

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 019 910**

51 Int. Cl.:

C12Q 1/68 (2008.01)

C12Q 1/6837 (2008.01)

C12Q 1/6841 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.10.2013** **E 21163827 (5)**

97 Fecha y número de publicación de la concesión europea: **12.03.2025** **EP 3901280**

54 Título: **Métodos y productos para optimizar la detección localizada o espacial de la expresión génica en una muestra de tejido**

30 Prioridad:

17.10.2012 GB 201218654

14.03.2013 GB 201304585

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.05.2025

73 Titular/es:

10X GENOMICS SWEDEN AB (100.00%)

Solnavägen 3 H

113 63 Stockholm, SE

72 Inventor/es:

FRISEN, JONAS;

STÅHL, PATRIK;

LUNDEBERG, JOAKIM y

SALMÉN, FREDRIK

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 3 019 910 T3

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

DESCRIPCIÓN

Métodos y productos para optimizar la detección localizada o espacial de la expresión génica en una muestra de tejido

5 La presente invención se refiere generalmente a la detección espacial del ARNm en una muestra de tejido o una porción de este. Particularmente, la presente invención proporciona métodos para detectar y/o analizar transcritos de ARNm, para obtener información espacial sobre la localización, distribución o expresión de genes en una muestra de tejido, por ejemplo, en una célula individual. La presente invención permite por lo tanto la transcriptómica espacial.

10 Más particularmente, la presente invención se refiere a un método para la detección espacial de transcritos de ARNm en una muestra de tejido o una porción de este, por ejemplo, para determinar y/o analizar un transcriptoma, y especialmente el transcriptoma global, de una muestra de tejido o una porción de este. En particular, el método se refiere a un método cuantitativo y/o cualitativo para analizar la distribución, ubicación o expresión de ARNm en una muestra de tejido en donde se conserva la expresión espacial o patrón de distribución o ubicación dentro de la muestra de tejido. Por lo tanto, el método proporciona un proceso para realizar "transcriptómica espacial", lo que permite al usuario determinar simultáneamente el patrón de expresión, o el patrón de ubicación/distribución de los genes expresados, presentes en una muestra de tejido o una porción de este. En la presente descripción se describen métodos para determinar las condiciones óptimas para la detección localizada o espacial de (por ejemplo, para capturar) ácidos nucleicos a partir de una muestra de tejido en un sustrato, lo que permite de esta manera capturar la cantidad máxima de moléculas de ácido nucleico, mientras se retiene el patrón de distribución o ubicación que se originó dentro de la muestra de tejido.

25 La invención se basa particularmente en la tecnología de matriz y puede acoplarse con tecnologías de secuenciación de ADN de alto rendimiento. Los métodos de la invención permiten que el ARNm en la muestra de tejido se capture en un sustrato objeto (por ejemplo, un portaobjetos o chip, que puede ser una matriz) y se etiqüete, que puede incluir la incorporación de una etiqueta posicional. Las moléculas etiquetadas pueden visualizarse para determinar o evaluar la eficacia de las condiciones usadas para capturar las moléculas de ácido nucleico. Alternativamente o adicionalmente, el ARNm capturado (o un subconjunto de este, por ejemplo, una porción del ARNm capturado a partir de la muestra de tejido) puede analizarse adicionalmente, por ejemplo, mediante análisis de secuencias. Por ejemplo, las moléculas de ARNm capturadas pueden usarse para moldear la síntesis de moléculas de ADN que se secuencian y analizan para determinar qué genes se expresan en toda o una o más partes de la muestra de tejido. Ventajosamente, el transcriptoma individual, separado y específico de cada célula en la muestra de tejido puede obtenerse al mismo tiempo. Por lo tanto, puede decirse que los métodos de la invención proporcionan firmas de transcriptomas integrales altamente paralelas de células individuales (o grupos de células) dentro de una muestra de tejido o una porción de este, sin perder información espacial dentro de dicha muestra de tejido investigada. Se describe en la presente descripción un sustrato objeto, tal como un chip o portaobjetos (por ejemplo, una matriz) para realizar el método de la invención y también se describen en la presente descripción métodos para fabricar el sustrato objeto, por ejemplo, chip o portaobjetos.

40 El cuerpo humano comprende más de 100 billones de células y se organiza en más de 250 órganos y tejidos diferentes. El desarrollo y la organización de órganos complejos, tales como el cerebro, están lejos de entenderse y existe la necesidad de diseccionar la expresión de genes expresados en tales tejidos mediante el uso de métodos cuantitativos para investigar y determinar los genes que controlan el desarrollo y la función de tales tejidos. Los órganos son en sí mismos una mezcla de células diferenciadas que permiten que todas las funciones corporales, tales como el transporte de nutrientes, la defensa, etc., se coordinen y mantengan. En consecuencia, la función celular depende de la posición de la célula dentro de una estructura tisular particular y de las interacciones que comparte con otras células dentro de ese tejido, tanto directa como indirectamente. Por lo tanto, existe la necesidad de desenredar cómo estas interacciones influyen en cada célula dentro de un tejido a nivel transcripcional.

50 Los hallazgos recientes mediante secuenciación profunda de ARN han demostrado que la mayoría de los transcritos pueden detectarse en una línea celular humana y que una gran fracción (75 %) de los genes que codifican proteínas humanas se expresan en la mayoría de los tejidos. De manera similar, un estudio detallado del 1 % del genoma humano mostró que los cromosomas se transcriben de manera ubicua y que la mayoría de todas las bases se incluyen en los transcritos primarios. Por lo tanto, la maquinaria de transcripción puede describirse como promiscua a nivel global.

60 Se conoce bien que los transcritos son simplemente un sustituto de la abundancia de proteínas, porque las tasas de traducción, degradación, etc. de ARN influirán en la cantidad de proteína producida a partir de cualquier transcrito. Con respecto a esto, un análisis reciente basado en anticuerpos de órganos y tejidos humanos sugiere que la especificidad tisular se logra mediante la regulación precisa de los niveles de proteínas en el espacio y el tiempo, y que los diferentes tejidos en el cuerpo adquieren sus características únicas al controlar no qué proteínas se expresan, sino cuánto de cada una se produce.

65 Sin embargo, en estudios globales posteriores se han comparado las correlaciones entre el transcriptoma y el proteoma, lo que demuestra que se ha demostrado que se expresa la mayoría de todos los genes. Curiosamente, se demostró que había una alta correlación entre los cambios en los niveles de ARN y proteínas para los productos

génicos individuales, lo que es indicativo de la utilidad biológica de estudiar el transcriptoma en células individuales en el contexto del papel funcional de las proteínas.

De hecho, el análisis de la histología y el patrón de expresión en tejidos es una piedra angular en la investigación y el diagnóstico biomédicos. La histología, mediante el uso de diferentes técnicas de tinción, estableció primero la organización estructural básica de órganos sanos y los cambios que tienen lugar en patologías comunes hace más de un centenario. Los desarrollos en este campo dieron como resultado la posibilidad de estudiar la distribución de proteínas mediante inmunohistoquímica y la expresión génica mediante hibridación *in situ*.

Sin embargo, el desarrollo en paralelo de técnicas histológicas y de expresión génica cada vez más avanzadas ha dado como resultado la separación del análisis de imágenes y transcriptomas y, hasta los métodos de la presente invención, no ha habido ningún método factible disponible para el análisis del transcriptoma global con resolución espacial.

Como alternativa, o adicionalmente, a las técnicas *in situ*, se han desarrollado métodos para el análisis *in vitro* de proteínas y ácidos nucleicos, es decir, mediante la extracción de moléculas de muestras de tejido completo, tipos de células individuales, o incluso células individuales, y la cuantificación de moléculas específicas en dichos extractos, por ejemplo, mediante ELISA, qPCR, etc.

Los recientes desarrollos en el análisis de la expresión génica han dado como resultado la posibilidad de evaluar el transcriptoma completo de los tejidos mediante el uso de micromatrices o secuenciación de ARN, y tales desarrollos han sido fundamentales en nuestra comprensión de los procesos biológicos y para el diagnóstico. Sin embargo, el análisis del transcriptoma se realiza típicamente en el ARNm extraído de tejidos completos (o incluso organismos completos), y los métodos para recolectar áreas de tejido más pequeñas o células individuales para el análisis del transcriptoma son típicamente laboriosos, costosos y tienen baja precisión.

Por lo tanto, la mayoría de los estudios de expresión génica basados en micromatrices o secuenciación de próxima generación de ARN usan una muestra representativa que contiene muchas células. Por lo tanto, los resultados representan los niveles de expresión promedio de los genes investigados. La separación de células que son fenotípicamente diferentes se ha usado en algunos casos junto con las plataformas de expresión génica globales (Tang F y otros, Nat Protoc. 2010; 5: 516-35; Wang D & Bodovitz S, Trends Biotechnol. 2010; 28:281-90) y dio como resultado información muy precisa sobre variaciones célula a célula. Sin embargo, hasta ahora no han estado disponibles métodos de alto rendimiento para estudiar la actividad transcripcional con alta resolución en tejidos intactos.

Por lo tanto, las técnicas existentes para el análisis de patrones de expresión génica proporcionan información transcripcional espacial solo para uno o un puñado de genes a la vez u ofrecen información transcripcional para todos los genes en una muestra a costa de perder información posicional. Por lo tanto, es evidente que se requieren métodos para determinar simultáneamente, por separado y específicamente el transcriptoma de cada célula en una muestra, es decir, para permitir el análisis global de la expresión génica en muestras de tejido que producen información transcriptómica con resolución espacial, y la presente invención aborda esta necesidad. La presente invención también puede observarse que proporciona métodos alternativos para el análisis de patrones de expresión génica que proporcionan información transcripcional espacial para uno o un puñado de genes.

El enfoque novedoso de los métodos de la presente invención utiliza ahora tecnología de matriz convencional y puede utilizar tecnología de secuenciación bien establecida, que puede producir información transcripcional para todos los genes en una muestra, mientras retiene información posicional para los transcritos. Será evidente para el experto en la técnica que esto representa un hito en las ciencias biológicas. La nueva tecnología abre un nuevo campo de la llamada "transcriptómica espacial", que probablemente tenga profundas consecuencias para nuestra comprensión del desarrollo tisular y la función tisular y celular en todos los organismos multicelulares. Será evidente que tales técnicas serán particularmente útiles en nuestra comprensión de la causa y el progreso de los estados de la enfermedad y en el desarrollo de tratamientos efectivos para tales enfermedades, por ejemplo, el cáncer. Los métodos de la invención también encontrarán usos en el diagnóstico de numerosas afecciones médicas.

La tecnología de matriz, particularmente micromatrices, surgió de la investigación en la Universidad de Stanford, donde pequeñas cantidades de oligonucleótidos de ADN se unieron con éxito a una superficie vítrea en una disposición ordenada, una llamada "matriz", y la usaron para monitorear la transcripción de 45 genes (Schna M y otros, Science. 1995; 270: 368-9, 371).

Desde entonces, investigadores de todo el mundo han publicado más de 30 000 artículos con el uso de la tecnología de micromatrices. Se han desarrollado múltiples tipos de micromatriz para diversas aplicaciones, por ejemplo, para detectar polimorfismos de un solo nucleótido (SNP) o genomas mutantes de genotipo o resecuencia, y un uso importante de la tecnología de micromatrices ha sido para la investigación de la expresión génica. De hecho, la micromatriz de expresión génica se creó como un medio para analizar el nivel de material genético expresado en una muestra particular, con la ganancia real que es la posibilidad de comparar los niveles de expresión de muchos genes

simultáneamente. Varias plataformas de micromatrices comerciales están disponibles para este tipo de experimentos, pero también ha sido posible crear matrices de expresión génica hechas a medida.

Aunque el uso de micromatrices en los estudios de expresión génica es ahora común, es evidente que las nuevas y más completas tecnologías de "secuenciación de ADN de nueva generación" (NGS) ha empezado a reemplazar a las micromatrices de ADN para muchas aplicaciones, por ejemplo, análisis del transcriptoma en profundidad.

El desarrollo de tecnologías NGS para la secuenciación de genoma ultrarrápida representa un hito en las ciencias biológicas (Pettersen E y otros, *Genomics*. 2009; 93: 105-11). Estas nuevas tecnologías han reducido drásticamente el coste de la secuenciación de ADN y han permitido la determinación del genoma de mayores organismos a una velocidad sin precedentes, incluidos los de individuos específicos (Wade CM y otros *Science*. 2009; 326: 865-7; Rubin J y otros, *Nature* 2010; 464: 587-91). Los nuevos avances en la genómica de alto rendimiento han dado forma al panorama de la investigación biológica y, además de una caracterización completa de los genomas, también es posible estudiar el transcriptoma completo de forma digital y cuantitativa. Las herramientas bioinformáticas para visualizar e integrar estos completos conjuntos de datos también se han mejorado significativamente en los últimos años.

Sin embargo, se ha descubierto sorprendentemente que una combinación única de técnicas histológicas y de micromatriz, que también pueden acoplarse con técnicas de NGS, puede producir información transcripcional completa a partir de múltiples células en una muestra de tejido cuya información se caracteriza por una resolución espacial bidimensional. Por lo tanto, en un extremo, los métodos de la presente invención pueden usarse para analizar la expresión de un gen único en una célula única en una muestra, mientras se retiene la célula dentro de su contexto en la muestra de tejido. En el otro extremo, y en un aspecto preferido de la invención, los métodos pueden usarse para determinar la expresión de cada gen en cada una de las células, o sustancialmente todas las células, en una muestra de tejido (o porción de este) simultáneamente, es decir, el patrón de expresión espacial global de una muestra de tejido o porción de este. Será evidente que los métodos de la invención también permiten realizar análisis intermedios. Por ejemplo, los métodos pueden usarse para determinar o cuantificar la actividad transcripcional de las células en una muestra de tejido (por ejemplo, la abundancia relativa de transcritos en diferentes tipos de células o tejidos), lo que permitiría que el análisis del transcriptoma se centre en regiones específicas de una muestra de tejido, por ejemplo, regiones o porciones de muestras de tejidos con actividad transcripcional alta (o baja).

Será evidente que la eficacia de la etapa de capturar moléculas de ácido nucleico a partir de muestras de tejido puede depender de la fuente del tejido y/o de los métodos usados para preparar la muestra de tejido. En consecuencia, los métodos de la invención pueden usarse para determinar las condiciones óptimas para capturar las moléculas de ácido nucleico a partir de una muestra de tejido en un sustrato objeto, por ejemplo, una matriz.

En su forma más simple, la invención puede ilustrarse mediante el siguiente resumen. La invención requiere que los cebadores de transcripción inversa (RT) se inmovilicen en un sustrato objeto, por ejemplo, un portaobjetos o chip de vidrio. Las secciones de tejido, se colocan en el sustrato y se realiza una reacción de transcripción inversa en la sección de tejido del sustrato. Los cebadores de RT, a los que se une el ARNm en la muestra de tejido (o se hibrida), se extienden mediante el uso del ARNm unido como una plantilla para obtener ADNc, que por lo tanto se une a la superficie del sustrato. La parte sintetizada del ADNc se marca, por ejemplo, con una etiqueta visiblemente detectable, tal como un nucleótido etiquetado con fluorescencia que puede incorporarse en las moléculas de ADNc sintetizadas. Una consecuencia de marcar el ADNc sintetizado es que cada hebra de ADNc proporciona una señal detectable que corresponde a la presencia de una molécula de ARN en la sección de tejido, y la ubicación del ADNc en la superficie del sustrato corresponde a su ubicación original en la sección de tejido. Se detecta la señal del ADNc etiquetado, por ejemplo, se obtienen imágenes de la superficie del sustrato y, en consecuencia, puede detectarse y cuantificarse la abundancia relativa de transcrito presente en cada parte de la sección de tejido, por ejemplo, cada célula. Opcionalmente, la sección de tejido puede visualizarse u obtenerse imágenes, por ejemplo, teñirse y fotografiarse, antes o después de la etapa de síntesis de ADNc para permitir que la molécula de ADNc etiquetada se correlacione con una posición dentro de la muestra de tejido.

Por lo tanto, en un aspecto, el método puede verse como un método para capturar y/o etiquetar el transcriptoma de una muestra de tejido en un sustrato objeto y será evidente a partir de la descripción más abajo que este método puede encontrar una variedad de utilidades, particularmente en métodos para detectar y/o analizar el transcriptoma de una muestra de tejido o una porción de este.

Por ejemplo, cuantificar la abundancia relativa del ADNc etiquetado puede usarse para determinar la actividad transcripcional de diferentes regiones de una muestra de tejido, por ejemplo, para identificar células con actividad transcripcional alta o nula. El método puede usarse para determinar las condiciones óptimas de reacción para detectar, por ejemplo, capturar, el transcriptoma de una muestra de tejido en un sustrato objeto, por ejemplo, al repetir el método mediante el uso de diferentes condiciones (por ejemplo, condiciones para permeabilizar el tejido para permitir que los ácidos nucleicos en la muestra de tejido interactúen con los cebadores inmovilizados en el sustrato), comparar la intensidad y/o resolución de la señal obtenida a partir de moléculas de ADNc etiquetadas en los sustratos fotografiados y opcionalmente seleccionar las condiciones que proporcionan la intensidad y/o resolución óptimas de la imagen.

En algunas modalidades, el método puede verse como un método para o detección espacial, por ejemplo, cuantificación, de la abundancia relativa de uno o más transcritos de ARNm en una muestra de tejido, por ejemplo, el

cebador de RT puede ser específico para uno o más transcritos, lo que permite de esta manera que un transcrito específico (o conjunto de transcritos) se capture y marque en la superficie del objeto, en donde la detección subsecuente, por ejemplo, imágenes, de la intensidad y distribución de la señal del transcrito etiquetado es representativa de la cantidad de transcrito (o conjunto de transcritos) presentes en la muestra de tejido en ubicaciones específicas.

En una modalidad relacionada, el método puede verse como un método para la detección y/o análisis espacial del transcriptoma de una o más porciones de una muestra de tejido, por ejemplo, el cebador de RT puede ser capaz de capturar todos los transcritos de ARNm en la superficie del objeto, que luego se marcan como se describió anteriormente. La detección subsecuente, por ejemplo, obtener imágenes, de la intensidad y distribución de la señal del transcriptoma etiquetado puede usarse para seleccionar una o más porciones del sustrato para un análisis adicional, por ejemplo, análisis de secuencias. Los transcritos de ARNm etiquetados en porciones del sustrato que no se seleccionan para un análisis adicional pueden eliminarse de la superficie del sustrato, por ejemplo, mediante destrucción láser, y desecharse. El subconjunto restante de transcritos de ARNm etiquetados puede analizarse, por ejemplo, secuenciarse, y la información de secuencia puede correlacionarse con la(s) porción(ones) del sustrato del cual no se eliminaron los transcritos de ARNm etiquetados y que corresponden(n) a una posición en la muestra de tejido.

En aún otro aspecto de la invención, los cebadores de transcripción inversa (RT) comprenden también etiquetas posicionales únicas (dominios) y los cebadores de RT pueden disponerse en el sustrato objeto, por ejemplo, un portaobjetos o chip de vidrio, para generar una "matriz". Las etiquetas posicionales únicas corresponden a la ubicación de los cebadores de RT en la matriz (los elementos de la matriz). Las secciones de tejido finas se colocan en la matriz y se realiza una reacción de transcripción inversa en la sección de tejido en el sustrato. Los cebadores de RT, a los que se une el ARN en la muestra de tejido (o se hibrida), se extienden mediante el uso del ARN unido como una plantilla para obtener ADNc, que por lo tanto se une a la superficie de la matriz. La parte sintetizada del ADNc se etiqueta, por ejemplo, una etiqueta detectable, tal como un nucleótido etiquetado con fluorescencia, puede incorporarse en el ADNc sintetizado. Como consecuencia de las etiquetas posicionales únicas en los cebadores de RT, cada hebra de ADNc lleva información sobre la posición del ARN plantilla en la sección de tejido. Se obtienen imágenes del ADNc etiquetado, lo que permite evaluar la eficacia de la etapa de captura del transcriptoma. En algunas modalidades, la visualización de las moléculas de ADNc también permite que las áreas del sustrato sean objetivo para la eliminación del material unido a la superficie, por ejemplo, mediante destrucción láser, de manera que solo los transcritos de porciones de muestra de tejido que son de interés pueden analizarse, además, como se describió anteriormente. La sección de tejido puede visualizarse u obtenerse imágenes, por ejemplo, teñirse y fotografiarse, antes o después de la etapa de síntesis de ADNc para permitir que la etiqueta posicional en la molécula de ADNc se correlacione con una posición dentro de la muestra de tejido. El ADNc se secuencia, lo que da como resultado un transcriptoma con información posicional exacta. Los datos de secuencias pueden entonces coincidir con una posición en la muestra de tejido, lo que permite la visualización, por ejemplo, mediante el uso de un ordenador, de los datos de secuencias junto con la sección de tejido, por ejemplo, para mostrar el patrón de expresión de cualquier gen de interés a través del tejido. De manera similar, sería posible marcar diferentes áreas de la sección de tejido en la pantalla del ordenador y obtener información sobre genes expresados diferencialmente entre cualquier área de interés seleccionada. Será evidente que los métodos de la invención pueden resultar en datos que están en contraste con los datos obtenidos mediante el uso de métodos actuales para estudiar las poblaciones de ARNm. Por ejemplo, los métodos basados en la hibridación *in situ* proporcionan solo información relativa de transcritos de ARNm únicos. Por lo tanto, los métodos de la presente invención tienen claras ventajas sobre las tecnologías *in situ* actuales. La información global de expresión génica que puede obtenerse a partir de los métodos de la invención también permite la información de coexpresión y estimaciones cuantitativas de la abundancia de transcritos. Será evidente que esta es una estrategia generalmente aplicable disponible para el análisis de cualquier tejido en cualquier especie, por ejemplo, animal, planta, hongos.

Se observará a partir de la explicación anterior que existe un inmenso valor en el acoplamiento de la información posicional con la información del transcriptoma. Por ejemplo, permite el mapeo de expresión génica global a alta resolución, lo que resultará útil en numerosas aplicaciones, que incluyen, por ejemplo, la investigación y el diagnóstico del cáncer.

Además, es evidente que los métodos descritos en la presente descripción difieren significativamente de los métodos descritos anteriormente para el análisis del transcriptoma global de una muestra de tejido y estas diferencias resultan en numerosas ventajas. La presente invención se basa en el descubrimiento sorprendente de que el uso de secciones de tejido no interfiere con la síntesis de ADN (por ejemplo, ADNc) preparado por cebadores (por ejemplo, cebadores de transcripción inversa) que se acoplan a la superficie de un sustrato objeto, por ejemplo, una matriz. Además, el etiquetado del ADNc sintetizado en el sustrato objeto permite seleccionar porciones específicas de una muestra de tejido para su análisis posterior, lo que permite que los recursos, por ejemplo, recursos de secuenciación, se centren en el análisis de tipos de células o tejidos específicos dentro de una muestra de tejido. Esto puede resultar en costos reducidos y un conjunto de datos menos complejo que puede analizarse de manera más eficiente y robusta que un conjunto de datos a partir del análisis del transcriptoma de la muestra de tejido completo.

La presente invención proporciona un método para la detección y/o análisis espacial del ARNm en una muestra de tejido o una porción de este, que comprende:

- a) proporcionar un sustrato objeto en el que al menos una especie de sonda de captura, en donde la al menos una especie de sonda de captura comprende un dominio de captura, en donde el dominio de captura se diseña para la captura selectiva de ARNm, se inmoviliza directa o indirectamente de manera que la al menos una especie de sonda de captura se orienta para tener un extremo 3' libre para permitir que la al menos una sonda de captura funcione como un cebador de RT;
- b) poner en contacto el sustrato objeto con una muestra de tejido y permitir que el ARNm de la muestra de tejido se hibride con la al menos una especie de sonda de captura;
- c) generar molécula(s) de ADNc a partir de la(s) molécula(s) de ARNm capturada(s) mediante el uso de la al menos una especie de sonda de captura como cebador de RT;
- d) etiquetar la(s) molécula(s) de ADNc generada(s) en la etapa (c), en donde el etiquetado es simultáneo o subsecuente a la etapa de generación, en donde la etiqueta no es la sonda de captura o una parte de esta;
- e) obtener imágenes de la muestra de tejido para detectar una señal de la(s) molécula(s) de ADNc etiquetadas;
- f) liberar al menos parte de la(s) molécula(s) de ADNc de la superficie del sustrato objeto; y
- g) analizar la secuencia de la(s) molécula(s) de ADNc liberadas.

En otro aspecto, la presente invención proporciona un método para la detección y/o análisis espacial del ARNm y la proteína en una muestra de tejido o una porción de este, que comprende:

- a) proporcionar un sustrato objeto en el que al menos una especie de sonda de captura, en donde la al menos una especie de sonda de captura comprende un dominio de captura, en donde el dominio de captura se diseña para la captura de ARNm, se inmoviliza directa o indirectamente de manera que la al menos una especie de sonda de captura se orienta para tener un extremo 3' libre para permitir que la al menos una especie de sonda de captura funcione como un cebador de transcriptasa inversa (RT);
- b) poner en contacto el sustrato objeto con una muestra de tejido y permitir que el ARNm de la muestra de tejido se hibride con la al menos una especie de sonda de captura;
- c) generar molécula(s) de ADNc a partir de la(s) molécula(s) de ARNm capturada(s) mediante el uso de la al menos una especie de sonda de captura como cebador de RT;
- d) etiquetar la(s) molécula(s) de ADNc generada(s) en la etapa (c), en donde la etapa de etiquetado puede ser simultánea con, o subsecuente a, la etapa de generación, en donde la etiqueta no es la sonda de captura o una parte de esta;
- e) poner en contacto la muestra de tejido con uno o más anticuerpos etiquetados que se unen a una proteína diana de la muestra de tejido; y
- f) detectar: (i) una señal de la(s) molécula(s) de ADNc etiquetada y (ii) el uno o más anticuerpos etiquetados.

En un aspecto adicional, la presente invención proporciona un método para la detección y/o análisis espacial del ARNm en una muestra de tejido o una porción de este, que comprende:

- a) proporcionar un sustrato objeto en el que al menos una especie de sonda de captura, en donde la al menos una especie de sonda de captura comprende un dominio de captura, en donde el dominio de captura se diseña para la captura selectiva de ARNm, se inmoviliza directa o indirectamente de manera que la al menos una especie de sonda de captura se orienta para tener un extremo 3' libre para permitir que la al menos una especie de sonda de captura funcione como un cebador de RT;
- b) poner en contacto el sustrato objeto con una muestra de tejido y permitir que el ARNm de la muestra de tejido se hibride con la al menos una especie de sonda de captura;
- c) generar la primera hebra(s) de molécula(s) de ADNc a partir de la(s) molécula(s) de ARNm capturada(s) mediante el uso de la al menos una especie de sonda de captura como cebador de RT;
- d) etiquetar la primera hebra(s) de molécula(s) de ADNc subsecuente a la generación en la etapa (c), en donde la etapa de etiquetado puede ser simultánea con, o subsecuente a, la etapa de generación (c), en donde la etiqueta no es la sonda de captura o una parte de esta;
- e) obtener imágenes de la muestra de tejido para detectar una señal de la(s) primera(s) hebra(s) etiquetada(s) de molécula(s) de ADNc;
- f) generar segunda(s) hebra(s) complementarias a la primera hebra(s) de molécula(s) de ADNc mediante el uso de la primera hebra(s) como plantilla y/o un cebador de cambio de plantilla;
- g) liberar la(s) segunda(s) hebra(s) del sustrato objeto; y
- h) analizar la secuencia de la(s) segunda(s) hebra(s) liberada(s).

Por lo tanto, en su primer y más amplio aspecto, la presente descripción de la invención proporciona un método para la detección y/o análisis localizado o espacial del ARN, y particularmente de transcritos, en una muestra de tejido o una porción de este, que comprende:

- (a) proporcionar un sustrato objeto en el que al menos una especie de sonda de captura se inmoviliza directa o indirectamente de manera que las sondas se orientan para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT);
- (b) poner en contacto dicho sustrato con una muestra de tejido y permitir que el ARN de la muestra de tejido se hibride con las sondas de captura;
- (c) generar moléculas de ADNc a partir de las moléculas de ARN capturado mediante el uso de dichas sondas de captura como cebadores de RT;

(d) etiquetar las moléculas de ADNc generadas en la etapa (c), en donde dicha etapa de etiquetado puede ser simultánea o subsecuente a dicha etapa de generación, preferentemente en donde la etiqueta se incorpora en la parte sintetizada de las moléculas de ADNc;

(e) detectar una señal de las moléculas de ADNc etiquetadas, por ejemplo, obtener imágenes del sustrato de manera que se detecte la señal de las moléculas de ADNc etiquetadas; y opcionalmente

(f) obtener imágenes de la muestra de tejido en donde de la muestra de tejido se obtienen imágenes antes o después de la etapa (c).

El método puede verse alternativamente o adicionalmente como un método para determinar y/o analizar un transcriptoma de una muestra de tejido o una porción de este o un método para capturar y/o etiquetar el transcriptoma de una muestra de tejido en un sustrato objeto, por ejemplo, una matriz.

Por lo tanto, en un segundo aspecto más particular de la descripción, puede observarse que la presente descripción proporciona un método para determinar las condiciones óptimas para la detección localizada o espacial del ARN (por ejemplo, transcritos) en una muestra de tejido en un sustrato objeto, que comprende las etapas (a)-(e), y opcionalmente la etapa (f), descrita anteriormente en un primer sustrato objeto y las etapas adicionales:

(g) repetir las etapas (a)-(e), y opcionalmente la etapa (f), con un segundo sustrato objeto, mediante el uso de diferentes condiciones en la etapa (b);

(h) comparar la intensidad y/o resolución de la señal de las moléculas de ADNc etiquetadas inmovilizadas en dicho primer y segundo sustrato objeto; y opcionalmente

(i) seleccionar las condiciones que proporcionan la intensidad y/o resolución óptima de la señal de las moléculas de ADNc etiquetadas.

En un tercer aspecto, puede observarse que la presente descripción proporciona un método para determinar y/o analizar ARN o un transcriptoma de una muestra de tejido o una porción de este que comprende las etapas (a)-(e), y opcionalmente la etapa (f), descrita anteriormente y las etapas adicionales:

(g') retirar el ADNc etiquetado de al menos una porción de la superficie del sustrato objeto;

(h') amplificar opcionalmente las moléculas de ADNc restantes inmovilizadas en la superficie del sustrato objeto;

(i') liberar al menos parte de las moléculas de ADNc restantes y/u opcionalmente sus amplicones de la superficie del sustrato objeto, en donde dichas moléculas liberadas pueden ser una primera hebra y/o segunda hebra de la molécula de ADNc o un amplicón de la misma;

(j') analizar directa o indirectamente la secuencia de las moléculas liberadas.

Se debe entender que este tercer aspecto permite determinar y/o analizar una parte del ARN o transcriptoma, y en particular analizarse selectivamente; al eliminar el ADNc etiquetado de al menos una porción de la superficie del sustrato objeto, puede seleccionarse cuál del ADNc (y por lo tanto cuál del ARN capturado) se analizará. La etapa de eliminar se discute más adelante, pero se debe entender que esto podría lograrse al eliminar una porción de la muestra de tejido del sustrato (que eliminará simultáneamente el ADNc etiquetado)

En una modalidad particularmente preferida del tercer aspecto de la descripción, el sustrato objeto es una matriz y cada una de las sondas de captura comprende un dominio posicional que corresponde a la posición de cada sonda de captura en la matriz. En consecuencia, el método puede verse como que comprende:

(a") proporcionar un sustrato objeto (por ejemplo, una matriz) en el que múltiples especies de sondas de captura se inmovilizan directa o indirectamente de manera que cada especie ocupa una posición distinta en el sustrato objeto y se orienta para tener un extremo libre 3' para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT), en donde cada especie de dicha sonda de captura comprende una molécula de ácido nucleico con 5' a 3':

(i) un dominio posicional que corresponde a la posición de la sonda de captura en el sustrato objeto, y

(ii) un dominio de captura;

(b") poner en contacto dicho sustrato objeto con una muestra de tejido de manera que la posición de una sonda de captura en el sustrato objeto puede correlacionarse con una posición en la muestra de tejido y permitir que el ARN de la muestra de tejido se hibride con el dominio de captura en dichas sondas de captura;

(c") generar moléculas de ADNc a partir de las moléculas de ARN capturado mediante el uso de dichas sondas de captura como cebadores de RT,

(d") etiquetar las moléculas de ADNc generadas en la etapa (c'), en donde dicha etapa de etiquetado puede ser simultánea con, o subsecuente a, dicha etapa de generación, preferentemente en donde la etiqueta se incorpora en la parte sintetizada de las moléculas de ADNc;

(e") detectar una señal de las moléculas de ADNc etiquetadas;

(f") opcionalmente obtener imágenes de la muestra de tejido en donde de la muestra de tejido se obtienen imágenes antes o después de la etapa (c").

(g") opcionalmente retirar el ADNc etiquetado de al menos una porción de la superficie del sustrato objeto;

(h") opcionalmente amplificar las moléculas de ADNc inmovilizadas en la superficie del sustrato objeto;

(i") liberar al menos parte de las moléculas de ADNc y/u opcionalmente sus amplicones de la superficie del sustrato objeto, en donde dichas moléculas liberadas pueden ser una primera hebra y/o segunda hebra de la molécula de ADNc o un amplicón de la misma y en donde dicha parte incluye el dominio posicional o un complemento de la misma;

(j") analizar directa o indirectamente la secuencia de las moléculas liberadas.

Los métodos permiten que la abundancia de los transcritos de una muestra de tejido se visualice directamente, por ejemplo, mediante fluorescencia, similar a una micromatriz estándar. Sin embargo, a diferencia de una micromatriz estándar, la abundancia de los transcritos puede correlacionarse directamente con su posición en la muestra de tejido. Ventajosamente, la detección de las moléculas de ADNc etiquetadas *in situ* en la superficie del sustrato objeto (es decir, de manera que su distribución el sustrato objeto corresponde directamente a su distribución en la muestra de tejido) también permite que una porción o porciones de la muestra de tejido de interés se seleccionen para el análisis, lo que minimiza de esta manera la cantidad de datos a evaluar. Además, la detección de las moléculas de ADNc etiquetadas *in situ* en la superficie del sustrato objeto permite la provisión de un método para determinar las condiciones óptimas para capturar el transcriptoma de una muestra de tejido en un sustrato objeto, en donde la distribución espacial de los transcritos en la muestra de tejido se transfiere directamente a la superficie del sustrato objeto. Los sustratos objeto usados para optimizar las condiciones para capturar un transcriptoma (por ejemplo, en donde las sondas de captura no contienen un dominio posicional y/o las sondas se inmovilizan uniformemente en la superficie del sustrato objeto) son relativamente baratos en comparación con un sustrato objeto en el que se disponen las sondas de captura que comprenden un dominio posicional, de manera que cada elemento comprende una especie de sonda de captura con un dominio posicional único. Sin embargo, las condiciones óptimas determinadas mediante el uso del sustrato objeto barato pueden usarse para realizar análisis mediante el uso de las matrices costosas. Por lo tanto, también puede verse que los métodos reducen los costos sobre otros métodos de transcriptómica espacial. Por lo tanto, este aspecto del método puede considerarse que proporciona o habilita una llamada etapa de "control de calidad" (QC) para determinar las condiciones óptimas o más apropiadas, por ejemplo, rentables, en las que realizar un análisis más detallado mediante el uso de matrices más costosas.

Los métodos de la invención también representan una mejora significativa con respecto a otros métodos para transcriptómica espacial conocidos en la técnica. Por ejemplo, los métodos descritos en la presente descripción pueden resultar en un perfil global y espacial de todos los transcritos en la muestra de tejido o una porción de este. Además, los métodos pueden permitir cuantificar la expresión de cada gen para cada posición o elemento en una matriz, lo que permite realizar una multiplicidad de análisis en base a datos de un solo ensayo. Por lo tanto, los métodos de la presente invención hacen posible detectar y/o cuantificar la expresión espacial de todos los genes en una única muestra de tejido o una porción de estos. En algunos aspectos de la invención, la abundancia de los transcritos también puede visualizarse directa e indirectamente. Cuando los métodos incluyen métodos de detección indirecta, por ejemplo, análisis de secuencia, es posible medir la expresión de genes en una única muestra simultáneamente incluso en donde dichos transcritos están presentes a concentraciones muy diferentes en la misma muestra.

Como se describió con más detalle más abajo, cualquier método de análisis de ácidos nucleicos puede usarse en la etapa de análisis (j). Típicamente, esto puede implicar la secuenciación, pero no es necesario realizar una determinación de secuencia real. Por ejemplo, pueden usarse métodos de análisis específico de secuencia. Por ejemplo, puede realizarse una reacción de amplificación específica de secuencia, por ejemplo, mediante el uso de cebadores que son específicos para el dominio posicional y/o para una secuencia diana específica, por ejemplo, un ADN diana particular a detectar (es decir, correspondiente a un ADNc/ARN particular, etc.). Un método de análisis ilustrativo es una reacción de PCR específica de secuencia.

La información del análisis de secuencias obtenida en la etapa (j) puede usarse para obtener información espacial en cuanto al ARN en la muestra o una porción de la muestra. En otras palabras, la información del análisis de secuencias puede proporcionar información sobre la ubicación del ARN en la muestra de tejido. Esta información espacial puede derivarse de la naturaleza de la información del análisis de secuencias determinada, por ejemplo, puede revelar la presencia de un ARN particular que puede ser en sí misma espacialmente informativa en el contexto de la muestra de tejido usada, y/o la información espacial (por ejemplo, localización espacial) puede derivarse de la posición de la muestra de tejido en la matriz, junto con la información de secuencia. Por lo tanto, el método puede implicar simplemente correlacionar la información del análisis de secuencias con una posición en la muestra de tejido, por ejemplo, en virtud de la etiqueta posicional y su correlación con una posición en la muestra de tejido. Sin embargo, como se describió anteriormente, la información espacial puede obtenerse convenientemente mediante la correlación de los datos de expresión, por ejemplo, la intensidad de la señal del ADNc etiquetado detectado, por ejemplo, en imágenes, en la etapa (e) o los datos de análisis de secuencias obtenidos de la etapa (j), con una imagen de la muestra de tejido y esto representa una modalidad preferida de la descripción. En consecuencia, en una modalidad preferida el método también incluye una etapa de:

(f) obtener imágenes de la muestra de tejido en donde de la muestra de tejido se obtienen imágenes antes o después de la etapa (c), preferentemente en donde la imagen del ADNc etiquetado se correlaciona con una imagen de dicha muestra de tejido.

Por lo tanto, el método descrito en el tercer aspecto de la descripción puede comprender una etapa de:

(k) correlacionar dicha información del análisis de secuencias con una imagen de dicha muestra de tejido, en donde de la muestra de tejido se obtienen imágenes antes o después de la etapa (c).

En su sentido más amplio, el método de la invención puede usarse para la detección espacial del ARNm en una muestra de tejido. Por lo tanto, en una modalidad, el método de la invención puede usarse para determinar y/o analizar todo el transcriptoma de una muestra de tejido, por ejemplo, el transcriptoma global de una muestra de tejido. Sin embargo, el método no se limita a esto y abarca determinar y/o analizar todo o parte del transcriptoma de una muestra

de tejido o una porción de este. Por lo tanto, el método puede implicar determinar y/o analizar una parte o subconjunto del transcriptoma, por ejemplo, un transcriptoma correspondiente a un gen o un subconjunto de genes, por ejemplo, un conjunto de genes particulares, por ejemplo, relacionados con una enfermedad o afección particular, tipo de tejido, etc. Alternativamente o adicionalmente, el método puede implicar determinar y/o analizar todo el transcriptoma de una porción de una muestra de tejido.

Visto desde otro aspecto, las etapas del método establecidas anteriormente pueden verse como que proporcionan un método para obtener un transcriptoma definido espacialmente, y en particular el transcriptoma global definido espacialmente de una muestra de tejido o porción de este.

Alternativamente, el método de la invención puede verse como un método para la detección espacial del ARNm, en una muestra de tejido o una porción de este, o para la determinación espacial y/o el análisis del ARNm en una muestra de tejido o una porción de este. En particular, el método puede usarse para la detección o determinación y/o análisis localizado o espacial de la expresión génica en una muestra de tejido o una porción de este. La detección/determinación/análisis localizado/espacial significa que el ARN puede localizarse en su posición o ubicación nativa dentro de una célula o tejido en la muestra de tejido. Por lo tanto, por ejemplo, el ARN puede localizarse en una célula o grupo de células, o tipo de células en la muestra, o en regiones particulares de áreas dentro de una muestra de tejido. Puede determinarse la ubicación o posición nativa del ARN (o, en otras palabras, la ubicación o posición del ARN en la muestra de tejido), por ejemplo, un gen expresado.

La descripción también puede verse como que proporciona métodos para determinar las condiciones óptimas para la detección y/o análisis localizados de ácidos nucleicos en una muestra de tejido en un sustrato objeto, es decir, para capturar y/o etiquetar ácidos nucleicos a partir de una muestra de tejido en un sustrato objeto, por ejemplo, un portaobjetos o chip.

La descripción también puede verse que proporciona un sustrato objeto, por ejemplo, un portaobjetos o chip, para su uso en los métodos de la invención que comprenden un sustrato en el que una o más especies de sonda de captura se inmoviliza directa o indirectamente de manera que cada sonda se orienta para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT).

Opcionalmente, las sondas se inmovilizan uniformemente en el sustrato objeto, es decir, las sondas no se disponen como elementos distintos. Por lo tanto, en algunas modalidades el sustrato objeto descrito en la presente descripción puede verse como una matriz sin elementos, es decir, una micromatriz sin elementos distintos, que se definen más abajo. En una modalidad particular de la invención, una especie de sonda de captura se inmoviliza en la superficie del objeto, es decir, las sondas de captura son idénticas.

En algunas modalidades de la invención, las sondas son capaces de hibridar (es decir, capturar) todo el ARNm, es decir, moléculas de ARN con una cola poliA. Por lo tanto, en modalidades particularmente preferidas de la invención, las sondas pueden comprender secuencias de nucleótidos dTTP o dUTP consecutivos, por ejemplo, oligoT y/u oligoU, como se describe en más detalle más abajo. En una modalidad preferida de la invención, el sustrato objeto descrito en la presente descripción es para su uso en métodos para determinar las condiciones óptimas para la detección localizada o espacial de transcritos a partir de una muestra de tejido en un sustrato objeto, por ejemplo, una matriz.

En algunas modalidades de la invención, las sondas pueden ser capaces de hibridar con (es decir, capturar) tipos específicos de ARNm, es decir, ARN expresado a partir de un gen específico del conjunto de genes. Por lo tanto, en algunas modalidades de la invención, las sondas pueden comprender secuencias específicas de genes o secuencias que se degeneran para una familia de genes. En una modalidad preferida de la invención, el sustrato objeto descrito en la presente descripción es para su uso en métodos para analizar uno o más transcritos a partir de una muestra de tejido o una porción de este.

Por lo tanto, se observará que el sustrato objeto descrito en la presente descripción puede usarse para capturar ARN, por ejemplo, ARNm, de una muestra de tejido, que se pone en contacto con dicha matriz. La matriz también puede usarse para determinar y/o analizar un transcriptoma parcial o global de una muestra de tejido o para obtener un transcriptoma parcial o global definido espacialmente de una muestra de tejido. El sustrato objeto puede ser para uso en los métodos de la invención que pueden considerarse como métodos para determinar, por ejemplo, cuantificar la expresión espacial de uno o más genes en una muestra de tejido o porción de este. Expresado de otra manera, el sustrato objeto puede ser para uso en los métodos usados para detectar la expresión espacial de uno o más genes en una muestra de tejido o porción de este. Aún de otra manera, el sustrato objeto puede ser para uso en los métodos usados para determinar simultáneamente la expresión de uno o más genes en una o más posiciones dentro de una muestra de tejido o una porción de este. Aún más, el sustrato objeto puede ser para uso en los métodos para el análisis parcial o global del transcriptoma de una muestra de tejido o porción de este con resolución espacial bidimensional.

El ARN puede ser cualquier molécula de ARN que puede aparecer en una célula. Por lo tanto, puede ser ARNm, ARNt, ARNr, ARN viral, ARN pequeño nuclear (ARNpn), ARN nucleolar pequeño (ARNnp), microARN (miARN), ARN de interferencia pequeño (ARNip), ARN de interacción con piwi (ARNpi), ARN ribozimal, ARN antisentido o ARN no codificante. Preferentemente, sin embargo, es ARNm.

La etapa (c) en los métodos anteriores para generar ADNc a partir del ARN capturado se observará como relacionada con la síntesis del ADNc. Esto implica una etapa de transcripción inversa del ARN capturado, que extiende la sonda de captura, que funciona como el cebador de RT, mediante el uso del ARN capturado como plantilla. Tal etapa genera el llamado ADNc de la primera hebra. Como se describirá con más detalle más abajo, la síntesis de ADNc de la segunda hebra puede tener lugar opcionalmente en la matriz, o puede tener lugar en una etapa separada, después de la liberación de ADNc de la primera hebra de la matriz. Como se describe con más detalle más abajo, en determinadas modalidades la síntesis de la segunda hebra puede ocurrir en la primera etapa de amplificación de una molécula de ADNc de la primera hebra liberada.

La etapa (d) en los métodos anteriores para etiquetar las moléculas de ADNc generadas en la etapa (c) se verá como relacionada con cualquier método adecuado para etiquetar las moléculas de ADNc. En particular, la etiqueta se proporcionará, por ejemplo, como parte de, o sobre, la molécula de ADNc generada. En modalidades preferidas de la invención, la etapa de etiquetado se realiza simultáneamente con, es decir, al mismo tiempo que, la etapa de generación. Por lo tanto, el etiquetado de puede implicar la incorporación de nucleótidos etiquetados en la molécula de ADNc sintetizada directamente. Como se describe en más detalle más abajo, los nucleótidos pueden etiquetarse con moléculas que proporcionan señales directa o indirectamente. Por ejemplo, las etiquetas que dan señales directamente pueden ser moléculas fluorescentes, es decir, los nucleótidos etiquetados pueden ser nucleótidos etiquetados con fluorescencia. Las etiquetas que dan señales indirectamente pueden ser, por ejemplo, moléculas de biotina, es decir, los nucleótidos etiquetados pueden ser nucleótidos etiquetados con biotina, que requieren etapas adicionales para proporcionar una señal, por ejemplo, la adición de estreptavidina conjugada a una enzima que puede actuar sobre un sustrato químico para proporcionar una señal detectable, por ejemplo, un cambio de color visiblemente detectable.

En algunas modalidades, el ADNc generado en la etapa (c) puede etiquetarse después de su síntesis, por ejemplo, teñido. Varios métodos para etiquetar moléculas de ácido nucleico se conocen en la técnica y podrían emplearse en los métodos de la invención. Por ejemplo, el ADNc puede teñirse con una tinción de ácido nucleico. Si la etapa de generación de ADNc solo crea una primera hebra de ADNc, puede ser ventajoso usar una cepa capaz de detectar ácido nucleico monocatenario tal como SYBR Gold® o GelStar®, ya que la plantilla de ARN puede degradarse. Sin embargo, si se genera una segunda hebra de ADNc, puede usarse una tinción capaz de detectar ácido nucleico bicatenario, tal como bromuro de etidio o SYBR Green®. En algunas modalidades puede ser ventajoso eliminar la muestra de tejido del sustrato objeto antes de etiquetar el ADNc, por ejemplo, para evitar señales de fondo del material de ácido nucleico que permanece en la muestra de tejido, por ejemplo, ADN genómico, etc. En modalidades donde es conveniente obtener imágenes de la muestra de tejido para poder correlacionar la detección, por ejemplo, obtener imágenes, del ADNc etiquetado con una imagen de la muestra de tejido, puede ser conveniente obtener imágenes de la muestra de tejido antes de la etapa de generación del ADNc y particularmente antes de la etapa de etiquetado del ADNc.

El “sustrato objeto” o “sustrato” descrito en la presente descripción puede ser cualquier sustrato sólido en el que las moléculas de ácido nucleico pueden inmovilizarse directa o indirectamente, por ejemplo, un portaobjetos o chip. En modalidades preferidas, el sustrato objeto puede verse como un sustrato de la matriz, es decir, cualquier sustrato que pueda usarse para generar una matriz de ácidos nucleicos, por ejemplo, un sustrato de micromatriz. En muchas modalidades, las sondas de captura pueden inmovilizarse en el sustrato objeto en forma de una matriz, es decir, en algunas modalidades el sustrato objeto es una matriz, por ejemplo, una micromatriz. En otras modalidades, las sondas de captura no se inmovilizan en el sustrato objeto en un formato de matriz, es decir, el sustrato puede no tener elementos distintos y las sondas de captura pueden inmovilizarse en el sustrato objeto de manera uniforme. Por lo tanto, en algunas modalidades el sustrato objeto no es una matriz. Alternativamente, el sujeto objeto puede verse como una matriz inexistente o alternativamente como una matriz que comprende un único elemento largo. Los sustratos de matriz, es decir, los sustratos objeto, para su uso en el contexto del análisis de ácidos nucleicos se discuten y describen más abajo.

Como se usa en la presente descripción, el término “múltiple” significa dos o más, o al menos dos, por ejemplo 3, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 400, 500, 1000, 2000, 5000, 10 000 o más, etc. Por lo tanto, por ejemplo, el número de sondas de captura distintas (es decir, sondas de captura con diferentes secuencias, por ejemplo, diferentes dominios posicionales) puede ser cualquier número entero en cualquier intervalo entre dos cualquiera de los números mencionados anteriormente. Sin embargo, se apreciará que se prevé que puedan usarse matrices de tipo convencional con muchos cientos, miles, decenas de miles, cientos de miles o incluso millones de sondas de captura.

Por lo tanto, los métodos descritos en la presente descripción pueden utilizar, pero no se limitan a, matrices de ácidos nucleicos de alta densidad que comprenden “sondas de captura” para capturar y etiquetar transcritos de todas las células individuales dentro de una muestra de tejido, por ejemplo, un corte de muestra de tejido delgado, o “sección”. Las muestras de tejido o secciones para el análisis pueden producirse de una manera altamente paralela, de manera que se retenga la información espacial en la sección. Las moléculas de ARN capturadas (preferentemente ARNm) para cada célula, o “transcriptomas”, se transcriben a ADNc y las moléculas de ADNc resultantes se etiquetan y detectan y/o analizan. En primer lugar, las moléculas de ADNc etiquetadas se detectan y/o analizan mediante la detección de la señal de las moléculas etiquetadas, por ejemplo, mediante la obtención de imágenes del sustrato

objeto de manera que la señal de la etiqueta puede detectarse, por ejemplo, cuantificarse. Las moléculas de ADNc etiquetadas, o una porción de estas, pueden someterse a análisis, por ejemplo, mediante secuenciación de alto rendimiento. Los datos resultantes de la primera y/o subsecuente detección y/o análisis pueden correlacionarse con imágenes de las muestras de tejido originales. Por ejemplo, puede usarse una superposición de las dos imágenes para determinar áreas de expresión alta, baja o nula, lo que puede permitir que se seleccionen porciones de la muestra de tejido para su análisis posterior. Los datos de los análisis adicionales pueden correlacionarse con imágenes de la muestra de tejido mediante, por ejemplo, las llamadas secuencias de código de barras (o etiquetas de ID, definidas en la presente descripción como dominios posicionales) incorporadas en las sondas de ácido nucleico dispuestas.

Las matrices o micromatrices de ácidos nucleicos de alta densidad forman un componente central de algunos de los métodos de etiquetado del transcriptoma espacial descritos en la presente descripción. Una micromatriz es una tecnología en formato múltiple usada en biología molecular. Una micromatriz típica consiste en una serie de puntos microscópicos de oligonucleótidos (cientos de miles de puntos, generalmente decenas de miles, pueden incorporarse en una única matriz). La posición distinta de cada punto de ácido nucleico (oligonucleótido) (cada especie de oligonucleótido/molécula de ácido nucleico) se conoce como un "elemento" (y por lo tanto en algunos de los métodos expuestos anteriormente cada especie de sonda de captura puede verse como un elemento específico de la matriz; cada elemento ocupa una posición distinta en la matriz), y típicamente cada elemento separado contiene en la región de picomoles (10-12 moles) una secuencia de ADN específica (una "especie"), que se conoce como "sondas" (o "indicadores"). Típicamente, estos pueden ser una sección corta de un gen u otro elemento de ácido nucleico al cual una muestra de ADNc o ARNc (o "diana") puede hibridarse en condiciones de hibridación de alta rigurosidad. Sin embargo, como se describe a continuación, las sondas descritas en la presente descripción y/o su distribución en la matriz pueden diferir de las sondas de micromatrices estándar.

En las micromatrices de expresión génica, la hibridación sonda-diana se detecta y cuantifica usualmente mediante la detección de la señal visual, por ejemplo, un fluoróforo, ion de plata o etiquetado de quimioluminiscencia, que se ha incorporado en todas las dianas antes de que las dianas se pongan en contacto con la matriz. La intensidad de la señal visual se correlaciona con la abundancia relativa de cada ácido nucleico diana en la muestra, pero no proporciona ninguna información espacial sobre el origen del ácido nucleico diana en la muestra. Dado que una matriz puede contener decenas de miles de sondas, un experimento de micromatriz puede lograr muchas pruebas genéticas en paralelo.

En micromatrices estándar, las sondas se unen a una superficie sólida o sustrato mediante un enlace covalente a una matriz química, por ejemplo, epoxi-silano, amino-silano, lisina, poliacrilamida, etc. El sustrato típicamente es un chip o portaobjetos de vidrio, plástico o silicio, aunque se conocen otras plataformas de micromatrices, por ejemplo, perlas microscópicas.

Las sondas pueden unirse al sustrato objeto, por ejemplo, matriz, descrita en la presente descripción por cualquier medio adecuado. En una modalidad preferida las sondas se inmovilizan a un sustrato mediante inmovilización química. Esta puede ser una interacción entre el sustrato (material de soporte) y la sonda en base a una reacción química. Una reacción química de este tipo típicamente no depende de la entrada de energía por medio de calor o luz, sino que puede mejorarse mediante la aplicación de calor, por ejemplo, una cierta temperatura óptima para una reacción química, o luz de cierta longitud de onda. Por ejemplo, puede tener lugar una inmovilización química entre los grupos funcionales en el sustrato y los elementos funcionales correspondientes en las sondas. Tales elementos funcionales correspondientes en las sondas pueden ser un grupo químico inherente de la sonda, por ejemplo, un grupo hidroxilo o introducirse adicionalmente. Un ejemplo de dicho grupo funcional sobre es un grupo amina. Típicamente, la sonda a inmovilizar comprende un grupo amina funcional o se modifica químicamente para comprender un grupo amina funcional. Los medios y métodos para una modificación química de este tipo se conocen bien.

La localización de dicho grupo funcional dentro de la sonda a inmovilizar puede usarse para controlar y dar forma al comportamiento de unión y/u orientación de la sonda, por ejemplo, el grupo funcional puede colocarse en el extremo 5' o 3' de la sonda o dentro de la secuencia de la sonda. Un sustrato típico para inmovilizar una sonda comprende restos que son capaces de unirse a tales sondas, por ejemplo, a ácidos nucleicos funcionalizados con amina. Ejemplos de tales sustratos son sustratos carboxi, aldehído o epoxi. Tales materiales son conocidos por los expertos en la técnica. Los grupos funcionales, que imparten una reacción de conexión entre sondas que son químicamente reactivas mediante la introducción de un grupo amina, y los sustratos de matriz son conocidos por los expertos en la técnica.

Los sustratos alternativos en los que las sondas pueden inmovilizarse pueden tener que activarse químicamente, por ejemplo, mediante la activación de grupos funcionales, disponibles en el sustrato objeto, por ejemplo, sustratos de matriz. El término "sustrato activado" se refiere a un material en el que se establecieron o habilitaron mediante procedimientos de modificación química, grupos funcionales químicos que interactúan o reactivos como conocen los expertos en la técnica. Por ejemplo, un sustrato que comprende grupos carboxilo tiene que activarse antes de su uso. Además, hay sustratos disponibles que contienen grupos funcionales que pueden reaccionar con restos específicos ya presentes en las sondas de ácido nucleico.

En algunas modalidades, las sondas pueden inmovilizarse en perlas, por ejemplo, micropérlas de plástico, que pueden usarse para inmovilizar las sondas en el sustrato. Las técnicas adecuadas para inmovilizar las moléculas de ácido

nucleico en perlas pueden seleccionarse de las técnicas descritas anteriormente o seleccionarse de métodos conocidos en la técnica. Las perlas pueden ponerse en contacto con, e inmovilizarse sobre, un sustrato objeto, lo que inmoviliza de esta manera indirectamente las sondas en la superficie del sustrato. Por ejemplo, después de poner en contacto las perlas con el sustrato, el sustrato puede tratarse al entrecruzar las perlas entre sí y/o la superficie del sustrato, por ejemplo, el sustrato puede calentarse para fundir parcialmente las perlas que pueden solidificarse, para generar un sustrato objeto en el que las sondas se inmovilizan indirectamente.

Alternativamente, las sondas pueden sintetizarse directamente en el sustrato. Los expertos en la técnica conocen métodos adecuados para un enfoque de este tipo. Los ejemplos son técnicas de fabricación desarrolladas por Agilent Inc., Affymetrix Inc., Roche Nimblegen Inc. o Flexgen BV. Típicamente, se usan láseres y un conjunto de espejos que activan específicamente los puntos donde tendrán lugar las adiciones de nucleótidos. Un enfoque de este tipo puede proporcionar, por ejemplo, tamaños de puntos (es decir, elementos) de aproximadamente 30 μm o más. Sin embargo, en algunas modalidades las sondas pueden inmovilizarse uniformemente en el sustrato, es decir, una distribución uniforme, consistente u homogénea de las sondas a través de la superficie del sustrato. Por lo tanto, puede ser necesario simplemente activar una porción o área del sustrato en la que se inmovilizarán las sondas. Una "porción" del sustrato se describe más abajo.

Por lo tanto, el sustrato objeto puede ser cualquier sustrato adecuado conocido por el experto en la técnica. El sustrato puede tener cualquier forma o formato adecuado, por ejemplo, puede ser plano, curvado, por ejemplo, curvado de forma convexa o cóncava hacia el área donde tiene lugar la interacción entre la muestra de tejido y el sustrato. Se prefiere particularmente donde el sustrato es un plano, es decir plano, chip o portaobjetos.

Típicamente, el sustrato es un soporte sólido y de esta manera permite un posicionamiento exacto y rastreable de las sondas en el sustrato. Un ejemplo de un sustrato es un material sólido o un sustrato que comprende grupos químicos funcionales, por ejemplo, grupos amina o grupos funcionalizados con amina. Un sustrato previsto por la presente descripción es un sustrato no poroso. Los sustratos no porosos preferidos son vidrio, silicio, material recubierto de poli-L-lisina, nitrocelulosa, poliestireno, copolímeros de olefina cíclica (COC), polímeros de olefina cíclica (COP), polipropileno, polietileno y policarbonato.

Puede usarse cualquier material adecuado conocido por el experto en la técnica. Típicamente, se usa vidrio o poliestireno. El poliestireno es un material hidrófobo adecuado para unir macromoléculas cargadas negativamente porque normalmente contiene pocos grupos hidrófilos. Para los ácidos nucleicos inmovilizados sobre portaobjetos de vidrio, se conoce además que al aumentar la hidrofobicidad de la superficie vítrea puede aumentarse la inmovilización de ácidos nucleicos. Tal mejora puede permitir una formación relativamente más densamente empaquetada, lo que es ventajoso cuando las sondas se disponen en un formato de matriz. Además de un recubrimiento o tratamiento de la superficie con poli-L-lisina, el sustrato, en particular vidrio, puede tratarse mediante silanación, por ejemplo, con epoxi-silano o amino-silano o mediante silanación o mediante un tratamiento con poliacrilamida.

Una serie de matrices estándar y sustratos de matriz están disponibles comercialmente y tanto el número como el tamaño de los elementos pueden variar. En la presente invención, cuando las sondas se distribuyen uniformemente sobre la superficie del sustrato, la concentración de las sondas inmovilizadas puede alterarse para corresponder con el tamaño y/o la densidad de las células presentes en diferentes tejidos u organismos. De manera similar, cuando las sondas están en un formato de matriz el arreglo de los elementos puede alterarse para que corresponda al tamaño y/o la densidad de las células presentes en diferentes tejidos u organismos. Por ejemplo, las células animales típicamente tienen una sección transversal en la región de 1-100 μm , mientras que la sección transversal de las células vegetales típicamente puede variar de 1-10 000 μm . Por lo tanto, en modalidades donde las sondas se disponen en el sustrato, matrices Nimblegen®, que están disponibles con hasta 2,1 millones de elementos, o 4,2 millones de elementos, y tamaños de elementos de 13 micrómetros, pueden preferirse para muestras de tejido de un animal u hongo, mientras que otros formatos, por ejemplo, con 8x130k elementos, pueden ser suficientes para muestras de tejido vegetal. Las matrices comerciales también están disponibles o se conocen para su uso en el contexto del análisis de secuencias y en particular en el contexto de las tecnologías NGS. Tales matrices también pueden usarse como el sustrato, por ejemplo, sustrato de la matriz en el contexto de la presente descripción, por ejemplo, una matriz de perlas de Illumina. Además de las matrices disponibles comercialmente, que pueden personalizarse en sí mismas, es posible fabricar matrices "internas" personalizadas o no estándar y los métodos para generar matrices están bien establecidos. Los métodos de la invención pueden utilizar matrices estándar y no estándar que comprenden sondas como se define más abajo.

Las sondas sobre un sustrato pueden inmovilizarse, es decir, unirse o enlazarse, al sustrato, por ejemplo, a la matriz a través del extremo 5' o 3', en dependencia de la matriz química de la matriz. Típicamente, para las matrices disponibles comercialmente, las sondas se unen a través de un enlace 3', de esta manera se deja un extremo 5' libre. Sin embargo, están disponibles sustratos, por ejemplo, matrices que comprenden sondas unidas al sustrato a través de un enlace 5', de esta manera se deja un extremo 3' libre, y pueden sintetizarse mediante el uso de técnicas estándar que se conocen bien en la técnica y se describen en otra parte en la presente descripción.

El enlace covalente usado para acoplar una sonda de ácido nucleico a un sustrato, puede verse tanto como un enlace directo como indirecto, dado que, aunque la sonda está unida por un enlace covalente "directo", puede haber un resto

químico o enlazador que separa el "primer" nucleótido de la sonda de ácido nucleico del sustrato, por ejemplo, vidrio o silicio, es decir, un enlace indirecto. Para los propósitos de la presente invención, se observa generalmente que las sondas que se inmovilizan al sustrato mediante un enlace covalente y/o enlazador químico se inmovilizan o se unen directamente al sustrato.

Como se describirá con más detalle más abajo, las sondas de captura descritas en la presente descripción pueden inmovilizarse sobre, o interactuar con, el sustrato, por ejemplo, matriz, directa o indirectamente. Por lo tanto, no es necesario que las sondas de captura se unan directamente al sustrato, pero pueden interactuar indirectamente, por ejemplo, mediante la unión a una molécula que se une por sí misma directa o indirectamente a la matriz (por ejemplo, la sonda de captura puede interactuar con (por ejemplo, unirse o hibridarse con) un compañero de unión de la sonda de captura, es decir, una sonda de superficie, que se une por sí misma al sustrato directa o indirectamente). En términos generales, sin embargo, la sonda de captura estará, directa o indirectamente (por uno o más intermediarios), unida a, o inmovilizada sobre, el sustrato.

El método y sustrato objeto, por ejemplo, portaobjetos o chip, descritos en la presente descripción pueden comprender sondas que se inmovilizan a través de su extremo 5' o 3'. Sin embargo, cuando la sonda de captura se inmoviliza directamente al sustrato de la matriz, puede inmovilizarse solo de manera que el extremo 3' de la sonda de captura sea libre de extenderse, por ejemplo, se inmoviliza por su extremo 5'. La sonda de captura puede inmovilizarse indirectamente, de manera que tenga un extremo 3' libre, es decir, extensible.

Por extremo 3' extendido o extensible, se entiende que pueden añadirse nucleótidos adicionales al nucleótido más 3' de la molécula de ácido nucleico, por ejemplo, sonda de captura, para extender la longitud de la molécula de ácido nucleico, es decir, la reacción de polimerización estándar utilizada para extender moléculas de ácido nucleico, por ejemplo, polimerización con plantilla catalizada por una polimerasa.

Por lo tanto, en una modalidad, el sustrato, por ejemplo, la matriz comprende sondas que se inmovilizan directamente a través de su extremo 3', las llamadas sondas de superficie, que se definen más abajo. Cada especie de sonda de superficie comprende una región de complementariedad con cada especie de sonda de captura, de manera que la sonda de captura puede hibridarse con la sonda de superficie, lo que da como resultado la sonda de captura que comprende un extremo 3' libre extensible. En un aspecto preferido de la descripción, cuando el sustrato comprende sondas de superficie, las sondas de captura se sintetizan *in situ* en el sustrato.

Las sondas pueden estar formadas por ribonucleótidos y/o desoxirribonucleótidos, así como también residuos de nucleótidos sintéticos que son capaces de participar en interacciones de tipo Watson-Crick o de pares de bases análogas. Por lo tanto, el dominio de ácido nucleico puede ser ADN o ARN o cualquier modificación del mismo, por ejemplo, PNA u otros derivados que contienen hebras principales no nucleotídicas. Sin embargo, la sonda de captura, por ejemplo, el dominio de captura de la sonda de captura debe ser capaz de cebar una reacción de transcripción inversa para generar ADNc que es complementario a las moléculas de ARN capturadas.

En una modalidad preferida de la invención al menos el dominio de captura de la sonda de captura comprende o consiste en desoxirribonucleótidos (dNTP). En una modalidad particularmente preferida, la totalidad de la sonda de captura comprende o consiste en desoxirribonucleótidos.

En una modalidad preferida de la invención, las sondas de captura se inmovilizan en el sustrato directamente, es decir, por su extremo 5', lo que da como resultado un extremo 3' libre extensible.

Las sondas de captura descritas en la presente descripción comprenden al menos un dominio, un dominio de captura, que es capaz de interactuar con (es decir, unirse o hibridar con) el ARN del tejido, es decir, capturar el ARN. En algunas modalidades en las que las sondas de captura se disponen en el sustrato, las sondas comprenden preferentemente al menos dos dominios, un dominio de captura y un dominio posicional (o una etiqueta o dominio de identificación de elemento; el dominio posicional puede definirse alternativamente como un dominio o etiqueta de identificación (ID), o como una etiqueta posicional). La sonda de captura puede comprender además un dominio universal como se define adicionalmente más abajo. Cuando la sonda de captura se une indirectamente a la superficie de la matriz a través de hibridación a una sonda de superficie, la sonda de superficie requiere una secuencia (por ejemplo, una porción o dominio) que sea complementaria a la sonda de captura. Tal secuencia complementaria puede ser complementaria a un dominio posicional/de identificación (si está presente en la sonda de captura) y/o un dominio universal en la sonda de superficie. En otras palabras, el dominio posicional y/o el dominio universal pueden constituir la región o porción de la sonda que es complementaria a la sonda de superficie. Sin embargo, la sonda de captura también puede comprender un dominio (o región, porción o secuencia) adicional que es complementario a la sonda de superficie. Para facilitar la síntesis, como se describe en más detalle más abajo, tal región complementaria a la sonda de superficie puede proporcionarse como parte de, o como una extensión del dominio de captura (tal como una parte o extensión que no se usa para, o es capaz de, unirse al ácido nucleico diana, por ejemplo, ARN).

Por lo tanto, en su forma más simple, las sondas de captura para su uso en la invención pueden comprender o consistir en un dominio de captura. Sin embargo, en algunas modalidades de la invención, las sondas de captura pueden comprender o consistir en: (i) un dominio de captura y un dominio posicional; (ii) un dominio de captura y un dominio

universal; (iii) un dominio de captura y un dominio que es complementario a una sonda de superficie; (iv) un dominio de captura, dominio posicional y un dominio universal, etcétera.

En algunas modalidades, una sola especie de sonda de captura se inmoviliza al sustrato, preferentemente en donde la sonda de captura se inmoviliza en el sustrato de manera uniforme. Sin embargo, una sola especie de sonda de captura puede disponerse en el sustrato, de manera que cada elemento en la matriz comprende la misma sonda. En algunas modalidades, se inmovilizan múltiples especies de sonda de captura al sustrato, preferentemente en donde cada especie de sonda de captura se inmoviliza en una posición diferente en el sustrato (es decir, cada especie forma un elemento en una matriz), aunque una única sonda de captura puede inmovilizarse en algunas modalidades en más de una posición (puede usarse una sola especie de sonda de captura para formar más de un elemento). En algunas modalidades pueden combinarse múltiples especies de sonda de captura para formar una mezcla que se inmoviliza en el sustrato de manera uniforme, es decir, de manera que hay una distribución uniforme de cada especie de sonda de captura en la superficie del sustrato.

El dominio de captura se ubica típicamente en el extremo 3' de la sonda de captura y comprende un extremo 3' libre que puede extenderse, por ejemplo, mediante polimerización dependiente de plantilla. El dominio de captura comprende una secuencia de nucleótidos que es capaz de hibridarse con ácido nucleico, por ejemplo, ARN (preferentemente ARNm), presente en las células de la muestra de tejido, en contacto con la matriz.

Ventajosamente, el dominio de captura puede seleccionarse o diseñarse para unirse (o dicho más generalmente puede ser capaz de unirse) selectiva o específicamente al ácido nucleico particular, por ejemplo, ARN, que se desea detectar o analizar. Por ejemplo, el dominio de captura puede seleccionarse o diseñarse para la captura selectiva de ARNm. Como se conoce bien en la técnica, esto puede basarse en la hibridación con la cola poli-A del ARNm. Por lo tanto, en una modalidad preferida el dominio de captura comprende un oligonucleótido de ADN poli-T, es decir, una serie de residuos consecutivos de desoxitimidina unidos por enlaces fosfodiéster, que es capaz de hibridarse con la cola poli-A del ARNm. Alternativamente, el dominio de captura puede comprender nucleótidos que son funcional o estructuralmente análogos a poli-T, es decir, son capaces de unirse selectivamente a poli-A, por ejemplo, un oligonucleótido poli-U o un oligonucleótido que comprende análogos de desoxitimidina, en donde dicho oligonucleótido retiene la propiedad funcional de unirse específicamente a poli-A. En una modalidad particularmente preferida el dominio de captura, o más particularmente el elemento poli-T del dominio de captura, comprende al menos 10 nucleótidos, preferentemente al menos 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 nucleótidos. En tales modalidades, el elemento poli-T del dominio de captura puede comprender hasta 30 o hasta 35 nucleótidos. En una modalidad adicional, el dominio de captura, o más particularmente el elemento poli-T del dominio de captura comprende al menos 25, 30 o 35 nucleótidos. Por ejemplo, el elemento poli-T del dominio de captura puede comprender 10-14 nucleótidos, 15-25 nucleótidos o 25-35 nucleótidos.

Las secuencias aleatorias también pueden usarse en la captura de ácido nucleico, como se conoce en la técnica, por ejemplo, hexámeros aleatorios o secuencias similares, y por lo tanto tales secuencias aleatorias pueden usarse para formar todo o una parte del dominio de captura. Por ejemplo, las secuencias aleatorias pueden usarse junto con las secuencias poli-T (o análogo poli-T, etc.). Por lo tanto, cuando un dominio de captura comprende un oligonucleótido poli-T (o "similar a poli-T"), también puede comprender una secuencia de oligonucleótidos aleatoria. Esto puede ubicarse, por ejemplo, 5' o 3' respecto a la secuencia poli-T, por ejemplo, en el extremo 3' de la sonda de captura, pero el posicionamiento de dicha secuencia aleatoria no es crítico. Un constructo de este tipo puede facilitar la captura de la parte inicial del poli-A del ARNm. Alternativamente, el dominio de captura puede ser una secuencia completamente aleatoria. También pueden usarse dominios de captura redundantes, de acuerdo con los principios conocidos en la técnica.

El dominio de captura puede ser capaz de unirse selectivamente a un subtipo o subconjunto de ácido nucleico deseado, por ejemplo, ARN, por ejemplo, un tipo particular de ARN tal como ARNm o ARNr, etc. como se mencionó anteriormente, o a un subconjunto particular de un tipo dado de ARN, por ejemplo, una especie de ARNm particular, por ejemplo, correspondiente a un gen o grupo de genes particular. Una sonda de captura de este tipo puede seleccionarse o diseñarse en base a la secuencia del ARN que se desea capturar. Por lo tanto, puede ser una sonda de captura específica de secuencia, específica para una diana o grupo de dianas de ARN particular (grupo diana, etc.). Por lo tanto, puede basarse en una secuencia génica particular o secuencia de motivo o secuencia común/conservada particular, etc., de acuerdo con principios bien conocidos en la técnica.

En modalidades donde la sonda de captura se inmoviliza en el sustrato indirectamente, por ejemplo, a través de la hibridación con una sonda de superficie, el dominio de captura puede comprender además una secuencia aguas arriba (5' con respecto a la secuencia que se hibrida con el ácido nucleico, por ejemplo, ARN de la muestra de tejido) que es capaz de hibridarse con el extremo 5' de la sonda de superficie. Solo, el dominio de captura de la sonda de captura puede verse como un oligonucleótido del dominio de captura, que puede usarse en la síntesis de la sonda de captura en modalidades donde la sonda de captura se inmoviliza sobre la matriz indirectamente.

El dominio posicional (dominio o etiqueta de identificación del elemento) de la sonda de captura, si está presente, se ubica directa o indirectamente aguas arriba, es decir, más cerca del extremo 5' de la molécula de ácido nucleico de sonda de captura, del dominio de captura. Preferentemente, el dominio posicional es directamente adyacente al

dominio de captura, es decir, no hay secuencia intermedia entre el dominio de captura y el dominio posicional. En algunas modalidades el dominio posicional forma el extremo 5' de la sonda de captura, que puede inmovilizarse directa o indirectamente en el sustrato de la matriz.

Como se describió anteriormente, cuando las sondas de captura se disponen en el sustrato, cada elemento (posición distinta) de la matriz comprende un punto de una especie de sonda de ácido nucleico, en donde preferentemente el dominio posicional en cada elemento es único. En algunas modalidades, el mismo dominio posicional puede usarse para un grupo de elementos, preferentemente un grupo de elementos muy próximas entre sí. Por ejemplo, cuando se usan múltiples especies de sondas de captura que contienen diferentes dominios de captura en el mismo sustrato, puede ser ventajoso usar el mismo dominio posicional para cada tipo de dominio de captura que se inmoviliza en un grupo de elementos directa o indirectamente adyacentes.

Los elementos directamente adyacentes son elementos vecinos en la matriz, por ejemplo, si los elementos se disponen en una formación de rejilla estándar, cualquier elemento único tendrá 8 elementos que son directamente adyacentes. Los elementos indirectamente adyacentes pueden verse como elementos que están muy próximas entre sí, es decir, se inmovilizan dentro de un área específica en la superficie de la matriz, pero pueden tener uno o más elementos que las separan, por ejemplo, 1, 2, 3, 4 o 5 elementos pueden separarse indirectamente los elementos adyacentes. En algunos casos, por ejemplo, en matrices de alta densidad, los elementos indirectamente adyacentes pueden separarse por hasta 10, 20, 30, 40 o 50 elementos.

Una "especie" de la sonda de captura se define con referencia a la secuencia del dominio de captura y, si está presente, su dominio posicional; una sola especie de sonda de captura tendrá una secuencia única del dominio de captura y/o una combinación única de la secuencia del dominio de captura y la secuencia del dominio posicional. Sin embargo, no se requiere que cada miembro de una especie de sonda de captura tenga la misma secuencia en su totalidad. En particular, dado que el dominio de captura puede ser o puede comprender una secuencia aleatoria o redundante, los dominios de captura de sondas individuales dentro de una especie pueden variar. En consecuencia, en algunas modalidades donde los dominios de captura de las sondas de captura son los mismos, cada elemento comprende una única secuencia de sonda. Sin embargo, en otras modalidades donde varía el dominio de captura, los miembros de una especie de sonda no tendrán exactamente la misma secuencia, aunque la secuencia del dominio posicional de cada miembro en la especie será la misma. Lo que se requiere es que cada elemento o posición de la matriz lleve una sonda de captura de una sola especie (específicamente, cada elemento o posición lleva una sonda de captura que tiene una etiqueta posicional idéntica, es decir, hay un solo dominio posicional en cada elemento o posición, aunque las sondas inmovilizadas en elementos directa o indirectamente adyacentes pueden comprender el mismo dominio posicional si comprenden diferentes dominios de captura). Cada especie tiene un dominio de captura, dominio posicional y/o combinación diferente de dominio de captura y dominio posicional que identifica la especie. Sin embargo, cada miembro de una especie, en algunos casos, como se describe en más detalle en la presente descripción, tiene un dominio de captura diferente, ya que el dominio de captura puede ser aleatorio o redundante o puede tener un componente aleatorio o redundante. Esto significa que, dentro de un elemento o posición dada, el dominio de captura de las sondas puede diferir.

Por lo tanto, en algunas, pero no necesariamente en todas las modalidades, la secuencia de nucleótidos de cualquier molécula de sonda inmovilizada en un elemento particular es la misma que las otras moléculas de sonda inmovilizadas en el mismo elemento, pero la secuencia de nucleótidos de las sondas en cada elemento es diferente, distinta o distinguible de las sondas inmovilizadas en cada uno de los otros elementos. Preferentemente cada elemento comprende una especie de sonda diferente. Sin embargo, en algunas modalidades puede ser ventajoso para un grupo de elementos adyacentes para comprender la misma especie de sonda, es decir, de manera efectiva producir un elemento que cubra un área de la matriz que sea mayor que un único elemento, por ejemplo, para disminuir la resolución de la matriz. En otras modalidades, la secuencia de nucleótidos del dominio posicional de cualquier molécula de sonda inmovilizada en un elemento particular puede ser la misma que las otras moléculas de sonda inmovilizadas en el mismo elemento, pero el dominio de captura puede variar. No obstante, el dominio de captura puede diseñarse para capturar el mismo tipo de molécula, por ejemplo, ARNm, en general.

El dominio posicional (o etiqueta) cuando esté presente en la sonda de captura comprende la secuencia que es única para cada elemento o un grupo de elementos directa o indirectamente adyacentes y actúa como un marcador posicional o espacial (la etiqueta de identificación). De esta manera, cada región o dominio de la muestra de tejido, por ejemplo, cada célula en el tejido, puede identificarse mediante resolución espacial a través de la matriz, que enlaza el ácido nucleico, por ejemplo, ARN (por ejemplo, los transcritos) de una determinada célula a una secuencia de dominio posicional único en la sonda de captura. El dominio posicional de una sonda de captura en la matriz es un aspecto de los métodos de la invención que permite que un transcrito específico (o grupo de transcritos) se correlacione con una posición en la muestra de tejido, por ejemplo, puede correlacionarse con una célula en la muestra. Por lo tanto, el dominio posicional del dominio de captura puede verse como una etiqueta de ácido nucleico (etiqueta de identificación) y permite que la posición de un transcrito se correlacione con una posición en la muestra indirectamente, por ejemplo, mediante el análisis de la secuencia del transcrito capturado.

En algunas modalidades de la presente invención, los métodos permiten que un transcrito específico (o grupo de transcritos) se correlacione directamente con una posición en la muestra de tejido, es decir sin la necesidad de análisis

de secuencias. Por ejemplo, una especie de sonda de captura puede inmovilizarse en un sustrato (ya sea en un formato de matriz o uniformemente), en donde el dominio de captura de la sonda de captura es específico para un transcrito (o grupo de transcritos). Solo el/los transcrito(s) específicos interactuarán con (hibridan a) las sondas de captura. Los transcritos se capturarán en el sustrato en una posición con relación a donde el transcrito se expresa en la muestra de tejido. Las etapas de generar y etiquetar las moléculas de ADNc mediante el uso del transcrito capturado como una plantilla y la detección subsecuente, por ejemplo, obtener imágenes, del ADNc etiquetado permitirán que el patrón de expresión espacial del transcrito se determine directamente. La información obtenida a partir de la detección espacial, por ejemplo, imagen, del ADNc etiquetado puede correlacionarse con una imagen de la muestra de tejido para determinar las áreas precisas de expresión en la muestra de tejido.

Será evidente que podrían usarse múltiples especies de sonda de captura, cada una específica para un transcrito diferente para determinar el patrón de expresión de múltiples genes simultáneamente. En una modalidad preferida la especie de sonda de captura se dispone en el sustrato y las sondas de captura comprenden un dominio posicional. En algunas modalidades, los grupos de elementos adyacentes (es decir, que definen un área pequeña en la matriz) pueden comprender sondas de captura con el mismo dominio posicional y diferentes dominios de captura. La etapa de detectar, por ejemplo, obtener imágenes, el ADNc etiquetado proporcionará un análisis preliminar de dónde se expresan los transcritos en la muestra de tejido. Si se requiere o es conveniente un análisis de expresión más específico, las moléculas de ADNc inmovilizadas pueden procesarse y analizarse adicionalmente, como se describe a continuación. El análisis de secuencias del ADNc inmovilizado permite que la posición de un transcrito particular se correlacione con una posición en la muestra de tejido, por ejemplo, una célula, en virtud del dominio posicional en la sonda de captura.

Alternativamente o adicionalmente, el ADNc etiquetado puede usarse como un marcador para seleccionar áreas de ADNc inmovilizadas en el sustrato (que son representativas de la actividad de expresión en la muestra de tejido) para su análisis posterior. Por ejemplo, las moléculas de ADNc inmovilizadas pueden eliminarse de todas las áreas del sustrato excepto el área de interés, por ejemplo, mediante destrucción láser, y las moléculas de ADNc inmovilizadas restantes pueden procesarse y analizarse como se describe en más detalle a continuación. La eliminación de moléculas de ADNc inmovilizadas de la superficie del sustrato, por ejemplo, de regiones que corresponden a células o tejidos en la muestra de tejido que no son de interés o de regiones donde la señal de las moléculas de ADNc etiquetadas indica que los transcritos se expresan por encima o por debajo de un nivel umbral específico, reduce la cantidad de análisis adicional requerido, es decir, menos moléculas de ADNc para el análisis significa que se requiere menos análisis de secuencia, lo que puede resultar en una reducción en la cantidad de reactivos y/o tiempo requerido para realizar el análisis.

Cualquier secuencia adecuada puede usarse como dominio posicional en las sondas de captura descritas en la presente descripción. Por una secuencia adecuada, se entiende que el dominio posicional no debe interferir con (es decir, inhibir o distorsionar) la interacción entre el ARN de la muestra de tejido y el dominio de captura de la sonda de captura. Por ejemplo, el dominio posicional debe diseñarse de manera que las moléculas de ácidos nucleicos en la muestra de tejido no se hibriden específicamente con el dominio posicional. Preferentemente, la secuencia de ácido nucleico del dominio posicional de las sondas de captura tiene menos de 80 % de identidad de secuencia con las secuencias de ácido nucleico en la muestra de tejido. Preferentemente, el dominio posicional de la sonda de captura tiene menos de 70 %, 60 %, 50 % o menos de 40 % de identidad de secuencia en una parte sustancial de las moléculas de ácidos nucleicos en la muestra de tejido. La identidad de secuencia puede determinarse mediante cualquier método apropiado conocido en la técnica, por ejemplo, mediante el uso del algoritmo de alineación BLAST.

En una modalidad preferida el dominio posicional de cada especie de sonda de captura contiene una secuencia de código de barras única. Las secuencias de código de barras pueden generarse mediante el uso de la generación de secuencias aleatorias. A las secuencias generadas aleatoriamente puede seguir un filtro estricto mediante el mapeo a los genomas de todas las especies de referencia comunes y con intervalos de Tm preestablecidos, contenido de GC y una distancia definida de diferencia con las otras secuencias de código de barras para garantizar que las secuencias de código de barras no interfieran con la captura del ácido nucleico, por ejemplo, ARN de la muestra de tejido y se distinguirán entre sí sin dificultad.

Como se mencionó anteriormente, en algunas modalidades, la sonda de captura puede comprender un dominio universal (o dominio enlazador o etiqueta). El dominio universal de la sonda de captura se ubica directa o indirectamente aguas arriba, es decir, más cerca del extremo 5' de la molécula de ácido nucleico de la sonda de captura, del dominio de captura o, si está presente, del dominio posicional. Preferentemente, el dominio posicional es directamente adyacente al dominio de captura o, si está presente, el dominio posicional, es decir, no hay secuencia intermediaria entre el dominio de captura y el dominio universal o el dominio posicional y el dominio universal. En modalidades donde la sonda de captura comprende un dominio universal, el dominio formará el extremo 5' de la sonda de captura, que puede inmovilizarse directa o indirectamente en el sustrato de la matriz.

El dominio universal puede utilizarse de varias maneras en los métodos de la invención. Por ejemplo, en algunas modalidades los métodos de la invención comprenden una etapa de liberar (por ejemplo, eliminar) al menos parte del ácido nucleico sintetizado (es decir, extendido), por ejemplo, moléculas de ADNc de la superficie de la matriz. Como se describe en otra parte de la presente descripción, esto puede lograrse de varias maneras, de las cuales una

comprende escindir el ácido nucleico, por ejemplo, la molécula de ADNc de la superficie de la matriz. Por lo tanto, el dominio universal puede comprender en sí mismo un dominio de escisión, es decir, una secuencia que puede escindirse específicamente, ya sea química o preferentemente enzimáticamente.

Por lo tanto, el dominio de escisión puede comprender una secuencia que se reconoce por una o más enzimas capaces de escindir una molécula de ácido nucleico, es decir, capaz de romper el enlace fosfodiéster entre dos o más nucleótidos. Por ejemplo, el dominio de escisión puede comprender una secuencia de reconocimiento de endonucleasa de restricción (enzima de restricción). Las enzimas de restricción cortan el ADN bicatenario o monocatenario en secuencias de nucleótidos de reconocimiento específicas conocidas como sitios de restricción y las enzimas adecuadas se conocen bien en la técnica. Por ejemplo, es particularmente ventajoso usar enzimas de restricción de corte poco frecuente, es decir, enzimas con un sitio de reconocimiento largo (al menos 8 pares de bases de longitud), para reducir la posibilidad de escindir en otra parte en el ácido nucleico inmovilizado, por ejemplo, molécula de ADNc. Con respecto a esto, se observará que eliminar o liberar al menos parte del ácido nucleico, por ejemplo, la molécula de ADNc requiere liberar una parte que comprende el dominio de captura del ácido nucleico, por ejemplo, ADNc, y toda la secuencia aguas abajo del dominio de captura, por ejemplo, toda la secuencia que es 3' al primer nucleótido en el dominio de captura. Por lo tanto, la escisión del ácido nucleico, por ejemplo, la molécula de ADNc debe tener lugar en 5' al dominio de captura. En modalidades preferidas, eliminar o liberar al menos parte del ácido nucleico, por ejemplo, ADNc, la molécula requiere liberar una parte que comprende el dominio posicional del ácido nucleico, por ejemplo, ADNc y toda la secuencia aguas abajo del dominio posicional, por ejemplo, toda la secuencia que es 3' al primer nucleótido en el dominio posicional. Por lo tanto, la escisión del ácido nucleico, por ejemplo, la molécula de ADNc debe tener lugar en 5' al dominio posicional.

A manera de ejemplo, el dominio de escisión puede comprender una secuencia de poli-U que puede escindirse mediante una mezcla de uracil ADN glucosilasa (UDG) y la endonucleasa ADN glucosilasa-lasa VIII, conocida comercialmente como la enzima USER™.

Un ejemplo adicional de un dominio de escisión puede utilizarse en modalidades donde la sonda de captura se inmoviliza al sustrato de la matriz indirectamente, es decir, a través de una sonda de superficie definida más abajo. El dominio de escisión puede comprender uno o más nucleótidos desapareados, es decir, cuando las partes complementarias de la sonda de superficie y la sonda de captura no son 100 % complementarias. Tal error de apareamiento se reconoce, por ejemplo, por las enzimas endonucleasa I de MutY y T7, lo que da como resultado la escisión de la molécula de ácido nucleico en la posición del error de apareamiento.

En algunas modalidades de la invención, el dominio de captura de la sonda de captura comprende un dominio de escisión, en donde dicho dominio de escisión se ubica en el extremo 5' del dominio de captura. Este dominio de escisión puede verse como un dominio universal o parte del dominio universal. En algunas modalidades de la invención, el dominio posicional de la sonda de captura comprende un dominio de escisión, en donde dicho dominio de escisión se ubica en el extremo 5' del dominio posicional. Este dominio de escisión puede verse como un dominio universal o parte del dominio universal.

El dominio universal puede comprender además un dominio de amplificación. Esto puede ser adicional o en lugar de un dominio de escisión. En algunas modalidades de la invención, como se describe en otra parte de la presente descripción, puede ser ventajoso amplificar las moléculas de ADNc, por ejemplo, después de que se hayan liberado (por ejemplo, eliminado o escindido) del sustrato. Sin embargo, se apreciará que el ciclo inicial de amplificación, o de hecho cualquiera o todos los ciclos de amplificación adicionales, también puede tener lugar *in situ* en el sustrato. El dominio de amplificación comprende una secuencia distinta a la que puede hibridarse un cebador de amplificación. El dominio de amplificación del dominio universal de la sonda de captura es preferentemente idéntico para cada especie de sonda de captura. Por lo tanto, una única reacción de amplificación será suficiente para amplificar todas las moléculas de ácido nucleico, por ejemplo, ADNc (que pueden o no liberarse del sustrato antes de la amplificación).

Cualquier secuencia adecuada puede usarse como dominio de amplificación en las sondas de captura descritas en la presente descripción. Por una secuencia adecuada, se entiende que el dominio de amplificación no debe interferir con (es decir, inhibir o distorsionar) la interacción entre el ácido nucleico, por ejemplo, ARN, de la muestra de tejido, y el dominio de captura de la sonda de captura. Además, el dominio de amplificación debe comprender una secuencia que no sea la misma o sustancialmente la misma que cualquier secuencia en el ácido nucleico, por ejemplo, ARN de la muestra de tejido, de manera que el cebador usado en la reacción de amplificación pueda hibridarse solamente al dominio de amplificación en las condiciones de amplificación de la reacción.

Por ejemplo, el dominio de amplificación debe diseñarse de manera que las moléculas de ácidos nucleicos en la muestra de tejido no se hibriden específicamente con el dominio de amplificación o la secuencia complementaria del dominio de amplificación. Preferentemente, la secuencia de ácido nucleico del dominio de amplificación de las sondas de captura y el complemento de las mismas tiene menos de 80 % de identidad de secuencia con las secuencias de ácido nucleico en la muestra de tejido. Preferentemente, el dominio posicional de la sonda de captura tiene menos de 70 %, 60 %, 50 % o menos de 40 % de identidad de secuencia en una parte sustancial de las moléculas de ácidos nucleicos en la muestra de tejido. La identidad de secuencia puede determinarse mediante cualquier método apropiado conocido en la técnica, por ejemplo, el uso del algoritmo de alineación BLAST.

Solo, el dominio universal de la sonda de captura puede verse como un oligonucleótido del dominio universal, que puede usarse en la síntesis de la sonda de captura en modalidades donde la sonda de captura se inmoviliza sobre la matriz indirectamente.

- 5 En algunas modalidades, el dominio de captura de la sonda de captura puede usarse como un dominio de amplificación. Por ejemplo, en modalidades en las que la sonda de captura no contiene un dominio posicional o dominio universal. En una modalidad representativa, la sonda de captura puede usarse para capturar ARNm a partir de una muestra de tejido, por ejemplo, mediante el uso de un oligonucleótido poli-T. La señal del ADNc etiquetado puede detectarse, por ejemplo, obtener imágenes, y una porción de la matriz puede seleccionarse para un análisis adicional.
- 10 Por lo tanto, las moléculas de ADNc no deseadas pueden eliminarse del sustrato, por ejemplo, mediante el uso de destrucción láser, y las moléculas de ADNc inmovilizadas restantes pueden liberarse y/o amplificarse mediante el uso del dominio de captura como un sitio de cebador, es decir, amplificarse mediante el uso de un cebador de oligonucleótidos poli-A. El análisis de secuencias puede proporcionar información posicional incluso cuando las sondas de captura no contienen dominios posicionales porque la información de secuencia se deriva de moléculas de ácido nucleico amplificadas solo de una región específica del sustrato, que se correlaciona con una región o porción específica de la muestra de tejido.
- 15

En una modalidad representativa de la invención solo el dominio posicional de cada especie de sonda de captura es único. Por lo tanto, los dominios de captura y dominios universales (si están presentes) son en una modalidad los mismos para cada especie de sonda de captura para cualquier matriz particular para garantizar que la captura del ácido nucleico, por ejemplo, ARN, de la muestra de tejido sea uniforme en toda la matriz. Sin embargo, como se describió anteriormente, en algunas modalidades los dominios de captura pueden diferir en virtud de incluir secuencias aleatorias o redundantes o secuencias de genes específicos.

20

En modalidades donde la sonda de captura se inmoviliza en el sustrato indirectamente, por ejemplo, a través de hibridación con una sonda de superficie, la sonda de captura puede sintetizarse en el sustrato como se describe a continuación.

25

Por ejemplo, las sondas de superficie se inmovilizan en el sustrato directamente mediante o en, por ejemplo, su extremo 3'. En modalidades donde las sondas se disponen en el sustrato, cada especie de sonda de superficie puede ser única para cada elemento (posición distinta) o grupos de elementos (elementos directa o indirectamente adyacentes) de la matriz y es parcialmente complementaria a la sonda de captura, definida anteriormente.

30

Por lo tanto, la sonda de superficie comprende en su extremo 5' un dominio (dominio de captura complementario) que es complementario a una parte del dominio de captura que no se une al ácido nucleico, por ejemplo, ARN de la muestra de tejido. En otras palabras, comprende un dominio que puede hibridarse con al menos parte de un oligonucleótido de dominio de captura. La sonda de superficie comprende además un dominio (dominio posicional complementario o dominio de identificación de elemento complementario) que es complementario al dominio posicional de la sonda de captura, si está presente. El dominio posicional complementario se ubica directa o indirectamente aguas abajo (es decir, en el extremo 3') del dominio de captura complementario, es decir, puede haber una secuencia intermedia o enlazadora que separa el dominio posicional complementario y el dominio de captura complementario. En modalidades donde la sonda de captura se sintetiza en la superficie de la matriz, las sondas de superficie de la matriz siempre comprenden un dominio (dominio universal complementario) en el extremo 3' de la sonda de superficie, es decir, directa o indirectamente aguas abajo del dominio posicional (si está presente), que es complementario al dominio universal de la sonda de captura. En otras palabras, comprende un dominio que puede hibridarse con al menos parte de un oligonucleótido de dominio universal.

35

40

45

En algunas modalidades descritas en la presente descripción, la secuencia de la sonda de superficie muestra 100 % de complementariedad o identidad de secuencia con el dominio posicional (si está presente) y el dominio universal y con la parte del dominio de captura que no se une al ácido nucleico, por ejemplo, ARN de la muestra de tejido. En otras modalidades, la secuencia de la sonda de superficie puede mostrar menos del 100 % de identidad de secuencia con los dominios de la sonda de captura, por ejemplo, menos del 99 %, 98 %, 97 %, 96 %, 95 %, 94 %, 93 %, 92 %, 91 % o 90 %. En una modalidad particularmente preferida, el dominio universal complementario comparte menos del 100 % de identidad de secuencia con el dominio universal de la sonda de captura.

50

55

En una modalidad de la invención, la sonda de captura puede sintetizarse o generarse en el sustrato. En una modalidad representativa (véase la figura 1), el sustrato comprende sondas de superficie como se definió anteriormente. Los oligonucleótidos que corresponden al dominio de captura y al dominio universal de la sonda de captura se ponen en contacto con el sustrato y se dejan hibridar con los dominios complementarios de las sondas de superficie. El exceso de oligonucleótidos puede eliminarse mediante el lavado del sustrato en condiciones de hibridación estándar. El sustrato resultante comprende sondas monocatenarias parcialmente, en donde los extremos 5' y 3' de las sondas de superficie son bicatenarios y el dominio posicional complementario es monocatenario. El sustrato puede tratarse con una enzima polimerasa para extender el extremo 3' del oligonucleótido de dominio universal, de una manera dependiente de la plantilla, para sintetizar el dominio posicional de la sonda de captura. El extremo 3' del dominio posicional sintetizado se liga después, por ejemplo, mediante el uso de una enzima ligasa, al extremo 5' del oligonucleótido de dominio de captura para generar la sonda de captura. Se entenderá con respecto a esto que el

60

65

extremo 5' del oligonucleótido de dominio de captura se fosforila para permitir que tenga lugar la ligazón. Como cada especie de sonda de superficie comprende un dominio posicional complementario único, cada especie de sonda de captura comprenderá un dominio posicional único.

Será evidente que en modalidades de la invención donde la sonda de captura no comprende un dominio posicional, puede omitirse la etapa de extensión con polimerasa descrita anteriormente. Por lo tanto, puede permitirse que el dominio de captura y el dominio universal se hibriden con los dominios complementarios de las sondas de superficie. El exceso de oligonucleótidos puede eliminarse mediante el lavado del sustrato en condiciones de hibridación estándar. El extremo 3' del dominio universal se liga después, por ejemplo, mediante el uso de una enzima ligasa, al extremo 5' del oligonucleótido de dominio de captura para generar la sonda de captura.

El término "hibridación" o "se hibrida", como se usa en la presente descripción, se refiere a la formación de un híbrido entre secuencias de nucleótidos que son suficientemente complementarias para formar híbridos a través de apareamiento de bases de tipo Watson-Crick. Dos secuencias de nucleótidos son "complementarias" entre sí cuando esas moléculas comparten homología de organización de pares de bases. Las secuencias de nucleótidos "complementarias" se combinarán con especificidad para formar un híbrido estable en condiciones de hibridación apropiadas. Por ejemplo, dos secuencias son complementarias cuando una sección de una primera secuencia puede unirse a una sección de una segunda secuencia en un sentido antiparalelo en donde el extremo 3' de cada secuencia se une al extremo 5' de la otra secuencia y cada A, T(U), G y C de una secuencia se alinea después con una T(U), A, C y G, respectivamente, de la otra secuencia. Las secuencias de ARN también pueden incluir pares de bases complementarias G=U o U=G. Por lo tanto, dos secuencias no necesitan tener homología perfecta para ser "complementarias" bajo la invención. Usualmente dos secuencias son suficientemente complementarias cuando al menos aproximadamente el 90 % (preferentemente al menos aproximadamente el 95 %) de los nucleótidos comparten la organización de pares de bases en una longitud definida de la molécula. Los dominios de las sondas de captura y de superficie contienen por lo tanto una región de complementariedad. Además, el dominio de captura de la sonda de captura contiene una región de complementariedad para el ácido nucleico, por ejemplo, ARN (preferentemente ARNm) de la muestra de tejido.

La sonda de captura puede sintetizarse en el sustrato mediante el uso de extensión con polimerasa (de manera similar a como se describió anteriormente) y una enzima de transferasa terminal para añadir una "cola" que puede constituir el dominio de captura. Esto se describe adicionalmente en el Ejemplo 5 más abajo. El uso de transferasas terminales para añadir secuencias de nucleótidos al final de un oligonucleótido se conoce en la técnica, por ejemplo, para introducir una cola homopolimérica, por ejemplo, una cola poli-T. En consecuencia, en una síntesis de este tipo, un oligonucleótido que corresponde al dominio universal de la sonda de captura puede ponerse en contacto con el sustrato y dejarse hibridar con el dominio complementario de las sondas de superficie. El exceso de oligonucleótidos puede eliminarse mediante el lavado del sustrato en condiciones de hibridación estándar. El sustrato resultante comprende sondas monocatenarias parcialmente, en donde los extremos 3' de las sondas de superficie son bicatenarios y el dominio posicional complementario es monocatenario. El sustrato puede tratarse con una enzima polimerasa para extender el extremo 3' del oligonucleótido de dominio universal, de una manera dependiente de la plantilla, para sintetizar el dominio posicional de la sonda de captura. El dominio de captura, por ejemplo, que comprende una secuencia poli-T puede introducirse después mediante el uso de una transferasa terminal para añadir una cola poli-T al dominio posicional para generar la sonda de captura. Como se describió anteriormente, en modalidades donde la sonda de captura no comprende un dominio posicional, la etapa de extensión con polimerasa puede omitirse y el dominio de captura puede introducirse mediante el uso de una transferasa terminal para añadir una cola poli-T al dominio universal para generar la sonda de captura.

El sustrato objeto (a menudo simplemente denominado "sustrato") de, y para su uso en los métodos de, la invención puede contener múltiples puntos, o "elementos". En consecuencia, en algunas modalidades de la invención, el sustrato objeto puede ser una matriz, es decir, un sustrato objeto, por ejemplo, portaobjetos o chip, en el que las sondas inmovilizadas se disponen en la superficie del sustrato. Un elemento puede definirse como un área o posición distinta en el sustrato de la matriz en la que se inmoviliza una única especie de sonda de captura. Por lo tanto, cada elemento comprenderá una multiplicidad de moléculas de sonda, de la misma especie. Se entenderá en este contexto que, aunque se abarca que cada sonda de captura de la misma especie puede tener la misma secuencia, este no es necesariamente el caso. Cada especie de sonda de captura tendrá el mismo dominio posicional (es decir, cada miembro de una especie y, por lo tanto, cada sonda en un elemento se "etiquetará" en forma idéntica), pero la secuencia de cada miembro del elemento (especie) puede diferir, porque la secuencia de un dominio de captura puede diferir. Como se describió anteriormente, pueden usarse dominios de captura aleatorios, degenerados o específicos de secuencia. Por lo tanto, las sondas de captura dentro de un elemento pueden comprender diferentes secuencias aleatorias o redundantes. El número y densidad de los elementos en el sustrato, por ejemplo, matriz, determinará la resolución de la matriz, es decir, el nivel de detalle en el que puede analizarse el transcriptoma de la muestra de tejido. Por lo tanto, una mayor densidad de elementos típicamente aumentará la resolución de la matriz.

Como se describió anteriormente, el tamaño y el número de los elementos en el sustrato, por ejemplo, matriz, descritos en la presente descripción dependerán de la naturaleza de la muestra de tejido y de la resolución requerida. Por lo tanto, si es conveniente determinar un transcriptoma solo para regiones de células dentro de una muestra de tejido (o la muestra contiene células grandes), entonces el número y/o densidad de elementos en la matriz puede reducirse (es

decir, menor que el número máximo posible de elementos) y/o el tamaño de los elementos puede aumentarse (es decir, el área de cada elemento puede ser mayor que el elemento más pequeño posible), por ejemplo, una matriz que comprende pocos elementos grandes. Alternativamente, si es conveniente determinar un transcriptoma de células individuales dentro de una muestra, puede ser necesario usar el número máximo de elementos posibles, lo que

En algunas modalidades, las sondas de captura inmovilizadas en el sustrato no están en un formato de matriz, es decir, las sondas de captura pueden distribuirse uniformemente en el sustrato. En estas modalidades, no es necesario que las sondas de captura comprendan un dominio posicional, porque las sondas de captura no se inmovilizan en posiciones específicas en el sustrato. Sin embargo, las sondas de captura pueden comprender un dominio universal.

Por distribuido uniformemente en el sustrato se entiende que las sondas de captura se inmovilizan en al menos una porción del sustrato y hay una cantidad uniforme de sonda de captura inmovilizada en cualquier área específica en esa porción, es decir, la cantidad media de sonda inmovilizada por unidad de área es consistente a través de la porción de sustrato en la que se inmoviliza la sonda. La cantidad media de sonda inmovilizada por unidad de área puede controlarse por la concentración de la sonda en contacto con el sustrato para inmovilizar la sonda o las condiciones usadas para inmovilizar la sonda. Por ejemplo, cuando la sonda de captura se inmoviliza al sustrato indirectamente, por ejemplo, mediante hibridación con una sonda de superficie, las condiciones de hibridación y/o lavado pueden modificarse de manera que no todas las sondas de superficie se unan a una sonda de captura, es decir, en algunas modalidades no todas las sondas de superficie se ocupan por (es decir se unen a) una sonda de captura. Los sustratos en los que la cantidad de sonda inmovilizada por unidad de área es alta tendrán una mayor resolución.

Una porción del sustrato puede ser al menos el 10 % del área total del sustrato. En algunas modalidades, una porción puede ser al menos 20, 30, 40, 50, 60, 70, 80, 90, 95, 99 o 100 % del área total del sustrato. La porción del sustrato en la que las sondas de captura pueden inmovilizarse dependerá del tamaño del sustrato y del tamaño de la muestra de tejido que se pone en contacto con el sustrato. Ventajosamente, el área de la porción del sustrato en la que se inmoviliza la sonda será mayor que el área total de la muestra de tejido. Preferentemente, el área de la porción será al menos 1,2, 3, 4 o 5 % mayor que el área de la muestra de tejido. En algunas modalidades, el área de la porción será al menos 10, 15, 20, 30, 40, 50 % mayor que el área de la muestra de tejido.

Aunque la resolución de células individuales puede ser un elemento preferido y ventajoso de la presente invención, no es esencial lograr esto, y la resolución a nivel del grupo celular también es de interés, por ejemplo, para detectar o distinguir un tipo de célula o región de tejido particular, por ejemplo, células normales frente a tumorales.

En modalidades representativas de la invención, cuando un sustrato es una matriz, la matriz puede contener al menos 2, 5, 10, 50, 100, 500, 750, 1000, 1500, 3000, 5000, 10 000, 20 000, 40 000, 50 000, 75 000, 100 000, 150 000, 200 000, 300 000, 400 000, 500 000, 750 000, 800 000, 1 000 000, 1 200 000, 1 500 000, 1 750 000, 2 000 000, 2 100 000, 3 000 000, 3 500 000, 4 000 000, 4 200 000 o 4 300 000 elementos. Aunque 4 300 000 representa el número máximo de elementos actualmente disponibles en una matriz comercial, se prevé que se puedan preparar matrices con elementos que excedan de esto y tales matrices sean de interés en la presente invención. Por ejemplo, las matrices disponibles comercialmente permiten que se proporcione más de una matriz en un solo sustrato, por ejemplo, portaobjetos o chip. Por lo tanto, el número de elementos en un sustrato puede ser múltiples de las figuras anteriores. Por ejemplo, el sustrato, por ejemplo, matriz, puede comprender al menos 8600000, 12 900 000 o 17 200 000 elementos. Dado que la tecnología de matriz se desarrolla continuamente, la matriz podría contener un número aún mayor de elementos, por ejemplo, se ha postulado que puede ser posible incluir hasta $3,3 \times 10^9$ elementos en una matriz. Como se indicó anteriormente, el tamaño del elemento puede disminuirse y esto puede permitir que se acomoden un mayor número de elementos dentro de la misma o un área similar. A manera de ejemplo, estos elementos pueden estar comprendidos en un área de menos de aproximadamente 20 cm², 10 cm², 5 cm², 1 cm², 1 mm², o 100 μm².

Por lo tanto, en algunas modalidades de la invención, el área de cada elemento puede ser de aproximadamente 1 μm², 2 μm², 3 μm², 4 μm², 5 μm², 10 μm², 12 μm², 15 μm², 20 μm², 50 μm², 75 μm², 100 μm², 150 μm², 200 μm², 250 μm², 300 μm², 400 μm², o 500 μm². En algunas modalidades, el área de un elemento puede ser menor que 1 μm², por ejemplo, menor que 0,5, 0,4, 0,2 o 0,1 μm². En algunas modalidades, por ejemplo, cuando el sustrato objeto comprende un único elemento, el elemento (por ejemplo, el área en la que se inmoviliza una sonda de captura en el sustrato con una distribución homogénea, es decir, uniformemente) puede ser de aproximadamente 100 μm²-20 cm², 1 mm²-10 cm², 0,5 cm²-5 cm², 0,6 cm²-4 cm², 0,7 cm²- 3 cm², 0,8 cm²-2 cm² o 0,9 cm²-1 cm², por ejemplo, al menos de aproximadamente 0,5 cm², 0,6 cm², 0,7 cm², 0,8 cm², 0,9 cm², 1 cm², 2 cm², 3 cm², 4 cm² o 5 cm².

Será evidente que una muestra de tejido de cualquier organismo podría usarse en los métodos de la invención, por ejemplo, planta, animal u hongo. El sustrato, por ejemplo, matriz, descrito en la presente descripción permite la captura de cualquier ácido nucleico, por ejemplo, moléculas de ARNm, que están presentes en células que son capaces de realizar la transcripción y/o traducción. Los sustratos, por ejemplo, matrices, descritos en la presente descripción y los métodos de la invención son particularmente adecuados para aislar y analizar el transcriptoma de las células dentro

de una muestra, en donde la resolución espacial de los transcriptomas es conveniente, por ejemplo, cuando las células están interconectadas o en contacto directo con células adyacentes. Sin embargo, será evidente para un experto en la técnica que los métodos de la invención también pueden ser útiles para el análisis del transcriptoma de diferentes células o tipos de células dentro de una muestra incluso si dichas células no interactúan directamente, por ejemplo, una muestra de sangre. En otras palabras, las células no necesitan estar presentes en el contexto de un tejido y pueden aplicarse a la matriz como células individuales (por ejemplo, células aisladas de un tejido no fijado). Tales células individuales, aunque no se fijan necesariamente a una determinada posición en un tejido, se aplican sin embargo a una determinada posición en el sustrato, por ejemplo, matriz y pueden identificarse individualmente. Por lo tanto, en el contexto del análisis de células que no interactúan directamente, o que no están presentes en un contexto tisular, las propiedades espaciales de los métodos descritos pueden aplicarse para obtener o recuperar información del transcriptoma única o independiente de células individuales. Por lo tanto, una muestra de tejido puede definirse como una muestra de tejido que comprende una o más células. Puede incluir una suspensión de células.

Por lo tanto, la muestra puede ser una muestra de tejido recolectado o de biopsia, o posiblemente una muestra cultivada. Las muestras representativas incluyen muestras clínicas, por ejemplo, sangre completa o productos derivados de la sangre, células sanguíneas, tejidos, biopsias o tejidos o células cultivadas, etc., que incluyen suspensiones celulares. En algunas modalidades, la muestra puede enriquecerse para uno o más tipos de células, por ejemplo, tipos específicos de células sanguíneas o células tumorales. Las técnicas para el aislamiento o enriquecimiento celular se conocen en la técnica, y pueden incluir una selección positiva o negativa en base a la expresión de marcadores celulares particulares. Los tejidos artificiales pueden prepararse, por ejemplo, a partir de la suspensión celular (lo que incluye, por ejemplo, células sanguíneas). Las células pueden capturarse en una matriz (por ejemplo, una matriz de gel, por ejemplo, agar, agarosa, etc.) y después pueden seccionarse de una manera convencional. Tales procedimientos se conocen en la técnica en el contexto de la inmunohistoquímica (ver, por ejemplo, Andersson y otros 2006, J. Histochem. Cytochem. 54(12): 1413-23. Epub 6 de septiembre de 2006).

Los métodos de la invención pueden encontrar particular utilidad en la identificación de células tumorales, especialmente activas transcripcionalmente, por ejemplo, células tumorales, metastásicas o células tumorales con potencial metastásico. A este respecto, las células tumorales pueden encontrarse en la sangre. Aunque aproximadamente el 90 % de la población de células tumorales circulantes puede mostrar una actividad génica nula o pequeña, debido a que las vías de apoptosis o necrosis se activan en esta población, el 10 % restante es transcripcionalmente activo. Es probable que al menos algunas de las células transcripcionalmente activas den lugar a metástasis y sería útil poder identificar estas células. En consecuencia, la muestra de tejido puede ser una suspensión de células que contienen células tumorales (por ejemplo, una muestra de células tumorales aisladas a partir de una muestra de sangre, o una muestra de sangre en la que las células tumorales se han enriquecido, o una biopsia o muestra de tejido, o células tumorales enriquecidas a partir de dicha muestra). El método puede realizarse como se describió anteriormente, en donde la detección del ADNc etiquetado se correlacionará con células tumorales transcripcionalmente activas, por ejemplo, metastásicas o potencialmente metastásicas. Por lo tanto, en algunas modalidades, la invención puede considerarse que proporciona un método para la identificación de células tumorales, especialmente transcripcionalmente activas, por ejemplo, células tumorales potencialmente metastásicas.

El modo de preparación del tejido, y cómo se manipula la muestra resultante puede afectar el análisis transcriptómico de los métodos de la invención. Además, varias muestras de tejido, tendrán diferentes características físicas y está bien dentro del conocimiento de un experto en la técnica realizar las manipulaciones necesarias para producir una muestra de tejido, por ejemplo, muestra de tejido, para su uso con los métodos de la invención. Sin embargo, es evidente a partir de las descripciones en la presente descripción que cualquier método de preparación de muestras puede usarse para obtener una muestra de tejido que es adecuada para su uso en los métodos de la invención. Por ejemplo, cualquier capa de células con un grosor de aproximadamente 1 célula o menos puede usarse en los métodos de la invención. En una modalidad, el grosor de la muestra de tejido, puede ser menor que 0,9, 0,8, 0,7, 0,6, 0,5, 0,4, 0,3, 0,2 o 0,1 de la sección transversal de una célula. Sin embargo, dado que como se indicó anteriormente, la presente invención no se limita a la resolución de células individuales y, por lo tanto, no es un requisito que la muestra de tejido, tenga un grosor del diámetro de una célula o menos; pueden usarse muestras de tejido más delgadas, si se desea. Por ejemplo, pueden usarse secciones de criostato, que pueden tener, por ejemplo, 10-20 μm de grosor.

La muestra de tejido, puede prepararse de cualquier manera conveniente o deseada y la invención no se limita a ningún tipo particular de preparación de tejido. Pueden usarse tejidos frescos, congelados, fijados o no fijados. Cualquier procedimiento conveniente deseado puede usarse para fijar o incrustar la muestra de tejido, como se describe y se conoce en la técnica. Por lo tanto, puede usarse cualquier material de fijación o de incrustación conocido. Como un primer ejemplo representativo de una muestra de tejido para su uso en la invención, el tejido puede prepararse mediante congelación profunda a una temperatura adecuada para mantener o preservar la integridad (es decir, las características físicas) de la estructura tisular, por ejemplo, menos de $-20\text{ }^{\circ}\text{C}$ y preferentemente menos de -25 , -30 , -40 , -50 , -60 , -70 o $-80\text{ }^{\circ}\text{C}$. La muestra de tejido congelado puede seccionarse, es decir, cortarse finamente, en la superficie del sustrato por cualquier medio adecuado. Por ejemplo, una muestra de tejido puede prepararse mediante el uso de un micrótopo frío, un criostato, 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, por ejemplo, a menos de $-15\text{ }^{\circ}\text{C}$ y preferentemente menos de -20 o $-25\text{ }^{\circ}\text{C}$. Por lo tanto, la muestra debe tratarse para minimizar la degeneración o degradación del ácido nucleico, por ejemplo, ARN, en el tejido. Tales condiciones están

bien establecidas en la técnica y la extensión de cualquier degradación puede monitorearse mediante extracción de ácidos nucleicos, por ejemplo, extracción de ARN total, y análisis de calidad subsecuente en varias etapas de la preparación de la muestra de tejido.

En un segundo ejemplo representativo, un tejido puede prepararse mediante el uso de métodos estándar de fijación con formalina e incrustación en parafina (FFPE), que están bien establecidos en la técnica. Después de la fijación de la muestra de tejido y la incrustación en un bloque de parafina o resina, las muestras de tejido pueden cortarse en secciones, es decir, cortarse finamente, en el sustrato, por ejemplo, matriz. Como se indicó anteriormente, pueden usarse otros fijadores y/o materiales de incrustación.

Será evidente que será necesario tratar la sección de muestra de tejido para eliminar el material de incrustación, por ejemplo, para desparafinar, es decir, para eliminar la parafina o resina, de la muestra antes de llevar a cabo los métodos de la invención. Esto puede lograrse mediante cualquier método adecuado y la eliminación de parafina o resina u otro material de las muestras de tejido está bien establecido en la técnica, por ejemplo, mediante la incubación de la muestra (en la superficie de la matriz) en un solvente apropiado, por ejemplo, xileno, por ejemplo, dos veces durante 10 minutos, seguido de un enjuague con etanol, por ejemplo, etanol al 99,5 % durante 2 minutos, etanol al 96 % durante 2 minutos y etanol al 70 % durante 2 minutos.

Será evidente para el experto que es más probable que el ARN en las secciones de tejido preparadas mediante el uso de métodos de FFPE u otros métodos de fijación e incrustación se degrade parcialmente que en el caso del tejido congelado. Sin embargo, sin desear limitarse a ninguna teoría particular, se cree que esto puede ser ventajoso en los métodos de la invención. Por ejemplo, si el ARN en la muestra se degrada parcialmente, la longitud promedio de los polinucleótidos de ARN será menor y más aleatoria que en una muestra no degradada. Por lo tanto, se postula que el ARN parcialmente degradado daría como resultado menos sesgo en las diversas etapas de procesamiento, descritas en otra parte de la presente descripción, por ejemplo, ligazón de adaptadores (dominios de amplificación), amplificación de las moléculas de ADNc y secuenciación de las mismas.

Por lo tanto, en una modalidad de la invención, la muestra de tejido, es decir, la sección de la muestra de tejido en contacto con el sustrato, por ejemplo, matriz, se prepara mediante el uso de FFPE u otros métodos de fijación e incrustación. En otras palabras, la muestra puede fijarse, por ejemplo, fijarse e incrustarse. En una modalidad alternativa de la invención, la muestra de tejido se prepara mediante congelación profunda. En otra modalidad puede usarse una impresión táctil de un tejido, de acuerdo con procedimientos conocidos en la técnica. En otras modalidades puede usarse una muestra no fijada.

El grosor de la sección de muestra de tejido, para su uso en los métodos de la invención puede depender del método usado para preparar la muestra y las características físicas del tejido. Por lo tanto, cualquier grosor de sección adecuado puede usarse en los métodos de la invención. En modalidades representativas de la invención, el grosor de la sección de muestra de tejido será al menos 0,1 μm , aún más preferentemente al menos 0,2, 0,3, 0,4, 0,5, 0,7, 1,0, 1,5, 2, 3, 4, 5, 6, 7, 8, 9 o 10 μm . En otras modalidades, el grosor de la sección de muestra de tejido es al menos 10, 12, 13, 14, 15, 20, 30, 40 o 50 μm . Sin embargo, el grosor no es crítico y estos son solo valores representativos. Las muestras más gruesas pueden usarse si se desea o es conveniente, por ejemplo, 70 o 100 μm o más. Típicamente, el grosor de la sección de muestra de tejido está entre 1-100 μm , 1-50 μm , 1-30 μm , 1-25 μm , 1-20 μm , 1-15 μm , 1-10 μm , 2-8 μm , 3-7 μm o 4-6 μm , pero como se mencionó anteriormente pueden usarse muestras más gruesas.

En el contacto de la sección de muestra de tejido con el sustrato, por ejemplo, después de la eliminación del material de incrustación, por ejemplo, desparafinación, el ácido nucleico, por ejemplo, moléculas de ARN en la muestra de tejido se unirá a las sondas de captura inmovilizadas en el sustrato. En algunas modalidades puede ser ventajoso facilitar la hibridación del ácido nucleico, por ejemplo, moléculas de ARN a las sondas de captura. Típicamente, facilitar la hibridación comprende modificar las condiciones en las que se ocurre la hibridación. Las condiciones primarias que pueden modificarse son el tiempo y la temperatura de la incubación de la sección de tejido, en la matriz antes de la etapa de transcripción inversa, que se describe en otra parte en la presente descripción.

Será evidente que las muestras de tejido de diferentes fuentes pueden requerir diferentes tratamientos para permitir que los ácidos nucleicos interactúen, es decir, se hibriden con, las sondas de captura inmovilizadas en el sustrato. Por ejemplo, puede ser útil permeabilizar la muestra de tejido para facilitar la transferencia de ácido nucleico a la superficie del sustrato. Si la muestra de tejido no se permeabiliza lo suficiente, la cantidad de ácido nucleico capturada en el sustrato puede ser demasiado baja para permitir un análisis adicional (véase la Figura 2A), es decir, la señal de las moléculas de ADNc etiquetadas puede ser de baja intensidad. Por el contrario, si la muestra de tejido es demasiado permeable, el ácido nucleico puede difundirse lejos de su origen en la muestra de tejido. Por lo tanto, el ácido nucleico puede capturarse en el sustrato, pero puede no correlacionarse con precisión con su distribución espacial original en la muestra de tejido (véase la Figura 2C), es decir, la señal de las moléculas de ADNc etiquetadas puede tener una resolución espacial baja. Por tanto, es conveniente un equilibrio entre permeabilizar la muestra de tejido lo suficiente como para obtener una buena intensidad de señal, al tiempo que se mantiene aún la resolución espacial de la distribución del ácido nucleico en la muestra de tejido (véase la figura 2B). Los métodos usados para fijar la muestra de tejido también pueden afectar la transferencia de ácido nucleico desde la muestra de tejido al sustrato.

Los métodos y agentes adecuados para permeabilizar y/o fijar células y tejidos, se conocen bien en la técnica y cualquier método apropiado puede seleccionarse para su uso en los métodos de la invención. Con respecto a esto, los métodos de la invención pueden usarse para determinar las condiciones óptimas, por ejemplo, la combinación óptima de agentes permeabilizantes y/o fijadores, para la captura de ácidos nucleicos a partir de una muestra de tejido particular. Los inventores han descubierto que algunas proteasas son particularmente útiles para permeabilizar células en una muestra de tejido, por ejemplo, pepsina. Los fijadores particularmente útiles incluyen, por ejemplo, metanol. Por lo tanto, los métodos de la invención y el sustrato descritos en la presente descripción son particularmente útiles para determinar las condiciones óptimas para la detección localizada o espacial del transcriptoma de una muestra de tejido. Con respecto a esto, la etapa de etiquetado del ADNc generado en el sustrato es particularmente importante porque permite evaluar la eficacia de la captura de molécula de ácido nucleico. El etiquetado del ADNc permite detectarlo, por ejemplo, visualizarlo, directamente y puede cuantificarse la intensidad y/o resolución de la señal detectada. La intensidad y/o resolución de la señal puede compararse con la señal obtenida a partir del ADNc generado en un sustrato preparado mediante el uso de diferentes métodos, por ejemplo, métodos de permeabilización y/o fijación de tejidos. Los métodos que dan como resultado la mejor intensidad y/o resolución de la señal pueden seleccionarse para su uso en futuros análisis y/o pueden optimizarse más. Los métodos no requieren que las sondas de captura comprendan dominios posicionales y/o dominios universales (aunque las sondas de captura pueden comprender estos dominios en algunas modalidades). Además, las sondas de captura no necesitan disponerse en el sustrato (aunque las sondas de captura pueden disponerse en algunas modalidades). Consecuentemente, los métodos de la invención y los sustratos descritos en la presente descripción son particularmente útiles para ventajosos porque pueden realizarse de manera barata y mediante el uso de instrumentos o aparatos comúnmente disponibles. Una vez que se han determinado las condiciones óptimas para capturar moléculas de ácido nucleico en un sustrato a partir de un tipo particular de muestra de tejido y/o muestra de tejido preparada mediante el uso de un método particular, las condiciones óptimas pueden usarse para analizar muestras de tejido similares, en donde puede realizarse un análisis adicional, tal como el análisis de secuencias. Los métodos de la invención por lo tanto obvian la necesidad de usar sustratos costosos para optimizar las condiciones de captura de ácido nucleico, es decir, matrices que comprenden múltiples especies de sondas de captura inmovilizadas, que contienen dominios posicionales y/o universales.

Como las condiciones para la detección localizada o espacial, de moléculas de ácidos nucleicos de una muestra de tejido en un sustrato varían en dependencia de la muestra de tejido, por ejemplo, muestra de tejido, se describe un intervalo típico de parámetros en la presente descripción. Por ejemplo, al poner en contacto la sección de muestra de tejido, con el sustrato, por ejemplo, una matriz, el sustrato puede incubarse durante al menos 1 hora para permitir que el ácido nucleico, por ejemplo, ARN, se hibride con las sondas de captura. Esto puede ser particularmente útil para las muestras de tejido con FFPE antes de que se elimine la parafina. Preferentemente, el sustrato puede incubarse durante al menos 2, 3, 5, 10, 12, 15, 20, 22 o 24 horas o hasta que la sección de muestra de tejido se haya secado. En otras modalidades, por ejemplo, para muestras de tejido con FFPE después de que la parafina se haya retirado o tejido congelado fresco, el sustrato puede incubarse por menos tiempo, por ejemplo, al menos 5 minutos, por ejemplo, al menos 10, 15, 20, 25 o 30 minutos. El tiempo de incubación del sustrato no es crítico y puede usarse cualquier tiempo conveniente o deseado. Las incubaciones de sustrato típico, por ejemplo, matriz, pueden ser de hasta 72 horas. Por lo tanto, la incubación puede ocurrir a cualquier temperatura adecuada, por ejemplo, a temperatura ambiente, aunque en una modalidad preferida la sección de muestra de tejido, se incuba en la matriz, a una temperatura de al menos 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36 o 37 °C. Las temperaturas de incubación de hasta 55 °C, por ejemplo, 40, 45 o 50°C son comunes en la técnica. En una modalidad particularmente preferida, la sección de muestra de tejido se deja secar en el sustrato, por ejemplo, matriz a 37 °C durante 24 horas. En una modalidad preferida, la sección de muestra de tejido se deja secar en el sustrato, por ejemplo, matriz, a 50 °C durante 15 minutos. Por lo tanto, se entenderá que las condiciones y métodos precisos para poner en contacto la muestra de tejido con la matriz no son críticos y pueden variar de acuerdo con la naturaleza de la muestra y la fijación. Una vez que la sección de muestra de tejido, ha secado el sustrato, puede almacenarse a temperatura ambiente antes de realizar la etapa de transcripción inversa. Se debe entender que, si la sección de muestra de tejido se deja secar sobre la superficie del sustrato, tendrá que rehidratarse antes de que pueda lograrse una manipulación adicional del ácido nucleico capturado, por ejemplo, la etapa de transcribir inversamente el ARN capturado.

Por lo tanto, el método de la invención puede comprender una etapa adicional de rehidratar la muestra de tejido, después de poner en contacto la muestra con el sustrato, por ejemplo, matriz.

En algunas modalidades, la muestra de tejido, por ejemplo, la sección de tejido, puede tratarse o modificarse antes de la etapa de poner en contacto la muestra de tejido con el sustrato y/o antes de generar las moléculas de ADNc en el sustrato, por ejemplo, para seleccionar una o más porciones de la muestra de tejido para el análisis. Por ejemplo, la muestra de tejido puede diseccionarse para aislar una o más porciones para el análisis. Alternativamente visto, la muestra de tejido puede diseccionarse para descartar una o más porciones para las que no se requiere análisis. Cualquier método adecuado para diseccionar la muestra de tejido puede utilizarse en los métodos de la invención. En algunas modalidades, la muestra de tejido se disecciona mediante el uso de microdisección de captura láser (LCM). En consecuencia, el método de la invención puede comprender una etapa de disección de la muestra de tejido. Este aspecto del método comprende una etapa de retener una o más porciones de muestra de tejido para el análisis y/o desechar una o más porciones de muestra de tejido.

En un aspecto preferido de la invención, la muestra de tejido puede diseccionarse en el sustrato. Por ejemplo, la muestra de tejido puede ponerse en contacto con (por ejemplo, fijarse a) una membrana de LCM. El compuesto de muestra de membrana de LCM:tejido puede ponerse en contacto adicionalmente con el sustrato objeto para formar un "sándwich", en el que la muestra de tejido es la capa central en el sándwich. Una o más porciones de la muestra de tejido pueden diseccionarse mediante el uso de un láser, en donde la una o más porciones de la muestra de tejido que no se requieren para el análisis se retiran del sustrato (por ejemplo, se desprenden junto con la membrana de LCM) y se desechan opcionalmente. La membrana de LCM puede retirarse de la una o más porciones restantes de muestra de tejido en el sustrato y el método de la invención puede realizarse como se describe en la presente descripción. En algunas modalidades, la una o más porciones de muestra de tejido retiradas del sustrato pueden usarse para el análisis. Por lo tanto, en algunas modalidades las sondas de captura pueden inmovilizarse en la superficie de la membrana de LCM y/o el sustrato. Alternativamente vista, la membrana de LCM puede ser un sustrato objeto como se describe en la presente descripción. Por lo tanto, la etapa de diseccionar la muestra de tejido puede comprender poner en contacto la muestra de tejido con más de un sustrato, en donde el ARN se captura de una o más porciones de la muestra de tejido en cada sustrato.

En algunas modalidades puede ser ventajoso bloquear (por ejemplo, enmascarar o modificar) las sondas de captura antes de poner en contacto la muestra de tejido, con el sustrato, particularmente cuando el ácido nucleico en la muestra de tejido, o la muestra de tejido en sí, está sujeta a un proceso de modificación antes de su captura en el sustrato. Específicamente, puede ser ventajoso bloquear o modificar el extremo 3' libre de la sonda de captura. Puede ser necesario bloquear o modificar las sondas de captura particularmente el extremo 3' libre de la sonda de captura antes de poner en contacto la muestra de tejido con el sustrato para evitar la modificación de las sondas de captura, por ejemplo, para evitar la eliminación o modificación o el grupo 3' OH libre en el extremo de las sondas de captura. Preferentemente, la incorporación de un dominio de bloqueo puede incorporarse en la sonda de captura cuando se sintetiza. Sin embargo, el dominio de bloqueo puede incorporarse a la sonda de captura después de su síntesis.

En algunas modalidades, las sondas de captura pueden bloquearse por cualquier medio adecuado y reversible que evitaría la modificación de los dominios de captura durante el proceso de modificación del ácido nucleico y/o de la muestra de tejido, que ocurre después de que la muestra de tejido, se ha puesto en contacto con el sustrato. En otras palabras, las sondas de captura pueden enmascararse o modificarse reversiblemente de manera que el dominio de captura de la sonda de captura no comprenda un extremo 3' libre, es decir, de manera que el extremo 3' se elimine o modifique, o se hace inaccesible de manera que la sonda de captura no sea susceptible al proceso o reacción que se usa para modificar el ácido nucleico y/o la muestra de tejido. Alternativamente, si las sondas de captura no se bloquean antes del proceso o reacción usada para modificar el ácido nucleico y/o la muestra de tejido, la sonda de captura puede modificarse después del proceso o reacción para revelar y/o restaurar el extremo 3' del dominio de captura de la sonda de captura.

Por ejemplo, las sondas de bloqueo pueden hibridarse con las sondas de captura para enmascarar el extremo 3' libre del dominio de captura, por ejemplo, sondas en horquilla o sondas bicatenarias parcialmente, cuyos ejemplos adecuados se conocen en la técnica. El extremo 3' libre del dominio de captura puede bloquearse mediante modificación química, por ejemplo, la adición de un grupo azidometilo como un resto protector químicamente reversible, de manera que las sondas de captura no comprenden un extremo 3' libre. Los restos protectores alternativos adecuados se conocen bien en la técnica, por ejemplo, el nucleótido terminal del dominio de captura podría ser un nucleótido terminador reversible, que podría incluirse en la sonda de captura durante o después de la síntesis de la sonda.

Alternativamente o adicionalmente, el dominio de captura de la sonda de captura podría modificarse para permitir la eliminación de cualquier modificación de la sonda de captura, por ejemplo, nucleótidos adicionales, que ocurre cuando se modifican las moléculas de ácido nucleico y/o la muestra de tejido. Por ejemplo, las sondas de captura pueden comprender una secuencia adicional aguas abajo del dominio de captura, es decir, 3' respecto al dominio de captura, específicamente un dominio de bloqueo. Este podría tener la forma de, por ejemplo, una secuencia de reconocimiento de endonucleasa de restricción o una secuencia de nucleótidos escindibles por actividades enzimáticas específicas, por ejemplo, uracilo. Después de la modificación del ácido nucleico y/o la muestra de tejido, las sondas de captura podrían someterse a una escisión enzimática, lo que permitiría la eliminación del dominio de bloqueo y cualquiera de los nucleótidos adicionales que se añaden al extremo 3' de la sonda de captura durante el proceso de modificación. La eliminación del dominio de bloqueo revelaría y/o restauraría el extremo 3' libre del dominio de captura de la sonda de captura. El dominio de bloqueo podría sintetizarse como parte de la sonda de captura o podría añadirse a la sonda de captura *in situ* (es decir, como una modificación de un sustrato existente), por ejemplo, mediante la ligazón del dominio de bloqueo.

Las sondas de captura pueden bloquearse mediante el uso de cualquier combinación de los mecanismos de bloqueo descritos anteriormente.

Una vez que el ácido nucleico y/o muestra de tejido, se ha modificado o procesado para permitirle que el ácido nucleico se hibride con el dominio de captura de la sonda de captura, la sonda de captura debe desbloquearse, por ejemplo, mediante la disociación del oligonucleótido de bloqueo, la eliminación del resto protector y/o el dominio de bloqueo.

Para correlacionar la información del transcriptoma, por ejemplo, intensidad de la señal y/o resolución del ADNc etiquetado, análisis de secuencia, etc., obtenido del sustrato, por ejemplo, al obtener imágenes del sustrato y/o análisis de secuencias del ADNc inmovilizado en uno o más elementos de la matriz, con la región (es decir, un área o célula) de la muestra de tejido, la muestra de tejido se orienta en relación con las sondas inmovilizadas en el sustrato, por ejemplo, se orienta en relación con los elementos de la matriz. En otras palabras, la muestra de tejido se coloca en el sustrato, por ejemplo, matriz, de manera que la posición de una sonda de captura en el sustrato, por ejemplo, matriz puede correlacionarse con una posición en la muestra de tejido. Por lo tanto, puede identificarse en qué lugar de la muestra de tejido, corresponde la posición de cada especie de sonda de captura (o cada elemento de la matriz). En otras palabras, puede identificarse a qué ubicación en la muestra de tejido, corresponde la posición de cada especie de sonda de captura. Esto puede hacerse en virtud de marcadores posicionales presentes sobre la matriz, como se describe a continuación.

Convenientemente, pero no necesariamente, puede obtenerse una imagen de la muestra de tejido, después de su contacto con la matriz. Esto puede realizarse antes o después de que se procese el ácido nucleico de la muestra de tejido, por ejemplo, antes o después de la etapa de generación de ADNc del método, en particular la etapa de generar el ADNc de la primera hebra mediante transcripción inversa. En algunas modalidades, se obtiene una imagen de la muestra de tejido antes de que se marque el ADNc inmovilizado. Sin embargo, la muestra de tejido puede obtener imágenes al mismo tiempo que se obtienen imágenes del ADNc etiquetado. En modalidades en las que el ADNc se libera de la superficie del sustrato, puede obtenerse imágenes de la muestra de tejido antes de la liberación del ADNc capturado y sintetizado (es decir, extendido) del sustrato, por ejemplo, matriz. En una modalidad particularmente preferida, se obtiene una imagen del tejido después de que se ha procesado el ácido nucleico de la muestra de tejido, por ejemplo, después de la etapa de transcripción inversa, y cualquier tejido residual se retira (por ejemplo, se lava) de la matriz antes de detectar, por ejemplo, obtener imágenes, el ADNc etiquetado y/o la liberación de moléculas del sustrato, por ejemplo, matriz. En algunas modalidades, la etapa de procesar el ácido nucleico capturado, por ejemplo, la etapa de transcripción inversa, puede actuar para eliminar el tejido residual de la superficie de la matriz, por ejemplo, cuando se usa preparación de tejido mediante congelación profunda. En tal caso, la obtención de imágenes de la muestra de tejido puede tener lugar antes de la etapa de procesamiento, por ejemplo, la etapa de síntesis de ADNc. En términos generales, la obtención de imágenes puede tener lugar en cualquier momento después de poner en contacto la muestra de tejido con el sustrato, pero antes de cualquier etapa que degrade o elimine la muestra de tejido. Como se indicó anteriormente, esto puede depender de la muestra de tejido.

Ventajosamente, el sustrato, por ejemplo, matriz, puede comprender marcadores para facilitar la orientación de la muestra de tejido, o la imagen del mismo en relación con las sondas de captura inmovilizadas en el sustrato, por ejemplo, los elementos de la matriz. Cualquier medio adecuado para marcar la matriz puede usarse de manera que sean detectables cuando se obtienen imágenes de la muestra de tejido. Por ejemplo, una molécula, por ejemplo, una molécula fluorescente, que genera una señal, preferentemente una señal visible, puede inmovilizarse directa o indirectamente sobre la superficie de la matriz. Preferentemente, la matriz, comprende al menos dos marcadores en posiciones distintas en la superficie del sustrato, aún más preferentemente al menos 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20, 30, 40, 50, 60, 70, 80, 90 o 100 marcadores. Convenientemente pueden usarse varios cientos o incluso varios miles de marcadores. En algunas modalidades, pueden usarse decenas de miles de marcadores. Los marcadores pueden proporcionarse en un patrón, por ejemplo, los marcadores pueden formar un borde exterior de la porción del sustrato en la que se inmovilizan las sondas de captura, por ejemplo, los marcadores pueden ser una hilera de elementos en el borde exterior de la porción en la que las sondas de captura se inmovilizan en el sustrato, por ejemplo, una hilera exterior completa de elementos en una matriz. Pueden usarse otros patrones informativos, por ejemplo, líneas que seccionan la matriz. Esto puede facilitar la alineación de una imagen de la muestra de tejido con la señal detectada a partir de las moléculas de ADNc etiquetadas, (por ejemplo, la imagen de las moléculas de ADNc etiquetadas) es decir, con la porción en la que las sondas de captura se inmovilizan en el sustrato, por ejemplo, una matriz, o de hecho generalmente en la correlación de los elementos de la matriz con la muestra de tejido. Por lo tanto, el marcador puede ser una molécula inmovilizada a la que una molécula generadora de señal puede interactuar para generar una señal. En un ejemplo representativo, el sustrato, por ejemplo, matriz, puede comprender un elemento marcador, por ejemplo, una sonda de ácido nucleico inmovilizada en el sustrato, por ejemplo, matriz, a la que puede hibridarse un ácido nucleico etiquetado. Por ejemplo, la molécula de ácido nucleico etiquetada, o el ácido nucleico marcador, puede estar unido o acoplado a un resto químico capaz de fluorescer cuando se somete a luz de una longitud de onda específica (o intervalo de longitudes de onda), es decir, se excita. Tal molécula de ácido nucleico marcadora puede ponerse en contacto con la matriz antes, al mismo tiempo con, o después de que la muestra de tejido se tiñe para visualizar o tomar imágenes de la muestra de tejido. Sin embargo, el marcador debe detectarse cuando se obtienen imágenes de la muestra de tejido. Por lo tanto, en una modalidad preferida, el marcador puede detectarse mediante el uso de las mismas condiciones de obtención de imágenes usadas para visualizar la muestra de tejido. Además, es ventajoso que el marcador sea detectable cuando se detecta el ADNc etiquetado, por ejemplo, se obtienen imágenes. Por lo tanto, en algunas modalidades el marcador puede detectarse mediante el uso de las mismas condiciones, por ejemplo, condiciones de obtención de imágenes, usadas para detectar la señal del ADNc etiquetado.

En una modalidad particularmente preferida descrita en la presente descripción, el sustrato, por ejemplo, matriz, comprende elementos marcadores a los que se hibrida una molécula de ácido nucleico marcadora, por ejemplo, oligonucleótido, etiquetada, preferentemente etiquetada con fluorescencia.

La etapa de detectar, por ejemplo, de obtener imágenes, el ADNc etiquetado puede usar cualquier medio conveniente conocido en la técnica, pero típicamente comprenderá microscopía, por ejemplo, microscopía de luz, campo claro, campo oscuro, contraste de fase, fluorescencia, reflexión, interferencia, confocal o una de sus combinaciones. Sin embargo, el método usado dependerá del método usado para etiquetar el ADNc sintetizado en la superficie del sustrato. Numerosos métodos para etiquetar moléculas de ácido nucleico, tanto moléculas monocatenarias como bicatenarias se conocen en la técnica. La etiqueta debe dar lugar a una señal visiblemente detectable. Aunque la etiqueta no tiene que ser directamente señal de entrega, esto se prefiere ya que reduce el número de etapas de procesamiento que requeridas para generar una señal. Si se requieren varias etapas para generar una señal a partir del ADNc etiquetado, la señal resultante puede ser inconsistente, es decir, las señales de diferentes áreas del sustrato pueden ser no uniformes.

Una etiqueta directamente detectable es una que puede detectarse directamente sin el uso de reactivos adicionales, mientras que una etiqueta indirectamente detectable es una que es detectable mediante el empleo de uno o más reactivos adicionales, por ejemplo, cuando la etiqueta es un miembro de un sistema productor de señales formado por dos o más componentes. En muchas modalidades, la etiqueta es una etiqueta directamente detectable, donde las etiquetas directamente detectables de interés incluyen, pero no se limitan a: etiquetas fluorescentes, etiquetas coloreadas, etiquetas radioisotópicas, etiquetas quimioluminiscentes y similares. Puede usarse cualquier etiqueta detectable espectrofotométrica u ópticamente. En otras modalidades la etiqueta puede proporcionar una señal indirectamente, es decir, puede requerir la adición de componentes adicionales para generar señal. Por ejemplo, la etiqueta puede ser capaz de unirse a una molécula que se conjuga con una molécula que da la señal.

La etiqueta se incorpora en la parte sintetizada de las moléculas de ADNc, es decir, como parte de las moléculas sintetizadas, por ejemplo, un nucleótido etiquetado, o se une a la parte recién sintetizada de la molécula de ácido nucleico. Por lo tanto, la sonda de captura, o una parte de esta (tal como un dominio posicional o universal), no es una etiqueta para los propósitos de este aspecto de la invención. La función de la etiqueta es indicar áreas en el sustrato en el que se ha capturado el transcrito y se ha sintetizado el ADNc. En consecuencia, una sonda de captura, o parte de esta, no puede lograr esta función ya que su presencia en la superficie del sustrato no está condicionada a la presencia de transcritos capturados en la superficie del sustrato.

En modalidades preferidas, el ADNc se etiqueta mediante la incorporación de un nucleótido etiquetado cuando el ADNc se sintetiza. El nucleótido etiquetado puede incorporarse en la síntesis de la primera y/o segunda hebra. En una modalidad particularmente preferida, el nucleótido etiquetado es un nucleótido etiquetado con fluorescencia. Por lo tanto, el ADNc etiquetado puede obtener imágenes mediante microscopía de fluorescencia. Aunque las etiquetas fluorescentes requieren excitación para proporcionar una señal detectable, dado que la fuente de excitación se deriva del instrumento/aparato usado para detectar la señal, las etiquetas fluorescentes pueden verse como etiquetas que dan señales directamente. Las moléculas fluorescentes que pueden usarse para etiquetar nucleótidos se conocen bien en la técnica, por ejemplo, fluoresceína, los colorantes de cianina, tales como Cy3, Cy5, Alexa 555, Bodipy 630/650, y similares. También pueden emplearse otras etiquetas, tales como las descritas a continuación, como se conoce en la técnica. En modalidades preferidas, el CTP etiquetado con fluorescencia (tal como Cy3-CTP, Cy5-CTP) se incorpora en las moléculas de ADNc sintetizadas en la superficie de la matriz.

Como se mencionó anteriormente, las etiquetas pueden incorporarse en el ADNc sintetizado mediante la unión a las moléculas, por ejemplo, a través de intercalación. Las moléculas detectables representativas que pueden ser útiles en tales modalidades incluyen tinciones de ácidos nucleicos fluorescentes, tales como colorantes de fenantridinio, que incluyen monómeros u homodímeros o heterodímeros de los mismos, que dan una fluorescencia mejorada cuando forman complejos con ácidos nucleicos. Los ejemplos de colorantes de fenantridinio incluyen homodímero de etidio, bromuro de etidio, yoduro de propidio y otros colorantes de fenantridinio sustituidos con alquilo. En otra modalidad de la invención, la tinción de ácido nucleico es o incorpora un colorante de acridina, o un homodímero o heterodímero del mismo, tal como naranja de acridina, homodímero de acridina, heterodímero de etidio-acridina o 9-amino-6-cloro-2-metoxiacridina. Aún en otra modalidad de la invención, la tinción de ácido nucleico es un colorante indólico o imidazólico, tal como Hoechst 33258, Hoechst 33342, Hoechst 34580 (BIOPROBES 34, Molecular Probes, Inc. Eugene, Oreg., (mayo de 2000)) DAPI (4',6-diamidino-2-fenilindol) o DIPI (4',6-(diimidazolin-2-il)-2-fenilindol). Otras tinciones de ácido nucleico permitidas incluyen, pero no se limitan a, 7-aminoactinomicina D, hidroxiestilbamidina, LDS 751, psoralenos seleccionados (furocumarinas), colorantes de estirilo, complejos con metales tales como complejos de rutenio y complejos con metales de transición (que incorporan Tb^{3+} y Eu^{3+} , por ejemplo). En ciertas modalidades de la invención, la tinción de ácido nucleico es un colorante de cianina o un homodímero o heterodímero de un colorante de cianina que da una fluorescencia potenciada cuando se asocia con ácidos nucleicos. Cualquiera de los colorantes descritos en la patente de Estados Unidos núm. 4,883,867 de Lee (1989), la patente de Estados Unidos núm. 5,582,977 de Yue y otros (1996), la patente de Estados Unidos núm. 5,321,130 de Yue y otros (1994), y la patente de Estados Unidos núm. 5,410,030 de Yue y otros (1995) puede usarse, que incluye las tinciones de ácidos nucleicos disponibles comercialmente bajo las marcas comerciales TOTO, BOBO, POPO, YOYO, TO-PRO, BO-PRO, PO-PRO y YO-PRO de Molecular Probes, Inc., Eugene, Oreg. Cualquiera de los colorantes descritos en la patente de Estados Unidos núm. 5,436,134 de Haugland y otros (1995), la patente de Estados Unidos núm. 5,658,751 de Yue y otros (1997), y la patente de Estados Unidos núm. 5,863,753 de Haugland y otros (1999) puede usarse, que incluye las tinciones de ácidos nucleicos disponibles comercialmente bajo las marcas comerciales SYBR Green, SYBR Gold, EvaGreen, SYTO, SYTOX, PICOGREEN, OLIGREEN y RIBOGREEN de Molecular Probes, Inc., Eugene, Oreg. En

aún otras modalidades de la invención, la tinción de ácido nucleico es un colorante de cianina monomérico, homodimérico o heterodimérico que incorpora un heterociclado de aza- o poliazabenzazolio, tal como un azabenzoxazol, azabencimidazol o azabenzotiazol, que da una fluorescencia mejorada cuando se asocia con ácidos nucleicos, que incluye tinciones de ácidos nucleicos disponibles comercialmente bajo las marcas comerciales SYTO, SYTOX, JOJO, JO-PRO, LOLO, LO-PRO de Molecular Probes, Inc., Eugene, Oreg. El tipo de tinción de ácido nucleico puede seleccionarse en base a su capacidad para unirse al ácido nucleico de hebra única o de doble hebra. En modalidades donde se etiqueta la primera hebra de ADNc, puede ser preferible usar tinciones de ácidos nucleicos capaces de etiquetar moléculas de ácidos nucleicos de hebra única cuando el transcrito de ARN capturado en el sustrato y usado para la síntesis de ADNc con plantilla puede degradarse parcial o completamente.

En una modalidad particularmente ventajosa de la invención, los métodos pueden incluir una etapa de eliminar una porción de las moléculas de ácido nucleico inmovilizadas en la superficie del sustrato. Esta etapa puede ser particularmente ventajosa para analizar el transcriptoma de parte o porción de una muestra de tejido, por ejemplo, un área de interés, célula particular o tipo de tejido, etc. Eliminar moléculas de ácido nucleico inmovilizadas en la superficie del sustrato que no son de interés puede reducir los costos y/o el tiempo requerido para etapas de análisis adicionales. Se necesitan menos reactivos para realizar el análisis de secuencias en moléculas de ADNc a partir de una porción de la muestra de tejido en comparación con los reactivos necesarios para realizar el análisis de secuencias en moléculas de ADNc a partir de toda la muestra de tejido. En consecuencia, se requiere menos análisis de secuencias para las moléculas de ADNc de una porción de la muestra de tejido. Un beneficio adicional derivado de eliminar una porción de las moléculas de ácido nucleico de la superficie del sustrato es que no es necesario incluir un dominio posicional en la sonda de captura para correlacionar el análisis de secuencias con una posición en la muestra de tejido. Con respecto a esto, debido a que las moléculas de ADNc que no son de interés (es decir, de áreas de la muestra de tejido que no son de interés) se han eliminado del sustrato, las secuencias analizadas se habrán derivado del área específica de la muestra de tejido que se correlaciona con la porción de moléculas de ADNc que no se eliminaron del sustrato. En consecuencia, no es necesario inmovilizar las sondas de captura en el sustrato en forma de una matriz. Sin embargo, en muchas modalidades de la invención, las sondas de captura comprenden un dominio posicional y/o se inmovilizan en el sustrato en un formato de matriz, es decir, el sustrato es una matriz.

La señal obtenida a partir de la etapa de detectar, por ejemplo, obtener imágenes, las moléculas de ADNc etiquetadas permiten seleccionar áreas de la muestra de tejido para su análisis posterior. En su forma más simple, una única porción discreta del ADNc inmovilizado en el sustrato (correspondiente a una porción de la muestra de tejido) puede seleccionarse para un análisis adicional y pueden eliminarse todas las demás moléculas de ácido nucleico inmovilizadas en la superficie del sustrato. Será evidente que más de una porción puede seleccionarse para un análisis adicional, por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más porciones. En algunas modalidades pueden seleccionarse 15, 20, 25, 30, 40, 50 o más porciones. En consecuencia, las moléculas de ADNc de una o más porciones pueden seleccionarse para su eliminación, por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más porciones. En algunas modalidades pueden seleccionarse 15, 20, 25, 30, 40, 50 o más porciones. En algunas modalidades, las porciones seleccionadas para el análisis (o eliminación) adicional pueden basarse en tipos de células o tejidos, por referencia una imagen de la muestra de tejido y correlacionar la posición en la muestra de tejido con la posición en el sustrato. En aún modalidades adicionales, las porciones seleccionadas para análisis (o eliminación) adicionales pueden basarse en la cantidad de ADNc inmovilizado en un área particular. Por ejemplo, las porciones del sustrato en las que la intensidad de la señal de las moléculas de ADNc etiquetadas inmovilizadas está por encima y/o por debajo de los límites umbrales específicos pueden dirigirse para su eliminación (o análisis adicional). Por lo tanto, las porciones en el sustrato que se correlacionan con partes de la muestra de tejido con niveles altos de transcripción pueden eliminarse para enriquecer la proporción del transcriptoma analizado para partes de la muestra de tejido con niveles moderados o bajos de transcripción. De manera similar, el análisis puede centrarse en partes de la muestra de tejido con altos niveles de transcripción.

La etapa de retirar una porción de las moléculas de ácido nucleico inmovilizadas en la superficie del sustrato puede lograrse mediante el uso de cualquier medio conveniente. En algunas modalidades, las moléculas de ADNc inmovilizadas pueden eliminarse mediante destrucción láser, por ejemplo, la(s) porción(es) del sustrato del cual se van a eliminar las moléculas de ADNc pueden identificarse mediante la detección, por ejemplo, obtener imágenes, las moléculas de ADNc etiquetadas y aquellas áreas sometidas a tratamiento con un láser que es suficiente para eliminar las moléculas de ADNc de la superficie del sustrato. Ventajosamente, la destrucción láser también puede retirar la muestra de tejido de las áreas etiquetadas. En algunas modalidades, la muestra de tejido puede eliminarse del sustrato, como se describe en otra parte en la presente descripción, antes de la eliminación de una porción de las moléculas de ADNc inmovilizadas. Los instrumentos y aparatos adecuados para eliminar moléculas de ADNc de la superficie de una matriz se conocen en la técnica, por ejemplo, un instrumento de corte celular MMI (Molecular Machines and Industries AG, Glattnburg, Suiza). Otros medios para eliminar moléculas de ADNc de la superficie del sustrato pueden incluir escisión, por ejemplo, escisión enzimática. Por ejemplo, la(s) porción(es) del sustrato para un análisis adicional pueden enmascarse (por ejemplo, mediante el uso de un aparato de máscara de matriz estándar) y las áreas no enmascaradas del sustrato sometidas a un agente de escisión, por ejemplo, una enzima, para liberar las sondas de captura (y moléculas de ADNc unidas) de la superficie del sustrato, como se describió anteriormente. La porción del sustrato seleccionado para un análisis adicional puede desenmascarse y el análisis subsecuente puede realizarse de acuerdo con las etapas del método restantes descritas en la presente descripción.

La etapa de obtención de imágenes del tejido puede usar cualquier medio histológico conveniente conocido en la técnica, por ejemplo, microscopía de luz, campo claro, campo oscuro, contraste de fase, fluorescencia, reflexión, interferencia, confocal o una de sus combinaciones. Típicamente, la muestra de tejido, se tiñe antes de la visualización para proporcionar contraste entre las diferentes regiones, por ejemplo, células, de la muestra de tejido. El tipo de tinción usado dependerá del tipo de tejido, y la región de las células a teñir. Tales protocolos de tinción se conocen en la técnica. En algunas modalidades puede usarse más de una tinción para visualizar (mediante imagen) diferentes aspectos de la muestra de tejido, por ejemplo, diferentes regiones de la muestra de tejido, estructuras celulares específicas (por ejemplo, organelos) o diferentes tipos de células. En otras modalidades, puede visualizarse u obtenerse imágenes de la muestra de tejido, sin teñir la muestra, por ejemplo, si la muestra de tejido ya contiene pigmentos que proporcionan suficiente contraste o si se usan formas particulares de microscopía.

En una modalidad preferida, se visualiza o se obtienen imágenes de la muestra de tejido mediante el uso de microscopía de fluorescencia. En consecuencia, en algunas modalidades, la muestra de tejido y el ADNc etiquetado pueden visualizarse u obtener imágenes al mismo tiempo o secuencialmente mediante el uso del mismo aparato de obtención de imágenes.

La muestra de tejido, es decir, cualquier tejido residual que permanezca en contacto con el sustrato después de la etapa de transcripción inversa y detecte, por ejemplo, obtener imágenes, el ADNc etiquetado y opcionalmente obtener imágenes de la muestra de tejido si se desea obtener imágenes de la muestra de tejido y no se llevó a cabo antes de la transcripción inversa, preferentemente se retira antes de la etapa de liberar las moléculas de ADNc del sustrato. Por lo tanto, los métodos de la invención pueden comprender una etapa de lavado del sustrato. La eliminación de la muestra de tejido residual puede realizarse mediante el uso de cualquier técnica adecuada y dependerá de la muestra de tejido. En la modalidad más simple, el sustrato puede lavarse con agua. El agua puede contener varios aditivos, por ejemplo, surfactantes (por ejemplo, detergentes), enzimas, etc. para facilitar la eliminación del tejido. En algunas modalidades, el sustrato se lava con una solución que comprende una enzima proteínasa (y tampón adecuado) por ejemplo, proteínasa K. En otras modalidades, la solución puede comprender, además o alternativamente, enzimas de celulasa, hemicelulasa o quitinasa, por ejemplo, si la muestra de tejido es de una planta o fuente fúngica. En modalidades adicionales, la temperatura de la solución usada para lavar el sustrato puede ser, por ejemplo, al menos 30 °C, preferentemente al menos 35, 40, 45, 50 o 55 °C. Será evidente que la solución de lavado debe minimizar la interrupción de las moléculas de ácido nucleico inmovilizadas. Por ejemplo, en algunas modalidades las moléculas de ácido nucleico pueden inmovilizarse en el sustrato indirectamente, por ejemplo, a través de la hibridación de la sonda de captura y el ARN y/o la sonda de captura y la sonda de superficie, por lo tanto, la etapa de lavado no debe interferir con la interacción entre las moléculas inmovilizadas en el sustrato, es decir, no debe provocar que las moléculas de ácido nucleico se desnaturalicen.

Después de la etapa de poner en contacto el sustrato con una muestra de tejido, en condiciones suficientes para permitir que se produzca la hibridación entre el ácido nucleico, por ejemplo, ARN (preferentemente ARNm), de la muestra de tejido a las sondas de captura, tiene lugar la etapa de asegurar (adquirir) el ácido nucleico hibridado. Asegurar o adquirir el ácido nucleico capturado implica extender la sonda de captura para producir una copia del ácido nucleico capturado, por ejemplo, generar ADNc a partir del ARN capturado (hibridado). Se debe entender que esto se refiere a la síntesis de una hebra complementaria del ácido nucleico hibridado, por ejemplo, generar ADNc en base a la plantilla de ARN capturado (el ARN hibridado al dominio de captura de la sonda de captura). Por lo tanto, en una etapa inicial de extensión de la sonda de captura, por ejemplo, la generación de ADNc, el ácido nucleico capturado (hibridado), por ejemplo, ARN, actúa como una plantilla para la etapa de extensión, por ejemplo, transcripción inversa. En algunas modalidades, asegurar o adquirir el ácido nucleico de captura puede verse como etiquetar o marcar el ácido nucleico capturado con el dominio posicional específico del elemento en el que se captura el ácido nucleico. En muchas modalidades, la etapa de asegurar o adquirir el ácido nucleico capturado implica incorporar directamente una etiqueta que da la señal en la copia sintetizada de la molécula de ácido nucleico capturada.

La transcripción inversa se refiere a la etapa de síntesis de ADNc (ADN complementario o copia) a partir de ARN, preferentemente ARNm (ARN mensajero), mediante transcriptasa inversa. Por lo tanto, el ADNc puede considerarse una copia del ARN presente en una célula en el momento en que se tomó la muestra de tejido, es decir, representa todos o algunos de los genes que se expresaron en dicha célula en el momento del aislamiento.

La sonda de captura, específicamente el dominio de captura de la sonda de captura, actúa como un cebador para producir la hebra complementaria del ácido nucleico hibridado a la sonda de captura, por ejemplo, un cebador para la transcripción inversa. Por lo tanto, las moléculas de ácido nucleico, por ejemplo, ADNc, generadas por la reacción de extensión, por ejemplo, reacción de transcripción inversa, incorporan la secuencia de la sonda de captura, es decir, la reacción de extensión, por ejemplo, reacción de transcripción inversa. Ventajosamente, las moléculas generadas por la reacción de extensión incorporan directamente una etiqueta de manera que la cantidad de transcrito en la muestra de tejido que está en contacto con el sustrato puede determinarse, por ejemplo, al medir la intensidad de la señal generada por la etiqueta. Como se mencionó anteriormente, en algunas modalidades, cada especie de sonda de captura puede comprender un dominio posicional (etiqueta de identificación del elemento) que representa una secuencia única para cada elemento de la matriz. Por lo tanto, en algunas modalidades todas las moléculas de ácidos nucleicos, por ejemplo, ADNc, sintetizadas en un elemento específico comprenderán la misma "etiqueta" de ácido nucleico.

Las moléculas de ácido nucleico, por ejemplo, ADNc sintetizadas en una posición o área específica sobre la superficie del sustrato, por ejemplo, cada elemento de una matriz, puede representar los genes expresados a partir de la región o área de la muestra de tejido en contacto con esa posición o área, por ejemplo, elemento. Por ejemplo, un tejido o tipo de célula o grupo o subgrupo de estos, y puede representar además genes expresados en condiciones específicas, por ejemplo, en un momento particular, en un ambiente específico, en una etapa de desarrollo o en respuesta a estímulos, etc. Por lo tanto, el ADNc en cualquier posición o área única, por ejemplo, elemento, puede representar los genes expresados en una célula única, o si la posición o área, por ejemplo, elemento, está en contacto con la muestra en una unión celular, el ADNc puede representar los genes expresados en más de una célula. De manera similar, si una célula única está en contacto con un área grande del sustrato, por ejemplo, múltiples elementos, entonces cada posición dentro del área, por ejemplo, cada elemento, puede representar una proporción de los genes expresados en dicha célula.

La etapa de extensión de la sonda de captura, por ejemplo, la transcripción inversa, puede realizarse mediante el uso de cualquier enzima y protocolo adecuados de los que existen muchos en la técnica, como se describe en detalle más abajo. Sin embargo, será evidente que no es necesario proporcionar un cebador para la síntesis de la primera hebra de ácido nucleico, por ejemplo, ADNc, porque el dominio de captura de la sonda de captura actúa como el cebador, por ejemplo, cebador de transcripción inversa.

Preferentemente, en el contexto de la presente invención, el ácido nucleico asegurado (es decir, el ácido nucleico unido covalentemente a la sonda de captura), por ejemplo, ADNc, se trata para comprender ADN bicatenario. El tratamiento del ácido nucleico capturado para producir ADN bicatenario puede lograrse en una única reacción para generar solo un segundo ADN, por ejemplo, ADNc, hebra, es decir, para producir moléculas de ADN bicatenario sin aumentar el número de moléculas de ADN bicatenario, o en una reacción de amplificación para generar múltiples copias de la segunda hebra, que puede estar en forma de ADN monocatenario (por ejemplo, amplificación lineal) o ADN bicatenario, por ejemplo, ADNc (por ejemplo, amplificación exponencial).

La etapa de síntesis de ADN de la segunda hebra, por ejemplo, ADNc, puede tener lugar *in situ* en el sustrato, por ejemplo, ya sea como una etapa discreta de síntesis de la segunda hebra, por ejemplo, mediante el uso de cebadores aleatorios como se describe con más detalle más abajo, o en la etapa inicial de una reacción de amplificación. Alternativamente, el ADN de la primera hebra, por ejemplo, ADNc (la hebra que comprende, es decir que incorpora, la sonda de captura) puede liberarse de la matriz y la síntesis de la segunda hebra, ya sea como una etapa discreta o en una reacción de amplificación puede ocurrir subsecuentemente, por ejemplo, en una reacción llevada a cabo en la solución.

Cuando la síntesis de la segunda hebra tiene lugar en el sustrato (es decir, *in situ*), el método puede incluir una etapa opcional de eliminación del ácido nucleico capturado, por ejemplo, ARN antes de la síntesis de la segunda hebra, por ejemplo, mediante el uso de una enzima de digestión de ARN (ARNasa), por ejemplo, ARNasa H. Los procedimientos para esto se conocen bien y se describen en la técnica. Sin embargo, esto generalmente no es necesario, y en la mayoría de los casos el ARN se degrada naturalmente. La eliminación de la muestra de tejido de la matriz generalmente eliminará el ARN de la matriz. La ARNasa H puede usarse si se desea para aumentar la robustez de la eliminación del ARN. La eliminación del ARN puede ser útil en modalidades donde el ADNc se etiqueta después de que se ha generado, por ejemplo, se etiqueta con una tinción de ácido nucleico. La eliminación del ARN puede proporcionar una diana consistente a la que la tinción de ácido nucleico puede interactuar (aglutinarse), es decir, todas las moléculas inmovilizadas serán monocatenarias después de la eliminación del ARN. Antes de una etapa de eliminación del ARN, las moléculas inmovilizadas pueden ser una mezcla de moléculas completas o parcialmente bicatenarias (híbridos de ARN:ADN) y moléculas monocatenarias (donde el ARN ya se ha degradado). Algunas cepas de ácidos nucleicos pueden proporcionar una señal más fuerte cuando interactúan con ácido nucleico bicatenario, en comparación con la señal del ácido nucleico monocatenario. Por lo tanto, es preferible cuando se usa una tinción de ácido nucleico para etiquetar el ADNc inmovilizado que las moléculas sean todas monocatenarias o bicatenarias.

En muestras de tejido que comprenden grandes cantidades de ARN, la etapa de generación del ADNc bicatenario puede producir una cantidad suficiente de ADNc que puede secuenciarse directamente (después de la liberación del sustrato). En este caso, la síntesis de ADNc de segunda hebra puede lograrse por cualquier medio conocido en la técnica y como se describe más abajo. La reacción de síntesis de la segunda hebra puede realizarse en el sustrato directamente, es decir, mientras el ADNc se inmoviliza en el sustrato, o preferentemente después de que el ADNc se ha liberado del sustrato, como se describe a continuación.

En otras modalidades puede ser necesario mejorar, es decir amplificar, la cantidad de ácido nucleico asegurado, por ejemplo, ADNc sintetizado, para producir cantidades que sean suficientes para la secuenciación de ADN. En esta modalidad, la primera hebra de las moléculas de ácido nucleico, por ejemplo, ADNc, que comprenden además la sonda de captura del sustrato, actúa como una plantilla para la reacción de amplificación, por ejemplo, una reacción en cadena de la polimerasa. El primer producto de la reacción de la amplificación será una segunda hebra de ADN, por ejemplo, ADNc, que actuará como una plantilla para ciclos adicionales de la reacción de amplificación.

En cualquiera de las modalidades descritas anteriormente, la segunda hebra de ADN, por ejemplo, ADNc comprenderá un complemento de la sonda de captura. Si la sonda de captura comprende un dominio universal, y particularmente

un dominio de amplificación dentro del dominio universal entonces esto puede usarse para la subsecuente amplificación del ADN, por ejemplo, ADNc, por ejemplo, la reacción de amplificación puede comprender un cebador con la misma secuencia que el dominio de amplificación, es decir, un cebador que es complementario (es decir, se hibrida) al complemento del dominio de amplificación. En vista del hecho de que el dominio de amplificación está aguas arriba del dominio posicional (si está presente) de la sonda de captura (en el ácido nucleico asegurado, por ejemplo, la primera hebra de ADNc), el complemento del dominio posicional (si el dominio de posición está presente en la sonda de captura) se incorporará en la segunda hebra de las moléculas de ADN, por ejemplo, ADNc.

En modalidades donde la segunda hebra de ADN, por ejemplo, ADNc, se genera en una única reacción, la síntesis de la segunda hebra puede lograrse por cualquier medio adecuado. Por ejemplo, el ADNc de la primera hebra, preferentemente, pero no necesariamente, liberado del sustrato, puede incubarse con cebadores aleatorios, por ejemplo, cebadores hexaméricos, y una ADN polimerasa, preferentemente una polimerasa de desplazamiento de hebra, por ejemplo, klenow (exo⁻), en condiciones suficientes para que ocurra la síntesis de ADN con plantilla. Este proceso producirá moléculas de ADNc de doble hebra de longitudes variables y es poco probable que produzca moléculas de ADNc de longitud completa, es decir moléculas de ADNc que correspondan a todo el ARNm a partir del cual se sintetizaron. Los cebadores aleatorios se hibridarán con las moléculas de ADNc de la primera hebra en una posición aleatoria, es decir, dentro de la secuencia en lugar del extremo de la secuencia.

Si es conveniente generar moléculas de ADN de longitud completa, por ejemplo, ADNc, es decir moléculas que corresponden a la totalidad de la molécula de ácido nucleico capturada, por ejemplo, ARN, (si el ácido nucleico, por ejemplo, ARN, se degrada parcialmente en la muestra de tejido entonces las moléculas de ácido nucleico capturadas, por ejemplo, ARN, no serán transcritos de "longitud completa"), entonces el extremo 3' de las moléculas de ácido nucleico asegurado, por ejemplo, ADNc de la primera hebra, pueden modificarse. Por ejemplo, un enlazador o adaptador puede ligarse al extremo 3' de las moléculas de ADNc. Esto puede lograrse mediante el uso de enzimas de ligazón de hebras únicas tales como T4 ARN ligasa o Circuligase™ (Epicentre Biotechnologies).

Alternativamente, una sonda auxiliar (una molécula de ADN bicatenario parcialmente capaz de hibridarse con el extremo 3' de la molécula de ADNc de la primera hebra), puede ligarse al extremo 3' de la molécula de ácido nucleico asegurado, por ejemplo, molécula de ADNc de la primera hebra, mediante el uso de una enzima de ligazón de doble hebra tal como la T4 ADN ligasa. Otras enzimas apropiadas para la etapa de ligazón se conocen en la técnica e incluyen, por ejemplo, Tth ADN ligasa, Taq ADN ligasa, ADN ligasa de *Thermococcus* sp. (cepa 9°N) (9°N™ ADN ligasa, New England Biolabs) y Ampligase™ (Epicentre Biotechnologies). La sonda auxiliar comprende además una secuencia específica a partir de la cual la síntesis de ADN de la segunda hebra, por ejemplo, ADNc, puede cebarse mediante el uso de un cebador que es complementario a la parte de la sonda auxiliar que se liga al ácido nucleico asegurado, por ejemplo, la primera hebra de ADNc. Una alternativa adicional comprende el uso de una enzima activa transferasa terminal para incorporar una cola de polinucleótido, por ejemplo, una cola poli-A, en el extremo 3' del ácido nucleico asegurado, por ejemplo, primera hebra de moléculas de ADNc. La síntesis de la segunda hebra puede cebarse mediante el uso de un cebador poli-T, que también puede comprender un dominio de amplificación específico para la amplificación adicional. Otros métodos para generar moléculas de ADN bicatenario de "longitud completa", por ejemplo, ADNc (o síntesis de la segunda hebra de longitud máxima) están bien establecidos en la técnica.

En algunas modalidades, la síntesis de la segunda hebra puede usar un método de cambio de plantilla, por ejemplo, mediante el uso de la tecnología SMART™ de Clontech®. La tecnología SMART (mecanismo de cambio en el extremo 5' de la plantilla de ARN) está bien establecida en la técnica y se basa en el descubrimiento de enzimas transcriptasas inversas, por ejemplo, Superscript® II (Invitrogen), que son capaces de añadir algunos nucleótidos en el extremo 3' de una molécula de ADNc extendida, es decir, para producir un híbrido de ADN/ARN con un saliente de ADN monocatenario en el extremo 3'. El saliente de ADN puede proporcionar una secuencia diana a la que puede hibridarse una sonda oligonucleotídica para proporcionar una plantilla adicional para la extensión adicional de la molécula de ADNc. Ventajosamente, la sonda oligonucleotídica que se hibrida con el saliente del ADNc contiene una secuencia de dominio de amplificación, cuyo complemento se incorpora en el producto de ADNc de primera hebra sintetizado. Los cebadores que contienen la secuencia del dominio de amplificación, que se hibridarán con la secuencia del dominio de amplificación complementaria incorporada en la primera hebra del ADNc, pueden añadirse a la mezcla de reacción para cebar la síntesis de la segunda hebra mediante el uso de una enzima polimerasa adecuada y la primera hebra de ADNc como plantilla. Este método evita la necesidad de ligar adaptadores al extremo 3' de la primera hebra de ADNc. Aunque el cambio de plantilla se desarrolló originalmente para ARNm de longitud completa, que tienen una estructura de caperuza 5', desde entonces se ha demostrado que funciona igualmente bien con ARNm truncados sin la estructura de caperuza. Por lo tanto, el cambio de plantilla puede usarse en los métodos de la invención para generar moléculas de ADNc de longitud completa y/o parciales o truncadas. Por lo tanto, en una modalidad preferida de la invención, la síntesis de la segunda hebra puede utilizarse, o se logra, mediante cambio de plantilla. En una modalidad particularmente preferida, la reacción de cambio de plantilla, es decir, la extensión adicional de la primera hebra de ADNc para incorporar el dominio de amplificación complementario, se realiza *in situ* (mientras que la sonda de captura esté aún unida directa o indirectamente al sustrato, por ejemplo, matriz). Preferentemente, la reacción de síntesis de la segunda hebra también se realiza *in situ*.

Como se mencionó anteriormente, en algunas modalidades el ADNc inmovilizado puede etiquetarse al incorporar la etiqueta en la segunda hebra del ADNc, por ejemplo, al incorporar nucleótidos etiquetados en la segunda hebra del ADNc. Esto puede ser adicional o alternativa a, incorporar nucleótidos etiquetados en la primera hebra de ADNc.

En modalidades donde puede ser necesario o ventajoso mejorar, enriquecer o amplificar las moléculas de ADN, por ejemplo, ADNc, los dominios de amplificación pueden incorporarse en las moléculas de ADN, por ejemplo, ADNc. Como se describió anteriormente, un primer dominio de amplificación puede incorporarse en las moléculas de ácido nucleico asegurado, por ejemplo, la primera hebra de las moléculas de ADNc, cuando la sonda de captura comprende un dominio universal que comprende un dominio de amplificación. En estas modalidades, la síntesis de la segunda hebra puede incorporar un segundo dominio de amplificación. Por ejemplo, los cebadores usados para generar el ADNc de la segunda hebra, por ejemplo, cebadores hexaméricos aleatorios, cebador poli-T, el cebador que es complementario a la sonda auxiliar puede comprender en su extremo 5' un dominio de amplificación, es decir, una secuencia de nucleótidos a la que puede hibridarse un cebador de amplificación. Por lo tanto, el ADN bicatenario resultante puede comprender un dominio de amplificación en o hacia cada extremo 5' de las moléculas de ADN bicatenario, por ejemplo, ADNc. Estos dominios de amplificación pueden usarse como dianas para cebadores usados en una reacción de amplificación, por ejemplo, PCR. Alternativamente, el enlazador o adaptador que se liga al extremo 3' de las moléculas de ácido nucleico asegurado, por ejemplo, las moléculas de ADNc de la primera hebra, pueden comprender un segundo dominio universal que comprende un segundo dominio de amplificación. De manera similar, un segundo dominio de amplificación se incorpora en las moléculas de ADNc de la primera hebra mediante cambio de plantilla.

En modalidades donde la sonda de captura no comprende un dominio universal, particularmente, que comprende un dominio de amplificación, la segunda hebra de las moléculas de ADNc puede sintetizarse de acuerdo con la descripción anterior. Las moléculas de ADN bicatenario resultantes pueden modificarse para incorporar un dominio de amplificación en el extremo 5' del primer ADN, por ejemplo, ADNc, hebra (un primer dominio de amplificación) y, si no se incorpora en el ADN de la segunda hebra, por ejemplo, etapa de síntesis de ADNc, en el extremo 5' del segundo ADN, por ejemplo, ADNc, hebra (un segundo dominio de amplificación). Tales dominios de amplificación pueden incorporarse, por ejemplo, mediante la ligazón de adaptadores de doble hebra a los extremos de las moléculas de ADN, por ejemplo, ADNc. Las enzimas apropiadas para la etapa de ligazón se conocen en la técnica e incluyen, por ejemplo, Tth ADN ligasa, Taq ADN ligasa, ADN ligasa de *Thermococcus* sp. (cepa 9°N) (9°NTM ADN ligasa, New England Biolabs), AmpligaseTM (Epicentre Biotechnologies) y T4 ADN ligasa. En una modalidad preferida el primer y segundo dominios de amplificación comprenden diferentes secuencias.

A partir de lo anterior, es por lo tanto evidente que los dominios universales que pueden comprender un dominio de amplificación pueden añadirse a las moléculas de ADN aseguradas (es decir, extendidas) por ejemplo, a las moléculas de ADNc, o sus complementos (por ejemplo, segunda hebra) mediante varios métodos y técnicas y combinaciones de tales técnicas conocidas en la técnica, por ejemplo mediante el uso de cebadores que incluyen tal dominio, ligazón de adaptadores, uso de enzimas transferasas terminales y/o mediante métodos de cambio de plantilla. Como está claro a partir de la descripción en la presente descripción, tales dominios pueden añadirse antes o después de la liberación de las moléculas de ADN de la matriz.

Será evidente a partir de la descripción anterior que todas las moléculas de ADN, por ejemplo, ADNc de un único sustrato que se han sintetizado mediante los métodos de la invención pueden comprender el mismo primer y segundo dominios de amplificación. En consecuencia, una única reacción de amplificación, por ejemplo, PCR, puede ser suficiente para amplificar todas las moléculas de ADN, por ejemplo, ADNc. Por lo tanto, en una modalidad preferida, el método de la invención puede comprender una etapa de amplificación de moléculas de ADN, por ejemplo, ADNc. En una modalidad, la etapa de amplificación se realiza después de la liberación de las moléculas de ADN, por ejemplo, ADNc del sustrato. En otras modalidades, la amplificación puede realizarse en el sustrato (es decir, *in situ* en el sustrato). Se conoce en la técnica que las reacciones de amplificación pueden llevarse a cabo en sustratos, tales como matrices, y existen termocicladores en chip para llevar a cabo tales reacciones. Por lo tanto, en una modalidad las matrices que se conocen en la técnica como plataformas de secuenciación o para su uso en cualquier forma de análisis de secuencias (por ejemplo, en o mediante tecnologías de secuenciación de nueva generación) pueden usarse como base de los sustratos descritos en la presente descripción (por ejemplo, matrices de perlas de Illumina, etc.)

Para la síntesis de la segunda hebra de ADN, por ejemplo ADNc, es preferible usar una polimerasa de desplazamiento de hebra (por ejemplo, ϕ 29 ADN polimerasa, Bst (exo⁻) ADN polimerasa, klenow (exo⁻) ADN polimerasa) si el ADNc liberado del sustrato de la matriz comprende una molécula de ácido nucleico bicatenario parcialmente. Por ejemplo, los ácidos nucleicos liberados serán al menos parcialmente bicatenarios (por ejemplo, híbrido ADN:ARN) en modalidades donde la sonda de captura se inmoviliza indirectamente en el sustrato de la matriz a través de una sonda de superficie y la etapa de liberar las moléculas de ADN, por ejemplo, ADNc comprenden una etapa de escisión. La polimerasa de desplazamiento de hebra es necesaria para garantizar que la síntesis de ADNc de segunda hebra incorpore el complemento de la sonda de captura que incluye el dominio posicional (dominio de identificación de elementos), si está presente, en el segundo ADN, por ejemplo, hebra de ADNc.

Será evidente que la etapa de liberar al menos parte de las moléculas de ADN, por ejemplo, ADNc o sus amplicones de la superficie del sustrato puede lograrse mediante el uso de una serie de métodos. En algunas modalidades, será

evidente que el objetivo principal de la etapa de liberación es producir moléculas en las que se incorpora (o incluye) el dominio posicional de la sonda de captura (o su complemento), de manera que las moléculas de ADN, por ejemplo, ADNc o sus amplicones se “etiquetan” de acuerdo con su elemento (o posición) en la matriz. Sin embargo, como se describió anteriormente, un dominio posicional no es esencial, particularmente cuando solo una porción de las moléculas de ácido nucleico se libera debido a que las otras moléculas de ácido nucleico se han eliminado de la superficie del sustrato (y desechado) en una etapa anterior. La etapa de liberación elimina por lo tanto las moléculas de ADN, por ejemplo, ADNc o amplicones de estas del sustrato, cuyas moléculas de ADN, por ejemplo, ADNc o amplicones incluyen la información posicional que puede correlacionarse con la muestra de tejido. Por ejemplo, en algunas modalidades el ADN liberado comprende un dominio posicional o su complemento (en virtud de que se ha incorporado al ácido nucleico asegurado, por ejemplo, el ADNc de la primera hebra mediante, por ejemplo, la extensión de la sonda de captura, y opcionalmente copiado en el ADN de la segunda hebra si la síntesis de la segunda hebra tiene lugar en la matriz, o copiado en amplicones si la amplificación tiene lugar en la matriz). Por lo tanto, para producir datos de análisis de secuencias que puedan correlacionarse específicamente con las diversas regiones en la muestra de tejido es ventajoso que las moléculas liberadas comprendan el dominio posicional de la sonda de captura (o su complemento). Sin embargo, si las moléculas de ADN se han liberado de una porción específica del sustrato, será evidente que el análisis de secuencias puede correlacionarse con la región(es) en la muestra de tejido que no se eliminaron del sustrato.

Dado que la molécula liberada puede ser una molécula de ADN de primera y/o segunda hebra, por ejemplo, ADNc o amplicón, y dado que la sonda de captura puede inmovilizarse indirectamente en el sustrato, se entenderá que aunque la etapa de liberación puede comprender una etapa de escindir una molécula de ADN, por ejemplo, ADNc de la matriz, la etapa de liberación no requiere una etapa de escisión del ácido nucleico; una molécula de ADN, por ejemplo, ADNc o un amplicón puede simplemente liberarse al desnaturalizar una molécula bicatenaria, por ejemplo, liberar la segunda hebra de ADNc de la primera hebra de ADNc, o liberar un amplicón de su plantilla o liberar la molécula de ADNc de la primera hebra (es decir, la sonda de captura extendida) de una sonda de superficie. En consecuencia, una molécula de ADN, por ejemplo, ADNc, puede liberarse del sustrato mediante escisión del ácido nucleico y/o por desnaturalización (por ejemplo, mediante calentamiento para desnaturalizar una molécula bicatenaria). Cuando la amplificación se lleva a cabo *in situ* en el sustrato, esto abarcará, por supuesto, la liberación de amplicones por desnaturalización en la reacción de ciclo.

En algunas modalidades, las moléculas de ADN, por ejemplo, ADNc, se liberan mediante escisión enzimática de un dominio de escisión, que puede ubicarse en el dominio universal o dominio posicional de la sonda de captura. Como se mencionó anteriormente, en algunas modalidades el dominio de escisión debe ubicarse aguas arriba (en el extremo 5') del dominio posicional de manera que las moléculas de ADN liberadas, por ejemplo, ADNc comprenden el dominio posicional (identificación). Las enzimas adecuadas para la escisión de ácidos nucleicos incluyen endonucleasas de restricción, por ejemplo, RsaI. Otras enzimas, por ejemplo, una mezcla de uracil ADN glucosilasa (UDG) y la ADN glucosilasa-liasa Endonucleasa VIII (enzima USER™) o una combinación de las enzimas MutY y T7 endonucleasa I, son modalidades preferidas de los métodos de la invención.

En una modalidad alternativa, las moléculas de ADN, por ejemplo, ADNc pueden liberarse de la superficie del sustrato por medios físicos. Por ejemplo, en modalidades donde la sonda de captura se inmoviliza indirectamente en el sustrato, por ejemplo, mediante hibridación con la sonda de superficie, puede ser suficiente interrumpir la interacción entre las moléculas de ácido nucleico. Los métodos para interrumpir la interacción entre moléculas de ácido nucleico, por ejemplo, moléculas de ácido nucleico bicatenario desnaturalizado se conocen bien en la técnica. Un método directo para liberar las moléculas de ADN, por ejemplo, ADNc, (es decir, de desprender el sustrato, por ejemplo, matriz, de las moléculas de ADN sintetizado, por ejemplo, ADNc) es usar una solución que interfiera con los enlaces de hidrógeno de las moléculas bicatenarias. En una modalidad preferida de la invención, las moléculas de ADNc pueden liberarse al aplicar agua calentada, por ejemplo, agua o tampón de al menos 85 °C, preferentemente al menos 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 °C. Como alternativa o adicional al uso de una temperatura suficiente para interrumpir la unión de hidrógeno, la solución puede comprender sales, surfactantes, etc. que pueden desestabilizar aún más la interacción entre las moléculas de ácido nucleico, lo que da como resultado la liberación de las moléculas de ADN, por ejemplo, ADNc.

Se debe entender que la aplicación de una solución de alta temperatura, por ejemplo, agua a 90-99 °C puede ser suficiente para interrumpir un enlace covalente usado para inmovilizar la sonda de captura o sonda de superficie al sustrato. Por lo tanto, en una modalidad preferida, las moléculas de ADN, por ejemplo, ADNc pueden liberarse mediante la aplicación de agua caliente al sustrato para interrumpir la captura inmovilizada covalentemente o sondas de superficie.

Es implícito que las moléculas de ADN liberadas, por ejemplo, ADNc (la solución que comprende las moléculas de ADN liberadas, por ejemplo, ADNc) se recojan para su manipulación adicional, por ejemplo, síntesis y/o amplificación de la segunda hebra. Sin embargo, puede observarse que el método de la invención comprende una etapa de recoger o recuperar las moléculas de ADNc liberadas. Como se indicó anteriormente, en el contexto de la amplificación *in situ*, las moléculas liberadas pueden incluir amplicones del ácido nucleico asegurado, por ejemplo, ADNc.

En modalidades de métodos de la invención, puede ser conveniente retirar cualquier sonda de captura no extendida. Esto puede ser, por ejemplo, después de la etapa de liberar moléculas de ADN del sustrato. Cualquier método deseado o conveniente puede usarse para tal eliminación que incluye, por ejemplo, el uso de una enzima para degradar las sondas no extendidas, por ejemplo, exonucleasa.

Las moléculas de ADN, por ejemplo, ADNc, o amplicones, que se han liberado del sustrato, que pueden haberse modificado como se describió anteriormente, se analizan para investigar (por ejemplo, determinar su secuencia, aunque como se indicó anteriormente no se requiere la determinación de la secuencia real - puede usarse cualquier método para analizar la secuencia). Por lo tanto, puede usarse cualquier método de análisis de ácidos nucleicos. La etapa de análisis de secuencias puede identificar el dominio posicional y, por lo tanto, permitir que la molécula analizada se localice en una posición en la muestra de tejido. De manera similar, puede determinarse la naturaleza o identidad de la molécula analizada. De esta manera puede determinarse el ácido nucleico, por ejemplo, ARN, en una posición en el sustrato y, por lo tanto, en la muestra de tejido. Por lo tanto, la etapa de análisis puede incluir o usar cualquier método que identifique la molécula analizada (y por lo tanto la molécula "diana") y su dominio posicional. Generalmente tal método será un método específico de secuencia. Por ejemplo, el método puede usar cebadores o sondas específicos de secuencias, particularmente cebadores o sondas específicos para el dominio posicional y/o para detectar o analizar una molécula de ácido nucleico específica, por ejemplo, una molécula de ADN correspondiente a un ácido nucleico, por ejemplo, molécula de ARN o ADNc a detectar. Típicamente, en un método de este tipo pueden usarse cebadores de amplificación específicos de secuencia, por ejemplo, cebadores de PCR.

En algunas modalidades puede ser conveniente analizar un subconjunto o familia de moléculas relacionadas diana, por ejemplo, todas las secuencias que codifican un grupo particular de proteínas que comparten similitud de secuencia y/o dominios conservados, por ejemplo, una familia de receptores. Por lo tanto, los métodos de amplificación y/o análisis descritos en la presente descripción pueden usar cebadores o sondas redundantes o específicos de familias de genes que se hibridan con un subconjunto de ácidos nucleicos capturados o ácidos nucleicos derivados de estos, por ejemplo, amplicones. En una modalidad particularmente preferida, los métodos de amplificación y/o análisis pueden utilizar un cebador universal (es decir, un cebador común a todas las secuencias capturadas) en combinación con un cebador redundante o específico de familias de genes específico para un subconjunto de moléculas diana.

Por lo tanto, en una modalidad, se usan métodos de análisis de secuencias basados en amplificación, especialmente basados en PCR.

Sin embargo, las etapas de modificación y/o amplificación de las moléculas de ADN liberadas, por ejemplo, ADNc pueden introducir componentes adicionales en la muestra, por ejemplo, enzimas, cebadores, nucleótidos, etc. Por lo tanto, los métodos de la invención pueden comprender además una etapa de purificación de la muestra que comprende las moléculas de ADN liberadas, por ejemplo, ADNc o amplicones antes del análisis de secuencias, por ejemplo, para eliminar cebadores de oligonucleótidos, nucleótidos, sales, etc. que pueden interferir con las reacciones de secuenciación. Puede usarse cualquier método adecuado para purificar las moléculas de ADN, por ejemplo, ADNc.

Como se indicó anteriormente el análisis de secuencias de las moléculas de ADN liberadas puede ser directo o indirecto. Por lo tanto, el material de análisis de secuencias o sustrato (que puede verse como las moléculas que se somete a la etapa o proceso de análisis de secuencias) pueden ser directamente las moléculas que se liberan del sustrato objeto, por ejemplo, matriz o puede ser una molécula que se deriva de la misma. Por lo tanto, por ejemplo, en el contexto de la etapa de análisis de secuencias que implica una reacción de secuenciación, la plantilla de secuenciación puede ser la molécula que se libera del sustrato objeto, por ejemplo, matriz o puede ser una molécula derivada de esta. Por ejemplo, una molécula de ADN de primera y/o segunda hebra, por ejemplo, ADNc, liberada del sustrato, por ejemplo, matriz puede someterse directamente al análisis de secuencias (por ejemplo, secuenciación), es decir, puede participar directamente en la reacción o proceso de análisis de secuencias (por ejemplo, la reacción de secuenciación o proceso de secuenciación, o ser la molécula que se secuencia o identifica de cualquier otra manera). En el contexto de la amplificación *in situ*, la molécula liberada puede ser un amplicón. Alternativamente, la molécula liberada puede someterse a una etapa de síntesis o amplificación de la segunda hebra antes del análisis de secuencias (por ejemplo, secuenciación o identificación por otros medios). El sustrato de análisis de secuencias (por ejemplo, plantilla) puede ser por lo tanto un amplicón o una segunda hebra de una molécula que se libera directamente del sustrato objeto, por ejemplo, matriz.

Ambas hebras de una molécula de doble hebra pueden someterse a análisis de secuencias (por ejemplo, secuenciarse) pero la invención no se limita a esto y pueden analizarse moléculas de hebra única (por ejemplo, ADNc) (por ejemplo, secuenciarse). Por ejemplo, pueden usarse diversas tecnologías de secuenciación para la secuenciación de moléculas individuales, por ejemplo, las tecnologías Helicos o Pacbio, o tecnologías de secuenciación de nanoporos que están en desarrollo. Por lo tanto, en una modalidad la primera hebra de ADN, por ejemplo, ADNc, puede someterse a secuenciación. Puede ser necesario modificar el ADN de la primera hebra, por ejemplo, ADNc en el extremo 3' para permitir la secuenciación de moléculas individuales. Esto puede hacerse mediante procedimientos análogos a aquellos para manipular la segunda hebra de ADN, por ejemplo, ADNc. Tales procedimientos se conocen en la técnica.

En un aspecto preferido de la invención, el análisis de secuencias identificará o revelará una porción de secuencia de ácido nucleico capturado, por ejemplo, ARN y opcionalmente, la secuencia del dominio posicional. En algunas modalidades, la secuencia del dominio posicional (o etiqueta) identificará el elemento en el que se capturó la molécula de ácido nucleico, por ejemplo, ARNm. La secuencia de la molécula de ácido nucleico capturada, por ejemplo, ARN, puede compararse con una base de datos de secuencias del organismo a partir de la cual se originó la muestra, para determinar el gen al cual corresponde. Al determinar qué región (por ejemplo, célula) de la muestra de tejido estaba en contacto con la posición o área, por ejemplo, elemento, del sustrato del cual se liberó el ácido nucleico capturado, es posible determinar qué región de la muestra de tejido estaba expresado dicho gen. Este análisis puede lograrse para todas las moléculas de ADN, por ejemplo, ADNc, generadas por los métodos de la invención, lo que produce un transcriptoma espacial de la muestra de tejido. Sin embargo, en algunas modalidades solo puede capturarse una selección de los transcritos presentes en la muestra de tejido en el sustrato (por ejemplo, si el dominio de captura de la sonda de captura comprende una secuencia específica para un gen o conjunto de genes) y/o solo puede seleccionarse una porción de las moléculas capturadas para un análisis adicional, por ejemplo, análisis de secuencias (por ejemplo, una porción de las moléculas capturadas puede eliminarse del sustrato antes de la etapa de análisis de secuencias).

Por medio de un ejemplo representativo, los datos de secuenciación pueden analizarse para clasificar las secuencias en especies específicas de sonda de captura, por ejemplo, de acuerdo con la secuencia del dominio posicional. Esto puede lograrse, por ejemplo, mediante el uso del kit de herramientas FastX de la herramienta divisora de código de barras FASTQ para clasificar las secuencias en archivos individuales para las respectivas secuencias de dominio posicional (etiqueta) de la sonda de captura. Las secuencias de cada especie, es decir, de cada elemento pueden analizarse para determinar la identidad de los transcritos. Por ejemplo, las secuencias pueden identificarse mediante el uso de, por ejemplo, el programa informático Blastn, para comparar las secuencias con una o más bases de datos de genoma, preferentemente la base de datos para el organismo del cual se obtuvo la muestra de tejido. La identidad de la secuencia de la base de datos con la mayor similitud con la secuencia generada por los métodos de la invención se asignará a dicha secuencia. En general, solo los aciertos con una certeza de al menos $1e^{-6}$, preferentemente $1e^{-7}$, $1e^{-8}$ o $1e^{-9}$ se considerarán identificados con éxito.

Será evidente que cualquier método de secuenciación de ácido nucleico puede utilizarse en los métodos de la invención. Sin embargo, las denominadas técnicas de "secuenciación de nueva generación" tendrán particular utilidad en la presente invención. La secuenciación de alto rendimiento es particularmente útil en los métodos de la invención porque permite secuenciar parcialmente un gran número de ácidos nucleicos en un período de tiempo muy corto. En vista de la explosión reciente en el número de genomas secuenciados total o parcialmente, no es esencial secuenciar la longitud completa de las moléculas de ADN, por ejemplo, ADNc generadas para determinar el gen al que corresponde cada molécula. Por ejemplo, los primeros 100 nucleótidos de cada extremo de las moléculas de ADN, por ejemplo, ADNc, deben ser suficientes para identificar el gen expresado y, en modalidades las que las sondas de captura se disponen en el sustrato, el elemento en el que se capturó la molécula de ácido nucleico, por ejemplo, ARNm, (es decir, su ubicación sobre la matriz). En algunas modalidades, la reacción de secuencia del "extremo de la sonda de captura" de las moléculas de ADN, por ejemplo, ADNc, produce la secuencia del dominio posicional y al menos aproximadamente 20 bases, preferentemente 30 o 40 bases de datos de secuencias específicas de transcritos. Sin embargo, en modalidades en las que la sonda de captura no contiene un dominio posicional, la reacción de secuencia del "extremo de la sonda de captura" del ADN, puede producir al menos aproximadamente 70 bases, preferentemente 80, 90 o 100 bases de datos de secuencias específicas de transcritos. La reacción de secuencia del "extremo que no es de sonda de captura" puede producir al menos aproximadamente 70 bases, preferentemente 80, 90 o 100 bases de datos de secuencias específicas de transcritos.

Como ejemplo representativo, la reacción de secuenciación puede basarse en terminadores-colorantes reversibles, tal como se usa en la tecnología Illumina™. Por ejemplo, las moléculas de ADN se unen primero a los cebadores en, por ejemplo, un portaobjetos de vidrio o silicio y se amplifican de manera que se formen colonias clonales locales (amplificación puente). Se añaden cuatro tipos de ddNTP y los nucleótidos no incorporados se eliminan por lavado. A diferencia de la pirosecuenciación, el ADN solo puede extenderse un nucleótido a la vez. Una cámara toma imágenes de los nucleótidos etiquetados con fluorescencia, después el colorante junto con el bloqueador terminal 3' se elimina químicamente del ADN, lo que permite un siguiente ciclo. Esto puede repetirse hasta que se obtengan los datos de secuencia requeridos. Mediante el uso de esta tecnología, miles de ácidos nucleicos pueden secuenciarse simultáneamente en un solo portaobjetos.

Otras técnicas de secuenciación de alto rendimiento pueden ser igualmente adecuadas para los métodos de la invención, por ejemplo, la pirosecuenciación. En este método, el ADN se amplifica dentro de gotitas de agua en una solución oleosa (PCR en emulsión), con cada gotita que contiene una única plantilla de ADN unida a una única perla recubierta con cebador que después forma una colonia clonal. La máquina de secuenciación contiene muchos pocillos de volumen de picolitros, cada uno de los cuales contiene una única perla y enzimas de secuenciación. La pirosecuenciación usa luciferasa para generar luz para la detección de los nucleótidos individuales añadidos al ADN naciente y los datos combinados se usan para generar lecturas de secuencias.

Un ejemplo de una tecnología en desarrollo se basa en la detección de iones de hidrógeno que se liberan durante la polimerización del ADN. Un micropocillo que contiene una hebra de ADN plantilla a secuenciar se inunda con un único

tipo de nucleótido. Si el nucleótido introducido es complementario al nucleótido plantilla delantero, este se incorpora en la hebra complementaria en crecimiento. Esto provoca la liberación de un ion de hidrógeno que activa un sensor de iones hipersensible, lo que indica que ha ocurrido una reacción. Si hay repeticiones de homopolímeros presentes en la secuencia plantilla, se incorporarán múltiples nucleótidos en un único ciclo. Esto conduce a un número correspondiente de iones de hidrógeno que se liberan y una señal electrónica proporcionalmente mayor.

Por lo tanto, está claro que pueden hacerse paulatinamente disponibles futuros formatos de secuenciación, y con tiempos de corrida más cortos como una de los principales elementos de esas plataformas, será evidente que otras tecnologías de secuenciación serán útiles en los métodos de la invención, por ejemplo, métodos de secuenciación de nanoporos.

Un elemento esencial de la presente invención, como se describió anteriormente, es una etapa de asegurar una hebra complementaria de las moléculas de ácidos nucleicos capturadas a la sonda de captura, por ejemplo, transcribir inversamente las moléculas de ARN capturadas. La reacción de transcripción inversa se conoce bien en la técnica y en las reacciones de transcripción inversa representativas, la mezcla de reacción incluye una transcriptasa inversa, dNTP y un tampón adecuado. La mezcla de reacción puede comprender otros componentes, por ejemplo, inhibidor(es) de ARNasa. Los cebadores y la plantilla son el dominio de captura de la sonda de captura y las moléculas de ARN capturadas, como se describió anteriormente. En los métodos de la invención, cada dNTP estará típicamente presente en una cantidad que varía de aproximadamente 10 a 5000 μM , normalmente de aproximadamente 20 a 1000 μM . Será evidente que puede realizarse una reacción equivalente para generar una hebra complementaria de una molécula de ADN capturada, mediante el uso de una enzima con actividad de ADN polimerasa. Las reacciones de este tipo se conocen bien en la técnica y se describen con más detalle más abajo.

En algunas modalidades, un dNTP etiquetado puede estar presente en la mezcla de reacción, al incorporar de esta manera una etiqueta en la molécula de ADN sintetizada. En una modalidad representativa, el dNTP etiquetado es un dNTP etiquetado con fluorescencia, por ejemplo, Cy3-dCTP.

La actividad de transcriptasa inversa deseada puede proporcionarse por una o más enzimas distintas, en donde los ejemplos adecuados son: las enzimas M-MLV, MuLV, AMV, HIV, ArrayScript™, MultiScribe™, ThermoScript™ y SuperScript® I, II y III.

La reacción de la transcriptasa inversa puede llevarse a cabo a cualquier temperatura adecuada, que dependerá de las propiedades de la enzima. Típicamente, las reacciones de transcriptasa inversa se realizan entre 37-55 °C, aunque también pueden ser apropiadas temperaturas fuera de este intervalo. El tiempo de reacción puede ser tan pequeño como 1, 2, 3, 4 o 5 minutos o tanto como 48 horas. Típicamente, la reacción se llevará a cabo durante entre 3-12 horas, tal como 5-120 minutos, 5-60, 5-45 o 530 minutos o 1-10 o 1-5 minutos de acuerdo con la elección. El tiempo de reacción no es crítico y puede usarse cualquier tiempo de reacción deseado. Por ejemplo, las incubaciones durante la noche son comunes.

Como se indicó anteriormente, ciertas modalidades de los métodos incluyen una etapa de amplificación, donde se aumenta el número de copias de moléculas de ADN, por ejemplo, ADNc generadas, por ejemplo, para enriquecer la muestra para obtener una mejor representación de los ácidos nucleicos, por ejemplo, transcritos, capturados de la muestra de tejido. La amplificación puede ser lineal o exponencial, según se desee, donde los protocolos de amplificación representativos de interés incluyen, pero no se limitan a: reacción en cadena de la polimerasa (PCR); amplificación isotérmica, etc.

La reacción en cadena de la polimerasa (PCR) se conoce bien en la técnica, que se describe en las patentes de Estados Unidos núms.: 4,683,202; 4,683,195; 4,800,159; 4,965,188 y 5,512,462. En reacciones de amplificación de PCR representativas, la mezcla de reacción que incluye las moléculas de ADN liberadas, por ejemplo, ADNc anteriores del sustrato, por ejemplo, matriz, que se combinan con uno o más cebadores que se emplean en la reacción de extensión del cebador, por ejemplo, los cebadores de PCR que se hibridan con el primer y/o segundo dominios de amplificación (tales como cebadores directos e inversos empleados en amplificación geométrica (o exponencial) o un único cebador empleado en una amplificación lineal). Los cebadores de oligonucleótidos con los que se ponen en contacto las moléculas de ADN liberadas, por ejemplo, ADNc (en lo sucesivo denominadas ADN plantilla por conveniencia) serán de longitud suficiente para proporcionar la hibridación al ADN plantilla complementario en condiciones de hibridación (descritas en mayor detalle más abajo). La longitud de los cebadores dependerá de la longitud de los dominios de amplificación, pero generalmente será de al menos 10 pb de longitud, usualmente al menos 15 pb de longitud y más usualmente al menos 16 pb de longitud y puede ser de hasta 30 pb de longitud o más, donde la longitud de los cebadores generalmente variará de 18 a 50 pb de longitud, usualmente de aproximadamente 20 a 35 pb de longitud. El ADN plantilla puede ponerse en contacto con un único cebador o un conjunto de dos cebadores (cebadores directos e inversos), en dependencia de si se desea la extensión del cebador, la amplificación lineal o exponencial del ADN plantilla.

Además de los componentes anteriores, la mezcla de reacción producida en los métodos de la invención incluye típicamente una polimerasa y desoxirribonucleósidos trifosfatados (dNTP). La actividad de la polimerasa deseada puede proporcionarse por una o más enzimas distintas de la polimerasa. En muchas modalidades, la mezcla de

reacción incluye al menos una polimerasa de la familia A, donde las polimerasas de la familia A representativas de interés incluyen, pero no se limitan a: Las polimerasas de *Thermus aquaticus*, que incluyen la polimerasa de origen natural (Taq) y derivados y homólogos de la misma, tales como KlenTaq (como se describe en Barnes y otros, Proc. Natl. Acad. Sci. USA (1994) 91:2216-2220); polimerasas de *Thermus thermophilus*, que incluyen la polimerasa de origen natural (Tth) y derivados y homólogos de la misma, y similares. En ciertas modalidades donde la reacción de amplificación que se lleva a cabo es una reacción de alta fidelidad, la mezcla de reacción puede incluir además una enzima polimerasa que tiene actividad exonucleasa 3'-5', por ejemplo, como puede proporcionarse por una polimerasa de la Familia B, donde las polimerasas de la Familia B de interés incluyen, pero no se limitan a: ADN polimerasa de *Thermococcus litoralis* (Vent) como se describe en Perler y otros, Proc. Natl. Acad. Sci. USA (1992) 89:5577-5581; GB-D de la especie *Pyrococcus* (Deep Vent); ADN polimerasa de *Pyrococcus furiosus* (Pfu) como se describe en Lundberg y otros, Gene (1991) 108:1-6, *Pyrococcus woesei* (Pwo) y similares. Cuando la mezcla de reacción incluye tanto una polimerasa de la Familia A como de la Familia B, la polimerasa de la Familia A puede estar presente en la mezcla de reacción en una cantidad mayor que la polimerasa de la Familia B, donde la diferencia de actividad será usualmente al menos 10 veces, y más usualmente al menos aproximadamente 100 veces. Usualmente, la mezcla de reacción incluirá cuatro tipos diferentes de dNTP correspondientes a las cuatro bases de origen natural presentes, es decir, dATP, dTTP, dCTP y dGTP. En los métodos sujetos, cada dNTP estará típicamente presente en una cantidad que varía de aproximadamente 10 a 5000 μM , usualmente de aproximadamente 20 a 1000 μM .

Las mezclas de reacción preparadas en las etapas de transcriptasa inversa y/o amplificación de los métodos de la invención pueden incluir además un medio tampón acuoso que incluye una fuente de iones monovalentes, una fuente de cationes divalentes y un agente tampón. Puede emplearse cualquier fuente conveniente de iones monovalentes, tales como KCl, acetato de K, acetato de NH_4 , glutamato de K, NH_4Cl , sulfato de amonio y similares. El catión divalente puede ser magnesio, manganeso, zinc y similares, donde el catión será típicamente magnesio. Puede emplearse cualquier fuente conveniente de catión de magnesio, que incluye MgCl_2 , acetato de Mg y similares. La cantidad de Mg^{2+} presente en el tampón puede variar de 0,5 a 10 mM, pero variará preferentemente de aproximadamente 3 a 6 mM, e idealmente estará a aproximadamente 5 mM. Los agentes tampones representativos o sales que pueden estar presentes en el tampón incluyen Tris, Tricina, HEPES, MOPS y similares, donde la cantidad de agente tampón variará típicamente de aproximadamente 5 a 150 mM, usualmente de aproximadamente 10 a 100 mM, y más usualmente de aproximadamente 20 a 50 mM, donde en ciertas modalidades preferidas el agente tampón estará presente en una cantidad suficiente para proporcionar un pH que varía de aproximadamente 6,0 a 9,5, donde lo más preferido es el pH 7,3 a 7,2 °C. Otros agentes que pueden estar presentes en el medio tampón incluyen agentes quelantes, tales como EDTA, EGTA y similares.

En la preparación de la mezcla de reacción de transcriptasa inversa, extensión de ADN o amplificación de las etapas de los métodos de la invención, los diversos componentes constituyentes pueden combinarse en cualquier orden conveniente. Por ejemplo, en la reacción de amplificación el tampón puede combinarse con cebador, polimerasa y después ADN plantilla, o todos los diversos componentes constituyentes pueden combinarse al mismo tiempo para producir la mezcla de reacción.

Como se describió anteriormente, una modalidad preferida de la invención las moléculas de ADN, por ejemplo, ADNc pueden modificarse mediante la adición de dominios de amplificación a los extremos de las moléculas de ácido nucleico, que pueden implicar una reacción de ligazón. También se requiere una reacción de ligazón para la síntesis *in situ* de la sonda de captura en el sustrato, cuando la sonda de captura se inmoviliza indirectamente sobre la superficie del sustrato.

Como se conoce en la técnica, las ligasas catalizan la formación de un enlace fosfodiéster entre los terminales 3'-hidroxilo y 5'-fosfato yuxtapuestos de dos ácidos nucleicos inmediatamente adyacentes. Puede emplearse cualquier ligasa conveniente, donde las ligasas representativas de interés incluyen, pero no se limitan a: Ligasas termoestables y sensibles a la temperatura. Las ligasas sensibles a la temperatura incluyen, pero no se limitan a, ADN ligasa del bacteriófago T4, ligasa del bacteriófago T7 y ligasa de *E. coli*. Las ligasas termoestables incluyen, pero no se limitan a, ligasa de Taq, ligasa de Tth y ligasa de Pfu. La ligasa termoestable puede obtenerse de organismos termófilos o hipertermófilos, que incluyen, pero no se limitan a, organismos procariotas, eucariotas o arqueas. Ciertas ligasas de ARN también pueden emplearse en los métodos de la invención.

En esta etapa de ligazón, una ligasa adecuada y cualquier reactivo que sea necesario y/o conveniente se combinan con la mezcla de reacción y se mantienen en condiciones suficientes para que ocurra la ligazón de los oligonucleótidos relevantes. Las condiciones de reacción de ligazón se conocen bien por los expertos en la técnica. Durante la ligazón, la mezcla de reacción en ciertas modalidades puede mantenerse a una temperatura que varía de aproximadamente 4 °C a aproximadamente 50 °C, tal como de aproximadamente 20 °C a aproximadamente 37 °C durante un período de tiempo que varía de aproximadamente 5 segundos a aproximadamente 16 horas, tal como de aproximadamente 1 minuto a aproximadamente 1 hora. Aún en otras modalidades, la mezcla de reacción puede mantenerse a una temperatura que varía de aproximadamente 35 °C a aproximadamente 45 °C, tal como de aproximadamente 37 °C a aproximadamente 42 °C, por ejemplo, a o aproximadamente 38 °C, 39 °C, 40 °C o 41 °C, durante un período de tiempo que varía de aproximadamente 5 segundos a aproximadamente 16 horas, tal como de aproximadamente 1 minuto a aproximadamente 1 hora, que incluye de aproximadamente 2 minutos a aproximadamente 8 horas. En una modalidad representativa, la mezcla de reacción de ligazón incluye Tris 50 mM pH7,5, MgCl_2 10 mM, DTT 10 mM, ATP 1 mM,

BSA 25 mg/ml, inhibidor de ARNasa 0,25 unidades/ml y ADN ligasa de T4 a 0,125 unidades/ml. Aún en otra modalidad representativa, se emplean iones de magnesio 2,125 mM, inhibidor de ARNasa 0,2 unidades/ml; y ADN ligasa 0,125 unidades/ml. La cantidad de adaptador en la reacción dependerá de la concentración del ADN, por ejemplo, ADNc en la muestra y generalmente estará presente a entre 10-100 veces la cantidad molar de ADN, por ejemplo, ADNc.

5 Por medio de un ejemplo representativo, el método descrito en la presente descripción puede comprender las siguientes etapas:

- (a) poner en contacto un sustrato objeto con una muestra de tejido, en donde al menos una especie de sonda de captura se inmoviliza directa o indirectamente sobre dicho sustrato de manera que las sondas se orientan para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT), de manera que el ARN de la muestra de tejido se hibride con dichas sondas de captura;
- 10 (b) obtener imágenes de la muestra de tejido en el sustrato;
- (c) transcribir inversamente las moléculas de ARNm capturadas para generar moléculas de ADNc, en donde los nucleótidos etiquetados se incorporan en la parte sintetizada de las moléculas de ADNc;
- (d) lavar el sustrato para eliminar el tejido residual;
- 15 (e) obtener imágenes del sustrato de manera que se detecte la señal de las moléculas de ADNc etiquetadas;
- (f) retirar el ADNc etiquetado de al menos una porción de la superficie del sustrato;
- (g) liberar al menos parte de las moléculas de ADNc etiquetadas restantes de la superficie de la matriz;
- (h) realizar la síntesis de ADNc de la segunda hebra en las moléculas de ADNc liberadas;
- y
- 20 (i) analizar la secuencia de (por ejemplo, secuenciación) las moléculas de ADNc.

Por medio de un ejemplo representativo alternativo, el método descrito en la presente descripción puede comprender las siguientes etapas:

- (a) poner en contacto un sustrato objeto con una muestra de tejido, en donde al menos una especie de sonda de captura se inmoviliza directa o indirectamente sobre dicho sustrato de manera que las sondas se orientan para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT), de manera que el ARN de la muestra de tejido se hibride con dichas sondas de captura;
- 25 (b) opcionalmente rehidratar la muestra de tejido;
- (c) transcribir inversamente las moléculas de ARNm capturadas para generar moléculas de ADNc;
- 30 (d) obtener imágenes de la muestra de tejido en el sustrato;
- (e) lavar el sustrato para eliminar el tejido residual;
- (f) etiquetar las moléculas de ADNc con una tinción de ácido nucleico;
- (g) obtener imágenes del sustrato de manera que se detecte la señal de las moléculas de ADNc etiquetadas;
- (h) retirar el ADNc etiquetado de al menos una porción de la superficie del sustrato;
- 35 (i) liberar al menos parte de las moléculas de ADNc etiquetadas restantes de la superficie de la matriz;
- (j) amplificar las moléculas de ADNc liberadas;
- (k) purificar opcionalmente las moléculas de ADNc para eliminar componentes que pueden interferir con la reacción de secuenciación;
- y
- 40 (l) analizar la secuencia de (por ejemplo, secuenciación) las moléculas de ADNc.

Por medio de aún un ejemplo representativo adicional, el método descrito en la presente descripción puede comprender las siguientes etapas:

- (a) poner en contacto un sustrato objeto con una muestra de tejido, en donde al menos una especie de sonda de captura se inmoviliza directa o indirectamente sobre dicho sustrato de manera que las sondas se orientan para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT), de manera que el ARN de la muestra de tejido se hibride con dichas sondas de captura;
- 45 (b) obtener imágenes de la muestra de tejido en el sustrato;
- (c) transcribir inversamente las moléculas de ARNm capturadas para generar moléculas de ADNc, en donde los nucleótidos etiquetados se incorporan en la parte sintetizada de las moléculas de ADNc;
- 50 (d) lavar el sustrato para eliminar el tejido residual;
- (e) obtener imágenes del sustrato de manera que se detecte la señal de las moléculas de ADNc etiquetadas;
- (f) repetir las etapas (a)-(e), con un segundo sustrato objeto, mediante el uso de diferentes condiciones en la etapa (a);
- 55 (g) comparar la intensidad y/o resolución de la señal de las moléculas de ADNc etiquetadas inmovilizadas en dicho primer y segundo sustrato objeto; y
- (h) seleccionar las condiciones que proporcionan la intensidad y/o resolución óptima de la señal de las moléculas de ADNc etiquetadas.

60 Por medio de un ejemplo representativo alternativo adicional, el método descrito en la presente descripción puede comprender las siguientes etapas:

- (a) poner en contacto un sustrato objeto con una muestra de tejido, en donde múltiples especies de sondas de captura se inmovilizan directa o indirectamente de manera que cada especie ocupa una posición distinta en dicho sustrato y se orienta para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT), en donde cada especie de dicha sonda de captura comprende una molécula de ácido nucleico con 5' a 3':
- 65

- (i) un dominio posicional que corresponde a la posición de la sonda de captura en el sustrato, y
- (ii) un dominio de captura;
- de manera que el ARN de la muestra de tejido se hibride con dichas sondas de captura;
- (b) opcionalmente rehidratar la muestra de tejido;
- (c) obtener imágenes de la muestra de tejido en el sustrato;
- (d) transcribir inversamente las moléculas de ARNm capturadas para generar moléculas de ADNc, en donde los nucleótidos etiquetados se incorporan en la parte sintetizada de las moléculas de ADNc;
- (e) lavar el sustrato para eliminar el tejido residual;
- (f) obtener imágenes del sustrato de manera que se detecte la señal de las moléculas de ADNc etiquetadas;
- (g) retirar el ADNc etiquetado de al menos una porción de la superficie del sustrato;
- (h) liberar al menos parte de las moléculas de ADNc etiquetadas restantes de la superficie de la matriz;
- (i) amplificar las moléculas de ADNc liberadas;
- (j) purificar opcionalmente las moléculas de ADNc para eliminar componentes que pueden interferir con la reacción de secuenciación;
- y
- (k) analizar la secuencia de (por ejemplo, secuenciación) las moléculas de ADNc.

En la presente descripción se describe cualquier combinación adecuada de las etapas en los métodos descritos anteriormente.

En la presente descripción se describen métodos, por ejemplo, donde la amplificación se realiza *in situ* en el sustrato, por ejemplo, matriz. También se abarcan métodos que omiten la etapa de obtener imágenes de la muestra de tejido.

En la presente descripción se describe un método para fabricar o producir un sustrato objeto (i) para su uso en la captura de ARNm a partir de una muestra de tejido que se pone en contacto con dicho sustrato; (ii) para su uso en la determinación y/o análisis de un transcriptoma (por ejemplo, el parcial o global) de una muestra de tejido; o (iii) para su uso en la determinación de las condiciones óptimas para la detección localizada o espacial de ácido nucleico a partir de una muestra de tejido en contacto con un sustrato, dicho método comprende inmovilizar, directa o indirectamente, al menos una especie de sonda de captura a un sustrato de manera que cada sonda se orienta para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT).

Opcionalmente, las sondas se inmovilizan uniformemente en el sustrato objeto, es decir, las sondas no se disponen como elementos distintos. En una modalidad particular de la invención, las sondas son idénticas.

En algunas modalidades de la invención, las sondas son capaces de hibridar (es decir, capturar) todo el ARNm, es decir, moléculas de ARN con una cola poliA. Por lo tanto, en modalidades particularmente preferidas de la invención, las sondas pueden comprender regiones de nucleótidos dTTP o dUTP consecutivos, por ejemplo, nucleótidos oligoT y/u oligoU, como se describió en más detalle anteriormente.

El método para inmovilizar las sondas de captura en el sustrato objeto puede lograrse mediante el uso de cualquier medio adecuado como se describe en la presente descripción. Cuando las sondas de captura se inmovilizan en la matriz indirectamente, la sonda de captura puede sintetizarse en el sustrato objeto. Dicho método puede comprender una cualquiera o más de las siguientes etapas:

- (a) inmovilizar directa o indirectamente múltiples sondas de superficie a un sustrato, en donde las sondas de superficie comprenden:
 - (i) un dominio capaz de hibridarse con parte del oligonucleótido de dominio de captura (una parte no implicada en capturar el ácido nucleico, por ejemplo, ARN); y
 - (ii) un dominio universal complementario;
- (b) hibridar con las sondas de superficie inmovilizadas en los oligonucleótidos del dominio de captura del sustrato y los oligonucleótidos del dominio universal;
- (c) unir el dominio universal al oligonucleótido de dominio de captura para producir el oligonucleótido de captura.

Por lo tanto, en una modalidad particular el método puede verse como un método para fabricar o producir un sustrato objeto que comprende un sustrato en el que una o más especies de sonda de captura, que comprende un dominio de captura, se inmoviliza directa o indirectamente de manera que cada sonda se orienta para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT), en donde las sondas se inmovilizan en el sustrato objeto con una distribución homogénea y dicho sustrato es para su uso en:

- (i) capturar ARN a partir de una muestra de tejido que se pone en contacto con dicho sustrato objeto; o
 - (ii) la detección localizada o espacial del ARN en una muestra de tejido,
- dicho método que comprende:
- (a) inmovilizar directa o indirectamente múltiples sondas de superficie a un sustrato, en donde las sondas de superficie comprenden:
 - (i) un dominio capaz de hibridarse con parte de un oligonucleótido del dominio de captura; y
 - (ii) un dominio que es complementario a un oligonucleótido de dominio universal;
 - (b) hibridar con las sondas de superficie inmovilizadas en el sustrato, capturar oligonucleótidos del dominio y oligonucleótidos del dominio universal;

(c) unir los oligonucleótidos de dominio universal a los oligonucleótidos del dominio de captura para producir las sondas de captura.

Los elementos del sustrato objeto producido por el método anterior para producir la matriz descrita en la presente descripción pueden definirse además de acuerdo con la descripción anterior.

Por lo tanto, en una modalidad la invención proporciona un sustrato objeto para su uso en la detección localizada o espacial del ARN en una muestra de tejido, que comprende un sustrato plano en el que una o más especies de sonda de captura, que comprende un dominio de captura, se inmoviliza directa o indirectamente de manera que cada sonda se orienta para tener un extremo 3' libre para permitir que dicha sonda funcione como un cebador de transcriptasa inversa (RT) en donde las sondas se inmovilizan en el sustrato objeto con una distribución homogénea y en donde la sonda de captura se selecciona de un oligonucleótido que comprende una secuencia de oligonucleótidos poli-T, poli-U y/o aleatorios.

La invención se describirá adicionalmente con referencia a los siguientes ejemplos no limitantes con referencia a los siguientes dibujos en los que:

La Figura 1 muestra la composición y síntesis de la sonda de superficie 3' a 5' de sondas de captura orientadas 5' a 3' que se inmovilizan indirectamente en la superficie de la matriz.

La Figura 2 muestra cómo pueden usarse los métodos de la invención para determinar las condiciones óptimas para capturar ARN a partir de una muestra de tejido en un sustrato, en donde: (A) muestra una muestra de tejido que no se permeabiliza lo suficiente como para permitir la captura de ARN en el sustrato; (B) muestra una muestra de tejido que se permeabiliza para permitir la captura de ARN en el sustrato mientras retiene la información espacial; y (C) muestra una muestra de tejido que es demasiado permeable de manera que se ha permitido que el ARN se difunda fuera de su origen en la muestra de tejido y el ARN capturado no se correlaciona con precisión con su distribución espacial original en la muestra de tejido.

La Figura 3 muestra un gráfico de barras que demuestra la eficiencia de la escisión enzimática (USER o RsaI) de matrices fabricadas internamente y por 99 °C de agua de matrices fabricadas por Agilent, medida mediante hibridación de sondas etiquetadas con fluorescencia a la superficie de la matriz después de la liberación de la sonda.

La Figura 4 muestra una tabla que enumera las lecturas ordenadas por su origen a través de la matriz de captura de ADN fabricada internamente a baja densidad como se ve en la representación esquemática.

La Figura 5 muestra una ilustración esquemática del principio del método descrito en el Ejemplo 4, es decir, el uso de micromatrices con oligos de ADN inmovilizados (sondas de captura) que llevan secuencias de etiquetas de etiquetado espacial (dominios posicionales). Cada elemento de los oligos de la micromatriz lleva 1) una etiqueta de etiquetado única (dominio posicional) y 2) una secuencia de captura (dominio de captura).

La Figura 6 muestra la composición de sondas de captura orientadas de 5' a 3' usadas en matrices de captura de alta densidad.

La Figura 7 muestra una visualización de Matlab de los transcritos capturados del ARN total extraído del bulbo olfativo de ratón.

La Figura 8 muestra los transcritos de Olfr (receptor olfativo) visualizados a través de la matriz de captura mediante el uso de la visualización de Matlab después de la captura a partir del tejido del bulbo olfativo de ratón.

La Figura 9 muestra un patrón de impresión para micromatrices internas con 41-ID-etiqueta.

La Figura 10 muestra una visualización de Matlab de los transcritos etiquetados con ID capturados a partir de tejido de bulbo olfativo de ratón en matrices internas 41-ID-etiqueta superpuestas con la imagen de tejido. Para mayor claridad, los elementos específicos en los que se identificaron genes particulares se han rodeado con un círculo.

La Figura 11 muestra un patrón de síntesis de ADNc en la superficie de la matriz a partir de una sección del bulbo olfativo del cerebro de ratón, visualizado por un escáner de micromatriz Agilent.

La Figura 12 muestra un patrón de síntesis de ADNc en la superficie de la matriz a partir de una sección de la corteza cerebral de ratón, visualizada por un escáner de micromatriz Agilent.

La Figura 13 muestra un patrón de síntesis de ADNc en la superficie de la matriz a partir de una sección del bulbo olfativo del cerebro de ratón, antes de la destrucción (a) y después de la destrucción de áreas no deseadas (b), como se muestra en la imagen en el instrumento de corte celular MMI.

La Figura 14 muestra la biblioteca resultante hecha a partir de una biblioteca de ADNc etiquetado con fluorescencia escindido de áreas no extirpadas visualizadas con un chip de ADN de alta sensibilidad en un bioanalizador Agilent.

La Figura 15 muestra un patrón de síntesis de ADNc en una superficie de matriz de elementos de alta densidad a partir de una sección del bulbo olfativo del cerebro de ratón. Un marco (visible en la parte superior y a la derecha en la imagen) que consiste en elementos que contienen una única secuencia de sonda de ADN se etiquetó mediante la hibridación de un oligonucleótido complementario etiquetado con Cy3.

La Figura 16 muestra un gráfico de barras de las intensidades de Cy3 en matrices que usan diferentes cantidades de dCTP etiquetadas con Cy3.

La Figura 17 muestra imágenes de la síntesis de ADNc en superficies de la matriz e imágenes correspondientes de las secciones de tejido en la superficie de la matriz mediante el uso de varios tipos de muestras alternativas, en donde: (A) representa la imagen de contraste de fase (izquierda) y la correspondiente huella de ADNc etiquetado con Cy3 (derecha) de una muestra de pez cebra; (B) representa la imagen de contraste de fase (izquierda) y la correspondiente huella de ADNc etiquetado con Cy3 (derecha) de una muestra de mosca de la fruta (*Drosophila*); (C) representa la imagen de contraste de fase (izquierda) y la correspondiente huella de ADNc etiquetado con Cy3 (derecha) de una sección de tumor de próstata; y (D) representa la imagen de contraste de fase (izquierda) y la correspondiente huella de ADNc etiquetado con Cy3 (derecha) de células de fibroblasto de ratón.

Ejemplos

Ejemplo 1

Preparación de la matriz

Los siguientes experimentos demuestran cómo las sondas de oligonucleótidos pueden unirse a un sustrato de la matriz por el extremo 5' o 3' para producir una matriz con sondas de captura capaces de hibridarse con el ARNm.

Preparación de una micromatriz impresa internamente con sondas orientadas de 5' a 3'
20 oligonucleótidos de captura de ARN con secuencias de etiquetas individuales (Etiqueta 1-20, Tabla 1 se detectaron en portaobjetos de vidrio para funcionar como sondas de captura. Las sondas se sintetizaron con un amino enlazador de extremo 5'-terminal con un espaciador C6. Todas las preguntas se sintetizaron por Sigma-Aldrich (St. Louis, MO, EE. UU.). Las sondas de captura de ARN se suspendieron a una concentración de 20 μ M en fosfato de sodio 150 mM, pH 8,5 y se detectaron mediante el uso de un Nanoplotter NP2.1/E (Gesim, Grosserkmannsdorf, Alemania) en portaobjetos de micromatriz activados por CodeLink™ (7,5 cm x 2,5 cm; Surmodics, Eden Prairie, MN, EE. UU.). Después de la impresión, el bloqueo de la superficie se realizó de acuerdo con las instrucciones del fabricante. Las sondas se imprimieron en 16 matrices idénticas en el portaobjetos, y cada matriz contenía un patrón de impresión predefinido. Las 16 submatrices se separaron durante la hibridación mediante una máscara de 16 almohadillas (ChipClip™ Schleicher & Schuell BioScience, Keene, NH, EE. UU.).

Table 1

[illegible]

5
10
15
20
25
30
35
40
45
50
55
60
65

Captura LP_Poli-4TVN	CTGATCTCCACTCCCACTCTGCAATTTTTTTTTTTTTTTT (SEQ ID NO: 34)	Fcsoforlado	45
Captura LP_Poli-424T	CTGATCTCCACTCCCACTCTGCAATTTTTTTTTTTTTTTT (SEQ ID NO: 35)	Fcsoforlado	47
Mangos de amplificación universal secundarios adicionales			
Mango D de emp. Illumina	AGACGTGTGTCTGTCTTCCGATCT (SEQ ID NO: 36)		21
Mango X de emp. universal	ACGTCTGTGAAATAGCGGCAT (SEQ ID NO: 37)		20
Mango B_R6 (o X)	AGACGTGTGTCTCTTCCGATCTTTTTTTTTTTTTTTT (SEQ ID NO: 38)		27 (26)
Mango B_R8 (o X)	AGACGTGTGTCTCTTCCGATCTTTTTTTTTTTTTTTT (SEQ ID NO: 39)		23 (28)
E_poliTVN (o X)	AGACGTGTGTCTCTTCCGATCTTTTTTTTTTTTTTTT (SEQ ID NO: 40)		43 (42)
E_poli24T (o X)	AGACGTGTGTCTCTTCCGATCTTTTTTTTTTTTTTTT (SEQ ID NO: 41)		45 (44)
Mango de amplificación para incorporar un mango A en los productos del mango P			
Mango A_P	ACACCTCTTCCCTACACGACGCTCTTCCGATCTATCTCCGATCTCGAA (SEQ ID NO: 42)		53

Preparación de una micromatriz impresa interna con sondas orientadas de 3' a 5' y síntesis de sondas de captura orientadas de 5' a 3'

La impresión de oligonucleótidos de sonda de superficie se realizó como en el caso de sondas orientadas 5' a 3' anteriores, con un enlazador amino-C7 en el extremo 3', como se muestra en la Tabla 1.

Para hibridar los cebadores para la síntesis de la sonda de captura, la solución de hibridación que contiene 4xSSC y SDS al 0,1 %, el cebador de extensión 2 μ M (el oligonucleótido de dominio universal) y el cebador de unión de rosca 2 μ M (el oligonucleótido de dominio de captura) se incubó durante 4 min a 50 °C. Mientras tanto, la matriz interna se unió a un ChipClip (Whatman). La matriz se incubó subsecuentemente a 50 °C durante 30 min a 300 rpm de agitación con 50 μ l de solución de hibridación por pocillo.

Después de la incubación, la matriz se retiró del ChipClip y se lavó con las 3 etapas siguientes: 1) solución de 2xSSC a 50 °C con SDS al 0,1 % durante 6 min a agitación de 300 rpm; 2) 0,2xSSC durante 1 min a agitación de 300 rpm; y 3) 0,1xSSC durante 1 min a agitación de 300 rpm. La matriz se centrifugó luego y se colocó de nuevo en el ChipClip. Para la reacción de extensión y ligazón (para generar el dominio posicional de la sonda de captura) se pipetearon a cada pocillo 50 μ l de mezcla enzimática que contenía tampón de 10xAmpligase, fragmento de Stoffel de ADN polimerasa AmpliTaq 2,5 U (Applied Biosystems), Ampligase 10 U (Epicentre Biotechnologies), dNTP 2 mM cada uno (Fermentas) y agua. La matriz se incubó subsecuentemente a 55 °C durante 30 min. Después de la incubación, la matriz se lavó de acuerdo con el método de lavado de matriz descrito anteriormente, pero la primera etapa tiene la duración de 10 min en lugar de 6 min.

El método se representa en la Figura 1.

Preparación de tejidos

Los siguientes experimentos demuestran cómo pueden prepararse las secciones de muestra de tejido para su uso en los métodos de la invención.

Preparación de tejido congelado fresco y seccionamiento en matrices de sondas de captura

El tejido cerebral de ratón fresco no fijado se recortó si fuera necesario y se congeló en isopentano frío a -40 °C y subsecuentemente se montó para seccionarse con un criostato a 10 μ m. Se aplicó un segmento de tejido sobre cada matriz de sonda de captura que se va a usar.

Preparación de tejido fijado en formalina e incluido en parafina (FFPE)

El tejido cerebral de ratón se fijó en formalina al 4 % a 4 °C durante 24 h. Después de eso se incubó de la siguiente manera: incubación 3x en etanol al 70 % durante 1 hora; incubación 1x en etanol al 80 % durante 1 hora; incubación 1x en etanol al 96 % durante 1 hora; incubación 3x en etanol al 100 % durante 1 hora; e incubación 2x en xileno a temperatura ambiente durante 1 h.

Luego, las muestras deshidratadas se incubaron en parafina líquida de baja fusión 52-54 °C durante hasta 3 horas, durante las cuales la parafina se cambió una vez para lavar el xileno residual. Los bloques de tejidos terminados se almacenaron a TA. A continuación, las secciones se cortaron a 4 μ m en parafina con un micrótopo en cada matriz de sonda de captura que se va a usar.

Las secciones se secaron a 37 °C en los portaobjetos de la matriz durante 24 horas y se almacenaron a TA.

Desparafinación del tejido FFPE

Las secciones de 10 μ m de cerebro de ratón parafinado fijadas con formalina unidas a los portaobjetos CodeLink se desparafinaron en xileno dos veces durante: 10 min, 99,5 % de etanol durante 2 min; 96 % de etanol durante 2 min; 70 % de etanol durante 2 min; y luego se secaron al aire.

Síntesis de ADNc

Los siguientes experimentos demuestran que el ARNm capturado en la matriz a partir de las secciones de muestra de tejido pueden usarse como plantilla para la síntesis de ADNc.

Síntesis de ADNc en chip

Se unieron un soporte para máscara de 16 pocillos y portaobjetos Chip Clip de Whatman a un portaobjetos CodeLink. El sistema de RT-PCR de una etapa SuperScript™III con ADN polimerasa de Invitrogen se usó para realizar la síntesis de ADNc. Para cada reacción se mezclaron 25 μ l de mezcla de reacción 2x (sistema de RT-PCR de una etapa

SuperScript™III con Platinum®Taq ADN polimerasa, Invitrogen), 22,5 µl de H₂O y 0,5 µl de 100xBSA y se calentaron hasta 50 °C. Se añadió la mezcla enzimática Superscript III/Platinum Taq a la mezcla de reacción, 2 µl por reacción y se añadió 50 µl de la mezcla de reacción a cada pocillo del chip. El chip se incubó a 50 °C durante 30 min (Thermomixer Comfort, Eppendorf).

La mezcla de reacción se retiró de los pocillos y el portaobjetos se lavó con: 2xSSC, SDS al 0,1 % a 50 °C durante 10 min; 0,2xSSC a temperatura ambiente durante 1 min; y 0,1xSSC a temperatura ambiente durante 1 min. A continuación, el chip se secó por centrifugación.

En el caso de las secciones de tejido FFPE, las secciones ahora podrían teñirse y visualizarse antes de extraer el tejido, véase la sección más abajo sobre visualización.

Visualización

Hibridación de sondas marcadoras fluorescentes antes de la tinción

Antes de la aplicación de tejido, las sondas marcadoras fluorescentes se hibridaron con elementos que comprenden oligonucleótidos marcadores impresos en la matriz de sondas de captura. Las sondas marcadoras fluorescentes ayudan en la orientación de la imagen resultante después de la visualización del tejido, lo que hace posible combinar la imagen con los perfiles de expresión resultantes para las secuencias "etiqueta" (dominio posicional) individuales de la sonda de captura obtenidas después de la secuenciación. Para hibridar sondas fluorescentes, se incubó una solución de hibridación que contenía 4xSSC y SDS al 0,1 %, sonda de detección (P) 2 µM durante 4 min a 50 °C. Mientras tanto, la matriz interna se unió a un ChipClip (Whatman). La matriz se incubó subsecuentemente a 50 °C durante 30 min a 300 rpm de agitación con 50 µl de solución de hibridación por pocillo.

Después de la incubación, la matriz se retiró del ChipClip y se lavó con las 3 etapas siguientes: 1) solución de 2xSSC a 50 °C con SDS al 0,1 % durante 6 min a agitación de 300 rpm; 2) 0,2xSSC durante 1 min a agitación de 300 rpm; y 3) 0,1xSSC durante 1 min a agitación de 300 rpm. La matriz se centrifugó luego en seco.

Tinción histológica general de secciones de tejido FFPE antes o después de la síntesis de ADNc

Las secciones de tejido FFPE inmovilizadas en matrices de sondas de captura se lavaron y rehidrataron después de la desparafinación antes de la síntesis de ADNc como se describió anteriormente, o se lavaron después de la síntesis de ADNc como se describió anteriormente. A continuación, se tratan de la siguiente manera: incubar durante 3 minutos en hematoxilina; enjuagar con agua desionizada; incubar 5 minutos en agua del grifo; sumergir rápidamente de 8 a 12 veces en etanol ácido; enjuagar 2x1 minuto en agua del grifo; enjuagar 2 minutos en agua desionizada; incubar 30 segundos en eosina; lavar 3x5 minutos en etanol al 95 %; lavar 3x5 minutos en etanol al 100 %; lavar 3x10 minutos en xileno (puede hacerse durante toda la noche); colocar el cubreobjetos en los portaobjetos mediante DPX; secar los portaobjetos en la campana durante la noche.

Tinción inmunohistoquímica general de una proteína diana en secciones de tejido FFPE antes o después de la síntesis de ADNc

Las secciones de tejido FFPE inmovilizadas en matrices de sondas de captura se lavaron y rehidrataron después de la desparafinación antes de la síntesis de ADNc como se describió anteriormente, o se lavaron después de la síntesis de ADNc como se describió anteriormente. Luego se trataron de la siguiente manera sin dejar que se secan durante todo el proceso de tinción: las secciones se incubaron con el anticuerpo primario (anticuerpo primario diluido en solución de bloqueo que comprende solución salina tamponada 1xTris (Tris 50 mM, NaCl 150 mM, pH 7,6), suero de burro al 4 % y tritón-x al 0,1 %) en una cámara húmeda durante toda la noche a TA; se enjuagaron tres veces con 1xTBS; se incubaron en sección con el anticuerpo secundario correspondiente conjugado a un fluorocromo (FITC, Cy3 o Cy5) en una cámara húmeda a TA durante 1 hora. Enjuagar 3x con 1xTBS, eliminar todo lo posible de TBS y la sección de montaje con Prolong Gold +DAPI (Invitrogen) y analizar con microscopio de fluorescencia y conjuntos de filtros coincidentes.

Eliminación del tejido residual

Tejido congelado

Para el tejido cerebral de ratón recién congelado, la etapa de lavado directamente después de la síntesis de ADNc fue suficiente para eliminar completamente el tejido.

Tejido FFPE

Los portaobjetos con secciones de tejido cerebral de ratón parafinadas fijadas en formalina se unieron a soportes de portaobjetos ChipClip y máscaras de 16 pocillos (Whatman). Para cada 150 µl de tampón de digestión con Proteinasa

K del kit RNeasy FFPE (Qiagen), se añadió 10 µl de solución de Proteínasa K (Qiagen). Se añadieron 50 µl de la mezcla final a cada pocillo y el portaobjetos se incubó a 56 °C durante 30 min.

Liberación de sonda de captura (ADNc)

Liberación de la sonda de captura con la mezcla enzimática de escisión de uracilo USER en tampón de PCR (sondas unidas covalentemente)

Se unió una máscara de 16 pocillos y un portaobjetos CodeLink al soporte ChipClip (Whatman). Se calentó a 37 °C 50 µl de una mezcla que contenía tampón de reacción de alta fidelidad 1x FastStart con MgCl₂ 1,8 mM (Roche), dNTP 200 µM (New England Biolabs) y 0,1 U/µl enzima USER (New England Biolabs) y se añadió a cada pocillo y se incubó a 37 °C durante 30 min con mezcla (3 segundos a 300 rpm, 6 segundos en reposo) (Confort del mezclador térmico; Eppendorf). La mezcla de reacción que contenía el ADNc liberado y las sondas se recuperaron de los pocillos con una pipeta.

Liberación de la sonda de captura con la mezcla enzimática de escisión de uracilo USER en tampón TdT (transferasa terminal) (sondas unidas covalentemente)

Se calentó a 37 °C, 50 µl de una mezcla que contiene: tampón TdT 1x (tris-acetato 20 mM (pH 7,9), acetato de potasio 50 mM y acetato de magnesio 10 mM) (New England Biolabs, www.neb.com); BSA 0,1 µg/µl (New England Biolabs); y enzima USER 0,1 U/µl (New England Biolabs) y se añadió a cada pocillo y se incubó a 37 °C durante 30 min con mezcla (3 segundos a 300 rpm, 6 segundos en reposo) (Thermomixer comfort; Eppendorf). La mezcla de reacción que contenía el ADNc liberado y las sondas se recuperaron de los pocillos con una pipeta.

Liberación de la sonda de captura con agua caliente hirviendo (sondas acopladas covalentemente)

Se unió una máscara de 16 pocillos y un portaobjetos CodeLink al soporte ChipClip (Whatman). Se pipetearon 50 µl de agua a 99 °C en cada pocillo. Se dejó reaccionar el agua a 99 °C durante 30 minutos. La mezcla de reacción que contenía el ADNc liberado y las sondas se recuperaron de los pocillos con una pipeta.

Liberación de la sonda de captura con tampón de PCR calentado (sondas de captura sintetizadas *in situ* híbridadas, es decir, sondas de captura hibridadas a sondas de superficie)

Se precalentó a 95 °C, 50 µl de una mezcla que contiene: tampón TdT 1x (tris-acetato 20 mM (pH 7,9), acetato de potasio 50 mM y acetato de magnesio 10 mM) (New England Biolabs, www.neb.com); BSA 0,1 µg/µl (New England Biolabs); y enzima USER 0,1 U/µl (New England Biolabs). Luego la mezcla se añadió a cada pocillo y se incubó durante 5 minutos a 95 °C con mezcla (3 segundos a 300 rpm, 6 segundos en reposo) (Thermomixer comfort; Eppendorf). La mezcla de reacción que contenía las sondas liberadas se recuperó de los pocillos.

Liberación de la sonda de captura con tampón TdT calentado (transferasa terminal) (sondas de captura sintetizadas *in situ* híbridadas, es decir, sondas de captura hibridadas a sondas de superficie)

Se precalentó a 95 °C, 50 µl de una mezcla que contiene: tampón TdT 1x (trisacetato 20 mM (pH 7,9), acetato de potasio 50 mM y acetato de magnesio 10 mM) (New England Biolabs, www.neb.com); BSA 0,1 µg/µl (New England Biolabs); y enzima USER 0,1 U/µl (New England Biolabs). Luego la mezcla se añadió a cada pocillo y se incubó durante 5 minutos a 95 °C con mezcla (3 segundos a 300 rpm, 6 segundos en reposo) (Thermomixer comfort; Eppendorf). La mezcla de reacción que contenía las sondas liberadas se recuperó de los pocillos.

La eficacia de tratar la matriz con la enzima USER y agua calentada a 99 °C puede observarse en la Figura 3. La escisión enzimática mediante el uso de la enzima USER y la enzima RsaI se realizó mediante el uso de las matrices "internas" descritas anteriormente (Figura 3). La liberación mediada por agua caliente de sondas de superficie de ADN también se realizó mediante el uso de matrices comerciales fabricadas por Agilent.

Recopilación de la sonda e introducción del enlazador

Los experimentos demuestran que el ADNc de la primera hebra liberado de la superficie de la matriz puede modificarse para producir ADN bicatenario y subsecuentemente amplificarse.

Amplificación del transcriptoma completo por el kit de amplificación de genoma completo Picoplex (secuencias de sonda de captura que incluyen secuencias de dominio posicional (etiqueta) no retenidas en el borde del ADNc resultante)

Las sondas de captura se liberaron con una mezcla enzimática de escisión de uracilo USER en tampón de PCR (sondas de captura acopladas covalentemente) o con tampón de PCR calentado (sondas de captura sintetizadas *in situ* híbridadas, es decir, sondas de captura hibridadas a sondas de superficie).

El ADNc liberado se amplificó mediante el uso del método de amplificación del genoma completo del cebador aleatorio Picoplex (Rubicon Genomics), que se llevó a cabo de acuerdo con las instrucciones del fabricante.

Amplificación del transcriptoma completo por cola de dA con transferasa terminal (TdT) (secuencias de sonda de captura que incluyen secuencias de dominio posicional (etiqueta) retenidas en el extremo del ADNc resultante)

Las sondas de captura se liberaron con una mezcla enzimática de escisión de uracilo USER en tampón TdT (transferasa terminal) (sondas de captura unidas covalentemente) o con tampón TdT calentado (transferasa terminal) (sondas de captura sintetizadas *in situ* híbridadas, es decir, sondas de captura hibridadas a sondas de superficie).

Se colocaron 38 µl de mezcla de escisión en un tubo de PCR limpio de 0,2 ml. La mezcla contenía: tampón TdT 1x (trisacetato 20 mM (pH 7,9), acetato de potasio 50 mM y acetato de magnesio 10 mM) (New England Biolabs, www.neb.com); BSA 0,1 µg/µl (New England Biolabs); enzima USER 0,1 U/µl (New England Biolabs) (no para liberación calentada); ADNc liberado (extendido de sondas de superficie); y sondas de superficie liberadas. Al tubo de PCR, se añadieron 0,5 µl de ARNasa H (5 U/µl, concentración final de 0,06 U/µl), 1 µl de TdT (20 U/µl, concentración final de 0,5 U/µl) y 0,5 µl de dATP (100 mM, concentración final de 1,25 mM). Para la cola de dA, el tubo se incubó en un termociclador (Applied Biosystems) a 37 °C durante 15 min seguido de una inactivación de TdT a 70 °C durante 10 min. Después de cola de dA, se preparó una mezcla maestra de PCR. La mezcla contenía: tampón de PCR 1x Faststart HiFi (pH 8,3) con MgCl₂ 1,8 mM (Roche); 0,2 mM de cada dNTP (Fermentas); 0,2 µM de cada cebador, A (complementario al dominio de amplificación de la sonda de captura) y B₁(dT)₂₄ (Eurofins MWG Operon) (complementario a la cola de poli-A a añadir al extremo 3' de la primera hebra de ADNc); y 0,1 U/µl de ADN polimerasa Faststart HiFi (Roche). Se colocaron 23 µl de mezcla maestra de PCR en nueve tubos de PCR limpios de 0,2 ml. Se añadieron 2 µl de mezcla de cola de dA a ocho de los tubos, mientras que se añadió 2 µl de agua (sin ARNasa/ADNasa) al último tubo (control negativo). La amplificación por PCR se llevó a cabo con el siguiente programa: Inicio en caliente a 95 °C durante 2 minutos, síntesis de la segunda hebra a 50 °C durante 2 minutos y 72 °C durante 3 minutos, amplificación con 30 ciclos de PCR a 95 °C durante 30 segundos, 65 °C durante 1 minuto, 72 °C durante 3 minutos y una extensión final a 72 °C durante 10 minutos.

Limpieza y análisis posteriores a la reacción

Se combinaron cuatro productos de amplificación y se procesaron a través de una columna de purificación por PCR Qiaquick (Qiagen) y se eluyeron en 30 µl de EB (Tris-Cl 10 mM, pH 8,5). El producto se analizó en un bioanalizador (Agilent). Se usó un kit de ADN 1000 de acuerdo con las instrucciones del fabricante.

Secuenciación

Secuenciación de Illumina

La biblioteca de ADNc para la secuenciación de Illumina mediante el uso del indexado de muestras se llevó a cabo de acuerdo con las instrucciones del fabricante. La secuenciación se llevó a cabo en una plataforma HiSeq2000 (Illumina).

Bioinformática

Obtener información transcriptómica digital a partir de datos de secuenciación de bibliotecas de transcriptomas completas amplificadas mediante el uso del enfoque de transferasa terminal de cola de Da

Los datos de secuenciación se clasificaron a través del kit de herramientas FastX de la herramienta divisora de código de barras FASTQ en archivos individuales para las respectivas secuencias de dominio posicional (etiqueta) de la sonda de captura. Los datos de secuenciación etiquetados individualmente se analizaron a través del mapeo al genoma de ratón con la herramienta de mapeo Tophat. El archivo SAM resultante se procesó para los recuentos de transcritos a través del software de recuento HTseq.

Obtener información transcriptómica digital a partir de datos de secuenciación de bibliotecas de transcriptomas completas amplificadas mediante el uso del enfoque del kit de amplificación del genoma completo Picoplex

Los datos de secuenciación se convirtieron de formato FASTQ a formato FASTA mediante el uso del kit de herramientas FastX conversor FASTQ a FASTA. Las lecturas de secuenciación se alinearon con las secuencias de dominio posicional (etiqueta) de la sonda de captura mediante el uso de Blastn y las lecturas con aciertos mejores que 1 e⁻⁶ a una de las secuencias de etiquetas se clasificaron en archivos individuales para cada secuencia de etiquetas respectivamente. El archivo de lecturas de secuencias de etiquetas se alineó entonces mediante el uso de Blastn al transcriptoma de ratón, y se recogieron los aciertos.

Combinación de datos de visualización y perfiles de expresión

Los perfiles de expresión para las secuencias de dominio posicional (etiqueta) de la sonda de captura individual se combinan con la información espacial obtenida de las secciones de tejido a través de la tinción. De esta manera, los datos transcriptómicos de los compartimentos celulares de la sección de tejido pueden analizarse de manera directamente comparativa, con la disponibilidad de distinguir elementos de expresión distintas para diferentes subtipos celulares en un contexto estructural dado

Ejemplo 2

Se colocaron secciones de tejido cerebral de ratón FFPE teñidas (bulbo olfativo) sobre una matriz de captura de transcriptomas codificado por barras, de acuerdo con el procedimiento general descrito en el Ejemplo 1. En comparación con el experimento con tejido congelado fresco en el Ejemplo 1, se observó una mejor morfología con el tejido FFPE.

Ejemplo 3

Amplificación del transcriptoma completo por síntesis de la segunda hebra del cebador aleatorio seguido de amplificación del mango universal (secuencias de sonda de captura que incluyen secuencias etiqueta retenidas en el extremo del ADNc resultante)

Después de la liberación de la sonda de captura con la mezcla enzimática de escisión de uracilo USER en tampón de PCR (sondas unidas covalentemente)

O

Después de la liberación de la sonda de captura con tampón de PCR calentado (sondas de captura sintetizadas *in situ* híbridadas)

Se añadió 1 µl de ARNasa H (5 U/µl) a cada uno de los dos tubos, concentración final de 0,12 U/µl, que contenía 40 µl de tampón de PCR 1x Faststart HiFi (pH 8,3) con MgCl₂ 1,8 mM (Roche, www.roche-applied-science.com), 0,2 mM de cada dNTP (Fermentas, www.fermentas.com), BSA 0,1 µg/µl (New England Biolabs, www.neb.com), enzima USER 0,1 U/µl (New England Biolabs), ADNc liberado (extendido de sondas de superficie) y sondas de superficie liberadas. Los tubos se incubaron a 37 °C durante 30 min seguido de 70 °C durante 20 min en un termociclador (Applied Biosystems, www.appliedbiosystems.com). Se añadió 1 µl de fragmento de Klenow (3' a 5' exo menos) (Illumina, www.illumina.com) y 1 µl de cebador aleatorio acoplado con mango (10 µM) (Eurofins MWG Operon, www.eurofinsdna.com) a los dos tubos (B_R8 (octámero) a uno de los tubos y B_R6 (hexámero) al otro tubo), concentración final de 0,23 µM. Los dos tubos se incubaron a 15 °C durante 15 min, 25 °C durante 15 min, 37 °C durante 15 min y finalmente 75 °C durante 20 min en un termociclador (Applied Biosystems). Después de la incubación, se añadió 1 µl de cada cebador, A_P y B (10 µM) (Eurofins MWG Operon), a ambos tubos, concentración final de 0,22 µM cada uno. También se añadió 1 µl de ADN polimerasa Faststart HiFi (5 U/µl) (Roche) a ambos tubos, concentración final de 0,11 U/µl. La amplificación de PGR se llevó a cabo en un termociclador (Applied Biosystems) con el siguiente programa: Inicio en caliente a 94 °C durante 2 min, seguido de 50 ciclos a 94 °C durante 15 segundos, 55 °C durante 30 segundos, 68 °C durante 1 minuto y una extensión final a 68 °C durante 5 minutos. Después de la amplificación, se purificaron 40 µl de cada uno de los dos tubos con columnas de purificación por PCR Qiaquick (Qiagen, www.qiagen.com) y se eluyeron en 30 µl de EB (Tris-Cl 10 mM, pH 8,5). Los productos purificados se analizaron con un bioanalizador (Agilent, www.home.agilent.com), se usó el kit de ADN 7500. Este ejemplo demuestra el uso de síntesis de la segunda hebra de hexámero aleatorio y octámero aleatorio, seguido de amplificación para generar la población a partir de las moléculas de ADNc liberadas.

Ejemplo 4

Amplificación de productos específicos de ID y específicos de genes después de la síntesis de ADNc y la recolección de sondas

Después de la liberación de la sonda de captura con la mezcla enzimática de escisión de uracilo USER en el tampón de PGR (sondas unidas covalentemente).

El ADNc escindido se amplificó en volúmenes de reacción finales de plantilla escindida de 10 µl. 7 µl, cebador directo específico de ID de 1 µl (2 µM), cebador inverso específico de gen de 1 µl (2 µM) y mezcla enzimática de alta fidelidad FastStart 1 µl en tampón de reacción de alta fidelidad FastStart 1,4x con MgCl₂ 1,8 mM para dar una reacción final de 10 µl con tampón de reacción de alta fidelidad FastStart 1x con MgCl₂ 1,8 mM y mezcla enzimática de alta fidelidad FastStart 1 U. La amplificación de PGR se llevó a cabo en un termociclador (Applied Biosystems) con el siguiente programa: Inicio en caliente a 94 °C durante 2 min, seguido de 50 ciclos a 94 °C durante 15 segundos, 55 °C durante 30 segundos, 68 °C durante 1 minuto y una extensión final a 68 °C durante 5 minutos.

Secuencias de cebadores, resultantes en un producto de aproximadamente 250 pb de cebador de Beta-2 microglobulina (B2M)

5'-TGGGGGTGAGAATTGCTAAG-3' (SEQ ID NO: 43)

Cebador ID-1

5'-CCTTCTCCTTCTCCTTCACC-3' (SEQ ID NO: 44)

Cebador ID-5

5'-GTCCTCTATTCCGTCACCAT-3' (SEQ ID NO: 45)

Cebador ID-20

5'-CTGCTTCTTCCTGGAAGTCA-3' (SEQ ID NO: 46)

Los resultados muestran la amplificación exitosa de productos específicos de ID y específicos de genes mediante el uso de dos cebadores de ID diferentes (es decir, específicos para etiquetas de ID colocadas en diferentes ubicaciones en la micromatriz y el mismo cebador específico de genes de un tejido cerebral que cubre todas las sondas). En consecuencia, este experimento establece que los productos pueden identificarse mediante una reacción de amplificación específica de la etiqueta de ID o específica del ácido nucleico diana. Se establece además que pueden distinguirse diferentes etiquetas de ID. Un segundo experimento, con tejido que cubrió solo la mitad de las sondas de ID (es decir, sondas de captura) en la matriz dio como resultado un resultado positivo (producto de PCR) para los puntos que se cubrieron con tejido.

Ejemplo 5

Síntesis alternativa de sondas de captura orientadas 5' a 3' mediante el uso de extensión con polimerasa y cola de la transferasa terminal

Para hibridar los cebadores para la solución de hibridación de síntesis de sonda de captura que contenía 4xSSC y SDS al 0,1 % y cebador de extensión 2 μ M (A_cebador) se incubó durante 4 min a 50 °C. Mientras tanto, la matriz interna (véase el Ejemplo 1) se unió a un ChipClip (Whatman). La matriz se incubó subsecuentemente a 50 °C durante 30 min a agitación de 300 rpm con 50 μ l de solución de hibridación por pocillo.

Después de la incubación, la matriz se retiró del ChipClip y se lavó con las 3 etapas siguientes: 1) solución de 2xSSC a 50 °C con SDS al 0,1 % durante 6 min a agitación de 300 rpm; 2) 0,2xSSC durante 1 min a agitación de 300 rpm; y 3) 0,1xSSC durante 1 min a agitación de 300 rpm. La matriz se centrifugó luego y se colocó de nuevo en el ChipClip.

1 μ l de fragmento Klenow (3' a 5' exo menos) (Illumina, www.illumina.com) junto con tampón Klenow 10x, dNTP 2 mM cada uno (Fermentas) y agua, se mezcló en una reacción de 50 μ l y se pipeteó en cada pocillo.

La matriz se incubó a 15 °C durante 15 min, 25 °C durante 15 min, 37 °C durante 15 min y finalmente 75 °C durante 20 min en un termomezclador Eppendorf.

Después de la incubación, la matriz se retiró del ChipClip y se lavó con las 3 etapas siguientes: 1) solución de 2xSSC a 50 °C con SDS al 0,1 % durante 6 min a agitación de 300 rpm; 2) 0,2xSSC durante 1 min a agitación de 300 rpm; y 3) 0,1xSSC durante 1 min a agitación de 300 rpm. La matriz se centrifugó luego y se colocó de nuevo en el ChipClip.

Para la cola de dT se preparó una mezcla de reacción de 50 μ l que contenía tampón TdT 1x (tris-acetato 20 mM (pH 7,9), acetato de potasio 50mM y acetato de magnesio 10mM) (New England Biolabs, www.neb.com), BSA 0,1 μ g/ μ l (New England Biolabs), ARNasa H 0,5 μ l (5U/ μ l), TdT 1 μ l (20 U/ μ l) y dTTP 0,5 μ l (100 mM). La mezcla se añadió a la superficie de la matriz y la matriz se incubó en un termociclador (Applied Biosystems) a 37 °C durante 15 min seguido de una inactivación de TdT a 70 °C durante 10 min.

Ejemplo 6

Transcriptómica espacial mediante el uso de matrices de alta densidad de sonda 5' a 3' y tejido congelado fijado en formalina (FF-congelado) con escisión y amplificación del sistema USER a través de transferasa terminal

Preparación de la matriz

Las micromatrices de alta densidad prefabricadas se pidieron de Roche- Nimblegen (Madison, WI, EE. UU.). Cada matriz de sondas de captura contenía 135 000 elementos, de los cuales 132 640 elementos portaban una sonda de captura que comprende una secuencia de etiqueta de ID única (dominio posicional) y una región de captura (dominio de captura). Cada elemento tenía un tamaño de 13x13 μ m. Las sondas de captura se componen de 5' a 3' de un dominio universal que contiene cinco bases dUTP (un dominio de escisión) y un dominio de amplificación general, una etiqueta de ID (dominio posicional) y una región de captura (dominio de captura) (Figura 6 y Tabla 4). Cada matriz también se equipó con un marco de sondas marcadoras que portaban una secuencia genérica de 30 pb (Tabla 4) para permitir la hibridación de sondas fluorescentes para ayudar con la orientación durante la visualización de la matriz.

Preparación de tejidos: preparación de tejido congelado fijado en formalina

El animal (ratón) se perfundió con 50 ml de PBS y 100 ml de solución de formalina al 4 %. Después de la escisión del bulbo olfativo, el tejido se colocó en un baño de formalina al 4 % para su subsecuente fijación durante 24 horas. El tejido se trató entonces con sacarosa en sacarosa al 30 % disuelta en PBS durante 24 horas para estabilizar la morfología y eliminar el exceso de formalina. El tejido se congeló a una velocidad controlada hasta -40 °C y se mantuvo a -20 °C entre experimentos. Se llevó a cabo una preparación similar de tejido posfijado durante 3 horas o sin fijación posterior para una muestra paralela. La perfusión con formalina al 2 % sin fijación posterior también se usó con éxito. De manera similar, podría omitirse la etapa de tratamiento con sacarosa. El tejido se montó en un criostato para seccionar a 10 µm. Se aplicó un segmento de tejido sobre cada matriz de sonda de captura que se va a usar. Opcionalmente para una mejor adherencia al tejido, el chip de matriz se colocó a 50 °C durante 15 minutos.

Control opcional - preparación de ARN total a partir de tejido seccionado

El ARN total se extrajo de una única sección de tejido (10 µm) mediante el uso del kit de RNeasy FFPE (Qiagen) de acuerdo con las instrucciones del fabricante. El ARN total obtenido de la sección de tejido se usó en experimentos de control para una comparación con experimentos en los que el ARN se capturó en la matriz directamente de la sección de tejido. En consecuencia, en el caso donde se aplicó ARN total a la matriz, se omitió la tinción, visualización y degradación de las etapas del tejido.

Reacciones en chip

La hibridación de la sonda marcadora con las sondas del marco, la transcripción inversa, la tinción nuclear, la digestión del tejido y las reacciones de escisión de la sonda se realizaron todas en una junta de silicona de 16 pocillos (ArrayIt, Sunnyvale, CA, EE. UU.) con un volumen de reacción de 50 µl por pocillo. Para evitar la evaporación, los casetes se recubrieron con selladores de placas (In Vitro AB, Estocolmo, Suecia).

Opcional - permeabilización del tejido antes de la síntesis de ADNc

Para la permeabilización mediante el uso de la Proteinasa K, la proteinasa K (Qiagen, Hilden, Alemania) se diluyó a 1 µg/ml en PBS. La solución se añadió a los pocillos y el portaobjetos se incubó a temperatura ambiente durante 5 minutos, seguido de un aumento gradual a 80 °C durante 10 minutos. El portaobjetos se lavó brevemente en PBS antes de la reacción de transcripción inversa.

Alternativamente, para la permeabilización mediante el uso de microondas, después de la unión al tejido, el portaobjetos se colocó en el fondo de un frasco de vidrio que contenía 50 ml de 0,2xSSC (Sigma-Aldrich) y se calentó en un horno de microondas durante 1 minuto a 800 W. Justo después del tratamiento con microondas, el portaobjetos se colocó sobre un papel de tejido y se secó durante 30 minutos en una cámara protegida de la exposición innecesaria al aire. Después del secado, el portaobjetos se sumergió brevemente en agua (sin ARNasa/ADNasa) y finalmente se secó por centrifugación antes de iniciar la síntesis de ADNc.

Síntesis de ADNc

Para la reacción de transcripción inversa se usó el sistema de RT-PCR de un paso SuperScript III con Platinum Taq (Life Technologies/Invitrogen, Carlsbad, CA, EE. UU.). Las reacciones de transcripción inversa contenían una mezcla de reacción 1x, una mezcla de BSA 1x (New England Biolabs, Ipswich, MA, EE. UU.) y una mezcla de RT/Platinum Taq SuperScript III de 2 µl en un volumen final de 50 µl. Esta solución se calentó hasta 50 °C antes de la aplicación a las secciones de tejido y la reacción se realizó a 50 °C durante 30 minutos. La solución de transcripción inversa se retiró subsecuentemente de los pocillos y el portaobjetos se dejó secar al aire durante 2 horas.

Visualización de tejidos

Después de la síntesis de ADNc, se realizó simultáneamente la tinción nuclear y la hibridación de la sonda marcadora a las sondas del marco (sondas unidas al sustrato de la matriz para permitir la orientación de la muestra de tejido en el arreglo). Se preparó una solución con DAPI a una concentración de 300 nM y sonda marcadora a una concentración de 170 nM en PBS. Esta solución se añadió a los pocillos y el portaobjetos se incubó a temperatura ambiente durante 5 minutos, seguido de un breve lavado en PBS y secado por centrifugación.

Alternativamente, la sonda marcadora se hibridó con las sondas del marco antes de colocar el tejido en la matriz. Luego, la sonda marcadora se diluyó a 170 nM en tampón de hibridación (4xSSC, SDS al 0,1 %). Esta solución se calentó hasta 50 °C antes de la aplicación al chip y la hibridación se realizó a 50 °C durante 30 minutos a 300 rpm. Después de la hibridación, el portaobjetos se lavó en 2xSSC, SDS al 0,1 % a 50 °C y 300 rpm durante 10 minutos, 0,2xSSC a 300 rpm durante 1 minuto y 0,1xSSC a 300 rpm durante 1 minuto. En ese caso, la solución de tinción después de la síntesis de ADNc solo contenía la tinción de DAPI nuclear diluida a 300 nM en PBS. La solución se

ES 3 019 910 T3

aplicó a los pocillos y el portaobjetos se incubó a temperatura ambiente durante 5 minutos, seguido de un breve lavado en PBS y secado por centrifugación.

5 Las secciones se examinaron microscópicamente con un generador de imágenes Zeiss Axio Z2 y se procesaron con el software MetaSystems.

Eliminación de tejido

10 Las secciones de tejido se digirieron mediante el uso de Proteinasa K diluida a 1,25 $\mu\text{g}/\mu\text{l}$ en tampón PKD del kit de RNeasy FFPE (ambos de Qiagen) a 56 °C durante 30 minutos con una mezcla de intervalo a 300 rpm durante 3 segundos, luego 6 segundos de reposo. El portaobjetos se lavó subsecuentemente en 2xSSC, SDS al 0,1 % a 50 °C y 300 rpm durante 10 minutos, 0,2xSSC a 300 rpm durante 1 minuto y 0,1xSSC a 300 rpm durante 1 minuto.

Liberación de la sonda

15 El casete de hibridación de 16 pocillos con junta de silicona (ArrayIt) se precalentó a 37 °C y se unió al portaobjetos Nimblegen. Un volumen de 50 μl de mezcla de escisión precalentado a 37 °C, que consiste en tampón de lisis a una concentración desconocida (Takara), enzima USER (NEB) 0,1 U/ μl y BSA 0,1 $\mu\text{g}/\mu\text{l}$ se añadió a cada uno de los pocillos que contenían ADNc inmovilizado superficialmente. Después de eliminar las burbujas, el portaobjetos se selló y se incubó a 37 °C durante 30 minutos en una Thermomixer comfort con agitación cíclica a 300 rpm durante 3 segundos con 6 segundos de reposo en el medio. Después de la incubación, se recogió una mezcla de escisión de 45 μl de cada uno de los pocillos usados y se colocó en tubos de PCR de 0,2 ml.

Preparación de bibliotecas

25 Tratamiento con exonucleasa

30 Después de enfriar las soluciones en hielo durante 2 minutos, se añadió Exonucleasa I (NEB), para retirar las sondas de ADNc no extendidas, hasta un volumen final de 46,2 μl y una concentración final de 0,52 U/ μl . Los tubos se incubaron en un termociclador (Applied Biosystems) a 37 °C durante 30 minutos seguido de la inactivación de la exonucleasa a 80 °C durante 25 minutos.

Cola de dA por transferasa terminal

35 Después de la etapa de exonucleasa, se añadió a cada una de las muestras 45 μl de mezcla de cola poliA, de acuerdo con las instrucciones del fabricante que consistían en tampón TdT (Takara), dATP 3mM (Takara) y mezcla de enzimas TdT de los fabricantes (TdT y ARNasa H) (Takara). Las mezclas se incubaron en un termociclador a 37 °C durante 15 minutos seguido de la inactivación de TdT a 70 °C durante 10 minutos.

40 Síntesis de segunda hebra y amplificación por PCR

45 Después de la cola dA, se colocó una mezcla maestra de PCR de 23 μl en cuatro tubos de PCR nuevos de 0,2 ml por muestra, a cada tubo se añadió una muestra de 2 μl como plantilla. Las PCR finales consistieron en tampón Ex Taq 1x (Takara), 200 μM de cada dNTP (Takara), 600 nM A_cebador (MWG), 600 nM B_dT20VN_cebador (MWG) y 0,025 U/ μl Ex Taq polimerasa (Takara)(Tabla 4). Se creó una segunda hebra de ADNc al ejecutar un ciclo en un termociclador a 95 °C durante 3 minutos, 50 °C durante 2 minutos y 72 °C durante 3 minutos. Luego las muestras se amplificaron al ejecutar 20 ciclos (para la preparación de bibliotecas) o 30 ciclos (para confirmar la presencia de ADNc) a 95 °C durante 30 segundos, 67 °C durante 1 minuto y 72 °C durante 3 minutos, seguido de una extensión final a 72 °C durante 10 minutos.

Limpieza de bibliotecas

55 Después de la amplificación, las cuatro PCR (100 μl) se mezclaron con tampón de unión a 500 μl (Qiagen) y se colocaron en una columna de purificación por PCR Qiaquick (Qiagen) y se centrifugaron durante 1 minuto a 17 900 x g para unir el ADNc amplificado a la membrana. Luego, la membrana se lavó con tampón de lavado (Qiagen) que contenía etanol y finalmente se eluyó en 50 μl de Tris-Cl 10 mM, pH 8,5.

60 La muestra purificada y concentrada se purificó y concentró adicionalmente mediante purificación con CA (purificación por perlas superparamagnéticas conjugadas con ácido carboxílico) con un robot MBS (Magnetic Biosolutions). Se usó una concentración final de PEG del 10 % para eliminar fragmentos por debajo de 150-200 pb. Se dejó que el ADNc amplificado se uniera a las perlas de CA (Invitrogen) durante 10 min y luego se eluyó en 15 μl de Tris-Cl 10mM, pH 8,5.

65 Análisis de calidad de la biblioteca

Las muestras amplificadas durante 30 ciclos se analizaron con un bioanalizador Agilent (Agilent) para confirmar la presencia de una biblioteca de ADNc amplificada, el kit de alta sensibilidad de ADN o el kit de ADN 1000 se usaron en dependencia de la cantidad de material.

5 Preparación de bibliotecas de secuenciación

Indexado de bibliotecas

10 Las muestras amplificadas durante 20 ciclos se usaron además para preparar bibliotecas de secuenciación. Se preparó una mezcla maestra de PCR de índice para cada muestra y se colocó 23 µl en seis tubos de 0,2 ml. Se añadieron 2 µl del ADNc amplificado y purificado a cada una de las seis PCR como plantilla que fabricó las PCR que contenían una mezcla maestra de Phusion 1x (Fermentas), InPE1.0 500nM (Illumina), Índice 1-12 500 nM (Illumina) y InPE2.0 0,4 nM (Illumina). Las muestras se amplificaron en un termociclador durante 18 ciclos a 98 °C durante 30 segundos, 65 °C durante 30 segundos y 72 °C durante 1 minuto, seguido de una extensión final a 72 °C durante 5 minutos.

Limpieza de la biblioteca de secuenciación

20 Después de la amplificación, las seis PCR (150 µl) se mezclaron con tampón de unión a 750 µl y se colocaron en una columna de purificación por PCR Qiaquick y se centrifugaron durante 1 minuto a 17 900 x g para unir el ADNc amplificado a la membrana (debido al gran volumen de muestra (900 µl), la muestra se dividió en dos (cada uno 450 µl) y se unió en dos etapas separadas). Luego, la membrana se lavó con tampón de lavado que contenía etanol y finalmente se eluyó en 50 µl de Tris-Cl 10 mM, pH 8,5.

25 La muestra purificada y concentrada se purificó y concentró adicionalmente mediante purificación con CA con un robot MBS. Se usó una concentración final de PEG de 7,8 % para eliminar fragmentos por debajo de 300-350 pb. Se dejó que el ADNc amplificado se uniera a las perlas de CA durante 10 min y luego se eluyó en 15 µl de Tris-Cl 10 mM, pH 8,5. Las muestras se analizaron con un bioanalizador Agilent para confirmar la presencia y el tamaño de las bibliotecas terminadas, el kit de alta sensibilidad al ADN o el kit de ADN 1000 se usaron de acuerdo con las instrucciones del fabricante en dependencia de la cantidad de material.

Secuenciación

35 Las bibliotecas se secuenciaron en Illumina Hiseq2000 o Miseq en dependencia del rendimiento de datos deseado de acuerdo con las instrucciones del fabricante. Opcionalmente para la lectura 2, se usó un cebador de secuenciación personalizado B_r2 para evitar la secuenciación a través del estiramiento homopolímero de 20T.

Análisis de datos

40 La lectura 1 se recortó en 42 bases en el extremo 5'. La lectura 2 se recortaron 25 bases en el extremo 5' (opcionalmente no se recortaron bases de la lectura 2 si se usó el cebador personalizado). A continuación, las lecturas se mapearon con Bowtie al conjunto de genoma de *Mus musculus* 9 enmascarado repetido y la salida se formateó en el formato de archivo SAM. Las lecturas mapeadas se extrajeron y se anotaron con anotaciones génicas de refGene de UCSC. Los índices se recuperaron con 'indexFinder' (un software interno para la recuperación de índices). Luego se creó una base de datos mongo DB que contenía información sobre todos los transcritos atrapados y su respectiva posición de índice en el chip.

50 Se conectó una implementación de matlab a la base de datos y se permitió la visualización espacial y el análisis de los datos (Figura 7).

Opcionalmente, la visualización de datos se superpuso con la imagen microscópica mediante el uso de las sondas de trama etiquetadas con fluorescencia para una alineación exacta y permitir la extracción de datos transcriptómicos espaciales.

55 Ejemplo 7

Transcriptómica espacial mediante el uso de matrices de alta densidad de sonda 3' a 5' y tejido FFPE con escisión y amplificación del sistema MutY a través de TdT

60 Preparación de la matriz

65 Los chips de micromatrices prefabricadas de alta densidad se pidieron a Roche-Nimblegen (Madison, WI, EE. UU.). Cada matriz de sondas de captura usadas contenía 72 k elementos, de las cuales 66 022 contenían una secuencia complementaria única con ID-etiqueta. Cada elemento tenía un tamaño de 16x16 µm. Las sondas de captura se componen de 3' a 5' de la misma manera que las sondas usadas para las matrices impresas internamente de 3' a 5'

con la ejecución a 3 bases adicionales que se añaden al mango general superior (P') de la sonda para convertirla en una versión larga de P', LP' (Tabla 4). Cada matriz también se equipó con un marco de sondas que llevaban una secuencia genérica de 30 pb para permitir la hibridación de sondas fluorescentes para ayudar con la orientación durante la visualización de la matriz.

Síntesis de sondas de captura orientadas de 5' a 3'

La síntesis de sondas de captura orientadas de 5' a 3' en las matrices de alta densidad se llevó a cabo como en el caso de matrices impresas internas, con la excepción de que las etapas de extensión y ligazón se llevaron a cabo a 55 °C durante 15 min seguido de 72 °C durante 15 min. La sonda con mango A (tabla 4) incluía una discrepancia A/G para permitir la liberación subsecuente de sondas a través del sistema enzimático MutY descrito a continuación. La sonda P se reemplazó por una versión LP más larga para coincidir con las sondas más largas en la superficie.

Preparación de tejido fijado en formalina e incluido en parafina y desparafinación

Esto se llevó a cabo como se describió anteriormente en el protocolo interno.

Síntesis y tinción del ADNc

La síntesis y tinción de ADNc se llevó a cabo como en el protocolo para matrices de Nimblegen de alta densidad orientadas 5' a 3' con la excepción de que se añadieron dCTP y dATP etiquetados con biotina a la síntesis de ADNc junto con los cuatro dNTP regulares (cada uno estaba presente 25 veces más que los etiquetados con biotina).

Eliminación de tejido

La eliminación de tejido se llevó a cabo de la misma manera que en el protocolo para matrices de Nimblegen de alta densidad orientadas de 5' a 3' descritas en el Ejemplo 6.

Escisión de la sonda por MutY

Se precalentó una cámara de incubación de 16 pocillos con junta de silicona (ArrayIT) a 37 °C y se unió al portaobjetos Codelink. Un volumen de 50 µl de mezcla de escisión precalentado a 37 °C, que consiste en tampón de endonucleasa VIII 1x (NEB), MutY 10U/µl (Trevigen), Endonucleasa VIII 10U/µl (NEB), se añadió BSA 0,1 µg/µl a cada uno de los pocillos que contenían ADNc inmovilizado en superficie. Después de eliminar las burbujas, el portaobjetos se selló y se incubó a 37 °C durante 30 minutos en un Thermomixer comfort con agitación cíclica a 300 rpm durante 3 segundos con 6 segundos de reposo en el medio. Después de la incubación, se eliminó el sellador de placas y una mezcla de escisión de 40 µl de cada uno de los pocillos usados y se colocó en una placa de PCR.

Preparación de bibliotecas

Limpieza de biblioteca mediada por biotina-estreptavidina

Para eliminar sondas de ADNc no extendidas y cambiar el tampón, las muestras se purificaron mediante la unión del ADNc etiquetado con biotina a C1-perlas recubiertas de estreptavidina (Invitrogen) y el lavado de las perlas con NaOH 0,1 M (fabricado fresco). La purificación se llevó a cabo con un robot MBS (Magnetic Biosolutions), el ADNc etiquetado con biotina se dejó unir a C1-perlas durante 10 min y luego se eluyó en 20 µl de agua al calentar la solución de perla-agua a 80 °C para romper la unión biotina-estreptavidina.

Cola de dA por transferasa terminal

Después de la etapa de purificación, se colocaron 18 µl de cada muestra en nuevos tubos de PCR de 0,2 ml y se mezclaron con 22 µl de una mezcla maestra de cola poliA que dio lugar a una mezcla de reacción de 40 µl de acuerdo con las instrucciones del fabricante que consisten en tampón de lisis (Takara, kit de amplificación de transcriptoma completo Cellamp), tampón TdT (Takara), dATP 1,5 mM (Takara) y mezcla enzimática TdT (TdT y ARNasa H) (Takara). Las mezclas se incubaron en un termociclador a 37 °C durante 15 minutos seguido de la inactivación de TdT a 70 °C durante 10 minutos.

Síntesis de segunda hebra y amplificación por PCR

Después de la cola dA, se colocó una mezcla maestra de PCR de 23 µl en cuatro tubos de PCR nuevos de 0,2 ml por muestra, a cada tubo se añadió una muestra de 2 µl como plantilla. Las PCR finales consistieron en tampón Ex Taq 1x (Takara), 200 µM de cada dNTP (Takara), 600 nM A_cebador (MWG), 600 nM B_dT20VN_cebador (MWG) y 0,025 U/µl Ex Taq polimerasa (Takara). Se creó una segunda hebra de ADNc al ejecutar un ciclo en un termociclador a 95 °C durante 3 minutos, 50 °C durante 2 minutos y 72 °C durante 3 minutos. Luego las muestras se amplificaron al ejecutar 20 ciclos (para la preparación de bibliotecas) o 30 ciclos (para confirmar la presencia de ADNc) a 95 °C

durante 30 segundos, 67 °C durante 1 minuto y 72 °C durante 3 minutos, seguido de una extensión final a 72 °C durante 10 minutos.

Limpieza de bibliotecas

Después de la amplificación, las cuatro PCR (100 µl) se mezclaron con tampón de unión a 500 µl (Qiagen) y se colocaron en una columna de purificación por PCR Qiaquick (Qiagen) y se centrifugaron durante 1 minuto a 17 900 x g para unir el ADNc amplificado a la membrana. Luego, la membrana se lavó con tampón de lavado (Qiagen) que contenía etanol y finalmente se eluyó en 50 µl de Tris-HCl 10 mM, pH 8,5.

La muestra purificada y concentrada se purificó y concentró adicionalmente mediante purificación con CA (purificación por perlas superparamagnéticas conjugadas con ácido carboxílico) con un robot MBS (Magnetic Biosolutions). Se usó una concentración final de PEG del 10 % para eliminar fragmentos por debajo de 150-200 pb. Se dejó que el ADNc amplificado se uniera a las perlas de CA (Invitrogen) durante 10 min y luego se eluyó en 15 µl de Tris-HCl 10 mM, pH 8,5.

Segunda amplificación por PCR

Las PCR finales consistieron en tampón Ex Taq 1x (Takara), 200 µM de cada dNTP (Takara), A_cebador (MWG) 600 nM, B_cebador (MWG) 600 nM y polimerasa Ex Taq (Takara) 0,025 U/µl. Las muestras se calentaron hasta 95 °C durante 3 minutos y luego se amplificaron mediante la ejecución de 10 ciclos a 95 °C durante 30 segundos, 65 °C durante 1 minuto y 72 °C durante 3 minutos, seguido de una extensión final a 72 °C durante 10 minutos.

Limpieza de la segunda biblioteca

Después de la amplificación, las cuatro PCR (100µl) se mezclaron con tampón de unión a 500 µl (Qiagen) y se colocaron en una columna de purificación por PCR Qiaquick (Qiagen) y se centrifugaron durante 1 minuto a 17 900 x g para unir el ADNc amplificado a la membrana. Luego, la membrana se lavó con tampón de lavado (Qiagen) que contenía etanol y finalmente se eluyó en 50 µl de Tris-Cl 10 mM, pH 8,5.

La muestra purificada y concentrada se purificó y concentró adicionalmente mediante purificación con CA (purificación por perlas superparamagnéticas conjugadas con ácido carboxílico) con un robot MBS (Magnetic Biosolutions). Se usó una concentración final de PEG del 10 % para eliminar fragmentos por debajo de 150-200 pb. Se dejó que el ADNc amplificado se uniera a las perlas de CA (Invitrogen) durante 10 min y luego se eluyó en 15 µl de Tris-HCl 10 mM, pH 8,5.

Preparación de bibliotecas de secuenciación

indexado de bibliotecas

Las muestras amplificadas durante 20 ciclos se usaron además para preparar bibliotecas de secuenciación. Se preparó una mezcla maestra de PCR de índice para cada muestra y se colocó 23 µl en seis tubos de 0,2 ml. Se añadieron 2 µl del ADNc amplificado y purificado a cada una de las seis PCR como plantilla que fabricó las PCR que contenían una mezcla maestra de Phusion 1x (Fermentas), InPE1.0 500nM (Illumina), Índice 1-12 500 nM (Illumina) y InPE2.0 0,4 nM (Illumina). Las muestras se amplificaron en un termociclador durante 18 ciclos a 98 °C durante 30 segundos, 65 °C durante 30 segundos y 72 °C durante 1 minuto, seguido de una extensión final a 72 °C durante 5 minutos.

Limpieza de la biblioteca de secuenciación

Después de la amplificación, las muestras se purificaron y se concentraron mediante purificación con CA con un robot MBS. Se usó una concentración final de PEG de 7,8 % para eliminar fragmentos por debajo de 300-350 pb. Se dejó que el ADNc amplificado se uniera a las perlas de CA durante 10 min y luego se eluyó en 15 µl de Tris-HCl 10 mM, pH 8,5.

Se colocaron 10 µl de las muestras amplificadas y purificadas en un chip Caliper XT y se cortaron fragmentos entre 480 pb y 720 pb con Caliper XT (Caliper). Las muestras se analizaron con un bioanalizador Agilent para confirmar la presencia y el tamaño de las bibliotecas terminadas, se usó el kit de alta sensibilidad al ADN.

Secuenciación y análisis de datos

La secuenciación y bioinformática se llevaron a cabo de la misma manera que en el protocolo para matrices de Nimblegen de alta densidad orientadas 5' a 3' descritas en el Ejemplo 6. Sin embargo, en el análisis de datos, la lectura 1 no se usó en el mapeo de transcritos. Los transcritos de Olfr específicos podrían ordenarse mediante el uso de la herramienta de visualización Matlab (Figura 8).

Ejemplo 8

Transcriptómica espacial mediante el uso de una micromatriz de 41 etiquetas impresa internamente con sondas orientadas de 5' a 3' y tejido congelado fijado en formalina (FF-congelado) con permeabilización a través de Proteinasa K o microondas con escisión del sistema USER y amplificación a través de TdT

Preparación de la matriz

Las matrices internas se imprimieron como se describió anteriormente, pero con un patrón de 41 sondas de ID-etiquetas únicas con la misma composición que las sondas en la matriz de alta densidad orientada 5' a 3' en el Ejemplo 6 (Figura 9).

Todas las demás etapas se llevaron a cabo de la misma manera que en el protocolo descrito en el Ejemplo 6.

Ejemplo 9

Método alternativo para realizar la etapa de síntesis de ADNc

La síntesis de ADNc en chip como se describió anteriormente también puede combinarse con el cambio de plantilla para crear una segunda hebra mediante la adición de un cebador de cambio de plantilla a la reacción de síntesis de ADNc (Tabla 4). El segundo dominio de amplificación se introduce al acoplarlo a las bases terminales añadidas por la transcriptasa inversa en el extremo 3' de la primera hebra de ADNc, y prepara la síntesis de la segunda hebra. La biblioteca puede amplificarse fácilmente directamente después de la liberación del complejo bicatenario de la superficie de la matriz.

Ejemplo 10

Los siguientes experimentos demuestran cómo las moléculas de ADNc aseguradas (capturadas) pueden etiquetarse y detectarse en la superficie del sustrato objeto, por ejemplo, matriz.

Preparación de una micromatriz impresa interna con sondas orientadas de 5' a 3'

El oligonucleótido de captura de ARN (Tabla 2) se imprimió en portaobjetos de vidrio para funcionar como la sonda de captura. La sonda se sintetizó con un amino enlazador de extremo 5'-terminal con un espaciador C6. Todas las preguntas se sintetizaron por Sigma-Aldrich (St. Louis, MO, EE. UU.). La sonda de captura de ARN se suspendió a una concentración de 20 μ M en fosfato de sodio 150 mM, pH 8,5 y se mancharon mediante el uso de una pipeta sobre portaobjetos de micromatriz activados CodeLink™ (7,5 cm x 2,5 cm; Surmodics, Eden Prairie, MN, EE. UU.). Cada matriz se imprimió con 10 μ l de solución que contenía sonda de captura, y se dejó secar. Después de la impresión, se realizó el bloqueo de la superficie de acuerdo con las instrucciones del fabricante. Las sondas se imprimieron en 16 matrices en el portaobjetos. Las 16 submatrices se separaron durante las etapas de reacción mediante una máscara de 16 almohadillas (ArrayIT Corporation, Sunnyvale, CA, EE. UU.).

Preparación de tejido congelado fresco y seccionamiento en matrices de sondas de captura

El tejido cerebral de ratón fresco no fijado se recortó si fuera necesario y se congeló en isopentano frío a -40 °C y posteriormente se montó para seccionarse con un criostato a 10 μ m. Se aplicó un segmento de tejido sobre cada matriz de sonda.

Fijación de la sección de tejido mediante el uso de formalina

Se añadió 50 μ l de paraformaldehído al 4 % disuelto en PBS directamente a la matriz de sonda para cubrir la sección de tejido. El arreglo se incubó a temperatura ambiente durante 10 minutos y luego se lavó durante 10 segundos en PBS. A continuación, la matriz se incubó a 50 °C durante 15 minutos.

Permeabilización de la muestra de tejido mediante el uso de pepsina e HCl

La pepsina se diluyó al 0,1 % en HCl 0,1 M y se precalentó a 37 °C. La matriz se unió a una máscara de 16 pocillos y el soporte ArrayIt. Se añadieron 50 μ l de la mezcla de pepsina/HCl a cada pocillo. La matriz se incubó durante 10 minutos a 37 °C y los pocillos se lavaron con 100 μ l de SSC 0,1x mediante pipeteado.

Síntesis de ADNc con Cv3-dNTP

Para cada pocillo se preparó una mezcla de síntesis de ADNc (80 μ l) que contenía 4 μ l de cada uno de dATP/dGTP/dTTP (10 mM), 4 μ l de dCTP (2,5 mM), 4 μ l de Cy3-dCTP (1 mM), 4 μ l de DTT (0,1 M), 1x de BSA, 20 U/ μ l de Superscript III, 5 U/ μ l de ARNasaOUT, 1x de tampón de primera hebra (Superscript III, Invitrogen) y agua

MiliQ. Se añadieron 70 µl de la mezcla de reacción a cada pocillo. Las reacciones se recubrieron con un sellador plástico y se incubaron a 37 °C durante toda la noche.

Lavado

Después de la incubación, la matriz se retiró de la máscara y el soporte ArrayIt y se lavó mediante el uso de las siguientes etapas: 1) solución de 2xSSC a 50 °C con SDS al 0,1 % durante 10 min a agitación de 300 rpm; 2) 0,2xSSC durante 1 min a agitación de 300 rpm; y 3) 0,1xSSC durante 1 min a agitación de 300 rpm. La matriz se centrifugó en seco.

Eliminación de tejido

La matriz se unió a un soporte de portaobjetos y máscara de 16 pocillos ArrayIt (ArrayIt Corporation). Se añadió 10 µl de solución de proteinasa K (Qiagen) por cada 150 µl de tampón de digestión de proteinasa K del kit RNeasy FFPE (Qiagen). Se añadió 50 µl de la mezcla final a cada pocillo y la matriz se incubó a 56 °C durante 1 hora. La matriz se lavó como se describió anteriormente.

Obtención de imágenes

Se obtuvieron imágenes de la matriz a 532 nm mediante el uso de un escáner de micromatriz Agilent a una exposición del 100 % y una resolución de 5 µm (Figura 11 y Figura 12).

Tabla 2

Sonda 1

Amino-C6-UUACACTCTTTCCCTACACGACGCTCTTCCGATCT GTCCGATATGATTGCCGCTTTTTTTTTTTTTTTTTTTVN (SEQ ID NO: 47)

Ejemplo 11

Los siguientes experimentos demuestran cómo una porción de las moléculas de ADNc aseguradas (capturadas) puede eliminarse de partes seleccionadas de la superficie del sustrato objeto, por ejemplo, matriz. Esto permite el aislamiento del ADNc de partes limitadas de la sección de tejido.

El método descrito en el Ejemplo 10 se realizó hasta e incluyó la etapa de eliminación de tejido. Siga las etapas descritas más abajo.

Eliminación del ADNc mediante destrucción láser

La matriz se montó en un instrumento MMI Cellcut (Molecular Machines and Industries AG, Glattbrugg, Suiza) y las secciones del ADNc etiquetado con fluorescencia a eliminar se etiquetaron para destrucción mediante láser.

Obtención de imágenes

Se obtuvieron imágenes del chip de vidrio del enlace de código a 532 nm mediante el uso de un escáner de micromatriz Agilent a una exposición del 100 % y una resolución de 5 µm. Se verificó la eliminación de las áreas extirpadas (Figuras 13a y b). La matriz se lavó de acuerdo con el procedimiento descrito en el Ejemplo 10.

Liberación del ADNc remanente de la matriz

La matriz se unió a un soporte de portaobjetos ArrayIt y una máscara de 16 pocillos (ArrayIt Corporation). Se añadió a cada pocillo una mezcla de escisión (50 µl) que contenía tampón Exo I 1x (New England Biolabs, Ipswich, MA, EE. UU.), BSA 1x, agua libre de ARNasa/ADNasa, 5 U de mezcla de enzima USER (New England Biolabs). Las reacciones se recubrieron con un sellador plástico y se incubaron a 37 °C durante 1 hora mediante el uso de un intervalo de mezcla de 3 segundos a 300 rpm y 6 segundos de reposo. Después de la incubación, se recogió una mezcla de escisión de 45 µl de cada uno de los pocillos usados y se colocó en tubos de PCR de 0,2 ml.

Tratamiento con exonucleasa

Después de enfriar las soluciones en hielo durante 2 minutos, se añadió Exonucleasa I (NEB), para retirar las sondas de ADNc no extendidas, hasta un volumen final de 46,2 µl y una concentración final de 0,52 U/µl. Los tubos se incubaron en un termociclador (Applied Biosystems) a 37 °C durante 30 minutos seguido de la inactivación de la exonucleasa a 80 °C durante 25 minutos.

Cola de dA por transferasa terminal

Después de la etapa de exonucleasa, se añadió a cada una de las muestras una mezcla de 45 µl de cola poliA que consiste en tampón TdT (Takara), dATP 3 mM (Takara) y la mezcla enzimática de TdT (TdT y ARNasa H) (Takara) de acuerdo con las instrucciones de los fabricantes. Las mezclas se incubaron en un termociclador a 37 °C durante 15 minutos seguido de la inactivación de TdT a 70 °C durante 10 minutos.

Síntesis de segunda hebra y amplificación por PCR

Después de la cola dA, se colocó una mezcla maestra de PCR de 23 µl en cuatro tubos de PCR nuevos de 0,2 ml por muestra. Se añadieron 2 µl de muestra a cada tubo como plantilla. Las PCR finales consistieron en tampón Ex Taq 1x (Takara), 200 µM de cada dNTP (Takara), 600 nM A_cebador (MWG), 600 nM B_dt20VN_cebador (MWG) y 0,025 U/µl Ex Taq polimerasa (Takara) (Tabla 3). Se creó una segunda hebra de ADNc al ejecutar un ciclo en un termociclador a 95 °C durante 3 minutos, 50 °C durante 2 minutos y 72 °C durante 3 minutos. Las muestras se amplificaron al ejecutar 20 ciclos (para la preparación de bibliotecas) o 30 ciclos (para confirmar la presencia de ADNc) a 95 °C durante 30 segundos, 67 °C durante 1 minuto y 72 °C durante 3 minutos, seguido de una extensión final a 72 °C durante 10 minutos.

Limpieza de bibliotecas

Después de la amplificación, las cuatro PCR (100 µl) se mezclaron con tampón de unión a 500 µl (Qiagen) y se colocaron en una columna de purificación por PCR Qiaquick (Qiagen) y se centrifugaron durante 1 minuto a 17 900 x g para unir el ADNc amplificado a la membrana. Luego, la membrana se lavó con tampón de lavado (Qiagen) que contenía etanol y finalmente se eluyó en 50 µl de Tris-Cl 10 mM, pH 8,5.

La muestra purificada y concentrada se purificó y concentró adicionalmente mediante purificación con CA (purificación por perlas superparamagnéticas conjugadas con ácido carboxílico) con un robot MBS (Magnetic Biosolutions). Se usó una concentración final de PEG del 10 % para eliminar fragmentos por debajo de 150-200 pb. Se dejó que el ADNc amplificado se uniera a las perlas de CA (Invitrogen) durante 10 min y luego se eluyó en 15 µl de Tris-Cl 10mM, pH 8,5.

Análisis de calidad de la biblioteca

Las muestras amplificadas durante 30 ciclos se analizaron con un bioanalizador Agilent (Agilent) para confirmar la presencia de una biblioteca de ADNc amplificada, el kit de ADN de alta sensibilidad o el kit de ADN 1000 se usaron en dependencia de la cantidad de material (Figura 14).

Tabla 3

Síntesis de la segunda hebra y primeros mangos de amplificación por PCR	
A_cebador	ACACTCTTTCCCTACACGACGCTCTTCCGATCT (SEQ ID NO: 48)
B_dt20VN_cebador	AGACGTGTGCTCTTCCGATCTTTTTTTTTTTTTTTTTTTTTTTVN (SEQ ID NO: 49)

	Ejemplo 6	
	Matrices de 5' a 3' de Nimblegen con sondas de matriz de extremo 3' libres	
	5' a 3'	
Sonda1 (SEQ ID NO: 50)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTGTGCCGATGATGATGCGCGCTTTTTTTTTTTTTTTTTTTT	
Sonda2 (SEQ ID NO: 51)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTATGAGCCGGGTCATCTTTTTTTTTTTTTTTTTTTTTT	
Sonda3 (SEQ ID NO: 52)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTTGAGGCACCTCTGTGGGAUUTTTTTTTTTTTTTTTTTT	
Sonda4 (SEQ ID NO: 53)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTATGATTAGTCGCCATTGCTTTTTTTTTTTTTTTTTTTT	
Sonda5 (SEQ ID NO: 54)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTACTIGAGGGTAGATGTTTTTTTTTTTTTTTTTTTTTTT	
Sonda6 (SEQ ID NO: 55)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTAIGGCCAACTACTGTTAICTTTTTTTTTTTTTTTTTTTT	
Sonda7 (SEQ ID NO: 56)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTCGCTACCCTGATTCGACCTTTTTTTTTTTTTTTTTTTT	
Sonda8 (SEQ ID NO: 57)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTGCCCACTTCGCCGTAGTTTTTTTTTTTTTTTTTTTTT	
Sonda9 (SEQ ID NO: 58)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTAGCAACTTTGAGCAAGATTTTTTTTTTTTTTTTTTTTTT	
Sonda10 (SEQ ID NO: 59)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTGCCAATTCGGAAATCCGGTTTTTTTTTTTTTTTTTTTTT	
Sonda11 (SEQ ID NO: 60)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTTCGCCCAAGGTAATACATTTTTTTTTTTTTTTTTTTTTT	
Sonda12 (SEQ ID NO: 61)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTTCGCAATTCCTATTGAGTTTTTTTTTTTTTTTTTTTTT	
Sonda13 (SEQ ID NO: 62)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTTGCTAAATCTAACCGCCTTTTTTTTTTTTTTTTTTTTTT	
Sonda14 (SEQ ID NO: 63)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTGGAAATAATTCTGATGGTTTTTTTTTTTTTTTTTTTTTT	
Sonda15 (SEQ ID NO: 64)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTCATATACATAGGTGCTAAAGTTTTTTTTTTTTTTTTTTTTT	
Sonda16 (SEQ ID NO: 65)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTATGACTTGGCTCGCACCTTTTTTTTTTTTTTTTTTTTTT	
Sonda17 (SEQ ID NO: 66)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTATAGTATCTCCCAAGTCTTTTTTTTTTTTTTTTTTTTTT	
Sonda18 (SEQ ID NO: 67)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTGTGGGCTGTAAATCCGCATTTTTTTTTTTTTTTTTTTTTT	
Sonda19 (SEQ ID NO: 68)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTGGCCCACTCTTAGGTAGTTTTTTTTTTTTTTTTTTTTTT	
Sonda20 (SEQ ID NO: 69)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTATGCAAGTGAATGGCTTTTTTTTTTTTTTTTTTTTTTT	
Sonda21 (SEQ ID NO: 70)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTCCAAGCCAGTTTATAGTTTTTTTTTTTTTTTTTTTTTTT	
Sonda22 (SEQ ID NO: 71)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTACCTGATTCGTGTAATACTTTTTTTTTTTTTTTTTTTTTT	
Sonda23 (SEQ ID NO: 72)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTCAGCGCACTATCCTCTATTTTTTTTTTTTTTTTTTTTTTT	
Sonda24 (SEQ ID NO: 73)	UUUUUACACTCTTTCCCTACACGAGCGTCTTCGGATCTCCACCGGTAGGACTAGTTTTTTTTTTTTTTTTTTTTTTTT	

Sonda25 (SEQ ID NO: 74)	UUUUUACACTCTTTCCCTACACGACGCTCTTCCGATCTCGACTAAGTATGTAGCGCTTTTCTTTTCTTTT
Sonda de marto	
Capa1 (SEQ ID NO: 75)	AAATTTCGTCTGCTATCGGCTCTCTGTGTAAC
Sonda marcadora fluorescente	
PS_1 (SEQ ID NO: 76)	GGTACAGAGGCGGATAGCAG - Cy3
Síntesis de la segunda hebra y primeros mangos de amplificación por PCR	
A_cebador (SEQ ID NO: 77)	ACACTCTTCCCTACACGACGCTCTTCCGATCT
D_dt20VN_cebador (SEQ ID NO: 78)	AGACGTGTGCTCTTCCGATCTTTTTTTTTTTTTTTT
Cebador de secuenciación personalizado	
B_r2 (SEQ ID NO: 79)	TCAGACGCTGCTCTTCCGATCTTTTTTTTTTTTTTTT
Ejemplo 7	
Matrices de 3' a 5' de Nimblegen con sondas de extremo 5' libres	
b' a 3'	
Sonda1 (SEQ ID NO: 80)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda2 (SEQ ID NO: 81)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda3 (SEQ ID NO: 82)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda4 (SEQ ID NO: 83)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda5 (SEQ ID NO: 84)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda6 (SEQ ID NO: 85)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda7 (SEQ ID NO: 86)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda8 (SEQ ID NO: 87)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda9 (SEQ ID NO: 88)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda10 (SEQ ID NO: 89)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda11 (SEQ ID NO: 90)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda12 (SEQ ID NO: 91)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda13 (SEQ ID NO: 92)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT
Sonda14 (SEQ ID NO: 93)	GCCTTCAGAGTGGCAGTGGAGATCAGAGTGGGAAATCAATCGGACAGATCGGAGAGAGGAGGAGT

5
10
15
20
25
30
35
40
45
50
55
60
65

Sonda15 (SEQ ID NO: 94)	GCGTTCAGAGTGGCAGTCGAGATCACCTTAGCACCCTAIGTAATGAGATCGGAAGAGCGTAGTGTAG
Sonda16 (SEQ ID NO: 95)	GCGTTCAGAGTGGCAGTCGAGATCACCTTAGCAGCGCAAGTCAATAGATCGGAAGAGCGTAGTGTAG
Sonda17 (SEQ ID NO: 96)	GCGTTCAGAGTGGCAGTCGAGATCACGAACCTTGGGAGATACTATAGATCGGAAGAGCGTAGTGTAG
Sonda18 (SEQ ID NO: 97)	GCGTTCAGAGTGGCAGTCGAGATCACGCGGATACAGGCGGCACAGATCGGAAGAGCGTAGTGTAG
Sonda19 (SEQ ID NO: 98)	GCGTTCAGAGTGGCAGTCGAGATCACGCTAUCCTAAGAGTGGGAGATCGGAAGAGCGTAGTGTAG
Sonda20 (SEQ ID NO: 99)	GCGTTCAGAGTGGCAGTCGAGATCACAGCCCAATCACTTCCATAAGATCGGAAGAGCGTAGTGTAG
Sonda21 (SEQ ID NO: 100)	GCGTTCAGAGTGGCAGTCGAGATCACCCCTATAAACCGTCCCTCCACATCGGAAGAGCGTAGTGTAG
Sonda22 (SEQ ID NO: 101)	GCGTTCAGAGTGGCAGTCGAGATCACGTTATACAGCAATCAGGTAGATCGGAAGAGCGTAGTGTAG
Sonda23 (SEQ ID NO: 102)	GCGTTCAGAGTGGCAGTCGAGATCACAGGATAGATCGGCTGAGATCGGAAGAGCGTAGTGTAG
Sonda24 (SEQ ID NO: 103)	GCGTTCAGAGTGGCAGTCGAGATCACACTAGTCTCAGCGTGGAAAGATCGGAAGAGCGTAGTGTAG
Sonda25 (SEQ ID NO: 104)	GCGTTCAGAGTGGCAGTCGAGATCACGGGCTACATACTTAGTCCGAGATCGGAAGAGCGTAGTGTAG
Sonda de marco	
Capa1 (SEQ ID NO: 105)	AAATTTCCTCTCTATCCCGCTCTCTATCC
Sonda de captura	
LP_Poli-dTVN (SEQ ID NO: 106)	GTGATCTCGACTGCACTCTGMATTTTTTTTTTTTTTTTVN
Sonda del mango de amplificación	
Mango A (SEQ ID NO: 107)	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Síntesis de la secuencia hebra y primeros mangos de amplificación por PCR	
A_cebador (SEQ ID NO: 108)	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
R_d20VN_diluido (SEQ ID NO: 109)	AGACGTGCTCTTCCGATCTTTTTTTTTTTTTTTTTVN
Segunda PCR.	
A_cebador (SEQ ID NO: 110)	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
B_cebador (SEQ ID NO: 111)	GTGACTGGAGTTCAGAGGTGTGCTCTTCCGATCT
Ejemplo 9	
Cambio de plantilla	
Cambio de plantilla_largoB (SEQ ID NO: 112)	GTGACTGGAGTTCAGAGGTGTGCTCTTCCGATCTATCTCTCT

Ejemplo 12

El siguiente experimento demuestra que la etapa de etiquetar las moléculas de ADNc aseguradas (capturadas) puede realizarse mediante el uso de matrices que comprenden una alta densidad de sondas de captura que comprenden dominios posicionales únicos.

Preparación de la matriz

Las micropartículas de alta densidad prefabricadas se pidieron de Roche- Nimblegen (Madison, WI, EE. UU.). Cada chip contenía múltiples matrices de sondas, cada una con 270 000 elementos de los cuales 135 000 elementos llevaban una sonda de captura que comprende una secuencia de ID-etiqueta única (dominio posicional) y una región de captura (dominio de captura). Cada elemento tenía un tamaño de 13x13 μm . Las sondas de captura se componen de 5' a 3' de: un dominio universal que contiene cinco bases dUTP (un dominio de escisión) y un dominio de amplificación general 5'; una etiqueta de ID (dominio posicional); y una región de captura (dominio de captura) (véase Tabla 5). A cada matriz también se le equipó con un marco de sondas marcadoras que portaban una secuencia genérica de 30 pb (Tabla 5) para permitir la hibridación de sondas fluorescentes para ayudar con la orientación durante la visualización de la matriz.

Las matrices de sondas se separaron durante las etapas de reacción mediante una máscara de 16 almohadillas (Arrayit Corporation, Sunnyvale, CA, EE. UU.).

El método descrito en el Ejemplo 10 se realizó hasta e incluyó la etapa de eliminación de tejido.

Obtención de imágenes

Se preparó una solución con sonda marcadora de marco (Tabla 5) a una concentración de 170 nM en PBS. Esta solución se añadió a los pocillos y el portaobjetos se incubó a temperatura ambiente durante 5 minutos, seguido de un breve lavado en PBS y secado por centrifugación.

Se obtuvieron imágenes del chip de vidrio de la micromatriz de alta densidad a 532 nm mediante el uso de un escáner de micromatriz Agilent a una exposición del 100 % y una resolución de 5 μm (Figura 15).

5
10
15
20
25
30
35
40
45
50
55
60
65

Tabla 5 matrices de E' a 3' de Nimblegen con sondas de matriz de extremo 3' libres	
	5' a 3'
Sonda1 (SEQ ID NO: 113)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTGTCCGATATGATTGCCGCTTTT
Sonda2 (SEQ ID NO: 114)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTATGAGCCGGTTCATCTTTT
Sonda3 (SEQ ID NO: 115)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTGTGAGGCACCTGTGTGGGATTT
Sonda4 (SEQ ID NO: 116)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTATGATTAGTCCGCAATCCGTTT
Sonda5 (SEQ ID NO: 117)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTACTTGAGGGTAGATGTTT
Sonda6 (SEQ ID NO: 118)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTATGCCCAATACGTTAAGTTT
Sonda7 (SEQ ID NO: 119)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTCCGTACCTGATTCGACCTTT
Sonda8 (SEQ ID NO: 120)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTGCCGACCTTCCGCGTAGTTT
Sonda9 (SEQ ID NO: 121)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTAGCAACTTTGAGCAAGATTT
Sonda10 (SEQ ID NO: 122)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTGCCAATTCGGAATCCGGTTT
Sonda11 (SEQ ID NO: 123)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTTCCGCCAAGGTAATACATTT
Sonda12 (SEQ ID NO: 124)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTTCCGATTCCTATTCGAGTTT
Sonda13 (SEQ ID NO: 125)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTTTGCTMACTACCCGCTTTT
Sonda14 (SEQ ID NO: 126)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTCGGAATTAATTCATGTTT
Sonda15 (SEQ ID NO: 127)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTCATTACATAGGTCCTAAGTTT
Sonda16 (SEQ ID NO: 128)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTATGACTTGGCTCGGACCTTT
Sonda17 (SEQ ID NO: 129)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTATAGTATCTCCCAAGTCTTT
Sonda18 (SEQ ID NO: 130)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTGTGGGCCGTGTAATCCGCAATTT
Sonda19 (SEQ ID NO: 131)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTCCGCCACTCTTTAGGTAGTTT
Sonda20 (SEQ ID NO: 132)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTTATGCCAAGTATGCGCTTTT
Sonda21 (SEQ ID NO: 133)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTCCAAGCCACGTTATACGTTT
Sonda22 (SEQ ID NO: 134)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTACCTGATTCCTATAACTTTT
Sonda23 (SEQ ID NO: 135)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTCAGGCGCATCTATCTCTATTT
Sonda24 (SEQ ID NO: 136)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTTCCAGCGGTAGGACTAGTTT
Sonda25 (SEQ ID NO: 137)	UUUUUACACTCTTCCCTACACGACGGCTCTCCGATCTTCCGACIAAGATTTGAGCGCTTT
Sonda de marca (SEQ ID NO: 138)	AAATTTGGTCTGCTATCGGGCTTCTGTACC
Sonda marcadora fluorescente (SEQ ID NO: 138)	CGTACACAAGCCCGCATAGCCAG - Dy0

Ejemplo 13

El siguiente experimento demuestra que la etapa de etiquetar las moléculas de ADNc aseguradas (capturadas) es efectiva a concentraciones menores de nucleótidos etiquetados con fluorescencia.

El método descrito en el Ejemplo 10 se realizó hasta la etapa de síntesis de ADNc, que se reemplazó por la etapa descrita más abajo.

Síntesis de ADNc con Cv3-dNTP

Para cada pocillo se preparó una mezcla de síntesis de ADNc de 80 μ l que contenía 4 μ l de cada uno de dATP/dGTP/dTTP (10 mM), 4 μ l de DTT (0,1 M), 1x de BSA, 20U/ μ l de Superscript III, 5U/ μ l de ARNasaOUT, 1x de tampón de la primera hebra (Superscript III, Invitrogen) y agua MilliQ. La concentración de dCTP y Cy3-dCTP se varió de manera que en cinco experimentos paralelos se usaron 1/1, 1/2, 1/4, 1/8 y 1/16 de la concentración original de 2,5 mM para dCTP y 1 mM para Cy3-dCTP. Se pipetearon 4 μ l de dCTP y 4 μ l de Cy3-dCTP con estas concentraciones en cada mezcla de síntesis respectiva. Se añadieron 70 μ l de la mezcla de reacción a cada pocillo. Las reacciones se recubrieron con un sellador plástico y se incubaron a 37 °C durante toda la noche.

Las etapas de lavado, eliminación de tejido e imágenes se realizaron de acuerdo con el Ejemplo 10.

Cuantificación de la señal

Las intensidades de señal resultantes de la obtención de imágenes del chip de vidrio del enlace de código en el escáner de micromatriz Agilent se analizaron mediante el uso del software Genepix pro. Las intensidades promedio de la señal se calcularon a partir de múltiples áreas seleccionadas dentro de cada huella Cy3 y se representaron en un diagrama (Figura 16).

Ejemplo 14

El siguiente experimento demuestra que puede usarse una variedad de tejidos en los métodos de la invención.

El método descrito en el Ejemplo 10 se realizó hasta la etapa de obtención de imágenes. Las etapas de preparación, fijación y obtención de imágenes del tejido se reemplazaron por las etapas descritas más abajo.

Preparación de tejido congelado fresco y seccionamiento en matrices de sondas de captura

El tejido fresco de Drosophila o pez cebra no fijado se congeló con hielo seco y subsecuentemente se montó para seccionarse con un criostato a 10 μ m. Se aplicó un segmento de tejido sobre cada conjunto de sondas que se va a usar.

Fijación de la sección de tejido mediante el uso de metanol

Las secciones se fijaron en un baño de metanol preenfriado a -20 °C durante 10 min. Después de la fijación, el portaobjetos se lavó brevemente durante 10 segundos 1xPBS y se secó por calor a 50 °C durante 15 min en un termomezclador eppendorf.

Obtención de imágenes

Se obtuvieron imágenes del chip de la matriz de sondas con tejido de pez cebra (Figura 17A) o Drosophila (Figura 17B) antes de retirarlo mediante el uso de obtención de imágenes de contraste de fase. Las imágenes después de retirar el tejido usan el canal compatible con Cy3 en un instrumento MMI Cellcut montado en un microscopio Olympus IX 81.

Ejemplo 15

El siguiente experimento demuestra además que puede usarse una variedad de tejidos en los métodos de la invención.

El método descrito en el Ejemplo 10 se realizó hasta la etapa de obtención de imágenes. Las etapas de preparación, fijación, permeabilización, eliminación de tejido y obtención de imágenes del tejido se reemplazaron por las etapas descritas a continuación.

Preparación de tejido congelado fresco y seccionamiento en matrices de sondas de captura

El tejido del cáncer de próstata fresco se recortó si fuera necesario, se incrustó en Neg- 50 (Thermo Scientific) y se congeló a presión en nitrógeno líquido. El tejido se montó subsecuentemente para seccionarse con un criostato a 10 μ m. Se aplicó un segmento de tejido sobre cada conjunto de sondas que se va a usar.

Fijación de la sección de tejido mediante el uso de formalina

El chip se unió a un soporte y máscara de 16 pocillos Arraylt. Se añadió 70 µl de paraformaldehído al 4 % disuelto en PBS al conjunto de sondas para cubrir la sección de tejido. La reacción se incubó a temperatura ambiente durante 10 minutos. La máscara se retiró y el chip se lavó durante 10 segundos en PBS. Luego, el chip se incubó a 50 °C durante 15 minutos.

Permeabilización mediante el uso de pepsina e HCl

La pepsina se diluyó al 0,1 % en HCl 0,1 M y se precalentó a 37 °C. El chip se unió a un soporte y máscara de 16 pocillos Arraylt. Se añadieron 70 µl de la mezcla de pepsina/HCl a cada pocillo. La reacción se incubó durante 10 minutos a 37 °C. Los pocillos se lavaron con 100 µl de SSC 0,1x mediante pipeteado.

Eliminación de tejido

La matriz se unió a un soporte de portaobjetos y máscara de 16 pocillos Arraylt (Arraylt Corporation). Se añadió 20 µl de solución de proteinasa K (Qiagen) por cada 150 µl de tampón de digestión de proteinasa K del kit RNeasy FFPE (Qiagen). Se añadió 50 µl de la mezcla final a cada pocillo y la matriz se incubó a 56 °C durante 1 hora. La matriz se lavó como se describió en el Ejemplo 10.

Obtención de imágenes

Se obtuvieron imágenes del chip de la matriz de sondas con tejido prostático antes de la extracción mediante el uso de imágenes de contraste de fase. Las imágenes después de la eliminación de tejido usan el canal compatible con Cy3 en un instrumento MMI Cellcut montado en un microscopio Olympus IX 81 (Figura 17C).

Ejemplo 16

El siguiente experimento demuestra que una muestra de células, es decir, una suspensión de células, puede usarse como la muestra de tejido en los métodos de la invención.

El método descrito en el Ejemplo 10 se realizó hasta la etapa de obtención de imágenes. Las etapas de preparación, fijación, permeabilización de tejidos, síntesis de ADNc, eliminación de tejidos y obtención de imágenes se reemplazaron por las etapas descritas más abajo.

Preparación de tejidos: Aplicación de células en matrices de sondas de captura

Aproximadamente 1000-2000 células de fibroblastos de ratón (línea celular (NIH 3T3)) en un volumen de 5 µl (en SSC 0,1x) se pipetearon sobre el conjunto de sondas y se distribuyeron mediante el uso de la punta de pipeta. A continuación, el chip se incubó a 37 °C durante 5,5 minutos.

Fijación de células mediante el uso de formalina

El chip se unió a un soporte y máscara de 16 pocillos Arraylt. Se añadió 100 µl de paraformaldehído al 4 % disuelto en PBS al pocillo de la matriz de sondas para cubrir las células. La reacción se incubó a temperatura ambiente durante 10 minutos. El pocillo de la matriz de sondas se lavó una vez con 100µl 0,1x SSC mediante pipeteado. Luego, el chip se incubó a 50 °C durante 15 minutos.

Permeabilización mediante el uso de pepsina e HCl

La pepsina se diluyó al 0,1 % en HCl 0,1 M y se precalentó a 37 °C. El chip se unió a un soporte y máscara de 16 pocillos Arraylt. Se añadieron 70 µl de la mezcla de pepsina/HCl a cada pocillo. La reacción se incubó durante 1 minuto a 37 °C. Los pocillos se lavaron con 100 µl de SSC 0,1x mediante pipeteado.

Síntesis de ADNc con Cy3-dNTP

Para cada pocillo se preparó una mezcla de síntesis de ADNc de 80 µl que contenía 4 µl de cada uno de dATP/dGTP/dTTP (10 mM), 2 µl de dCTP (2,5 mM), 2 µl de Cy3-dCTP (1 mM), 4 µl de DTT (0,1 M), 1x de BSA, 20U/µl de Superscript III, 5U/µl de ARNasaOUT, 1x de tampón de primera hebra (Superscript III, Invitrogen) y agua MiliQ. Se añadió 70 µl de la mezcla de reacción a cada pocillo. Las reacciones se recubrieron con un sellador plástico y se incubaron a 37 °C durante toda la noche.

Obtención de imágenes

Se obtuvieron imágenes del chip de matriz de sondas con células antes de su eliminación mediante el uso de imágenes de contraste de fase. Las imágenes después de la eliminación de tejido usan el canal compatible con Cy3 en un instrumento MMI Cellcut montado en un microscopio Olympus IX 81 (Figura 17D).

5

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Un método para la detección y/o análisis espacial del ARNm y la proteína en una muestra de tejido o una porción de este, que comprende:
 - a) proporcionar un sustrato objeto en el que al menos una especie de sonda de captura, en donde la al menos una especie de sonda de captura comprende un dominio de captura, en donde el dominio de captura se diseña para la captura de ARNm, se inmoviliza directa o indirectamente de manera que la al menos una especie de sonda de captura se orienta para tener un extremo 3' libre para permitir que la al menos una especie de sonda de captura funcione como un cebador de transcriptasa inversa (RT);
 - b) poner en contacto el sustrato objeto con la muestra de tejido y permitir que el ARNm de la muestra de tejido se hibride con la al menos una especie de sonda de captura, en donde la muestra de tejido es una sección de tejido fijada en formalina e incluida en parafina;
 - c) generar molécula(s) de ADNc a partir de la(s) molécula(s) de ARNm capturada(s) mediante el uso de la al menos una especie de sonda de captura como cebador de RT;
 - d) etiquetar la(s) molécula(s) de ADNc generada(s) en la etapa (c), en donde la etapa de etiquetado es simultánea a la etapa de generación, en donde la etiqueta no es la sonda de captura o una parte de esta;
 - e) después de la etapa d) poner en contacto la muestra de tejido con uno o más anticuerpos primarios que se unen a una proteína diana de la muestra de tejido; y subsecuentemente poner en contacto la muestra de tejido con uno o más anticuerpos secundarios coincidentes conjugados a un fluoróforo; y
 - f) detectar: (i) una señal a partir de la(s) molécula(s) de ADNc etiquetada(s) y (ii) una señal a partir del uno o más anticuerpos etiquetados.
2. El método de la reivindicación 1, en donde el método comprende además obtener imágenes la muestra de tejido.
3. El método de la reivindicación 2, en donde la obtención de imágenes da como resultado la detección del uno o más anticuerpos secundarios en la muestra de tejido y/o da como resultado la detección de las moléculas de ADNc etiquetadas.
4. El método de la reivindicación 1, en donde el dominio de captura que se diseña para la captura selectiva de ARNm comprende una secuencia poli-T, una secuencia poli-U, una secuencia aleatoria o una secuencia redundante.
5. El método de la reivindicación 1, en donde la al menos una especie de sonda de captura comprende además un dominio universal que es 5' con relación al dominio de captura y/o un dominio posicional que es 5' con relación al dominio de captura, en donde el dominio universal comprende: (i) un dominio de amplificación, para amplificar la(s) molécula(s) de ADNc generada(s) y/o (ii) un dominio de escisión para liberar la(s) molécula(s) de ADNc generada(s) del sustrato objeto.
6. El método de la reivindicación 1, en donde el sustrato objeto comprende perlas, en donde la al menos una especie de sonda de captura se inmoviliza en las perlas.
7. Un método para la detección y/o análisis espacial del ARNm en una muestra de tejido o una porción de este, que comprende:
 - a) proporcionar un sustrato objeto en el que al menos una especie de sonda de captura, en donde la al menos una especie de sonda de captura comprende un dominio de captura, en donde el dominio de captura se diseña para la captura selectiva de ARNm, se inmoviliza directa o indirectamente de manera que la al menos una especie de sonda de captura se orienta para tener un extremo 3' libre para permitir que la al menos una especie de sonda de captura funcione como un cebador de RT;
 - b) poner en contacto el sustrato objeto con la muestra de tejido y permitir que el ARNm de la muestra de tejido se hibride con la al menos una especie de sonda de captura;
 - c) generar la primera hebra(s) de molécula(s) de ADNc a partir de la(s) molécula(s) de ARNm capturada(s) mediante el uso de la al menos una especie de sonda de captura como cebador de RT;
 - d) etiquetar la(s) primera(s) hebra(s) de molécula(s) de ADNc, en donde la etapa de etiquetado es simultánea a la etapa de generación (c), en donde la etiqueta no es la sonda de captura o una parte de esta;
 - e) obtener imágenes de la muestra de tejido para detectar una señal de la(s) primera(s) hebra(s) etiquetada(s) de molécula(s) de ADNc;
 - f) generar la(s) segunda(s) hebra(s) complementaria(s) a la primera hebra(s) de molécula(s) de ADNc mediante el uso de la(s) primera(s) hebra(s) como plantilla y/o un cebador de cambio de plantilla;
 - g) liberar la(s) segunda(s) hebra(s) del sustrato objeto; y
 - h) analizar la secuencia de la(s) segunda(s) hebra(s) liberada(s).
8. El método de la reivindicación 7, que comprende además amplificar la(s) segunda(s) hebra(s) antes de la etapa (h).
9. El método de la reivindicación 7 u 8, en donde el método comprende además (i) correlacionar la secuencia de la(s) segunda(s) hebra(s) liberada(s) de la superficie del sustrato objeto con una imagen de la muestra de tejido.

10. El método de una cualquiera de las reivindicaciones 7-9, en donde el dominio de captura que se diseña para la captura selectiva de ARNm comprende una secuencia poli-T, una secuencia poli-U, una secuencia aleatoria o una secuencia redundante.

11. El método de una cualquiera de las reivindicaciones 7-10, en donde la al menos una sonda de captura comprende además un dominio universal que es 5' con relación al dominio de captura y/o un dominio posicional que es 5' con relación al dominio de captura, en donde el dominio universal comprende: (i) un dominio de amplificación, para amplificar la(s) primera(s) hebra(s) generada(s) de molécula(s) de ADNc y/o (ii) un dominio de escisión para liberar la(s) primera(s) hebra(s) generada(s) de moléculas de ADNc del sustrato objeto.

12. El método de una cualquiera de las reivindicaciones 7-11, en donde:
(i) la muestra de tejido es una sección de tejido o una suspensión celular, y/o
(ii) el sustrato objeto comprende perlas, en donde las sondas de captura se inmovilizan en las perlas.

Figura 1

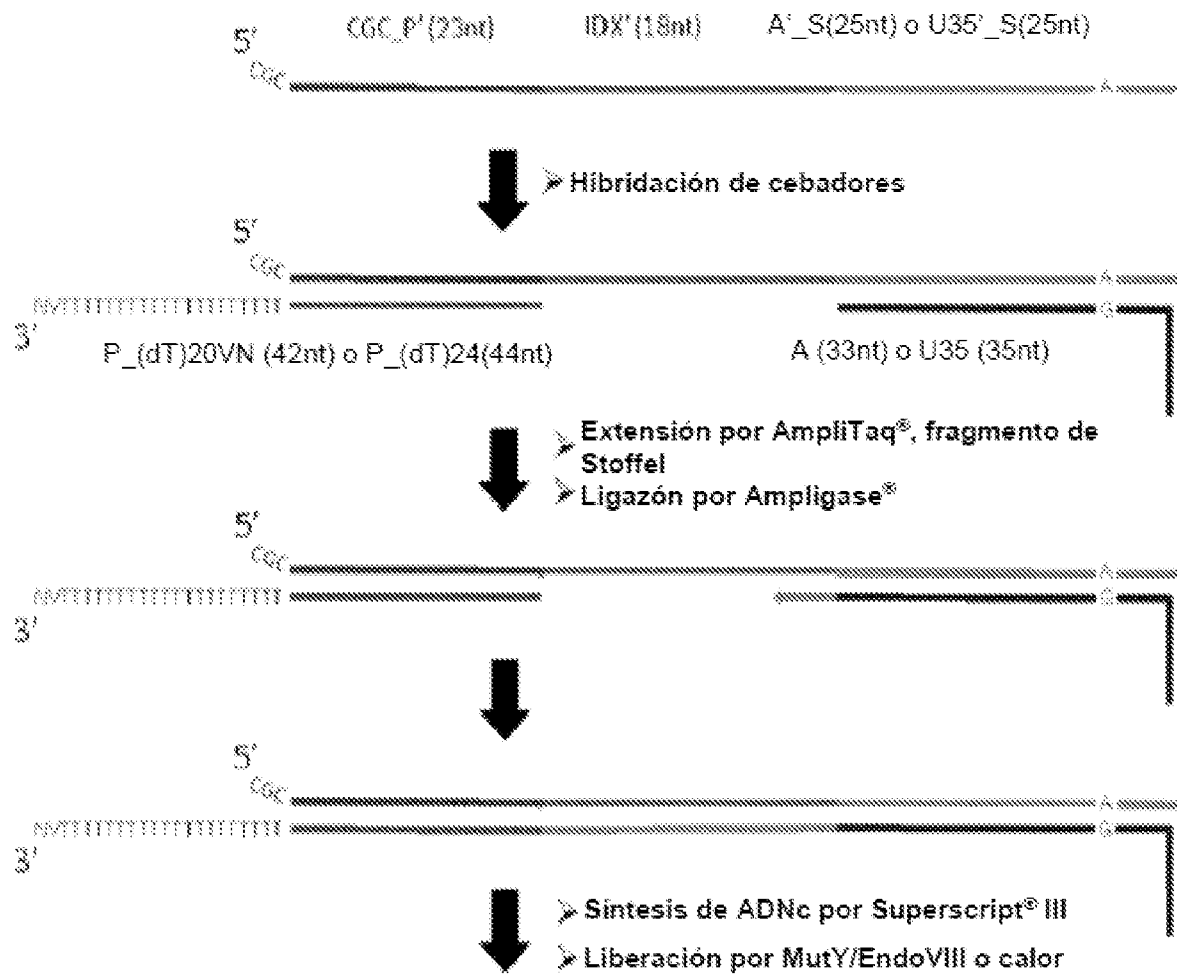
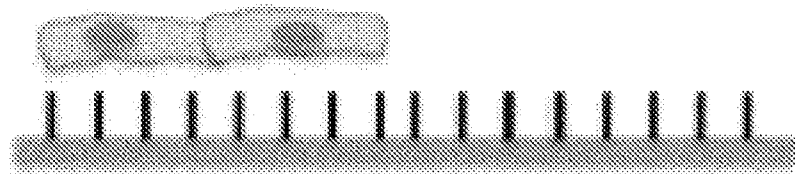
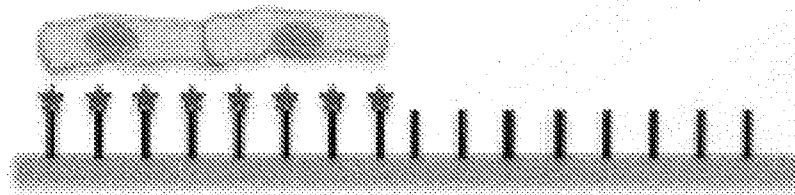


Figura 2

A Tejido con ARN de difusión



B Tejido con ARN de difusión



C Tejido con ARN de difusión

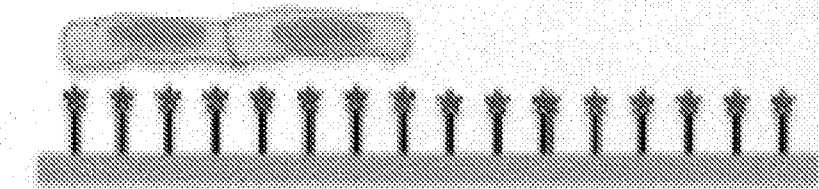


Figura 3

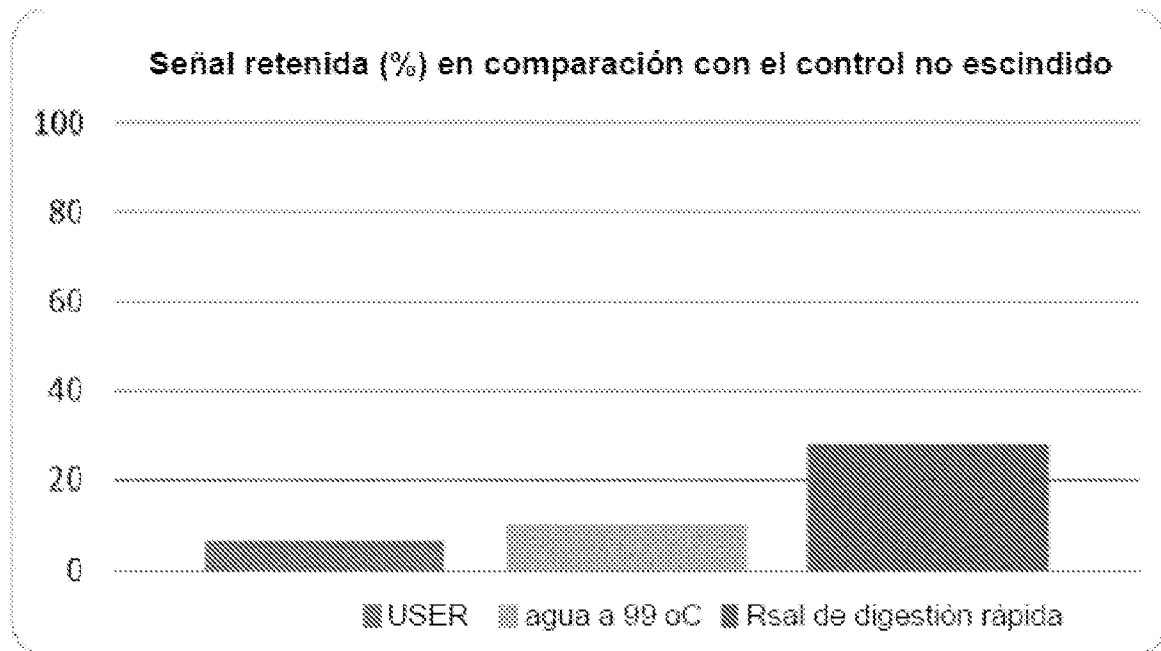


Figura 4

Camí 7 Resultados de blast																					
Valor E	1,00E-06																				
Código de barras	Aciertos	Etiqueta1	Etiqueta2	Etiqueta3	Etiqueta4	Etiqueta5	Etiqueta6	Etiqueta7	Etiqueta8	Etiqueta9	Etiqueta10	Etiqueta11	Etiqueta12	Etiqueta13	Etiqueta14	Etiqueta15	Etiqueta16	Etiqueta17	Etiqueta18	Etiqueta19	Etiqueta20
Etiqueta1	2165																				
Etiqueta2	1223																				
Etiqueta3	1338																				
Etiqueta4	3605																				
Etiqueta5	763																				
Etiqueta6	1083																				
Etiqueta7	1175																				
Etiqueta8	1350																				
Etiqueta9	616																				
Etiqueta10	710																				
Etiqueta11	1602																				
Etiqueta12	1359																				
Etiqueta13	1404																				
Etiqueta14	1124																				
Etiqueta15	1077																				
Etiqueta16	1238																				
Etiqueta17	1122																				
Etiqueta18	1647																				
Etiqueta19	1401																				
Etiqueta20	1354																				

Figura 5

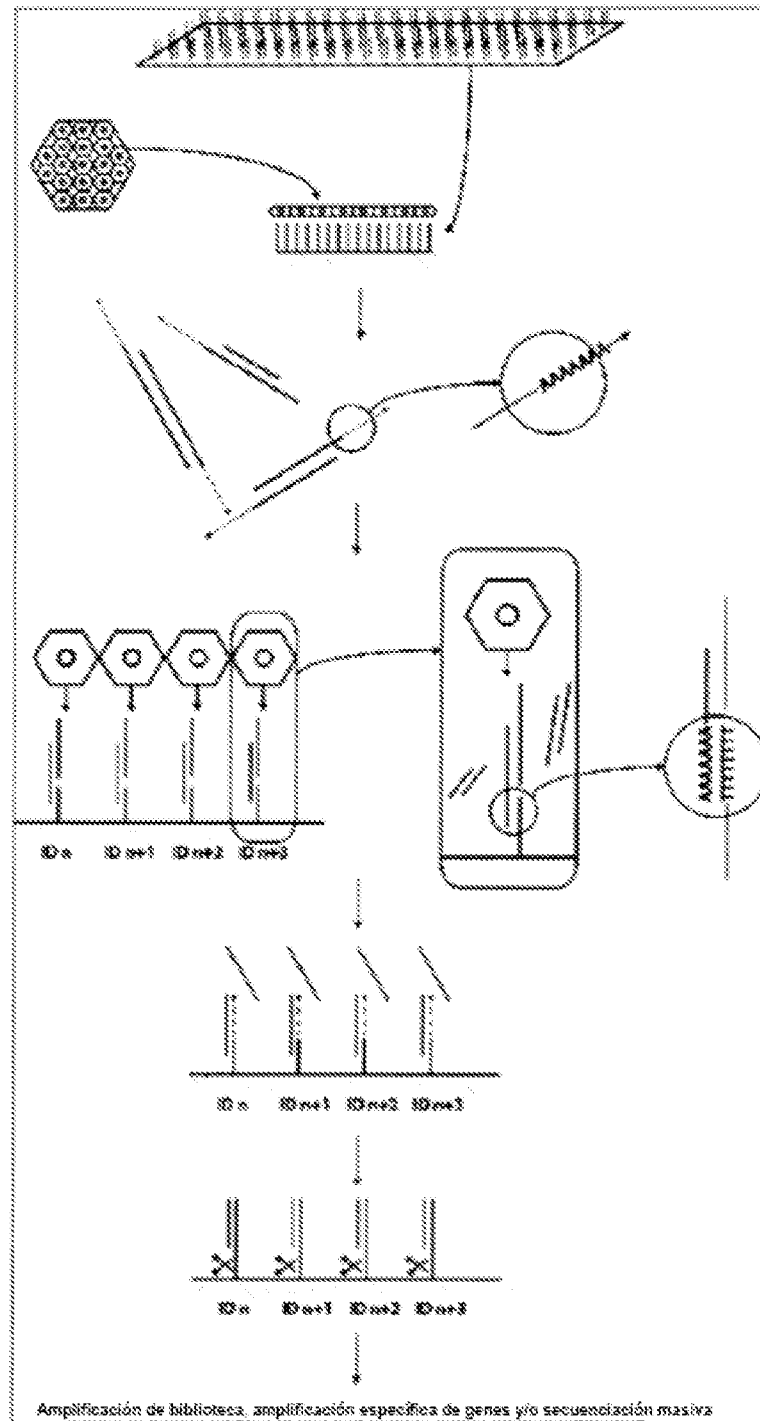


Figura 6



Figura 7

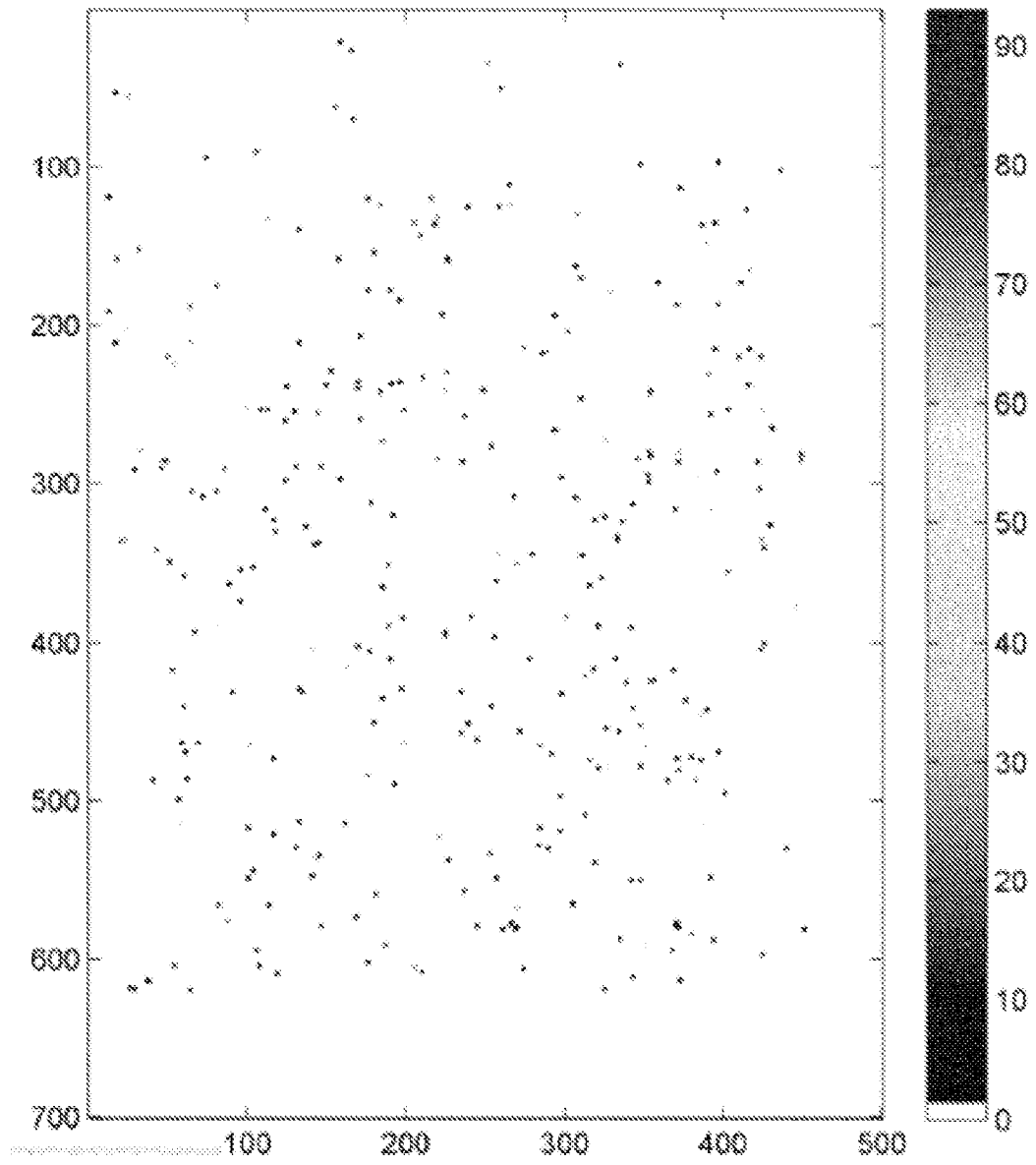


Figura 8

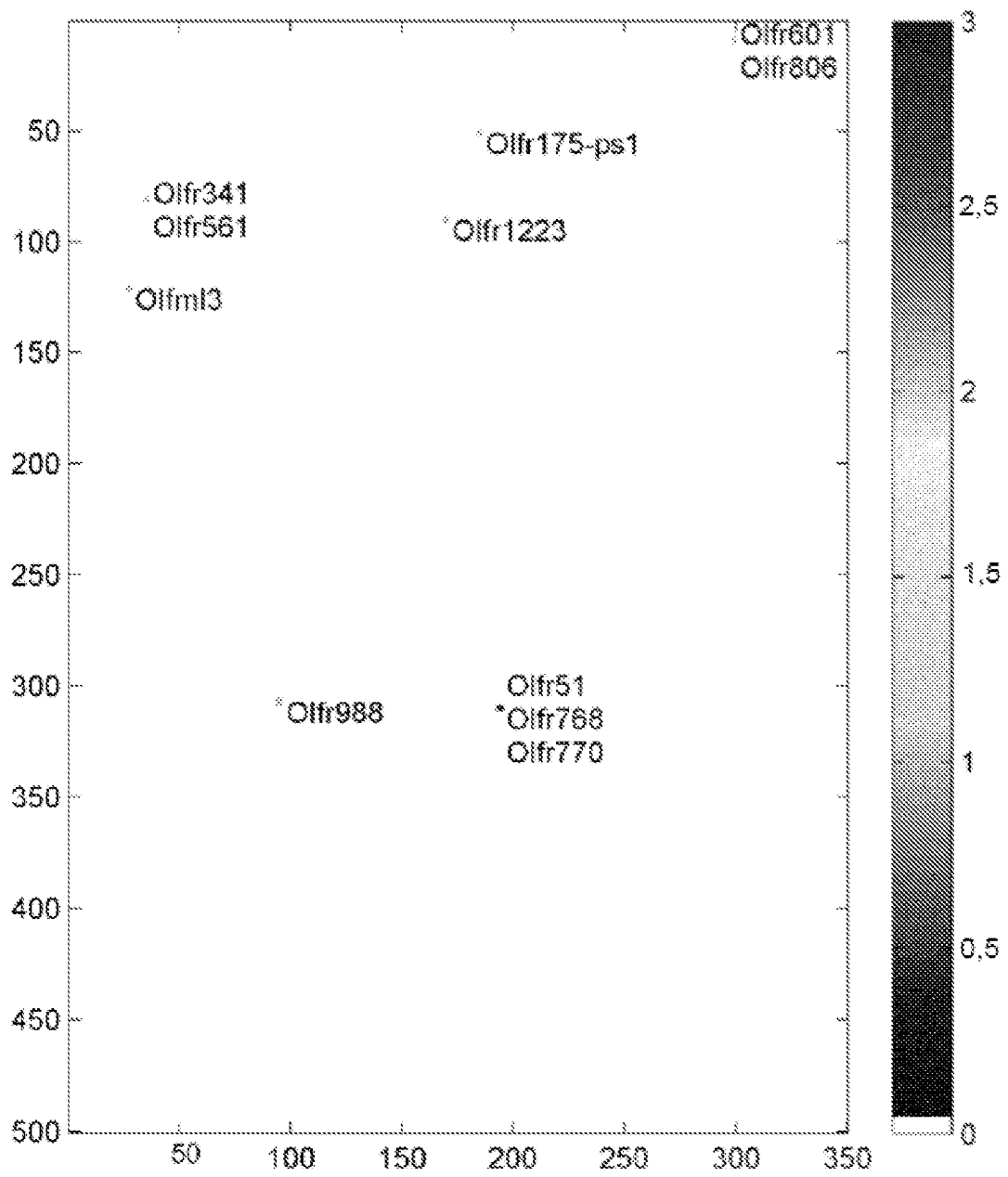


Figura 9

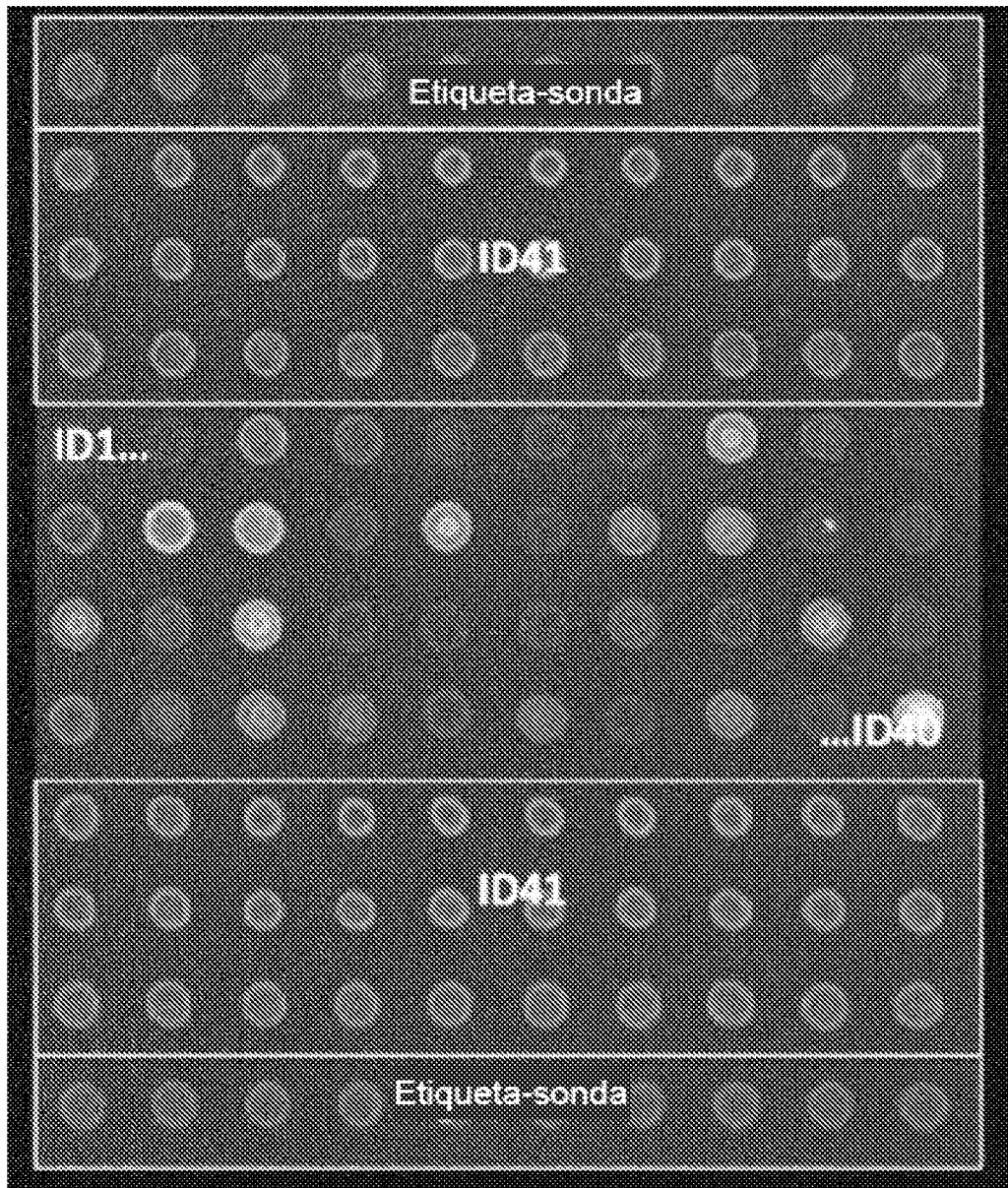


Figura 10

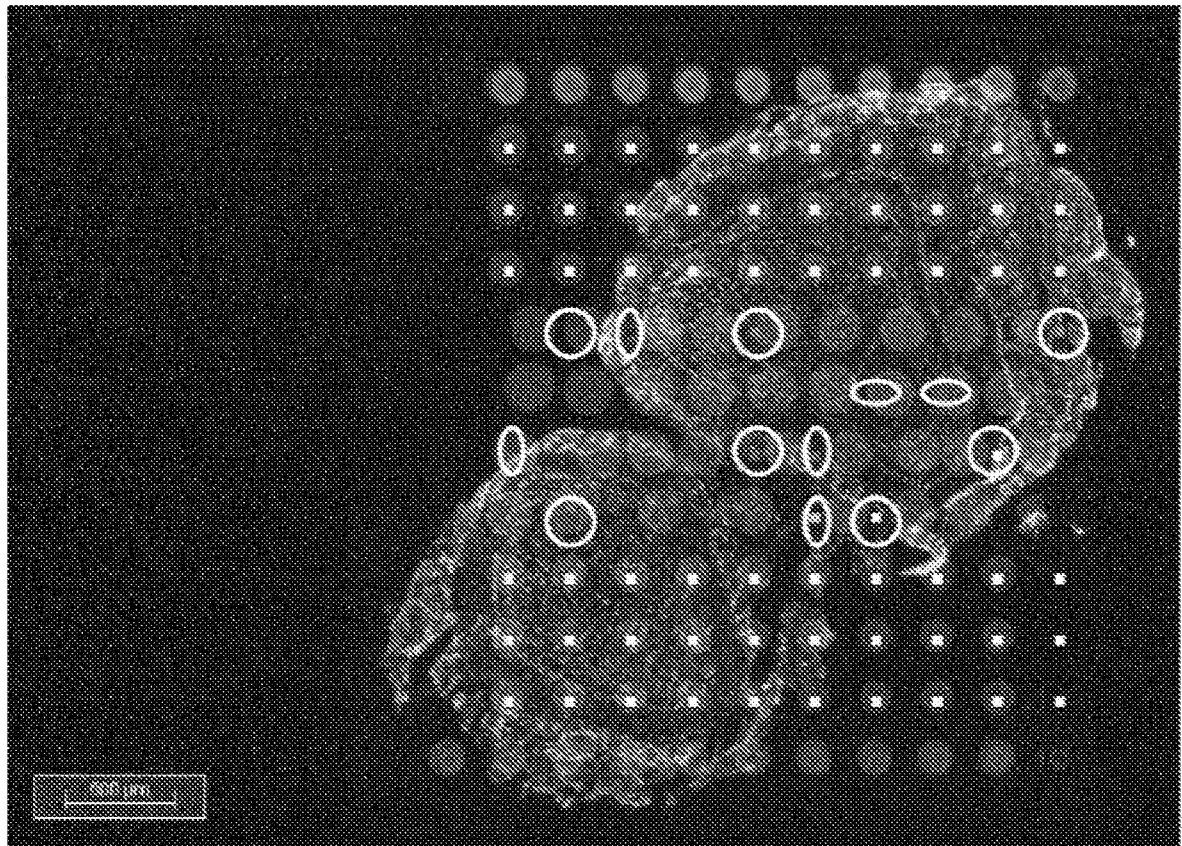


Figura 11

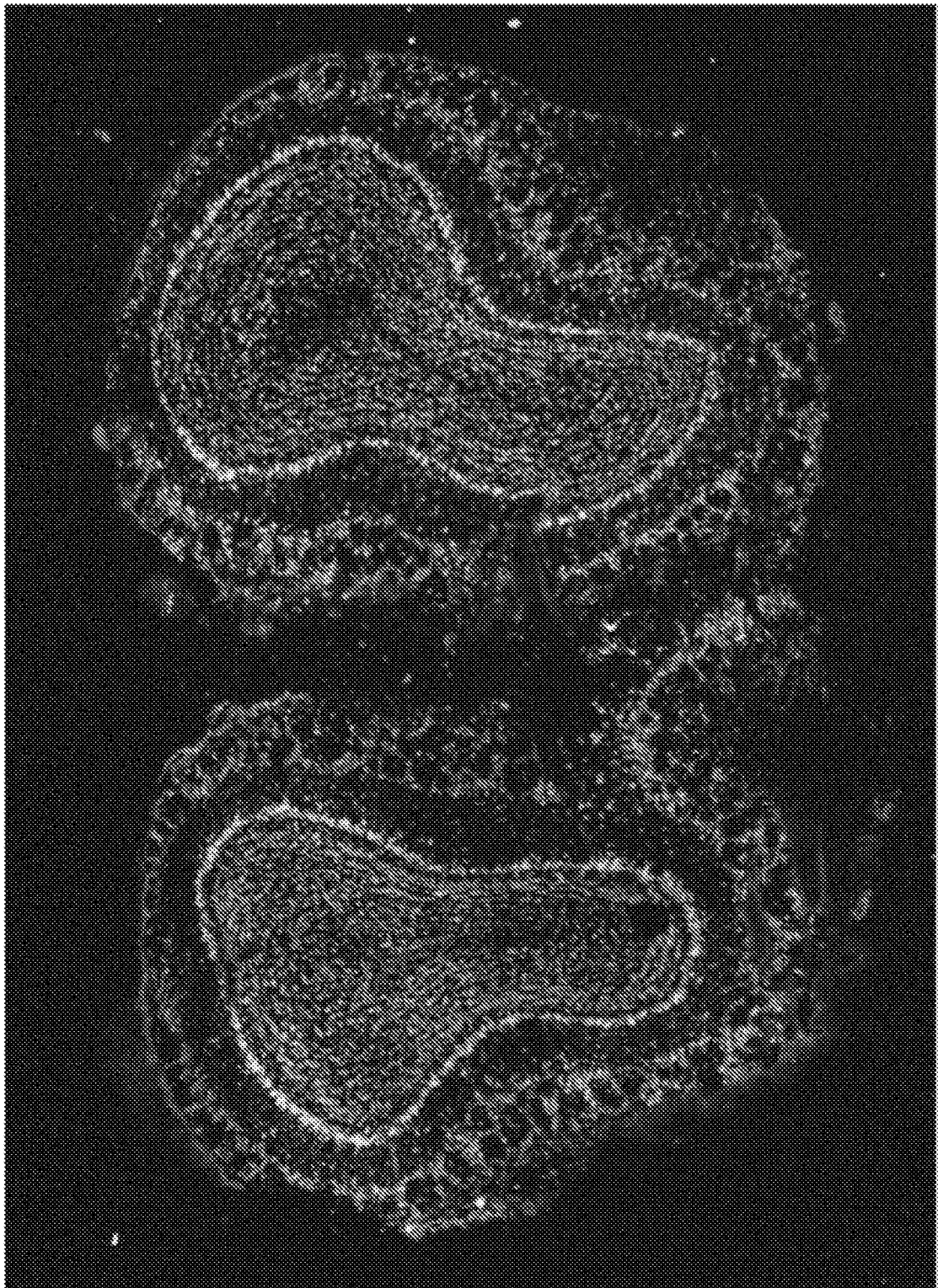


Figura 12

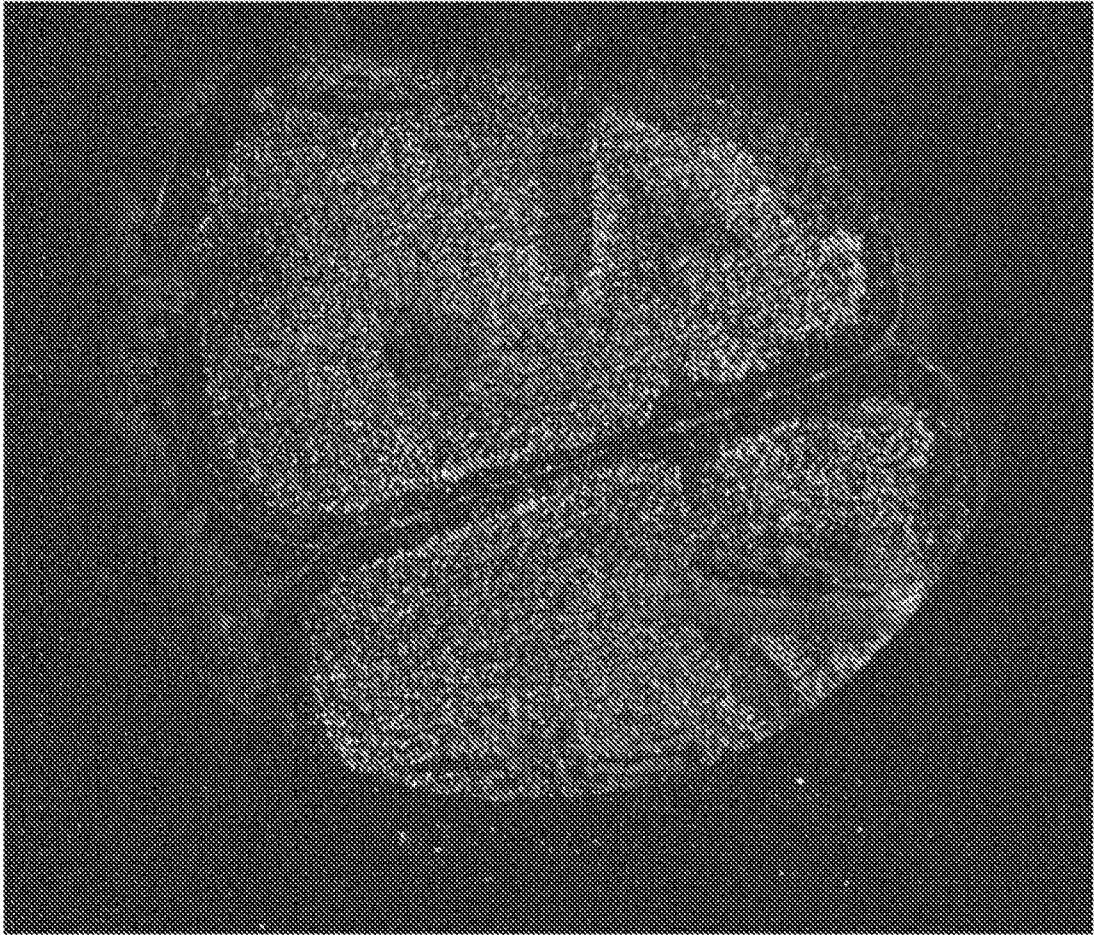


Figura 13



Figura 14

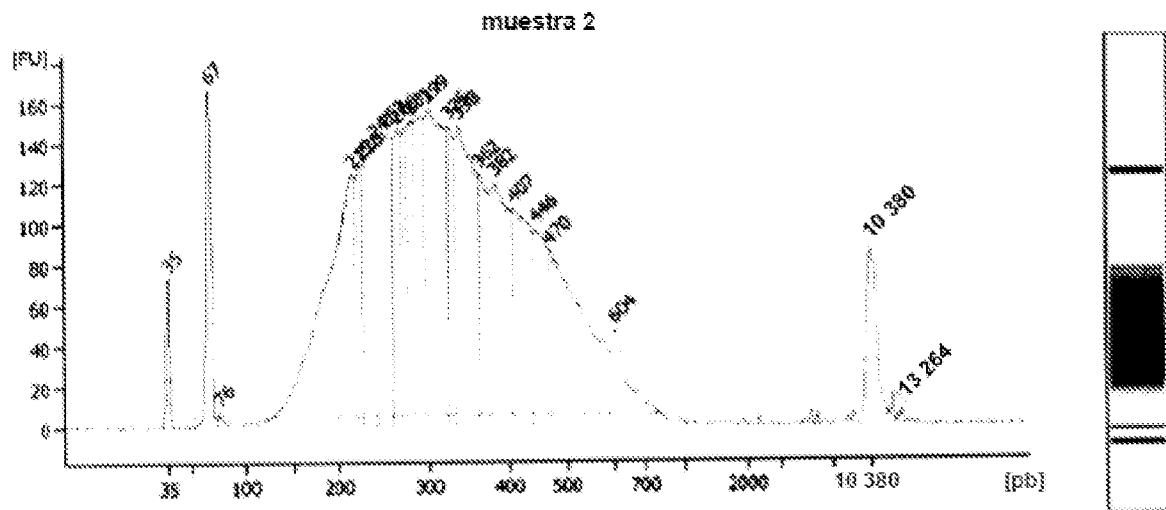


Figura 15

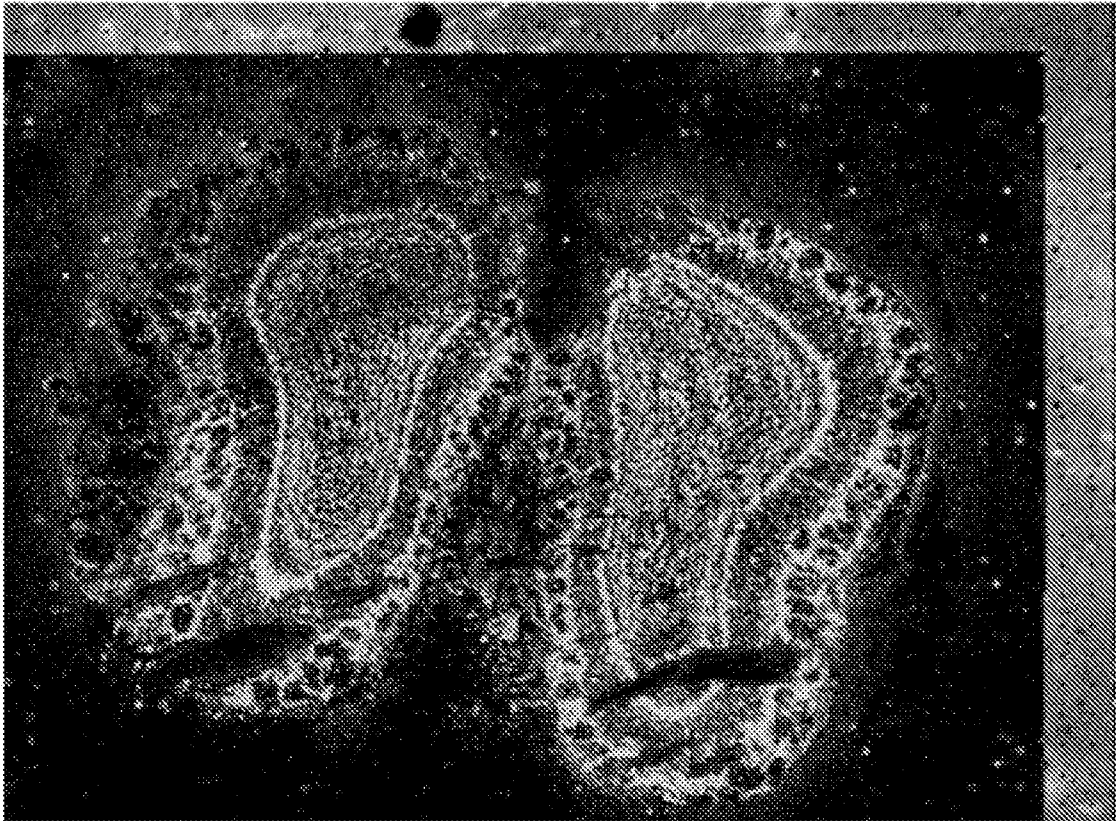


Figura 16

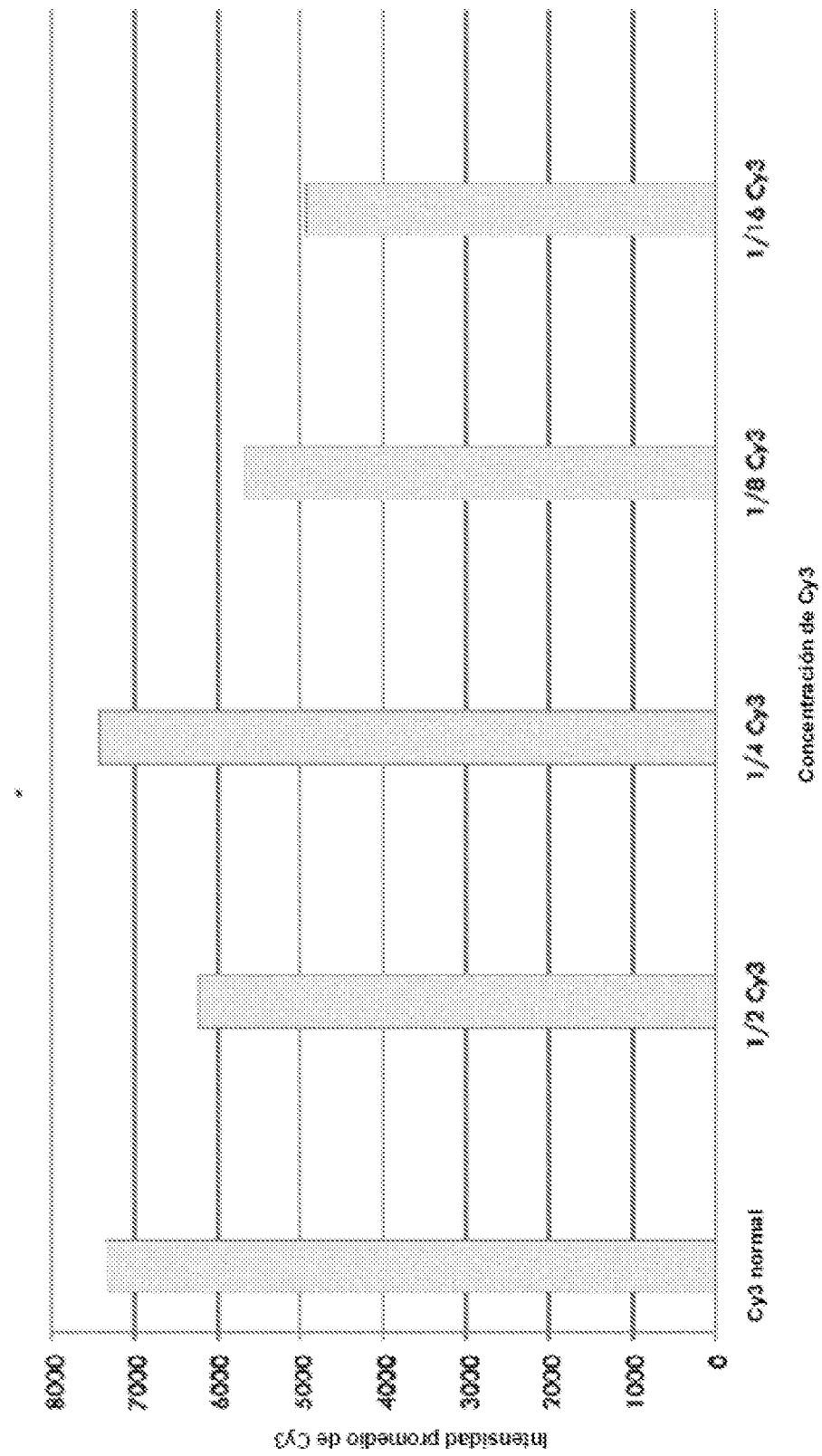


Figura 17

