

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 994 053**

51 Int. Cl.:

**C12Q 1/68**

(2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.04.2016** **E 23195081 (7)**

97 Fecha y número de publicación de la concesión europea: **05.06.2024** **EP 4282977**

54 Título: **Análisis de ácidos nucleicos en formato múltiple, distinguidos espacialmente, de especímenes biológicos**

30 Prioridad:

**10.04.2015 US 201562145874 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**17.01.2025**

73 Titular/es:

**10X GENOMICS SWEDEN AB (50.00%)**  
**Solnavägen 3 H**  
**113 63 Stockholm, SE y**  
**ILLUMINA, INC. (50.00%)**

72 Inventor/es:

**FRISÉN, JONAS;**  
**STÅHL, PATRIK;**  
**LUNDEBERG, JOAKIM;**  
**CANN, GORDON M;**  
**BAZARGAN, LEILA y**  
**ARAVANIS, ALEX**

74 Agente/Representante:

**GONZÁLEZ PECES, Gustavo Adolfo**

ES 2 994 053 T3

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

## DESCRIPCIÓN

Análisis de ácidos nucleicos en formato múltiple, distinguidos espacialmente, de especímenes biológicos

## Antecedentes

5 Uno de cada cuatro hombres morirá de cáncer. Otras estadísticas de la Sociedad Estadounidense contra el Cáncer predicen que una de cada cinco mujeres sufrirá el mismo destino. Se dispone de tratamientos para muchos cánceres. Sin embargo, el éxito para la mayoría depende de la detección temprana.

10 En la actualidad se dice que el cáncer es una enfermedad del genoma. Muchos oncólogos e investigadores oncológicos esperan que los avances en las herramientas de análisis genómico proporcionen una detección temprana y una vía de tratamiento. Sin embargo, estas herramientas son más destacadas en los laboratorios de investigación sin haber madurado aún hasta el nivel de estar fácilmente disponibles para la gran mayoría de los oncólogos. Se necesitan mejoras.

15 Se ha dicho que en el momento del diagnóstico, todos los pacientes con cáncer son mosaicos. Son mosaicos porque tienen al menos dos genomas distintos: el genoma con el que nacieron y el genoma que adquirieron de forma involuntaria por medio del cáncer. Además, a medida que los tumores crecen, se hacen evidentes poblaciones distintas de células cancerosas. Esto conduce a mosaicos aún más complejos dentro del tumor. Esta heterogeneidad de las células cancerosas a menudo da como resultado subpoblaciones de células que responden de forma diferente a las terapias contra el cáncer. El resultado final es a menudo una respuesta positiva inicial de una subpoblación de células, lo que da como resultado la observación de la reducción del tumor del paciente, solo para que le siga el nuevo crecimiento de tejido tumoral y, en algunos casos, metástasis. A pesar de la detección temprana del tumor, la incapacidad de identificar la subpoblación de las células que son resistentes al tratamiento puede dar como resultado la pérdida del tiempo necesario para tratar un cáncer agresivo. Esto crea consecuencias adversas para el paciente tanto emocional como físicamente.

20 El documento WO 2016/138496 A1 proporciona métodos para determinar el número de dianas distintas en ubicaciones espaciales distintas dentro de una muestra, los métodos comprenden codificar estocásticamente la pluralidad de dianas en la muestra mediante el uso de una pluralidad de códigos de barras estocásticos. El documento WO 2014/060483 describe la obtención de perfiles espaciales mediante el uso de una matriz que comprende sondas de captura con códigos de barras espaciales.

Aún existe la necesidad de herramientas genómicas que puedan distinguir las subpoblaciones de células cancerosas en los tumores. La presente descripción aborda esta necesidad y también proporciona otras ventajas.

## Breve resumen

La invención se define por las reivindicaciones adjuntas.

35 La presente descripción proporciona un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico. El método puede incluir las etapas de (a) proporcionar un soporte sólido que comprende una pluralidad de diferentes sondas de ácido nucleico que se ubican aleatoriamente sobre el soporte sólido, en donde cada una de las diferentes sondas de ácido nucleico incluye una secuencia de código de barras que es diferente de la secuencia de código de barras de otras sondas ubicadas aleatoriamente sobre el soporte sólido; (b) realizar una reacción de detección de ácidos nucleicos sobre el soporte sólido para ubicar las secuencias de código de barras sobre el soporte sólido; (c) poner en contacto un espécimen biológico con el soporte sólido que tiene las sondas ubicadas aleatoriamente; (d) hibridar las sondas ubicadas aleatoriamente con ácidos nucleicos diana de porciones del espécimen biológico que están cerca de las sondas ubicadas aleatoriamente; y (e) modificar las sondas ubicadas aleatoriamente que se hibridan con los ácidos nucleicos diana, de esta manera se producen sondas modificadas que incluyen las secuencias de código de barras y una modificación específica de diana, para de esta manera etiquetar espacialmente los ácidos nucleicos del espécimen biológico.

45 Esta descripción proporciona además un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico, el método incluye las etapas de (a) unir diferentes sondas de ácido nucleico a un soporte sólido para producir sondas ubicadas aleatoriamente sobre el soporte sólido, en donde cada una de las diferentes sondas de ácido nucleico incluye una secuencia de código de barras, y en donde cada una de las sondas ubicadas aleatoriamente incluye secuencias de código de barras diferentes a las de otras sondas ubicadas aleatoriamente sobre el soporte sólido; (b) realizar una reacción de detección de ácidos nucleicos sobre el soporte sólido para determinar las secuencias de código de barras de las sondas ubicadas aleatoriamente sobre el soporte sólido; (c) poner en contacto un espécimen biológico con el soporte sólido que tiene las sondas ubicadas aleatoriamente; (d) hibridar las sondas ubicadas aleatoriamente con ácidos nucleicos diana de porciones del espécimen biológico que están cerca de las sondas ubicadas aleatoriamente; y (e) extender las sondas ubicadas aleatoriamente para producir sondas extendidas que incluyen las secuencias de código de barras y las secuencias de los ácidos nucleicos diana, para de esta manera etiquetar espacialmente los ácidos nucleicos del espécimen biológico.

También se proporciona un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico que incluye las etapas de (a) proporcionar una pluralidad de cebadores de ácido nucleico unidos a un soporte sólido, en donde los cebadores de ácido nucleico en la pluralidad incluyen una secuencia de cebador universal que es común a los cebadores de ácido nucleico en la pluralidad; (b) unir una población de sondas de ácido nucleico a la pluralidad de cebadores de ácido nucleico, en donde las sondas de ácido nucleico incluyen una secuencia de unión a cebador universal que se hibrida con la secuencia de cebador universal, una secuencia de captura de diana y una secuencia de código de barras que difiere de las secuencias de código de barras de otras sondas de ácido nucleico en la población, de esta manera se unen las diferentes sondas de ácido nucleico en posiciones ubicadas aleatoriamente sobre el soporte sólido; (c) amplificar las diferentes sondas de ácido nucleico mediante la extensión de los cebadores de ácido nucleico, de esta manera se producen conglomerados de ácidos nucleicos que tienen copias de la secuencia de código de barras y la secuencia de captura de diana en las posiciones ubicadas aleatoriamente sobre el soporte sólido; (d) realizar una reacción de secuenciación para determinar las secuencias de código de barras en las posiciones ubicadas aleatoriamente sobre el soporte sólido; (e) poner en contacto un espécimen biológico con los conglomerados de ácidos nucleicos sobre el soporte sólido; (f) hibridar las secuencias de captura de diana de los conglomerados con los ácidos nucleicos diana de porciones del espécimen biológico que están cerca de los conglomerados; y (g) extender las secuencias de captura de diana para producir sondas extendidas que incluyen las secuencias de los ácidos nucleicos diana y las copias de las secuencias de código de barras, para de esta manera etiquetar los ácidos nucleicos del espécimen biológico.

Esta descripción proporciona además un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico, el método incluye las etapas de (a) proporcionar una matriz de perlas sobre un soporte sólido, en donde diferentes sondas de ácido nucleico se unen a diferentes perlas en la matriz, en donde cada una de las diferentes sondas de ácido nucleico incluye una secuencia de código de barras, en donde cada perla incluye una secuencia de código de barras diferente a la de otras perlas sobre el soporte sólido, y en donde cada una de las diferentes sondas de ácido nucleico incluye una secuencia de captura de diana; (b) realizar una reacción de hibridación de sonda decodificadora sobre el soporte sólido para determinar las secuencias de código de barras en las sondas ubicadas aleatoriamente sobre el soporte sólido; (c) poner en contacto un espécimen biológico con la matriz de perlas; (d) hibridar las diferentes sondas de ácido nucleico con los ácidos nucleicos diana de porciones del espécimen biológico que están cerca de las perlas; y (e) extender las diferentes sondas de ácido nucleico para producir sondas extendidas que incluyen secuencias de los ácidos nucleicos diana y las secuencias de código de barras, para de esta manera etiquetar los ácidos nucleicos del espécimen biológico.

#### Breve descripción de los dibujos

La Figura 1 muestra una representación esquemática de las etapas y reactivos que pueden usarse para generar sondas oligo dT con código de barras en una celda de flujo de Illumina, crear sondas con código de barras extendidas que tienen secuencias de ARNm y liberar las sondas extendidas de la celda de flujo.

La Figura 2 muestra datos que indican la disponibilidad de secuencias de captura de oligo dT en sondas después de la amplificación puente de las sondas y la digestión por enzima de restricción con BspH1 para eliminar uno de los sitios de unión a cebador usados para la amplificación puente.

La Figura 3 muestra métricas de secuenciación de la celda de flujo descrita en el Ejemplo 1 y mostrada en la Figura 2.

La Figura 4 es el número de códigos de barras únicos determinados en 21 mosaicos de la celda de flujo descrita en el Ejemplo 1 y mostrada en la Figura 2.

La Figura 5 muestra una imagen de células capturadas en una celda de flujo con patrón (Panel A) y datos de recuento de células (Panel B).

La Figura 6 muestra las células que permanecen adheridas a una celda de flujo en diferentes condiciones.

La Figura 7 muestra una representación esquemática de las etapas y reactivos usados para crear sondas unidas a un gel (Panel A), una representación esquemática de las etapas y reactivos usados para capturar los ácidos nucleicos diana mediante el uso de las sondas unidas al gel y marcar con fluorescencia las sondas (Panel B), y una imagen creada por los ácidos nucleicos diana marcados con fluorescencia después de la captura con las sondas y la eliminación del tejido del gel.

La Figura 8 muestra una representación esquemática de las etapas y reactivos usados para capturar los ácidos nucleicos diana mediante el uso de sondas unidas a BeadArray™ y marcar con fluorescencia las sondas (Panel A), y una imagen creada por los ácidos nucleicos diana marcados con fluorescencia después de la captura con las sondas y la eliminación del tejido de la BeadArray™.

#### Descripción detallada

La presente descripción proporciona composiciones, aparatos y métodos para conservar la información espacial cuando se realizan análisis de ácidos nucleicos en formato múltiple de especímenes biológicos. Se dispone de una

variedad de herramientas para análisis de ácidos nucleicos en formato múltiple que incluyen, por ejemplo, micromatrices de ácidos nucleicos y las denominadas plataformas de secuenciación de “nueva generación”. Tales herramientas permiten la detección en paralelo de colecciones muy grandes y complejas de ácidos nucleicos, que incluyen, por ejemplo, colecciones de ADN que representan todo o casi todo el material genético de un organismo (es decir, el ‘genoma’), colecciones de ARN (o ADNc) que representan todo o casi todo el complemento de genes expresados (es decir, el “transcriptoma”) de un organismo, y en algunos casos las colecciones pueden incluir varios genomas y/o transcriptomas de varios organismos diferentes (por ejemplo, un metaboloma o bioma de una comunidad o ecosistema). Aunque estas herramientas proporcionan una gran cantidad de información acerca de *qué* secuencias de ácidos nucleicos están presentes en un espécimen biológico en evaluación, no distinguen inherentemente *dónde* reside cualquier ácido nucleico particular en el espécimen biológico. De hecho, la gran mayoría de las muestras aplicadas a las herramientas de análisis de ácidos nucleicos en formato múltiple son homogeneizados derivados de mezclas de muchas células diferentes de un espécimen biológico. Como resultado, se pierde la información espacial y los resultados obtenidos de estas herramientas constituyen un transcriptoma promedio o genoma promedio del espécimen, con pérdida de diferencias importantes entre las células individuales.

En modalidades particulares, la presente descripción proporciona modificaciones nuevas y útiles a las herramientas de análisis de ácidos nucleicos en formato múltiple existentes para permitir la conservación de la información espacial de especímenes biológicos a partir de los cuales se obtienen los ácidos nucleicos. Por ejemplo, los soportes sólidos que se usan usualmente para técnicas de secuenciación por síntesis (SBS) en formato múltiple pueden modificarse para su uso en la captura y etiquetado espacial de ácidos nucleicos de un espécimen biológico. En un ejemplo alternativo, pueden usarse matrices de perlas, tales como las usadas para el genotipado o el análisis de la expresión génica, para capturar y etiquetar espacialmente los ácidos nucleicos de un espécimen biológico. Como se expone en los ejemplos más abajo, los soportes sólidos usados para una plataforma de SBS o BeadArray™ comercializada por Illumina (San Diego, CA) pueden modificarse para el etiquetado espacial. Sin embargo, se debe entender que cualquiera de una variedad de soportes sólidos puede fabricarse y usarse de acuerdo con las enseñanzas en la presente descripción. Los ácidos nucleicos etiquetados espacialmente pueden eliminarse del soporte sólido, combinarse y unirse a un segundo soporte sólido para la detección en cualquiera de una variedad de sistemas de análisis de ácidos nucleicos en formato múltiple que incluyen, por ejemplo, una plataforma de secuenciación o una plataforma de micromatriz expuesta en la presente descripción.

La información espacial proporcionada por un método, composición o aparato en la presente descripción puede incluir, por ejemplo, la ubicación de una o más células en un tejido (u otro espécimen) que tiene un alelo particular en uno o más locus (por ejemplo, un genotipo), tiene una variación estructural particular en el genoma (por ejemplo, fusión, inserción, eliminación, reordenamiento, etc.), tiene una firma epigenética particular (por ejemplo, metilación), expresa un gen particular, expresa un alelo particular de un gen, expresa una variante de corte y empalme particular de un gen o similar. Además de identificar ácidos nucleicos de acuerdo con su ubicación espacial en un espécimen biológico, un método, composición o aparato de la presente descripción puede usarse para cuantificar uno o más ácidos nucleicos de acuerdo con la ubicación espacial. Por ejemplo, la información espacial para una o más células en un tejido (u otro espécimen) puede incluir la cantidad de un alelo o región cromosómica particular en un genoma (por ejemplo, ploidía); la cantidad de modificación epigenética de un locus genético (por ejemplo, metilación); nivel de expresión de un gen, alelo o variante de corte y empalme particular; o similares. Las cantidades pueden ser cantidades absolutas o cantidades relativas de acuerdo con mediciones similares obtenidas en la técnica para muestras mixtas o no etiquetadas espacialmente.

Un método expuesto en la presente descripción puede usarse para la detección localizada de un ácido nucleico en un espécimen biológico. En algunas modalidades, puede usarse un método para identificar o caracterizar todo el transcriptoma o genoma de un espécimen biológico. Alternativamente, puede usarse un método para identificar o caracterizar solo una parte del transcriptoma o genoma de un espécimen. Un subconjunto de transcritos o genes evaluados en un método en la presente descripción puede relacionarse con una enfermedad o afección particular.

Un método expuesto en la presente descripción puede usarse para la detección localizada o espacial de ácidos nucleicos, ya sea ADN o ARN, en un espécimen biológico. Por lo tanto, una o más moléculas de ARN o ADN pueden ubicarse con respecto a su posición o ubicación nativa dentro de una célula o tejido u otro espécimen biológico. Por ejemplo, uno o más ácidos nucleicos pueden localizarse en una célula o grupo de células adyacentes, o tipo de célula, o en regiones particulares de áreas dentro de una muestra de tejido. La ubicación o posición nativa de moléculas de ARN o ADN individuales puede determinarse mediante el uso de un método, aparato o composición de la presente descripción.

Se entenderá que los términos usados en la presente descripción asumen su significado habitual en la técnica en cuestión a menos que se especifique de cualquier otra manera. Varios términos usados en la presente descripción y sus significados se exponen más abajo.

Como se usa en la presente, el término “amplicón”, cuando se usa en referencia a un ácido nucleico, significa el producto de la copia del ácido nucleico, en donde el producto tiene una secuencia de nucleótidos que es la misma o complementaria a al menos una porción de la secuencia de nucleótidos del ácido nucleico. Un amplicón puede producirse mediante cualquiera de una variedad de métodos de amplificación que usan el ácido nucleico, o un amplicón del mismo, como una plantilla que incluyen, por ejemplo, extensión por polimerasa, reacción en cadena de

la polimerasa (PCR), amplificación de círculo rodante (RCA), amplificación por desplazamiento múltiple (MDA), extensión de la ligazón, o reacción en cadena de la ligasa. Un amplicón puede ser una molécula de ácido nucleico que tiene una única copia de una secuencia de nucleótidos particular (por ejemplo, un producto de PCR) o múltiples copias de la secuencia de nucleótidos (por ejemplo, un producto concatamérico de RCA). Un primer amplicón de un ácido nucleico diana es típicamente una copia complementaria. Los amplicones subsecuentes son copias que se crean, después de la generación del primer amplicón, a partir del ácido nucleico diana o del primer amplicón. Un amplicón subsecuente puede tener una secuencia que es sustancialmente complementaria al ácido nucleico diana o sustancialmente idéntica al ácido nucleico diana.

Como se usa en la presente, el término "matriz" se refiere a una población de elementos o sitios que pueden diferenciarse entre sí de acuerdo con la ubicación relativa. Diferentes moléculas que están en diferentes sitios de una matriz pueden diferenciarse entre sí de acuerdo con las ubicaciones de los sitios en la matriz. Un sitio individual de una matriz puede incluir una o más moléculas de un tipo particular. Por ejemplo, un sitio puede incluir una única molécula de ácido nucleico diana que tiene una secuencia particular o un sitio puede incluir varias moléculas de ácidos nucleicos que tienen la misma secuencia (y/o secuencia complementaria, de la misma). Los sitios de una matriz pueden ser diferentes elementos ubicados en el mismo sustrato. Los elementos ilustrativos incluyen, sin limitación, pocillos en un sustrato, perlas (u otras partículas) en o sobre un sustrato, proyecciones desde un sustrato, crestas sobre un sustrato o canales en un sustrato. Los sitios de una matriz pueden ser sustratos separados, cada uno de los cuales tiene una molécula diferente. Pueden identificarse diferentes moléculas unidas a sustratos separados de acuerdo con las ubicaciones de los sustratos en una superficie a la que se asocian los sustratos o de acuerdo con las ubicaciones de los sustratos en un líquido o gel. Las matrices ilustrativas en las que se ubican sustratos separados sobre una superficie incluyen, sin limitación, las que tienen perlas en pocillos.

Como se usa en la presente, el término "unido" se refiere al estado de dos cosas que se unen, sujetan, adhieren, conectan o enlazan entre sí. Por ejemplo, un analito, tal como un ácido nucleico, puede unirse a un material, tal como un gel o soporte sólido, mediante un enlace covalente o no covalente. Un enlace covalente se caracteriza por compartir pares de electrones entre átomos. Un enlace no covalente es un enlace químico que no implica compartir pares de electrones y puede incluir, por ejemplo, enlaces de hidrógeno, enlaces iónicos, fuerzas van der Waals, interacciones hidrófilas e interacciones hidrófobas.

Como se usa en la presente, el término "secuencia de código de barras" se refiere a una serie de nucleótidos en un ácido nucleico que puede usarse para identificar el ácido nucleico, una característica del ácido nucleico, o una manipulación que se ha llevado a cabo en el ácido nucleico. La secuencia de código de barras puede ser una secuencia de origen natural o una secuencia que no aparece naturalmente en el organismo del que se obtuvo el ácido nucleico con código de barras. Una secuencia de código de barras puede ser única para una especie de ácido nucleico individual en una población o varias especies de ácidos nucleicos diferentes en una población pueden compartir una secuencia de código de barras. Por ejemplo, cada sonda de ácido nucleico en una población puede incluir secuencias de código de barras diferentes a las de todas las demás sondas de ácido nucleico en la población. Alternativamente, cada sonda de ácido nucleico en una población puede incluir secuencias de código de barras diferentes a las de algunas o la mayoría de las otras sondas de ácido nucleico en una población. Por ejemplo, cada sonda en una población puede tener un código de barras que está presente para varias sondas diferentes en la población, aunque las sondas con el código de barras común difieran entre sí en otras regiones de secuencia a lo largo de su longitud. En modalidades particulares, una o más secuencias de código de barras que se usan con un espécimen biológico no están presentes en el genoma, transcriptoma u otros ácidos nucleicos del espécimen biológico. Por ejemplo, las secuencias de código de barras pueden tener menos de 80 %, 70 %, 60 %, 50 % o 40 % de identidad de secuencia con las secuencias de ácido nucleico en un espécimen biológico particular.

Como se usa en la presente, el término "especimen biológico" se refiere a una o más células, tejidos, organismos o porciones de los mismos. Se puede obtener un espécimen biológico de cualquiera de una variedad de organismos. Los organismos ilustrativos incluyen, pero no se limitan a, un mamífero tal como un roedor, ratón, rata, conejo, cobaya, ungulado, caballo, oveja, cerdo, cabra, vaca, gato, perro, primate (es decir, primate humano o no humano); una planta tal como *Arabidopsis thaliana*, maíz, sorgo, avena, trigo, arroz, canola o soja; una alga tal como *Chlamydomonas reinhardtii*; un nemátodo tal como *Caenorhabditis elegans*; un insecto tal como *Drosophila melanogaster*, mosquito, mosca de fruta, abeja melífera o araña; un pez tal como un pez cebra; un reptil; un anfibio tal como una rana o *Xenopus laevis*; un *Dictyostelium discoideum*; un hongo tal como *Pneumocystis carinii*, *Takifugu rubripes*, levadura, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*; o un Plasmodium falciparum. Los ácidos nucleicos diana también pueden derivarse de un procarionta tal como una bacteria, *Escherichia coli*, *Staphylococci* o *Mycoplasma pneumoniae*; una arquea; un virus tal como el virus de la hepatitis C o el virus de la inmunodeficiencia humana; o un viroide. Los especímenes pueden derivarse de un cultivo o población homogénea de los organismos anteriores o alternativamente de una colección de varios organismos diferentes, por ejemplo, en una comunidad o ecosistema.

Como se usa en la presente, el término "sitio de escisión" se refiere a una ubicación en una molécula de ácido nucleico que es susceptible a la ruptura de enlaces. La ubicación puede ser específica para un proceso químico, enzimático o físico particular que da como resultado la ruptura de enlaces. Por ejemplo, la ubicación puede ser un nucleótido que es abásico o un nucleótido que tiene una base que es susceptible a eliminarse para crear un sitio abásico. Los ejemplos de nucleótidos que son susceptibles a eliminarse incluyen uracilo y 8-oxo-guanina como se expone con más detalle

en la presente descripción más abajo. La ubicación también puede estar en o cerca de una secuencia de reconocimiento para una endonucleasa de restricción tal como una enzima formadora de mellas.

Como se usa en la presente, el término “conglomerado”, cuando se usa en referencia a los ácidos nucleicos, se refiere a una población de los ácidos nucleicos que se une a un soporte sólido para formar un elemento o sitio. Generalmente los ácidos nucleicos son miembros de una sola especie, de esta manera forman un conglomerado monoclonal. Una “población monoclonal” de ácidos nucleicos es una población que es homogénea con respecto a una secuencia de nucleótidos particular. Los conglomerados no tienen por qué ser monoclonales. Más bien, para algunas aplicaciones, un conglomerado puede poblarse predominantemente con amplicones de un primer ácido nucleico y también puede tener un nivel bajo de amplicones contaminantes de un segundo ácido nucleico. Por ejemplo, cuando se va a usar una matriz de conglomerados en una aplicación de detección, un nivel aceptable de contaminación sería un nivel que no afecte la señal a ruido o resolución de la técnica de detección de una manera inaceptable. En consecuencia, la clonalidad aparente será generalmente relevante para un uso o aplicación particular de una matriz preparada mediante los métodos expuestos en la presente descripción. Los niveles ilustrativos de contaminación que pueden ser aceptables en un conglomerado individual incluyen, pero no se limitan a, a lo máximo, 0,1 %, 0,5 %, 1 %, 5 %, 10 %, 5 25 % o 35 % de amplicones contaminantes. Los ácidos nucleicos en un conglomerado generalmente se unen covalentemente a un soporte sólido, por ejemplo, por medio de sus extremos 5', pero en algunos casos son posibles otros medios de unión. Los ácidos nucleicos en un conglomerado pueden ser de una hebra o de doble hebra. En algunas modalidades, pero no en todas, los conglomerados se preparan mediante un método de amplificación en fase sólida conocido como amplificación puente. Las configuraciones ilustrativas de conglomerados y métodos para su producción se exponen, por ejemplo, en la patente de Estados Unidos núm. 5,641,658; la publicación de patente de Estados Unidos núm. 2002/0055100; la patente de Estados Unidos núm. 7,115,400; la publicación de patente de Estados Unidos núm. 2004/0096853; la publicación de patente de Estados Unidos núm. 2004/0002090; la publicación de patente de Estados Unidos núm. 2007/0128624; y la publicación de patente de Estados Unidos núm. 2008/0009420.

Como se usa en la presente, el término “diferente”, cuando se usa en referencia a ácidos nucleicos, significa que los ácidos nucleicos tienen secuencias de nucleótidos que no son iguales entre sí. Dos o más ácidos nucleicos pueden tener secuencias de nucleótidos que son diferentes a lo largo de toda su longitud. Alternativamente, dos o más ácidos nucleicos pueden tener secuencias de nucleótidos que son diferentes a lo largo de una porción sustancial de su longitud. Por ejemplo, dos o más ácidos nucleicos pueden tener porciones de secuencias de nucleótidos diana que son diferentes para las dos o más moléculas mientras que también tienen una porción de secuencia universal que es la misma en las dos o más moléculas. Dos perlas pueden ser diferentes entre sí en virtud de unirse a ácidos nucleicos diferentes.

Como se usa en la presente, el término “cada uno”, cuando se usa en referencia a una colección de objetos, pretende identificar un objeto individual en la colección pero no se refiere necesariamente a cada objeto en la colección. Pueden ocurrir excepciones si la descripción o el contexto explícitos lo indican claramente de cualquier otra manera.

Como se usa en la presente, el término “extender”, cuando se usa en referencia a un ácido nucleico, se refiere a la adición de al menos un nucleótido u oligonucleótido al ácido nucleico. En modalidades particulares uno o más nucleótidos pueden añadirse al extremo 3' de un ácido nucleico, por ejemplo, por medio de catálisis de polimerasa (por ejemplo, ADN polimerasa, ARN polimerasa o transcriptasa inversa). Pueden usarse métodos químicos o enzimáticos para añadir uno o más nucleótidos al extremo 3' o 5' de un ácido nucleico. Se pueden añadir uno o más oligonucleótidos al extremo 3' o 5' de un ácido nucleico, por ejemplo, por medio de métodos químicos o enzimáticos (por ejemplo, catálisis de ligasa). Un ácido nucleico puede extenderse de una manera dirigida por plantilla, de manera que el producto de extensión es complementario a un ácido nucleico plantilla que se hibrida con el ácido nucleico que se extiende.

Como se usa en la presente, el término “elemento” significa una ubicación en una matriz para una especie particular de molécula. Un elemento puede contener una sola molécula solamente o puede contener una población de varias moléculas de la misma especie. Los elementos de una matriz son típicamente discretos. Los elementos discretos pueden ser contiguos o pueden tener espacios entre sí. El tamaño de los elementos y/o la separación entre los elementos puede variar de manera que las matrices pueden ser de densidad alta, densidad media o menor densidad. Las matrices de densidad alta se caracterizan por tener sitios separados por menos de aproximadamente 15 µm. Las matrices de densidad media tienen sitios separados por aproximadamente 15 a 30 µm, mientras que las matrices de densidad baja tienen sitios separados por más de 30 µm. Una matriz útil en la presente descripción puede tener, por ejemplo, sitios que están separados por menos de 100 µm, 50 µm, 10 µm, 5 µm, 1 µm o 0,5 µm. Un aparato o método de la presente descripción puede usarse para detectar una matriz a una resolución suficiente para distinguir sitios en las densidades o intervalos de densidad anteriores.

Como se usa en la presente, el término “mezcla de fluidos” se refiere a dos o más objetos diferentes que están presentes simultáneamente en una solución. Típicamente, los dos o más objetos son libremente difusibles en la solución. Los dos o más objetos pueden ser diferentes tipos de objetos (por ejemplo, un ácido nucleico y una proteína que son diferentes tipos de moléculas) o pueden ser diferentes especies del mismo tipo de objetos (por ejemplo, dos moléculas de ácido nucleico que tienen secuencias diferentes). Los objetos ilustrativos que pueden estar en una mezcla de fluidos incluyen, pero no se limitan a, moléculas, células o perlas.

Como se usa en la presente, el término "celda de flujo" se refiere a un recipiente que tiene una cámara donde puede llevarse a cabo una reacción, una entrada para suministrar reactivos a la cámara y una salida para eliminar reactivos de la cámara. En algunas modalidades la cámara se configura para la detección de la reacción que ocurre en la cámara. Por ejemplo, la cámara puede incluir una o más superficies transparentes que permiten la detección óptica de especímenes biológicos, moléculas marcadas ópticamente, o similares en la cámara. Las celdas de flujo ilustrativas incluyen, pero no se limitan a las usadas en un aparato de secuenciación de ácido nucleico tal como celdas de flujo para las plataformas Genome Analyzer®, MiSeq®, NextSeq® o HiSeq® comercializadas por Illumina, Inc. (San Diego, CA); o para la plataforma de secuenciación SOLiD™ o Ion Torrent™ comercializada por Life Technologies (Carlsbad, CA). Las celdas de flujo y los métodos ilustrativos para su fabricación y uso también se describen, por ejemplo, en el documento WO 2014/142841 A1; la publicación de solicitud de patente de Estados Unidos núm. 2010/0111768 A1 y la patente de Estados Unidos núm. 8,951,781.

Como se usa en la presente, el término "gel" se refiere a un material semirrígido que es permeable a líquidos y gases. Típicamente, el material de gel puede hincharse cuando se absorbe el líquido y puede contraerse cuando se elimina el líquido mediante secado. Los geles ilustrativos incluyen, pero no se limitan a, aquellos que tienen una estructura coloidal, tal como agarosa; estructura de malla polimérica, tal como gelatina; o estructura polimérica reticulada, tal como poliacrilamida, SFA (ver, por ejemplo, la publicación de solicitud de patente de Estados Unidos núm. 2011/0059865 A1) o PAZAM (ver, por ejemplo, la publicación de solicitud de patente de Estados Unidos núm. 2014/0079923 A1). El material de gel particularmente útil se ajustará a la forma de un pocillo u otro elemento cóncavo donde reside.

Como se usa en la presente, los términos "ácido nucleico" y "nucleótido" pretenden ser consistentes con su uso en la técnica e incluir especies de origen natural o análogos funcionales de las mismas. Los análogos funcionales particularmente útiles de ácidos nucleicos son capaces de hibridarse con un ácido nucleico de una manera específica de secuencia o capaces de usarse como una plantilla para la replicación de una secuencia de nucleótidos particular. Los ácidos nucleicos de origen natural generalmente tienen una cadena principal que contiene enlaces fosfodiéster. Una estructura análoga puede tener un enlace de cadena principal alternativo que incluye cualquiera de una variedad de los conocidos en la técnica. Los ácidos nucleicos de origen natural generalmente tienen un azúcar desoxirribosa (por ejemplo, que se encuentra en el ácido desoxirribonucleico (ADN)) o un azúcar ribosa (por ejemplo, que se encuentra en el ácido ribonucleico (ARN)). Un ácido nucleico puede contener nucleótidos que tienen cualquiera de una variedad de análogos de estos restos de azúcar que se conocen en la técnica. Un ácido nucleico puede incluir nucleótidos nativos o no nativos. Con respecto a esto, un ácido desoxirribonucleico nativo puede tener una o más bases seleccionadas del grupo que consiste en adenina, timina, citosina o guanina y un ácido ribonucleico puede tener una o más bases seleccionadas del grupo que consiste en uracilo, adenina, citosina o guanina. En la técnica se conocen bases no nativas útiles que pueden incluirse en un ácido nucleico o nucleótido. Los términos "sonda" o "diana", cuando se usan en referencia a un ácido nucleico o secuencia de un ácido nucleico, se destinan como identificadores semánticos del ácido nucleico o secuencia en el contexto de un método o composición expuesto en la presente descripción y no limitan necesariamente la estructura o función del ácido nucleico o secuencia más allá de lo que se indique explícitamente de cualquier otra manera. Los términos "sonda" y "diana" pueden aplicarse de manera similar a otros analitos tales como proteínas, moléculas pequeñas, células o similares.

Como se usa en la presente, el término "paso", cuando se usa en referencia a los elementos de una matriz, se refiere a la separación de centro a centro para elementos adyacentes. Un patrón de elementos puede caracterizarse en términos de paso promedio. El patrón puede pedirse de manera que el coeficiente de variación alrededor del paso promedio sea pequeño o el patrón puede ser aleatorio en cuyo caso el coeficiente de variación puede ser relativamente grande. En cualquier caso, el paso promedio puede ser, por ejemplo, al menos aproximadamente 10 nm, 0,1 µm, 0,5 µm, 1 µm, 5 µm, 10 µm, 100 µm o más. Alternativa o adicionalmente, el paso promedio puede ser, por ejemplo, a lo máximo aproximadamente 100 µm, 10 µm, 5 µm, 1 µm, 0,5 µm, 0,1 µm o menos. Por supuesto, el paso promedio para un patrón particular de elementos puede estar entre uno de los valores inferiores y uno de los valores superiores seleccionados de los intervalos anteriores.

Como se usa en la presente, el término "poli T o poli A", cuando se usa en referencia a una secuencia de ácido nucleico, se refiere a una serie de dos o más bases de timina (T) o adenina (A), respectivamente. Una poli T o poli A puede incluir al menos aproximadamente 2, 5, 8, 10, 12, 15, 18, 20 o más de las bases T o A, respectivamente. Alternativa o adicionalmente, un poli T o poli A puede incluir a lo máximo aproximadamente, 30, 20, 18, 15, 12, 10, 8, 5 o 2 de las bases T o A, respectivamente.

Como se usa en la presente, el término "aleatorio" puede usarse para referirse a la

disposición o composición espacial de ubicaciones en una superficie. Por ejemplo, hay al menos dos tipos de orden para una matriz descrita en la presente descripción, el primero se relaciona con la separación y la ubicación relativa de los elementos (también denominados "sitios") y el segundo se relaciona con la identidad o el conocimiento predeterminado de la especie particular de molécula que está presente en un elemento particular. En consecuencia, los elementos de una matriz pueden separarse aleatoriamente de manera que los elementos vecinos más cercanos tengan una separación variable entre sí. Alternativamente, la separación entre elementos puede ordenarse, por ejemplo, que forme un patrón regular tal como una rejilla rectilínea o una rejilla hexagonal. En otro sentido, los elementos de una matriz pueden ser aleatorios con respecto a la identidad o el conocimiento predeterminado de la

especie de analito (por ejemplo, ácido nucleico de una secuencia particular) que ocupa cada elemento independientemente de si la separación produce un patrón aleatorio o un patrón ordenado. Una matriz expuesta en la presente descripción puede ser ordenada en un sentido y aleatoria en otro. Por ejemplo, en algunas modalidades expuestas en la presente descripción una superficie se pone en contacto con una población de ácidos nucleicos en condiciones donde los ácidos nucleicos se unen en sitios que se ordenan con respecto a sus ubicaciones relativas pero que 'se ubican aleatoriamente' con respecto al conocimiento de la secuencia para la especie de ácido nucleico presente en algún sitio particular. La referencia a los ácidos nucleicos "que se distribuyen aleatoriamente" en ubicaciones sobre una superficie pretende referirse a la ausencia de conocimiento o ausencia de predeterminación con respecto a qué ácido nucleico se capturará en qué ubicación (independientemente de si las ubicaciones se disponen en un patrón ordenado o no).

Como se usa en la presente, el término "soporte sólido" se refiere a un sustrato rígido que es insoluble en líquido acuoso. El sustrato puede ser no poroso o poroso. El sustrato puede ser capaz opcionalmente de absorber un líquido (por ejemplo, debido a la porosidad) pero será típicamente lo suficientemente rígido como para que el sustrato no se hinche sustancialmente cuando se absorba el líquido y no se contraiga sustancialmente cuando el líquido se elimine mediante secado. Un soporte sólido no poroso es generalmente impermeable a líquidos o gases. Los soportes sólidos ilustrativos incluyen, pero no se limitan a, vidrio y vidrio modificado o funcionalizado, plásticos (que incluyen acrílicos, poliestireno y copolímeros de estireno y otros materiales, polipropileno, polietileno, polibutileno, poliuretanos, Teflon™, olefinas cíclicas, poliimidas, etc.), nailon, cerámicas, resinas, Zeonor, sílice o materiales a base de sílice que incluyen silicio y silicio modificado, carbono, metales, vidrios inorgánicos, conjuntos de fibras ópticas y polímeros. Los soportes sólidos particularmente útiles para algunas modalidades se ubican dentro de un aparato de celda de flujo. Las celdas de flujo ilustrativas se exponen en más detalle en la presente descripción.

Como se usa en la presente, el término "etiqueta espacial" se refiere a un ácido nucleico que tiene una secuencia que es indicativa de una ubicación. Típicamente, el ácido nucleico es una molécula sintética que tiene una secuencia que no se encuentra en uno o más especímenes biológicos que se usarán con el ácido nucleico. Sin embargo, en algunas modalidades la molécula de ácido nucleico se puede obtener naturalmente o la secuencia del ácido nucleico puede ser de origen natural, por ejemplo, en un espécimen biológico que se usa con el ácido nucleico. La ubicación indicada por una etiqueta espacial puede ser una ubicación en o dentro de un espécimen biológico, en o sobre un soporte sólido o una de sus combinaciones. Una secuencia de código de barras puede funcionar como una etiqueta espacial.

Como se usa en la presente, el término "tejido" se refiere a una agregación de células y, opcionalmente, materia intercelular. Típicamente, las células en un tejido no flotan libremente en solución y, en cambio, se unen entre sí para formar una estructura multicelular. Los tipos de tejido ilustrativos incluyen tejidos musculares, nerviosos, epidérmicos y conectivos.

Como se usa en la presente, el término "secuencia universal" se refiere a una serie de nucleótidos que es común a dos o más moléculas de ácido nucleico incluso si las moléculas también tienen regiones de secuencia que difieren entre sí. Una secuencia universal que está presente en diferentes miembros de una colección de moléculas puede permitir la captura de múltiples ácidos nucleicos diferentes mediante el uso de una población de ácidos nucleicos de captura universales que son complementarios a la secuencia universal. De manera similar, una secuencia universal presente en diferentes miembros de una colección de moléculas puede permitir la replicación o amplificación de múltiples ácidos nucleicos diferentes mediante el uso de una población de cebadores universales que son complementarios a la secuencia universal. Por lo tanto, un ácido nucleico de captura universal o un cebador universal incluye una secuencia que puede hibridarse específicamente con una secuencia universal. Las moléculas de ácido nucleico diana pueden modificarse para unir adaptadores universales, por ejemplo, en uno o ambos extremos de las diferentes secuencias diana.

Las modalidades expuestas a continuación y mencionadas en las secciones pueden entenderse en vista de las definiciones anteriores.

La presente descripción proporciona un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico. El método puede incluir las etapas de (a) unir diferentes sondas de ácido nucleico a un soporte sólido para producir sondas ubicadas aleatoriamente sobre el soporte sólido, en donde cada una de las diferentes sondas de ácido nucleico incluye una secuencia de código de barras, y en donde cada una de las sondas ubicadas aleatoriamente incluye secuencias de código de barras diferentes a las de otras sondas ubicadas aleatoriamente sobre el soporte sólido; (b) realizar una reacción de detección de ácidos nucleicos sobre el soporte sólido para determinar las secuencias de código de barras de las sondas ubicadas aleatoriamente sobre el soporte sólido; (c) poner en contacto un espécimen biológico con el soporte sólido que tiene las sondas ubicadas aleatoriamente; (d) hibridar las sondas ubicadas aleatoriamente con ácidos nucleicos diana de porciones del espécimen biológico que están cerca de las sondas ubicadas aleatoriamente; y (e) extender las sondas ubicadas aleatoriamente para producir sondas extendidas que incluyen las secuencias de código de barras y las secuencias de los ácidos nucleicos diana, para de esta manera etiquetar espacialmente los ácidos nucleicos del espécimen biológico.

Cualquiera de una variedad de soportes sólidos puede usarse en un método, composición o aparato de la presente descripción. Los soportes sólidos particularmente útiles son los usados para matrices de ácidos nucleicos. Los



ejemplos incluyen vidrio, vidrio modificado, vidrio funcionalizado, vidrios inorgánicos, microesferas (por ejemplo, partículas inertes y/o magnéticas), plásticos, polisacáridos, nailon, nitrocelulosa, cerámicas, resinas, sílice, materiales a base de sílice, carbono, metales, una fibra óptica o conjuntos de fibras ópticas, polímeros y placas de múltiples pocillos (por ejemplo, de microtitulación). Los plásticos ilustrativos incluyen acrílicos, poliestireno, copolímeros de estireno y otros materiales, polipropileno, polietileno, polibutileno, poliuretanos y Teflon™. Los materiales a base de sílice ilustrativos incluyen silicio y varias formas de silicio modificado.

En modalidades particulares, un soporte sólido puede estar dentro de o ser parte de un recipiente tal como un pocillo, tubo, canal, cubeta, placa Petri, botella o similares. Un recipiente particularmente útil es una celda de flujo, por ejemplo, como se describe en el documento WO 2014/142841 A1; la publicación de solicitud de patente de Estados Unidos núm. 2010/0111768 A1 y la patente de Estados Unidos núm. 8,951,781 o Bentley y otros, *Nature* 456:53-59 (2008). Las celdas de flujo ilustrativas son aquellas que están disponibles comercialmente de Illumina, Inc. (San Diego, CA) para usar con una plataforma de secuenciación tal como una plataforma Genome Analyzer®, MiSeq®, NextSeq® o HiSeq®. Otro recipiente particularmente útil es un pocillo en una placa de múltiples pocillos o placa de microtitulación.

Opcionalmente, un soporte sólido puede incluir un recubrimiento de gel. La unión de ácidos nucleicos a un soporte sólido por medio de un gel se ejemplifica mediante celdas de flujo disponibles comercialmente de Illumina Inc. (San Diego, CA) o descritas en las solicitudes de patente de Estados Unidos núm. 2011/0059865 A1, 2014/0079923 A1 o 2015/0005447 A1; o la publicación PCT núm. WO 2008/093098. Los geles ilustrativos que pueden usarse en los métodos y aparatos expuestos en la presente descripción incluyen, pero no se limitan a, aquellos que tienen una estructura coloidal, tal como agarosa; estructura de malla polimérica, tal como gelatina; o estructura polimérica reticulada, tal como poliacrilamida, SFA (ver, por ejemplo, la publicación de solicitud de patente de Estados Unidos núm. 2011/0059865 A1) o PAZAM (ver, por ejemplo, las publicaciones de solicitud de patente de Estados Unidos núm. 2014/0079923 A1 o 2015/0005447 A1).

En algunas modalidades, un soporte sólido puede configurarse como una matriz de elementos a los que pueden unirse los ácidos nucleicos. Los elementos pueden estar presentes en cualquiera de una variedad de formatos deseados. Por ejemplo, los elementos pueden ser pocillos, hoyitos, canales, crestas, regiones elevadas, clavijas, postes o similares. En algunas modalidades, los elementos pueden contener perlas. Sin embargo, en modalidades particulares no es necesario que los elementos contengan una perla o partícula. Los elementos ilustrativos incluyen pocillos que están presentes en sustratos usados para plataformas de secuenciación comerciales vendidas por 454 LifeSciences (una subsidiaria de Roche, Basilea, Suiza) o Ion Torrent (una subsidiaria de Life Technologies, Carlsbad California). Otros sustratos que tienen pocillos incluyen, por ejemplo, fibra óptica grabada y otros sustratos descritos en las patentes de Estados Unidos núm. 6,266,459; 6,355,431; 6,770,441; 6,859,570; 6,210,891; 6,258,568; 6,274,320; las publicaciones de solicitud de patente de Estados Unidos núm. 2009/0026082 A1; 2009/0127589 A1; 2010/0137143 A1; 2010/0282617 A1 o la publicación PCT núm. WO 00/63437. En algunas modalidades, los pocillos de un sustrato pueden incluir material de gel (con o sin perlas) como se expone en la publicación de solicitud de patente de Estados Unidos núm. 2014/0243224 A1.

Los elementos de un soporte sólido pueden ser elementos metálicos en una superficie no metálica tal como vidrio, plástico u otros materiales ejemplificados anteriormente. Una capa metálica puede depositarse sobre una superficie mediante el uso de métodos conocidos en la técnica tales como grabado con plasma en húmedo, grabado con plasma en seco, deposición de capas atómicas, grabado por haz de iones, deposición de vapor químico, pulverización catódica al vacío o similares. Cualquiera de una variedad de instrumentos comerciales puede usarse según corresponda, que incluyen, por ejemplo, los sistemas FlexAL®, OpAL®, Ionfab 300plus® u Optofab 3000® (Oxford Instruments, Reino Unido). Una capa metálica también puede depositarse mediante evaporación por haz de electrones o pulverización catódica como se establece en Thornton, *Ann. Rev. Mater. Sci.* 7:239-60 (1977). Las técnicas de deposición de capas metálicas, tales como las ejemplificadas anteriormente, pueden combinarse con técnicas de fotolitografía para crear regiones o parches metálicos sobre una superficie. Los métodos ilustrativos para combinar técnicas de deposición de capas metálicas y técnicas de fotolitografía se proporcionan en la patente de Estados Unidos núm. 8,895,249 o la publicación de solicitud de patente de Estados Unidos núm. 2014/0243224 A1.

Los elementos pueden aparecer sobre un soporte sólido como una rejilla de puntos o parches. Los elementos pueden ubicarse en un patrón repetitivo o en un patrón irregular no repetitivo. Los patrones repetitivos particularmente útiles son patrones hexagonales, patrones rectilíneos, patrones de rejilla, patrones que tienen simetría reflectante, patrones que tienen simetría giratoria o similares. Los patrones asimétricos también pueden ser útiles. El paso puede ser el mismo entre diferentes pares de elementos vecinos más cercanos o el paso puede variar entre diferentes pares de elementos vecinos más cercanos.

Las matrices de densidad alta se caracterizan por tener un paso promedio de menos de aproximadamente 15 µm. Las matrices de densidad media tienen un paso promedio de aproximadamente 15 a 30 µm, mientras que las matrices de densidad baja tienen un paso promedio mayor que 30 µm. Una matriz útil en la invención puede tener un paso promedio que es menor que 100 µm, 50 µm, 10 µm, 5 µm, 1 µm o 0,5 µm. Se pretende que los valores de paso promedio y los intervalos expuestos anteriormente o en cualquier otro lugar en la presente descripción sean aplicables a matrices ordenadas o matrices aleatorias.

En modalidades particulares, cada uno de los elementos sobre un soporte sólido puede tener un área que es mayor que aproximadamente 100 nm<sup>2</sup>, 250 nm<sup>2</sup>, 500 nm<sup>2</sup>, 1 μm<sup>2</sup>, 2,5 μm<sup>2</sup>, 5 μm<sup>2</sup>, 10 μm<sup>2</sup>, 100 μm<sup>2</sup> o 500 μm<sup>2</sup>. Alternativa o adicionalmente, cada uno de los elementos puede tener un área que es menor que aproximadamente 1 mm<sup>2</sup>, 500 μm<sup>2</sup>, 100 μm<sup>2</sup>, 25 μm<sup>2</sup>, 10 μm<sup>2</sup>, 5 μm<sup>2</sup>, 1 μm<sup>2</sup>, 500 nm<sup>2</sup> o 100 nm<sup>2</sup>. Los intervalos anteriores pueden describir el área aparente de una perla u otra partícula sobre un soporte sólido cuando se observa o se obtienen imágenes desde arriba.

En modalidades particulares, un soporte sólido puede incluir una colección de perlas u otras partículas. Las partículas pueden suspenderse en una solución o pueden ubicarse en la superficie de un sustrato. Los ejemplos de matrices que tienen perlas ubicadas en una superficie incluyen aquellas en donde las perlas se ubican en pocillos tales como una matriz BeadChip (Illumina Inc., San Diego CA), sustratos usados en plataformas de secuenciación de 454 LifeSciences (una subsidiaria de Roche, Basilea Suiza) o sustratos usados en plataformas de secuenciación de Ion Torrent (una subsidiaria de Life Technologies, Carlsbad California). Otros soportes sólidos que tienen perlas ubicadas en una superficie se describen en las patentes de Estados Unidos núms. 6,266,459; 6,355,431; 6,770,441; 6,859,570; 6,210,891; 6,258,568; o 6,274,320; las publicaciones de solicitud de patente de Estados Unidos núm. 2009/0026082 A1; 2009/0127589 A1; 2010/0137143 A1; o 2010/0282617 A1 o la publicación PCT núm. WO 00/63437. Varias de las referencias anteriores describen métodos para unir sondas de ácido nucleico a perlas antes de cargar las perlas en o sobre un soporte sólido. Como tal, la colección de perlas puede incluir diferentes perlas, cada una de las cuales tiene una sonda única unida. Sin embargo, se entenderá que las perlas pueden prepararse de modo que incluyan cebadores universales, y las perlas pueden cargarse después en una matriz, para forma de esta manera matrices universales para su uso en un método expuesto en la presente descripción. Como se expuso anteriormente en la presente descripción, los soportes sólidos usados típicamente para matrices de perlas pueden usarse sin perlas. Por ejemplo, los ácidos nucleicos, tales como sondas o cebadores, pueden unirse directamente a los pocillos o al material de gel en los pocillos. Por lo tanto, las referencias anteriores son ilustrativas de materiales, composiciones o aparatos que pueden modificarse para su uso en los métodos y composiciones expuestos en la presente descripción.

En consecuencia, un soporte sólido usado en un método expuesto en la presente descripción puede incluir una matriz de perlas, en donde diferentes sondas de ácido nucleico se unen a diferentes perlas en la matriz. En esta modalidad, cada perla puede unirse a una sonda de ácido nucleico diferente y las perlas pueden distribuirse aleatoriamente sobre el soporte sólido para unir de manera efectiva las diferentes sondas de ácido nucleico al soporte sólido. Opcionalmente, el soporte sólido puede incluir pocillos que tienen dimensiones que no alojan más de una única perla. En una configuración de este tipo, las perlas pueden unirse a los pocillos debido a las fuerzas resultantes del ajuste de las perlas en los pocillos. También es posible usar químicos o adhesivos de unión para contener las perlas en los pocillos.

Las sondas de ácido nucleico que se unen a las perlas pueden incluir secuencias de código de barras. Una población de las perlas puede configurarse de manera que cada perla se una a solo un tipo de código de barras y muchas perlas diferentes cada una con un código de barras diferente están presentes en la población. En esta modalidad, la distribución aleatoria de las perlas a un soporte sólido dará como resultado la ubicación aleatoria de las sondas de ácido nucleico (y sus secuencias de código de barras respectivas) sobre el soporte sólido. En algunos casos puede haber múltiples perlas con la misma secuencia de código de barras de manera que haya redundancia en la población. La distribución aleatoria de una población redundante de perlas sobre un soporte sólido que tiene una capacidad que es mayor que el número de códigos de barras únicos en la población de perlas dará como resultado la redundancia de códigos de barras sobre el soporte sólido. Alternativamente, el número de códigos de barras diferentes en una población de perlas puede exceder la capacidad del soporte sólido para producir una matriz que no sea redundante con respecto a la población de códigos de barras sobre el soporte sólido. La capacidad del soporte sólido estará determinada en algunas modalidades por el número de elementos (por ejemplo, pocillos con capacidad para una perla) que unen o alojan de cualquier otra manera una perla.

Un soporte sólido puede incluir, o puede prepararse mediante los métodos expuestos en la presente descripción para unir, una pluralidad de diferentes sondas de ácido nucleico. Por ejemplo, un soporte sólido puede incluir al menos 10, 100, 1 x 10<sup>3</sup>, 1 x 10<sup>4</sup>, 1 x 10<sup>5</sup>, 1 x 10<sup>6</sup>, 1 x 10<sup>7</sup>, 1 x 10<sup>8</sup>, 1 x 10<sup>9</sup> o más sondas diferentes. Alternativa o adicionalmente, un soporte sólido puede incluir a lo máximo 1 x 10<sup>9</sup>, 1 x 10<sup>8</sup>, 1 x 10<sup>7</sup>, 1 x 10<sup>6</sup>, 1 x 10<sup>5</sup>, 1 x 10<sup>4</sup>, 1 x 10<sup>3</sup>, 100 o menos sondas diferentes. Se debe entender que cada una de las diferentes sondas puede estar presente en varias copias, por ejemplo, cuando las sondas se han amplificado para formar un conglomerado. Por lo tanto, los intervalos anteriores pueden describir el número de conglomerados de ácidos nucleicos diferentes sobre un soporte sólido. También se entenderá que los intervalos anteriores pueden describir el número de códigos de barras, secuencias de captura de diana u otros elementos de secuencia diferentes expuestos en la presente descripción como únicos para sondas de ácido nucleico particulares. Alternativa o adicionalmente, los intervalos pueden describir el número de sondas extendidas o sondas modificadas creadas sobre un soporte sólido mediante el uso de un método expuesto en la presente descripción.

Los elementos, pueden estar presentes sobre un soporte sólido antes de poner en contacto el soporte sólido con sondas de ácido nucleico. Por ejemplo, en modalidades donde las sondas se unen a un soporte por medio de hibridación a cebadores, los cebadores pueden unirse en los elementos, mientras que las áreas intersticiales fuera de los elementos carecen sustancialmente de ninguno de los cebadores. Las sondas de ácido nucleico pueden capturarse en elementos preformados sobre un soporte sólido, y opcionalmente amplificarse sobre el soporte sólido, mediante el uso de métodos expuestos en la patente de Estados Unidos núm. 8,895,249, la patente de Estados Unidos núm.

8,778,849, o la publicación de solicitud de patente de Estados Unidos núm. 2014/0243224 A1. Alternativamente, un soporte sólido puede tener un césped de cebadores o de cualquier otra manera puede carecer de elementos. En este caso, puede formarse un elemento en virtud de la unión de una sonda de ácido nucleico sobre el soporte sólido. Opcionalmente, la sonda de ácido nucleico capturada puede amplificarse sobre el soporte sólido de manera que el conglomerado resultante se convierta en un elemento. Aunque la unión se ejemplifica anteriormente como captura entre un cebador y una porción complementaria de una sonda, se debe entender que restos de captura distintos de los cebadores pueden estar presentes en elementos preformados o como un césped. Otros restos de captura ilustrativos incluyen, pero no se limitan a, restos químicos capaces de reaccionar con una sonda de ácido nucleico para crear un enlace covalente o receptores capaces de unirse no covalentemente a un ligando en una sonda de ácido nucleico.

Una etapa de unir sondas de ácido nucleico a un soporte sólido puede llevarse a cabo al proporcionar un fluido que contiene una mezcla de diferentes sondas de ácido nucleico y poner en contacto esta mezcla de fluidos con el soporte sólido. El contacto puede dar como resultado que la mezcla de fluidos esté en contacto con una superficie a la que se unirán muchas sondas de ácido nucleico diferentes de la mezcla de fluidos. Por lo tanto, las sondas tienen acceso aleatorio a la superficie (ya sea que la superficie tenga elementos preformados configurados para unirse a las sondas o una superficie uniforme configurada para la unión). En consecuencia, las sondas pueden ubicarse aleatoriamente sobre el soporte sólido.

El número total y la variedad de sondas diferentes que terminan unidas a una superficie pueden seleccionarse para una aplicación o uso particular. Por ejemplo, en modalidades donde una mezcla de fluidos de diferentes sondas de ácido nucleico se pone en contacto con un soporte sólido con el propósito de unir las sondas al soporte, el número de especies de sondas diferentes puede exceder la capacidad del soporte sólido para sondas. Por lo tanto, el número y variedad de sondas diferentes que se unen al soporte sólido puede ser equivalente a la capacidad del soporte sólido para sondas. Alternativamente, el número y variedad de especies de sondas diferentes sobre el soporte sólido puede ser menor que la capacidad (es decir, habrá redundancia de especies de sondas de manera que el soporte sólido puede contener múltiples elementos que tienen la misma especie de sonda). Tal redundancia puede lograrse, por ejemplo, al poner en contacto el soporte sólido con una mezcla de fluidos que contiene un número y variedad de especies de sonda que es sustancialmente menor que la capacidad del soporte sólido para sondas.

La unión de las sondas de ácido nucleico puede estar mediada por la hibridación de las sondas de ácido nucleico a cebadores complementarios que se unen al soporte sólido, la formación de enlaces químicos entre un resto reactivo en la sonda de ácido nucleico y el soporte sólido (se exponen ejemplos en la patente de Estados Unidos núm. 8,895,249, la patente de Estados Unidos núm. 8,778,849 o la publicación de solicitud de patente de Estados Unidos núm. 2014/0243224 A1), interacciones de afinidad de un resto en la sonda de ácido nucleico con un resto unido al soporte sólido (por ejemplo, entre pares de receptor-ligando conocidos tales como estreptavidina-biotina, anticuerpo-epitopo, lectina-carbohidrato y similares), interacciones físicas de las sondas de ácido nucleico con el soporte sólido (por ejemplo, unión de hidrógeno, fuerzas iónicas, fuerzas de van der Waals y similares), u otras interacciones conocidas en la técnica para unir ácidos nucleicos a superficies.

En algunas modalidades, la unión de una sonda de ácido nucleico es inespecífica con respecto a cualquier diferencia de secuencia entre la sonda de ácido nucleico y otras sondas de ácido nucleico que se unen o se unirán al soporte sólido. Por ejemplo, diferentes sondas pueden tener una secuencia universal que se complementa con los cebadores unidos a la superficie o las diferentes sondas pueden tener un resto común que media la unión a la superficie. Alternativamente, cada una de las diferentes sondas (o una subpoblación de diferentes sondas) puede tener una secuencia única que se complementa con un cebador único sobre el soporte sólido o pueden tener un resto único que interactúa con uno o más restos reactivos diferentes sobre el soporte sólido. En tales casos, los cebadores únicos o restos únicos pueden, opcionalmente, unirse en ubicaciones predefinidas para capturar selectivamente sondas particulares, o tipos particulares de sondas, en las ubicaciones predefinidas respectivas.

Cada uno de uno o más elementos sobre un soporte sólido puede incluir una sola molécula de una sonda particular. Los elementos pueden configurarse, en algunas modalidades, para alojar no más de una sola molécula de sonda de ácido nucleico. Sin embargo, ya sea que el elemento pueda o no alojar más de una molécula de sonda de ácido nucleico, el elemento puede no obstante incluir no más de una sola molécula de sonda de ácido nucleico. Alternativamente, un elemento individual puede incluir una pluralidad de moléculas de sonda de ácido nucleico, por ejemplo, un conjunto de moléculas de sonda de ácido nucleico que tienen la misma secuencia entre sí. En modalidades particulares, el conjunto puede producirse mediante amplificación a partir de una sola plantilla de sonda de ácido nucleico para producir amplicones, por ejemplo, como un conglomerado unido a la superficie.

Un método expuesto en la presente descripción puede usar cualquiera de una variedad de técnicas de amplificación. Las técnicas ilustrativas que pueden usarse incluyen, pero no se limitan a, reacción en cadena de la polimerasa (PCR), amplificación de círculo rodante (RCA), amplificación por desplazamiento múltiple (MDA) o amplificación de cebado aleatorio (RPA). En algunas modalidades la amplificación puede llevarse a cabo en solución, por ejemplo, cuando los elementos de una matriz son capaces de contener amplicones en un volumen que tiene una capacidad deseada. Preferentemente, una técnica de amplificación usada en un método de la presente descripción se llevará a cabo en fase sólida. Por ejemplo, una o más especies de cebadores (por ejemplo, cebadores universales para uno o más sitios de unión a cebadores universales presentes en una sonda de ácido nucleico) pueden unirse a un soporte sólido. En

modalidades de PCR, uno o ambos de los cebadores usados para la amplificación pueden unirse a un soporte sólido (por ejemplo, por medio de un gel). Los formatos que utilizan dos especies de cebadores unidos a un soporte sólido se denominan a menudo amplificación puente porque los amplicones de doble hebra forman una estructura tipo puente entre los dos cebadores unidos a la superficie que flanquean la secuencia plantilla que se ha copiado. Los reactivos y condiciones ilustrativos que pueden usarse para la amplificación puente se describen, por ejemplo, en las patentes de Estados Unidos núm. 5,641,658, 7,115,400 o 8,895,249; o las publicaciones de patente de Estados Unidos núm. 2002/0055100 A1, 2004/0096853 A1, 2004/0002090 A1, 2007/0128624 A1 o 2008/0009420 A1. La amplificación por PCR en fase sólida también puede llevarse a cabo con uno de los cebadores de amplificación unidos a un soporte sólido y el segundo cebador en solución. Un formato ilustrativo que usa una combinación de un cebador unido a la superficie y un cebador soluble es el formato usado en la PCR en emulsión como se describió, por ejemplo, en Dressman y otros, *Proc. Natl. Acad. Sci. USA* 100:8817-8822 (2003), WO 05/010145, o las publicaciones de solicitud de patente de Estados Unidos núm. 2005/0130173 A1 o 2005/0064460 A1. La PCR en emulsión es ilustrativa del formato y se debe entender que para los propósitos de los métodos expuestos en la presente descripción el uso de una emulsión es opcional y de hecho para varias modalidades no se usa una emulsión.

Las técnicas de RCA pueden modificarse para su uso en un método de la presente descripción. Los componentes ilustrativos que pueden usarse en una reacción de RCA y los principios mediante los cuales RCA produce amplicones se describen, por ejemplo, en Lizardi y otros, *Nat. Genet.* 19:225-232 (1998) y la publicación de solicitud de patente de Estados Unidos núm. 2007/0099208 A1. Los cebadores usados para RCA pueden estar en solución o unidos a un soporte sólido. Los cebadores pueden ser uno o más de los cebadores universales descritos en la presente descripción.

Las técnicas de MDA pueden modificarse para su uso en un método de la presente descripción. Algunos principios básicos y condiciones útiles para MDA se describen, por ejemplo, en Dean y otros, *Proc Natl. Acad. Sci. USA* 99:5261-66 (2002); Lage y otros, *Genome Research* 13:294-307 (2003); Walker y otros, *Molecular Methods for Virus Detection*, Academic Press, Inc., 1995; Walker y otros, *Nucl. Acids Res.* 20:1691-96 (1992); los documentos US 5,455,166; US 5,130,238; y US 6,214,587. Los cebadores usados para MDA pueden estar en solución o unidos a un soporte sólido en un sitio de amplificación. Nuevamente, los cebadores pueden ser uno o más de los cebadores universales descritos en la presente descripción.

En modalidades particulares puede usarse una combinación de las técnicas de amplificación ejemplificadas anteriormente. Por ejemplo, RCA y MDA pueden usarse en una combinación en donde RCA se usa para generar un amplicón concatamérico en solución (por ejemplo, mediante el uso de cebadores en fase de solución). Después el amplicón puede usarse como una plantilla para MDA mediante el uso de cebadores que se unen a un soporte sólido (por ejemplo, cebadores universales). En este ejemplo, los amplicones producidos después de las etapas combinadas de RCA y MDA se unirán al soporte sólido.

Las sondas de ácido nucleico que se usan en un método expuesto en la presente descripción o presentes en un aparato o composición de la presente descripción pueden incluir secuencias de código de barras, y para modalidades que incluyen una pluralidad de diferentes sondas de ácido nucleico, cada una de las sondas puede incluir una secuencia de código de barras diferente a las de otras sondas en la pluralidad. Las secuencias de código de barras pueden tener cualquiera de una variedad de longitudes. Las secuencias más largas generalmente pueden alojar un mayor número y variedad de códigos de barras para una población. Generalmente, todas las sondas en una pluralidad tendrán códigos de barras de la misma longitud (aunque con secuencias diferentes), pero también es posible usar códigos de barras de longitud diferente para diferentes sondas. Una secuencia de código de barras puede tener al menos 2, 4, 6, 8, 10, 12, 15, 20 o más nucleótidos de longitud. Alternativa o adicionalmente, la longitud de la secuencia de código de barras puede ser a lo máximo 20, 15, 12, 10, 8, 6, 4 o menos nucleótidos. Los ejemplos de secuencias de código de barras que pueden usarse se exponen, por ejemplo, en la publicación de solicitud de patente de Estados Unidos núm. 2014/0342921 A1 y la patente de Estados Unidos núm. 8,460,865.

Un método de la presente descripción puede incluir una etapa de realizar una reacción de detección de ácidos nucleicos sobre un soporte sólido para determinar secuencias de código de barras de sondas de ácido nucleico que se ubican sobre el soporte sólido. En muchas modalidades las sondas se ubican aleatoriamente sobre el soporte sólido y la reacción de detección de ácidos nucleicos proporciona información para ubicar cada una de las diferentes sondas. Los métodos de detección de ácidos nucleicos ilustrativos incluyen, pero no se limitan a la secuenciación de ácidos nucleicos de una sonda, hibridación de ácidos nucleicos con una sonda, ligazón de ácidos nucleicos que se hibridan con una sonda, extensión de ácidos nucleicos que se hibridan con una sonda, extensión de un primer ácido nucleico que se hibrida con una sonda seguido de la ligazón del ácido nucleico extendido a un segundo ácido nucleico que se hibrida con la sonda, u otros métodos conocidos en la técnica tales como los expuestos en las patentes de Estados Unidos núm. 8,288,103 u 8,486,625.

Las técnicas de secuenciación, tales como las técnicas de secuenciación por síntesis (SBS), son un método particularmente útil para determinar secuencias de código de barras. La SBS puede llevarse a cabo de la siguiente manera. Para iniciar un primer ciclo de SBS, uno o más nucleótidos marcados, ADN polimerasa, cebadores de SBS, etc., pueden ponerse en contacto con uno o más elementos sobre un soporte sólido (por ejemplo, elemento(s) donde las sondas de ácido nucleico se unen al soporte sólido). Se pueden detectar aquellos elementos donde la extensión del cebador de SBS provoca que se incorpore un nucleótido marcado. Opcionalmente, los nucleótidos pueden incluir

un resto de terminación reversible que termina con una extensión del cebador adicional una vez que se ha añadido un nucleótido al cebador de SBS. Por ejemplo, un análogo de nucleótido que tiene un resto terminador reversible puede añadirse a un cebador de manera que la extensión subsecuente no pueda ocurrir hasta que se suministre un agente de desbloqueo para eliminar el resto. Por lo tanto, para modalidades que usan terminación reversible, puede suministrarse un reactivo de desbloqueo al soporte sólido (antes o después de que ocurra la detección). Los lavados pueden llevarse a cabo entre las diversas etapas de suministro. Después el ciclo puede repetirse n veces para extender el cebador en n nucleótidos, de esta manera se detecta una secuencia de longitud n. Los procedimientos de SBS, sistemas de fluidos y plataformas de detección ilustrativos que pueden adaptarse fácilmente para su uso con una composición, aparato o método de la presente descripción se describen, por ejemplo, en Bentley y otros, *Nature* 456:53-59 (2008), las publicaciones PCT núm. WO 91/06678, WO 04/018497 o WO 07/123744; las patentes de Estados Unidos núm. 7,057,026, 7,329,492, 7,211,414, 7,315,019 o 7,405,281, y la publicación de solicitud de patente de Estados Unidos núm. 2008/0108082.

Pueden usarse otros procedimientos de secuenciación que usan reacciones cíclicas, tales como pirosecuenciación. La pirosecuenciación detecta la liberación de pirofosfato inorgánico (PPi) a medida que se incorporan nucleótidos particulares en una hebra de ácido nucleico naciente (Ronaghi, y otros, *Analytical Biochemistry* 242(1), 84-9 (1996); Ronaghi, *Genome Res.* 11(1), 3-11 (2001); Ronaghi y otros *Science* 281(5375), 363 (1998); o las patentes de Estados Unidos núm. 6,210,891, 6,258,568 o 6,274,320). En la pirosecuenciación, el PPi liberado puede detectarse al convertirse inmediatamente en trifosfato de adenosina (ATP) por la ATP sulfurilasa, y el nivel de ATP generado puede detectarse por medio de fotones producidos por luciferasa. Por lo tanto, la reacción de secuenciación puede monitorearse por medio de un sistema de detección de luminiscencia. Las fuentes de radiación de excitación usadas para sistemas de detección basados en fluorescencia no son necesarias para los procedimientos de pirosecuenciación. Los sistemas de fluidos, detectores y procedimientos útiles que pueden usarse para la aplicación de pirosecuenciación a aparatos, composiciones o métodos de la presente descripción se describen, por ejemplo, en la publicación de solicitud de patente PCT núm. WO2012/058096, la publicación de solicitud de patente de Estados Unidos núm. 2005/0191698 A1, o las patentes de Estados Unidos núm. 7,595,883 o 7,244,559.

Las reacciones de secuenciación por ligazón también son útiles, que incluyen, por ejemplo, las descritas en Shendure y otros *Science* 309:1728-1732 (2005); o las patentes de Estados Unidos núm. 5,599,675 o 5,750,341. Algunas modalidades pueden incluir procedimientos de secuenciación por hibridación como se describió, por ejemplo, en Bains y otros, *Journal of Theoretical Biology* 135(3), 303-7 (1988); Drmanac y otros, *Nature Biotechnology* 16, 54-58 (1998); Fodor y otros, *Science* 251(4995), 767-773 (1995); o la publicación de solicitud de patente de Estados Unidos núm. WO 1989/10977. En los procedimientos de secuenciación por ligazón y secuenciación por hibridación, los ácidos nucleicos diana (o amplicones de los mismos) que están presentes en los sitios de una matriz se someten a ciclos repetidos de suministro y detección de oligonucleótidos. Las composiciones, aparatos o métodos expuestos en la presente descripción o en las referencias mencionadas en la presente descripción pueden adaptarse fácilmente para procedimientos de secuenciación por ligazón o secuenciación por hibridación. Típicamente, los oligonucleótidos se marcan con fluorescencia y pueden detectarse mediante el uso de detectores de fluorescencia similares a los descritos con respecto a los procedimientos de SBS en la presente descripción o en las referencias mencionadas en la presente descripción.

Algunas modalidades de secuenciación pueden utilizar métodos que involucran el monitoreo en tiempo real de la actividad de la ADN polimerasa. Por ejemplo, las incorporaciones de nucleótidos pueden detectarse a través de interacciones de transferencia de energía de resonancia de fluorescencia (FRET) entre una polimerasa portadora de fluoróforo y nucleótidos marcados en  $\gamma$ -fosfato, o con guías de onda de modo cero (ZMW). Se describen técnicas y reactivos para la secuenciación basada en FRET, por ejemplo, en Levene y otros *Science* 299, 682-686 (2003); Lundquist y otros *Opt. Lett.* 33, 1026-1028 (2008); Korlach y otros *Proc. Natl. Acad. Sci. USA* 105, 1176-1181 (2008).

Algunas modalidades de secuenciación incluyen la detección de un protón liberado tras la incorporación de un nucleótido en un producto de extensión. Por ejemplo, la secuenciación basada en la detección de protones liberados puede usar un detector eléctrico y técnicas asociadas que están disponibles comercialmente en Ion Torrent (Guilford, CT, una subsidiaria de Life Technologies y Thermo Fisher) o métodos y sistemas de secuenciación descritos en las publicaciones de solicitud de patente de Estados Unidos núm. 2009/0026082 A1; 2009/0127589 A1; 2010/0137143 A1; o US 2010/0282617 A1.

Las técnicas de hibridación de ácido nucleico también son un método útil para determinar secuencias de código de barras. En algunos casos, pueden usarse métodos combinatorios de hibridación, tales como los usados para la decodificación de matrices de perlas en formato múltiple (ver, por ejemplo, la patente de Estados Unidos núm. 8,460,865). Tales métodos utilizan sondas decodificadoras de ácido nucleico marcadas que son complementarias a al menos una porción de una secuencia de código de barras. Una reacción de hibridación puede llevarse a cabo mediante el uso de sondas decodificadoras que tienen marcadores conocidos de manera que la ubicación donde los marcadores terminan en el soporte sólido identifica las sondas de ácido nucleico de acuerdo con las reglas de complementariedad de ácidos nucleicos. En algunos casos, se usan mezclas de muchas sondas diferentes con marcadores distinguibles, lo que permite de esta manera una operación de decodificación en formato múltiple. El número de códigos de barras diferentes determinado en una operación de decodificación puede exceder el número de marcadores usados para la operación de decodificación. Por ejemplo, la decodificación puede llevarse a cabo en varias etapas donde cada etapa constituye la hibridación con una mezcla diferente de sondas decodificadoras. Las

mismas sondas decodificadoras pueden estar presentes en diferentes mezclas, pero el marcador que está presente en cada sonda decodificadora puede diferir de mezcla a mezcla (es decir, cada sonda decodificadora está en un "estado" diferente cuando está en mezclas diferentes). Pueden usarse varias combinaciones de estos estados y etapas para expandir el número de códigos de barras que pueden decodificarse mucho más allá del número de marcadores distintos disponibles para la decodificación. Tales métodos combinatorios se exponen con más detalle en la patente de Estados Unidos núm. 8,460,865 o en Gunderson y otros, *Genome Research* 14:870-877 (2004).

Un método de la presente descripción puede incluir una etapa de poner en contacto un espécimen biológico con un soporte sólido que tiene sondas de ácido nucleico unidas a este. En algunas modalidades, las sondas de ácido nucleico se ubican aleatoriamente sobre el soporte sólido. La identidad y ubicación de las sondas de ácido nucleico pueden haberse decodificado antes de poner en contacto el espécimen biológico con el soporte sólido. Alternativamente, la identidad y ubicación de las sondas de ácido nucleico pueden determinarse después de poner en contacto el soporte sólido con el espécimen biológico.

En algunas modalidades, el espécimen biológico es una o más células. La(s) célula(s) puede(n) ser individual(es) y libre(s) de cualquier tejido o estructura multicelular en el momento en que se hace el contacto con el soporte sólido. Por ejemplo, la(s) célula(s) puede(n) estar presente(s) en un fluido (por ejemplo, cuando están presentes una pluralidad de diferentes células, el fluido puede ser una mezcla de fluidos de las diferentes células) y el fluido puede ponerse en contacto con el soporte sólido al que se unen las diferentes sondas. Puede usarse cualquiera de una variedad de células que incluyen, por ejemplo, las de un procarionta, arquea o eucarionta. Una o más células usadas en un método, composición o aparato de la presente descripción pueden ser organismos unicelulares o de un organismo multicelular. Los organismos ilustrativos a partir de los cuales se puede obtener una o más células incluyen, pero no se limitan a, un mamífero, planta, alga, nemátodo, insecto, pez, reptil, anfibio, hongo o *Plasmodium falciparum*. Las especies ilustrativas se exponen anteriormente en la presente descripción o se conocen en la técnica.

Las modalidades de la presente descripción también pueden usar uno o más componentes subcelulares como un espécimen biológico. Por ejemplo, una mezcla de fluidos puede incluir uno o más núcleos, aparatos de Golgi, mitocondrias, cloroplastos, fracciones de membrana, vesículas, retículo endoplasmático u otros componentes conocidos en la técnica. Otros tipos útiles de especímenes biológicos son uno o más virus o unos viroides.

Se debe entender que un espécimen biológico puede ser un cultivo o población homogénea de las células, componentes subcelulares, virus o viroides anteriores. Alternativamente, el espécimen biológico puede ser una colección no homogénea de células, componentes subcelulares, virus o viroides, por ejemplo, derivados de varios organismos diferentes en una comunidad o ecosistema. Una comunidad ilustrativa es la colección de bacterias presentes en el sistema digestivo, el pulmón u otro órgano de un organismo multicelular, tal como un mamífero.

Una o más células, componentes subcelulares, virus o viroides que se ponen en contacto con un soporte sólido en un método expuesto en la presente descripción pueden unirse al soporte sólido. La unión puede lograrse mediante el uso de métodos conocidos en la técnica tales como los ejemplificados en la presente descripción con respecto a la unión de ácidos nucleicos a un soporte sólido. En algunas modalidades, la unión es selectiva para tipos específicos de células, componentes subcelulares, virus o viroides. Por ejemplo, el soporte sólido puede incluir anticuerpos u otros receptores que son selectivos para epítopos o ligandos presentes en uno o un subconjunto de diferentes células, componentes subcelulares, virus o viroides presentes en una mezcla de fluidos. En otras modalidades, la unión de células, componentes subcelulares, virus o viroides puede estar mediada por restos no selectivos tales como restos químicos que son ampliamente reactivos.

En modalidades particulares, una o más células, componentes subcelulares, virus o viroides que se han puesto en contacto con un soporte sólido pueden lisarse para liberar ácidos nucleicos diana. La lisis puede llevarse a cabo mediante el uso de métodos conocidos en la técnica tales como aquellos que emplean uno o más de tratamiento químico, tratamiento enzimático, electroporación, calor, tratamiento hipotónico, sonicación o similares. Las técnicas de lisis ilustrativas se exponen en Sambrook y otros, *Molecular Cloning: A Laboratory Manual, Tercera Ed.*, Cold Spring Harbor Laboratory, Nueva York (2001) y en Ansubel y otros, *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999).

En algunas modalidades el espécimen biológico es una sección de tejido. El tejido puede derivarse de un organismo multicelular tal como los ejemplificados anteriormente con respecto a las células. Una sección de tejido puede ponerse en contacto con un soporte sólido, por ejemplo, al colocar el tejido sobre la superficie del soporte sólido. El tejido puede ser recién escindido de un organismo o puede haberse conservado anteriormente, por ejemplo, mediante congelación, incrustación en un material tal como parafina (por ejemplo, muestras incrustadas en parafina fijadas con formalina), fijación con formalina, infiltración, deshidratación o similares. Opcionalmente, una sección de tejido puede unirse a un soporte sólido, por ejemplo, mediante el uso de técnicas y composiciones ejemplificadas en la presente descripción con respecto a la unión de ácidos nucleicos, células, virus, perlas o similares a un soporte sólido. Como una opción adicional, un tejido puede permeabilizarse y las células del tejido se lisan cuando el tejido está en contacto con un soporte sólido. Cualquiera de una variedad de tratamientos puede usarse tales como los expuestos anteriormente con respecto a la lisis de células. Los ácidos nucleicos diana que se liberan de un tejido que se permeabiliza pueden capturarse mediante sondas de ácido nucleico en la superficie.

Un tejido puede prepararse de cualquier forma conveniente o deseada para su uso en un método, composición o aparato en la presente descripción. Pueden usarse tejidos frescos, congelados, fijados o sin fijar. Un tejido puede fijarse o incrustarse mediante el uso de métodos descritos en la presente descripción o conocidos en la técnica.

Una muestra de tejido para su uso en la presente descripción, puede fijarse mediante congelación por inmersión a una temperatura adecuada para mantener o conservar la integridad de la estructura del tejido, por ejemplo, menos de -20 °C. En otro ejemplo, un tejido puede prepararse mediante el uso de métodos de fijación con formalina e incrustación en parafina (FFPE) que se conocen en la técnica. Pueden usarse otros fijadores y/o materiales de incrustación según convenga. Una muestra de tejido fijada o incrustada puede cortarse en secciones, es decir, cortarse finamente, mediante el uso de métodos conocidos. Por ejemplo, una muestra de tejido puede cortarse en secciones mediante el uso de un micrótopo o criostato frío, establecido a una temperatura adecuada para mantener tanto la integridad estructural de la muestra de tejido como las propiedades químicas de los ácidos nucleicos en la muestra.

En algunas modalidades, una muestra de tejido se tratará para eliminar el material de incrustación (por ejemplo, para eliminar parafina o formalina) de la muestra antes de la liberación, captura o modificación de ácidos nucleicos. Esto puede lograrse al poner en contacto la muestra con un solvente apropiado (por ejemplo, lavados con xileno y etanol). El tratamiento puede ocurrir antes de poner en contacto la muestra de tejido con un soporte sólido expuesto en la presente descripción o el tratamiento puede ocurrir mientras la muestra de tejido está sobre el soporte sólido. Los métodos ilustrativos para manipular tejidos para su uso con soportes sólidos a los que se unen los ácidos nucleicos se exponen en la publicación de solicitud de patente de Estados Unidos núm. 2014/0066318 A1.

El grosor de una muestra de tejido u otro espécimen biológico que se pone en contacto con un soporte sólido en un método, composición o aparato expuesto en la presente descripción puede ser cualquier grosor adecuado deseado. En modalidades representativas, el grosor será al menos 0,1 µm, 0,25 µm, 0,5 µm, 0,75 µm, 1 µm, 5 µm, 10 µm, 50 µm, 100 µm o más grueso. Alternativa o adicionalmente, el grosor de un espécimen biológico que se pone en contacto con un soporte sólido no será más de 100 µm, 50 µm, 10 µm, 5 µm, 1 µm, 0,5 µm, 0,25 µm, 0,1 µm o más delgado.

Una fuente particularmente relevante para un espécimen biológico es un ser humano. El espécimen puede derivarse de un órgano, que incluye, por ejemplo, un órgano del sistema musculoesquelético tal como músculo, hueso, tendón o ligamento; un órgano del sistema digestivo tal como glándula salival, faringe, esófago, estómago, intestino delgado, intestino grueso, hígado, vesícula biliar o páncreas; un órgano del sistema respiratorio tal como laringe, tráquea, bronquios, pulmones o diafragma; un órgano del sistema urinario tal como riñón, uréter, vejiga o uretra; un órgano reproductivo tal como ovario, trompa de Falopio, útero, vagina, placenta, testículo, epidídimo, conductos deferentes, vesícula seminal, próstata, pene o escroto; un órgano del sistema endocrino tal como glándula pituitaria, glándula pineal, glándula tiroidea, glándula paratiroidea, o glándula suprarrenal; un órgano del sistema circulatorio tal como corazón, arteria, vena o capilar; un órgano del sistema linfático tal como vaso linfático, ganglio linfático, médula ósea, timo o bazo; un órgano del sistema nervioso central tal como cerebro, tronco encefálico, cerebelo, médula espinal, nervio craneal o nervio raquídeo; un órgano sensorial tal como ojo, oreja, nariz o lengua; o un órgano del tegumento tal como piel, tejido subcutáneo o glándula mamaria. En algunas modalidades, se obtiene un espécimen biológico de un fluido corporal o excreta tal como sangre, linfa, lágrimas, sudor, saliva, semen, secreción vaginal, cerumen, materia fecal u orina.

Un espécimen de un ser humano puede considerarse (o sospecharse) sano o enfermo cuando se usa. En algunos casos, pueden usarse dos especímenes: un primero que se considera enfermo y un segundo que se considera sano (por ejemplo, para su uso como un control sano). Puede evaluarse cualquiera de una variedad de afecciones, que incluyen pero sin limitarse a, una enfermedad autoinmunitaria, cáncer, fibrosis quística, aneuploidía, infección patógena, afección psicológica, hepatitis, diabetes, enfermedad de transmisión sexual, enfermedad cardíaca, ictus, enfermedad cardiovascular, esclerosis múltiple o distrofia muscular. Las afecciones particularmente relevantes son afecciones genéticas o afecciones asociadas con patógenos que tienen firmas genéticas identificables.

Como se estableció anteriormente, una celda de flujo proporciona un aparato conveniente para su uso en un método expuesto en la presente descripción. Por ejemplo, una celda de flujo es un aparato conveniente para alojar un soporte sólido que se tratará con múltiples reactivos fluidicos tales como los suministros de fluidos repetidos usados para algunos protocolos de secuenciación de ácidos nucleicos o algunos protocolos de hibridación de ácidos nucleicos. En algunas modalidades, un espécimen biológico puede suministrarse a un soporte sólido en una celda de flujo, por ejemplo, cuando una mezcla de fluidos de células, componentes subcelulares, virus o viroides se suministra al soporte sólido. En algunas modalidades puede ser preferible abrir una celda de flujo para exponer un soporte sólido dentro o para retirar el soporte sólido de la celda de flujo para permitir el suministro conveniente de un espécimen biológico al soporte sólido. Por ejemplo, abrir la celda de flujo o retirar el soporte sólido puede permitir a un usuario o dispositivo robótico colocar una sección de tejido sobre el soporte sólido. La abertura de una celda de flujo o la retirada de un soporte sólido de una celda de flujo puede ser temporal. Por lo tanto, la celda de flujo puede cerrarse subsecuentemente o el soporte sólido devolverse a la celda de flujo para continuar con una o más etapas subsecuentes de un método expuesto en la presente descripción.

En algunas modalidades, una celda de flujo puede tener una construcción que permita abrirla o desmontarla. Por ejemplo, la celda de flujo puede estar en un estado cerrado mientras se realiza una reacción de secuenciación, por ejemplo, para decodificar códigos de barras. Después la celda de flujo puede desmontarse de manera que el tejido



pueda colocarse en la superficie de la celda de flujo. La celda de flujo puede mantenerse unida por un adhesivo de manera que una o más superficies puedan retirarse para abrirla. Por ejemplo, una celda de flujo puede tener un separador con superficies adhesivas en la parte superior o inferior (similar a una cinta adhesiva de una o dos caras) y este separador puede aparecer entre dos soportes sólidos. Uno o ambos de los soportes sólidos pueden configurarse para unir ácidos nucleicos y brindar soporte a un espécimen biológico como se expone en la presente descripción. El separador puede tener regiones abiertas (por ejemplo, creadas por corte con láser del material separador) que crean canales de fluidos unidos por los dos soportes sólidos y el separador. Por lo tanto, uno o ambos de los soportes sólidos pueden adherirse de manera no permanente al separador para permitir que uno o ambos de ellos se retiren para permitir el acceso a la superficie cuando se coloca un tejido u otro espécimen sobre el mismo.

Una sonda de ácido nucleico usada en una composición, aparato o método expuesto en la presente descripción puede incluir un resto de captura de diana. En modalidades particulares, el resto de captura de diana es una secuencia de captura de diana. La secuencia de captura de diana es generalmente complementaria a una secuencia diana de manera que la captura de la diana ocurre mediante la formación de un complejo híbrido sonda-diana. Una secuencia de captura de diana puede tener cualquiera de una variedad de longitudes que incluyen, por ejemplo, las longitudes ejemplificadas anteriormente en el contexto de secuencias de código de barras.

En modalidades de formato múltiple, una pluralidad de diferentes sondas de ácido nucleico puede

incluir diferentes secuencias de captura de diana que se hibridan con diferentes secuencias de ácido nucleico diana de un espécimen biológico. Pueden usarse diferentes secuencias de captura de diana para unirse selectivamente a uno o más ácidos nucleicos diana deseados a partir de un espécimen biológico. En algunos casos, las diferentes sondas de ácido nucleico pueden incluir una secuencia de captura de diana que es común a todas o a un subconjunto de las sondas sobre un soporte sólido. Por ejemplo, las sondas de ácido nucleico sobre un soporte sólido pueden tener una secuencia poli A o poli T. Dichas sondas o amplicones de las mismas pueden hibridarse con moléculas de ARNm, moléculas de ADNc o amplicones de las mismas que tienen colas poli A o poli T. Aunque la especie de ARNm o ADNc tendrá diferentes secuencias diana, la captura estará mediada por las regiones de secuencias poli A o poli T comunes.

Cualquiera de una variedad de ácidos nucleicos diana puede capturarse y analizarse en un método expuesto en la presente descripción que incluye, pero no se limita a, ARN mensajero (ARNm), ADN copiado (ADNc), ADN genómico (ADNg), ARN ribosomal (ARNr) o ARN de transferencia (ARNt). Las secuencias diana particulares pueden seleccionarse de bases de datos y secuencias de captura apropiadas diseñadas mediante el uso de técnicas y bases de datos conocidas en la técnica.

Otros restos de captura de diana que son útiles incluyen, por ejemplo, los restos expuestos en la presente descripción como útiles para unir sondas de ácido nucleico a un soporte sólido.

Un método expuesto en la presente descripción puede incluir una etapa de hibridación de sondas de ácido nucleico, que están sobre un soporte sólido, con ácidos nucleicos diana que son de porciones del espécimen biológico que están cerca de las sondas. Generalmente, un ácido nucleico diana difundirá desde una región del espécimen biológico hacia un área del soporte sólido que está cerca de esa región del espécimen. Aquí el ácido nucleico diana interactuará con sondas de ácido nucleico que están cerca de la región del espécimen de la que se liberó el ácido nucleico diana. Un complejo híbrido sonda-diana puede formarse donde el ácido nucleico diana se encuentra con una secuencia de captura de diana complementaria en una sonda de ácido nucleico. La ubicación del complejo híbrido sonda-diana se correlacionará generalmente con la región del espécimen biológico de donde se derivó el ácido nucleico diana. En modalidades de formato múltiple, el soporte sólido incluirá una pluralidad de sondas de ácido nucleico, el espécimen biológico liberará una pluralidad de ácidos nucleicos diana y se formará una pluralidad de híbridos sonda-diana sobre el soporte sólido. Las secuencias de los ácidos nucleicos diana y sus ubicaciones en el soporte proporcionarán información espacial acerca del contenido de ácidos nucleicos del espécimen biológico. Aunque el ejemplo anterior se describe en el contexto de los ácidos nucleicos diana que se liberan de un espécimen biológico, se debe entender que no es necesario liberar los ácidos nucleicos diana. Más bien, los ácidos nucleicos diana pueden permanecer en contacto con el espécimen biológico, por ejemplo, cuando se unen a una superficie expuesta del espécimen biológico de manera que los ácidos nucleicos diana también pueden unirse a sondas de ácido nucleico apropiadas sobre el soporte sólido.

Un método de la presente descripción puede incluir una etapa de extensión de las sondas unidas al soporte sólido a las que se hibridan los ácidos nucleicos diana. En modalidades donde las sondas incluyen secuencias de código de barras, las sondas extendidas resultantes incluirán las secuencias de código de barras y secuencias de los ácidos nucleicos diana (aunque en forma complementaria). Las sondas extendidas son por lo tanto versiones etiquetadas espacialmente de los ácidos nucleicos diana del espécimen biológico. Las secuencias de las sondas extendidas identifican qué ácidos nucleicos están en el espécimen biológico y dónde se ubican los ácidos nucleicos diana en el espécimen biológico. Se debe entender que otros elementos de secuencia que están presentes en las sondas de ácido nucleico también pueden incluirse en las sondas extendidas. Dichos elementos incluyen, por ejemplo, sitios de unión a cebador, sitios de escisión, otras secuencias de etiquetas (por ejemplo, etiquetas de identificación de muestras), secuencias de captura, sitios de reconocimiento para proteínas de unión a ácidos nucleicos o enzimas de ácidos nucleicos, o similares.



La extensión de las sondas puede llevarse a cabo mediante el uso de métodos ejemplificados en la presente descripción o conocidos de cualquier otra manera en la técnica para la amplificación de ácidos nucleicos o la secuenciación de ácidos nucleicos. En modalidades particulares uno o más nucleótidos pueden añadirse al extremo 3' de un ácido nucleico, por ejemplo, por medio de catálisis de polimerasa (por ejemplo, ADN polimerasa, ARN polimerasa o transcriptasa inversa). Pueden usarse métodos químicos o enzimáticos para añadir uno o más nucleótidos al extremo 3' o 5' de un ácido nucleico. Se pueden añadir uno o más oligonucleótidos al extremo 3' o 5' de un ácido nucleico, por ejemplo, por medio de métodos químicos o enzimáticos (por ejemplo, catálisis de ligasa). Un ácido nucleico puede extenderse de una manera dirigida por plantilla, de manera que el producto de extensión es complementario a un ácido nucleico plantilla que se hibrida con el ácido nucleico que se extiende. En algunas modalidades, una transcriptasa inversa extiende un cebador de ADN mediante el uso de una plantilla de ARN, de esta manera se produce un ADNc. Por lo tanto, una sonda extendida preparada en un método expuesto en la presente descripción puede ser una molécula de ADN transcrita inversamente. Los métodos ilustrativos para extender los ácidos nucleicos se exponen en la publicación de solicitud de patente de Estados Unidos núm. US 2005/0037393 A1 o la patente de Estados Unidos núm. 8,288,103 u 8,486,625.

Todo o parte de un ácido nucleico diana que se hibrida con una sonda de ácido nucleico puede copiarse por extensión. Por ejemplo, una sonda extendida puede incluir al menos 1, 2, 5, 10, 25, 50, 100, 200, 500, 1000 o más nucleótidos que se copian de un ácido nucleico diana. La longitud del producto de extensión puede controlarse, por ejemplo, mediante el uso de nucleótidos terminados de manera reversible en la reacción de extensión y ejecución de un número limitado de ciclos de extensión. Los ciclos pueden ejecutarse como se ejemplifica para técnicas de SBS y no es necesario el uso de nucleótidos marcados. En consecuencia, una sonda extendida producida en un método expuesto en la presente descripción puede incluir no más de 1000, 500, 200, 100, 50, 25, 10, 5, 2 o 1 nucleótidos que se copian de un ácido nucleico diana. Por supuesto, las sondas extendidas pueden tener cualquier longitud dentro o fuera de los intervalos expuestos anteriormente.

Aunque los métodos de la presente descripción se ejemplifican mediante una modalidad donde las sondas que se hibridan con los ácidos nucleicos diana se extienden para copiar al menos una porción del ácido nucleico diana, se debe entender que las sondas pueden modificarse de maneras alternativas. Las sondas que se hibridan con los ácidos nucleicos diana pueden someterse a una reacción que crea una modificación específica de la diana de la sonda. Una modificación específica de la diana resultará solo cuando la sonda interactúe con un ácido nucleico diana, por ejemplo, por medio de hibridación basada en complementariedad. En muchas modalidades, la modificación específica de la diana será específica de la secuencia del ácido nucleico diana particular que interactúa con la sonda. Los ejemplos de modificaciones específicas de la diana útiles, incluyen pero no se limitan a, inserción o adición de una secuencia por ligazón o transposición (ver, por ejemplo, la publicación de solicitud de patente de Estados Unidos núm. 2010/0120098 A1), modificaciones químicas tales como entrecruzamiento con psoraleno o adición de un resto de etiqueta detectable, modificaciones por enzimas de ácidos nucleicos, ligazón de un enlazador en horquilla u otras modificaciones expuestas en los ensayos de ácidos nucleicos de la publicación de solicitud de patente de Estados Unidos núm. US 2005/0037393 A1 o la patente de Estados Unidos núm. 8,288,103 u 8,486,625.

Se debe entender que no es necesario que las sondas usadas en un método, composición o aparato expuesto en la presente descripción sean ácidos nucleicos. Pueden usarse otras moléculas tales como proteínas, carbohidratos, moléculas pequeñas, partículas o similares. Las sondas pueden ser una combinación de un componente de ácido nucleico (por ejemplo, que tiene un código de barras, sitio de unión a cebador, sitio de escisión y/u otro elemento de secuencia expuesto en la presente descripción) y otro resto (por ejemplo, un resto que captura o modifica un ácido nucleico diana).

Un método expuesto en la presente descripción puede incluir además una etapa de adquirir una imagen de un espécimen biológico que está en contacto con un soporte sólido. El soporte sólido puede estar en cualquiera de una variedad de estados expuestos en la presente descripción. Por ejemplo, el soporte sólido puede incluir sondas o conglomerados de ácidos nucleicos unidos derivados de sondas de ácido nucleico unidas. Alternativamente, el soporte sólido puede no incluir sondas de ácido nucleico, en cambio estará en un estado que precede a la unión de sondas de ácido nucleico o en un estado que sigue a la eliminación de sondas de ácido nucleico del soporte sólido. En consecuencia, se puede obtener una imagen en cualquiera de una variedad de puntos en un método expuesto en la presente descripción.

Se puede obtener una imagen mediante el uso de dispositivos de detección conocidos en la técnica. Los ejemplos incluyen microscopios configurados para imágenes de luz, campo claro, campo oscuro, contraste de fases, fluorescencia, reflexión, interferencia o confocal. Se puede teñir un espécimen biológico antes de la obtención de imágenes para proporcionar contraste entre diferentes regiones o células. En algunas modalidades, puede usarse más de una tinción para obtener imágenes de diferentes aspectos del espécimen (por ejemplo, diferentes regiones de un tejido, diferentes células, componentes subcelulares específicos o similares). En otras modalidades, se puede obtener una imagen de un espécimen biológico sin tinción.

En modalidades particulares, un microscopio de fluorescencia (por ejemplo, un microscopio de fluorescencia confocal) puede usarse para detectar un espécimen biológico que es fluorescente, por ejemplo, en virtud de un marcador fluorescente. Las imágenes de los especímenes fluorescentes también se pueden obtener mediante el uso de un dispositivo de secuenciación de ácidos nucleicos que tiene óptica para la detección de fluorescencia tal como un

dispositivo de plataforma Genome Analyzer®, MiSeq®, NextSeq® o HiSeq® comercializado por Illumina, Inc. (San Diego, CA); o una plataforma de secuenciación SOLiD™ comercializada por Life Technologies (Carlsbad, CA). Otras ópticas de obtención de imágenes que pueden usarse incluyen las que se encuentran en los dispositivos de detección descritos en Bentley y otros, *Nature* 456:53-59 (2008), publicaciones PCT núm. WO 91/06678, WO 04/018497 o WO 07/123744; las patentes de Estados Unidos núm. 7,057,026, 7,329,492, 7,211,414, 7,315,019 o 7,405,281, y la publicación de solicitud de patente de Estados Unidos núm. 2008/0108082.

Se puede obtener una imagen de un espécimen biológico a una resolución deseada, por ejemplo, para distinguir tejidos, células o componentes subcelulares. En consecuencia, la resolución puede ser suficiente para distinguir los componentes de un espécimen biológico que se separan por al menos 0,5 µm, 1 µm, 5 µm, 10 µm, 50 µm, 100 µm, 500 µm, 1 mm o más. Alternativa o adicionalmente, la resolución puede establecerse para distinguir los componentes de un espécimen biológico que se separan por al menos 1 mm, 500 µm, 100 µm, 50 µm, 10 µm, 5 µm, 1 µm, 0,5 µm o menos.

Un método expuesto en la presente descripción puede incluir una etapa de correlacionar ubicaciones en una imagen de un espécimen biológico con secuencias de código de barras de sondas de ácido nucleico que se unen a una superficie con la que el espécimen biológico se pone, se puso o se pondrá en contacto. En consecuencia, las características del espécimen biológico que son identificables en la imagen pueden correlacionarse con los ácidos nucleicos que se encuentran presentes en su cercanía. Puede usarse cualquiera de una variedad de características morfológicas en una correlación de este tipo, que incluyen por ejemplo, forma celular, tamaño celular, forma del tejido, patrones de tinción, presencia de proteínas particulares (por ejemplo, detectadas por tinciones inmunohistoquímicas) u otras características que se evalúan habitualmente en aplicaciones de patología o investigación. En consecuencia, el estado biológico de un tejido o sus componentes según se determina mediante observación visual puede correlacionarse con las características biológicas moleculares según se determina mediante análisis de ácidos nucleicos resueltos espacialmente.

Un soporte sólido sobre el cual se obtienen imágenes de un espécimen biológico puede incluir marcadores fiduciales para facilitar la determinación de la orientación del espécimen o la imagen de la misma en relación con sondas que se unen al soporte sólido. Los fiduciales ilustrativos incluyen, pero no se limitan a perlas (con o sin restos fluorescentes o restos tales como ácidos nucleicos a los que pueden unirse sondas marcadas), moléculas fluorescentes unidas a características conocidas o determinables, o estructuras que combinan formas morfológicas con restos fluorescentes. Los fiduciales ilustrativos se exponen en la publicación de solicitud de patente de Estados Unidos núm. 2002/0150909 A1 o la solicitud de patente de Estados Unidos con núm. de serie 14/530,299. Uno o más fiduciales son preferentemente visibles mientras se obtiene una imagen de un espécimen biológico. Preferentemente, el soporte sólido incluye al menos 2, 3, 4, 5, 10, 25, 50, 100 o más marcadores fiduciales. Los fiduciales pueden proporcionarse en un patrón, por ejemplo, a lo largo de un borde exterior de un soporte sólido o perímetro de una ubicación donde reside un espécimen biológico. En una modalidad preferida, se detectan uno o más fiduciales mediante el uso de las mismas condiciones de obtención de imágenes usadas para visualizar un espécimen biológico. Sin embargo, si se desea, se pueden obtener imágenes separadas (por ejemplo, una imagen del espécimen biológico y otra imagen de los fiduciales) y las imágenes pueden alinearse entre sí.

Opcionalmente, un espécimen biológico puede eliminarse de un soporte sólido después de que se ha obtenido una imagen y después de que los ácidos nucleicos diana se hayan capturado con sondas de ácido nucleico sobre el soporte sólido. Por lo tanto, un método de la presente descripción puede incluir una etapa de lavar un soporte sólido para eliminar células, tejido u otros materiales de un espécimen biológico. La eliminación del espécimen puede realizarse mediante el uso de cualquier técnica adecuada y dependerá de la muestra de tejido. En algunos casos, el soporte sólido puede lavarse con agua. El agua puede contener varios aditivos, tales como surfactantes (por ejemplo, detergentes), enzimas (por ejemplo, proteasas y colagenasas), reactivos de escisión o similares, para facilitar la eliminación del espécimen. En algunas modalidades, el soporte sólido se trata con una solución que comprende una enzima proteinasa. Alternativa o adicionalmente, la solución puede incluir enzimas celulasas, hemicelulasas o quitinasas (por ejemplo, si se desea eliminar una muestra de tejido de una fuente vegetal o fúngica). En algunos casos, la temperatura de una solución de lavado será al menos 30 °C, 35 °C, 50 °C, 60 °C o 90 °C. Las condiciones para la eliminación de un espécimen biológico pueden seleccionarse mientras no se desnaturalizan los complejos híbridos formados entre los ácidos nucleicos diana y las sondas de ácido nucleico unidas al soporte sólido.

Un método de la presente descripción puede incluir además una etapa de eliminar una o más sondas extendidas de un soporte sólido. En modalidades particulares, las sondas tendrán incluido un sitio de escisión de manera que el producto de extensión de las sondas también incluirá el sitio de escisión. Alternativamente, un sitio de escisión puede introducirse en una sonda durante una etapa de modificación. Por ejemplo, un sitio de escisión puede introducirse en una sonda extendida durante la etapa de extensión.

Los sitios de escisión ilustrativos incluyen, pero no se limitan a, restos que son susceptibles a un proceso químico, enzimático o físico que da como resultado la ruptura de enlaces. Por ejemplo, la ubicación puede ser una secuencia de nucleótidos que se reconoce por una endonucleasa. Las endonucleasas adecuadas y sus secuencias de reconocimiento se conocen bien en la técnica y en muchos casos incluso están disponibles comercialmente (por ejemplo, de New England Biolabs, Beverly MA; ThermoFisher, Waltham, MA o Sigma Aldrich, St. Louis MO). Una endonucleasa particularmente útil romperá un enlace en una hebra de ácido nucleico en un sitio que está remoto en

3' a su sitio de unión en el ácido nucleico, cuyos ejemplos incluyen endonucleasas de restricción de tipo II o tipo IIs. En algunas modalidades, una endonucleasa solo cortará una hebra en un ácido nucleico dúplex (por ejemplo, una enzima formadora de mellas). Los ejemplos de endonucleasas que escinden solo una hebra incluyen Nt.BstNBI y Nt.AlwI.

- 5 En algunas modalidades, un sitio de escisión es un sitio abásico o un nucleótido que tiene una base que es susceptible de eliminarse para crear un sitio abásico. Los ejemplos de nucleótidos que son susceptibles a eliminarse para formar un sitio abásico incluyen uracilo y 8-oxo-guanina. Los sitios básicos pueden crearse por hidrólisis de residuos de nucleótidos mediante el uso de reactivos químicos o enzimáticos. Una vez formados, los sitios básicos pueden escindirse (por ejemplo, mediante tratamiento con una endonucleasa u otra enzima de escisión de una sola hebra, exposición al calor o álcali), lo que proporciona un medio para la escisión sitio específica de un ácido nucleico. Un sitio abásico puede crearse en un nucleótido de uracilo en una hebra de un ácido nucleico. La enzima uracil ADN glucosilasa (UDG) puede usarse para eliminar la base de uracilo, lo que genera un sitio abásico en la hebra. La hebra de ácido nucleico que tiene el sitio abásico puede escindirse después en el sitio abásico mediante el tratamiento con endonucleasa (por ejemplo, endonucleasa EndoIV, AP liasa, FPG glucosilasa/AP liasa, EndoVIII glucosilasa/AP liasa), calor o álcali. En una modalidad particular, el reactivo USER<sup>TM</sup> disponible de New England Biolabs se usa para la creación de una interrupción de un único nucleótido en una base de uracilo en un ácido nucleico.

- Los sitios abásicos también pueden generarse en desoxirribonucleótidos no naturales/modificados distintos de uracilo y escindirse de manera análoga mediante tratamiento con endonucleasa, calor o álcali. Por ejemplo, la 8-oxo-guanina puede convertirse en un sitio abásico mediante la exposición a la FPG glucosilasa. La desoxiinosina puede convertirse en un sitio abásico mediante la exposición a la AlkA glucosilasa. Después, los sitios abásicos así generados pueden escindirse, típicamente mediante tratamiento con una endonucleasa adecuada (por ejemplo, endoIV o AP liasa).

Otros ejemplos de sitios de escisión y métodos que pueden usarse para escindir ácidos nucleicos se exponen, por ejemplo, en la patente de Estados Unidos núm. 7,960,120.

- Las sondas de ácido nucleico modificadas (por ejemplo, sondas de ácido nucleico extendidas) que se liberan de un soporte sólido pueden combinarse para formar una mezcla de fluidos. La mezcla puede incluir, por ejemplo, al menos 10, 100,  $1 \times 10^3$ ,  $1 \times 10^4$ ,  $1 \times 10^5$ ,  $1 \times 10^6$ ,  $1 \times 10^7$ ,  $1 \times 10^8$ ,  $1 \times 10^9$  o más sondas modificadas diferentes. Alternativa o adicionalmente, una mezcla de fluidos puede incluir a lo máximo  $1 \times 10^9$ ,  $1 \times 10^8$ ,  $1 \times 10^7$ ,  $1 \times 10^6$ ,  $1 \times 10^5$ ,  $1 \times 10^4$ ,  $1 \times 10^3$ , 100, 10 o menos sondas modificadas diferentes. La mezcla de fluidos puede manipularse para permitir la detección de las sondas de ácido nucleico modificadas. Por ejemplo, las sondas de ácido nucleico modificadas pueden separarse espacialmente en un segundo soporte sólido (es decir, diferente del soporte sólido del que se liberaron las sondas de ácido nucleico después de haberse puesto en contacto con un espécimen biológico y modificarse), o las sondas pueden separarse temporalmente en una corriente de fluido.

- Las sondas de ácido nucleico modificadas (por ejemplo, sondas de ácido nucleico extendidas) pueden separarse sobre un soporte sólido en un método de captura o detección comúnmente empleado para técnicas basadas en micromatrices o técnicas de secuenciación de ácidos nucleicos tales como las expuestas anteriormente en la presente descripción. Por ejemplo, las sondas modificadas pueden unirse a una micromatriz mediante hibridación con ácidos nucleicos complementarios. Las sondas modificadas pueden unirse a perlas o a la superficie de una celda de flujo y opcionalmente amplificarse como se lleva a cabo en muchas plataformas de secuenciación de ácidos nucleicos. Las sondas modificadas pueden separarse en una corriente de fluido mediante el uso de un dispositivo microfluídico, un dispositivo de manipulación de gotitas o un citómetro de flujo. Típicamente, la detección se lleva a cabo en estos dispositivos de separación, pero la detección no es necesaria en todas las modalidades.

- Un dispositivo de manipulación de gotitas particularmente útil es un actuador de gotitas como se describe, por ejemplo, en la patente de Estados Unidos núm. 8,637,242, la patente de Estados Unidos núm. 6,911,132, titulada "Apparatus for Manipulating Droplets by Electrowetting-Based Techniques", publicada el 28 de junio de 2005; Pamula y otros, la publicación de patente de Estados Unidos núm. 20060194331, titulada "Apparatuses and Methods for Manipulating Droplets on a Print Circuit Board", publicada el 31 de agosto de 2006; Pollack y otros, publicación de patente internacional núm. WO/2007/120241, titulada "Droplet-Based Biochemistry", publicada el 25 de octubre de 2007; Shenderov, patente de Estados Unidos núm. 6,773,566, titulada "Electrostatic actuators for Microfluidics and Methods for Using Same", publicada el 10 de agosto de 2004; Shenderov, patente de Estados Unidos núm. 25 6,565,727, titulada "Actuators for Microfluidics Without Moving Parts", publicada el 20 de mayo de 2003; Kim y otros, patente de Estados Unidos Pub. núm. 20030205632, titulada "Electrowettingdriven Micropumping", publicada el 6 de noviembre de 2003; Kim y otros, publicación de patente de Estados Unidos núm. 20060164490, titulada "Method and Apparatus for Promoting the Complete Transfer of Liquid Drops from a Nozzle", publicada el 27 de julio de 2006; Kim y otros, patente de Estados Unidos núm. 20070023292, titulada "Small Object Moving on Print Circuit Board", publicada el 1 de febrero de 2007; Shah y otros, publicación de patente de Estados Unidos núm. 20090283407, titulada "Method for Using Magnetic Particles in Droplet Microfluidics", publicada el 19 de noviembre de 2009; Kim y otros, publicación de patente de Estados Unidos núm. 20100096266, titulada "Method and Apparatus for Real-time Feedback Control of Electrical Manipulation of Droplets on Chip," publicada el 22 de abril de 2010; Velez, patente de Estados Unidos núm. 7,547,380, titulada "Droplet Transportation Devices and Methods Having a Fluid Surface," publicada el 16 de junio de 2009; Sterling y otros, patente de Estados Unidos núm. 7,163,612, titulada "Method, Apparatus and Article for Microfluidic Control via Electrowetting, for Chemical, Biochemical and Biological Assays and the Like," publicada el 16

de enero de 2007; Becker y otros, patente de Estados Unidos núm. 7,641,779, titulada "Method and Apparatus for Programmable Fluidic Processing," publicada el 5 de enero de 2010; Becker y otros, patente de Estados Unidos núm. 6,977,033, titulada "Method and Apparatus for Programmable Fluidic Processing," publicada el 20 de diciembre de 2005; Decre y otros, patente de Estados Unidos núm. 7,328,979, titulada "System for Manipulation of a Body of Fluid," publicada el 12 de febrero de 2008; Yamakawa y otros, publicación de patente de Estados Unidos núm. 15 20060039823, titulada "Chemical Analysis Apparatus", publicada el 23 de febrero de 2006; Wu, publicación de patente de Estados Unidos núm. 20110048951, titulada "Digital Microfluidics Based Apparatus for Heat-exchange Chemical Processes", publicada el 3 de marzo de 2011; Fouillet y otros, publicación de patente de Estados Unidos núm. 20090192044, titulada "Electrode Addressing Method", publicada el 30 de julio de 2009; Fouillet y otros, patente de Estados Unidos núm. 7,052,244, titulada "Device for Displacement of Small Liquid Volumes Along a Micro-catenary Line by Electrostatic Forces", publicada el 30 de mayo de 2006; Marchand y otros, publicación de patente de Estados Unidos núm. 20080124252, titulada "Droplet Microreactor", publicada el 29 de mayo de 2008; Adachi y otros, publicación de patente de Estados Unidos núm. 20090321262, titulada "Liquid Transfer Device", publicada el 31 de diciembre de 2009; Roux y otros, publicación de patente de Estados Unidos núm. 20050179746, titulada "Device for Controlling the Displacement of a Drop Between Two or Several Solid Substrates", publicada el 18 de agosto de 2005; y Dhindsa y otros, "Virtual Electrowetting Channels: Electronic Liquid Transport with Continuous Channel Functionality," Lab Chip, 10:832-836 (2010).

Pueden detectarse sondas modificadas (por ejemplo, sondas de ácido nucleico extendidas), por ejemplo, después de la separación de una mezcla de fluidos mediante el uso de métodos expuestos anteriormente o conocidos en la técnica. En modalidades particulares, las sondas modificadas que se separan en un segundo soporte sólido (es decir, un soporte sólido que es diferente del primer soporte sólido donde se realizó el contacto entre las sondas y el espécimen biológico) pueden detectarse mediante el uso de técnicas basadas en micromatrices o técnicas de secuenciación de ácidos nucleicos tales como las expuestas anteriormente en la presente descripción. Las sondas que se separan en una corriente de fluido pueden detectarse mediante el uso de detectores ópticos, eléctricos u otros que se usan para equipar dispositivos microfluidicos conocidos, dispositivos de manipulación de gotitas o citómetros de flujo. Un método de detección puede usarse para determinar secuencias de ácidos nucleicos diana, secuencias de código de barras u otras regiones de secuencias de sondas extendidas.

Se han ejemplificado varias modalidades con respecto a la eliminación de sondas modificadas del soporte sólido donde se produjeron las sondas. Sin embargo, se debe entender que las sondas sobre un soporte sólido pueden ponerse en contacto con un espécimen biológico, modificarse sobre el soporte sólido en presencia de ácidos nucleicos diana del espécimen y después las sondas modificadas pueden detectarse sobre el soporte sólido. En una modalidad de este tipo, el espécimen biológico puede eliminarse del soporte sólido antes de la etapa de detección.

En modalidades particulares, la presente descripción proporciona un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico que incluye las etapas de (a) proporcionar una pluralidad de cebadores de ácidos nucleicos unidos a un soporte sólido, en donde los cebadores de ácido nucleico en la pluralidad incluyen una secuencia de cebador universal que es común a los cebadores de ácido nucleico en la pluralidad; (b) unir una población de sondas de ácido nucleico a la pluralidad de cebadores de ácido nucleico, en donde las sondas de ácido nucleico incluyen una secuencia de unión a cebador universal que se hibrida con la secuencia de cebador universal, una secuencia de captura de diana y una secuencia de código de barras que difiere de las secuencias de código de barras de otras sondas de ácido nucleico en la población, de esta manera se unen las diferentes sondas de ácido nucleico en posiciones ubicadas aleatoriamente sobre el soporte sólido; (c) amplificar las diferentes sondas de ácido nucleico mediante la extensión de los cebadores de ácido nucleico, de esta manera se producen conglomerados de ácidos nucleicos que tienen copias de la secuencia de código de barras y la secuencia de captura de diana en las posiciones ubicadas aleatoriamente sobre el soporte sólido; (d) realizar una reacción de secuenciación para determinar las secuencias de código de barras en las posiciones ubicadas aleatoriamente sobre el soporte sólido; (e) poner en contacto un espécimen biológico con los conglomerados de ácidos nucleicos sobre el soporte sólido; (f) hibridar las secuencias de captura de diana de los conglomerados con los ácidos nucleicos diana de porciones del espécimen biológico que están cerca de los conglomerados; y (g) extender las secuencias de captura de diana para producir sondas extendidas que incluyen secuencias de los ácidos nucleicos diana y las copias de las secuencias de código de barras, para de esta manera etiquetar los ácidos nucleicos del espécimen biológico.

Como se ejemplificó anteriormente en la presente descripción, una pluralidad de cebadores de ácido nucleico puede unirse a un soporte sólido, en donde los cebadores de ácido nucleico en la pluralidad incluyen una secuencia de cebadores universales que es común a los cebadores de ácido nucleico en la pluralidad. En esta modalidad, una segunda pluralidad de cebadores de ácido nucleico puede unirse al soporte sólido, y los cebadores de ácido nucleico en la segunda pluralidad pueden tener una segunda secuencia de cebador universal que es común a los cebadores de ácido nucleico en la segunda pluralidad. En esta modalidad, una pluralidad de diferentes sondas de ácido nucleico que se ponen en contacto con el soporte puede incluir una secuencia de unión a cebador universal que se hibrida con el cebador universal sobre el soporte sólido, como se expuso anteriormente, y las diferentes sondas de ácido nucleico también pueden incluir una segunda secuencia de unión a cebador universal que se hibrida con la segunda secuencia de cebador universal. Esta configuración de cebadores universales y sitios de unión a cebador universales puede ser particularmente útil para amplificar las diferentes sondas de ácido nucleico mediante amplificación puente, en donde los cebadores de ácido nucleico en la primera y segunda pluralidad se extienden.

Típicamente, cuando una sonda de ácido nucleico contiene un primer y segundo sitios de unión a cebador universal, se ubicarán en los extremos de la sonda. En algunas modalidades puede ser conveniente eliminar al menos uno de los sitios de unión a cebador de la sonda de ácido nucleico o de los amplicones producidos a partir de la sonda. En consecuencia, las sondas de ácido nucleico pueden incluir opcionalmente un sitio de escisión entre la secuencia de captura de diana y una de la secuencia de unión a cebador universal. En este caso, puede realizarse una reacción de escisión para separar el sitio de unión a cebador universal de la secuencia de captura de diana. Generalmente, la porción de la sonda (o sus amplicones) que contiene la secuencia de captura de diana se unirá al soporte sólido lo que da como resultado la eliminación del sitio de unión a cebador del soporte sólido y la retención de la secuencia de captura de diana. Por lo tanto, la sonda escindida puede usarse para hibridar los ácidos nucleicos diana y la sonda escindida puede extenderse mediante el uso del método expuesto anteriormente en la presente descripción.

En algunas modalidades, una sonda de ácido nucleico incluirá dos sitios de escisión diferentes. Un primer sitio de escisión se ubicará entre un primer sitio de unión a cebador y uno o más de otros elementos de secuencia de la sonda. Un segundo sitio de escisión puede ubicarse entre un segundo sitio de unión a cebador y el uno o más de otros elementos de secuencia de la sonda. Los sitios de escisión deben ser reactivos a diferentes reacciones de escisión de manera que cada uno pueda escindirse selectivamente sin escindir necesariamente el otro. En consecuencia, el primer sitio de escisión puede escindirse antes de modificar la sonda (por ejemplo, antes de producir una sonda extendida), de esta manera se separa el primer sitio de unión a cebador del uno o más de otros elementos de secuencia que permanecen unidos a un soporte sólido. El segundo sitio de escisión puede escindirse después de modificar la sonda (por ejemplo, después de producir la sonda extendida), de esta manera se libera la sonda modificada para su detección subsecuente.

Alternativamente, una sonda de ácido nucleico puede incluir el primer sitio de escisión y un cebador que se usa para capturar o amplificar la sonda de ácido nucleico puede incluir el segundo sitio de escisión. En esta configuración, el primer sitio de escisión puede ubicarse entre un primer sitio de unión a cebador y uno o más de otros elementos de secuencia de la sonda de manera que la escisión separe el primer sitio de unión a cebador de uno o más de otros elementos de secuencia de la sonda que permanecen unidos a un soporte sólido. Nuevamente, esta primera etapa de escisión se llevará a cabo típicamente antes de modificar la sonda (por ejemplo, antes de producir una sonda extendida). Una segunda etapa de escisión puede llevarse a cabo para escindir el segundo sitio de escisión después de modificar la sonda (por ejemplo, después de producir la sonda extendida), de esta manera se libera la sonda modificada para su detección subsecuente.

Las dos modalidades anteriores ejemplifican un sitio de escisión ubicado entre un punto de unión de una sonda de ácido nucleico (o sonda de ácido nucleico modificada) y una o más secuencias de la sonda (o sonda modificada) que contienen información tal como un código de barras espacial o secuencia diana. Por lo tanto, este sitio de escisión es útil para la liberación de sondas modificadas (por ejemplo, sondas extendidas) para detectar la información de secuencia y determinar qué secuencias están presentes en un espécimen biológico y dónde están presentes las secuencias en el espécimen.

En algunas modalidades, una o más sondas que se ponen en contacto con un soporte sólido en un método expuesto en la presente descripción pueden incluir un sitio de unión a cebador de secuenciación. En consecuencia, una sonda modificada (por ejemplo, sonda extendida) puede detectarse en una técnica de secuenciación que incluye una etapa de hibridación de un cebador de secuenciación al sitio de unión a cebador de secuenciación. El sitio de unión a cebador de secuenciación puede ubicarse en la sonda de manera que la escisión de una versión modificada de la sonda (por ejemplo, una sonda extendida) producirá una sonda liberada que incluye el sitio de unión a cebador de secuenciación. El sitio de unión a cebador de secuenciación puede ser un sitio de unión a cebador de secuenciación universal de manera que una pluralidad de diferentes sondas (por ejemplo, que tienen secuencias de código de barras y/o diana diferentes) tendrán el mismo sitio de unión a cebador de secuenciación.

Esta descripción proporciona además un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico, el método incluye las etapas de (a) proporcionar una matriz de perlas sobre un soporte sólido, en donde diferentes sondas de ácido nucleico se unen a diferentes perlas en la matriz, en donde cada una de las diferentes sondas de ácido nucleico incluye una secuencia de código de barras, en donde cada perla incluye una secuencia de código de barras diferente a la de otras perlas sobre el soporte sólido, y en donde cada una de las diferentes sondas de ácido nucleico incluye una secuencia de captura de diana; (b) realizar una reacción de hibridación de sonda decodificadora sobre el soporte sólido para determinar las secuencias de código de barras en las sondas ubicadas aleatoriamente sobre el soporte sólido; (c) poner en contacto un espécimen biológico con la matriz de perlas; (d) hibridar las diferentes sondas de ácido nucleico con los ácidos nucleicos diana de porciones del espécimen biológico que están cerca de las perlas; y (e) extender las diferentes sondas de ácido nucleico para producir sondas extendidas que incluyen secuencias de los ácidos nucleicos diana y las secuencias de código de barras, para de esta manera etiquetar los ácidos nucleicos del espécimen biológico.

Se debe entender que las manipulaciones de soportes sólidos o de ácidos nucleicos unidos a soportes sólidos pueden llevarse a cabo mediante el uso de perlas como soportes sólidos. Las perlas pueden unirse a una superficie (por ejemplo, una matriz de pocillos como en una BeadArray™ de Illumina) antes o después de que se lleven a cabo dichas manipulaciones. Por ejemplo, las sondas de ácido nucleico pueden capturarse en perlas antes o después de que las

perlas se distribuyan en una matriz, las sondas de ácido nucleico pueden amplificarse para crear amplicones en perlas antes o después de que las perlas se distribuyan en una matriz, etc.

#### Ejemplo I

Etiquetado espacial del ARNm de una muestra de tejido mediante el uso de celdas de flujo de Illumina

5 En la Figura 1 se describe un método para generar conglomerados que contienen oligo-dT con código de barras, después revelar el oligo-dT con código de barras con una digestión por enzima de restricción seguida de secuenciación. Se preparó una biblioteca de fragmentos que contenían oligo-dA con código de barras de una hebra, P5', P7, sitio de unión a cebador de secuenciación SBS3 y un sitio de enzima de restricción BspHI (mostrados en el panel superior de la Figura 1) mediante síntesis de oligos (Integrated DNA Technologies). Los códigos de barras eran de 27 mers y se generaron aleatoriamente durante la síntesis. El sitio de unión para el cebador de secuenciación SBS3 se incluyó para la decodificación del código de barras mediante secuenciación. Se incluyó un tramo de oligo-dA para generar un sitio oligo dT después de la formación de conglomerados y la linealización. La amplificación puente y la formación de conglomerados se realizaron de acuerdo con la química de conglomerado estándar (Illumina TruSeq PE Cluster Kit v3 cBot P/N: 15037931) en una celda de flujo de GA de Illumina mediante el uso del protocolo recomendado por el fabricante.

Después de la amplificación puente y la formación de conglomerados, los conglomerados se linealizaron mediante escisión de 8-oxo-G en el cebador P7 mediante el uso de la enzima formamidopirimidina ADN glucosilasa (Fpg) proporcionada en el kit TruSeq PE Cluster. A esto le siguió una digestión por enzima de restricción con 200 unidades/ml de BspHI (NEB Cat # R0517L a 37 °C durante 15 min para eliminar P7' de la hebra anclada al adaptador P5 del conglomerado para revelar el tramo de oligo-dT para su extensión subsecuente en presencia de un ARNm. Las concentraciones de enzimas en el intervalo de 100-400 U/ml se han probado durante 15 o 30 min. La decodificación del código de barras se inició con el cebador de secuenciación SBS3.

Como se muestra en el panel inferior de la Figura 1, se usaron secuencias oligo-dT en el conglomerado para capturar el ARN poli A+ después de la decodificación del código de barras. El ADNc con código de barras se produjo mediante la extensión de la hebra oligo-dT del conglomerado mediante el uso del kit de preparación de muestras de ARN TruSeq (Illumina P/N: 15012997) y el kit de síntesis de ADNc de 1ra hebra mediante transcriptasa inversa del MMLV (Epicentre P/N: MM070150) de acuerdo con las condiciones recomendadas por el fabricante. El ARN capturado se usó como plantilla. El ADNc con código de barras se liberó de la secuencia P5 de la celda de flujo mediante el uso de reactivos de escisión específica de uracilo (USER) de Illumina (kit de conglomerados TruSeq PE de Illumina) que liberan una biblioteca de ADNc con código de barras que se usó para la secuenciación en una segunda celda de flujo de Illumina.

La disponibilidad de la secuencia de captura de oligo dT después de la digestión por enzima de restricción con BspHI se confirmó mediante la hibridación de los conglomerados linealizados con un poli A marcado con Cy5 (24 mer) como se representa en el panel A de la Figura 2. Brevemente, después de la digestión por enzima de restricción, los conglomerados se trataron con NaOH 0,1 N y se lavaron con tampón bajo en sal HT2 para eliminar la segunda hebra en la celda de flujo. Después, se hizo fluir 500 nM de oligo-dA con Cy5 (24 mer) sobre los conglomerados linealizados y desnaturalizados a una velocidad de 30 µl / min y se incubó a 40 °C durante 5 min y después se obtuvieron imágenes. La hibridación del poli A marcado con Cy5 al oligo dT se detectó en los carriles 2-7 de la celda de flujo de GA donde estaban presentes las bibliotecas BODT-1 que contienen oligo dT (ver la imagen de la celda de flujo mostrada en la Figura 2, Panel B). Como es evidente a partir de la imagen de la celda de flujo (Panel B), y el gráfico de barras (Panel C), las bibliotecas PhiX de control (carriles 1 y 8 de la celda de flujo) demostraron tener una fluorescencia muy baja en la señal de Cy5. Estos resultados demostraron que puede crearse un sitio oligo-dT en el conglomerado que después de la linealización puede unirse específicamente al poli A con Cy5 (24 mer).

Las métricas de secuenciación de la celda de flujo descrita anteriormente con 3,2 pM de la biblioteca BODT-1 se presentan en la tabla que se muestra en la Figura 3. Se detectaron millones de lecturas en 21 mosaicos de la secuenciación GA. Después de la secuenciación, se determinó el número de códigos de barras únicos como se representó gráficamente en la Figura 4. Esto se hizo al suponer que cada lectura de filtro de paso (PF) era un código de barras y determinar el número de lecturas únicas (códigos de barras) en cada carril. Se detectaron entre 5 y 11 millones de conglomerados con código de barras únicos después de la secuenciación de los mosaicos en comparación con las bibliotecas de control PhiX. Estos resultados demostraron que la decodificación de secuencias de una biblioteca de secuencias de oligo-dT con código de barras es factible y genera millones de códigos de barras únicos.

#### Ejemplo II

Adhesión celular en celdas de flujo de Illumina

Se capturaron células individuales en una celda de flujo con patrón (celda de flujo HiSeq X10, Illumina). Todas las etapas de flujo de reactivos se realizaron mediante el uso de una bomba peristáltica o el instrumento de generación de conglomerados cBOT (Illumina). Brevemente, se hizo fluir agua libre de nucleasas sobre todos los carriles de la celda de flujo con patrón seguido de una solución de Poli D lisina 30-70 K (100 µg/ml y 20 µg/ml) a un régimen de flujo de 100 µl/min durante 8 min. El suero bovino fetal inactivado por calor (Life Technologies #10082-139) también se probó como adhesivo. Los adhesivos se incubaron en los carriles de las celdas de flujo durante 1 h, seguido de un

lavado con PBS 1X + Pluronic F-68 al 0,5 % (Life Technologies #24040-032). A continuación, las células se adhirieron a las celdas de flujo recubiertas al hacer fluir de 5 a 50 células /  $\mu$ l o aproximadamente 100 – 1000 células por carril a una velocidad de 100  $\mu$ l / min, seguido de una etapa de incubación durante 60 min para unir las células. La celda de flujo se lavó con PBS 1X/Pluronic al 0,5 % a una velocidad de 75  $\mu$ l / min. Si las células se fijaron en la celda de flujo, el paraformaldehído (PFA) al 1 % se hizo fluir en la celda de flujo después de hacer fluir las células como se describió anteriormente y se incubó durante 15 min seguido de la etapa de PBS 1X/Pluronic al 0,5 %. La celda de flujo se retiró y el número de células por carril se contó mediante el uso de un microscopio.

La Figura 5, Panel A muestra una imagen de las celdas capturadas en la celda de flujo con patrón. Los datos del recuento de células mostrados en la Figura 5, Panel B confirmaron que las celdas de flujo recubiertas con poli D Lisina ayudaron a la adherencia celular en comparación con el control recubierto con BSA o sin tratamiento con adhesivo. Como se muestra en la Figura 6, las células adheridas pueden fijarse exitosamente con PFA al 1 %.

#### Ejemplo III

Captura localizada espacialmente del ARNm diana con sondas unidas a una superficie de gel

Este ejemplo describe la creación de un césped de sondas poli T en un portaobjetos recubierto con gel, la colocación de cortes de tejido sobre un césped de sondas poli T, la liberación de ARN de las secciones de tejido, la captura del ARNm liberado por las sondas poli T, la transcripción inversa para marcar las sondas poli T con Cy 3, la retirada del tejido y la obtención de imágenes del portaobjetos.

La Figura 7, Panel A muestra una representación esquemática de las etapas y reactivos usados para crear sondas unidas a un gel. Brevemente, se recubrió un portaobjetos de microscopio con acrilamida libre de silano (SFA), se unieron cebadores P5 y P7 (ver la publicación de solicitud de patente de Estados Unidos núm. 2011/0059865 A1), las sondas que tienen una secuencia poli A y una secuencia complementaria a P5 o P7 se hibridaron con los cebadores P5 y P7, respectivamente, y los cebadores P5 y P7 se extendieron para producir extensiones de secuencias poli T. Se realizó una etapa de control de calidad mediante la hibridación de oligonucleótidos poliA marcados con Cy5 con los cebadores extendidos y la obtención de imágenes de la superficie mediante el uso de un generador de imágenes Axon.

Como se muestra en el Panel B de la Figura 7, se colocó una sección de tejido sobre el gel que tiene los cebadores extendidos con poliT. El tejido se trató para liberar el ARNm y las colas poli A del ARNm liberado se hibridaron con secuencias poli T de los cebadores extendidos. Las secuencias poli T se extendieron mediante el uso de los ARNm capturados como plantillas y los cebadores extendidos se marcaron selectivamente con Cy3. El tejido se retiró del gel y se obtuvieron imágenes de la superficie para detectar la fluorescencia de Cy3.

Como se muestra en la imagen de la Figura 7, las áreas del gel que estaban cerca de las áreas del tejido que liberaron especies de ARNm aparecían fluorescentes mientras que las áreas que no liberaron ARNm aparecían oscuras en la imagen. Por lo tanto, el ARNm capturado creó una imagen tipo huella digital del tejido.

#### Ejemplo IV

Captura localizada espacialmente del ARNm diana con sondas unidas a una superficie BeadArray™

Este ejemplo describe la colocación de cortes de tejido sobre una BeadArray™ que tiene sondas poli T, la liberación de ARN de las secciones de tejido, la captura del ARNm liberado con las sondas poli T, la transcripción inversa para marcar las sondas poli T con Cy5, la retirada del tejido y la obtención de imágenes de la BeadArray™.

Como se muestra en el Panel A de la Figura 8, se colocó una sección de tejido olfativo de ratón en una BeadArray™ que tenía sondas poliT. El tejido se trató para liberar el ARNm y las colas de poli A del ARNm liberado se hibridaron con las secuencias poli T de las sondas. Las secuencias poli T se extendieron mediante el uso de los ARNm capturados como plantillas y los cebadores extendidos se marcaron selectivamente con Cy5. El tejido se retiró de la BeadArray™ y se obtuvieron imágenes de la BeadArray™ para detectar la fluorescencia de Cy5.

Como se muestra en el Panel B de la Figura 7, las áreas de la BeadArray™ que estaban cerca de las áreas del tejido que liberaron especies de ARNm aparecían fluorescentes mientras que las áreas que no liberaron ARNm aparecían oscuras en la imagen. Por lo tanto, el ARNm capturado creó una imagen tipo huella digital del tejido.

Se pretende que el término “que comprende” en la presente descripción sea de extremo abierto, que incluye no solo los elementos mencionados, sino que también abarca cualquier elemento adicional.

## REIVINDICACIONES

1. Un método para etiquetar espacialmente los ácidos nucleicos de un espécimen biológico, que comprende:
  - (a) proporcionar un soporte sólido que comprende ácidos nucleicos unidos a perlas distribuidas aleatoriamente en todo dicho soporte sólido, en donde los ácidos nucleicos unidos a perlas comprenden una secuencia de código de barras y una secuencia de captura, y en donde la secuencia de código de barras en una perla en una posición sobre el soporte sólido difiere de las secuencias de código de barras de ácidos nucleicos en perlas en otras posiciones sobre el soporte sólido;;
  - (b) realizar una reacción de secuenciación para determinar las secuencias de código de barras, o complementos de las mismas, de los ácidos nucleicos, para de esta manera determinar la posición de las secuencias de código de barras sobre el soporte sólido, en donde la reacción de secuenciación es una reacción de secuenciación por ligazón;
  - (c) poner en contacto un espécimen biológico con el soporte sólido que comprende los ácidos nucleicos;
  - (d) hibridar un ácido nucleico diana del espécimen biológico con una secuencia de captura en un ácido nucleico que está cerca del ácido nucleico diana; y
  - (e) extender el ácido nucleico para producir una sonda extendida que comprende la secuencia de código de barras, o un complemento de la misma, y todo o una porción del ácido nucleico diana, o un complemento de la misma, para de esta manera etiquetar espacialmente el ácido nucleico diana del espécimen biológico.
2. El método de la reivindicación 1, en donde los ácidos nucleicos comprenden además una secuencia de unión a cebador, en donde la secuencia de unión a cebador es la misma en todos los ácidos nucleicos.
3. El método de la reivindicación 1 o 2, en donde la secuencia de captura comprende una secuencia poliT o poliA.
4. El método de cualquiera de las reivindicaciones 1-3, en donde el soporte sólido comprende además uno o más marcadores fiduciales.
5. El método de cualquiera de las reivindicaciones 1-4, en donde el ácido nucleico diana se selecciona del grupo que consiste en ARNm, ARNr, ADNg y ADNc.
6. El método de cualquiera de las reivindicaciones 1-5, que comprende además amplificar la sonda extendida para producir una pluralidad de sondas extendidas amplificadas, y determinar la secuencia de los ácidos nucleicos diana etiquetados espacialmente, o complementos de los mismos, y las secuencias de código de barras, o complementos de las mismas, para de esta manera determinar la ubicación del ácido nucleico diana etiquetado espacialmente sobre el soporte sólido.
7. El método de la reivindicación 6, en donde la amplificación comprende una reacción en cadena de la polimerasa, amplificación de círculo rodante, amplificación por desplazamiento de múltiples hebras o amplificación de cebador aleatorio.
8. El método de cualquiera de las reivindicaciones 1-7, en donde el espécimen biológico es una sección de tejido, opcionalmente en donde la etapa (c) comprende además poner en contacto la sección de tejido con el soporte sólido y/o permeabilizar la sección de tejido para liberar los ácidos nucleicos diana de la sección de tejido.
9. El método de cualquiera de las reivindicaciones 1-8, que comprende además una etapa de correlacionar la secuencia de código de barras, o complemento de la misma, de la sonda extendida con una ubicación en el espécimen biológico.
10. El método de cualquiera de las reivindicaciones 1-9, en donde el soporte sólido comprende vidrio, vidrio modificado, vidrio funcionalizado o vidrio inorgánico.
11. El método de cualquiera de las reivindicaciones 1-10, en donde la población de perlas comprende un diámetro promedio de menos de 10 micras y opcionalmente en donde los ácidos nucleicos se unen a las perlas, en donde la etapa (a) comprende además distribuir aleatoriamente la población de perlas sobre el soporte sólido de manera que cada perla en la población de perlas ocupe una posición ubicada aleatoriamente sobre el soporte sólido.



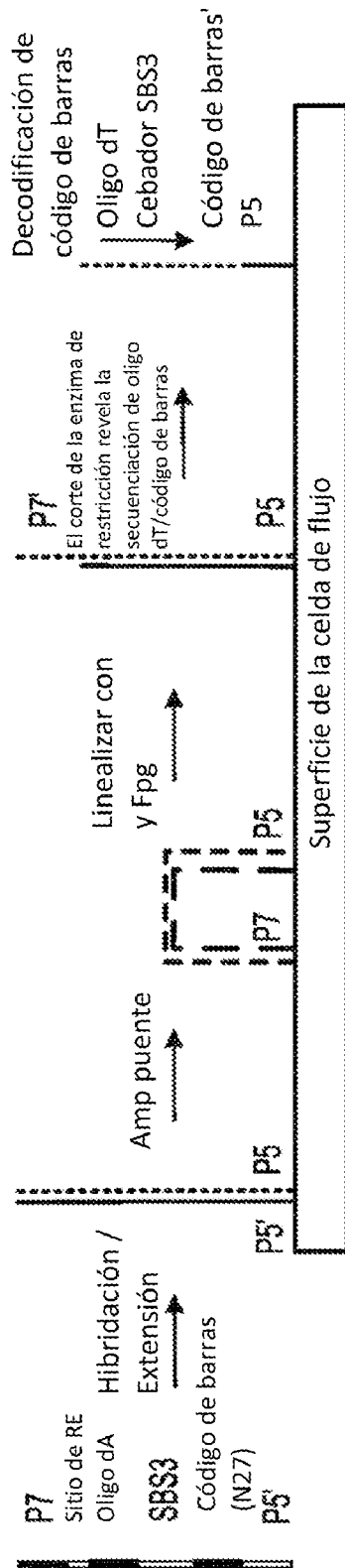


Figura 1A

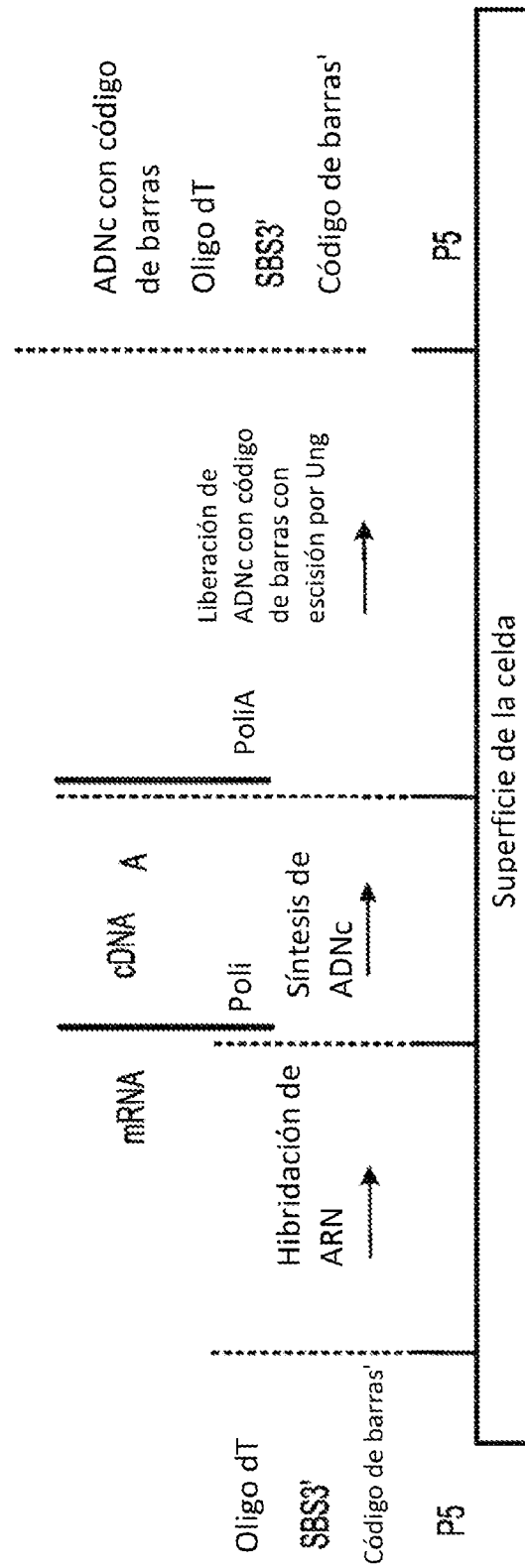


Figura 1B

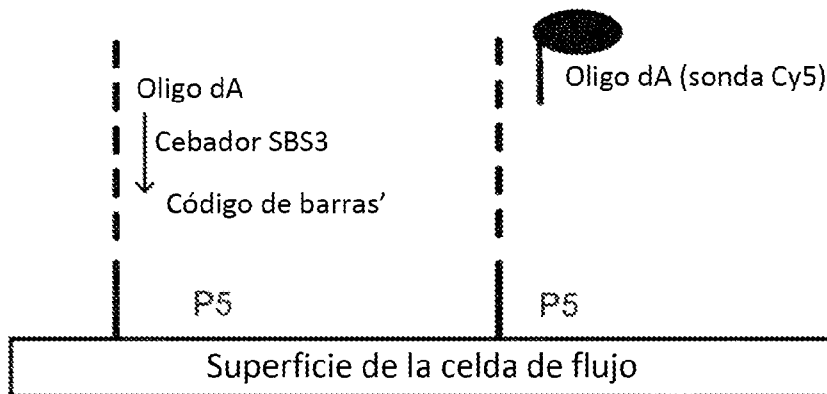


Figura 2A

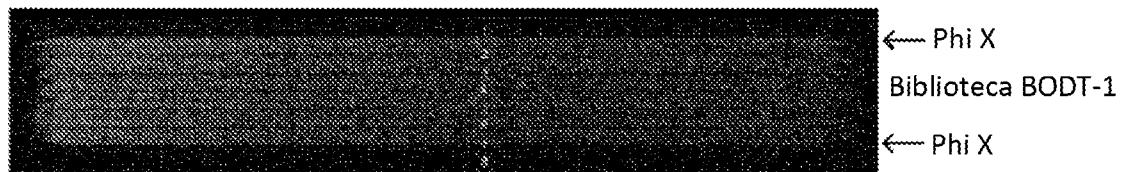


Figura 2B

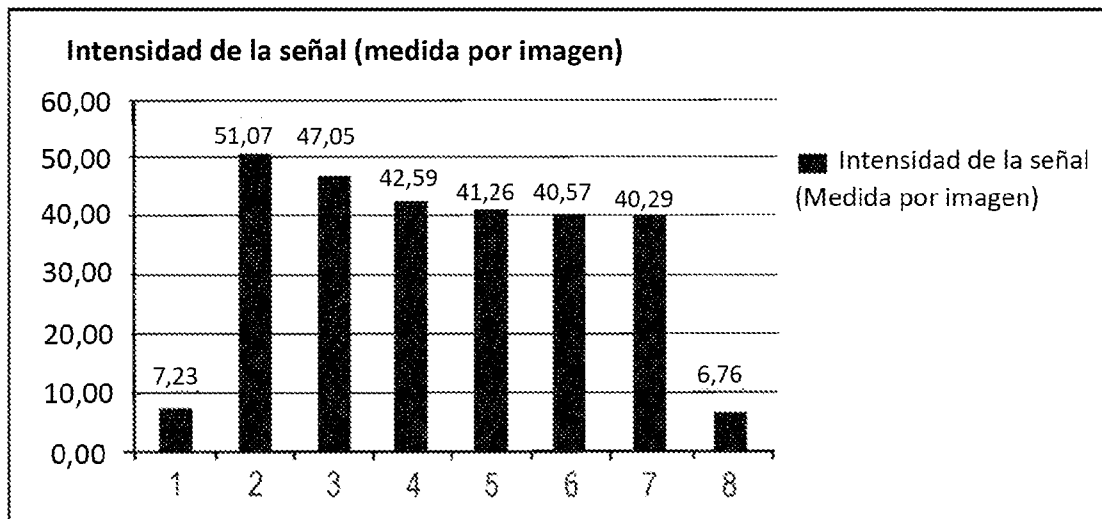


Figura 2C

Carril	Mosaicos	Densidad (K/mm2)	Conglomerado de PF (%)	Fas/Prefas (%)	Lecturas (M)	Lecturas de PF (M)	% $\geq \infty$ Q30
1) Phi X	21	904 +/- 62	92,1 +/- 1,0	0,797 / 0,097	9,83	9,06	94,6
2) BODT-1 a 3,2 nm	21	1026 +/- 171	48,2 +/- 36,6	0,948 / 0,151	11,15	5,59	72,9
3) BODT-1 a 3,2 nm	21	975 +/- 186	41,9 +/- 43,2	0,964 / 0,154	10,6	5,03	70,8
4) BODT-1 a 3,2 nm	21	1144 +/- 24	85,3 +/- 14,2	0,969 / 0,163	12,45	10,62	76,4
5) BODT-1 a 3,2 nm	21	1136 +/- 37	90,6 +/- 0,8	0,977 / 0,173	12,35	11,19	80,4
6) BODT-1 a 3,2 nm	21	1128 +/- 28	90,5 +/- 0,7	0,953 / 0,182	12,27	11,1	80,6
7) BODT-1 a 3,2 nm	21	1131 +/- 28	90,3 +/- 0,8	0,955 / 0,170	12,3	11,1	11,6
8) Phi X	21	763 +/- 41	91,3 +/- 0,4	0,756 / 0,135	8,3	7,58	94

Figura 3

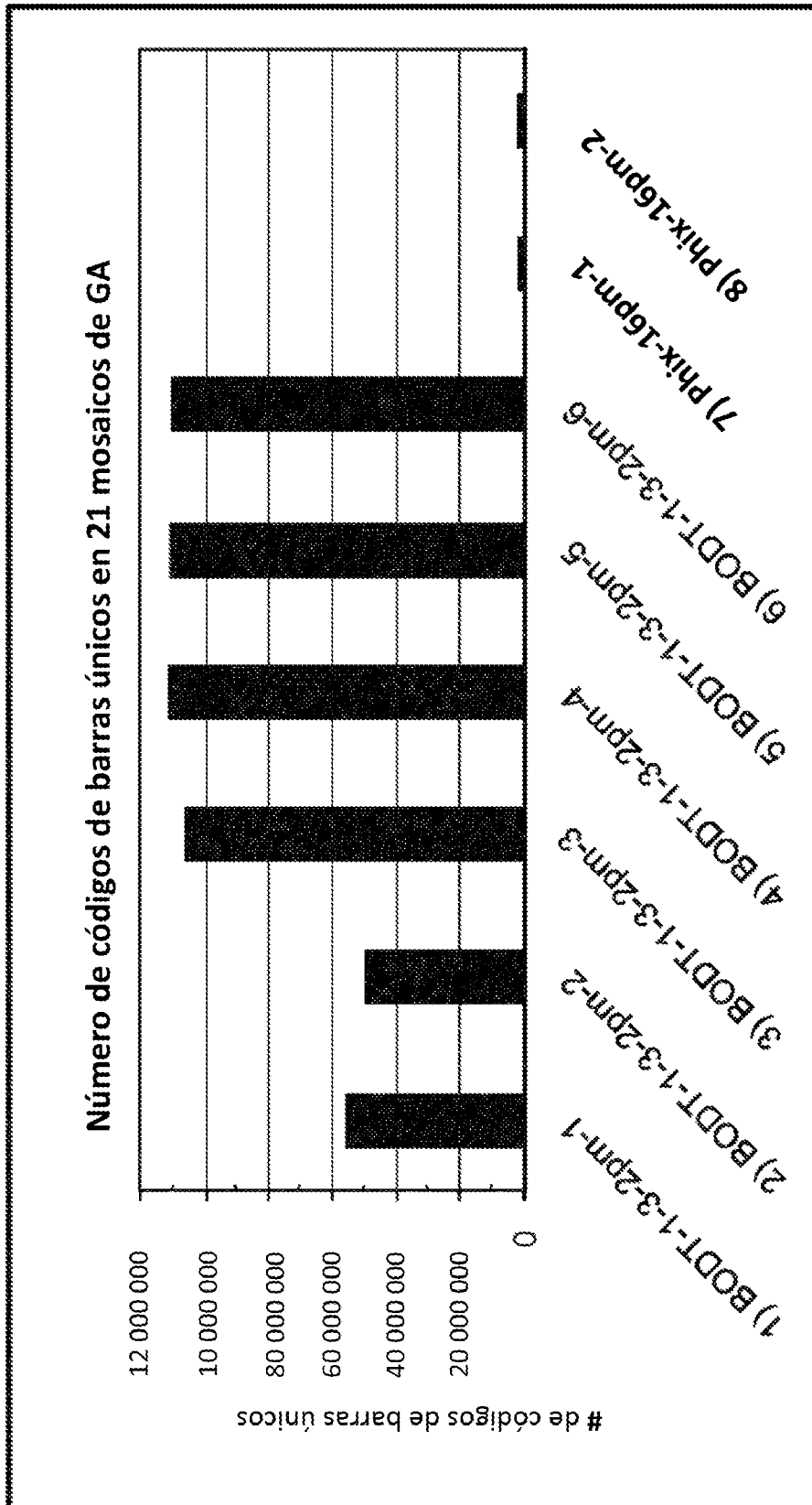


Figura 4

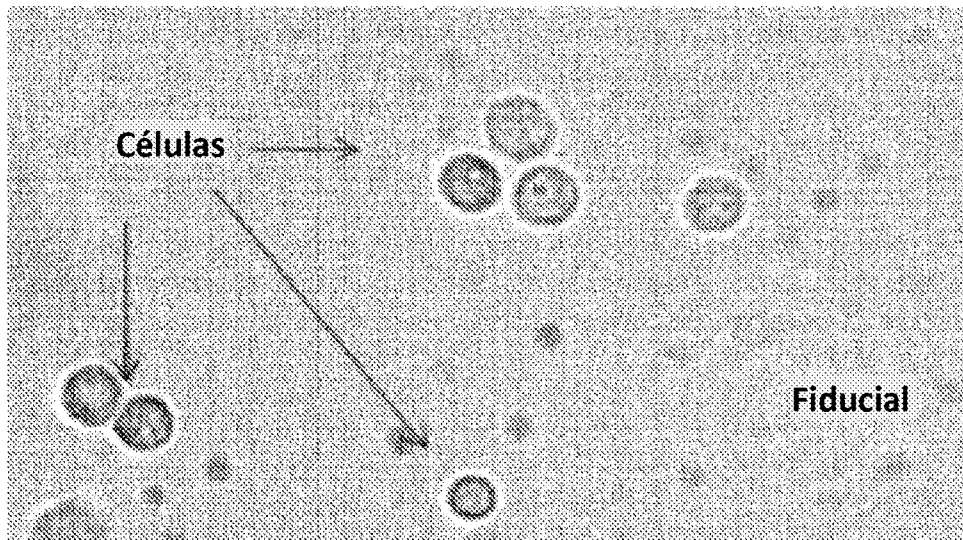


Figura 5A

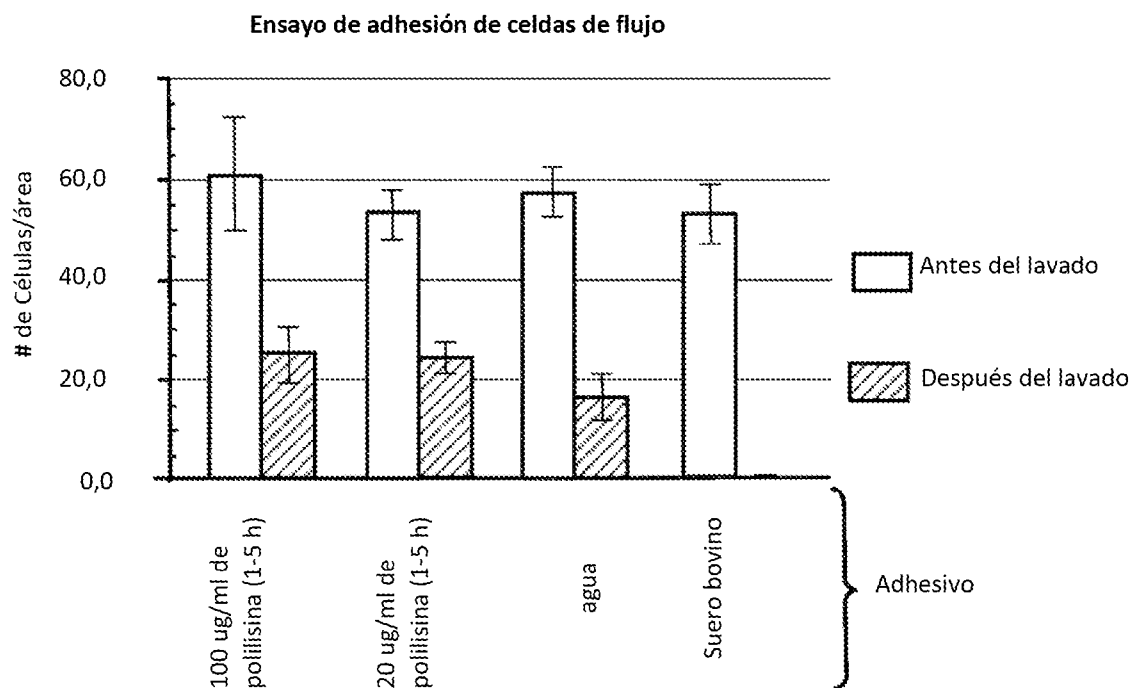


Figura 5B

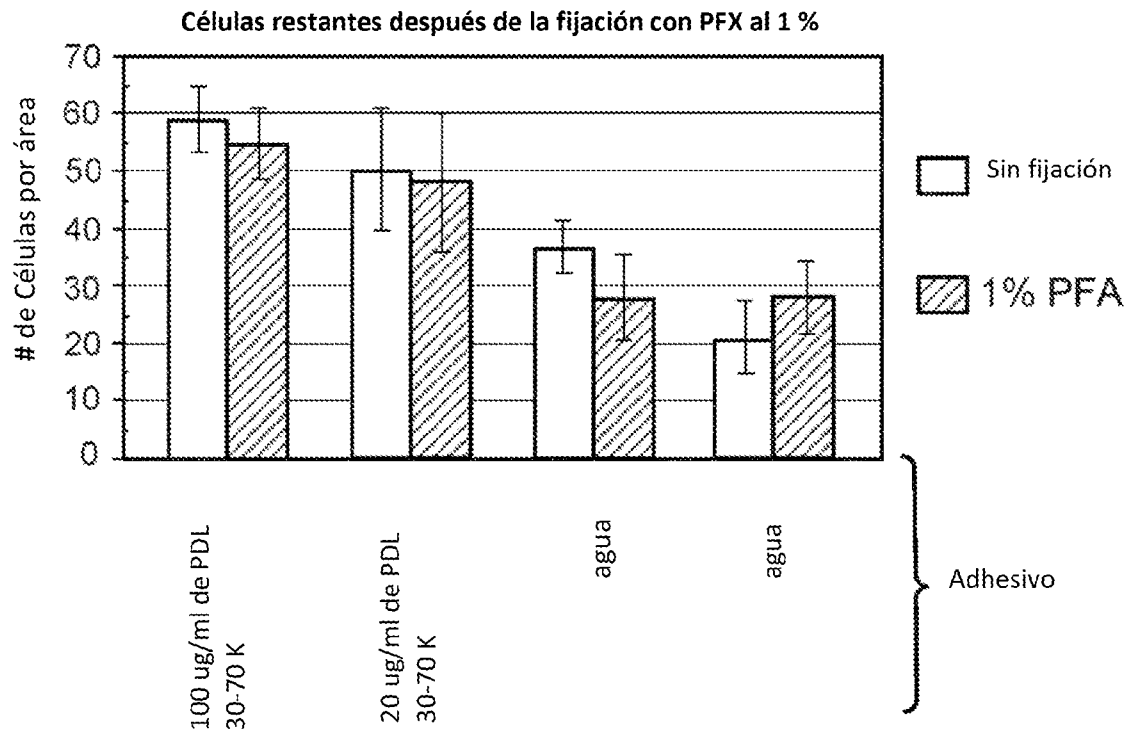


Figura 6

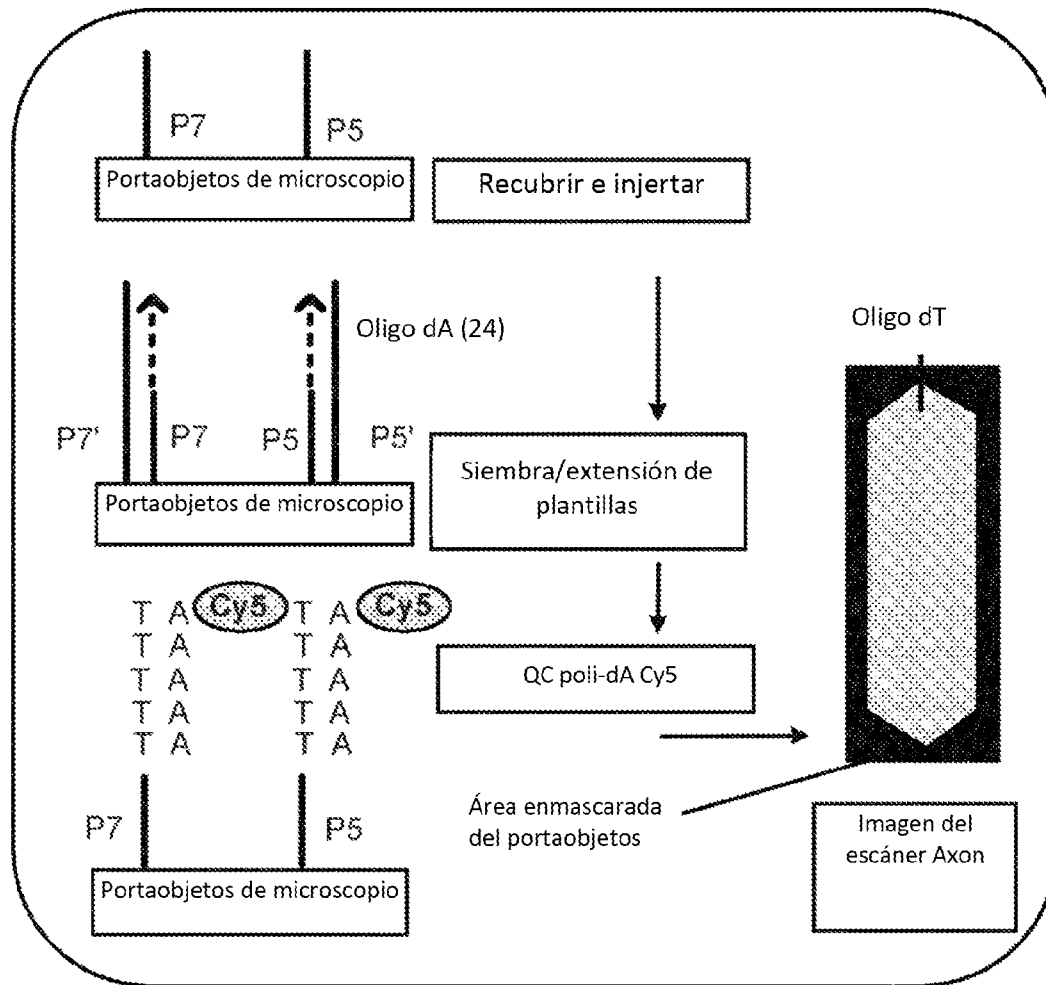


Figura 7A

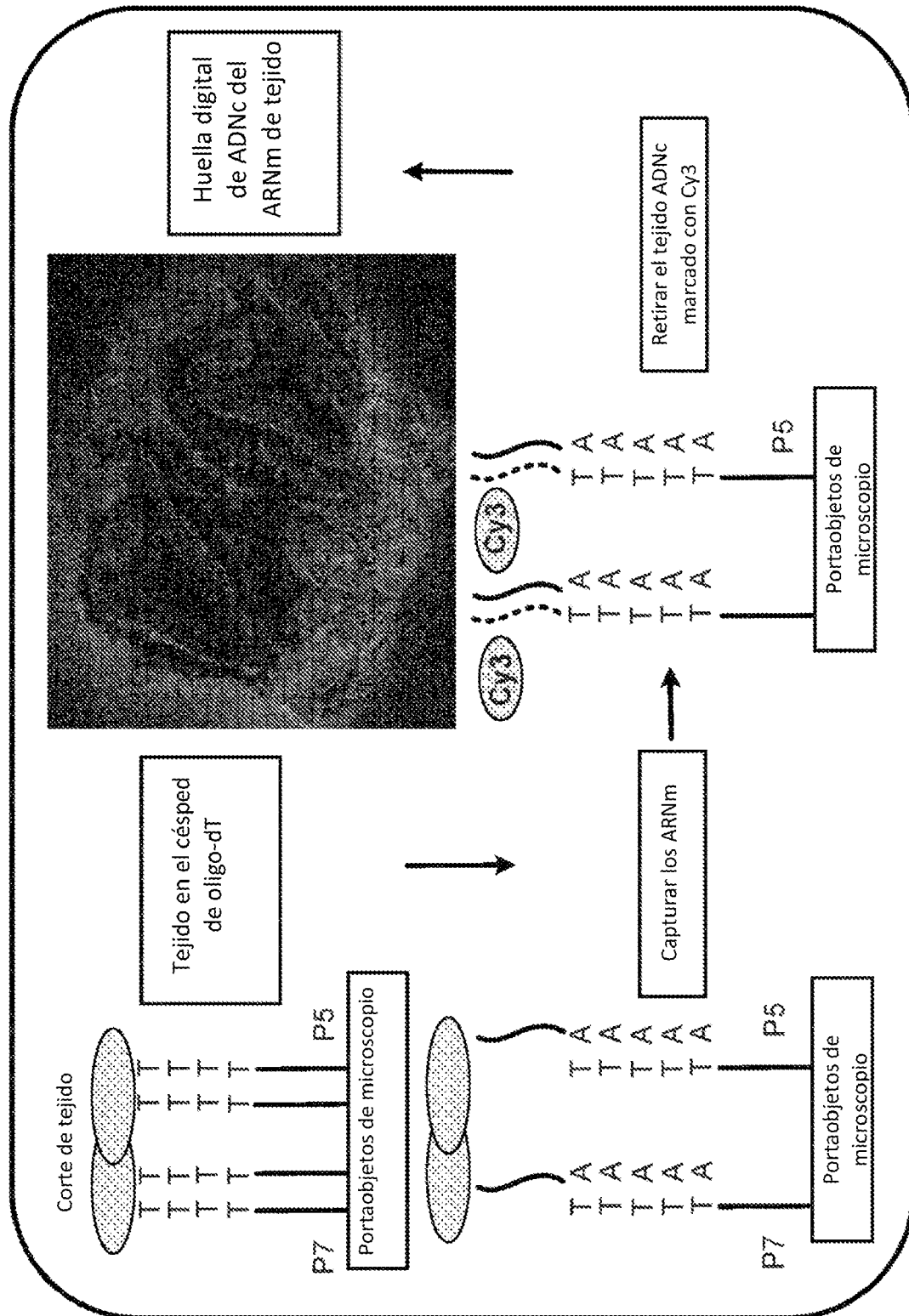


Figura 7B



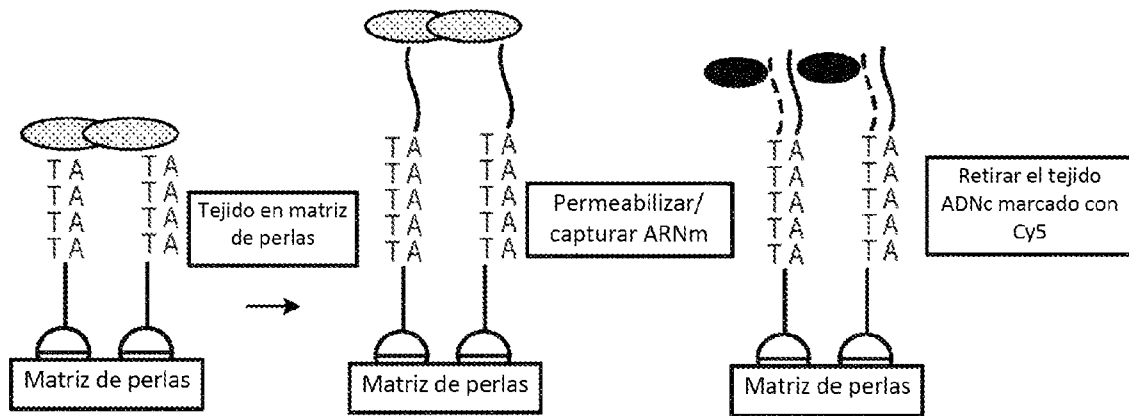


Figura 8A

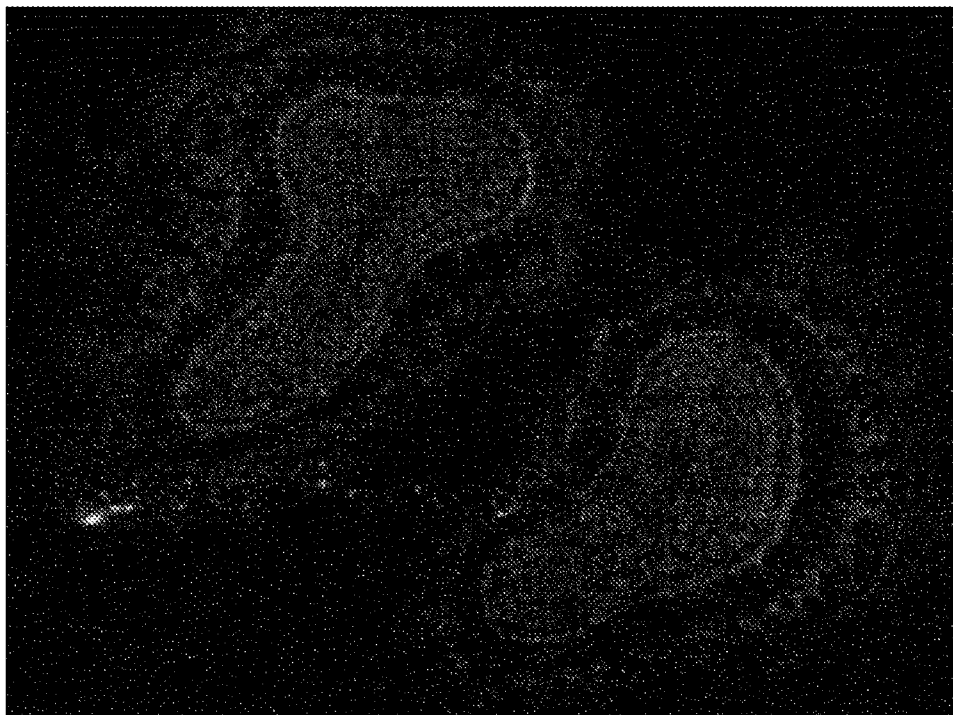


Figura 8B