo, análisis de transcriptoma o proteoma de células individuales).

En algunas realizaciones, el método 400 comprende: aislar una pluralidad de núcleos de una pluralidad de células usando una composición de aislamiento de núcleos. La pluralidad de núcleos aislados puede incluir un núcleo 408 que comprende una envoltura nuclear 412 ilustrada en las FIGS. 4A-4B. Los contenidos (por ejemplo, contenidos de ARNm o contenidos de proteínas) de un núcleo aislado, como moléculas de ARNm 416 del núcleo 408, pueden analizarse usando un método de células individuales (como el ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ))). La composición de aislamiento de núcleos puede comprender un reactivo de unión a núcleos. El reactivo de unión a núcleos puede ser capaz de unirse específicamente a uno o más componentes de un núcleo (por ejemplo, una proteína de envoltura nuclear 420 en la envoltura nuclear 412 del núcleo celular 408 de la célula 404 ilustrada en las FIGS. 4A-4B).

En diferentes implementaciones el número de núcleos aislados puede ser diferente. En algunas realizaciones, el número de núcleos aislados puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 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, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000,  $10^7$ ,  $10^8$ ,  $10^9$ ,  $10^{10}$ ,  $10^{11}$ ,  $10^{12}$ , o un número o un intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de núcleos aislados puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 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, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000,  $10^7$ ,  $10^8$ ,  $10^9$ ,  $10^{10}$ ,  $10^{11}$ , o  $10^{12}$ .

En diferentes implementaciones el número de reactivos de unión a núcleos en la composición de aislamiento de núcleos puede ser diferente. En algunas realizaciones, el número de reactivos de unión a núcleos en la composición de aislamiento de núcleos puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de reactivos de unión a núcleos en la composición de aislamiento de núcleos puede ser como mínimo, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

Los reactivos de unión a núcleos pueden unirse a uno o más componentes del núcleo. En algunas realizaciones, el número de reactivos de unión a núcleos que se unen a un componente del núcleo puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de reactivos de unión a núcleos que se unen a un componente del núcleo puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

En diferentes implementaciones el número de componentes del núcleo a los que pueden unirse los reactivos de unión a núcleos en la composición de aislamiento de núcleos puede ser diferente. En algunas realizaciones, el número de los componentes del núcleo a los que pueden unirse los reactivos de unión a núcleos en la composición de aislamiento de núcleos puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de los componentes del núcleo a los que pueden unirse los reactivos de unión a núcleos de la composición de aislamiento de núcleos puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,

49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

El método 400 puede incluir identificar con códigos de barras de una pluralidad de dianas (por ejemplo, las moléculas de ARNm 416 ilustradas en las FIGS. 4A-4B) en la pluralidad de núcleos usando una pluralidad de códigos de barras para generar una pluralidad de dianas identificadas con códigos de barras. Cada uno de la pluralidad de códigos de barras puede comprender una secuencia de marcador molecular y una región de unión a diana, y las secuencias de marcadores moleculares de por lo menos dos códigos de barras de la pluralidad de códigos de barras pueden comprender secuencias diferentes. El método 400 puede incluir obtener datos de secuenciación de la pluralidad de dianas identificadas con códigos de barras. El método 400 puede incluir estimar el número de cada una de la pluralidad de dianas en la pluralidad de células usando las secuencias de marcadores moleculares de la pluralidad de códigos de barras en los datos de secuenciación. En las FIGS. 4A-4B, la célula 404 puede incluir una o más mitocondrias 424 con proteína mitocondrial 428.

En algunas realizaciones, aislar la pluralidad de núcleos comprende: poner en contacto la pluralidad de núcleos de la pluralidad de células con la composición de aislamiento de núcleos (por ejemplo, un anticuerpo específico de la proteína de envoltura nuclear 432a-432c ilustrado en las FIGS. 4A-4B) para generar núcleos unidos al reactivo de unión a núcleos. Aislar la pluralidad de núcleos puede comprender: aislar los núcleos unidos al reactivo de unión a núcleos usando un reactivo capaz de unirse específicamente al reactivo de unión a núcleos (por ejemplo, un reactivo de unión a epítomos 444 ilustrado en las FIGS. 4A-4B).

#### Componentes de un núcleo y un reactivo de unión a núcleos

En algunas realizaciones, el reactivo de unión a núcleos está asociado con un primer epítipo (por ejemplo, el epítipo 436 ilustrado en las FIGS. 4A-4B), y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un primer reactivo de unión a epítipo. El primer epítipo puede comprender biotina, hapteno o una combinación de los mismos. El hapteno puede comprender digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender un anticuerpo antihapteno. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender avidina, estreptavidina, neutravidina o una combinación de los mismos. En algunas realizaciones, el primer epítipo y/o el primer reactivo de unión a epítipo puede ser una fracción de afinidad como biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, o cualquier combinación de los mismos. El primer epítipo y el primer reactivo de unión a epítipo pueden ser miembros de un par de unión, por ejemplo, biotina/estreptavidina.

En diferentes implementaciones el número de primeros epítomos diferentes asociados con el reactivo de unión a núcleos puede ser diferente. En algunas realizaciones, el número de primeros epítomos diferentes asociados con el reactivo de unión a núcleo puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de primeros epítomos diferentes asociados con el reactivo de unión a núcleos puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100.

En diferentes implementaciones el número de moléculas de un primer epítipo asociado con el reactivo de unión a núcleos puede ser diferente. En algunas realizaciones, el número de moléculas de un primer epítipo asociado con el reactivo de unión a núcleos puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, o un número o un intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de moléculas de un primer epítipo asociado con el reactivo de unión a núcleos puede ser como mínimo, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100.

En algunas realizaciones, el reactivo de unión a núcleos comprende un anticuerpo primario (por ejemplo, un anticuerpo específico de proteína de envoltura nuclear 432a-432c ilustrado en las FIGS. 4A-4B) capaz de unirse específicamente a uno o más componentes del núcleo, y el reactivo capaz de unirse específicamente al reactivo de

unión a núcleos comprende un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario.

En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a componentes de la superficie de la envoltura nuclear (por ejemplo, un anticuerpo específico de la proteína de la envoltura nuclear), y el reactivo de unión a núcleos puede ser capaz de unirse específicamente a uno o más componentes de la superficie de la envoltura nuclear.

En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a carbohidratos. El reactivo de unión a carbohidratos puede comprender una proteína de unión a carbohidratos. La proteína de unión a carbohidratos puede comprender una lectina. La lectina puede comprender una lectina de unión a manosa, una lectina de unión a galactosa, una lectina de unión a N-acetilgalactosamina, una lectina de unión a N-acetilglucosamina, una lectina de unión a ácido N-acetilneuramínico, una lectina de unión a fucosa o una combinación de las mismas. La lectina puede comprender Concanavalina A (ConA), lectina de lenteja (LCH), lectina de campanilla de invierno (GNA), aglutinina de *Ricinus communis* (RCA), aglutinina de cacahuete (PNA), jacalina (AEL), lectina de veza peluda (VVL), aglutinina de germen de trigo (WGA), lectina del saúco (SNA), leucoaglutinina de *Maackia amurensis* (MAL), hemoaglutinina de *Maackia amurensis* (MAH), aglutinina de *Ulex europaeus* (UEA), lectina de *Aleuria aurantia* (AAL), o una combinación de las mismas. La lectina puede ser, o comprender, una aglutinina. La aglutinina puede ser, o comprender, aglutinina de germen de trigo (WGA). La proteína de unión a carbohidratos puede proceder o derivarse de un animal, una bacteria, un virus, un hongo o una combinación de los mismos. La proteína de unión a carbohidratos puede proceder o derivarse de una planta. La planta puede ser *Canavalia ensiformis*, *Lens culinaris*, *Galanthus nivalis*, *Ricinus communis*, *Arachis hypogaea*, *Artocarpus integrifolia*, *Vicia villosa*, *Triticum vulgare*, *Sambucus nigra*, *Maackia amurensis*, *Ulex europaeus*, *Aleuria aurantia*, o una combinación de las mismas.

En algunas realizaciones, los uno o más componentes del núcleo comprenden un azúcar, un oligosacárido, un polisacárido, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender un monosacárido, un disacárido, un poliol, un malto-oligosacárido, un no-malto-oligosacárido, un almidón, un polisacárido no amiláceo, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender glucosa, galactosa, fructosa, xilosa, sacarosa, lactosa, maltosa, trehalosa, sorbitol, manitol, maltodextrina, rafinosa, estaquiosa, fructooligosacárido, amilosa, amilopectina, almidón modificado, glucógeno, celulosa, hemicelulosa, pectina, hidrocoloide, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo pueden comprender un residuo de  $\alpha$ -D-manosilo, un residuo de  $\alpha$ -D-glucosilo, una estructura  $\alpha$ -manosídica ramificada de tipo de alta  $\alpha$ -manosa, una estructura  $\alpha$ -manosídica ramificada de tipo híbrido y N-Glicano de tipo complejo biantenarico, una región central fucosilada de N-Glicano de tipo complejo bi- y triantenarico, una estructura  $\alpha$  1-3 y  $\alpha$  1-6 enlazada de alta manosa, Gal $\beta$ 1-4GalNAc $\beta$ 1-R, Gal $\beta$ 1-3GalNAc $\alpha$ 1-Ser/Thr, (Sia)Gal $\beta$ 1-3GalNAc $\alpha$ 1-Ser/Thr, GalNAc $\alpha$ -Ser/Thr, GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc, Neu5Ac (ácido siálico), Neu5Ac $\alpha$ 2-6Gal(NAc)-R, Neu5Ac/Gc $\alpha$ 2,3Gal $\beta$ 1,4Glc(NAc), Neu5Ac/Gc $\alpha$ 2,3Gal $\beta$ 1,3(Neu5Ac $\alpha$ 2,6)GalNAc, Fuca1-2Gal-R, Fuca1-2Gal $\beta$ 1-4(Fuca1-3/4)Gal $\beta$ 1-4GlcNAc, R2-GlcNAc $\beta$ 1-4(Fuca1-6)GlcNAc-R1, un derivado de los mismos, o una combinación de los mismos. Los uno o más componentes del núcleo puede comprender una glicoproteína, un glicolípido, o una combinación de los mismos.

En algunas realizaciones, los uno o más componentes del núcleo comprenden Lamin, Emerin, Nesprin, Nurim, UNC-83, Klar, ZYG-12, Kms1p, UNC-84, Klaroid, SUN-1, Sad1p, LBR, MAN1, LAP1, LAP2, LINK, un complejo de poro nuclear, una parte de los mismos, o una combinación de los mismos. En distintas implementaciones el número de los uno o más componentes del núcleo puede ser diferente. En algunas realizaciones, el número de los uno o más componentes del núcleo puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de los uno o más componentes del núcleo puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

#### Partícula de aislamiento del núcleo

En algunas realizaciones, el reactivo de unión a núcleos se asocia con una partícula de aislamiento del núcleo. El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede asociarse con la partícula de aislamiento del núcleo (por ejemplo, una partícula 448 con una superficie ilustrada en las FIGS. 4A-4B). El reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede inmovilizarse, o inmovilizarse parcialmente, en la partícula de aislamiento del núcleo. Por ejemplo, el reactivo capaz de unirse específicamente al reactivo de unión a núcleos (como el reactivo de unión a epítomos 444) puede asociarse de manera reversible, no reversible, covalente, no covalente, o una combinación de las mismas, a la partícula de aislamiento del núcleo. Como otro ejemplo, el reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede estar insertado, parcialmente insertado, no

insertado, encerrado, parcialmente encerrado, no encerrado, o una combinación de los mismos, en la partícula de aislamiento del núcleo. En algunas realizaciones, el primer reactivo de unión a epítomos se asocia (por ejemplo, se inmoviliza, se inmoviliza parcialmente, se encierra, se encierra parcialmente) con la partícula de aislamiento del núcleo mediante un conector escindible. El conector escindible puede enlazar operativamente el primer reactivo de unión a epítomo con la partícula de aislamiento de núcleo. El conector escindible puede comprender un enlace químicamente escindible, un enlace fotoescindible, un conector lábil al ácido, un enlace sensible al calor, un enlace escindible enzimáticamente, o una combinación de los mismos. En algunas realizaciones de los métodos y composiciones divulgados en la presente, se proporcionan partículas nulas. En algunas realizaciones, las partículas nulas no comprenden una pluralidad de códigos de barras, una propiedad magnética, y/o un reactivo capaz de unirse específicamente al reactivo de unión a núcleos (por ejemplo, un primer reactivo de unión a epítomos). Las partículas nulas pueden parecerse a las partículas de aislamiento de núcleo en todos los aspectos distintos de la ausencia de una pluralidad de códigos de barras, una propiedad magnética, y/o un primer reactivo de unión a epítomo. En algunas realizaciones de los métodos y composiciones divulgados en la presente se proporciona un primer epítomo libre. El primer epítomo libre puede comprender la totalidad o una parte del primer epítomo. En algunas realizaciones, el primer epítomo libre puede unirse a un reactivo de unión al primer epítomo no unido.

En algunas realizaciones, la partícula de aislamiento del núcleo comprende una perla de aislamiento del núcleo. La partícula de aislamiento del núcleo 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 antifluorocromo, o cualquier combinación de las mismas. La partícula de aislamiento del núcleo 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. La partícula de aislamiento del núcleo puede ser alterable. La partícula de aislamiento del núcleo puede comprender una partícula de hidrogel alterable de aislamiento del núcleo.

En algunas realizaciones, aislar los núcleos unidos al reactivo de unión a núcleos puede comprender: aislar la partícula de aislamiento del núcleo mediante extracción magnética, centrifugación, o cualquier combinación de las mismas. En algunas realizaciones, la pluralidad de códigos de barras está asociada a la partícula de aislamiento del núcleo. En la partícula de aislamiento del núcleo puede inmovilizarse por lo menos un código de barras de la pluralidad de códigos de barras. En la partícula de aislamiento del núcleo puede estar parcialmente inmovilizado por lo menos un código de barras de la pluralidad de códigos de barras. En la partícula de aislamiento del núcleo puede estar encerrado por lo menos un código de barras de la pluralidad de códigos de barras. En la partícula de aislamiento del núcleo puede estar parcialmente encerrado por lo menos un código de barras de la pluralidad de códigos de barras. En la partícula de aislamiento del núcleo puede no estar encerrado por lo menos un código de barras de la pluralidad de códigos de barras. En la partícula de aislamiento del núcleo puede estar insertado por lo menos un código de barras de la pluralidad de códigos de barras. En la partícula de aislamiento del núcleo puede estar parcialmente insertado por lo menos un código de barras de la pluralidad de códigos de barras. En la partícula de aislamiento del núcleo puede no estar insertado por lo menos un código de barras de la pluralidad de códigos de barras.

#### Oligonucleótido de indexación de núcleos

En algunas realizaciones, el reactivo de unión a núcleos puede comprender un oligonucleótido de indexación de núcleos (por ejemplo, un código de barras 440a asociado con el anticuerpo específico de proteína de envoltura nuclear 432a ilustrado en las FIGS. 4A-4B). El oligonucleótido de indexación de núcleos puede comprender una secuencia de indexación de núcleos. El método 400 puede comprender: identificar con códigos de barras los oligonucleótidos de indexación de núcleos usando la pluralidad de códigos de barras (por ejemplo, para indexación de muestras de núcleos celulares) para generar una pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras; y obtener datos de secuenciación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. La identificación con códigos de barras de los oligonucleótidos de indexación de núcleos puede comprender: identificar estocásticamente con códigos de barras usando la pluralidad de códigos de barras para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. Identificar con códigos de barras los oligonucleótidos de indexación de núcleos usando la pluralidad de códigos de barras puede comprender: poner en contacto la pluralidad de códigos de barras con los oligonucleótidos de indexación de núcleos para generar códigos de barras hibridados con los oligonucleótidos de indexación de núcleos; y extender los códigos de barras hibridados con los oligonucleótidos de indexación de núcleos para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. Extender los códigos de barras puede comprender: extender los códigos de barras usando una ADN polimerasa para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras. Extender los códigos de barras puede comprender: extender los códigos de barras usando una transcriptasa inversa para generar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras.

En diferentes implementaciones el número de oligonucleótidos de indexación de núcleos asociados con (por

ejemplo, unidos a, conjugados con, etc.) el reactivo de unión a núcleos puede ser diferente. En algunas realizaciones, el número de oligonucleótidos de indexación de núcleos asociados con el reactivo de unión a núcleos puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de oligonucleótidos de indexación de núcleos asociado con el reactivo de unión a núcleos puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

Los oligonucleótidos de indexación de núcleos asociados con el reactivo de unión a núcleos pueden tener secuencias idénticas (o secuencias de indexación de núcleos idénticas) o secuencias diferentes (o secuencias de indexación de núcleos diferentes). En algunas realizaciones, el número de oligonucleótidos de indexación de núcleos asociados con el reactivo de unión a núcleos con una secuencia idéntica (o una secuencia de indexación de núcleos idéntica) puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de oligonucleótidos de indexación de núcleos asociados con el reactivo de unión a núcleos con una secuencia idéntica (o una secuencia de indexación de núcleos idéntica) puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000. En algunas realizaciones, el número de oligonucleótidos de indexación de núcleos asociados con el reactivo de unión a núcleos con diferentes secuencias (o diferentes secuencias de indexación de núcleos) puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de oligonucleótidos de indexación de núcleos asociados con el reactivo de unión a núcleos con diferentes secuencias (o diferentes secuencias de indexación de núcleos) puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

En algunas realizaciones, el método 400 puede comprender: amplificar la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras para producir una pluralidad de amplicones de indexación de núcleos identificados con códigos de barras. La amplificación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras puede comprender: amplificar, mediante reacción en cadena de la polimerasa (PCR), por lo menos una parte de la secuencia del marcador molecular y por lo menos una parte del oligonucleótido de indexación de núcleos. La obtención de los datos de secuenciación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras puede comprender: obtener los datos de secuenciación de la pluralidad de amplicones de indexación de núcleos identificados con códigos de barras. La obtención de los datos de secuenciación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras puede comprender: secuenciar la por lo menos una parte de la secuencia del marcador molecular y la por lo menos una parte del oligonucleótido de indexación de núcleos.

#### Partícula de identificación con códigos de barras

En algunas realizaciones, la pluralidad de códigos de barras está asociada con una partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar inmovilizado en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente inmovilizado en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar encerrado en la partícula de identificación con códigos de barras. Por lo menos un código de barras de la pluralidad de códigos de barras puede estar parcialmente encerrado en la partícula de identificación con códigos de barras. La partícula de identificación con códigos de barras puede ser alterable. La partícula de identificación con códigos de barras puede comprender una perla de identificación con código de barras. La partícula de identificación con códigos de barras 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 antifluorocromo, o cualquier combinación de las mismas. La partícula de identificación con códigos de barras 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, Sepharose, celulosa, nylon, silicona, y cualquier combinación de los mismos. La partícula de identificación con códigos de barras puede comprender una partícula de hidrogel alterable.

#### Códigos de barras e identificación con códigos de barras

En algunas realizaciones, identificar con códigos de barras la pluralidad de dianas usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras comprende: poner en contacto copias de las dianas con regiones de unión a diana de los códigos de barras; y transcribir inversamente la pluralidad de dianas usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras. El método puede comprender: antes de obtener los datos de secuenciación de la pluralidad de dianas identificadas con códigos de barras, amplificar la pluralidad de dianas identificadas con códigos de barras para generar una pluralidad de dianas identificadas con códigos de barras amplificadas. Amplificar las dianas identificadas con códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras amplificadas puede comprender: amplificar, mediante reacción en cadena de la polimerasa (PCR), las dianas identificadas con códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras amplificadas. El método 400 puede comprender: amplificar la pluralidad de dianas identificadas con códigos de barras amplificadas para generar una pluralidad de amplicones de dianas identificadas con códigos de barras. Amplificar la pluralidad de dianas amplificadas con códigos de barras puede comprender: amplificar la secuencia del marcador molecular y la secuencia de una diana de la pluralidad de dianas, o una parte de la misma, para generar la pluralidad de amplicones de dianas identificadas con códigos de barras. Amplificar la pluralidad de dianas identificadas con códigos de barras amplificadas puede comprender: amplificar, mediante reacción en cadena de la polimerasa (PCR), la pluralidad de dianas identificadas con códigos de barras amplificadas para generar la pluralidad de amplicones de dianas identificadas con códigos de barras. Identificar con códigos de barras la pluralidad de dianas de la célula usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras puede comprender: identificar estocásticamente con códigos de barras la pluralidad de dianas de la célula usando una pluralidad de códigos de barras estocásticos para generar una pluralidad de dianas identificadas estocásticamente con códigos de barras.

En algunas realizaciones, cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador celular, un sitio de unión para un cebador universal, o cualquier combinación de los mismos, y las secuencias de marcador celular de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden una secuencia idéntica. La región de unión a la diana puede comprender una región poli(dT). Por lo menos 100 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Por lo menos 1000 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Por lo menos 10000 secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias diferentes. Las secuencias de marcadores moleculares de la pluralidad de códigos de barras pueden comprender secuencias aleatorias. La región de unión a la diana puede comprender una secuencia específica de gen, una secuencia oligo(dT), un multímero aleatorio o cualquier combinación de los mismos.

#### Lisis celular

En algunas realizaciones, el método 400 comprende: antes de identificar con códigos de barras la pluralidad de dianas en la pluralidad de núcleos usando la pluralidad de códigos de barras para generar la pluralidad de dianas identificadas con códigos de barras, lisar la pluralidad de núcleos. En algunas realizaciones, el método 400 comprende: antes de aislar la pluralidad de núcleos de la pluralidad de células usando la composición de aislamiento de núcleos, lisar la pluralidad de células sin lisar los núcleos de la pluralidad de células.

#### Composición de identificación con códigos de barras de los núcleos

En la presente se divulga, pero no se reivindica, una composición de identificación con códigos de barras. En algunas realizaciones, la composición de identificación con códigos de barras comprende: una composición de aislamiento de núcleos que comprende un reactivo de unión a núcleos, en donde el reactivo de unión a núcleos es capaz de unirse específicamente a uno o más componentes de un núcleo; y una pluralidad de códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador molecular y una región de unión a diana, y en donde las secuencias de marcadores moleculares de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden secuencias diferentes. En algunas realizaciones, la composición comprende: un reactivo capaz de unirse específicamente al reactivo de unión a núcleos. El reactivo de unión a núcleos puede estar asociado con un primer epítipo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender un primer reactivo de unión a epítipos. En algunas realizaciones, el primer reactivo de unión a epítipo está asociado con (por ejemplo, inmovilizado, parcialmente inmovilizado, encerrado, parcialmente



encerrado) la partícula de aislamiento de núcleo a través de un conector escindible. El conector escindible puede enlazar operativamente el primer reactivo de unión a epítipo a la partícula de aislamiento del núcleo. El conector escindible puede comprender un conector escindible químicamente, un enlace fotoescindible, un conector lábil al ácido, un enlace sensible al calor, un enlace escindible enzimáticamente, o una combinación de los mismos. En algunas realizaciones, el reactivo de unión a núcleos comprende un anticuerpo primario capaz de unirse específicamente al uno o más componentes del núcleo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario. En algunas de tales realizaciones, dicho anticuerpo secundario está enlazado a la partícula de aislamiento del núcleo a través de un conector escindible. En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a carbohidratos. El reactivo de unión a carbohidratos puede comprender una proteína de unión a carbohidratos. En algunas realizaciones, el reactivo de unión a núcleos está asociado con una partícula de aislamiento del núcleo. En algunas realizaciones, la pluralidad de códigos de barras está asociada con la partícula de aislamiento del núcleo. En algunas realizaciones, el reactivo de unión a núcleos puede comprender un oligonucleótido de indexación de núcleos, y en donde el oligonucleótido de indexación de núcleos comprende una secuencia de indexación de núcleos. La pluralidad de códigos de barras puede estar asociada con una partícula de identificación con códigos de barras.

#### Método y composición de eliminación de orgánulos

También se divulgan en la presente, pero no se reivindican, métodos, composiciones y sistemas para realizar la eliminación o el aislamiento de orgánulos para el análisis del transcriptoma de células individuales (u otros análisis ómicos o multiómicos de células individuales, como el análisis proteómico para determinar los perfiles de expresión de proteínas de células individuales). En algunas realizaciones, para realizar un paso de eliminación de orgánulos puede usarse un anticuerpo, como un anticuerpo asociado a un epítipo. Por ejemplo, para un paso de eliminación de orgánulos puede usarse un anticuerpo conjugado con uno o más epítopos fuertes. Con un paso de eliminación de orgánulos, puede obtenerse ARN citoplasmático y/o nuclear para la preparación de bibliotecas NGS con menos contaminantes. Por ejemplo, para la preparación de bibliotecas de NGS puede obtenerse ARN citoplasmático y/o nuclear purificado o más limpio. Un paso de eliminación de orgánulos puede ser útil para la preparación de bibliotecas de células individuales porque los ARN específicos de orgánulos, como los ARN mitocondriales, pueden contribuir a las lecturas de secuenciación (por ejemplo, la mayoría de las lecturas de secuenciación). En algunas realizaciones, un paso de eliminación de orgánulos incluye la eliminación de mitocondrias con un anticuerpo específico de mitocondrias (u otro reactivo de unión específico de mitocondrias). Por ejemplo, un paso de eliminación de orgánulos puede incluir la eliminación de mitocondrias con un anticuerpo específico de mitocondrias (u otro reactivo de unión específico de mitocondrias) en un único paso mientras se realizan uno o más procesos posteriores en un grupo de ARN sin mitocondrias. Los procesos posteriores pueden incluir la síntesis de ADNc para el análisis del transcriptoma y una reacción de extensión del ácido nucleico para el análisis proteómico de células individuales. En algunas realizaciones, para el análisis de células individuales (por ejemplo, análisis del transcriptoma completo o análisis proteómico) puede reducirse (por ejemplo, eliminarse) contaminación mitocondrial.

En algunas realizaciones, la identificación con códigos de barras de núcleos y el aislamiento pueden producirse en un único paso. En algunas realizaciones, el agotamiento del contenido mitocondrial (por ejemplo, ARN) puede producirse al principio de la preparación de la biblioteca. En algunas realizaciones, la codificación con códigos de barras y el aislamiento de los núcleos y la eliminación de los orgánulos pueden producirse en un único paso o cerca en el tiempo.

Las FIGS. 5A-5B muestran una ilustración esquemática ejemplar no limitativo de un flujo de trabajo de identificación con códigos de barras (por ejemplo, identificación con códigos de barras estocásticos) con eliminación de orgánulos (por ejemplo, mitocondrias). El método 500 puede incluir lisis celular (por ejemplo, usando un tampón de lisis ligero para lisar una célula sin lisar uno o más orgánulos), unión y eliminación de orgánulos (como la unión y eliminación de mitocondrias), seguido de análisis de células individuales (por ejemplo, análisis de perfil de expresión de células individuales). La eliminación selectiva de orgánulos como las mitocondrias mediante, por ejemplo, un anticuerpo asociado con (por ejemplo, conjugado con) un epítipo (por ejemplo, un epítipo fuerte como biotina, fluoresceína y DIG) puede reducir (por ejemplo, minimizar o evitar) la contaminación de la biblioteca NGS. El paso de eliminación de orgánulos puede formar parte de un flujo de trabajo de células individuales, como el método de identificación con códigos de barras de núcleos 400 descrito con referencia a las FIGS. 4A-4B o el ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ)).

En algunas realizaciones, el método 500 comprende: antes de aislar, o cuando se asila, la pluralidad de núcleos de la pluralidad de células usando la composición de aislamiento de núcleos, permeabilizar la pluralidad de células (por ejemplo, permeabilizar y/o lisar selectivamente la membrana plasmática mientras se mantienen intactas las membranas de uno o más orgánulos); y agotar uno o más orgánulos (por ejemplo, mitocondrias 524) de la pluralidad de células usando una composición de captura de orgánulos que comprende un reactivo de unión a orgánulos (por ejemplo, un anticuerpo 532a ilustrado en las FIGS. 5A-5B), en donde el reactivo de unión a orgánulos es capaz de unirse específicamente a uno o más componentes de los uno o más orgánulos de la pluralidad de células. El agotamiento de los uno o más orgánulos puede comprender: poner en contacto los uno o más orgánulos de la pluralidad de células con la composición de captura de orgánulos para generar uno o más orgánulos unidos al reactivo

de unión a componentes de orgánulos. El agotamiento de los uno o más orgánulos puede comprender: agotar los uno o más orgánulos unidos al reactivo de unión a orgánulos usando un reactivo capaz de unirse específicamente al reactivo de unión a orgánulos. El reactivo de unión a orgánulos puede estar asociado a un segundo epítipo (por ejemplo, un epítipo 536), y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos (por ejemplo, un reactivo de unión a epítipo 544) puede comprender un segundo reactivo de unión a epítopos.

En diferentes implementaciones el número de reactivos de unión a orgánulos en una composición de captura de orgánulos puede ser diferente. En algunas realizaciones, el número de reactivos de unión a orgánulos en una composición de captura de orgánulos puede ser, o puede ser de 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de reactivos de unión a orgánulos en una composición de captura de orgánulos puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

En diferentes implementaciones el número de reactivos de unión a orgánulos en una composición de captura de orgánulos que se unen a un orgánulo puede ser diferente. En algunas realizaciones, el número de reactivos de unión a orgánulos en una composición de captura de orgánulos que se unen a un orgánulo puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos de cualesquiera de estos valores. En algunas realizaciones, el número de reactivos de unión a orgánulos en una composición de captura de orgánulos que se unen a un orgánulo puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

En diferentes implementaciones el número de orgánulos a los que pueden unirse los reactivos de unión a orgánulos en una composición de captura de orgánulos puede ser diferente. En algunas realizaciones, el número de orgánulos a los que pueden unirse los reactivos de unión a orgánulos en una composición de captura de orgánulos puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos de estos valores. En algunas realizaciones, el número de orgánulos a los que pueden unirse los reactivos de unión a orgánulos en una composición de captura de orgánulos puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

En diferentes implementaciones el número de diferentes orgánulos eliminados o agotados puede ser diferente. En algunas realizaciones, el número de diferentes orgánulos eliminados o agotados puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el número de orgánulos diferentes eliminados o agotados puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

En diferentes realizaciones el número de moléculas de un orgánulo eliminadas o agotadas puede ser diferente. En algunas realizaciones, el número de moléculas de un orgánulo eliminadas o agotadas puede ser, o puede ser 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, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, o un número o intervalo entre

dos cualesquiera de estos valores. En algunas realizaciones, el número de moléculas de un orgánulo eliminadas o agotadas puede ser por lo menos, o puede ser 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, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000.

En algunas realizaciones, el segundo epítipo puede comprender biotina, hapteno o una combinación de los mismos. El hapteno puede comprender digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos. El reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un anticuerpo antihapteno. El reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender avidina, estreptavidina, neutravidina o una combinación de los mismos. En algunas realizaciones, el segundo epítipo y/o el segundo reactivo de unión a epítopos puede ser una fracción de afinidad como biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, o cualquier combinación de los mismos. El segundo epítipo y el segundo reactivo de unión a epítopos pueden ser miembros de un par de unión, por ejemplo, biotina/estreptavidina. El reactivo de unión a orgánulos puede comprender un anticuerpo primario capaz de unirse específicamente a uno o más componentes de uno o más orgánulos de la pluralidad de células, y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario.

En algunas realizaciones, el reactivo de unión a orgánulos puede comprender un reactivo de unión a componentes de superficie de orgánulos, los uno o más componentes de los uno o más orgánulos pueden comprender uno o más componentes de superficie de orgánulos, y el reactivo de unión a orgánulos puede ser capaz de unirse específicamente a los uno o más componentes de superficie de orgánulos.

#### Partículas de captura de orgánulos

En algunas realizaciones, el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos (por ejemplo, el segundo reactivo de unión a epítopos) se asocia con una partícula de captura de orgánulos (por ejemplo, una partícula 548 con una superficie ilustrada en las FIGS. 5A-5B). El reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede inmovilizarse, o inmovilizarse parcialmente, en la partícula de captura de orgánulos. Por ejemplo, el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede estar asociado reversiblemente, no reversiblemente, covalentemente, no covalentemente, o una combinación de los mismos, con la partícula de captura de orgánulos e. Como otro ejemplo, el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede estar insertado, parcialmente insertado, no insertado, encerrado, parcialmente encerrado, no encerrado, o una combinación de los mismos, en la partícula de captura de orgánulos.

En algunas realizaciones, la partícula de captura de orgánulos comprende una perla de captura de orgánulos. La partícula de captura de orgánulos puede comprender una perla de sefrosa, 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 antifluorocromo, o cualquier combinación de las mismas. La partícula de captura de orgánulos 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, Sepharose, celulosa, nylon, silicona, y cualquier combinación de los mismos.

En algunas realizaciones, el agotamiento de los orgánulos de la pluralidad de células usando la composición de captura de orgánulos puede comprender: el agotamiento de una o más partículas de captura de orgánulos mediante extracción magnética, centrifugación, o cualquier combinación de las mismas. Los orgánulos pueden comprender mitocondrias de la pluralidad de células. Los uno o más componentes de los uno o más orgánulos de la pluralidad de células puede comprender: ABCD3, ESR2, NOS3, ALB, HIF1A, NR3C1, ATP5A1, HK1, PGR, CASQ1, HSPA1A, PHB, CLTC, HSPD1, PLN, COX4I1, IFM1, SOD1, CPS1, LGALS3, TP53, Citocromo C Oxidasa, MAPT, TP5B, ERN1, MT-CO1, VDAC1, o una combinación de los mismos.

#### Composición de eliminación de orgánulos

En algunas realizaciones, la composición comprende una composición de captura de orgánulos que comprende un reactivo de unión a orgánulos, en donde el reactivo de unión a orgánulos es capaz de unirse específicamente a uno o más componentes de uno o más orgánulos. La composición puede comprender un reactivo capaz de unirse específicamente al reactivo de unión a orgánulos. El reactivo de unión a orgánulos puede asociarse con un segundo epítipo, y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos puede comprender un segundo reactivo de unión a epítipo. En algunas realizaciones, el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos está asociado a una partícula de captura de orgánulos. En algunas

realizaciones, el reactivo de unión a orgánulos comprende un reactivo de unión a componentes de superficie de orgánulos, los uno o más componentes de los uno o más orgánulos pueden comprender uno o más componentes de superficie de orgánulos, y el reactivo de unión a orgánulos puede ser capaz de unirse específicamente a los uno o más componentes de superficie de orgánulos.

## Secuenciación

En algunas realizaciones, la estimación del número de diferentes dianas identificadas con códigos de barras (por ejemplo, dianas identificadas con códigos de barras estocástico) puede comprender la determinación de las secuencias de las dianas marcadas, el marcador espacial, el marcador molecular, el marcador de la muestra, el marcador celular, o cualquier producto de los mismos (por ejemplo, amplicones marcados, o moléculas de ADNc marcadas). Una diana amplificada puede someterse a secuenciación. La determinación de la secuencia de una diana identificada con código de barras (por ejemplo, una diana identificada estocásticamente con código de barras) o cualquier producto de la misma puede comprender la realización de una reacción de secuenciación para determinar la secuencia de por lo menos una parte de un marcador de muestra, un marcador espacial, un marcador celular, un marcador molecular, por lo menos una parte de la diana marcada (por ejemplo, una diana marcada estocásticamente), un complemento de la misma, un complemento inverso de la misma, o cualquier combinación de los mismos.

La determinación de la secuencia de una diana identificada con código de barras o de una diana identificada estocásticamente con código de barras (por ejemplo, ácido nucleico amplificado, ácido nucleico marcado, copia de ADNc de un ácido nucleico marcado, etc.) puede realizarse usando una variedad de métodos de secuenciación que incluyen, pero no se limitan a, secuenciación por hibridación (SBH), secuenciación por ligadura (SBL), secuenciación por adición cuantitativa de nucleótidos fluorescentes incrementales (QIFNAS), ligadura y escisión por pasos, transferencia de energía por resonancia de fluorescencia (FRET), balizas moleculares, digestión con sonda informadora TaqMan, pirosecuenciación, secuenciación in situ fluorescente (FISSEQ), perlas FISSEQ, secuenciación oscilante, secuenciación multiplex, secuenciación de colonias polimerizadas (POLONY); secuenciación en círculo rodante en nanorejilla (ROLONY), ensayos de ligadura de oligo alelos específicos (por ejemplo, ensayo de ligadura de oligonucleótidos (OLA), OLA de molécula de plantilla única usando una sonda lineal ligada y una lectura de amplificación de círculo rodante (RCA), sondas de candado ligadas u OLA de molécula de plantilla única usando una sonda de candado circular ligada y una lectura de amplificación de círculo rodante (RCA) y similares.

En algunas realizaciones, la determinación de la secuencia de la diana identificada con código de barras (por ejemplo, diana identificada estocásticamente con código de barras) o cualquier producto de la misma comprende secuenciación de extremos pareados, secuenciación de nanoporos, secuenciación de alto rendimiento, secuenciación de escopeta, secuenciación de colorante-terminador, secuenciación de ADN de cebadores múltiples, caminata de cebadores, secuenciación dideoxilada de Sanger, secuenciación de Maxim-Gilbert, pirosecuenciación, secuenciación de molécula única verdadera, o cualquier combinación de las mismas. Alternativamente, la secuencia de la diana identificada con código de barras o cualquier producto de la misma puede determinarse mediante microscopía electrónica o una matriz de transistores de efecto de campo químico-sensibles (chemFET).

Pueden utilizarse métodos de secuenciación de alto rendimiento, como la secuenciación cíclica de matrices usando plataformas como Roche 454, Illumina Solexa, ABI-SOLiD, ION Torrent, Complete Genomics, Pacific Bioscience, Helicos o la plataforma Polonator. En algunas realizaciones, la secuenciación puede incluir la secuenciación MiSeq. En algunas realizaciones, la secuenciación puede incluir la secuenciación HiSeq.

Las dianas marcadas (por ejemplo, dianas marcadas estocásticamente) pueden comprender ácidos nucleicos que representan desde aproximadamente el 0,01% de los genes del genoma de un organismo hasta aproximadamente el 100% de los genes del genoma de un organismo. Por ejemplo, puede secuenciarse desde aproximadamente el 0,01% de los genes del genoma de un organismo hasta aproximadamente el 100% de los genes del genoma de un organismo usando una región complementaria diana que comprende una pluralidad de multímeros mediante la captura de los genes que contienen una secuencia complementaria de la muestra. En algunas realizaciones, las dianas identificadas con códigos de barras comprenden ácidos nucleicos que representan desde aproximadamente el 0,01% de los transcritos del transcriptoma de un organismo hasta aproximadamente el 100% de los transcritos del transcriptoma de un organismo. Por ejemplo, puede secuenciarse desde aproximadamente el 0,501% de los transcritos del transcriptoma de un organismo hasta aproximadamente el 100% de los transcritos del transcriptoma de un organismo usando una región complementaria diana que comprende una cola poli(T) mediante la captura de los ARNm de la muestra.

Determinar las secuencias de los marcadores espaciales y los marcadores moleculares de la pluralidad de códigos de barras (por ejemplo, códigos de barras estocásticos) puede incluir secuenciar el 0,00001%, 0,0001%, 0,001%, 0,01%, 0,1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 99%, 100%, o un número o un intervalo entre dos cualesquiera de estos valores, de la pluralidad de códigos de barras. Determinar las secuencias de los marcadores de la pluralidad de códigos de barras, por ejemplo los marcadores de muestra, los marcadores espaciales y los marcadores moleculares, puede incluir la secuenciación de 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ ,  $10^{10}$ ,  $10^{11}$ ,  $10^{12}$ ,  $10^{13}$ ,  $10^{14}$ ,  $10^{15}$ ,  $10^{16}$ ,  $10^{17}$ ,  $10^{18}$ ,  $10^{19}$ ,  $10^{20}$ ,

o un número o un intervalo entre dos cualesquiera de estos valores, de la pluralidad de códigos de barras. La secuenciación de una parte o de la totalidad de la pluralidad de códigos de barras puede incluir la generación de secuencias con longitudes de lectura de, de aproximadamente, de por lo menos, o de como máximo, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, o un número o un intervalo entre dos cualesquiera de estos valores, de nucleótidos o bases.

La secuenciación puede comprender secuenciar por lo menos, o por lo menos aproximadamente, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, o más nucleótidos o pares de bases de las dianas identificadas con códigos de barras. Por ejemplo, la secuenciación puede comprender la generación de datos de secuenciación con secuencias con longitudes de lectura de 50, 75, o 100, o más nucleótidos realizando amplificación por reacción en cadena de la polimerasa (PCR) de la pluralidad de dianas identificadas con códigos de barras. La secuenciación puede comprender la secuenciación de por lo menos, o por lo menos aproximadamente, 200, 300, 400, 500, 600, 700, 800, 900, 1.000, o más nucleótidos o pares de bases de las dianas identificadas con códigos de barras. La secuenciación puede comprender la secuenciación de por lo menos, o por lo menos aproximadamente, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, o 10000, o más nucleótidos o pares de bases de las dianas identificadas con códigos de barras.

La secuenciación puede comprender por lo menos aproximadamente 200, 300, 400, 500, 600, 700, 800, 900, 1.000, o más lecturas de secuenciación por ejecución. En algunas realizaciones, la secuenciación comprende secuenciar por lo menos, o por lo menos aproximadamente, 1500, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, o 10000, o más lecturas de secuenciación por ejecución. La secuenciación puede comprender menos de o igual a aproximadamente 1.600.000.000 lecturas de secuenciación por serie. La secuenciación puede comprender menos o igual a aproximadamente 200.000.000 de lecturas de secuenciación por serie.

#### Muestras

En algunas realizaciones, la pluralidad de dianas puede estar comprendida en una o más muestras. Una muestra puede comprender una o más células, o ácidos nucleicos de una o más células. Una muestra puede ser una sola célula o ácidos nucleicos de una sola célula. La una o más células pueden ser de uno o más tipos celulares. Por lo menos uno de los uno o más tipos celulares puede ser célula cerebral, célula cardiaca, célula cancerosa, célula tumoral circulante, célula de órgano, célula epitelial, célula metastásica, célula benigna, célula primaria, célula circulatoria, o cualquier combinación de las mismas.

Una muestra para uso en el método de la divulgación puede comprender una o más células. Una muestra puede referirse a una o más células. En algunas realizaciones, la pluralidad de células puede incluir uno o más tipos de células. Por lo menos uno de los uno o más tipos celulares puede ser célula cerebral, célula cardiaca, célula cancerosa, célula tumoral circulante, célula de órgano, célula epitelial, célula metastásica, célula benigna, célula primaria, célula circulatoria, o cualquier combinación de las mismas. En algunas realizaciones, las células son células cancerosas extirpadas de un tejido canceroso, por ejemplo, cáncer de mama, cáncer de pulmón, cáncer de colon, cáncer de próstata, cáncer de ovario, cáncer de páncreas, cáncer cerebral, melanoma y cánceres de piel no melanoma, y similares. En algunas realizaciones, las células se derivan de un cáncer pero se recogen de un fluido corporal (por ejemplo, células tumorales circulantes). Ejemplos no limitativos de cánceres pueden incluir adenoma, adenocarcinoma, carcinoma de células escamosas, carcinoma de células basales, carcinoma de células pequeñas, carcinoma indiferenciado de células grandes, condrosarcoma y fibrosarcoma. La muestra puede incluir un tejido, una monocapa celular, células fijadas, una sección de tejido o cualquier combinación de los mismos. La muestra puede incluir una muestra biológica, una muestra clínica, una muestra ambiental, un fluido biológico, un tejido o una célula de un sujeto. La muestra puede obtenerse de un humano, un mamífero, un perro, una rata, un ratón, un pez, una mosca, un gusano, una planta, un hongo, una bacteria, un virus, un vertebrado o un invertebrado.

En algunas realizaciones, las células son células que han sido infectadas con virus y contienen oligonucleótidos virales. En algunas realizaciones, la infección viral puede estar provocada por un virus como virus de ADN (por ejemplo, parvovirus) de cadena sencilla (cadena + o "sentido"), o virus de ARN (por ejemplo, reovirus) de cadena doble. En algunas realizaciones, las células son bacterias. Estas incluyen bacterias grampositivas o gramnegativas. En algunas realizaciones, las células son hongos. En algunas realizaciones, las células son protozoos u otros parásitos.

Como se usa en la presente, el término "célula" puede referirse a una o más células. En algunas realizaciones, las células son células normales, por ejemplo, células humanas en diferentes etapas de desarrollo, o células humanas de diferentes órganos o tipos de tejido. En algunas realizaciones, las células son células no humanas, por ejemplo, otros tipos de células de mamíferos (por ejemplo, ratón, rata, cerdo, perro, vaca o caballo). En algunas realizaciones, las células son otros tipos de células animales o vegetales. En otras realizaciones, las células pueden ser cualquier célula procariota o eucariota.

En algunas realizaciones las células se clasifican antes de asociar una célula con una perla. Por ejemplo, las células pueden clasificarse mediante clasificación celular activada por fluorescencia o clasificación celular activada magnéticamente, o más generalmente mediante citometría de flujo. Las células pueden filtrarse por tamaño. En

algunas realizaciones, un retentado contiene las células que se van a asociar a la perla. En algunas realizaciones, el flujo contiene las células que se van a asociar a la perla.

Una muestra puede referirse a una pluralidad de células. La muestra puede referirse a una monocapa de células. La muestra puede referirse a una sección delgada (por ejemplo, sección delgada de tejido). La muestra puede referirse a una colección sólida o semisólida de células que pueden colocarse en una dimensión en una matriz.

#### Métodos de eliminación de especies indeseables de ácidos nucleicos

El análisis de células individuales, como el análisis del transcriptoma completo (WTA) de células individuales, puede verse obstaculizado por la presencia o alta abundancia de dianas o especies de dianas no de interés (como ácido nucleico o especies de ácido nucleico no de interés, también denominadas en la presente ácido nucleico o especies de ácido nucleico indeseables o no deseadas). En muchas muestras, los ácidos nucleicos derivados de orgánulos no nucleares (por ejemplo, ARNm mitocondrial y ARN ribosómico) pueden estar presentes en gran abundancia y/o ser indeseables. Por ejemplo, para una muestra de ADNc heterogénea, la PCR se realiza habitualmente en exceso de ciclos para amplificar adecuadamente los expresadores bajos; en esos escenarios, el perfil de expresión génica nativa suele estar sesgado por los productos de PCR dominantes de los altos expresadores. Un método para corregir este sesgo en los productos de PCR es la Indexación Molecular; sin embargo, los altos expresadores, como los ARNm mitocondriales, suelen dominar la ejecución de la secuenciación con poca contribución a la interpretación experimental, lo que encarece el coste de la secuenciación para el recuento de la Indexación Molecular. Los esfuerzos anteriores para aumentar la abundancia relativa de especies de baja abundancia en una muestra de ácido nucleico han incluido métodos de normalización de bibliotecas. Algunas realizaciones divulgadas en la presente proporcionan métodos para eliminar especies de ácidos nucleicos indeseables (por ejemplo, ácidos nucleicos derivados de orgánulos no nucleares, ARNm, ARNr) de la pluralidad de moléculas de ácidos nucleicos.

En algunas realizaciones, los métodos divulgados en la presente comprenden proporcionar una muestra que comprenda una pluralidad de moléculas diana de ácido nucleico. La muestra puede, en algunas realizaciones, comprender una o más especies de ácidos nucleicos indeseables. Los expertos en la técnica apreciarían que la pluralidad de moléculas diana de ácido nucleico y/o las especies de ácido nucleico indeseables pueden comprender una variedad de moléculas diana de ácido nucleico. Por ejemplo, las moléculas diana de ácido nucleico y/o las especies de ácido nucleico indeseables pueden comprender moléculas de ADN, moléculas de ARN, moléculas de ADN genómico, moléculas de ADNc, moléculas de ARNm, moléculas de ARNr, ADNmt, moléculas de ARNip, o cualquier combinación de las mismas. La molécula diana de ácido nucleico puede ser de cadena doble o de cadena sencilla. En algunas realizaciones, la pluralidad de moléculas diana de ácido nucleico puede comprender ácidos nucleicos derivados de uno o más orgánulos no nucleares. En algunas realizaciones, la pluralidad de moléculas diana de ácido nucleico puede comprender moléculas de ARN polIA. En algunas realizaciones, la pluralidad de moléculas diana de ácido nucleico comprende por lo menos 100, por lo menos 1.000, por lo menos 10.000, por lo menos 20.000, por lo menos 30.000, por lo menos 40.000, por lo menos 50.000, por lo menos 100.000, por lo menos 1.000.000, o más especies de ácido nucleico.

La muestra puede, por ejemplo, comprender una pluralidad de moléculas diana de ácido nucleico, y una o más especies de ácido nucleico indeseables (por ejemplo, ácidos nucleicos derivados de orgánulos no nucleares). La una o más especies de ácidos nucleicos indeseables pueden estar presentes en la muestra en cantidades diferentes. Por ejemplo, la una o más especies de ácido nucleico indeseables pueden representar aproximadamente, o por lo menos aproximadamente, el 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, o más, o el 100%, o un intervalo entre dos de estos valores, del contenido de ácido nucleico de la muestra. En algunas realizaciones, las una o más especies de ácido nucleico indeseables representan por lo menos el 10% del contenido de ácido nucleico de la muestra. En algunas realizaciones, la especie de ácido nucleico indeseable representa por lo menos el 30% del contenido de ácido nucleico de la muestra. En algunas realizaciones, la especie de ácido nucleico indeseable representa por lo menos el 50% del contenido de ácido nucleico de la muestra. Las especies de ácido nucleico indeseables pueden ser, o comprender, varios tipos de moléculas de ácido nucleico. Por ejemplo, por lo menos una de las especies de ácido nucleico indeseables puede ser ARNm de proteína ribosómica, ARNm mitocondrial, ADN genómico, secuencia intrónica, secuencia de alta abundancia o una combinación de las mismas. En algunas realizaciones, las especies de ácidos nucleicos indeseables son, o comprenden, especies de ARNm. Las especies de ARNm pueden ser, por ejemplo, una o más de las especies de ARNm de genes constitutivos, especies de ARNm de genes ribosómicos, especies de ARNm de genes mitocondriales, especies de ARNm de genes de alta abundancia, o cualquier combinación de las mismas. En algunas realizaciones, las especies de ácido nucleico indeseables son, o comprenden, especies de ADN. Las especies de ADN pueden ser, por ejemplo, una o más de las especies de ADN de genes constitutivos, especies de ADN de genes ribosómicos, especies de ADN de genes mitocondriales, especies de ADN de genes de alta abundancia, o cualquier combinación de las mismas. En algunas realizaciones, las especies de ácido nucleico indeseables son, o comprenden, especies de ADN y especies de ARN.

En algunas realizaciones, los métodos de aislamiento de núcleos y/o agotamiento de orgánulos descritos en la presente pueden reducir significativamente la abundancia de una o más especies de ácidos nucleicos indeseables

(por ejemplo, ácidos nucleicos derivados de orgánulos no nucleares) en comparación con la pluralidad de moléculas diana de ácidos nucleicos en la muestra. En algunas realizaciones, los métodos y composiciones divulgados en la presente pueden reducir la abundancia de una o más especies de ácidos nucleicos indeseables en una muestra. Por ejemplo, los métodos y composiciones divulgados en la presente pueden reducir la abundancia de por lo menos 1, 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 50, por lo menos 100, por lo menos 200, por lo menos 500, por lo menos 1.000, o más, especies de ácidos nucleicos indeseables en la muestra. En algunas realizaciones, los métodos y composiciones divulgados en la presente pueden reducir la abundancia en por lo menos un 10%, por lo menos un 20%, por lo menos un 30%, por lo menos un 40%, por lo menos un 50%, por lo menos un 60%, por lo menos un 70%, por lo menos un 80%, por lo menos un 90%, por lo menos un 95%, por lo menos un 99%, o un 100% de cada una de las una o más especies de ácidos nucleicos indeseables en la muestra. En algunas realizaciones, la abundancia de la una o más especies de ácidos nucleicos indeseables puede reducirse en por lo menos un 10%, por lo menos un 20%, por lo menos un 30%, por lo menos un 40%, por lo menos un 55%, por lo menos un 50%, por lo menos un 65%, por lo menos un 60%, por lo menos un 65%, por lo menos un 70%, por lo menos un 75%, por lo menos un 80%, por lo menos un 85%, por lo menos un 90%, por lo menos un 95%, por lo menos un 99%, o más, o un intervalo entre dos cualesquiera de estos valores, en comparación con la abundancia de por lo menos una de las moléculas diana de ácido nucleico en la muestra de partida antes del agotamiento de orgánulos y/o el aislamiento de núcleos como se describe en la presente.

Se apreciará que en algunas realizaciones, los métodos y composiciones divulgados en la presente pueden reducir la abundancia de especies de ácido nucleico indeseables sin reducir significativamente el número de secuencias de marcadores moleculares diferentes asociadas con las otras moléculas diana de ácido nucleico. Por ejemplo, los métodos y composiciones divulgados en la presente pueden reducir la abundancia de especies de ácidos nucleicos indeseables mientras retienen por lo menos el 10%, por lo menos el 20%, por lo menos el 30%, por lo menos el 40%, por lo menos el 50%, por lo menos el 60%, por lo menos el 70%, por lo menos el 80%, por lo menos el 90%, por lo menos el 95%, o el 100% de las diferentes secuencias de marcadores moleculares asociadas con las otras moléculas diana de ácidos nucleicos. En algunas realizaciones, los métodos y composiciones divulgados en la presente pueden reducir la abundancia de especies de ácidos nucleicos indeseables en por lo menos un 10%, por lo menos un 20%, por lo menos un 30%, por lo menos un 40%, por lo menos un 50%, por lo menos un 60%, por lo menos un 70%, por lo menos un 80%, por lo menos un 90%, por lo menos un 95%, o por lo menos un 99% mientras conserva por lo menos el 10%, por lo menos el 20%, por lo menos el 30%, por lo menos el 40%, por lo menos el 50%, por lo menos el 60%, por lo menos el 70%, por lo menos el 80%, por lo menos el 90%, por lo menos el 95%, o el 100% de las diferentes secuencias moleculares de marcado asociadas con las otras moléculas diana de ácido nucleico. En algunas realizaciones, la reducción de la abundancia de especies de ácidos nucleicos indeseables no reduce significativamente el número de secuencias moleculares marcadas diferentes asociadas con las otras moléculas diana de ácido nucleico.

En algunas realizaciones, después de usar los métodos de aislamiento de núcleos y/o agotamiento de orgánulos descritos en la presente, las lecturas de secuenciación para las especies de ácido nucleico indeseables son menos del 50%, menos del 40%, menos del 30%, menos del 20%, menos del 10%, menos del 5%, o menos, del total de lecturas de secuenciación. En algunas realizaciones, las lecturas de secuenciación para las especies de ácidos nucleicos indeseables son de menos del 40% de las lecturas de secuenciación totales. En algunas realizaciones, las lecturas de secuenciación para las especies de ácidos nucleicos indeseables son de menos del 30% de las lecturas de secuenciación totales. En algunas realizaciones, las lecturas de secuenciación para las especies de ácidos nucleicos indeseables son de menos del 20% de las lecturas de secuenciación totales. En algunas realizaciones, las lecturas de secuenciación para las especies de ácidos nucleicos indeseables son de menos del 10% de las lecturas de secuenciación totales. En algunas realizaciones, después de usar los métodos de aislamiento de núcleos y/o agotamiento de orgánulos descritos en la presente, las lecturas de secuenciación para las especies de ácido nucleico indeseables se reducen a menos del 60%, menos del 50%, menos del 40%, menos del 30%, menos del 20%, menos del 10%, menos del 5% de las lecturas de secuenciación para el ácido nucleico indeseable sin usar los métodos de aislamiento de núcleos y/o agotamiento de orgánulos descritos en la presente. En algunas realizaciones, después de usar los métodos de aislamiento de núcleos y/o agotamiento de orgánulos descritos en la presente, las lecturas de secuenciación para las especies de ácido nucleico indeseables se reducen a, o aproximadamente, el 60%, 50%, 40%, 30%, 20%, 10%, 5%, 2%, 1%, 0,5%, o un intervalo entre dos cualesquiera de estos valores, de las lecturas de secuenciación para el ácido nucleico indeseable sin usar los métodos de aislamiento de núcleos y/o agotamiento de orgánulos descritos en la presente.

En algunas realizaciones, los métodos y composiciones divulgados en la presente pueden mejorar la eficiencia de secuenciación disminuyendo la relación lecturas de secuenciación: marcador molecular de una especie de ácido nucleico indeseable y/o aumentando la relación lecturas de secuenciación: marcador molecular de una molécula diana de ácido nucleico. Por ejemplo, la relación de lecturas de secuenciación: marcador molecular para una especie de ácido nucleico indeseable puede ser de menos de 20, de menos de 15, de menos de 10, de menos de 9, de menos de 8, de menos de 7, de menos de 6, de menos de 5, de menos de 4, de menos de 3, de menos de 2 o de menos de 1. En algunas realizaciones, la proporción entre las lecturas de secuenciación y el marcador molecular para una especie de ácido nucleico indeseable es 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, o un intervalo entre dos cualesquiera de estos valores.



## **EJEMPLOS**

Algunos aspectos de las realizaciones analizadas anteriormente se describen con más detalle en los siguientes ejemplos, que no se pretende en modo alguno que limiten el alcance de la presente divulgación.

### **Ejemplo 1**

#### **Captura de núcleos y flujo de trabajo de identificación con códigos de barras**

Este ejemplo demuestra un flujo de trabajo para la captura de núcleos e identificación con códigos de barras. Las FIGS. 6A-6F son una ilustración esquemática de un flujo de trabajo ejemplar no limitativo para el aislamiento de núcleos y la realización de perfiles de células individuales. El flujo de trabajo puede comprender lisis celular (por ejemplo, usando un tampón de lisis ligero para lisar la membrana plasmática sin lisar la membrana nuclear y/o uno o más orgánulos). El flujo de trabajo puede incluir el agotamiento selectivo de uno o más orgánulos (por ejemplo, mitocondrias). El flujo de trabajo puede incluir la asociación de cada núcleo con un oligonucleótido de indexación de núcleos que comprende una secuencia de indexación de núcleos (por ejemplo, para la indexación de muestras). El flujo de trabajo puede incluir aislar los núcleos de uno o más componentes celulares. El flujo de trabajo puede incluir dividir núcleos individuales en divisiones y la realización de lisis de núcleos. El flujo de trabajo puede incluir identificar con códigos de barras (por ejemplo, identificación con códigos de barras estocásticos) de moléculas de ácido nucleico diana y/u oligonucleótidos de indexación de núcleos para el análisis de células individuales (por ejemplo, ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ)), solución 3' de células individuales Chromium™ (10X Genomics (San Francisco, CA))).

#### **Composición de la muestra**

La muestra 601 puede comprender una pluralidad de células 602. Una muestra para su uso en el método de la divulgación puede ser una para la que es técnicamente difícil o imposible generar una suspensión de células individuales (por ejemplo, células congeladas, células fijadas, muestras de tejido, muestras de tumor, tejidos epiteliales, células fijadas en formalina e insertadas en parafina, y combinaciones de las mismas). Una célula puede comprender una membrana plasmática 604 y un núcleo celular 606. El núcleo celular puede comprender una envoltura nuclear 608 y una o más moléculas diana de ácido nucleico 612. La envoltura nuclear puede comprender uno o más componentes de superficie de la envoltura nuclear 610, como una proteína de envoltura nuclear (por ejemplo, una lámina). El núcleo puede comprender una o más moléculas de ácido nucleico diana (por ejemplo, ARN mensajero (ARNm), microARN (miARN), ARN no codificante largo (ARNnc largo o ARNlnc), ARN que interactúa con Piwi (ARNip)). La célula puede comprender uno o más orgánulos (por ejemplo, ribosomas, mitocondrias) como se describe en la presente, y los uno o más orgánulos pueden comprender uno o más componentes de la superficie de orgánulos. Las mitocondrias 616 pueden comprender un componente de superficie de orgánulos (por ejemplo, componente de superficie mitocondrial 618). La célula también puede comprender componentes celulares extranucleares 614. Los componentes celulares extranucleares pueden, por ejemplo, comprender componentes citoplasmáticos extranucleares no deseados (por ejemplo, moléculas que son innecesarias y/o perjudiciales para la realización de perfiles de expresión de células individuales). Los componentes celulares extranucleares pueden incluir fragmentos citoplasmáticos. Los componentes celulares extranucleares pueden incluir orgánulos. Los componentes celulares extranucleares pueden incluir mitocondrias. Los componentes celulares extranucleares pueden incluir ribosomas. Los componentes celulares extranucleares pueden incluir retículos endoplasmáticos rugosos. Los componentes celulares extranucleares pueden comprender uno o más ácidos nucleicos no deseados (por ejemplo, ARN ribosómico, ADN mitocondrial, ARN mitocondrial). Los componentes celulares extranucleares pueden comprender una o más macromoléculas que interferirían con los métodos de análisis de expresión de células individuales divulgados en la presente. Los componentes celulares extranucleares 614 pueden, por ejemplo, comprender mitocondrias, peroxisomas, citosol, vesículas, lisosomas, membranas plasmáticas, cloroplastos, matrices mitocondriales internas, membranas mitocondriales internas, espacios intermembrana, membranas mitocondriales externas, vesículas secretoras, retículos endoplásmicos lisos, retículos endoplásmicos rugosos, cuerpos de Golgi, fagosomas, endosomas, exosomas, membranas plasmáticas, microtúbulos, microfilamentos, filamentos intermedios, filopodios, remolinos, lamelipodios, sarcómeros, contactos focales, podosomas, ribosomas, microsomas, balsas lipídicas, paredes celulares y combinaciones de los mismos.

#### **Lisis de la membrana plasmática**

El flujo de trabajo puede comprender lisis de membrana plasmática (paso 600a) de la población de células 602 de la muestra 601. La lisis de la membrana plasmática puede generar un lisado que comprende una pluralidad de una población de núcleos, orgánulos y componentes celulares extranucleares (por ejemplo, ribosomas) como se representa en la FIG. 6A. El método puede comprender lisar la pluralidad de células sin lisar la envoltura nuclear. El método puede comprender lisar la pluralidad de células sin lisar las membranas de los orgánulos. El método puede comprender lisar la pluralidad de células y lisar las membranas de los orgánulos sin lisar la envoltura nuclear. La lisis de la membrana plasmática puede realizarse usando un homogeneizador (por ejemplo, un homogeneizador Dounce), un detergente o un método enzimático.



Contacto con el reactivo o reactivos de unión a componentes de la superficie de orgánulos

El flujo de trabajo puede comprender poner en contacto el lisado con uno o más reactivos de unión a componentes de superficie de orgánulos (paso 600b). Un segundo epítipo 622 puede asociarse (por ejemplo, inmovilizarse, inmovilizarse parcialmente, encerrarse, encerrarse parcialmente) con el reactivo de unión a componentes de superficie de orgánulos 620 (por ejemplo, reactivo de unión a orgánulos). El segundo epítipo 622 puede comprender un epítipo fuerte (por ejemplo, biotina, fluoresceína y/o DIG) como se describe en la presente. El reactivo de unión a componentes de la superficie de orgánulos 620 puede ser capaz de unirse específicamente a uno o más componentes de uno o más orgánulos (por ejemplo, mitocondrias) como, por ejemplo, el componente de superficie mitocondrial 618 (como se representa en la FIG. 6B). El contacto de uno o más orgánulos de la pluralidad de células con el reactivo de unión a componentes de superficie de orgánulos puede generar uno o más orgánulos unidos al reactivo de unión a componentes de orgánulos. El reactivo de unión a componentes de superficie de orgánulos 620 puede comprender un anticuerpo específico de mitocondria (u otro reactivo de unión específico de mitocondria).

Contacto con partículas de captura de orgánulos

El flujo de trabajo puede comprender poner en contacto el lisado con una o más partículas de captura de orgánulos (paso 600c). La partícula de captura de orgánulos 630 puede comprender un reactivo capaz de unirse específicamente al reactivo de unión a orgánulos (por ejemplo, un segundo reactivo de unión a epítipo 632 como se representa en la FIG. 6B) como se describe en la presente. El segundo reactivo de unión a epítipo 632 puede estar asociado (por ejemplo, inmovilizado, parcialmente inmovilizado, encerrado, parcialmente encerrado) con la partícula de captura de orgánulos 630. En algunas realizaciones, el reactivo de unión a componentes de la superficie de orgánulos 620 no comprende un segundo epítipo. En algunas de tales realizaciones, el reactivo de unión a componentes de la superficie de orgánulos puede comprender un anticuerpo primario capaz de unirse específicamente a uno o más componentes de uno o más orgánulos (por ejemplo, mitocondrias) y el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos comprende un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario. El paso 600c puede generar uno o más orgánulos unidos a las partículas de captura de orgánulos mediante el reactivo de unión a componentes de superficie de orgánulos.

Eliminación de partículas de captura de orgánulos

El flujo de trabajo puede comprender la eliminación de partículas de captura de orgánulos (paso 600d). La eliminación de las partículas de captura de orgánulos puede producirse por eliminación magnética, centrifugación, filtración o cualquier combinación de las mismas. El paso 600d puede generar un lisado desprovisto de uno o más orgánulos.

Contacto con reactivo o reactivos de unión a núcleos

El flujo de trabajo puede comprender poner en contacto el lisado con uno o más reactivos de unión a núcleos (paso 600e). El reactivo de unión a núcleos 640 puede comprender un primer epítipo 642. El primer epítipo 642 puede, por ejemplo, comprender un epítipo fuerte (por ejemplo, biotina, fluoresceína, y/o DIG) como se describe en la presente. El reactivo de unión a núcleos 640 puede ser capaz de unirse específicamente a uno o más componentes de los núcleos (por ejemplo, proteínas de la envoltura nuclear) como, por ejemplo, componentes de la superficie de la envoltura nuclear 610 (como se representa en la FIG. 6C). En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a carbohidratos (por ejemplo, una proteína de unión a carbohidratos, una lectina). El contacto de los núcleos con el reactivo de unión a núcleos puede generar núcleos unidos al reactivo de unión a núcleos. El reactivo de unión a núcleos 640 puede comprender el oligonucleótido de indexación de núcleos 644. El oligonucleótido de indexación de núcleos 644 puede comprender una secuencia de indexación de núcleos (por ejemplo, para indexación de muestras). El oligonucleótido de indexación de núcleos 644 puede permanecer asociado con los núcleos a lo largo de todo el flujo de trabajo. En algunas realizaciones, el flujo de trabajo puede comprender la agrupación de núcleos de diferentes muestras una vez que los núcleos están unidos por un reactivo de unión de núcleos que comprende un oligonucleótido de indexación de núcleos.

Contacto con partículas de aislamiento del núcleo y partículas nulas

El flujo de trabajo puede comprender poner en contacto el lisado con partículas de aislamiento de núcleo y partículas nulas (paso 600f). La partícula de aislamiento de núcleo 650 puede comprender un reactivo capaz de unirse específicamente al reactivo de unión a núcleo (por ejemplo, un primer reactivo de unión a epítipo 652 como se representa en la FIG. 6D) como se describe en la presente. El primer reactivo de unión a epítipo 652 puede estar asociado con (por ejemplo, inmovilizado, parcialmente inmovilizado, encerrado, parcialmente encerrado) la partícula de aislamiento del núcleo. El reactivo de unión a núcleos puede comprender un anticuerpo primario capaz de unirse específicamente a uno o más componentes del núcleo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario. Una pluralidad de códigos de barras 654 puede estar asociada (por ejemplo, inmovilizada, parcialmente inmovilizada,

encerrada, parcialmente encerrada) con la partícula de aislamiento del núcleo.

Como se representa en la FIG. 6D, el paso 600f puede generar núcleos unidos a partículas de aislamiento de núcleos a través del reactivo capaz de unirse específicamente al reactivo de unión de núcleos. En algunas realizaciones, una población de núcleos dentro de una muestra puede separarse de componentes celulares extranucleares de acuerdo con los métodos proporcionados en la presente. En algunas realizaciones, puede separarse individualmente unos de otros (por ejemplo, aislarse) una población de núcleos dentro de una muestra de acuerdo con los métodos proporcionados en la presente. Algunas realizaciones de las composiciones y métodos divulgados en la presente pueden usarse opcionalmente para reducir o eliminar la aglutinación de partículas de aislamiento de núcleos. Algunas realizaciones de los métodos y composiciones proporcionados en la presente pueden usarse para, por ejemplo, (1) reducir o eliminar la unión de una partícula de aislamiento de núcleo individual a dos o más núcleos y/o (2) reducir o eliminar la unión de un núcleo individual a múltiples partículas de aislamiento de núcleo.

En algunas realizaciones, el porcentaje de partículas de aislamiento de núcleo unidas a un único núcleo o sin núcleo después del contacto puede ser, o ser aproximadamente, del 0,000000001%, 0,00000001%, 0,0000001%, 0,000001%, 0,00001%, 0,0001%, 0,001%, 0,01%, 0,1%, 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%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el porcentaje de partículas de aislamiento de núcleo unidas a un único núcleo o a ningún núcleo después del contacto puede ser como mínimo, o como máximo, del 0,000000001%, 0,00000001%, 0,0000001%, 0,000001%, 0,00001%, 0,0001%, 0,001%, 0,01%, 0,1%, 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%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%.

En algunas realizaciones, el porcentaje de núcleos unidos a una única partícula de aislamiento de núcleo después del contacto puede ser, o ser aproximadamente, del 0,000000001%, 0,00000001%, 0,0000001%, 0,000001%, 0,00001%, 0,0001%, 0,001%, 0,01%, 0,1%, 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%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el porcentaje de núcleos unidos a una única partícula de aislamiento de núcleo después del contacto puede ser como mínimo, o como máximo, del 0,000000001%, 0,00000001%, 0,0000001%, 0,000001%, 0,00001%, 0,0001%, 0,001%, 0,01%, 0,1%, 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%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%.

En algunas realizaciones, la unión de núcleos individuales a partículas de aislamiento de núcleos puede seguir una distribución de Poisson. En algunas realizaciones, la unión de núcleos individuales a partículas de aislamiento de núcleos puede seguir una distribución no de Poisson. La probabilidad de que dos núcleos distintos de una muestra puedan unirse a la misma partícula de aislamiento de núcleo puede ser, o ser por lo menos, de  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$ , o  $10^{-1}$  o más. La probabilidad de que dos núcleos distintos de una muestra puedan unirse a la misma partícula de aislamiento del núcleo puede ser como máximo de  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$ , o  $10^{-1}$  o más. La probabilidad de que un único núcleo pueda unirse simultáneamente a dos o más partículas de aislamiento de núcleo distintas puede ser, o ser como mínimo, de  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$ , o  $10^{-1}$  o más. La probabilidad de que un único núcleo pueda unirse simultáneamente a dos o más partículas de aislamiento de núcleo distintas puede ser como máximo de  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$ ,  $10^{-3}$ ,  $10^{-2}$ , o  $10^{-1}$  o más.

En algunas realizaciones, el flujo de trabajo comprende poner en contacto el lisado con partículas de aislamiento de núcleos 650 y partículas nulas 656. En algunas realizaciones, las partículas nulas 656 no comprenden una pluralidad de códigos de barras, una propiedad magnética, y/o un reactivo capaz de unirse específicamente al reactivo de unión a núcleos (por ejemplo, un primer reactivo de unión a epítomos). Las partículas nulas pueden parecerse a las partículas de aislamiento de núcleos en todos los aspectos distintos de la ausencia de una pluralidad de códigos de barras, una propiedad magnética, y/o un reactivo capaz de unirse específicamente al reactivo de unión a núcleos (por ejemplo, un primer reactivo de unión a epítomos). En algunas realizaciones, el lisado se pone en contacto

con partículas nulas antes de ponerlo en contacto con partículas de aislamiento de núcleos. En algunas realizaciones, el lisado se pone en contacto con partículas nulas y partículas de aislamiento del núcleo simultáneamente. El contacto con partículas nulas puede (1) reducir o eliminar la unión de una partícula de aislamiento de núcleo individual a dos o más núcleos y/o (2) reducir o eliminar la unión de un núcleo individual a múltiples partículas de aislamiento de núcleo.

En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas varía de 1:100 a 100:1. En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas es como máximo de 10:1. En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas es como máximo de 100:1. En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas es como máximo de 1:1000. En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas es de por lo menos 1:10. En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas es de por lo menos 1:100. En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas es de por lo menos 1:1000.

En algunas realizaciones, la proporción de las partículas de aislamiento del núcleo con las partículas nulas puede ser, o ser aproximadamente, de 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 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, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 1:10000, o un número o intervalo comprendido entre dos cualesquiera de los valores. En algunas realizaciones, la proporción entre las partículas de aislamiento del núcleo y las partículas nulas puede ser como mínimo, o como máximo, 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 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, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000 o 1:10000.

Algunas realizaciones de los métodos y composiciones proporcionados en la presente emplean citometría de flujo para (1) reducir o eliminar la unión de una partícula de aislamiento de núcleo individual con dos o más núcleos y/o (2) reducir o eliminar la unión de un núcleo individual a múltiples partículas de aislamiento de núcleo. Por ejemplo, una serie de gotitas que comprenden cada una un único núcleo y una serie de gotitas que comprenden cada una un único núcleo pueden fusionarse para crear una serie de gotitas que comprenden cada una un único núcleo y una única partícula de aislamiento de núcleo. Después de la unión de los únicos núcleos a la partícula de aislamiento de núcleo único dentro de una gotita, cada gotita de la serie de gotitas puede fusionarse con otra serie de gotitas que comprenden un primer epítipo libre como se describe en la presente. Tras la saturación de los sitios de unión del primer epítipo en la partícula de aislamiento del núcleo, la pluralidad de gotitas puede agruparse y someterse a los métodos posteriores de la divulgación.

#### Contacto con el primer epítipo libre

El flujo de trabajo puede comprender poner en contacto el lisado con el primer epítipo libre (paso 600g). El primer epítipo libre 660 puede comprender la totalidad o una parte del primer epítipo 642. El primer epítipo libre puede unirse al primer reactivo de unión a epítipo no unido (por ejemplo, un primer reactivo de unión a epítipo 652 como se representa en la FIG. 6E) asociado con partículas de aislamiento de núcleo. Después del paso 600g, la mayoría o todos los sitios de unión del reactivo de unión a núcleos pueden saturarse (por ejemplo, unirse por un primer epítipo libre o un primer epítipo de un reactivo de unión a núcleos), reduciendo o impidiendo de este modo cualquier unión adicional de partículas de aislamiento de núcleo a núcleos. El contacto con el primer epítipo libre puede evitar la aglutinación de las partículas de aislamiento de núcleos cuando se separan de las partículas nulas en sentido descendente en el flujo de trabajo.

En algunas realizaciones, la proporción de primer epítipo libre con las partículas de aislamiento de núcleo puede ser, o ser aproximadamente, de 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 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, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000, 1:10000, o un número o intervalo entre dos cualesquiera de los valores. En algunas realizaciones, la proporción entre el primer epítipo libre y las partículas

de aislamiento del núcleo puede ser como mínimo, o como máximo, de 1:1, 1:1.1, 1:1.2, 1:1.3, 1:1.4, 1:1.5, 1:1.6, 1:1.7, 1:1.8, 1:1.9, 1:2, 1:2.5, 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, 1:21, 1:22, 1:23, 1:24, 1:25, 1:26, 1:27, 1:28, 1:29, 1:30, 1:31, 1:32, 1:33, 1:34, 1:35, 1:36, 1:37, 1:38, 1:39, 1:40, 1:41, 1:42, 1:43, 1:44, 1:45, 1:46, 1:47, 1:48, 1:49, 1:50, 1:51, 1:52, 1:53, 1:54, 1:55, 1:56, 1:57, 1:58, 1:59, 1:60, 1:61, 1:62, 1:63, 1:64, 1:65, 1:66, 1:67, 1:68, 1:69, 1:70, 1:71, 1:72, 1:73, 1:74, 1:75, 1:76, 1:77, 1:78, 1:79, 1:80, 1:81, 1:82, 1:83, 1:84, 1:85, 1:86, 1:87, 1:88, 1:89, 1:90, 1:91, 1:92, 1:93, 1:94, 1:95, 1:96, 1:97, 1:98, 1:99, 1:100, 1:200, 1:300, 1:400, 1:500, 1:600, 1:700, 1:800, 1:900, 1:1000, 1:2000, 1:3000, 1:4000, 1:5000, 1:6000, 1:7000, 1:8000, 1:9000 o 1:10000.

En algunas realizaciones, el porcentaje de partículas de aislamiento del núcleo que comprenden uno o más reactivos de unión al primer epítipo no unidos después del contacto con el primer epítipo libre puede ser, o ser aproximadamente, del 0,000000001%, 0,00000001%, 0,0000001%, 0,000001%, 0,00001%, 0,0001%, 0,001%, 0,01%, 0,1%, 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%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, o un número o intervalo entre dos cualesquiera de estos valores. En algunas realizaciones, el porcentaje de partículas de aislamiento del núcleo que comprenden uno o más reactivos de unión al primer epítipo no unidos después del contacto con el primer epítipo libre puede ser por lo menos, o ser como máximo, del 0,000000001%, 0,00000001%, 0,0000001%, 0,000001%, 0,0001%, 0,001%, 0,01%, 0,1%, 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%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%.

#### Aislamiento de partículas de aislamiento del núcleo

El flujo de trabajo puede comprender el aislamiento de partículas de aislamiento de núcleo (paso 600h). El aislamiento de partículas de aislamiento de núcleo (y los núcleos unidos a las mismas) puede comprender el aislamiento de la partícula de aislamiento de núcleo mediante extracción magnética, filtración, centrifugación, o cualquier combinación de las mismas. El paso 600h puede generar una pluralidad de partículas de aislamiento de núcleo aisladas (cada una unida a uno o cero núcleos) desprovistas de orgánulos y/o componentes celulares extranucleares.

#### División, lisis nuclear e identificación con códigos de barras

El flujo de trabajo puede comprender división, lisis nuclear e identificación con códigos de barras (paso 600i). La pluralidad de núcleos y las partículas de aislamiento de núcleo unidas a los mismos pueden dividirse en una pluralidad de divisiones. En algunas realizaciones, la pluralidad de núcleos se separa de las partículas de aislamiento del núcleo antes de la división de dichos núcleos en una pluralidad de divisiones. Una división puede ser un micropocillo, una gotita o una emulsión. Una división de la pluralidad de divisiones puede comprender un solo núcleo de la pluralidad de núcleos y su partícula de aislamiento de núcleo asociada. Una división de la pluralidad de divisiones puede comprender un único núcleo de la pluralidad de núcleos y una partícula de identificación con códigos de barras. Una división de la pluralidad de divisiones puede comprender un único núcleo de la pluralidad de núcleos, su partícula de aislamiento de núcleo asociada y una partícula de identificación con códigos de barras. En algunas realizaciones, las dimensiones del micropocillo (por ejemplo, la profundidad del micropocillo) se eligen para optimizar la eficiencia de captura de los núcleos y de la partícula de aislamiento del núcleo, a la vez que se proporciona un intercambio eficiente de los tampones de ensayo y otros reactivos contenidos dentro de los pocillos. Una vez divididos, los núcleos pueden lisarse de acuerdo con los métodos proporcionados en la presente. La identificación con códigos de barras puede comprender el uso de la pluralidad de códigos de barras de una partícula de aislamiento del núcleo y/o una partícula de identificación con códigos de barras para generar una pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras y/o dianas identificadas con códigos de barras de acuerdo con los métodos proporcionados en la presente. Los productos del Paso 600i (por ejemplo, oligonucleótidos de indexación de núcleos identificados con códigos de barras y/o dianas identificadas con códigos de barras) pueden someterse a los métodos posteriores proporcionados en la presente para generar un perfil de expresión de una célula individual.

En algunas realizaciones, uno o más pasos del flujo de trabajo comprenden centrifugación. En algunas realizaciones, uno o más pasos del flujo de trabajo no comprenden centrifugación. En algunas realizaciones, ninguno de los pasos del flujo de trabajo comprende centrifugación. En algunas realizaciones, el flujo de trabajo no comprende uno o más del paso 600a, paso 600b, paso 600c, paso 600d, paso 600e, paso 600f, paso 600g, paso 600h y/o paso 600i. Por ejemplo, en algunas realizaciones, el método no comprende el contacto con reactivo de unión de componentes de superficie de orgánulos (paso 600b), el contacto con partículas de captura de orgánulos (paso 600c), y/o la eliminación de partículas de captura de orgánulos (paso 600d). Opcionalmente, el paso 600a comprende lisar

la pluralidad de células y lisar las membranas de los orgánulos sin lisar la envoltura nuclear, y por lo tanto no se realizan el paso 600b, el paso 600c, y/o el paso 600d. En algunas realizaciones, la contaminación de orgánulos y/o la contaminación de componentes celulares extranucleares (por ejemplo, contaminación mitocondrial) puede reducirse (por ejemplo, eliminarse) para el análisis de células individuales (por ejemplo, análisis del transcriptoma completo o análisis proteómico) sin realizar el paso 600b, el paso 600c, y/o el paso 600d del flujo de trabajo. En algunas realizaciones, uno o más del paso 600a, paso 600b, paso 600c, paso 600d, paso 600e, paso 600f, paso 600g, paso 600h y/o paso 600i del flujo de trabajo se realizan simultáneamente. Por ejemplo, el contacto con el reactivo de unión del componente de superficie del orgánulo (paso 600b) y el contacto con las partículas de captura del orgánulo (paso 600c) pueden producirse a la vez. En algunas de tales realizaciones, el reactivo de unión al componente de superficie de orgánulos se asocia con una partícula de captura de orgánulos. El reactivo de unión al componente de superficie de orgánulos puede inmovilizarse, o inmovilizarse parcialmente, en la partícula de captura de orgánulos. Por ejemplo, el reactivo de unión del componente de superficie de orgánulos puede estar asociado de manera reversible, no reversible, covalente, no covalente, o una combinación de las mismas, con la partícula de captura de orgánulos. Como otro ejemplo, el reactivo de unión al componente de superficie de orgánulos puede estar insertado, parcialmente insertado, no insertado, encerrado, parcialmente encerrado, no encerrado, o una combinación de los mismos, en la partícula de captura de orgánulos.

En algunas realizaciones, el reactivo de unión a núcleos no comprende un primer epítipo. En algunas de tales realizaciones, el reactivo de unión a núcleos puede comprender un anticuerpo primario capaz de unirse específicamente a uno o más componentes de los núcleos (por ejemplo, proteínas de la envoltura nuclear) y las partículas de aislamiento de núcleos están asociadas con un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario. En algunas de tales realizaciones, el anticuerpo primario libre (en lugar del primer epítipo libre) puede ponerse en contacto con el lisado antes del aislamiento de las partículas de aislamiento del núcleo.

## Ejemplo 2

### Flujo de trabajo de captura de núcleos e identificación con códigos de barras

Este ejemplo demuestra un flujo de trabajo para la captura de núcleos y la identificación con códigos de barras. Las FIGS. 7A-7C son una ilustración esquemática de un flujo de trabajo ejemplar no limitativo para el aislamiento de núcleos y la realización de perfiles de células individuales. El flujo de trabajo puede comprender lisis celular (por ejemplo, usando un tampón de lisis ligero para lisar la membrana plasmática sin lisar la membrana nuclear y/o uno o más orgánulos). El flujo de trabajo puede incluir asociar cada núcleo con un oligonucleótido de indexación de núcleos que comprende una secuencia de indexación de núcleos (por ejemplo, para la indexación de muestras). El flujo de trabajo puede incluir aislar los núcleos de uno o más componentes y/u orgánulos celulares. El flujo de trabajo puede incluir la división de núcleos individuales en divisiones y la realización de lisis de núcleos. El flujo de trabajo puede incluir la identificación con códigos de barras (por ejemplo, identificación con códigos de barras estocásticos) de moléculas de ácido nucleico diana y/u oligonucleótidos de indexación de núcleos para el análisis de células individuales (por ejemplo, ensayo Rhapsody™ (Becton, Dickinson and Company (Franklin Lakes, NJ)), solución de cromo™ Single Cell 3' (10X Genomics (San Francisco, CA))).

### Composición de la muestra

La muestra 701 puede comprender una pluralidad de células 702. Una muestra para uso en el método de la divulgación puede ser una para la cual es técnicamente difícil o imposible generar una suspensión de células individuales (células congeladas, células fijadas, muestras de tejido, muestras de tumor, tejidos epiteliales, células fijadas en formalina e incluidas en parafina, y combinaciones de las mismas). Una célula puede comprender una membrana plasmática 704 y un núcleo celular 706. El núcleo celular puede comprender una envoltura nuclear 708 y una o más moléculas diana de ácido nucleico 712. La envoltura nuclear puede comprender uno o más componentes de superficie de envoltura nuclear 710, como una proteína de envoltura nuclear (por ejemplo, una lámina). El núcleo puede comprender una o más moléculas de ácido nucleico diana (por ejemplo, ARN mensajero (ARNm), microARN (miARN), ARN no codificante largo (ARNnc largo o ARNlnc), ARN que interactúa con Piwi (ARNip)). La célula puede comprender uno o más orgánulos (por ejemplo, ribosomas, mitocondrias) como se describe en la presente, y los uno o más orgánulos pueden comprender uno o más componentes de superficie de orgánulos. Las mitocondrias 716 pueden comprender un componente de superficie de orgánulos (por ejemplo, componente de superficie mitocondrial 718). La célula también puede comprender componentes celulares extranucleares 714. Los componentes celulares extranucleares pueden comprender componentes citoplasmáticos extranucleares no deseados (por ejemplo, moléculas que son innecesarias y/o perjudiciales para la realización del perfil de la expresión de células individuales). Los componentes celulares extranucleares pueden incluir fragmentos citoplasmáticos. Los componentes celulares extranucleares pueden incluir orgánulos. Los componentes celulares extranucleares pueden incluir mitocondrias. Los componentes celulares extranucleares pueden incluir ribosomas. Los componentes celulares extranucleares pueden incluir retículos endoplasmáticos rugosos. Los componentes celulares extranucleares pueden comprender uno o más ácidos nucleicos no deseados (por ejemplo, ARN ribosómico, ADN mitocondrial, ARN mitocondrial). Los componentes celulares extranucleares pueden comprender una o más macromoléculas que interferirían con los métodos de análisis de expresión de células individuales divulgados en la presente. Los componentes celulares extranucleares 714 pueden

comprender mitocondrias, peroxisomas, citosol, vesículas, lisosomas, membranas plasmáticas, cloroplastos, matrices mitocondriales internas, membranas mitocondriales internas, espacios intermembrana, membranas mitocondriales externas, vesículas secretoras, retículos endoplásmicos lisos, retículos endoplásmicos rugosos, cuerpos de Golgi, fagosomas, endosomas, exosomas, membranas plasmáticas, microtúbulos, microfilamentos, filamentos intermedios, filopodios, remolinos, lamelipodios, sarcómeros, contactos focales, podosomas, ribosomas, microsomas, balsas lipídicas, paredes celulares y combinaciones de los mismos.

#### Lisis de la membrana plasmática

El flujo de trabajo puede comprender la lisis de la membrana plasmática (paso 700a) de la población de células 702 de la muestra 701. La lisis de la membrana plasmática puede generar un lisado que comprende una pluralidad de una población de núcleos, orgánulos y componentes celulares extranucleares (por ejemplo, ribosomas) como se representa en la FIG. 7A. El método puede comprender lisar la pluralidad de células sin lisar la envoltura nuclear. El método puede comprender lisar la pluralidad de células sin lisar las membranas de los orgánulos. El método puede comprender lisar la pluralidad de células y lisar las membranas de los orgánulos sin lisar la envoltura nuclear. La lisis de la membrana plasmática puede realizarse usando un homogeneizador (por ejemplo, un homogeneizador Dounce), un detergente o un método enzimático. En algunas realizaciones, el flujo de trabajo comprende poner en contacto el lisado con el reactivo de unión de componentes de superficie de orgánulos, poner en contacto el lisado con partículas de captura de orgánulos y/o eliminar dichas partículas de captura de orgánulos (por ejemplo, paso 600b, paso 600c y/o paso 600d del Ejemplo 1) antes de realizar el paso 700b.

#### Contacto con reactivo o reactivos de unión a núcleos

El flujo de trabajo puede comprender poner en contacto el lisado con uno o más reactivos de unión a núcleos (paso 700b). El reactivo de unión a núcleos 740 puede comprender un primer epítipo 742. El primer epítipo 742 puede comprender un epítipo fuerte (por ejemplo, biotina, fluoresceína y/o DIG) como se describe en la presente. El reactivo de unión a núcleos 740 puede ser capaz de unirse específicamente a uno o más componentes de los núcleos (por ejemplo, proteínas de la envoltura nuclear) como, por ejemplo, componentes de la superficie de la envoltura nuclear 710 (como se representa en la FIG. 7B). En algunas realizaciones, el reactivo de unión a núcleos comprende un reactivo de unión a carbohidratos (por ejemplo, una proteína de unión a carbohidratos, una lectina). Poner en contacto los núcleos con el reactivo de unión a núcleos puede generar núcleos unidos al reactivo de unión a núcleos. El reactivo de unión a núcleos 740 puede comprender el oligonucleótido 744 de indexación de núcleos. El oligonucleótido de indexación de núcleos 744 puede comprender una secuencia de indexación de núcleos (por ejemplo, para indexación de muestras). El oligonucleótido de indexación de núcleos 744 puede permanecer asociado con los núcleos a lo largo de todo el flujo de trabajo. En algunas realizaciones, el flujo de trabajo puede comprender la agrupación de núcleos de diferentes muestras una vez que los núcleos se han unido por un reactivo de unión de núcleos que comprende un oligonucleótido de indexación de núcleos.

#### Contacto con las partículas de aislamiento del núcleo

El flujo de trabajo puede comprender poner en contacto el lisado con una o más partículas de aislamiento de núcleo (paso 700c). La partícula de aislamiento de núcleo 750 puede comprender un reactivo capaz de unirse específicamente al reactivo de unión a núcleos (por ejemplo, un primer reactivo de unión a epítipo 752 como se representa en la FIG. 7B) como se describe en la presente. El primer reactivo de unión a epítipo 752 puede asociarse con (por ejemplo, inmovilizarse, inmovilizarse parcialmente, encerrarse, encerrarse parcialmente) la partícula de aislamiento de núcleo a través de un conector escindible 780. El conector escindible 780 puede asociarse con la partícula de aislamiento de núcleo a través de un conector escindible 780. El conector escindible 780 puede enlazar operativamente con el primer reactivo de unión a epítipo 752 a la partícula de aislamiento de núcleo 750. El conector escindible puede comprender un conector escindible químicamente, un enlace fotoescindible, un conector lábil al ácido, un enlace sensible al calor, un enlace escindible enzimáticamente, o una combinación de los mismos. El reactivo de unión a núcleos puede comprender un anticuerpo primario capaz de unirse específicamente a uno o más componentes del núcleo, y el reactivo capaz de unirse específicamente al reactivo de unión a núcleos puede comprender un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario, en donde dicho anticuerpo secundario está unido a la partícula de aislamiento del núcleo mediante un conector escindible. Como se muestra en la FIG. 7B, el paso 700c puede generar núcleos unidos a partículas de aislamiento de núcleos mediante el reactivo capaz de unirse específicamente al reactivo de unión a núcleos.

#### Aislamiento de partículas de aislamiento del núcleo

El flujo de trabajo puede comprender el aislamiento de partículas de aislamiento de núcleo (paso 700d). El aislamiento de partículas de aislamiento de núcleo (y los núcleos unidos a las mismas) puede comprender el aislamiento de la partícula de aislamiento de núcleo mediante extracción magnética, filtración, centrifugación, o cualquier combinación de las mismas. El paso 700d puede generar una pluralidad de partículas de aislamiento de núcleo aisladas (y núcleos unidos a las mismas) desprovistas de orgánulos y/o componentes celulares extranucleares.

Escisión del conector y eliminación de las partículas de aislamiento del núcleo

El flujo de trabajo puede comprender la escisión del conector escindible (efectuando de este modo la separación de los núcleos de las partículas de aislamiento del núcleo) y la eliminación de las partículas de aislamiento del núcleo (paso 700e). La pluralidad de partículas de aislamiento de núcleo aisladas (y los núcleos unidos a ellas) pueden exponerse a las condiciones necesarias para efectuar la escisión del conector escindible. Tales condiciones pueden incluir un entorno electrostático (por ejemplo, pH alto, pH bajo), temperatura (por ejemplo, escindible por calor), una longitud de onda de luz seleccionada (por ejemplo, fotoescindible), un compuesto químico (por ejemplo, escindible químicamente), una enzima (por ejemplo, escindible enzimáticamente). Tras la escisión del conector escindible, los núcleos pueden dejar de estar asociados a las partículas de aislamiento del núcleo. La eliminación de las partículas de aislamiento de núcleos de los núcleos puede producirse por eliminación magnética, centrifugación, filtración o cualquier combinación de las mismas, generando de este modo una población de núcleos purificados.

División, lisis nuclear e identificación con códigos de barras

El flujo de trabajo puede comprender división, lisis nuclear e identificación con códigos de barras (paso 700f). La pluralidad de núcleos puede dividir a una pluralidad de divisiones. Una división puede ser un micropocillo, gotita o emulsión. Una división de la pluralidad de divisiones puede comprender un único núcleo de la pluralidad de núcleos. Una división de la pluralidad de divisiones puede comprender un único núcleo de la pluralidad de núcleos y una partícula de identificación con códigos de barras. En algunas realizaciones, las dimensiones del micropocillo (por ejemplo, la profundidad del micropocillo) se eligen para que optimicen la eficacia de captura de los núcleos y las partículas de identificación con códigos de barras, a la vez que se proporciona un intercambio eficaz de los tampones de ensayo y otros reactivos contenidos dentro de los pocillos.

En algunas realizaciones, una emulsión, micropocillo o pocillo contiene sólo un núcleo. En algunas realizaciones, de 1 a 2.000.000 de emulsiones, micropocillos o pocillos contienen sólo un núcleo cada uno. En algunas realizaciones, el método comprende distribuir como máximo un núcleo en cada emulsión, micropocillo o pocillo. En algunas realizaciones, en una emulsión, micropocillo o pocillo se distribuyen un único soporte sólido y un único núcleo. En algunas realizaciones, de 1 a 2.000.000 de emulsiones, micropocillos o pocillos tienen distribuidos en los mismos un núcleo y un soporte sólido. En algunas realizaciones, el método comprende distribuir como máximo un soporte sólido por emulsión, micropocillo o pocillo. En algunas realizaciones, el método comprende distribuir un soporte sólido y un núcleo a cada uno de 1 a 2.000.000 micropocillos, emulsiones o pocillos. En algunas realizaciones, la distribución de núcleos es aleatoria o no aleatoria. En algunas realizaciones, la distribución de los núcleos es estocástica. En algunas realizaciones, un núcleo se distribuye mediante un clasificador celular. En algunas realizaciones, un núcleo se distribuye poniendo en contacto uno o más pocillos, micropocillos o emulsiones con una solución diluida de núcleos diluida de tal manera que se distribuya como máximo un núcleo a uno o más pocillos, micropocillos o emulsiones.

En algunos casos, los núcleos de una población de núcleos pueden separarse (por ejemplo, aislarse) en pocillos de un sustrato de la divulgación. Por ejemplo, la población de núcleos puede diluirse antes de la separación. La población de núcleos puede diluirse de tal manera que por lo menos el 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 100% de los pocillos del sustrato reciban un solo núcleo. La población de núcleos puede diluirse de tal manera que como máximo el 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 100% de los pocillos del sustrato reciban un solo núcleo. La población de núcleos puede diluirse de tal manera que el número de núcleos de la población diluida sea, o sea por lo menos, el 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, o 100%, del número de pocillos del sustrato. La población de núcleos puede diluirse de tal manera que el número de núcleos en la población diluida sea, o sea por lo menos, el 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 100% del número de pocillos del sustrato. En algunos casos, la población de núcleos se diluye de tal manera que el número de núcleos sea aproximadamente el 10% del número de pocillos del sustrato.

La distribución de núcleos individuales en pocillos del sustrato puede seguir una distribución de Poisson. Por ejemplo, puede haber por lo menos un 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, o 10%, o más de probabilidad de que un pocillo del sustrato tenga más de un núcleo. Puede haber por lo menos un 0,1%, 0,5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, o 10%, o más de probabilidad de que un pocillo del sustrato tenga más de un núcleo. La distribución de los núcleos individuales en los pocillos del sustrato puede ser aleatoria. La distribución de los núcleos individuales en los pocillos del sustrato puede ser no aleatoria. Los núcleos pueden separarse de tal manera que un pocillo del sustrato reciba sólo un núcleo.

Una vez divididos, los núcleos pueden lisarse de acuerdo con los métodos proporcionados en la presente. La identificación con códigos de barras puede comprender el uso de la pluralidad de códigos de barras de la partícula de identificación con códigos de barras para generar una pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras y/o dianas identificadas con códigos de barras de acuerdo con los métodos proporcionados en la presente. Los productos del paso 700f (por ejemplo, oligonucleótidos de indexación de núcleos identificados con códigos de barras y/o dianas identificadas con códigos de barras) pueden someterse a los métodos



posteriores proporcionados en la presente para generar un perfil de expresión de células individuales.

En algunas realizaciones, uno o más pasos del flujo de trabajo comprenden centrifugación. En algunas realizaciones, uno o más pasos del flujo de trabajo no comprenden centrifugación. En algunas realizaciones, ninguno de los pasos del flujo de trabajo comprende centrifugación. En algunas realizaciones, el flujo de trabajo no comprende uno o más del paso 700a, paso 700b, paso 700c, paso 700d, paso 700e y/o paso 700f. Por ejemplo, el reactivo de unión a núcleos puede asociarse con (por ejemplo, inmovilizarse, inmovilizarse parcialmente, encerrarse, encerrarse parcialmente) una partícula de aislamiento del núcleo mediante un conector escindible. En algunas realizaciones, uno o más del paso 700a, paso 700b, paso 700c, paso 700d, paso 700e y/o paso 700f del flujo de trabajo se realizan simultáneamente. Por ejemplo, el contacto con el reactivo de unión a núcleos (paso 700b) y el contacto con las partículas de aislamiento de núcleos (paso 700c) pueden producirse al mismo tiempo.

En por lo menos algunas de las realizaciones descritas anteriormente, uno o más elementos usados en una realización pueden usarse indistintamente en otra realización a menos que tal sustitución no sea técnicamente factible. Los expertos en la técnica apreciarán que a los métodos y estructuras descritos anteriormente se les pueden realizar otras omisiones, adiciones y modificaciones sin apartarse del alcance de la materia reivindicada. Todas estas modificaciones y cambios se incluyen en el ámbito de aplicación del objeto, tal como se define en las reivindicaciones adjuntas.

Con respecto al uso de sustancialmente cualquier término en 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 singular/plural pueden exponerse expresamente en la presente en aras de la claridad. Como se usan en esta memoria descriptiva y en las reivindicaciones adjuntas, las formas singulares "un", "uno" y "el" incluyen referencias plurales a menos que el contexto indique claramente lo contrario. Se pretende que cualquier referencia a "o" en la presente abarque "y/o" a menos que se indique lo contrario.

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) se entienden generalmente 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 expertos en la técnica entenderán además que si se pretende un número específico de una reivindicación introducida, tal intención se enunciará explícitamente en la reivindicación, y en ausencia de tal enunciado no hay tal intención. 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 los enunciados de las reivindicaciones. Sin embargo, no debe interpretarse que el uso de tales frases implica que la introducción de un enunciado de reivindicación mediante los artículos indefinidos "un" o "uno" limita cualquier reivindicación particular que contenga dicho enunciado de reivindicación introducido a realizaciones que contengan solo uno de tales enunciados, incluso cuando la misma reivindicación incluya las frases introductorias "uno o más" o "por lo menos uno" y artículos indefinidos como "un" o "uno" (por ejemplo, "un" y/o "uno" deben interpretarse en el sentido de "por lo menos uno" o "uno o más"); lo mismo ocurre con el uso de artículos definidos usados para introducir los enunciados de las reivindicaciones. Además, incluso si se enuncia explícitamente un número específico de un enunciado de reivindicación introducida, los expertos en la técnica reconocerán que dicho enunciado debe interpretarse en el sentido de que significa por lo menos el número enunciado (por ejemplo, el simple enunciado de "dos enunciados", sin otros modificadores, significa por lo menos dos enunciados, o dos o más enunciados). 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 entiende en el sentido en el que un experto en la técnica entendería la convención (por ejemplo, "un sistema que tenga por lo menos uno de A, B y C" incluiría, entre otros, sistemas que tengan 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 entiende en el sentido en el 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 se limitaría 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 entenderán además que debe entenderse que virtualmente 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 que se incluyan uno de los términos, cualquiera de los términos, o ambos términos.

Además, cuando las características o aspectos de la divulgación se describen en términos de grupos de 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 de Markush.

Como comprenderán los expertos 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 subintervalos de los mismos. Cualquier intervalo enumerado puede reconocerse fácilmente como suficientemente descriptivo y permite descomponer el mismo intervalo en mitades, tercios, cuartos, quintos, décimos, etc., por lo menos iguales. A modo de 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 también comprenderán los expertos en la técnica, todos los términos como "hasta", "por lo menos", "mayor que", "menor que" y similares incluyen el número mencionado y se refieren a intervalos que pueden dividirse posteriormente en subintervalos, como se ha indicado anteriormente. Por último, como comprenderán los expertos en la técnica, un intervalo incluye a cada miembro individual. Así, por ejemplo, un grupo que tiene 1-3 artículos se refiere a grupos que tienen 1, 2 o 3 artículos. De manera similar, un grupo que tiene 1-5 artículos se refiere a grupos que tienen 1, 2, 3, 4 o 5 artículos, y así sucesivamente.

Aunque en la presente se han divulgado varios aspectos y realizaciones, otros aspectos y realizaciones serán evidentes para los expertos en la técnica. Los varios aspectos y realizaciones divulgados en la presente tienen propósitos de ilustración y no se pretende que sean limitativos, su verdadero alcance indicándose en las reivindicaciones siguientes.

## REIVINDICACIONES

1. Un método para determinar el número de dianas en una pluralidad de células, que comprende:

aislar una pluralidad de núcleos de una pluralidad de células usando una composición de aislamiento de núcleos, en donde la composición de aislamiento de núcleos comprende un reactivo de unión a núcleos, en donde el reactivo de unión a núcleos comprende un oligonucleótido de indexación de núcleos, y en donde el oligonucleótido de indexación de núcleos comprende una secuencia de indexación de núcleos, y en donde el reactivo de unión a núcleos es capaz de unirse específicamente a uno o más componentes de un núcleo;

identificar con códigos de barras una pluralidad de dianas en la pluralidad de núcleos usando una pluralidad de códigos de barras para generar una pluralidad de dianas identificadas con códigos de barras, en donde cada uno de la pluralidad de códigos de barras comprende una secuencia de marcador molecular y una región de unión a la diana, y

en donde las secuencias de el marcador molecular de por lo menos dos códigos de barras de la pluralidad de códigos de barras comprenden secuencias diferentes;

obtener datos de secuenciación de la pluralidad de dianas identificadas con códigos de barras;

identificar con códigos de barras los oligonucleótidos de indexación de núcleos usando la pluralidad de códigos de barras para generar una pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras; y

obtener datos de secuenciación de la pluralidad de oligonucleótidos de indexación de núcleos identificados con códigos de barras; y

estimar el número de cada uno de la pluralidad de dianas en la pluralidad de células usando las secuencias de marcadores moleculares de la pluralidad de códigos de barras en los datos de secuenciación.

2. El método de la reivindicación 1, en donde:

(a) aislar la pluralidad de núcleos comprende poner en contacto la pluralidad de núcleos de la pluralidad de células con la composición de aislamiento de núcleos para generar núcleos unidos al reactivo de unión a núcleos; y opcionalmente

(b) aislar la pluralidad de núcleos comprende aislar los núcleos unidos al reactivo de unión a núcleos usando un reactivo capaz de unirse específicamente al reactivo de unión a núcleos; y opcionalmente

(c) el reactivo de unión a núcleos está asociado con un primer epítipo, y en donde el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un primer reactivo de unión a epítipo; y opcionalmente

(d) el primer epítipo comprende biotina, hapteno o una combinación de los mismos; opcionalmente, en donde:

(i) el hapteno comprende digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos; y/o

(ii) el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un anticuerpo antihapteno; o

(iii) el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende avidina, estreptavidina, neutravidina o una combinación de las mismas.

3. El método de la reivindicación 1 o 2, en donde el reactivo de unión a núcleos comprende un grupo funcional seleccionado del grupo que consiste en biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, y una combinación de los mismos.

4. El método de la reivindicación 2 o 3, en donde el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un grupo funcional seleccionado del grupo que consiste en biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, y una combinación de los mismos.

5. El método de la reivindicación 2(b), en donde el reactivo de unión a núcleos comprende:

(a) un anticuerpo primario capaz de unirse específicamente a uno o más componentes del núcleo, y en donde el reactivo capaz de unirse específicamente al reactivo de unión a núcleos comprende un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario; o

(b) un reactivo de unión a carbohidratos, opcionalmente en donde el reactivo de unión a carbohidratos comprende una proteína de unión a carbohidratos, opcionalmente en donde la proteína de unión a carbohidratos comprende una lectina, opcionalmente en donde la lectina comprende:

(i) una lectina de unión a manosa, una lectina de unión a galactosa, una lectina de unión a N-acetilgalactosamina, una lectina de unión a N-acetilglucosamina, una lectina de unión a ácido N-acetilneuramínico, una lectina de unión a fucosa o una combinación de las mismas, o

(ii) Concanavalina A (ConA), lectina de lenteja (LCH), lectina de campanilla de invierno (GNA), aglutinina de

Ricinus communis (RCA), aglutinina de cacahuete (PNA), jacalina (AIL), lectina de veza peluda (VVL), aglutinina de germen de trigo (WGA), lectina de saúco (SNA), leucoaglutinina de Maackia amurensis (MAL), hemoaglutinina de Maackia amurensis (MAH), aglutinina de Ulex europaeus (UEA), lectina de Aleuria aurantia (AAL), o una combinación de las mismas.

6. El método de cualquiera de las reivindicaciones 2-5, en donde el reactivo de unión a núcleos está asociado con una partícula de aislamiento del núcleo, opcionalmente en donde:

- (a) el reactivo capaz de unirse específicamente al reactivo de unión a núcleos está asociado con la partícula de aislamiento del núcleo, opcionalmente en donde el reactivo capaz de unirse específicamente al reactivo de unión a núcleos está inmovilizado en la partícula de aislamiento del núcleo;
- (b) el reactivo capaz de unirse específicamente al reactivo de unión a núcleos está asociado con la partícula de aislamiento del núcleo a través de un conector escindible, opcionalmente en donde el conector escindible comprende un enlace químicamente escindible, un enlace fotoescindible, un conector lábil a los ácidos, un enlace sensible al calor, un enlace escindible enzimáticamente o una combinación de los mismos;
- (c) la partícula de aislamiento del núcleo comprende una perla de aislamiento del núcleo;
- (d) la partícula de aislamiento del núcleo comprende 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 antil fluorocromo, una perla de hidrogel o cualquier combinación de las mismas;
- (e) la partícula de aislamiento del núcleo 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;
- (f) la partícula de aislamiento del núcleo es alterable;
- (g) la partícula de aislamiento del núcleo comprende una partícula de hidrogel alterable de aislamiento del núcleo; y/o
- (h) aislar los núcleos unidos al reactivo de unión a núcleos comprende poner en contacto los núcleos unidos al reactivo de unión a núcleos con una pluralidad de partículas de aislamiento de núcleos.

7. El método de la reivindicación 6, en donde:

- (a) aislar los núcleos unidos al reactivo de unión a núcleos comprende aislar la partícula de aislamiento del núcleo mediante extracción magnética, centrifugación o cualquier combinación de las mismas, comprendiendo opcionalmente, antes de aislar la pluralidad de partículas de aislamiento de núcleos, poner en contacto la pluralidad de núcleos unidos a las partículas de aislamiento de núcleos con el primer epítipo libre, en donde el primer epítipo libre comprende la totalidad o una parte del primer epítipo;
- (b) la pluralidad de códigos de barras está asociada con la partícula de aislamiento del núcleo;
- (c) por lo menos un código de barras de la pluralidad de códigos de barras está inmovilizado en la partícula de aislamiento del núcleo;
- (d) por lo menos un código de barras de la pluralidad de códigos de barras está parcialmente inmovilizado en la partícula de aislamiento del núcleo;
- (e) por lo menos un código de barras de la pluralidad de códigos de barras está encerrado en la partícula de aislamiento del núcleo; y/o
- (f) por lo menos un código de barras de la pluralidad de códigos de barras está parcialmente encerrado en la partícula de aislamiento del núcleo.

8. El método de cualquiera de las reivindicaciones 1-7, que comprende:

antes de aislar la pluralidad de núcleos de la pluralidad de células usando la composición de aislamiento de núcleos, lisar la membrana plasmática de la pluralidad de células; y agotar uno o más orgánulos de la pluralidad de células usando una composición de captura de orgánulos que comprende un reactivo de unión a orgánulos, en donde el reactivo de unión a orgánulos es capaz de unirse específicamente a uno o más componentes de uno o más orgánulos de la pluralidad de células.

9. El método de la reivindicación 8, en donde:

- (a) el reactivo de unión a orgánulos está asociado a un segundo epítipo, y en donde el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos comprende un segundo reactivo de unión a epítipo, opcionalmente en donde el segundo epítipo comprende biotina, hapteno o una combinación de los mismos, opcionalmente en donde el hapteno comprende digoxigenina, 2,4-dinitrofenol, fluoresceína o una combinación de los mismos; y/o

(b) el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos comprende un anticuerpo antihapteno, opcionalmente en donde el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos comprende avidina, estreptavidina, neutravidina o una combinación de los mismos.

5 10. El método de la reivindicación 8 o 9, en donde:

(a) el agotamiento de uno o más orgánulos comprende poner en contacto uno o más orgánulos de la pluralidad de células con la composición de captura de orgánulos para generar uno o más orgánulos unidos al reactivo de unión a componentes orgánulos, opcionalmente en donde el agotamiento de uno o más orgánulos comprende el agotamiento de uno o más orgánulos unidos al reactivo de unión a orgánulos usando un reactivo capaz de unirse específicamente al reactivo de unión a orgánulos, opcionalmente, en donde el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos comprende un grupo funcional seleccionado del grupo que consiste en biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas, y cualquier combinación de los mismos; y/o

b) el reactivo de unión a orgánulos comprende un grupo funcional seleccionado del grupo formado por biotina, estreptavidina, heparina, un aptámero, una fracción de química de clic, digoxigenina, amina o aminas primarias, carboxilo o carboxilos, hidroxilo o hidroxilos, aldehído o aldehídos, cetona o cetonas y cualquier combinación de los mismos.

11. El método de la reivindicación 10,

en donde el reactivo de unión a orgánulos comprende un anticuerpo primario capaz de unirse específicamente a los uno o más componentes de los uno o más orgánulos de la pluralidad de células, y

en donde el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos comprende un anticuerpo secundario capaz de unirse específicamente al anticuerpo primario.

12. El método de cualquiera de las reivindicaciones 8-10, en donde el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos está asociado con una partícula de captura de orgánulos.

13. El método de la reivindicación 12, en donde:

(a) el reactivo capaz de unirse específicamente al reactivo de unión a orgánulos se inmoviliza en la partícula de captura de orgánulos;

(b) la partícula de captura de orgánulos comprende una perla de captura de orgánulos;

(c) la partícula de captura de orgánulos comprende 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 antil fluorocromo o cualquier combinación de las mismas;

(d) la partícula de captura de orgánulos comprende un material seleccionado del grupo consistente en polidimetilsiloxano (PDMS), poliestireno, vidrio, polipropileno, agarosa, gelatina, hidrogel, paramagnético, cerámica, plástico, vidrio, metil estireno, polímero acrílico, titanio, látex, Sefarosa, celulosa, nailon, silicona y cualquier combinación de los mismos; y/o

(e) el agotamiento de los orgánulos de la pluralidad de células usando la composición de captura de orgánulos comprende el agotamiento de las una o más partículas de captura de orgánulos mediante extracción magnética, centrifugación o cualquier combinación de las mismas.

14. El método de cualquiera de las reivindicaciones 8-13, en donde:

(a) los orgánulos comprenden mitocondrias de la pluralidad de células; y/o

(b) los uno o más componentes de los uno o más orgánulos de la pluralidad de células comprenden: ABCD3, ESR2, NOS3, ALB, HIF1A, NR3C1, ATP5A1, HK1, PGR, CASQ1, HSPA1A, PHB, CLTC, HSPD1, PLN, COX4I1, IFM1, SOD1, CPS1, LGALS3, TP53, citocromo C oxidasa, MAPT, TP5B, ERN1, MT-C01, VDAC1, o una combinación de los mismos; y/o

(c) el reactivo de unión a orgánulos comprende un reactivo de unión a componentes de superficie de orgánulos, en donde los uno o más componentes de los uno o más orgánulos comprenden uno o más componentes de superficie de orgánulos, y en donde el reactivo de unión a orgánulos es capaz de unirse específicamente a los uno o más componentes de superficie de orgánulos.

15. El método de cualquiera de las reivindicaciones 1-14, en donde la pluralidad de células comprende:

(a) una muestra de tejido;

(b) muestra de tejido epitelial;

(c) células congeladas;

- (d) células fijadas;
- (e) células fijadas en formol e insertadas en parafina;
- (f) células tumorales;
- (g) células tumorales fijadas;
- 5 (h) células tumorales congeladas; y/o
- (i) células tumorales fijadas en formol e insertadas en parafina.

10

15

20

25

30

35

40

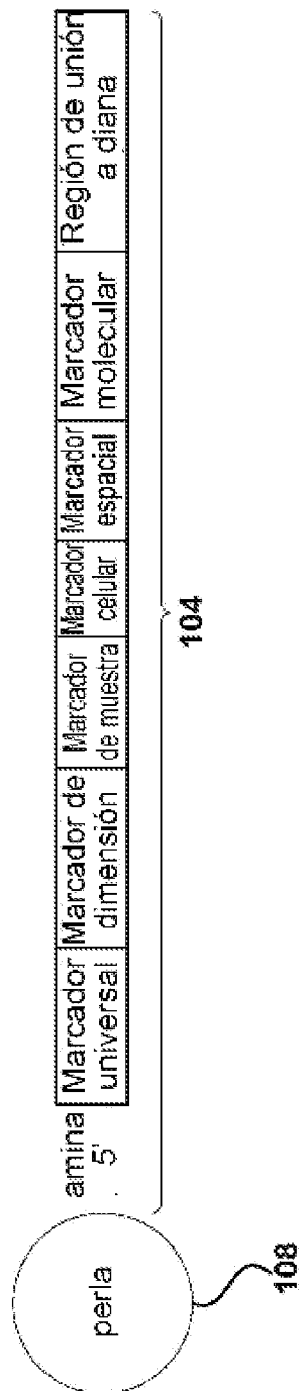
45

50

55

60

65



**FIG. 1**

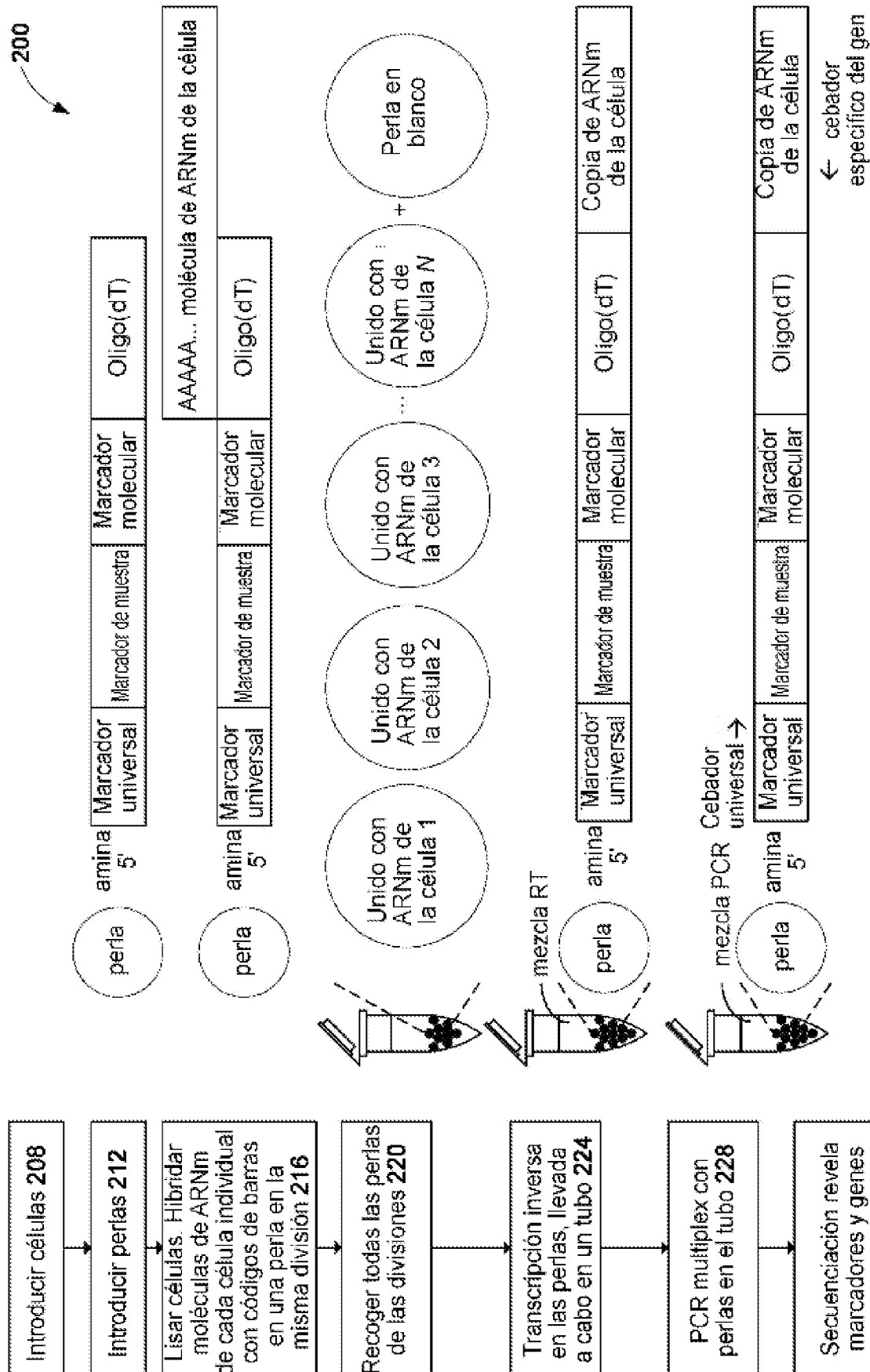
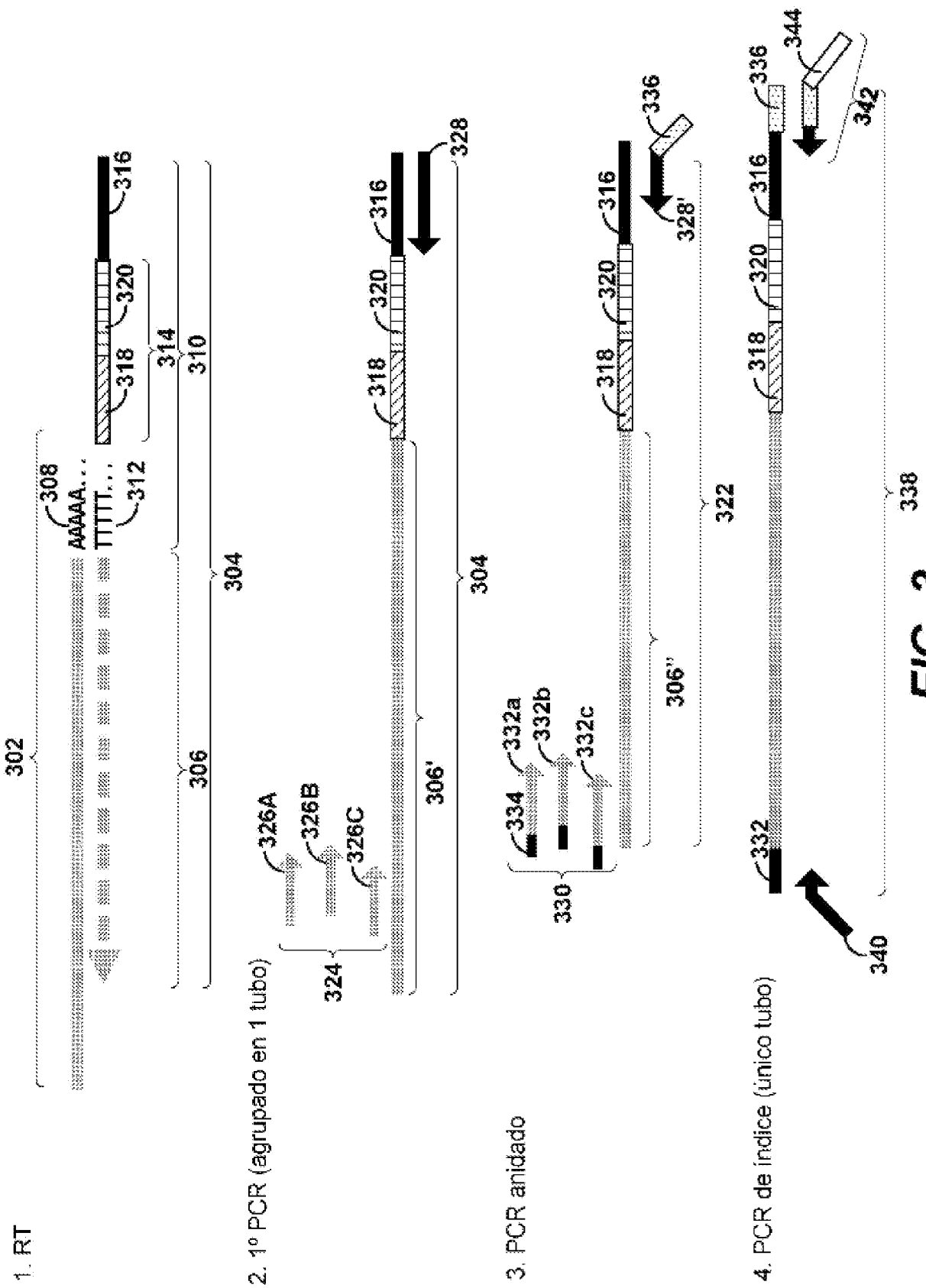
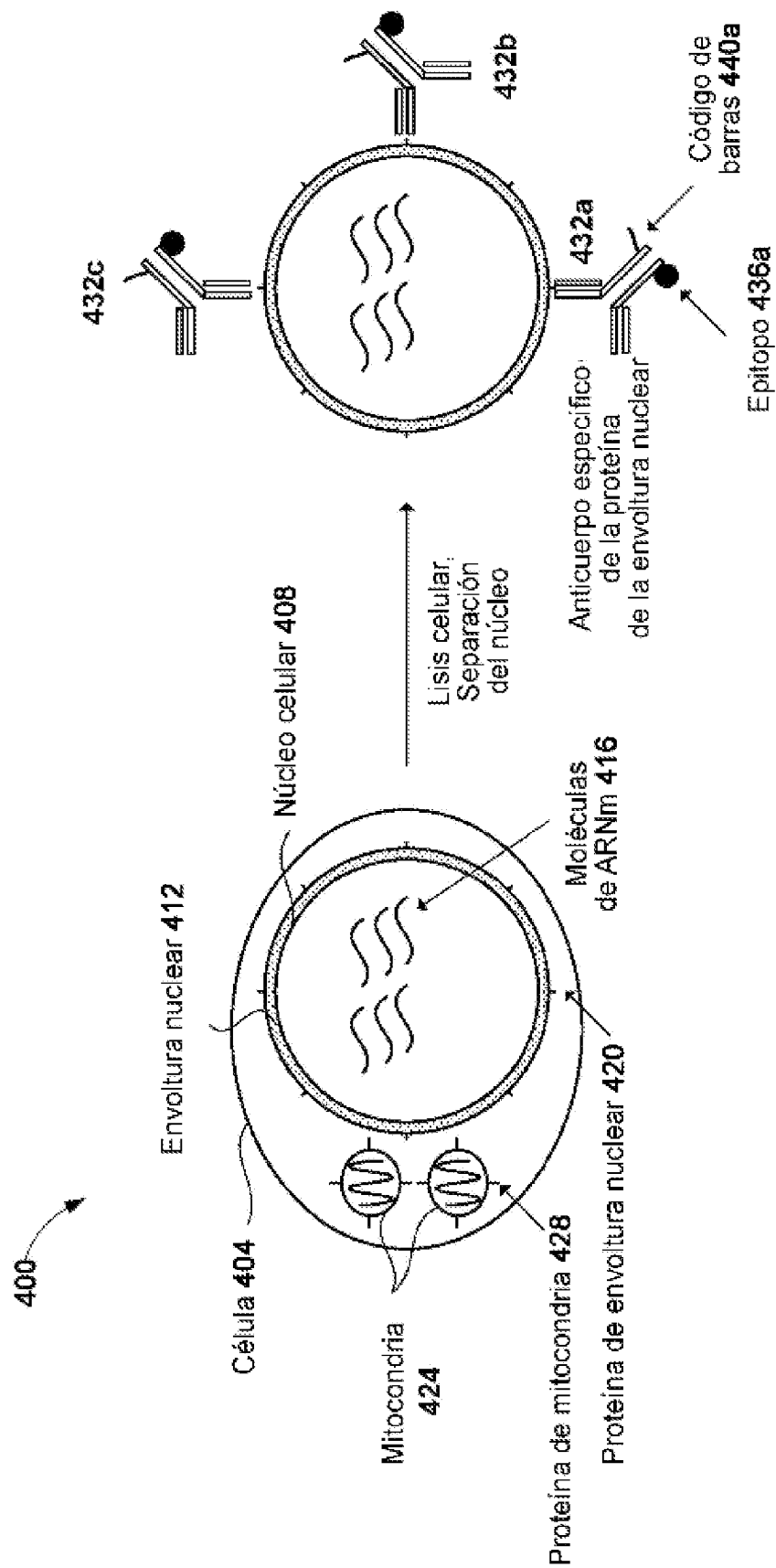


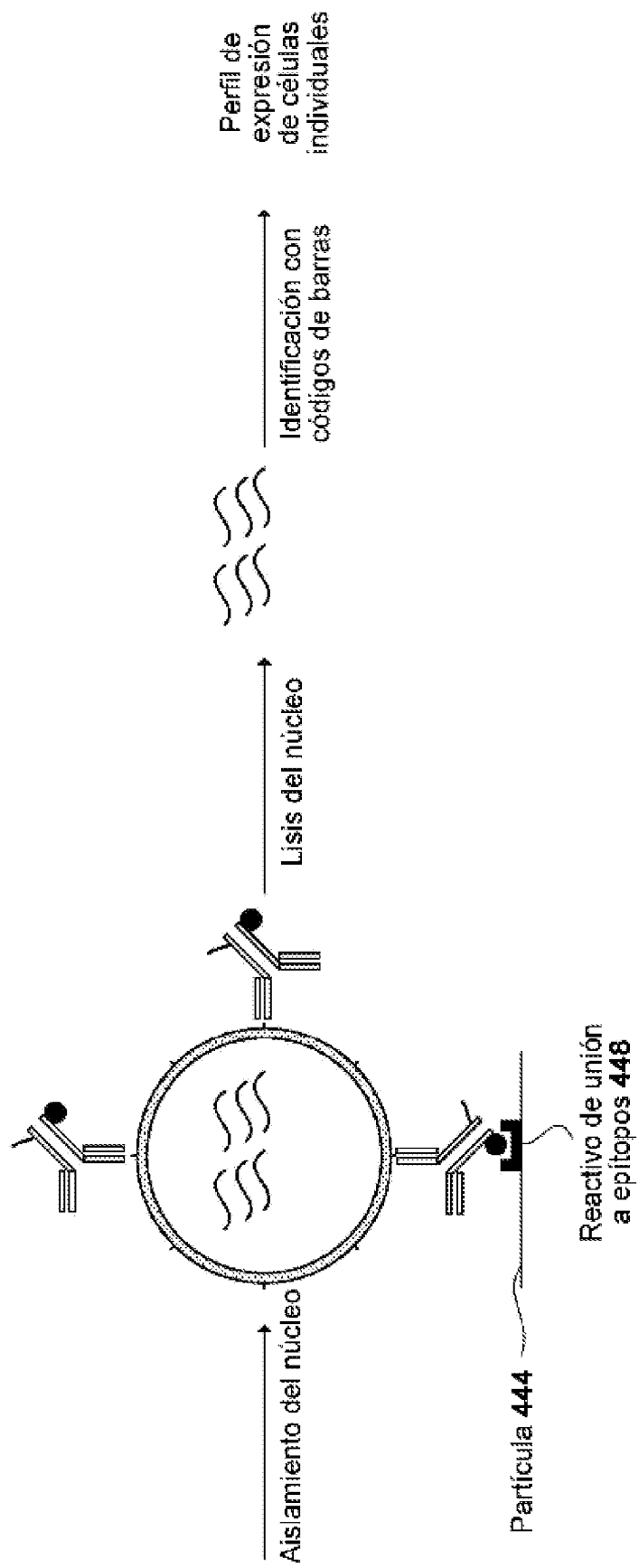
FIG. 2



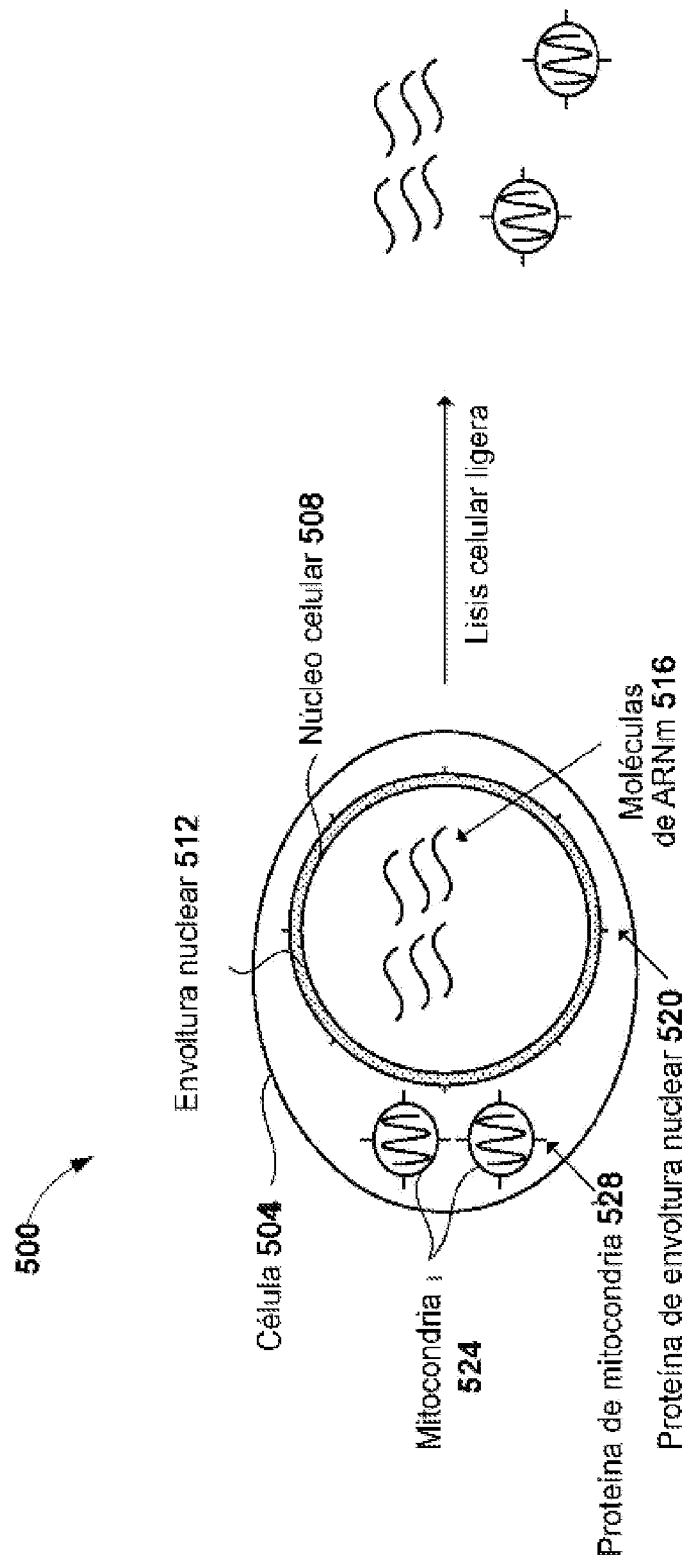




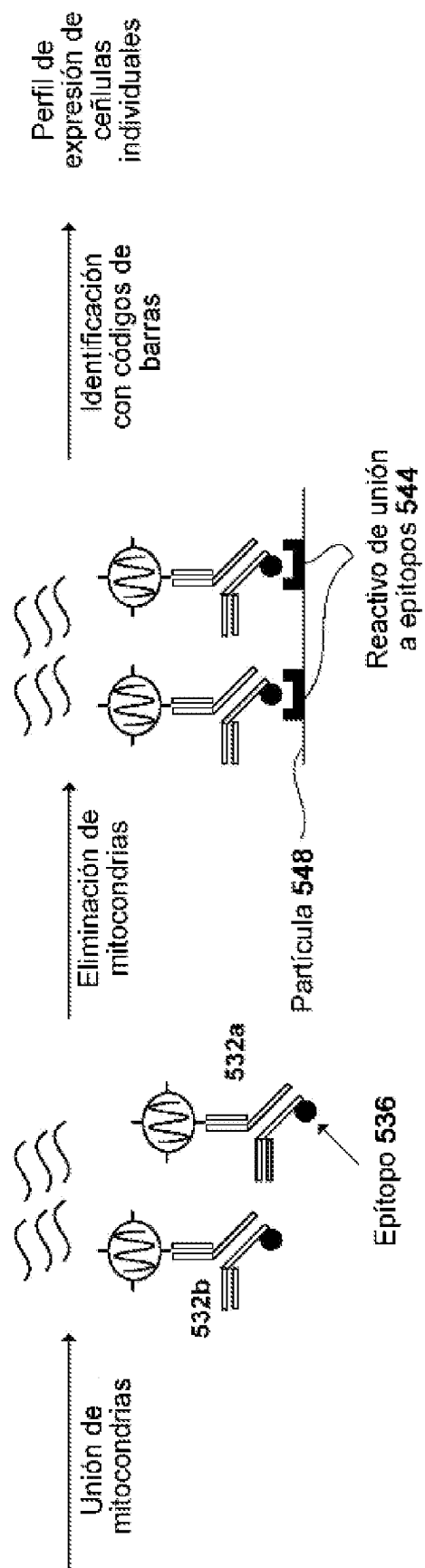
**FIG. 4A**



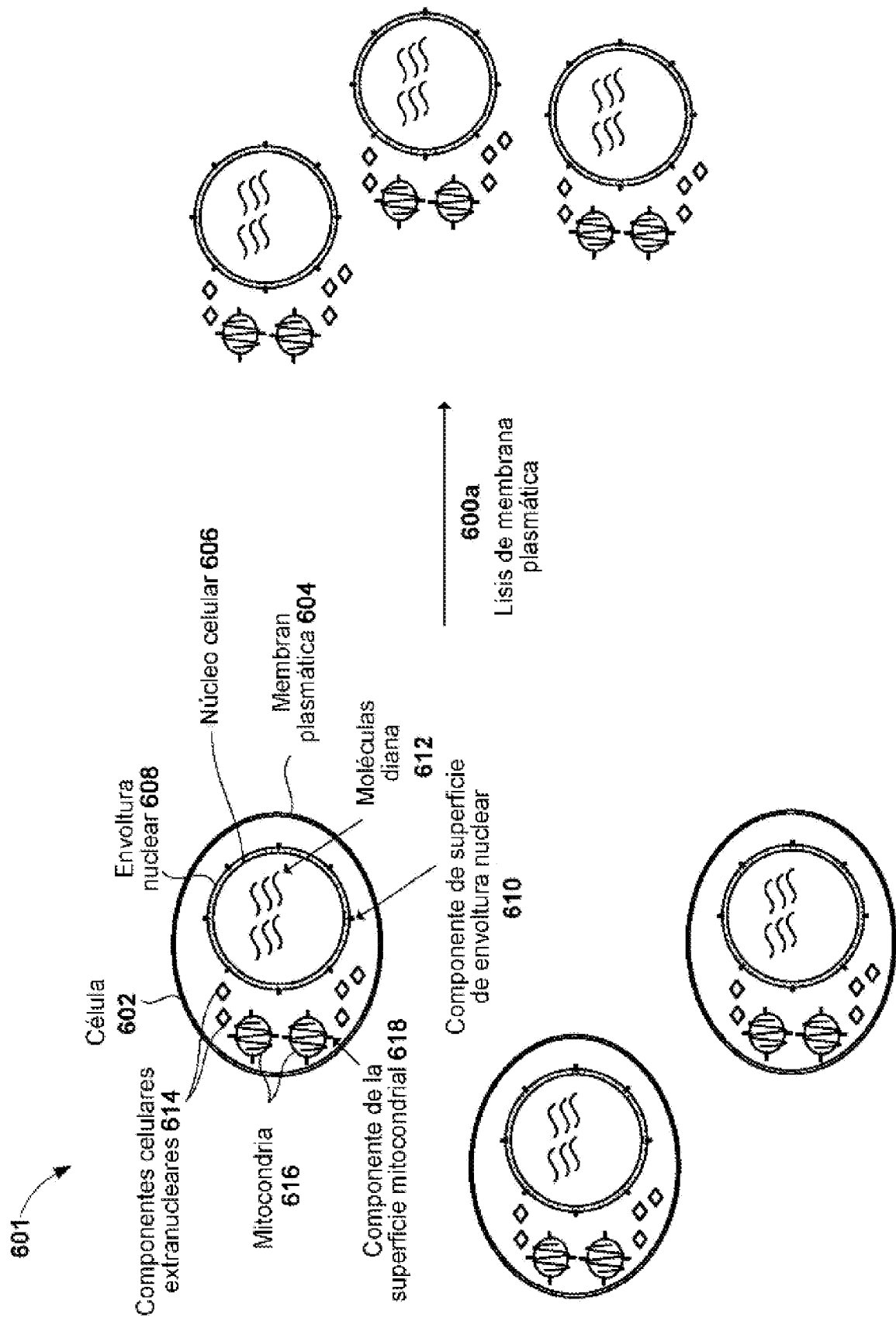
**FIG. 4B**



**FIG. 5A**



**FIG. 5B**



**FIG. 6A**

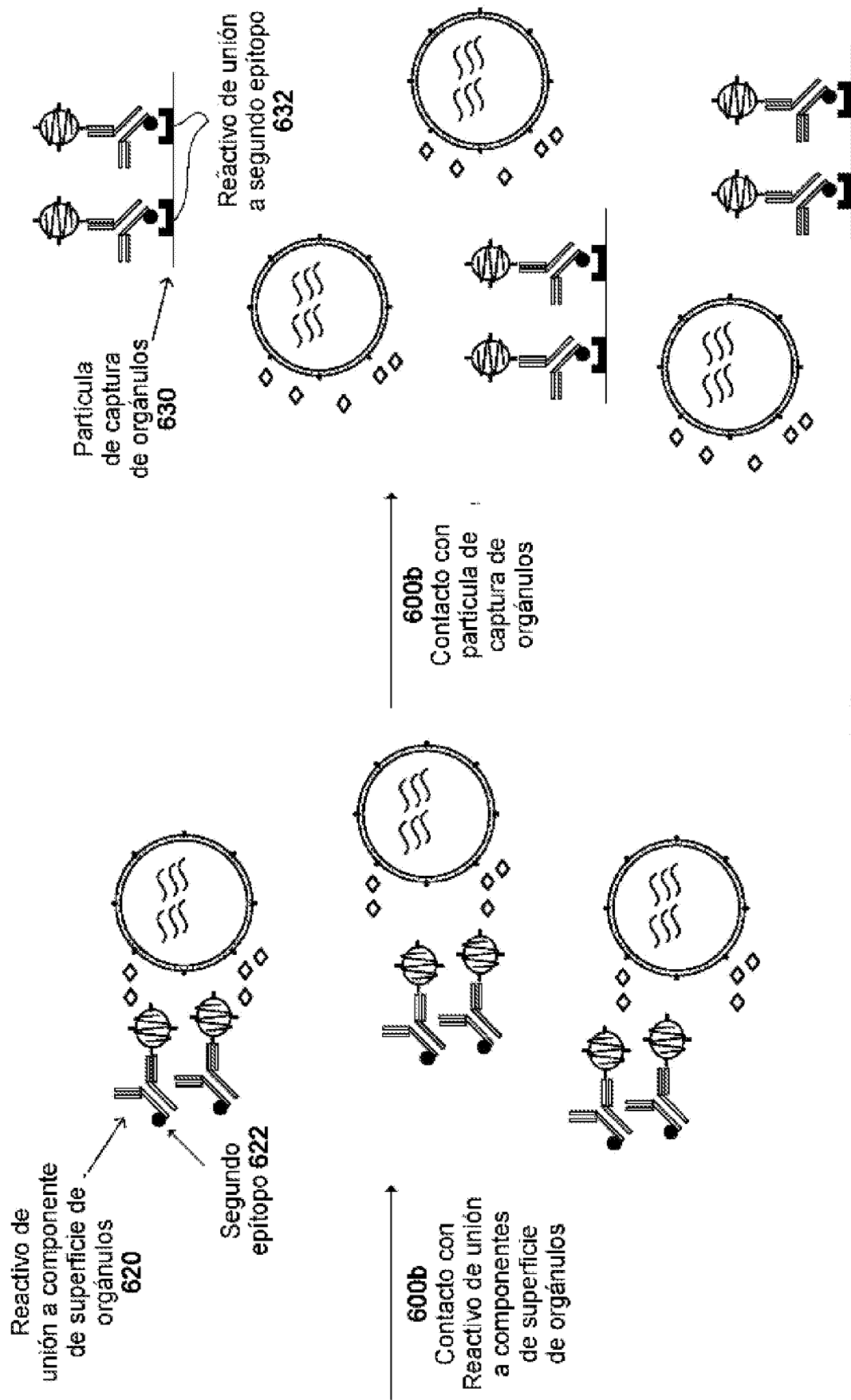


FIG. 6B

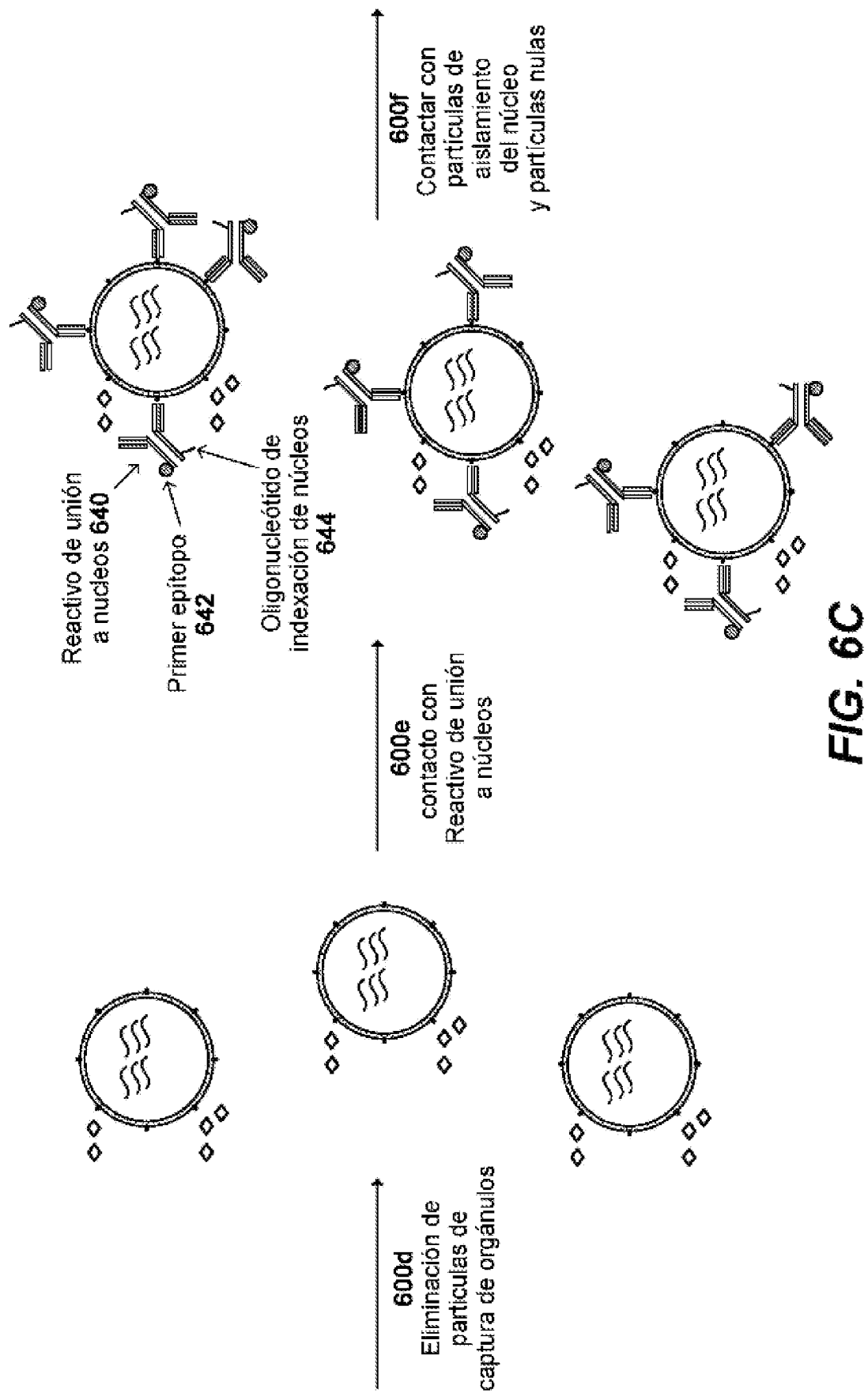
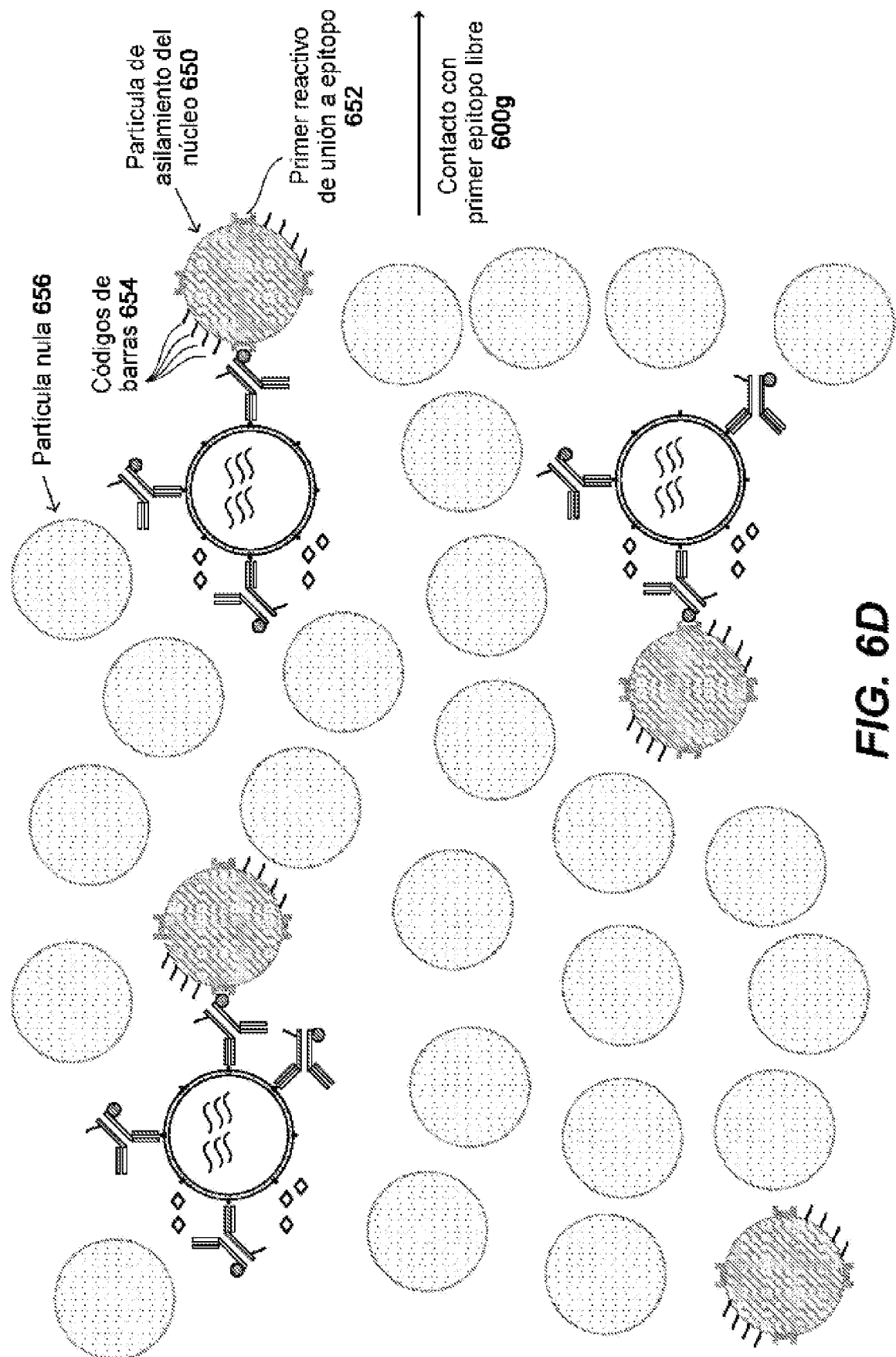
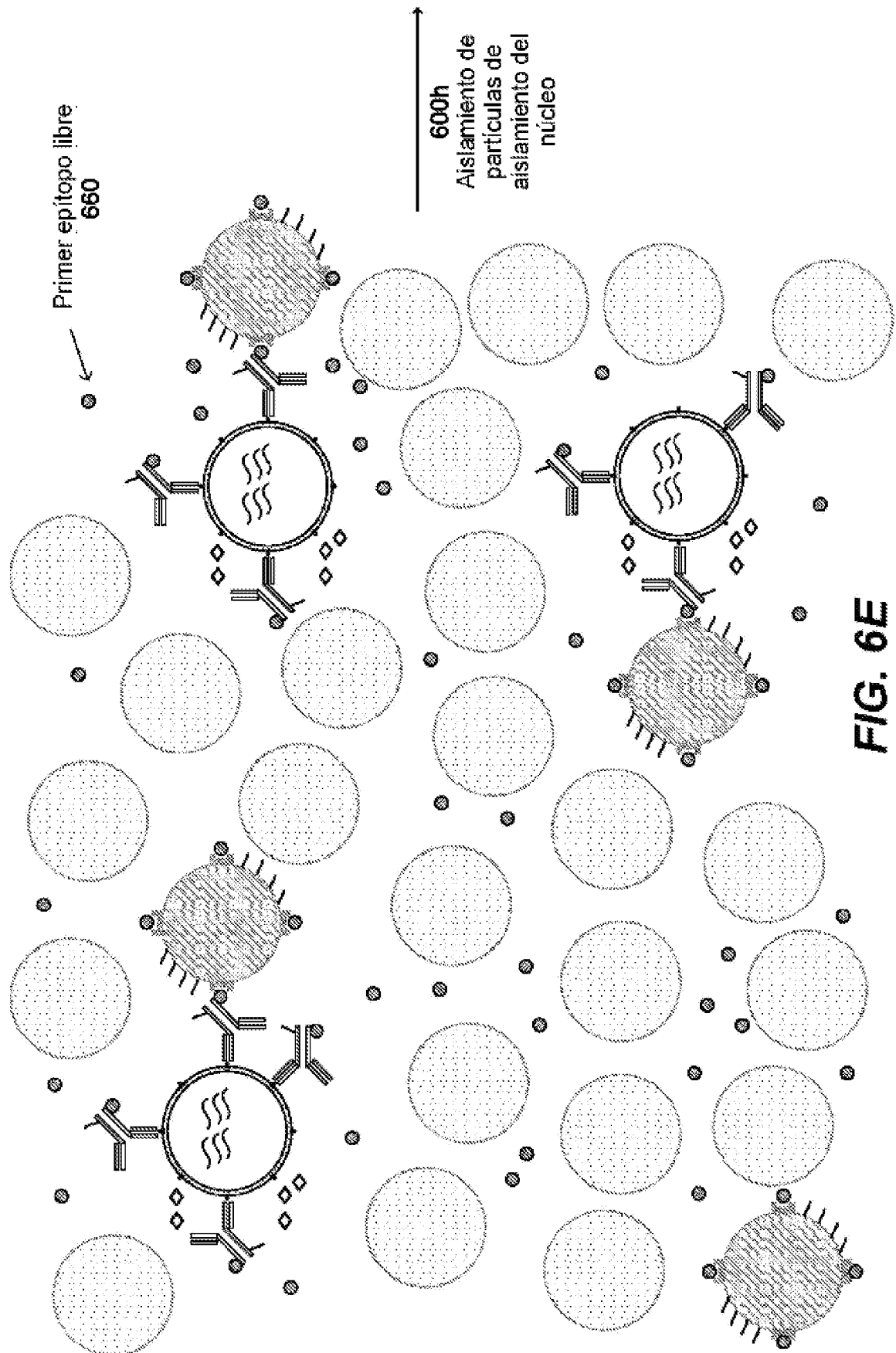
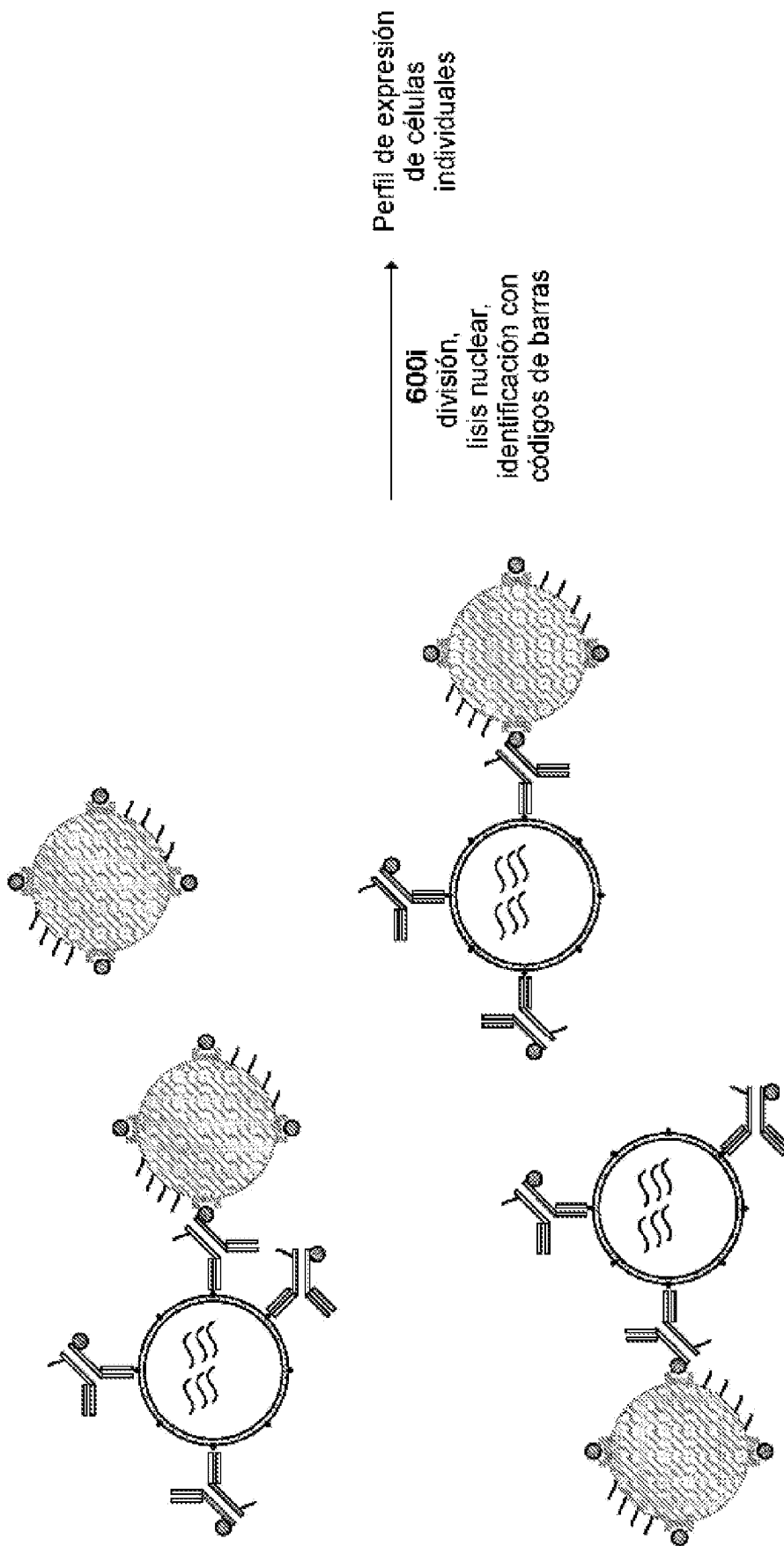


FIG. 6C

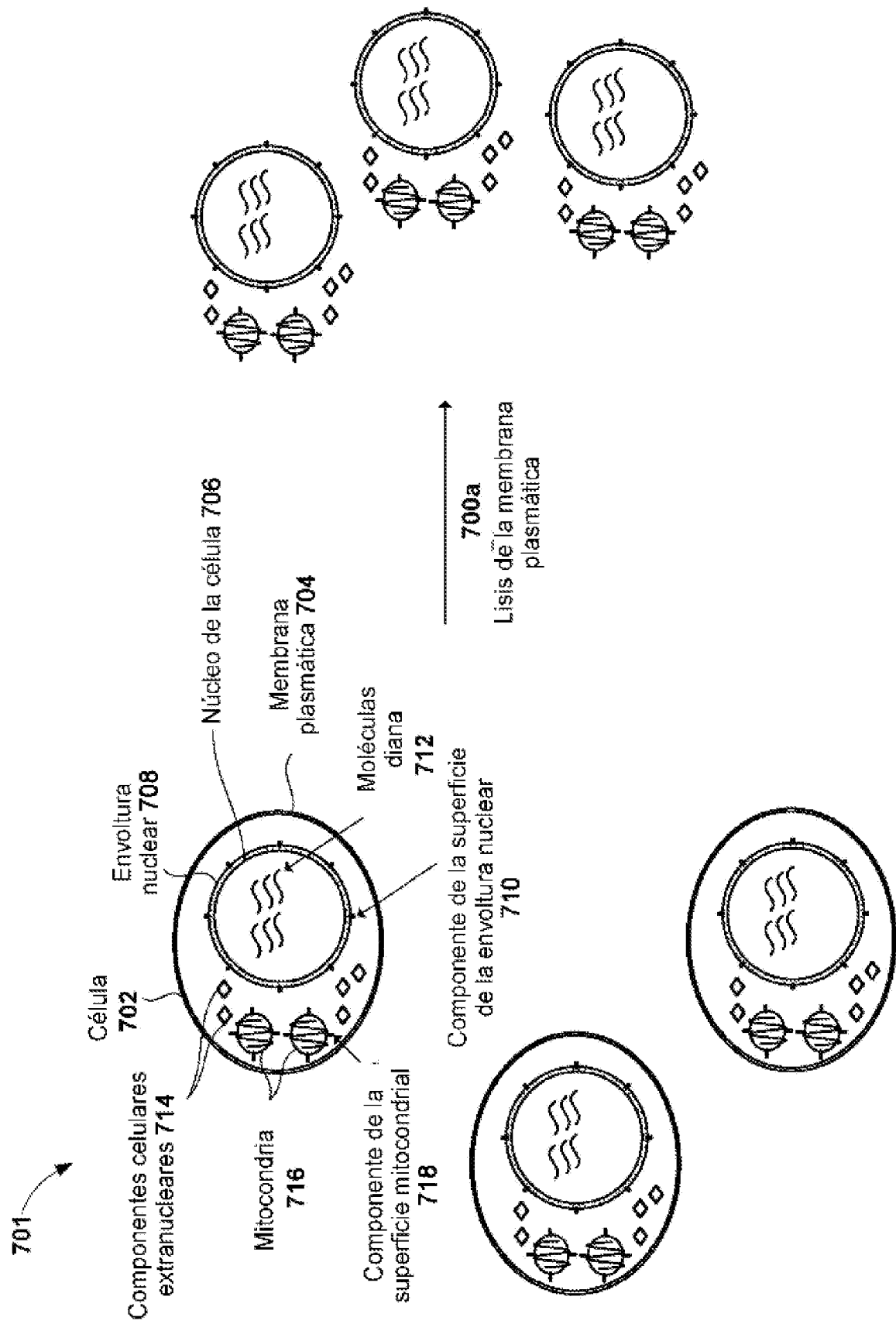




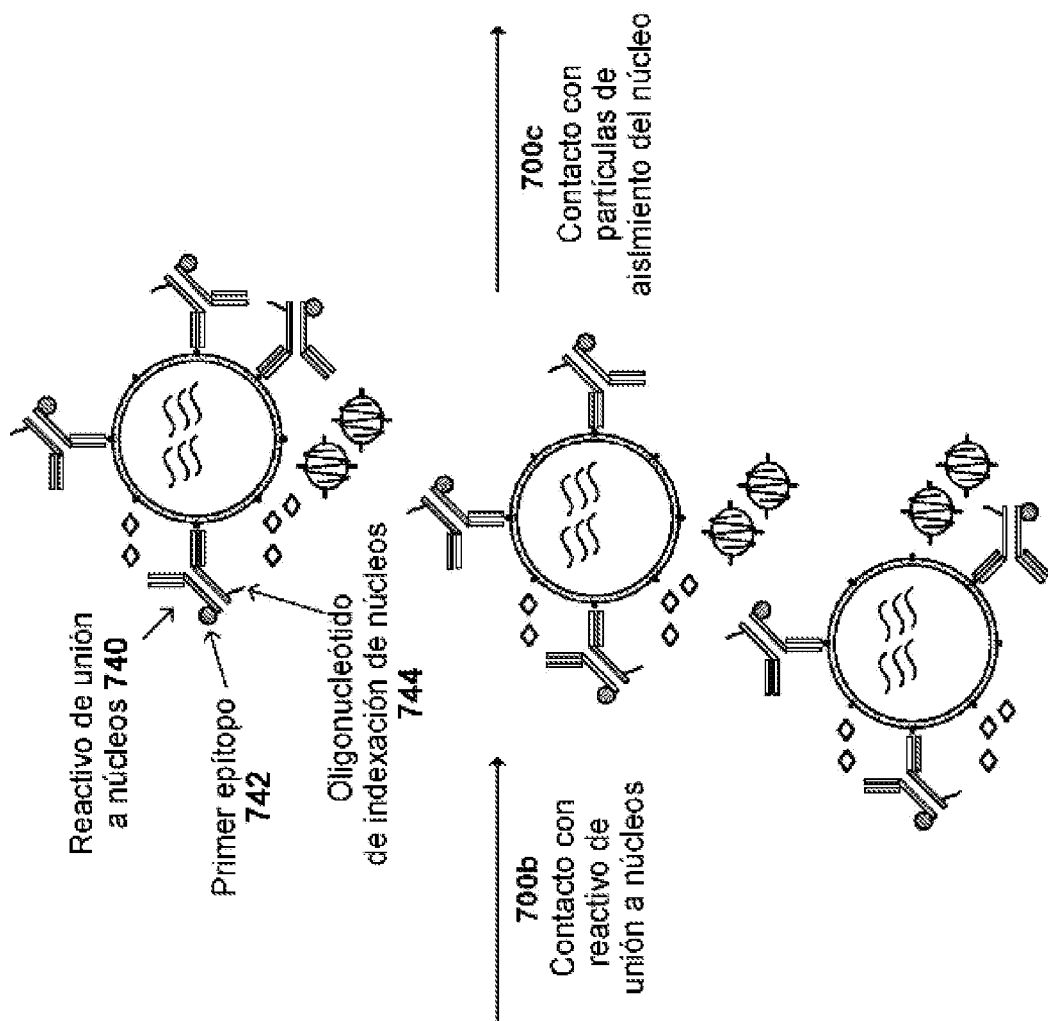
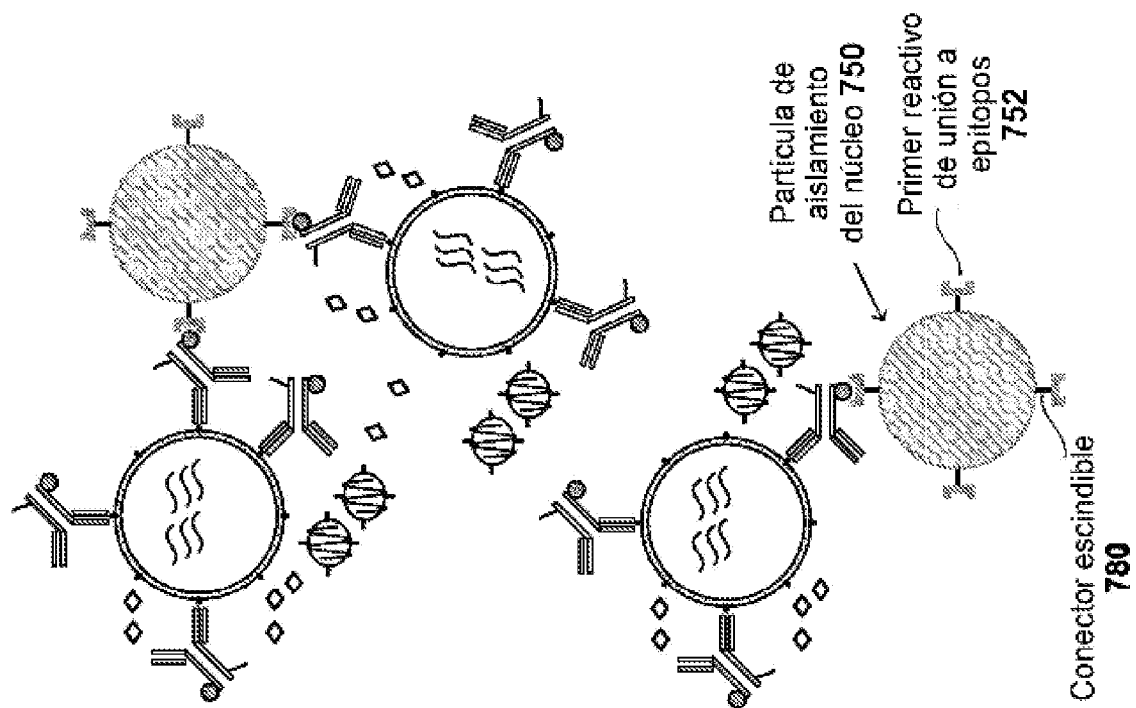




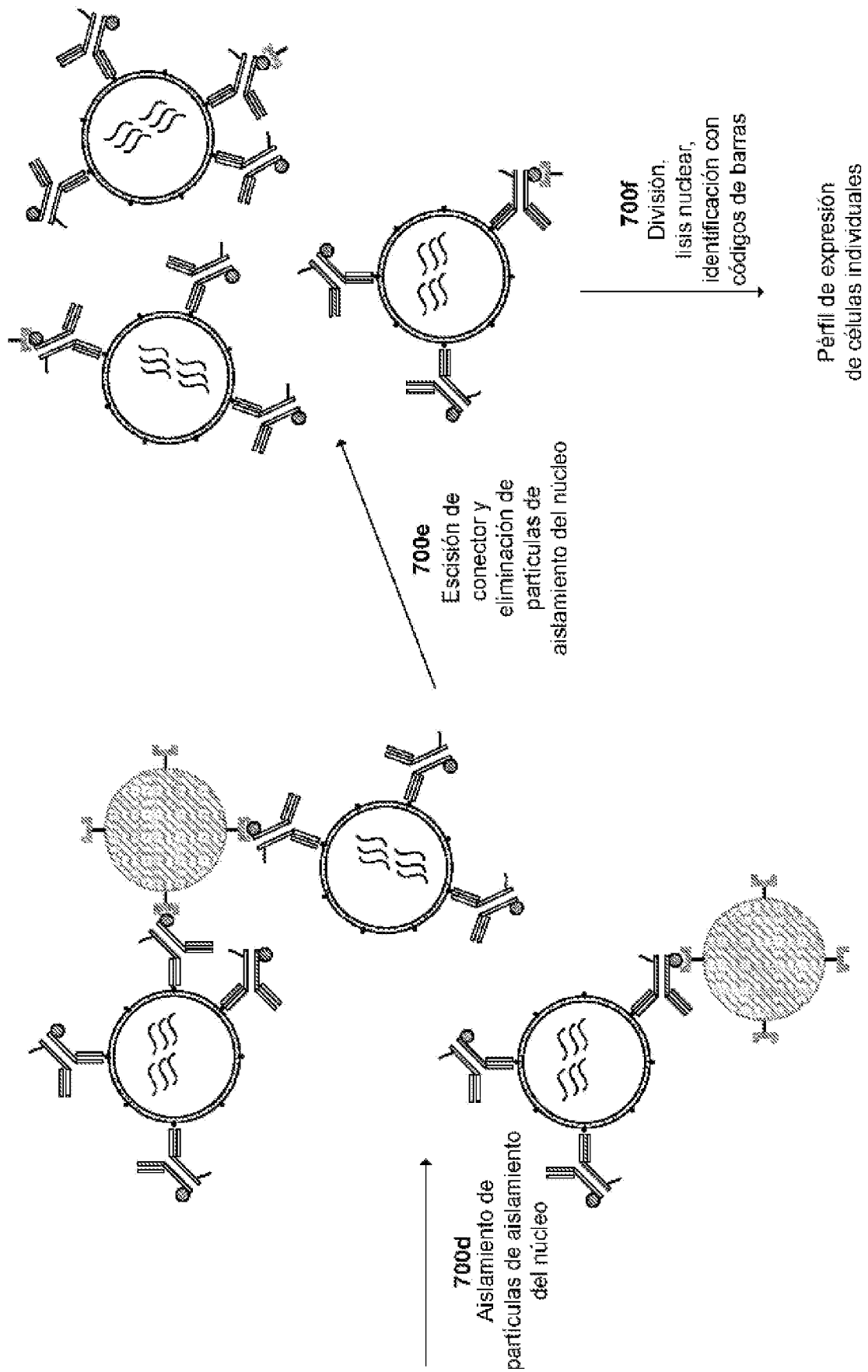
**FIG. 6F**



**FIG. 7A**



**FIG. 7B**



**FIG. 7C**