

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 981 292**

51 Int. Cl.:

C12Q 1/6834 (2008.01)

C12Q 1/6841 (2008.01)

C12Q 1/6874 (2008.01)

C12Q 1/6823 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.11.2017** **E 22198482 (6)**

97 Fecha y número de publicación de la concesión europea: **10.04.2024** **EP 4148145**

54 Título: **Método para etiquetar espacialmente y analizar ácidos nucleicos en un espécimen biológico**

30 Prioridad:

17.11.2016 GB 201619458

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.10.2024

73 Titular/es:

10X GENOMICS SWEDEN AB (100.0%)
Solnavägen 3 H
113 63 Stockholm, SE

72 Inventor/es:

FRISÉN, JONAS;
STÅHL, PATRIK;
LUNDEBERG, JOAKIM;
SALMÉN, FREDRIK y
VICKOVIC, SANJA

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 981 292 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

Método para etiquetar espacialmente y analizar ácidos nucleicos en un espécimen biológico

La invención se define por las reivindicaciones adjuntas.

Específicamente, la presente invención se refiere a un método para etiquetar espacialmente ácidos nucleicos de un espécimen biológico, por ejemplo, la detección localizada o espacial de ácidos nucleicos en un espécimen biológico, tal como una muestra de tejido (denominada "genómica espacial" y "transcriptómica espacial"). En particular, la presente invención se refiere a una mejora en dichos métodos para aumentar el número y la diversidad de ácidos nucleicos únicos capturados y etiquetados de cada espécimen biológico.

Los métodos genómicos y transcriptómicos espaciales utilizan sondas de ácido nucleico inmovilizadas a un sustrato sólido (por ejemplo, una matriz) para capturar ácidos nucleicos (ADN o ARN) de un espécimen biológico y etiquetar subsecuentemente dichos ácidos nucleicos (por ejemplo, con una secuencia de "código de barras" de ácido nucleico) en base a su ubicación dentro del espécimen biológico (ver, por ejemplo, el documento WO 2012/140224). Los ácidos nucleicos "etiquetados" se liberan subsecuentemente del sustrato sólido y se analizan para generar información sobre la localización, distribución y/o expresión de genes, o de hecho sobre la localización o distribución de cualquier variación genómica (no necesariamente en un gen) en un espécimen biológico, tal como una muestra de tejido.

A manera de ejemplo, la transcriptómica espacial utiliza cebadores de transcripción inversa (RT) que comprenden etiquetas posicionales únicas (dominios, tales como secuencias de código de barras), que se inmovilizan sobre un sustrato sólido, por ejemplo, un portaobjetos 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). Los especímenes biológicos, tales como las secciones de tejido, se colocan en la matriz y se realiza una reacción de transcripción inversa en la sección de tejido en la matriz. 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. 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. Después, el ADNc se libera de la superficie de la matriz y se analiza, por ejemplo, se secuencia, lo que da como resultado un transcriptoma con información posicional exacta. Los datos de secuencia pueden entonces coincidir con una posición en la muestra de tejido. Por ejemplo, pueden obtenerse imágenes de o visualizarse la sección de tejido, por ejemplo, teñirse y fotografiarse, antes o después de la etapa de síntesis de ADNc para facilitar la correlación de la etiqueta posicional en la molécula de ADNc con una posición dentro de la muestra de tejido y los datos de secuencia pueden superponerse en una imagen del espécimen de tejido, por ejemplo mediante el uso de un ordenador, para mostrar el patrón de expresión de cualquier gen de interés en todo el tejido. Sin embargo, la etapa de visualización no es esencial ya que la molécula de ADNc etiquetada puede correlacionarse con una posición en la muestra de tejido mediante el uso de otros medios, por ejemplo, el perfil único de las moléculas capturadas con la misma etiqueta posicional, es decir, en la misma ubicación (elemento) en la matriz, puede permitir que la molécula se correlacione con una posición dentro de la muestra de tejido en base a los elementos de expresión conocidas de las células o áreas de células dentro de la muestra de tejido.

La sensibilidad de la transcriptómica y la genómica espacial está limitada por varios factores, que incluyen: (i) la eficiencia de la hibridación de las moléculas de ácidos nucleicos con las sondas de captura; (ii) la eficiencia de la reacción de extensión subsecuente; (iii) la densidad de las sondas de captura en la matriz, es decir, la resolución de la matriz; y (iv) la extensión inespecífica de moléculas de ácidos nucleicos que no se unen a sondas de captura, por ejemplo, reacciones de extensión cebadas por ácidos nucleicos dentro del espécimen biológico, tal como ADN y ARN fragmentados a causa de la preparación del espécimen biológico, por ejemplo, fijación de una sección de tejido. Estas limitaciones significan que algunos transcritos pueden estar infrarrepresentados en el análisis subsecuente. Por ejemplo, algunos transcritos pueden no capturarse en matrices de baja densidad porque el tejido en el que se expresan no hace contacto con un elemento de la matriz. Las limitaciones anteriores pueden ser particularmente problemáticas para los transcritos de baja abundancia.

Los métodos para mejorar la densidad de las sondas de captura en la matriz se han propuesto en el documento WO 2016/162309, que utiliza la tecnología de la celda de flujo y la matriz de perlas para generar matrices "aleatorias" con una alta densidad de elementos, es decir, matrices en las que la ubicación de las sondas de captura no está predeterminada. El documento WO 2016/1668825 describe métodos para realizar perfiles espaciales de moléculas biológicas. El documento WO 2012/140224 describe un método y producto para la detección localizada o espacial de ácido nucleico en una muestra de tejido.

Sin embargo, hay margen para mejorar la sensibilidad de los métodos transcriptómicos y genómicos espaciales para superar las limitaciones de los métodos conocidos en la técnica. En este sentido, los inventores han descubierto sorprendentemente que es posible aumentar el número y la diversidad de moléculas de ácidos nucleicos únicas capturadas por las sondas de captura al combinar la etapa de liberación de las sondas de la superficie del sustrato sólido, por ejemplo, matriz, con la etapa de extensión de las sondas mediante el uso de los ácidos nucleicos capturados como plantillas para la extensión, por ejemplo, las etapas de liberación y extensión se realizan simultáneamente.

En particular, los inventores han descubierto que es posible combinar la enzima usada para liberar las sondas de captura de la superficie del sustrato sólido (es decir, la enzima de escisión) con la mezcla de reacción usada en la etapa de extensión del ácido nucleico sin una pérdida sustancial de la actividad enzimática de escisión. Además, y contrario a las expectativas, la liberación y extensión combinadas (por ejemplo simultáneas) de las sondas de captura no elimina la captura localizada de moléculas de ácidos nucleicos, es decir, el etiquetado espacial de los ácidos nucleicos en el espécimen biológico se retiene en los métodos de la invención. Por otra parte, como se describe en detalle en los Ejemplos más abajo, los inventores han determinado ventajosamente que la liberación y extensión simultáneas de las sondas de captura aumentan el número y diversidad de moléculas de ácidos nucleicos únicas capturadas mediante los métodos de la invención, de esta manera se mejora la sensibilidad de los métodos transcriptómicos y genómicos espaciales.

Aunque no se desea limitarse a la teoría, se plantea la hipótesis de que la combinación de las etapas de liberación y síntesis de ácidos nucleicos puede dar como resultado la liberación localizada de algunas o todas las sondas de captura antes de su extensión, es decir, las sondas pueden liberarse en la cercanía local de su posición original en la matriz. Se cree que esto puede mejorar la sensibilidad de los métodos al permitir que las sondas de captura (que no se han hibridado con un ácido nucleico del espécimen biológico en contacto entre el espécimen y la matriz) difundan en el espécimen, de esta manera aumenta la probabilidad de que la sonda de captura se hibride con una secuencia diana. Alternativa o adicionalmente, la liberación y extensión simultáneas de las sondas de captura pueden mejorar la reacción de extensión, por ejemplo, la reacción de síntesis de ADN, porque las sondas de captura ya no están físicamente restringidas, es decir, inmovilizadas sobre sustrato sólido, y/o porque el entorno del microrreactor formado entre el espécimen biológico y la superficie de la matriz es particularmente adecuado para una reacción de extensión de ácido nucleico con plantilla. Por lo tanto, la reacción de extensión puede continuar de manera más eficiente una vez que las sondas de captura, que se hibridan con las secuencias diana cuando todavía estaban inmovilizadas sobre el sustrato sólido, se liberan de la superficie de la matriz.

En consecuencia, la presente invención proporciona un método para etiquetar espacialmente ácidos nucleicos de un espécimen biológico que comprende:

(a) proporcionar un sustrato sólido en el que se inmovilizan múltiples especies de sondas de captura de manera que cada especie ocupe una posición distinta en el sustrato sólido, en donde cada sonda de captura de las múltiples especies es un cebador para una reacción de extensión y en donde la sonda de captura comprende una molécula de ácido nucleico que comprende:

- (i) un dominio de escisión para liberar la sonda de captura de la superficie del sustrato sólido,
- (ii) un dominio posicional que corresponde a una posición de la sonda de captura en el sustrato sólido, y
- (iii) un dominio de captura;

(b) poner en contacto el sustrato sólido con el espécimen biológico; (c) liberar simultáneamente las múltiples especies de sondas de captura del sustrato sólido en condiciones que permitan que un ácido nucleico del espécimen biológico se hibride con el dominio de captura de la sonda de captura como una plantilla de extensión para producir una sonda de captura extendida, en donde la etapa (c) comprende poner en contacto el sustrato sólido con una mezcla de reacción acuosa que comprende:

(i) una transcriptasa inversa capaz de extender la sonda de captura mediante el uso del ácido nucleico hibridado a la sonda de captura como una plantilla de extensión para producir la sonda de captura extendida, en donde la sonda de captura extendida comprende además un saliente de ADN de hebra única;

(ii) un oligonucleótido de cambio de plantilla que comprende una secuencia capaz de hibridarse con el saliente de ADN de hebra única y el dominio de amplificación, en donde la transcriptasa inversa es capaz de extender aún más la sonda de captura extendida mediante el uso del oligonucleótido de cambio de plantilla como una plantilla de extensión;

(iii) una polimerasa capaz de extender un cebador hibridado al dominio de amplificación incorporado en la sonda de captura extendida como una plantilla de extensión para producir una sonda de captura extendida de doble hebra; y

(iv) un medio para liberar las múltiples especies de sondas de captura del sustrato sólido.

Los métodos de la invención representan una mejora significativa con respecto a otros métodos de transcriptómica y genómica espacial conocidos en la técnica, por ejemplo, los documentos WO 2012/140224, WO 2014/060483 y WO 2016/162309. Los métodos descritos en la presente descripción dan como resultado un aumento en el número y la diversidad de transcritos únicos capturados del espécimen biológico, de esta manera se proporciona un transcriptoma etiquetado espacialmente que es más representativo del perfil de expresión del espécimen biológico. Por otra parte, como las sondas de captura pueden liberarse de la superficie de la matriz antes de su extensión, no es necesario unir las sondas a la matriz por medio de sus extremos 5' en los métodos de la presente invención. En este sentido, las sondas de captura pueden unirse a la matriz por medio de sus extremos 3' y la liberación de las sondas de captura de

la superficie de la matriz expondrá los extremos 3' de las sondas, lo que permite que las sondas funcionen como cebadores para las reacciones de extensión de ácidos nucleicos.

Como se describe en más detalle más abajo, el método de la invención puede comprender una etapa adicional de análisis de las sondas extendidas. En este sentido, es evidente que la combinación de etiquetado espacial de los ácidos nucleicos de un espécimen biológico y el análisis subsecuente de dichos ácidos nucleicos etiquetados facilita la detección localizada de un ácido nucleico en un espécimen biológico, por ejemplo, muestra de tejido. Por lo tanto, en una modalidad, el método de la invención puede usarse para determinar y/o analizar todo el transcriptoma o genoma de un espécimen biológico, por ejemplo, el transcriptoma o genoma global de un espécimen biológico. Sin embargo, el método no se limita a esto y abarca determinar y/o analizar todo o parte del transcriptoma o genoma. Por lo tanto, el método puede implicar determinar y/o analizar una parte o subconjunto del transcriptoma o genoma, por ejemplo, un transcriptoma correspondiente a un subconjunto de genes, por ejemplo, un conjunto de genes particulares, por ejemplo relacionados con una enfermedad o afección particular, tipo de tejido, etc.

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

Alternativamente, el método de la invención puede verse como un método para la detección localizada o espacial de ácido nucleico, ya sea ADN o ARN, en un espécimen biológico, por ejemplo, muestra de tejido, o para la determinación y/o análisis localizado o espacial de ácido nucleico (ADN o ARN) en una muestra de tejido. En particular, el método puede usarse para la detección o determinación y/o análisis localizado o espacial de la expresión génica o variación genómica en una muestra de tejido. La detección/determinación/análisis localizado/espacial significa que el ARN o ADN 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 o ADN 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. La ubicación o posición nativa del ARN o ADN (o en otras palabras, la ubicación o posición del ARN o ADN en la muestra de tejido), por ejemplo, un gen o locus genómico expresado, puede determinarse.

Como se describe con más detalle más abajo, cualquier método de análisis de ácidos nucleicos puede usarse en la etapa de análisis (etapa (d)). Típicamente esto puede implicar la secuenciación, es decir, el análisis de la secuencia de las sondas extendidas, pero no es necesario realizar una determinación de secuencia real. Por ejemplo, pueden usarse métodos de análisis específicos 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 o gen particular, etcétera). Un método de análisis ilustrativo es una reacción de PCR específica de secuencia.

La información del análisis de secuencias obtenida en la etapa (d) puede usarse para obtener información espacial en cuanto al ARN o ADN en el espécimen biológico, por ejemplo, muestra de tejido. En otras palabras, la información del análisis de secuencias puede proporcionar información sobre la ubicación del ARN o ADN en el espécimen biológico, por ejemplo, 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 o ADN particular que puede ser en sí misma espacialmente informativa en el contexto del espécimen biológico, por ejemplo, muestra de tejido, usada y/o la información espacial (por ejemplo, localización espacial) puede derivarse de la posición del espécimen biológico, por ejemplo, muestra de tejido, sobre el sustrato sólido, por ejemplo, matriz, junto con la información de secuenciación. Por lo tanto, el método puede implicar simplemente correlacionar la información del análisis de secuencias con una posición en el espécimen biológico, por ejemplo, muestra de tejido, por ejemplo, en virtud de la etiqueta posicional y su correlación con una posición en el espécimen biológico, por ejemplo, muestra de tejido. Sin embargo, en algunas modalidades, la información espacial puede obtenerse convenientemente mediante la correlación de los datos de análisis de secuencias con una imagen del espécimen biológico, por ejemplo, muestra de tejido. En consecuencia, en una modalidad preferida el método también incluye una etapa de:

(e) correlacionar dicha información del análisis de secuencias con una imagen de dicho espécimen biológico, en donde se obtienen imágenes del espécimen biológico antes o después de la etapa (c), es decir después de la etapa (b).

Por lo tanto, se observará que la matriz puede usarse para capturar ARN, por ejemplo ARNm, o ADN de un espécimen biológico, por ejemplo muestra de tejido, que se pone en contacto con dicho sustrato sólido, por ejemplo matriz. Los métodos de la invención pueden considerarse por lo tanto como métodos para cuantificar la expresión espacial de uno o más genes en una muestra de tejido. Expresado de otra manera, los métodos de la presente invención pueden usarse para detectar la expresión espacial de uno o más genes en un espécimen biológico, por ejemplo, muestra de tejido. Aún de otra manera, los métodos de la presente invención pueden usarse para determinar simultáneamente la expresión de uno o más genes en una o más posiciones dentro de un espécimen biológico, por ejemplo, muestra de tejido. Aún más, los métodos pueden verse como métodos para el análisis parcial o global del transcriptoma o genoma de un espécimen biológico, por ejemplo, muestra de tejido, con resolución espacial bidimensional.

- La etapa (c) en el método se verá como relacionada con la combinación de las reacciones de liberación y extensión, es decir, combinar los medios para lograr la liberación de la sonda de captura y la extensión con plantilla de ácido nucleico diana de dichas sondas de captura. Como se describió anteriormente, la mejora en la sensibilidad del método puede ser una consecuencia de la liberación de sondas de captura que no se han unido a una secuencia diana en la etapa (b), lo que puede facilitar la unión de otras moléculas de ácidos nucleicos diana en el espécimen biológico. Alternativa o adicionalmente, la mejora en la sensibilidad del método puede ser el resultado de la reducción del impedimento estérico en el complejo sonda de captura-ácido nucleico diana o la formación de un microentorno que es particularmente propicio para las reacciones de extensión de ácido nucleico, de esta manera se mejora la eficiencia de la reacción de extensión.
- En consecuencia, la liberación y extensión simultáneas de las sondas de captura no significa que todas las sondas de captura se liberen y extiendan al mismo tiempo, sino que los medios para liberar y extender las sondas de captura se aplican al sustrato sólido, por ejemplo, matriz, simultáneamente, es decir, sustancialmente al mismo tiempo.
- En particular, como se mencionó anteriormente, las sondas de captura no se limitan a una orientación particular en la matriz, a diferencia de los métodos en los documentos WO 2012/140224, WO 2014/060483 y WO 2016/162309, que requieren que las sondas de captura se inmovilicen sobre el sustrato sólido de manera que tengan un extremo 3' libre capaz de funcionar como un cebador de extensión. En este sentido, la combinación de las etapas de liberación y extensión elimina el requisito de una orientación particular de las sondas de captura sobre el sustrato sólido. Sin embargo, en algunas modalidades, se prefiere que las sondas de captura se inmovilicen sobre el sustrato sólido de manera que tengan un extremo 3' libre capaz de funcionar como un cebador de extensión.
- Además, ya que las sondas de captura pueden orientarse sobre el sustrato sólido de manera que el dominio de captura no esté libre o disponible para interactuar con (es decir, unirse o hibridarse con) las moléculas de ácidos nucleicos en el espécimen biológico (es decir, las sondas de captura pueden inmovilizarse por medio de sus extremos 3'), no es necesario poner en contacto el espécimen biológico con el sustrato sólido en condiciones que permitan (es decir, son adecuadas o facilitan) que los ácidos nucleicos del espécimen biológico se hibriden con el dominio de captura en dichas sondas de captura. Sin embargo, en algunas modalidades (por ejemplo, donde las sondas de captura se inmovilizan sobre el sustrato sólido de manera que tengan un extremo 3' libre capaz de funcionar como un cebador de extensión, por ejemplo, por medio de sus extremos 5') la etapa (b) puede realizarse en condiciones que permitan que los ácidos nucleicos del espécimen biológico se hibriden con el dominio de captura en dichas sondas de captura.
- Por lo tanto, en modalidades donde las sondas de captura se inmovilizan sobre el sustrato sólido de manera que tengan un extremo 3' libre capaz de funcionar como un cebador de extensión (por ejemplo, por su extremo 5'), algunas sondas de captura pueden liberarse del sustrato sólido antes de su extensión, es decir, algunas sondas de captura se liberan y subsecuentemente se extienden. Por otra parte, algunas sondas de captura pueden extenderse al mismo tiempo que se liberan del sustrato sólido, es decir, algunas sondas de captura se extienden y liberan del sustrato sólido simultáneamente. En consecuencia, en modalidades donde las sondas de captura se inmovilizan sobre el sustrato sólido de manera que tengan un extremo 3' libre capaz de funcionar como un cebador de extensión (por ejemplo, por su extremo 5'), la etapa (c) puede comprender reacciones de liberación y extensión simultáneas y reacciones de liberación y extensión secuenciales, es decir, liberación y extensión subsecuente de las sondas de captura. Las reacciones simultáneas pueden ocurrir particularmente cuando la etapa (b) se realiza en condiciones que permiten que los ácidos nucleicos del espécimen biológico se hibriden con el dominio de captura en dichas sondas de captura, es decir, de manera que algunos ácidos nucleicos puedan hibridarse con los dominios de captura de las sondas de captura antes de las reacciones de liberación y extensión combinadas, es decir antes de la etapa (c).
- En consecuencia, en modalidades donde las sondas de captura se inmovilizan sobre el sustrato sólido de manera que no tengan un extremo 3' libre capaz de funcionar como un cebador de extensión (por ejemplo, por su extremo 3'), todas las sondas de captura pueden liberarse del sustrato sólido antes de su extensión, es decir, todas las sondas de captura se liberan y subsecuentemente se extienden. En consecuencia, en modalidades donde las sondas de captura se inmovilizan sobre el sustrato sólido de manera que no tengan un extremo 3' libre capaz de funcionar como un cebador de extensión (por ejemplo, por su extremo 3'), la etapa (c) puede comprender reacciones de liberación y extensión secuenciales, es decir, liberación y extensión subsecuente de las sondas de captura, a pesar de la presencia de ambos componentes requeridos para las reacciones de liberación y extensión.
- En el método de la invención, la enzima polimerasa y los medios para liberar dichas sondas de captura de la superficie del sustrato sólido, por ejemplo, la enzima de escisión, se combinan en una única mezcla de reacción, que se pone en contacto con la matriz. Por lo tanto, la etapa (c) comprende poner en contacto el sustrato sólido con una mezcla de reacción (es decir, una mezcla de reacción acuosa) que comprende:
- (i) una transcriptasa inversa capaz de extender la sonda de captura mediante el uso del ácido nucleico hibridado a la sonda de captura como una plantilla de extensión para producir la sonda de captura extendida, en donde la sonda de captura extendida comprende además un saliente de ADN de hebra única;
 - (ii) un oligonucleótido de cambio de plantilla que comprende una secuencia capaz de hibridarse con el saliente de ADN de hebra única y un dominio de amplificación, en donde la transcriptasa inversa es capaz de extender

aún más la sonda de captura extendida mediante el uso del oligonucleótido de cambio de plantilla como una plantilla de extensión;

(iii) una polimerasa capaz de extender un cebador hibridado al dominio de amplificación incorporado en la sonda de captura extendida como una plantilla de extensión para producir una sonda de captura extendida de doble hebra; y

(iv) un medio para liberar las múltiples especies de sondas de captura del sustrato sólido.

En consecuencia, se proporciona el uso de una mezcla de reacción (es decir, una mezcla de reacción acuosa) que comprende:

(i) una enzima polimerasa capaz de extender sondas de captura (como se definió anteriormente) mediante el uso de las moléculas de ácidos nucleicos hibridadas a las sondas de captura como plantillas de extensión; y

(ii) medios para liberar dichas sondas de captura de la superficie de un sustrato sólido,

en un método para etiquetar espacialmente ácidos nucleicos de un espécimen biológico, tal como los métodos definidos en la presente descripción.

La etapa de liberación de las sondas de captura de la superficie del sustrato sólido puede lograrse de varias maneras. En este sentido, como se mencionó anteriormente, las sondas de captura comprenden un dominio de escisión para facilitar la liberación de las sondas de captura. En consecuencia, "medios para liberar las sondas de captura de la superficie del sustrato sólido" incluye cualquier componente que sea adecuado para este propósito, por ejemplo, medios químicos o enzimáticos. Como se describe más abajo, el método preferido para liberar las sondas de captura de la superficie del sustrato sólido comprende la escisión enzimática de la sonda de captura en el dominio de escisión. En este sentido, liberar la sonda a través del dominio de escisión asegura que las sondas de captura liberadas (que incluyen cualquier sonda de captura extendida o parcialmente extendida) comprendan el dominio posicional, que contiene información relacionada con la ubicación de las sondas de captura en la matriz.

El término "dominio de escisión" se refiere a un dominio dentro de la sonda de captura que puede escindirse específicamente para liberar la sonda de captura (o parte de la sonda de captura que comprende el dominio posicional y la secuencia 3' respecto al dominio posicional (si está presente), es decir, el dominio de captura y el producto de extensión (si está presente)) de la superficie del sustrato sólido.

Como se describe más abajo, las sondas de captura pueden unirse al sustrato sólido por cualquier medio adecuado, tal como mediante inmovilización química, por ejemplo mediante un reticulador químico. Por lo tanto, en algunas modalidades, el dominio de escisión de la sonda de captura se refiere a un reticulador químico escindible que une la sonda de captura a la superficie del sustrato sólido. Por lo tanto, en algunas modalidades, el dominio de captura no forma parte de la secuencia de nucleótidos de la sonda de captura. Por ejemplo, la sonda de captura puede comprender un dominio posicional de ácido nucleico y un dominio de captura y un dominio de escisión de ácido nucleico (por ejemplo, un reticulador químico escindible) conjugado al extremo 5' o 3' de la sonda de captura (en dependencia de la orientación de la sonda unida al sustrato sólido). Por lo tanto, la escisión de dicho dominio da como resultado la liberación de la totalidad de la sonda de captura de ácido nucleico, es decir, que comprende el dominio posicional y la secuencia 3' respecto al dominio posicional desde la superficie del sustrato sólido. Los reticuladores químicos escindibles capaces de inmovilizar moléculas de ácidos nucleicos a un sustrato sólido, como se define en la presente descripción, se conocen en la técnica.

En modalidades preferidas, el dominio de escisión de la sonda de captura se refiere a una secuencia de nucleótidos dentro de la sonda de captura que puede escindirse específicamente, ya sea química o preferentemente enzimáticamente. La ubicación del dominio de escisión dentro de la sonda de captura dependerá de si las sondas de captura se inmovilizan o no en el sustrato sólido de manera que tengan un extremo 3' libre capaz de funcionar como un cebador de extensión (por ejemplo, por su extremo 5' o 3'). Por ejemplo, si las sondas de captura se inmovilizan por su extremo 5', el dominio de escisión se ubicará 5' respecto al dominio posicional y la escisión de dicho dominio da como resultado la liberación de parte de la sonda de captura que comprende el dominio posicional y la secuencia 3' respecto al dominio posicional, y opcionalmente parte del dominio de escisión, de la superficie del sustrato sólido. Alternativamente, si las sondas de captura se inmovilizan por su extremo 3', el dominio de escisión se ubicará 3' respecto al dominio de captura (y al dominio posicional) y la escisión de dicho dominio da como resultado la liberación de parte de la sonda de captura que comprende el dominio posicional y la secuencia 3' respecto al dominio posicional desde la superficie del sustrato sólido. Preferentemente, la escisión da como resultado la eliminación completa del dominio de escisión, particularmente cuando las sondas de captura se inmovilizan por medio de su extremo 3' ya que la presencia de una parte del dominio de escisión puede interferir con la hibridación del dominio de captura y el ácido nucleico diana y/o su extensión subsecuente.

La "escisión" se define en sentido general en la presente descripción para incluir cualquier medio para romper un enlace covalente. En algunas modalidades, la escisión implica la escisión de un enlace covalente en una hebra de nucleótidos (es decir, escisión de la hebra o escisión de la hebra), por ejemplo mediante la escisión de un enlace fosfodiéster.

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 de doble hebra o de hebra única 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 del ácido nucleico, por ejemplo, molécula de ADNc. Como se mencionó anteriormente, se observará que la liberación de al menos parte de las sondas de captura de la superficie del sustrato sólido requiere liberar una parte que comprende el dominio posicional y el dominio de captura de la sonda de captura y, en algunas modalidades, cualquier secuencia producida por la extensión con plantilla del dominio de captura. Por lo tanto, si la sonda de captura se extiende parcial o completamente mientras aún está unida al sustrato sólido (es decir, inmovilizada sobre este), la escisión del dominio de escisión liberará la sonda de captura extendida, que comprende el dominio posicional, el dominio de captura y el producto de extensión. Por lo tanto, en algunas modalidades preferidas, es decir, cuando las sondas de captura se inmovilizan por medio de su extremo 5', la escisión de la sonda de captura debe tener lugar 5' respecto 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-liasas 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 sólido indirectamente, es decir, por medio 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.

Otro ejemplo de un dominio de escisión que puede utilizarse en modalidades donde la sonda de captura se inmoviliza indirectamente al sustrato sólido, es decir por medio de una sonda de superficie, comprende un sitio o secuencia de reconocimiento de nickasa. En este sentido, las enzimas nickasas escinden solo una hebra en un dúplex de ácido nucleico. Las nickasas son endonucleasas que escinden solo una sola hebra de un dúplex de ADN. Por lo tanto, el dominio de escisión puede comprender un sitio de reconocimiento de nickasa cerca del extremo 5' de la sonda de superficie (y/o el extremo 5' de la sonda de captura) de manera que la escisión de la sonda de superficie o sonda de captura desestabiliza el dúplex entre la sonda de superficie y la sonda de captura para liberar de esta manera la sonda de captura del sustrato sólido.

Las enzimas nickasas también podrían usarse en modalidades donde la sonda de captura se inmoviliza directamente al sustrato sólido. Por ejemplo, el sustrato sólido puede ponerse en contacto con una molécula de ácido nucleico que se hibrida con el dominio de escisión de la sonda de captura para proporcionar o reconstituir un sitio de reconocimiento de nickasa, es decir, una sonda auxiliar de escisión. Por lo tanto, el contacto con una enzima nickasa dará como resultado la escisión del dominio de escisión para liberar de esta manera la sonda de captura del sustrato sólido. Será evidente que tales sondas auxiliares de escisión podrían usarse para proporcionar o reconstituir sitios de reconocimiento de escisión para otras enzimas de escisión, por ejemplo, enzimas de restricción, como se describió anteriormente.

Por lo tanto, en algunas modalidades, el método también puede comprender una etapa de poner en contacto el sustrato sólido con una sonda auxiliar de escisión, es decir, en condiciones que permitan que la sonda auxiliar de escisión se hibride con el dominio de escisión de la sonda de captura para proporcionar un sitio de reconocimiento de escisión, por ejemplo, un sitio de reconocimiento de enzima nickasa o un sitio de reconocimiento de enzima de restricción.

Algunas nickasas introducen mellas de hebra única solo en sitios particulares en una molécula de ADN, al unirse y reconocer una secuencia de reconocimiento de nucleótidos particular. Se ha descubierto un número de nickasas de origen natural, de las cuales se han determinado en la actualidad las propiedades de reconocimiento de secuencias para al menos cuatro. Las nickasas se describen en la patente de Estados Unidos núm. 6,867,028, y cualquier nickasa adecuada puede usarse en los métodos de la invención.

En modalidades que utilizan una enzima nickasa, la enzima nickasa puede eliminarse del ensayo o inactivarse antes de la etapa (d) para evitar la escisión no deseada de las sondas extendidas o los amplicones de las mismas.

Por lo tanto, en una modalidad preferida, "medios para liberar dichas sondas de captura de la superficie de un sustrato sólido" comprende una enzima de escisión, tal como una enzima como se definió anteriormente. Será evidente a partir de la descripción anterior que una enzima de escisión puede referirse a una combinación o mezcla de enzimas que tienen actividad de escisión. Por ejemplo, en una modalidad preferida, la enzima de escisión comprende una mezcla de uracil ADN glucosilasa (UDG) y una ADN glucosilasa-liasas, tal como la endonucleasa VIII. Se prefiere

particularmente una mezcla disponible comercialmente de estas enzimas, conocida como la enzima USER™. Por lo tanto, en algunas modalidades, el dominio de escisión comprende una secuencia de poli-U.

Como se mencionó anteriormente, la enzima de escisión se pone en contacto con la matriz al mismo tiempo, o antes, que la enzima polimerasa, en una única mezcla de reacción. En consecuencia, la enzima de escisión debe ser funcional en las mismas condiciones que la enzima polimerasa, por ejemplo, funcional en las mismas condiciones de tampón, sal, temperatura usadas para la enzima polimerasa.

Por "funcional" se entiende que la enzima de escisión puede mostrar cierta actividad reducida en la escisión del dominio de escisión de la sonda de captura con relación a la actividad de la enzima de escisión en el dominio de escisión en condiciones que son óptimas para la enzima, por ejemplo, en las condiciones de tampón, sal y temperatura recomendadas por el fabricante. Por lo tanto, la enzima de escisión puede considerarse funcional si tiene al menos 50 %, por ejemplo, al menos 60, 70, 80, 85, 90, 95, 96, 97, 98, 99 o 100 %, actividad con relación a la actividad de la enzima de escisión en condiciones que son óptimas para la enzima.

Adicionalmente, la enzima polimerasa debe ser funcional en las mismas condiciones que la enzima de escisión, por ejemplo, funcional en las mismas condiciones de tampón, sal y temperatura usadas para la enzima de escisión.

Por "funcional" se entiende que la enzima polimerasa puede mostrar cierta actividad reducida en la extensión con plantilla de la diana de la sonda de captura con relación a la actividad de la enzima polimerasa en la misma extensión con plantilla de la diana en condiciones que son óptimas para la enzima, por ejemplo, en las condiciones de tampón, sal y temperatura recomendadas por el fabricante. Por lo tanto, la enzima polimerasa puede considerarse funcional si tiene al menos 50 %, por ejemplo, al menos 60, 70, 80, 85, 90, 95, 96, 97, 98, 99 o 100 %, de actividad con relación a la actividad de la enzima polimerasa en condiciones que son óptimas para la enzima.

Aunque será evidente a partir de la descripción en la presente descripción que el método de la invención puede usarse para capturar ADN (por ejemplo ADN genómico) y ARN, en aras de la simplicidad, la invención se describirá en detalle en el contexto de la captura de ARN, que forma un aspecto preferido de la invención. Sin embargo, más abajo se presentan modalidades particulares en relación con la captura de ADN.

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 el método anterior implica extender las sondas de captura mediante el uso de las moléculas de ácidos nucleicos hibridadas a las sondas de captura (es decir, "capturadas" por las sondas de captura) como plantillas de extensión para producir sondas extendidas para de esta manera etiquetar espacialmente los ácidos nucleicos del espécimen biológico. Por lo tanto, en el contexto del ARN, este aspecto de la etapa (c) puede verse como la generación de ADNc (particularmente ADNc etiquetado) a partir del ARN capturado, es decir, relacionado con la síntesis del ADNc. Esto implica una etapa de transcripción inversa (RT) 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 primera hebra, es decir, una sonda extendida. Como se describirá con más detalle más abajo, la síntesis de ADNc de segunda hebra puede tener lugar en una etapa separada, antes de la etapa de análisis de las sondas extendidas (por ejemplo, la secuencia de las sondas extendidas) o puede tener lugar como parte de la etapa de análisis. Por lo tanto, por ejemplo, la síntesis de la segunda hebra puede ocurrir en la primera etapa de amplificación de una molécula de ADNc de primera hebra liberada. En algunas modalidades, la síntesis de la segunda hebra puede ocurrir simultáneamente con la síntesis de la primera hebra o puede realizarse inmediatamente después de la reacción de síntesis de la primera hebra. Por ejemplo, la síntesis de la segunda hebra puede ocurrir simultáneamente con la reacción de síntesis de la primera hebra cuando se usa una reacción de cambio de plantilla para la síntesis de la segunda hebra. Las reacciones de cambio de plantilla se describen en detalle más abajo.

Por lo tanto, en algunas modalidades (es decir, cuando el método se usa para capturar ARN), la reacción de extensión se realiza mediante el uso de una enzima transcriptasa inversa, es decir, la enzima polimerasa es una enzima transcriptasa inversa. La actividad de transcriptasa inversa deseada puede proporcionarse por una o más enzimas transcriptasas inversas distintas, en donde los ejemplos adecuados son: las enzimas M-MLV, MuLV, AMV, HIV, ArrayScript™, MultiScribe™, ThermoScript™ y SuperScript® I, II y III. Como se usa en la presente descripción, el término "transcriptasa inversa" incluye no solo enzimas de origen natural sino también todos estos derivados modificados, que incluyen además derivados de enzimas transcriptasas inversas de origen natural.

Las enzimas transcriptasas inversas particularmente preferidas para su uso en los métodos de la presente solicitud incluyen enzimas transcriptasas inversas de M-MLV, MuLV, AMV y HIV y derivados, por ejemplo, derivados de secuencia modificada, o mutantes de las mismas.

Los derivados de secuencia modificada o mutantes de las enzimas transcriptasas inversas de M-MLV, MuLV, AMV y HIV incluyen mutantes que retienen al menos parte de la actividad funcional, por ejemplo, transcriptasa inversa, de la secuencia de tipo silvestre. Las mutaciones pueden afectar el perfil de actividad de las enzimas, por ejemplo, mejorar o reducir la velocidad de polimerización, en diferentes condiciones de reacción, por ejemplo, temperatura,

concentración de plantilla, concentración de cebador, etc. Las mutaciones o modificaciones de secuencias también pueden afectar la actividad de la RNasa y/o la termoestabilidad de la enzima. La enzima transcriptasa inversa puede proporcionarse como parte de una composición que comprende otros componentes, por ejemplo, componentes estabilizantes, que aumentan o mejoran la actividad de la enzima transcriptasa inversa, tal como inhibidor(es) de RNasa, inhibidores de la síntesis de ADN dependiente del ADN, por ejemplo, actinomicina D. Muchos derivados de secuencia modificada o mutantes de enzimas transcriptasas inversas, por ejemplo de M-MLV, y composiciones que comprenden enzimas no modificadas y modificadas se conocen en la técnica y están disponibles comercialmente, por ejemplo, enzimas ArrayScript™, MultiScribe™, ThermoScript™ y SuperScript® I, II, III y IV, y todas estas enzimas se consideran útiles en los métodos de la invención.

Por lo tanto, en una modalidad, la enzima transcriptasa inversa preferida está codificada por una secuencia de nucleótidos que comprende la SEQ ID NO: 1 (transcriptasa inversa de M-MLV) o una secuencia de nucleótidos que tiene al menos 80, 85 o 90 % de identidad de secuencia con la misma. Alternativa o adicionalmente, la enzima transcriptasa inversa preferida comprende la secuencia polipeptídica de SEQ ID NO: 2 o una secuencia polipeptídica que tiene al menos 80, 85 o 90 % de identidad de secuencia con la misma. Por lo tanto, la enzima transcriptasa inversa puede ser una variante o derivado natural de la enzima transcriptasa inversa de M-MLV, por ejemplo, de otra cepa de M-MLV.

Preferentemente, dicha secuencia de nucleótidos o polipeptídica es al menos 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % idéntica a la secuencia con la que se compara.

La identidad de secuencia de moléculas de nucleótidos puede determinarse, por ejemplo, mediante la búsqueda FASTA mediante el uso de paquetes de GCG, con valores predeterminados y un factor pam variable, y la penalización por creación de brechas se establece en 12,0 y la penalización por extensión de brechas se establece en 4,0 con una ventana de 6 nucleótidos.

Preferentemente tales moléculas de ácidos nucleicos relacionadas con la identidad de secuencia son funcionalmente equivalentes a la molécula de ácido nucleico que se expone en la SEQ ID NO: 1. Tales moléculas de ácidos nucleicos funcionalmente equivalentes pueden tomar la forma de derivados y se consideran funcionalmente equivalentes si codifican polipéptidos que se considerarían equivalentes funcionales de acuerdo con las pruebas de actividad de la transcriptasa inversa conocidas en la técnica. Los equivalentes funcionales preferidos son aquellos que codifican los polipéptidos preferidos como se estableció anteriormente.

La identidad de secuencia de moléculas polipeptídicas puede determinarse mediante, por ejemplo, mediante el uso del banco de datos de secuencia de proteína SWISS-PROT mediante el uso de FASTA pep-cmp con un factor fam variable, y la penalización por creación de brechas se establece en 12,0 y la penalización por extensión de brechas se establece en 4,0, y una ventana de 2 aminoácidos. Preferentemente, dicha comparación se realiza en la longitud completa de la secuencia, pero puede realizarse en una ventana de comparación más pequeña, por ejemplo, menos de 600, 500, 400, 300, 200, 100 o 50 aminoácidos contiguos.

Preferentemente tales polipéptidos relacionados con la identidad de secuencia son funcionalmente equivalentes al polipéptido que se expone en la SEQ ID NO: 2. Como tal, el polipéptido con una secuencia como se expone en la SEQ ID NO: 2 puede modificarse sin afectar la secuencia del polipéptido.

Las modificaciones que no afectan la secuencia del polipéptido incluyen, por ejemplo, modificación química, que incluye por desglicosilación o glicosilación. Tales polipéptidos pueden prepararse mediante modificación posterior a la síntesis/aislamiento del polipéptido sin afectar la funcionalidad, por ejemplo, cierta glicosilación, metilación, etc. de residuos particulares.

Como se hace referencia en la presente descripción, para lograr la "equivalencia funcional", el polipéptido puede mostrar alguna eficacia aumentada o reducida en la actividad de la transcriptasa inversa con relación a la molécula original (es decir, la molécula de la que se derivó, por ejemplo, por sustitución de aminoácidos), pero preferentemente es tan eficiente o más eficiente. En algunas modalidades, parte de la actividad enzimática nativa puede reducirse o inhibirse en el polipéptido funcionalmente equivalente, por ejemplo, la actividad de la RNasa. Por lo tanto, la equivalencia funcional se refiere a un polipéptido que tiene actividad de transcriptasa inversa capaz de extender los dominios de captura de las sondas de captura en una reacción de extensión con plantilla de ARN. Esto puede probarse mediante la comparación de la actividad de la transcriptasa inversa del polipéptido derivado con relación al polipéptido del cual se deriva de manera cuantitativa, como se describió anteriormente. El derivado es preferentemente al menos 30, 50, 70 o 90 % tan efectivo como el polipéptido original en los métodos de la invención.

Las proteínas funcionalmente equivalentes que se relacionan con o se derivan de la proteína de origen natural, pueden obtenerse mediante la modificación de la secuencia de aminoácidos nativa mediante sustitución, adición y/o delección de uno o múltiples aminoácidos (siempre que satisfagan los requisitos de identidad de secuencia mencionados anteriormente), pero sin destruir la función de la molécula. Preferentemente, la secuencia nativa tiene menos de 20 sustituciones, adiciones o delecciones, por ejemplo, menos de 10, 5, 4, 3, 2 o 1 de tales modificaciones. Tales proteínas están codificadas por "moléculas de ácidos nucleicos funcionalmente equivalentes" que se generan mediante la sustitución, adición y/o delección apropiada de una o más bases.

Como se usa en la presente descripción, el término "múltiples" 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, 100 000, 1 000 000, 10 000 000, 15 000 000 o más, etc. Por lo tanto, por ejemplo, el número de sondas de captura puede ser cualquier número entero en cualquier intervalo entre cualquiera de dos números mencionados anteriormente.

Como se describió anteriormente, las sondas de captura se inmovilizan sobre un sustrato sólido y, por lo tanto, el sustrato sólido puede verse por tanto como una matriz o que comprende una o más matrices, por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más matrices. Por ejemplo, un sustrato sólido puede comprender 20, 25, 30, 35, 40 o más matrices. Las matrices para su uso en el contexto del análisis de ácidos nucleicos en general, y el análisis de ADN en particular, se analizan y describen a continuación. Los detalles y modalidades específicos descritos en la presente descripción en relación con matrices y sondas de captura para su uso en el contexto del ARN, se aplican igualmente (cuando sea apropiado) a todas estas matrices, que incluyen aquellas para su uso con ADN. En este sentido, se apreciará sin embargo que se prevé que las matrices de tipo convencional que comprenden sondas de captura como se definen en la presente descripción puedan usarse en los métodos de la invención.

Por lo tanto, los métodos descritos en la presente descripción pueden utilizar matrices de ácidos nucleicos de alta densidad que comprenden "sondas de captura" como se define en la presente descripción para capturar y etiquetar transcritos de todas las células individuales dentro de un espécimen biológico, por ejemplo, muestra de tejido, tal como un corte, o "sección" de muestra de tejido delgado. Las muestras o secciones de tejido 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 analizan, por ejemplo mediante secuenciación de alto rendimiento. Los datos resultantes pueden correlacionarse con imágenes del espécimen biológico, por ejemplo, muestras de tejido a través de los llamados dominios posicionales, por ejemplo, secuencias de nucleótidos únicas dentro de las sondas de captura, por ejemplo, secuencias de código de barras o etiquetas de ID).

Las matrices o micromatrices de ácidos nucleicos de alta densidad son un componente central de los métodos de marcaje y etiquetado espacial del transcriptoma y el genoma descritos en la presente descripción. Una micromatriz es una tecnología en formato múltiple usada en biología molecular. Una micromatriz típica consiste en una serie ordenada de puntos microscópicos (elementos o sitios) de oligonucleótidos (millones de puntos pueden incorporarse en una única matriz y varias matrices pueden presentarse en un único sustrato sólido). La posición distinta (discreta) de cada punto de ácido nucleico (oligonucleótido) (cada especie de oligonucleótido/molécula de ácido nucleico) se conoce como una "elemento" o "sitio" (y por lo tanto en 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 (discreta) 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 "sonda" (o "indicador"). 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 más abajo, las sondas difieren de las sondas de micromatrices estándar.

Por lo tanto, el término "matriz" se refiere a una población de elementos o sitios de ácidos nucleicos sobre un sustrato sólido que pueden diferenciarse entre sí de acuerdo con la ubicación relativa. Diferentes moléculas de ácidos nucleicos que están en diferentes sitios o elementos de una matriz pueden diferenciarse entre sí de acuerdo con las ubicaciones de los sitios o elementos en la matriz. Un sitio o elemento individual de una matriz puede incluir una o más moléculas de un tipo particular (por ejemplo, especie de sonda de captura). Por ejemplo, un sitio o elemento puede incluir una única molécula de ácido nucleico que tiene una secuencia particular o un sitio puede incluir varias moléculas de ácidos nucleicos que tienen la misma secuencia.

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

En una modalidad preferida las sondas se inmovilizan al sustrato sólido, por ejemplo, matriz, mediante inmovilización química, es decir mediante un enlace covalente. Esta puede ser una interacción entre el sustrato sólido 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 un grupo funcional introducido en la sonda de captura es un grupo amina. Típicamente, la sonda a inmovilizar comprende un grupo amina funcional o se modifica químicamente para comprender un grupo amina funcional. Como se describió anteriormente, en algunas modalidades la sonda puede comprender un enlazador químico escindible que une la sonda al sustrato sólido, de manera que forma

un dominio escindible de la sonda de captura. 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 en el que puede inmovilizarse 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 sólido. El término "sustrato activado" se refiere a un material en el que se establecieron o habilitaron mediante procedimientos de modificación química, grupos funcionales químicos reactivos o que interactúan 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 sintetizarse directamente sobre el sustrato sólido. Los expertos en la técnica conocen métodos adecuados para un enfoque de este tipo. Por ejemplo, los elementos (es decir, grupos de sondas) pueden prepararse mediante un método de amplificación de fase sólida conocido como amplificación puente, como se describe en más detalle más abajo. Otros ejemplos incluyen 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.

Por lo tanto, el sustrato sólido puede ser cualquier sustrato adecuado conocido por el experto en la técnica, es decir, adecuado para su uso como una matriz. 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 el espécimen biológico, por ejemplo, muestra de tejido, y el sustrato. Se prefiere particularmente el lugar donde el sustrato es un plano, es decir plano, chip o portaobjetos. En algunas modalidades, el sustrato puede comprender varias estructuras tales como pocillos, perlas (u otras partículas) en o sobre un sustrato, proyecciones desde el sustrato, crestas sobre el sustrato o canales en un sustrato. En algunas modalidades, cada una de las estructuras de un sustrato, por ejemplo, pocillos o perlas, pueden tener una sonda de captura diferente. Pueden identificarse diferentes sondas de captura unidas a cada estructura de acuerdo con las ubicaciones de las estructuras en o sobre la superficie del sustrato. Las matrices ilustrativas en las que se ubican estructuras separadas sobre el sustrato incluyen, sin limitación, las que tienen perlas en los pocillos.

El término "sustrato sólido" se refiere típicamente a un soporte rígido que es insoluble en líquido acuoso y que permite un posicionamiento exacto y rastreado de las sondas sobre 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. El sustrato puede ser no poroso o poroso. Por lo tanto, el sustrato puede ser capaz opcionalmente de absorber un líquido (por ejemplo, debido a la porosidad) pero será típicamente lo suficientemente rígido como para que el sustrato no se hinche sustancialmente cuando se absorba el líquido y no se contraiga sustancialmente cuando el líquido se retire mediante secado. Un soporte sólido no poroso es generalmente impermeable a líquidos o gases. Los soportes sólidos ilustrativos incluyen, pero no se limitan a, vidrio y vidrio modificado o funcionalizado, plásticos (que incluyen acrílicos, poliestireno y copolímeros de estireno y otros materiales, polipropileno, polietileno, polibutileno, poliuretanos, TeflonTM, olefinas cíclicas, poliimidas, etc.), nailon, cerámica, resinas, Zeonor, sílice o materiales a base de sílice que incluyen silicio y silicio modificado, carbono, metales, vidrios inorgánicos, conjuntos de fibra óptica, y polímeros, tales como poliestireno, copolímeros de olefina cíclica (COC), polímeros de olefina cíclica (COP), polipropileno, polietileno y policarbonato. Un sustrato preferido previsto por la presente invención es un sustrato no poroso. En algunas modalidades, el sustrato sólido se ubica dentro de un aparato de celda de flujo.

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. 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 silinación o mediante un tratamiento con poliacrilamida.

Una serie de matrices estándar están disponibles comercialmente y tanto el número como el tamaño de los elementos pueden variar. En la presente invención, la disposición de los elementos puede alterarse para que corresponda al tamaño y/o densidad de las células presentes en diferentes especímenes biológicos, por ejemplo, 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, las 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 la superficie de la matriz en el contexto de la presente invención, por ejemplo, una matriz de perlas o celda de flujo de Illumina, en la que las sondas se producen mediante el uso de amplificación puente. Además de las matrices disponibles comercialmente, que pueden personalizarse en sí mismas, es posible fabricar matrices "domésticas" 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.

El término "celda de flujo" pretende referirse a un recipiente que tiene una cámara donde puede llevarse a cabo una reacción, una entrada para suministrar reactivos a la cámara y una salida para retirar reactivos de la cámara. En algunas modalidades la cámara se configura para la detección de la reacción que ocurre en la cámara. Por ejemplo, la cámara puede incluir una o más superficies transparentes que permiten la detección óptica de especímenes biológicos, moléculas marcadas ópticamente o similares en la cámara. Las celdas de flujo ilustrativas incluyen, pero no se limitan a las usadas en un aparato de secuenciación de ácido nucleico tal como celdas de flujo para las plataformas Genome Analyzer®, MiSeq®, NextSeq® o HiSeq® comercializadas por Illumina, Inc. (San Diego, CA); o para la plataforma de secuenciación SOLiD™ o Ion Torrent™ comercializada por Life Technologies (Carlsbad, CA). Las celdas de flujo y los métodos ilustrativos para su fabricación y uso también se describen, por ejemplo, en los documentos WO 2016/162309, WO 2014/142841, la publicación de solicitud de patente de Estados Unidos núm. 2010/0111768 y la patente de Estados Unidos núm. 8,951,781.

El término "matriz de perlas" se refiere a un soporte sólido que comprende una colección de perlas u otras partículas. Las partículas pueden suspenderse en una solución o pueden ubicarse en la superficie de un sustrato sólido. Los ejemplos de matrices que tienen perlas ubicadas en una superficie incluyen aquellas en donde las perlas se ubican en pocillos tales como una matriz BeadChip (Illumina Inc., San Diego CA), sustratos usados en plataformas de secuenciación de 454 LifeSciences (una subsidiaria de Roche, Basel Suiza) o sustratos usados en plataformas de secuenciación de Ion Torrent (una subsidiaria de Life Technologies, Carlsbad California). Otros soportes sólidos que tienen perlas ubicadas en una superficie se describen en las patentes de Estados Unidos núms. 6,266,459; 6,355,431; 6,770,441; 6,859,570; 6,210,891; 6,258,568; o 6,274,320; las publicaciones de solicitud de patente de Estados Unidos núm. 2009/0026082 A1; 2009/0127589 A1; 2010/0137143 A1; o 2010/0282617 A1 o la publicación PCT núm. WO 00/63437 y WO 2016/162309. Varias de las referencias anteriores describen métodos para unir sondas de ácido nucleico a perlas antes de cargar las perlas en o sobre un soporte sólido. Como tal, la colección de perlas puede incluir diferentes perlas, cada una tiene una sonda de captura única unida. Sin embargo, se entenderá que las perlas pueden prepararse de modo que incluyan cebadores universales, y las perlas pueden cargarse después sobre una matriz, lo que permite de esta manera la síntesis de sondas de captura en la matriz, por ejemplo mediante el uso de métodos de amplificación puente. Como se expuso anteriormente en la presente descripción, los soportes sólidos usados típicamente para matrices de perlas pueden usarse sin perlas. Por ejemplo, los ácidos nucleicos, tales como sondas de captura, pueden unirse directamente a los pocillos o al material de gel en los pocillos. Por lo tanto, las referencias anteriores son ilustrativas de materiales, composiciones o aparatos que pueden modificarse para su uso en los métodos y composiciones expuestos en la presente descripción.

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

Las sondas sobre un sustrato sólido, por ejemplo, matriz, pueden inmovilizarse, es decir, unirse o enlazarse, a la matriz preferentemente por medio del extremo 5' o 3', en dependencia de la matriz química de la matriz. Típicamente, para las matrices disponibles comercialmente, las sondas se unen por medio de un enlace 3', de esta manera se deja un extremo 5' libre. Sin embargo, están disponibles matrices que comprenden sondas unidas al sustrato por medio 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 sólido, por ejemplo una matriz, 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 mencionó anteriormente, en algunas

modalidades el enlazador químico es escindible, por ejemplo, enzimática y/o químicamente, de manera que forma el dominio de escisión de la sonda de captura.

Como se describirá con más detalle más abajo, las sondas de captura pueden inmovilizarse sobre, o interactuar con, el sustrato sólido, por ejemplo, matriz, directa o indirectamente. Por lo tanto, no es necesario que las sondas de captura se unan directamente a la matriz, pero pueden interactuar indirectamente, por ejemplo, mediante la unión a una molécula o entidad (por ejemplo, perla) que se une por sí misma directa o indirectamente al sustrato sólido, por ejemplo, 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 a la matriz 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 sólido, por ejemplo, una matriz.

El método de la invención puede comprender sondas que se inmovilizan por medio de su extremo 5' o 3'. Sin embargo, cuando la sonda de captura se inmoviliza directamente al sustrato sólido, por ejemplo, puede preferirse que se inmovilice de manera que el extremo 3' de la sonda de captura sea libre de extenderse, por ejemplo, se inmoviliza por su extremo 5'. En algunas modalidades, 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 ácidos nucleicos, por ejemplo, polimerización con plantilla catalizada por una polimerasa, por ejemplo, transcriptasa inversa.

Por lo tanto, en una modalidad, la matriz comprende sondas que se inmovilizan directamente por medio 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' extensible libre. Sin embargo, será evidente que, en algunas modalidades, las sondas de superficie pueden inmovilizarse directamente por medio de su extremo 5', de manera que se requiere la escisión del dominio de escisión para liberar el extremo 3' de la sonda de captura, es decir, para exponer el dominio de captura para su extensión. En un aspecto preferido de la invención, cuando el sustrato sólido, por ejemplo la matriz, comprende sondas de superficie, las sondas de captura se sintetizan *in situ* sobre la matriz.

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 cadenas principales no nucleotídicas. Sin embargo, en el contexto del análisis del transcriptoma, 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. Como se describe más abajo con más detalle, en el contexto del análisis del genoma, el dominio de captura de la sonda de captura debe ser capaz de unirse a los fragmentos de ADN, lo que puede comprender unirse a un dominio de unión que se ha añadido al ADN fragmentado. En algunas modalidades, el dominio de captura de la sonda de captura puede cebar una reacción de extensión de ADN (polimerasa) para generar ADN que es complementario a las moléculas de ADN capturadas.

En 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 sobre el sustrato de la matriz directamente, es decir, por su extremo 5', lo que da como resultado un extremo 3' extensible libre.

Las sondas de captura comprenden al menos tres 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) y un dominio de escisión, como se definió anteriormente. La sonda de captura puede comprender además un dominio de amplificación como se define adicionalmente más abajo. Cuando la sonda de captura se une indirectamente a la superficie de la matriz por medio 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. Por ejemplo, cuando la sonda de superficie se inmoviliza sobre el sustrato sólido por medio de su extremo 3', una secuencia complementaria de este tipo puede ser complementaria a un dominio posicional/de identificación y/o un dominio de amplificación y/o dominio de escisión en la sonda de captura. En otras palabras, el dominio posicional y/o el dominio de amplificación y/o el dominio de escisión pueden constituir la región o porción de la sonda de captura 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.

Alternativamente, cuando la sonda de superficie se inmoviliza sobre el sustrato sólido por medio de su extremo 5', una secuencia complementaria de este tipo puede ser complementaria a un dominio de escisión y/o dominio de captura y/o dominio posicional/de identificación y/o un dominio de amplificación, la sonda de captura, preferentemente el dominio de escisión. En otras palabras, el dominio de escisión y/o el dominio de captura y/o el dominio posicional/de identificación y/o un dominio de amplificación pueden constituir la región o porción de la sonda de captura 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.

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 mediante polimerización dependiente de plantilla. Sin embargo, será evidente que en algunas modalidades, por ejemplo, cuando la sonda de captura se inmoviliza directa o indirectamente por su extremo 3', el dominio de escisión puede ubicarse en el extremo 3' y el dominio de captura se ubica directa o indirectamente aguas arriba de (es decir en el extremo 5' de) el dominio de escisión. 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 del espécimen biológico, por ejemplo muestra de tejido, en contacto con el sustrato sólido, por ejemplo, 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 de poli-A del ARNm. Por lo tanto, en una modalidad preferida el dominio de captura comprende un oligonucleótido poli-T, es decir, una serie de residuos consecutivos de desoxitimidina unidos por enlaces fosfodiéster, que es capaz de hibridarse con la cola de 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 o sus combinaciones, en donde dicho oligonucleótido retiene la propiedad funcional de unirse a poli-A. En una modalidad particularmente preferida el dominio de captura, o más particularmente el elemento poli-T y/o -U 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 una modalidad adicional, el dominio de captura, o más particularmente el elemento poli-T y/o -U del dominio de captura comprende al menos 25, 30 o 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 oligonucleotídica aleatoria. Esto puede ubicarse, por ejemplo, 5' o 3' respecto a la secuencia de 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 el dominio de captura comprende una secuencia específica para un gen o grupo de genes diana particular, puede ser conveniente sintetizar las sondas de captura sobre el sustrato sólido directamente, por ejemplo, mediante el uso de reacciones de ligazón. En particular, puede ser útil sintetizar el dominio de captura de la sonda de captura sobre el sustrato sólido directamente.

Por ejemplo, puede proporcionarse un sustrato sólido que comprende sondas de captura inmovilizadas por su extremo 5' en el que el dominio de captura se reemplaza con un dominio universal que es común a todas las sondas en el sustrato sólido (tales sondas pueden verse como un oligonucleótido de dominio posicional). El sustrato se pone en contacto con una o más sondas auxiliares de ligazón que comprenden de 5' a 3':

(i) un dominio que es complementario a una secuencia en el 5' del dominio de captura; y

(ii) un dominio que es complementario a un dominio universal en las sondas inmovilizadas sobre el sustrato sólido,

en condiciones que permitan que la sonda auxiliar de ligazón se hibride con las sondas inmovilizadas sobre el sustrato sólido. El sustrato también se pone en contacto con una o más especies de sonda que contienen un dominio de captura (tales sondas pueden verse como un oligonucleótido de dominio de captura), en donde las sondas comprenden de 5' a 3':

- 5 (i) un dominio que es complementario a una secuencia en el 5' de la sonda auxiliar; y
- (ii) un dominio de captura que comprende una secuencia específica para un gen diana particular,

en condiciones que permitan que las sondas que comprenden el dominio de captura se hibriden con las sondas auxiliares.

10 Por último, el sustrato se pone en contacto con componentes (por ejemplo, una enzima de ligasa y tampones, sales, etc. apropiados) que facilitan la ligazón de las sondas que contienen el dominio de captura a las sondas inmovilizadas sobre el sustrato sólido mediante el uso de las sondas auxiliares de ligazón como plantilla de ligazón para formar sondas de captura. Por lo tanto, si una especie de sonda que contiene el dominio de captura se liga a las sondas inmovilizadas sobre el sustrato sólido, cada elemento de la matriz será capaz de capturar un único ácido nucleico diana del espécimen biológico, por ejemplo, para analizar la expresión de un único gen. Sin embargo, si más de una especie de sonda que contiene el dominio de captura se liga a las sondas inmovilizadas sobre el sustrato sólido (es decir, cada especie de sonda contiene un dominio de captura que comprende una secuencia específica para un gen diana diferente), cada elemento de la matriz será capaz de capturar múltiples ácidos nucleicos diana del espécimen biológico, por ejemplo, para analizar la expresión de un subconjunto de genes. Será evidente que las sondas auxiliares de ligazón, las sondas que contienen el dominio de captura y los componentes que facilitan la ligazón de las sondas pueden ponerse en contacto con el sustrato sólido por separado (por ejemplo, secuencialmente) o simultáneamente.

15 Aunque un sustrato sólido que comprende sondas de captura que son específicas para un único gen o grupos de genes puede producirse por cualquier medio adecuado, será evidente que el método descrito anteriormente proporciona un método conveniente para generar múltiples sustratos específicos de genes en base a una única plataforma de matriz. En este sentido, un sustrato común o universal que comprende sondas de captura en las que el dominio de captura se reemplaza con un dominio universal que es común a todas las sondas en el sustrato sólido puede fabricarse y usarse para generar múltiples sustratos, cada uno específico para un gen diana o grupo de genes diana particular. Tales sustratos pueden encontrar particular utilidad en ensayos de diagnóstico, por ejemplo, para analizar la expresión de genes o grupos de genes que se asocian con una o más enfermedades, por ejemplo, cáncer.

Por lo tanto, la etapa (a) del método descrito anteriormente puede comprender:

30 (1) proporcionar un sustrato sólido en el que se inmovilizan múltiples especies de sondas de manera que cada especie ocupe una posición distinta en el sustrato sólido y se orienta de modo que tenga un extremo 3' libre, en donde dicha sonda es para una reacción de ligazón con plantilla, en donde cada especie de dicha sonda comprende una molécula de ácido nucleico con de 5' a 3':

- (i) un dominio de escisión para liberar la sonda de captura de la superficie del sustrato sólido,
- 35 (ii) un dominio posicional que corresponde a la posición de la sonda de captura en el sustrato sólido, y opcionalmente

(iii) un dominio universal;

(2) poner en contacto dicho sustrato sólido con:

- (i) una sonda auxiliar de ligazón como se definió anteriormente;
- 40 (ii) una o más especies de sonda que contienen un dominio de captura como se definió anteriormente; y
- (iii) componentes que facilitan la ligazón de las sondas en (ii) a las sondas en (1) mediante el uso de la sonda auxiliar en (2)(i) como plantilla de ligazón,

45 en condiciones adecuadas, por ejemplo, condiciones que forman hibridación adecuada de los dominios complementarios de las sondas y la ligazón de las sondas en (ii) a las sondas en (1) para formar sondas de captura, que son para una reacción de extensión del cebador.

Las enzimas y condiciones adecuadas para realizar las reacciones de ligazón se describen en detalle más abajo y pueden aplicarse al método descrito anteriormente.

50 El dominio posicional (dominio o etiqueta de identificación de elementos) 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 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. Por ejemplo, el dominio posicional puede formar el

extremo 5' de la sonda de captura cuando el dominio de escisión es un reticulador escindible que une la sonda de captura al sustrato sólido. Adicional o alternatively, el dominio posicional puede formar el extremo 5' de la sonda de captura cuando la sonda de captura se inmoviliza al sustrato por medio de su extremo 3'.

Como se describió anteriormente, cada elemento (posición distinta) del sustrato sólido, por ejemplo, matriz, comprende un punto de una especie de sonda de ácido nucleico, en donde el dominio posicional en cada elemento (o grupos de elementos adyacentes) es único. Por lo tanto, una "especie" de sonda de captura se define con referencia a su dominio posicional; una única especie de sonda de captura tendrá el mismo 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 la sonda 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 sitio (o grupos de elementos o sitios adyacentes) del sustrato sólido, por ejemplo, matriz, tenga una sonda de captura de una única especie (específicamente, cada elemento o posición tiene una sonda de captura que tiene una etiqueta posicional idéntica, es decir, hay un único dominio posicional en cada elemento o posición o grupos de elementos o posiciones adyacentes). Cada especie tiene un dominio posicional diferente que identifica la especie. Sin embargo, cada miembro de una especie, en algunos casos, como se describe en más detalle en la presente descripción, puede tener un dominio de captura diferente, ya que el dominio de captura puede ser aleatorio o redundante o puede tener un componente aleatorio o redundante o, como se describió anteriormente, pueden producirse dominios de captura específicos para subconjuntos de genes en cada elemento. 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 (o grupos de elementos adyacentes) 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 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. Por lo tanto, por ejemplo, cuando los elementos comprenden sondas de captura que contienen múltiples dominios de captura, por ejemplo, dominios de captura, cada uno específico para un gen particular, cada elemento es capaz de capturar los mismos ácidos nucleicos diana.

El dominio posicional (o etiqueta) de la sonda de captura comprende la secuencia que es única para cada elemento (o grupo de elementos adyacentes) y actúa como un marcador posicional o espacial (la etiqueta de identificación). De esta manera, cada región o dominio del espécimen biológico, por ejemplo, muestra de tejido, por ejemplo, cada célula en el tejido, se identificará mediante resolución espacial en el sustrato sólido, por ejemplo, matriz, que enlaza el ácido nucleico, por ejemplo ARN (por ejemplo, los transcritos) de una determinada célula a una secuencia de dominio posicional única en la sonda de captura. En virtud del dominio posicional, una sonda de captura en la matriz puede correlacionarse 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).

Cualquier secuencia adecuada puede usarse como dominio posicional en las sondas de captura. Por una secuencia adecuada, se entiende que el dominio posicional no debe interferir con (es decir, inhibir o distorsionar) la interacción entre el ácido nucleico, por ejemplo ARN, del espécimen biológico, por ejemplo 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 el espécimen biológico, por ejemplo, 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 el espécimen biológico, por ejemplo, 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.

En una modalidad preferida el dominio posicional de cada especie de sonda de captura contiene una secuencia de código de barras. 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.

En algunas modalidades de la invención, como se describió en otra parte de la presente descripción, puede ser ventajoso amplificar las sondas extendidas, por ejemplo moléculas de ADNc etiquetadas, producidas mediante los métodos de la invención. Por lo tanto, en algunas modalidades, la sonda de captura comprende un dominio de amplificación, es decir, un dominio que es común a todas las sondas de captura y facilita la amplificación de todas las sondas extendidas en una única reacción. El dominio de amplificación 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 sonda de captura, del dominio posicional. En algunas modalidades, por ejemplo, cuando la sonda de captura se inmoviliza sobre la matriz por medio de su extremo 5', el dominio de amplificación puede ubicarse directa o indirectamente aguas abajo, es decir, más cerca del extremo 3' de la molécula de ácido nucleico de sonda de captura, del dominio de escisión, es decir, el dominio de amplificación puede ubicarse entre el dominio de escisión y el dominio posicional. Por lo tanto, en algunas modalidades el dominio de amplificación es directamente adyacente al dominio posicional, es decir, no hay secuencia intermedia entre el dominio posicional y el dominio de amplificación.

Sin embargo, en algunas modalidades, por ejemplo, cuando la sonda de captura se inmoviliza en la matriz por medio de su extremo 3', el dominio de amplificación puede ubicarse indirectamente aguas arriba, es decir, más cerca del extremo 5' de la molécula de ácido nucleico de sonda de captura, del dominio de escisión. Por lo tanto, en algunas modalidades donde la sonda de captura comprende un dominio de amplificación, el dominio puede formar el extremo 5' de la sonda de captura. Por ejemplo, el dominio de amplificación puede formar el extremo 5' de la sonda de captura cuando el dominio de escisión es un reticulador escindible que une la sonda de captura al sustrato sólido, por ejemplo, matriz. Alternativamente, el dominio de amplificación puede formar el extremo 5' de la sonda de captura cuando la sonda de captura se inmoviliza por medio de su extremo 5', por ejemplo, cuando el dominio de escisión se ubica en el extremo 3' de la sonda de captura.

Se apreciará que la amplificación de las sondas extendidas puede tener lugar sobre el sustrato sólido, por ejemplo, matriz, es decir, las sondas extendidas no necesitan retirarse o recolectarse (por ejemplo, mezclarse) del sustrato, por ejemplo, transferirse a un recipiente o contenedor separado, antes de realizar la etapa de amplificación. Sin embargo, en algunas modalidades puede ser conveniente retirar o recolectar (por ejemplo, mezclar) las sondas extendidas antes de la amplificación (o síntesis de la segunda hebra, como se describe más abajo). En este sentido, el método puede comprender una etapa de mezcla de las sondas extendidas de la etapa (c) (por ejemplo, para formar una mezcla de las sondas extendidas). La etapa de mezcla de las sondas extendidas puede verse como retirar, recolectar y/o transferir la solución que comprende las sondas extendidas del sustrato sólido, por ejemplo, matriz, a un recipiente o contenedor separado, por ejemplo un tubo de PCR.

El dominio de amplificación comprende una secuencia distinta a la que puede hibridarse un cebador de amplificación. El dominio de amplificación 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 sondas extendidas, por ejemplo, moléculas de ADNc etiquetadas.

Cualquier secuencia adecuada puede usarse como dominio de amplificación en las sondas de captura. 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, del espécimen biológico, por ejemplo 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, del espécimen biológico, por ejemplo 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 el espécimen biológico, por ejemplo, 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.

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, los dominios de escisión y los dominios de amplificación (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, del espécimen biológico, por ejemplo muestra de tejido, la liberación de las sondas de captura y la amplificación de las sondas extendidas 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 dominios de captura específicos para subconjuntos de genes específicos.

En modalidades donde la sonda de captura se inmoviliza sobre el sustrato de la matriz indirectamente, por ejemplo, por medio de hibridación con una sonda de superficie, la sonda de captura puede sintetizarse sobre la matriz como se describe más abajo.

Por ejemplo, las sondas de superficie pueden inmovilizarse sobre el sustrato de la matriz directamente mediante o en, por ejemplo, su extremo 3'. Cada especie de sonda de superficie es única para cada elemento (o grupo de elementos adyacentes) (posición distinta) de la matriz y es parcialmente complementaria a la sonda de captura, definida anteriormente.

Por lo tanto, en algunas modalidades, 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, del espécimen biológico, por ejemplo, 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 y un dominio (dominio de escisión complementario) que es complementario al dominio de escisión de la sonda de captura. 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. El dominio de escisión complementario se ubica directa o indirectamente aguas abajo (es decir, en el extremo 3') del dominio posicional complementario, es decir, puede haber una secuencia intermedia o enlazadora que separa el dominio de escisión complementario y el dominio posicional complementario. Cuando la sonda de captura comprende un dominio de amplificación, la sonda de superficie también puede contener un dominio (dominio de amplificación complementario) que es complementario al dominio de amplificación y se ubica directa o indirectamente aguas abajo (es decir, en el extremo 3') del dominio posicional complementario, es decir, puede haber una secuencia intermedia o enlazadora que separa el dominio de amplificación complementario y el dominio posicional complementario. Sin embargo, el dominio de amplificación complementario se ubicará aguas arriba (es decir, en el extremo 5') del dominio de escisión complementario, es decir, para garantizar que la sonda de captura extendida contenga el dominio de amplificación, el dominio posicional y la secuencia extendida, es decir, la secuencia en el extremo 3' del dominio de captura.

En modalidades donde la sonda de captura se sintetiza sobre la superficie del sustrato sólido, por ejemplo, matriz, las sondas de superficie del sustrato, por ejemplo matriz, pueden comprender un dominio (por ejemplo, dominio de amplificación complementario) en el extremo 3' de la sonda de superficie, es decir, directa o indirectamente aguas abajo del dominio de escisión complementario, que es complementario a un dominio (por ejemplo, dominio de amplificación) de la sonda de captura.

En algunas modalidades de la invención, la secuencia de la sonda de superficie muestra 100 % de complementariedad o identidad de secuencia con los dominios posicional, de escisión y amplificación (si está presente) y con la parte del dominio de captura que no se une al ácido nucleico, por ejemplo ARN, del espécimen biológico, por ejemplo, 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 de la invención, el dominio de escisión complementario comparte menos del 100 % de identidad de secuencia con el dominio de escisión de la sonda de captura.

Como se indicó anteriormente, en algunas modalidades de la invención, la sonda de captura puede sintetizarse o generarse sobre el sustrato de la matriz.

En una modalidad representativa, la matriz comprende sondas de superficie como se definió anteriormente. Los oligonucleótidos que corresponden al dominio de captura y al menos al dominio de escisión de la sonda de captura (que incluye opcionalmente además el dominio de amplificación) se ponen en contacto con el sustrato sólido, por ejemplo, matriz, y se dejan hibridar con los dominios complementarios de las sondas de superficie. El exceso de oligonucleótidos puede eliminarse mediante el lavado del sustrato sólido, por ejemplo, matriz, en condiciones de hibridación estándar. El sustrato resultante, por ejemplo, matriz, comprende sondas de hebra única parcial, en donde tanto los extremos 5' como 3' de la sonda de superficie son de doble hebra y el dominio posicional complementario (y opcionalmente también el dominio de amplificación complementario) es de hebra única. El sustrato sólido, por ejemplo, matriz, puede tratarse con una enzima polimerasa para extender el extremo 3' del oligonucleótido de dominio de escisión, de una manera dependiente de la plantilla, para sintetizar el dominio posicional (y opcionalmente el dominio de amplificación) 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 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.

El término "hibridación" o "se hibrida", como se usa en la presente descripción, se refiere a la formación de un dúplex entre secuencias de nucleótidos que son suficientemente complementarias para formar dúplex por medio 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 dúplex 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), del espécimen biológico, por ejemplo, muestra de tejido.

En una modalidad representativa adicional, la sonda de captura puede sintetizarse sobre el sustrato sólido, por ejemplo, matriz, 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. 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 de poli-T. En consecuencia, en una síntesis de este tipo, un oligonucleótido que corresponde al dominio de escisión de la sonda de captura (un oligonucleótido del dominio de escisión) puede ponerse en contacto con la matriz y dejarse hibridar con el dominio complementario de las sondas de superficie. El exceso de oligonucleótidos puede eliminarse mediante el lavado de la matriz en condiciones de hibridación estándar. La matriz resultante comprende sondas de hebra única parcial, en donde los extremos 3' de las sondas de superficie son de doble hebra y el dominio posicional complementario es de hebra única. La matriz puede tratarse con una enzima polimerasa para extender el extremo 3' del oligonucleótido del dominio de escisión, de una manera dependiente de la plantilla, para sintetizar el dominio posicional (y opcionalmente el dominio de amplificación) de la sonda de captura. El dominio de captura, por ejemplo, que comprende una secuencia de poli-T puede introducirse después mediante el uso de una transferasa terminal para añadir una cola de poli-T para generar la sonda de captura.

En aún otro ejemplo representativo, las sondas de captura pueden sintetizarse sobre el sustrato sólido, por ejemplo, matriz, mediante el uso de técnicas de amplificación, tales como la amplificación puente. Por ejemplo, una primera pluralidad de cebadores de ácido nucleico se une al soporte sólido, en donde los cebadores de ácido nucleico incluyen una primera secuencia de cebador universal que es común a dichos primeros cebadores de ácido nucleico y una segunda pluralidad de cebadores de ácido nucleico se une al soporte sólido, en donde los cebadores de ácido nucleico comprenden una segunda secuencia de cebadores universales que es común a dichos segundos cebadores de ácido nucleico. El soporte sólido que comprende dichos primer y segundo cebadores de ácido nucleico se pone en contacto con una pluralidad de diferentes sondas que comprenden secuencias que son complementarias al dominio de captura, dominio posicional, dominio de escisión y dominio de amplificación (si está presente), una primera secuencia de unión a cebador universal que se hibrida con los primeros cebadores universales en el soporte sólido y una secuencia idéntica a la segunda secuencia de cebador universal. Las sondas se hibridan con los primeros cebadores universales, que se extienden mediante el uso de las sondas de ácido nucleico como plantillas, lo que genera de esta manera una primera pluralidad de moléculas de ácidos nucleicos extendidas.

La primera pluralidad de moléculas de ácidos nucleicos extendidas comprende segundas secuencias de unión a cebadores universales que se hibridan con los segundos cebadores universales en el soporte sólido, que se extienden mediante el uso de la primera pluralidad de moléculas de ácidos nucleicos extendidas como plantillas, lo que genera de esta manera una segunda pluralidad de moléculas de ácidos nucleicos extendidas que son idénticas a las sondas en contacto con el sustrato sólido. La segunda pluralidad de moléculas de ácidos nucleicos extendidas comprende las primeras secuencias de unión a cebadores universales que se hibridan con los primeros cebadores universales en el soporte sólido y, por lo tanto, son capaces de servir como plantillas para la extensión de dichos primeros cebadores universales. La extensión continua de los cebadores da como resultado la formación de grupos en la superficie del sustrato sólido que comprende moléculas de ácidos nucleicos con secuencias que son idénticas o complementarias a la pluralidad de diferentes sondas en contacto con el soporte sólido.

Las sondas de ácido nucleico en contacto con el sustrato sólido comprenden un sitio o dominio de escisión entre la secuencia que es idéntica a la segunda secuencia de cebador universal y la secuencia que es complementaria al dominio de captura. La escisión de este sitio o dominio de escisión da como resultado la liberación del dominio que comprende la segunda secuencia de cebador universal. Será evidente que, en algunas modalidades, la liberación del dominio que comprende la segunda secuencia de cebador universal puede revelar el dominio de captura de las sondas de captura y liberar la segunda pluralidad de moléculas de ácidos nucleicos extendidas de la superficie de la matriz. En algunas modalidades, la liberación del dominio que comprende la segunda secuencia de cebador universal puede revelar el dominio posicional o dominio de amplificación de las sondas de captura, y liberar la segunda pluralidad de moléculas de ácidos nucleicos extendidas de la superficie de la matriz.

Por lo tanto, el método descrito anteriormente da como resultado grupos aleatorios de sondas de captura unidas al soporte sólido que pueden usarse para capturar ácidos nucleicos de un espécimen biológico mediante el uso del

método expuesto anteriormente en la presente descripción. Ventajosamente, la segunda pluralidad de moléculas de ácidos nucleicos extendidas liberadas puede eliminarse, por ejemplo, mediante el lavado del sustrato sólido, antes de realizar el método de la invención.

En este sentido, será evidente que el método descrito anteriormente puede modificarse para usar una pluralidad de sondas que comprenden el dominio de captura, el dominio posicional y el dominio de escisión, en lugar de sus complementos, por ejemplo, mediante la inclusión del sitio o dominio de escisión entre la primera secuencia de unión a cebador universal y el dominio de captura. De hecho, el orden de los dominios de las sondas en contacto con el sustrato sólido dependerá de si la sonda de captura se inmoviliza por medio de su extremo 3' o extremo 5' y el experto podría modificar fácilmente el orden del dominio de las sondas para generar un sustrato sólido, por ejemplo, matriz, que comprende sondas de captura adecuadas para su uso en los métodos de la invención. En particular, el método para sintetizar las sondas de captura en las matrices mediante el uso de amplificación puente puede ser particularmente útil para generar sustratos sólidos en forma de celdas de flujo, como se describió anteriormente.

Como se usa en la presente descripción, el término "aleatorio" puede usarse para referirse a la disposición espacial de las sondas de captura en el sustrato sólido, por ejemplo, matriz. En este sentido, hay al menos dos parámetros relacionados con la disposición espacial de las sondas de captura (es decir, elementos) en el sustrato sólido, por ejemplo, matriz, descrita en la presente descripción. El primer parámetro se refiere a la separación y ubicación relativa de los elementos y el segundo se refiere a la identidad o el conocimiento predeterminado de la especie particular de molécula que está presente en un elemento particular. En consecuencia, en algunas modalidades, los elementos de una matriz pueden separarse aleatoriamente de manera que los elementos adyacentes tengan una separación variable entre sí. Alternativamente, en algunas modalidades, la separación entre elementos puede ordenarse, por ejemplo, que forme un patrón regular tal como una rejilla rectilínea o una rejilla hexagonal. En otro aspecto, los elementos de una matriz pueden ser aleatorias con respecto a la identidad o el conocimiento predeterminado de la especie de sonda de captura que ocupa cada elemento independientemente de si la separación produce un patrón aleatorio o un patrón ordenado. Por ejemplo, en algunas modalidades expuestas en la presente descripción un sustrato sólido puede ponerse en contacto con una población de ácidos nucleicos en condiciones donde los ácidos nucleicos se unan en sitios que se ordenan con respecto a sus ubicaciones relativas pero que "se ubican aleatoriamente" con respecto al conocimiento de la secuencia para la especie de ácidos nucleico presente en algún sitio particular. La referencia a los ácidos nucleicos "que se distribuyen aleatoriamente" en ubicaciones sobre una superficie pretende referirse a la ausencia de conocimiento o ausencia de predeterminación con respecto a qué ácido nucleico se capturará en qué ubicación (independientemente de si las ubicaciones se disponen en un patrón ordenado o no).

Será evidente que las matrices aleatorias pueden producirse mediante el uso de cualquier medio conveniente. Sin embargo, en algunas modalidades, las matrices aleatorias se producen mediante la amplificación de sondas de captura en la superficie del sustrato sólido mediante el uso de cebadores que se unen a la superficie del sustrato sólido de manera que la ubicación de las sondas de captura no esté predeterminada. En algunas modalidades, las matrices aleatorias se producen mediante el uso de las llamadas matrices de perlas, en donde diferentes sondas de captura se unen a diferentes perlas en la matriz, que se distribuyen aleatoriamente en la superficie de la matriz de manera que cada perla, y por lo tanto cada especie de sonda de captura, se representa solo una vez sobre el sustrato sólido. En este sentido, será evidente que las perlas pueden disponerse en una disposición espacial ordenada, por ejemplo, en donde el sustrato sólido comprende pocillos en una disposición espacial ordenada que tiene dimensiones que alojan no más de una única perla. Sin embargo, tales matrices de perlas pueden verse como matrices aleatorias en la medida en que exista una ausencia de conocimiento o ausencia de predeterminación con respecto a dónde se ubicará cada sonda de captura en el sustrato sólido.

Como los métodos de la invención dependen de la capacidad de correlacionar la secuencia de una sonda extendida que comprende un dominio posicional con una ubicación particular en el sustrato sólido, es evidente que las matrices aleatorias deben analizarse antes de su uso en los métodos de la invención para determinar dónde se ubica cada especie de sonda de captura (es decir, donde cada dominio posicional) en el sustrato sólido. Esto puede lograrse mediante cualquier medio adecuado conocido en la técnica, tal como secuenciación o uso de sondas de hibridación.

Por lo tanto, en algunas modalidades, el método de la invención puede comprender una etapa de realizar una reacción de detección de ácido nucleico en el sustrato sólido para determinar las secuencias de dominio posicional de las sondas ubicadas aleatoriamente en el sustrato sólido. Por ejemplo, la reacción de detección de ácido nucleico puede ser una reacción de secuenciación para determinar las secuencias de dominio posicional de las sondas de captura ubicadas aleatoriamente sobre el sustrato sólido. En algunas modalidades, la reacción de detección de ácido nucleico puede ser una reacción de hibridación de la sonda decodificadora en el sustrato sólido para determinar las secuencias de dominio posicional de las sondas de captura ubicadas aleatoriamente en el sustrato sólido. Estas etapas pueden realizarse antes de la etapa (b) del método de la invención.

A manera de ejemplo, métodos combinatorios de hibridación, tales como los usados para la decodificación de matrices de perlas en formato múltiple (ver, por ejemplo, la patente de Estados Unidos núm. 8,460,865), puede usarse para decodificar una matriz "aleatoria". Tales métodos utilizan sondas decodificadoras de ácido nucleico marcadas que son complementarias a al menos una porción de la secuencia de dominio posicional (por ejemplo, secuencia de código de barras). Una reacción de hibridación puede llevarse a cabo mediante el uso de sondas decodificadoras que tienen marcadores conocidos de manera que la ubicación donde los marcadores terminan en el sustrato sólido identifica las

sondas de ácido nucleico de acuerdo con las reglas de complementariedad de ácidos nucleicos. En algunos casos, se usan mezclas de muchas sondas diferentes con marcadores distinguibles, lo que permite de esta manera una operación de decodificación en formato múltiple. El número de códigos de barras diferentes determinado en una operación de decodificación puede exceder el número de marcadores usados para la operación de decodificación. Por ejemplo, la decodificación puede llevarse a cabo en varias etapas donde cada etapa constituye la hibridación con una mezcla diferente de sondas decodificadoras. Las mismas sondas decodificadoras pueden estar presentes en diferentes mezclas, pero el marcador que está presente en cada sonda decodificadora puede diferir de mezcla a mezcla (es decir, cada sonda decodificadora está en un "estado" diferente cuando está en mezclas diferentes). Pueden usarse varias combinaciones de estos estados y etapas para expandir el número de dominios posicionales (por ejemplo, códigos de barras) que pueden decodificarse mucho más allá del número de marcadores distintos disponibles para la decodificación.

Las técnicas de secuenciación, tales como las técnicas de secuenciación por síntesis (SBS), son un método particularmente útil para determinar secuencias de dominio posicional (por ejemplo, código de barras) en matrices aleatorias. A manera de ejemplo, la SBS puede llevarse a cabo de la siguiente manera. Para iniciar un primer ciclo de SBS, uno o más nucleótidos marcados, ADN polimerasa, cebadores de SBS, etc., pueden ponerse en contacto con uno o más elementos en el sustrato sólido (por ejemplo, elemento(s) donde las sondas de captura se unen al soporte sólido). Pueden detectarse aquellos elementos donde la extensión del cebador de SBS provoca que se incorpore un nucleótido marcado. Opcionalmente, los nucleótidos pueden incluir un resto de terminación reversible que termina con una extensión del cebador adicional una vez que se ha añadido un nucleótido al cebador de SBS. Por ejemplo, un análogo de nucleótidos que tiene un resto terminador reversible puede añadirse a un cebador de manera que la extensión subsecuente no pueda ocurrir hasta que se suministre un agente de desbloqueo para eliminar el resto. Por lo tanto, para modalidades que usan terminación reversible, puede suministrarse un reactivo de desbloqueo al sustrato sólido (antes o después de que ocurra la detección). Los lavados pueden llevarse a cabo entre las diversas etapas de suministro. Después el ciclo puede repetirse n veces para extender el cebador en n nucleótidos, de esta manera se detecta una secuencia de longitud n. Los procedimientos de SBS, sistemas fluidicos y plataformas de detección ilustrativos que pueden adaptarse fácilmente para su uso con el método de la presente invención se describen, por ejemplo, en los documentos WO 91/06678, WO 04/018497, WO 07/123744; las patentes de Estados Unidos núms. 7,057,026, 7,329,492, 7,211,414, 7,315,019 o 7,405,281; y la solicitud de patente de Estados Unidos núm. 2008/0108082.

Por lo tanto, en algunas modalidades, las sondas de captura también pueden comprender un dominio que comprende un sitio de unión al cebador de secuenciación. En particular, el sitio de unión al cebador de secuenciación puede ubicarse entre el dominio de captura y el dominio posicional de la sonda de captura. Un dominio que comprende un sitio de unión al cebador de secuenciación ubicado entre el dominio de captura y el dominio posicional de la sonda de captura puede ser particularmente útil para la secuenciación de las sondas de captura en una matriz aleatoria como se describió anteriormente.

Por lo tanto, será evidente a partir de la descripción anterior que la sonda de captura puede comprender uno o más dominios universales, es decir dominios que son comunes a cada especie de sonda de captura. En particular, el(los) dominio(s) universal(es) pueden ser sitios de unión al cebador útiles para sintetizar las sondas de captura en el sustrato sólido, por ejemplo, matriz, y/o secuenciar la sonda de captura en matrices aleatorias.

También será evidente que la sonda de captura puede comprender dos sitios o dominios de escisión diferentes. Por ejemplo, un primer dominio de escisión puede ubicarse aguas arriba (es decir, en el extremo 5') del dominio posicional, por ejemplo, cuando la sonda de captura se inmoviliza por su extremo 5'. Sin embargo, en algunas modalidades, el primer dominio de escisión puede ubicarse aguas abajo (es decir, en el extremo 3') del dominio de captura, por ejemplo, cuando la sonda de captura se inmoviliza por su extremo 3'. Un segundo dominio de escisión puede localizarse entre un primer sitio de unión al cebador universal usado en la síntesis de la sonda de captura en el sustrato sólido, por ejemplo, matriz, y el dominio de captura o complemento del mismo. Los sitios o dominios de escisión deben ser reactivos a diferentes reacciones de escisión (por ejemplo, contener diferentes sitios de reconocimiento de enzimas de escisión) de manera que cada uno pueda escindirse selectivamente sin escindir necesariamente el otro. Por ejemplo, el segundo sitio o dominio de escisión puede escindirse antes de la etapa (c) (es decir, antes de liberar y extender la sonda de captura), de esta manera se separa o retira el primer sitio de unión al cebador universal de los otros dominios de la sonda de captura que permanecen unidos a un sustrato sólido. El primer sitio o dominio de escisión se escinde después de poner en contacto el sustrato sólido con el espécimen biológico, es decir simultáneamente con la reacción de extensión.

Un sustrato sólido típico, por ejemplo, una matriz, para su uso en los métodos de la invención puede contener múltiples puntos o "elementos". 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 o redundantes o pueden usarse dominios de captura específicos para un

subconjunto de genes. Por lo tanto, las sondas de captura dentro de un elemento pueden comprender diferentes secuencias aleatorias o redundantes o diferentes secuencias específicas de genes. El número y densidad de los elementos en el sustrato sólido, por ejemplo, matriz, determinará la resolución de la matriz, es decir, el nivel de detalle en el que puede analizarse el transcriptoma o genoma 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 la matriz dependerán de la naturaleza de la muestra de tejido y de la resolución requerida. Por lo tanto, si es conveniente determinar un transcriptoma o genoma solo para regiones de células dentro de un espécimen biológico, por ejemplo, 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. Como se mencionó anteriormente, en algunas modalidades, los elementos adyacentes pueden comprender las mismas sondas de captura, es decir, sondas de captura que contienen el mismo dominio posicional, para aumentar de manera efectiva el tamaño de un elemento, es decir, para disminuir la resolución de la matriz. Alternativamente, si es conveniente determinar un transcriptoma o genoma de células individuales dentro de una muestra, puede ser necesario usar el número máximo de elementos posibles, lo que necesitaría usar el tamaño de elemento más pequeño posible, por ejemplo, una matriz que comprende muchos elementos pequeños.

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, una 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, 4 500 000, 5 000 000, 6 000 000, 7 000 000, 8 000 000, 9 000 000, 10 000 000, 12 000 000 o 15 000 000 de elementos. Como se indicó anteriormente, el tamaño del elemento y/o la distancia promedio entre los elementos (es decir, el paso de los elementos) pueden disminuirse y esto puede permitir que se aloje 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.

Se establece bien en la técnica que un sustrato sólido puede comprender más de una "matriz", lo que permite de esta manera el análisis en formato múltiple de especímenes biológicos en un único sustrato. Por lo tanto, en algunas modalidades, el sustrato sólido comprende 2 o más matrices, por ejemplo 3, 4, 5, 6, 7, 8, 9, 10 o más matrices, por ejemplo 12, 15, 20, 25 o más matrices.

Por lo tanto, en algunas modalidades de la invención, el área de cada elemento puede ser de aproximadamente 0,1 μm², 0,5 μm², 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.

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

El término "especimen biológico" pretende referirse a una o más células, tejidos, organismos o porciones de los mismos. Será evidente que un espécimen biológico de cualquier organismo podría usarse en los métodos de la invención, por ejemplo, planta, animal u hongo. El método de la invención permite la captura de cualquier ácido nucleico, por ejemplo ARNm, moléculas, que están presentes en células que son capaces de realizar la transcripción y/o traducción. El método de la invención es particularmente adecuado para aislar y analizar el transcriptoma o genoma de las células dentro de un espécimen biológico, por ejemplo, una muestra de tejido, en donde la resolución espacial de los transcriptomas o genomas 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 o genoma 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 fijo, por ejemplo, una muestra de sangre). Tales células individuales, aunque no se fijan necesariamente a una determinada posición en un tejido, se aplican sin embargo a una determinada posición en la 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 única o independiente del transcriptoma o genoma de células individuales.

Por lo tanto, el espécimen biológico 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. 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).

El modo de preparación del espécimen biológico, por ejemplo, muestra de tejido, y cómo se manipula la muestra resultante puede afectar el análisis transcriptómico o genómico de los métodos de la invención. Además, varias especímenes biológicos, por ejemplo, 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 un espécimen biológico, 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 un espécimen biológico, por ejemplo 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 del espécimen biológico, por ejemplo 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 el espécimen biológico, por ejemplo, muestra de tejido, tenga un grosor del diámetro de una célula o menos; pueden usarse especímenes biológicos más gruesos, por ejemplo, muestras de tejido, si se desea. Por ejemplo, pueden usarse secciones de criostato, que pueden tener, por ejemplo, 10-20 µm de grosor.

La espécimen biológico, por ejemplo 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 el espécimen biológico, por ejemplo 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 un espécimen biológico, por ejemplo, una muestra de tejido para su uso en la invención, un 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 °C y preferentemente menos de -25, -30, -40, -50, -60, -70 o -80 °C. La muestra de tejido congelado puede seccionarse, es decir, cortarse finamente, sobre el sustrato sólido, por ejemplo, matriz, superficie 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 °C y preferentemente menos de -20 o -25 °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 de formalina e incrustación en parafina (FFPE), que están bien establecidos en la técnica. Después de la fijación de una muestra de tejido y la incrustación en un bloque de parafina o resina, la muestra de tejido puede cortarse en secciones, es decir, cortarse finamente, sobre el sustrato sólido, por ejemplo, una matriz. Como se indicó anteriormente, pueden usarse otros fijadores y/o materiales de incrustación.

Será evidente que será necesario tratar el espécimen biológico, por ejemplo, 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 establecida en la técnica, por ejemplo, mediante la incubación de la muestra (en la superficie del sustrato sólido, por ejemplo, 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, el espécimen biológico, por ejemplo muestra de tejido, es decir, la sección de la muestra de tejido en contacto con el sustrato sólido, 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.

Aunque el espécimen biológico puede prepararse para su uso en los métodos de la invención mediante el uso de cualquier técnica o método de preparación adecuados, en algunas modalidades particularmente preferidas, las moléculas de ácidos nucleicos del espécimen (particularmente moléculas de ARN) no se modifican antes del contacto con el sustrato sólido (por ejemplo, matriz). Por ejemplo, en algunas modalidades preferidas las moléculas de ácidos nucleicos (por ejemplo ARN) del espécimen biológico no se tratan para mejorar o facilitar su interacción con las sondas de captura antes de ponerlas en contacto con el sustrato sólido. En particular, en algunas modalidades, las moléculas de ácidos nucleicos (por ejemplo, ARN) del espécimen biológico no se ponen en contacto ni se hibridan con ácidos nucleicos (por ejemplo, aptámeros, oligonucleótidos o etiquetas de ácidos nucleicos) que facilitan la interacción entre las moléculas de ácidos nucleicos del espécimen biológico y las sondas de captura) antes de poner en contacto el espécimen biológico con el sustrato sólido. Por ejemplo, en algunas modalidades, el espécimen biológico no se pone en contacto ni se hibrida con ácidos nucleicos que funcionan como un intermediario entre el ácido nucleico diana del espécimen biológico y las sondas de captura, es decir, que funcionan para unir o enlazar indirectamente el ácido nucleico diana del espécimen biológico a la sonda de captura, antes de poner en contacto el espécimen biológico con el sustrato (por ejemplo, matriz).

El grosor del espécimen biológico, por ejemplo, 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 del espécimen biológico, por ejemplo, sección de muestra de tejido, con el sustrato sólido, por ejemplo, después de la eliminación del material de incrustación, por ejemplo, desparafinización, el ácido nucleico, por ejemplo ARN, moléculas en el espécimen biológico, por ejemplo muestra de tejido, se unirá a las sondas de captura inmovilizadas en la matriz. En algunas modalidades puede ser ventajoso facilitar la hibridación del ácido nucleico, por ejemplo ARN, de moléculas 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 del espécimen biológico, por ejemplo, la sección de muestra 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 los especímenes biológicos, por ejemplo 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. Por ejemplo, puede ser útil permeabilizar el espécimen biológico, por ejemplo, muestra de tejido, para facilitar la interacción entre el ácido nucleico y las sondas de captura. Si la muestra de tejido no se permeabiliza lo suficiente, la cantidad de ácido nucleico capturada por las sondas de captura puede ser demasiado baja para permitir un análisis adicional. Por el contrario, si el espécimen biológico, por ejemplo, la muestra de tejido, es demasiado permeable, el ácido nucleico puede difundirse lejos de su origen en el espécimen biológico, por ejemplo, la muestra de tejido, es decir, el ácido nucleico puede capturarse con las sondas de captura puede no correlacionarse con exactitud con su distribución espacial original en el espécimen biológico, por ejemplo, la muestra de tejido. Por lo tanto, debe haber un equilibrio entre la permeabilización del espécimen biológico, por ejemplo, muestra de tejido, suficiente para permitir una interacción eficiente entre los ácidos nucleicos y las sondas de captura mientras se mantiene la resolución espacial de la distribución de ácido nucleico en el espécimen biológico, por ejemplo, muestra de tejido. Los métodos usados para fijar el espécimen biológico, por ejemplo, muestra de tejido, también pueden afectar la interacción entre el ácido nucleico del espécimen biológico, por ejemplo, muestra de tejido, y las sondas de captura.

Por lo tanto, en algunas modalidades, el método comprende además una etapa de permeabilizar el espécimen biológico. Esta etapa puede realizarse después de la etapa de poner en contacto el espécimen con el sustrato sólido. En algunas modalidades, la etapa de permeabilizar el espécimen biológico puede tener lugar antes o simultáneamente con la etapa de liberación y extensión de las sondas de captura.

Los métodos y agentes adecuados para permeabilizar y/o fijar especímenes biológicos, por ejemplo células y tejidos, se conocen bien en la técnica y cualquier método apropiado puede seleccionarse para su uso en los métodos de la

invención. Algunas proteasas son particularmente útiles para permeabilizar células en un espécimen biológico, por ejemplo, pepsina. Los fijadores particularmente útiles incluyen, por ejemplo, metanol.

Como las condiciones para el etiquetado espacial, por ejemplo, la detección localizada o espacial, de moléculas de ácidos nucleicos de un espécimen biológico, por ejemplo, muestra de tejido, en un sustrato sólido varían en dependencia del espécimen biológico, por ejemplo, muestra de tejido, se describe un intervalo típico de parámetros en la presente descripción. Por ejemplo, al poner en contacto un espécimen biológico, por ejemplo, sección de muestra de tejido, con el sustrato sólido, por ejemplo, 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. Preferentemente, el sustrato sólido, por ejemplo, matriz, 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. 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 el espécimen biológico, por ejemplo sección de muestra de tejido, se incuba en el sustrato, por ejemplo, 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 son comunes en la técnica. En una modalidad particularmente preferida, la sección de muestra de tejido se deja secar en la matriz a 37 °C durante 24 horas. Una vez que el espécimen biológico, por ejemplo sección de muestra de tejido, ha secado el sustrato, por ejemplo, matriz, puede almacenarse a temperatura ambiente antes de realizar la etapa combinada de escisión y extensión, por ejemplo, transcripción inversa. Se debe entender que si se deja secar el espécimen biológico, por ejemplo, sección de muestra de tejido, en la superficie del sustrato, por ejemplo, matriz, tendrá que rehidratarse antes de que pueda lograrse una manipulación adicional del ácido nucleico capturado, por ejemplo, la etapa de escisión de las sondas de captura del sustrato sólido y simultáneamente o subsecuentemente transcribir inversamente el ARN capturado.

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

En algunas modalidades puede ser ventajoso bloquear (por ejemplo, enmascarar o modificar) las sondas de captura antes de poner en contacto el espécimen biológico, por ejemplo, muestra de tejido, con el sustrato, por ejemplo, matriz, particularmente cuando el ácido nucleico en el espécimen biológico, por ejemplo, muestra de tejido, está sujeto a un proceso de modificación antes de su captura en la matriz. Específicamente, puede ser ventajoso bloquear o modificar el extremo 3' libre de la sonda de captura. En una modalidad particular, el ácido nucleico en el espécimen biológico, por ejemplo, muestra de tejido, por ejemplo ADN genómico fragmentado, puede modificarse de manera que pueda capturarse con la sonda de captura. Por ejemplo, y como se describe con más detalle más abajo, una secuencia adaptadora (que comprende un dominio de unión capaz de unirse al dominio de captura de la sonda de captura) puede añadirse al extremo del ácido nucleico, por ejemplo, ADN genómico fragmentado. Esto puede lograrse, por ejemplo, mediante la ligazón de un adaptador o extensión del ácido nucleico, por ejemplo, mediante el uso de una enzima para incorporar nucleótidos adicionales en el extremo de la secuencia, por ejemplo, una cola de poli-A. Es necesario bloquear o modificar las sondas de captura, particularmente el extremo 3' libre de la sonda de captura, antes de poner en contacto el espécimen biológico, por ejemplo muestra de tejido, con la matriz para evitar la modificación de las sondas de captura, por ejemplo, para evitar la adición de una cola de poli-A al extremo 3' libre 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 del espécimen biológico, por ejemplo, muestra de tejido, que ocurre después de que el espécimen biológico, por ejemplo muestra de tejido, se ha puesto en contacto con la matriz. 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 que se usa para modificar el ácido nucleico del espécimen biológico, por ejemplo, ligazón o extensión, o los nucleótidos adicionales pueden eliminarse 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 de doble hebra parcial, 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.

Alternativa 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 ácidos nucleicos del espécimen biológico, por ejemplo, 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 un dominio de escisión como se definió anteriormente, 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 del espécimen biológico, por ejemplo, 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 (por ejemplo, como se describió anteriormente con respecto a la síntesis de las sondas en el sustrato sólido mediante el uso de amplificación puente) o podría añadirse a la sonda de captura *in situ* (es decir, como una modificación de una matriz 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. En este sentido, será evidente que el dominio de captura puede "bloquearse" mediante la inmovilización de las sondas de captura por su extremo 3'.

Una vez que el ácido nucleico del espécimen biológico, por ejemplo, muestra de tejido, por ejemplo, ADN genómico fragmentado, se ha modificado para permitirle hibridarse 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 el análisis de secuencias o la información del transcriptoma o genoma obtenida de cada elemento de la matriz con la región (es decir, un área o célula) del espécimen biológico, por ejemplo, muestra de tejido, el espécimen biológico se orienta en relación con los elementos de la matriz. En otras palabras, la muestra de tejido se coloca en la matriz de manera que la posición de una sonda de captura en la matriz puede correlacionarse con una posición en el espécimen biológico, por ejemplo, muestra de tejido. Por lo tanto, puede identificarse en qué lugar del espécimen biológico, por ejemplo, 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 el espécimen biológico, por ejemplo, muestra de tejido, corresponde la posición de cada especie de sonda de captura. Por ejemplo, esto puede hacerse en virtud de marcadores posicionales presentes en el sustrato sólido, por ejemplo, matriz, como se describe más abajo.

Convenientemente, pero no necesariamente, puede obtenerse una imagen del espécimen biológico, por ejemplo, muestra de tejido, después de su contacto con la matriz. Esto puede realizarse antes o después de que se procese el ácido nucleico del espécimen biológico, por ejemplo, muestra de tejido, por ejemplo, antes o después de la etapa combinada de liberación y extensión. En algunas modalidades, se obtiene una imagen del espécimen biológico, por ejemplo, muestra de tejido, antes de la liberación y extensión combinadas de las sondas de captura del sustrato sólido, por ejemplo, matriz. En una modalidad particularmente preferida, se obtiene una imagen del espécimen biológico, por ejemplo, muestra de tejido, después de la liberación y extensión combinadas de las sondas de captura del sustrato sólido, por ejemplo, matriz. No es necesario eliminar ningún tejido residual del sustrato sólido, por ejemplo, matriz, antes de la mezcla, es decir, recolección, de las moléculas liberadas (sondas extendidas), del sustrato ya que las sondas extendidas son capaces de difundir a través del tejido residual. Además, la etapa de procesar el ácido nucleico capturado, por ejemplo, liberación y extensión combinadas de las sondas de captura, puede actuar para eliminar algún tejido de la superficie de la matriz, por ejemplo, cuando se usa preparación de tejidos mediante congelación profunda. En tal caso, puede ser necesario separar el tejido residual recolectado cuando se mezclan las sondas extendidas del sustrato sólido, si se realiza, antes de la etapa de análisis de las sondas extendidas. Esto puede lograrse por cualquier medio adecuado, por ejemplo, centrifugación.

Por lo tanto, en algunas modalidades, el método puede comprender una etapa de separación del tejido de las sondas extendidas mezcladas.

Cuando la etapa de procesamiento del ácido nucleico de captura actúa para eliminar algún tejido de la superficie de la matriz, la obtención de imágenes del espécimen biológico, por ejemplo, muestra de tejido, puede tener lugar antes de la etapa de procesamiento, por ejemplo, la etapa combinada de liberación y extensión. En términos generales, las imágenes pueden tener lugar en cualquier momento después de poner en contacto el espécimen biológico, por ejemplo, muestra de tejido, con el sustrato sólido, pero antes de cualquier etapa que degrade o elimine el espécimen biológico, por ejemplo, muestra de tejido. Como se indicó anteriormente, esto puede depender del espécimen biológico, por ejemplo, muestra de tejido.

Ventajosamente, el sustrato sólido, por ejemplo, matriz, puede comprender marcadores para facilitar la orientación del espécimen biológico, por ejemplo, muestra de tejido, o la imagen del mismo en relación con los elementos del sustrato sólido, por ejemplo, matriz. Cualquier medio adecuado para marcar el sustrato sólido, por ejemplo, matriz, puede usarse de manera que sean detectables cuando se obtienen imágenes del espécimen biológico, por ejemplo, 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 del sustrato, por ejemplo, matriz. Preferentemente, el sustrato, por ejemplo, matriz, comprende al menos dos marcadores en posiciones distintas en la superficie de la matriz, 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. Los marcadores pueden proporcionarse en un patrón, por ejemplo, conforman un borde exterior de la matriz, por ejemplo, una hilera exterior completa de los elementos de una matriz. Pueden usarse otros patrones informativos, por ejemplo, líneas que seccionan la matriz. Esto puede facilitar la alineación de una imagen del espécimen biológico, por ejemplo, muestra de tejido, con una matriz, o de hecho generalmente en la correlación de los elementos de la matriz con el espécimen biológico, por ejemplo, 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 sólido, 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 marcado. Por ejemplo, la molécula de ácido nucleico marcada, o el ácido nucleico marcador, puede enlazarse o acoplarse 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. Una molécula de ácido nucleico marcadora de este tipo puede ponerse en contacto con el sustrato, por ejemplo, matriz, antes, simultáneamente con o después de teñir el espécimen biológico, por ejemplo, muestra de tejido, para visualizar u obtener imágenes del espécimen biológico, por ejemplo, muestra de tejido. Sin embargo, el marcador debe detectarse cuando se obtienen imágenes del espécimen biológico, por ejemplo, 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 el espécimen biológico, por ejemplo, muestra de tejido.

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

Sin embargo, será evidente que el uso de marcadores posicionales no es esencial para correlacionar el análisis de secuencias o la información del transcriptoma o genoma obtenida de cada elemento de la matriz con la región (es decir, un área o célula) del espécimen biológico, es decir, para alinear el espécimen biológico con los elementos de la matriz. En este sentido, puede ser posible alinear la topología del espécimen biológico con los datos generados a partir del análisis de las sondas extendidas. Por ejemplo, una grieta en una sección de muestra de tejido correspondería a la ausencia de sondas extendidas de elementos particulares en la matriz y esto puede ser suficiente para alinear el espécimen biológico con los elementos de la matriz. Además, en algunas modalidades, los elementos en la superficie del sustrato sólido son visibles, particularmente después de la tinción del espécimen biológico, mediante el uso de imágenes de alta resolución, lo que permite de esta manera que el espécimen biológico se alinee con los elementos de la matriz. Por lo tanto, en algunas modalidades, el elemento de la matriz puede verse como marcadores posicionales.

La etapa de obtención de imágenes del espécimen biológico, por ejemplo, muestra de tejido, puede usar cualquier medio histológico conveniente conocido en la técnica, por ejemplo luz, campo claro, campo oscuro, contraste de fase, fluorescencia, reflexión, interferencia, microscopía confocal o una de sus combinaciones. Típicamente, el espécimen biológico, por ejemplo, muestra de tejido, se tiñe antes de la visualización para proporcionar contraste entre las diferentes regiones, por ejemplo, células, del espécimen biológico, por ejemplo, muestra de tejido. El tipo de tinción usado dependerá del tipo de espécimen biológico, por ejemplo, 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 del espécimen biológico, por ejemplo, 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 del espécimen biológico, por ejemplo 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 algunas modalidades, la extensión de la sonda de captura comprende 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.

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 el espécimen biológico, por ejemplo, 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 ácidos nucleicos, 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, que puede verse como una forma de marcar o etiquetar indirectamente el ácido nucleico, por ejemplo, transcritos, del espécimen biológico, por ejemplo, muestra de tejido, que están en contacto con cada elemento de la matriz o dentro de los alrededores del elemento de manera que las sondas de captura liberadas sean capaces de unirse a los transcritos de una manera localizada. Como se mencionó anteriormente, cada especie de sonda de captura comprende un dominio posicional (etiqueta de identificación del elemento) que representa una secuencia única para cada elemento (o grupo de elementos adyacentes) de la matriz. Por lo tanto, todas las moléculas de ácidos nucleicos, por ejemplo, ADNc, sintetizadas en un elemento específico (o en los alrededores de un elemento específico) comprenderán la misma "etiqueta" de ácido nucleico.

Las moléculas de ácidos nucleicos, por ejemplo, ADNc, sintetizadas en cada elemento (o en los alrededores de un elemento específico) de la matriz pueden representar el genoma de, o genes expresados a partir de, la región o área del espécimen biológico, por ejemplo, muestra de tejido, en contacto con ese elemento (o en los alrededores de ese elemento), por ejemplo, un tejido o tipo de célula o grupo o subgrupo del mismo, y puede representar además genes expresados en condiciones específicas, por ejemplo, en un momento particular, en un entorno específico, en una etapa de desarrollo o en respuesta a estímulos, etc. Por lo tanto, el ADNc en cualquier elemento individual (es decir, capturado por una especie específica de sonda de captura) puede representar los genes expresados en una célula individual, o si el elemento está en contacto con (o en los alrededores de) la muestra en una unión celular, el ADNc puede representar los genes expresados en más de una célula. De manera similar, si una célula individual está en contacto con (o en los alrededores de) múltiples elementos, entonces cada elemento (es decir, una especie específica de sonda de captura extendida) puede representar una proporción de los genes expresados en dicha célula. De manera similar, en modalidades en las que el ácido nucleico capturado es ADN, cualquier elemento individual (es decir, una especie específica de sonda de captura prolongada) puede ser representativa del genoma de una célula individual o más de una célula. Alternativamente, el genoma de una célula individual puede representarse por múltiples elementos (es decir, múltiples especies de sonda de captura extendida).

La extensión de la sonda de captura, por ejemplo, la reacción de 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 la presente descripción. 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, las sondas extendidas, por ejemplo, ADNc, se tratan de modo que comprendan ADN de doble hebra. El tratamiento de las sondas extendidas para producir ADN de doble hebra puede lograrse en una única reacción para generar solo una segunda hebra de ADN, por ejemplo, ADNc, es decir, para producir moléculas de ADN de doble hebra sin aumentar el número de moléculas de ADN de doble hebra, o en una reacción de amplificación para generar múltiples copias de la segunda hebra, que puede estar en forma de ADN de hebra única (por ejemplo, amplificación lineal) o ADN de doble hebra, por ejemplo, ADNc (por ejemplo, amplificación exponencial).

La etapa de síntesis de ADN de segunda hebra, por ejemplo, ADNc, puede tener lugar sobre el sustrato sólido, por ejemplo, matriz, (es decir, antes de mezclar las moléculas extendidas), 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 primera hebra (es decir, sondas extendidas), por ejemplo, ADNc (la hebra que comprende, es decir que incorpora, la sonda de captura) puede mezclarse (por ejemplo, recolectarse) 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 un recipiente o contenedor separado, por ejemplo, tubo de PCR.

Cuando la síntesis de la segunda hebra tiene lugar sobre el sustrato sólido, por ejemplo, matriz, 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 (RNasa), por ejemplo RNasa 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 RNasa H puede usarse si se desea para aumentar la robustez de la eliminación del ARN.

Por ejemplo, en especímenes biológicos, por ejemplo muestras de tejido, que comprenden grandes cantidades de ARN, la etapa de generación del ADNc de doble hebra puede producir una cantidad suficiente de ADNc que puede secuenciarse directamente (después de la etapa de mezcla de las sondas extendidas). 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 sobre el sustrato sólido, es decir, antes de mezclar las sondas extendidas (por ejemplo, transferir las sondas extendidas a un recipiente o contenedor separado, por ejemplo, tubo de PCR), o después de haber mezclado las sondas extendidas, por ejemplo ADNc.

En otras modalidades puede ser necesario mejorar, es decir amplificar, las sondas extendidas, por ejemplo, ADNc sintetizado, para producir cantidades que sean suficientes para el análisis, por ejemplo, secuenciación de ADN. En esta modalidad, la primera hebra de las sondas extendidas, por ejemplo moléculas de ADNc, que comprenden además la sonda de captura de los elementos de la matriz, 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 de amplificación, entonces esto puede usarse para la subsecuente amplificación y análisis de secuencia 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 de la sonda de captura (en la sonda extendida, por ejemplo, la primera hebra de ADNc), el complemento del dominio posicional se incorporará en la segunda hebra del ADN, por ejemplo, moléculas de 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 primera hebra (es decir, sonda extendida) puede incubarse con cebadores aleatorios, por ejemplo, cebadores hexámeros, 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 primera hebra (es decir, sondas extendidas) 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" o la misma longitud que los fragmentos iniciales de ADN genómico), entonces el extremo 3' de las sondas extendidas, por ejemplo, moléculas de ADNc de 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 cadenas únicas tales como ARN ligasa T4 o Circuligase™ (Epicentre Biotechnologies).

Alternativamente, una segunda sonda auxiliar de síntesis de hebra (una molécula de ADN de doble hebra parcial capaz de hibridarse con el extremo 3' de la molécula de ADNc de primera hebra), puede ligarse al extremo 3' de la sonda extendida, por ejemplo, molécula de ADNc de primera hebra, mediante el uso de una enzima de ligazón de doble hebra tal como la ligasa de ADN de T4. Otras enzimas apropiadas para la etapa de ligazón se conocen en la técnica e incluyen, por ejemplo, ADN ligasa Tth, ADN ligasa Taq, ADN ligasa de *Thermococcus* sp. (cepa 9°N) (ADN ligasa 9°N™, New England Biolabs) y Ampligase™ (Epicentre Biotechnologies). La segunda sonda auxiliar de síntesis de hebra comprende además una secuencia específica a partir de la cual la síntesis de ADN de 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 a la sonda extendida, 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 de poli-A, en el extremo 3' de la sonda extendida, por ejemplo, moléculas de primera hebra 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 de doble hebra de "longitud completa", por ejemplo ADNc, (o síntesis de la segunda hebra de longitud máxima) están bien establecidos en la técnica.

En el método de la invención, la síntesis de la segunda hebra usa 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 de hebra única 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 la invención, la síntesis de la segunda hebra utiliza, 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 en el sustrato sólido (es decir, antes de mezclar las sondas extendidas). Preferentemente, la reacción de síntesis de la segunda hebra también se realiza en solución sobre el sustrato sólido.

En algunas modalidades puede ser ventajoso tratar las sondas extendidas de doble hebra para eliminar cualquier sonda de captura no extendida antes de la amplificación y/o el análisis, por ejemplo, el análisis de secuencias. Esto puede lograrse mediante cualquier medio adecuado conocido en la técnica, por ejemplo, mediante el uso de una enzima para degradar las sondas no extendidas, por ejemplo, una enzima exonucleasa o columnas de purificación. Por lo tanto, en algunas modalidades, el método comprende además una etapa de tratar las sondas extendidas para eliminar las sondas de captura no extendidas. Esto puede ser, por ejemplo, después de la etapa de mezcla de las sondas extendidas.

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 sondas extendidas, por ejemplo, la primera hebra de las moléculas de ADNc, cuando la sonda de captura 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 segunda hebra, por ejemplo, cebadores hexámeros aleatorios, cebadores poli-T, el cebador que es complementario a la sonda auxiliar de síntesis de la segunda hebra, pueden 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 de doble hebra resultante puede comprender un dominio de amplificación en o hacia cada extremo 5' de las moléculas de ADN de doble hebra, 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 sondas extendidas, por ejemplo, las moléculas de ADNc de primera hebra, pueden comprender un segundo dominio de amplificación. En el método de la invención, un dominio de amplificación se incorpora en las moléculas de ADNc de primera hebra mediante cambio de plantilla.

En modalidades donde la sonda de captura no 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 de doble hebra resultantes pueden modificarse para incorporar un dominio de amplificación en el extremo 5' del primer ADN, por ejemplo, hebra de ADNc (un primer dominio de amplificación) y, si no se incorpora en la segunda hebra de ADN, por ejemplo, etapa de síntesis de ADNc, en el extremo 5' del segundo ADN, por ejemplo, hebra de ADNc (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, ADN ligasa Tth, ADN ligasa Taq, ADN ligasa de *Thermococcus* sp. (cepa 9^oN) (ADN ligasa 9^oNTM, New England Biolabs), AmpligaseTM (Epicentre Biotechnologies) y ADN ligasa de T4. 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 de amplificación pueden añadirse a las sondas extendidas, es decir, moléculas de ADN, 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 presente descripción, tales dominios pueden añadirse antes o después de mezclar las sondas extendidas, es decir, de transferirlas a un recipiente o contenedor separado, por ejemplo, un tubo de PCR.

Será evidente a partir de la descripción anterior que todas las moléculas de ADN, por ejemplo, ADNc de un único sustrato sólido, por ejemplo, matriz, 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 del ADN, por ejemplo, moléculas de ADNc. En una modalidad la etapa de amplificación se realiza después de haber mezclado las sondas extendidas, por ejemplo moléculas de ADNc. En otras modalidades, la amplificación puede realizarse antes de mezclar las sondas extendidas, es decir, sobre el sustrato sólido, por ejemplo, matriz. Se conoce en la técnica que las reacciones de amplificación pueden llevarse a cabo en sustratos sólidos, tales como matrices, mediante la formación de pocillos en la superficie del sustrato para contener la solución o inmersión del sustrato en una solución. De hecho, 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 las matrices (por ejemplo, matrices de perlas de Illumina, celdas de flujo, etc. como se describió anteriormente)

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 las sondas extendidas comprenden una molécula de ácido nucleico de doble hebra parcial. Por ejemplo, las sondas extendidas pueden ser al menos parcialmente de doble hebra (por ejemplo híbrido ADN:ADN, ADN:ARN o ADN:ADN/ARN) en modalidades donde la sonda de captura se inmoviliza indirectamente sobre el sustrato de la matriz por medio de una sonda de superficie. La polimerasa de desplazamiento de hebra es necesaria para garantizar que la síntesis de ADNc de segunda hebra incorpore el complemento del dominio posicional (dominio de identificación de elementos) en el segundo ADN, por ejemplo, hebra de ADNc.

Como se mencionó anteriormente, puede verse que el método de la invención comprende una etapa de mezcla, por ejemplo, recolección o recuperación, de las sondas extendidas, por ejemplo moléculas de ADNc (que incluyen moléculas de ADNc de doble hebra) y/o sus amplicones. Como se indicó anteriormente, en el contexto de amplificación sobre el sustrato sólido, las sondas extendidas pueden incluir amplicones de las sondas extendidas, por ejemplo, ADNc.

Las sondas extendidas, es decir, ADN, por ejemplo moléculas de ADNc, o amplicones, que se han mezclado, 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 el espécimen biológico, por ejemplo, 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 dada en la matriz y, por lo tanto, en el espécimen biológico, por ejemplo, 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 las sondas extendidas o ácidos nucleicos derivados de estas, 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 sondas extendidas, por ejemplo moléculas de 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 sondas extendidas, por ejemplo, moléculas de ADNc, o amplicones de las mismas 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 sondas extendidas, es decir ADN, por ejemplo moléculas de ADNc.

El análisis de secuencias de las moléculas de ADN liberadas puede ser directo o indirecto. Por lo tanto, el sustrato de análisis de secuencias (que puede verse como la molécula que se somete a la etapa o proceso de análisis de secuencias) puede ser directamente la sonda extendida 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 sonda extendida o puede ser una molécula derivada de esta. Por ejemplo, un ADN de primera y/o segunda hebra, por ejemplo molécula de ADNc, 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). Alternativamente, la sonda extendida 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 sonda extendida.

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 sonda extendida, es decir, la primera hebra de ADN, por ejemplo ADNc, puede someterse a secuenciación. Puede ser necesario modificar el ADN de 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 al menos una porción de la sonda extendida, que incluye al menos parte de la secuencia de ácido nucleico capturada, por ejemplo ARN, y la secuencia del dominio posicional. La secuencia del dominio posicional (o etiqueta) identificará el elemento (o grupo de elementos adyacentes) desde el cual se liberó la sonda de captura y, por lo tanto, la ubicación o cercanía en el sustrato sólido 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 el espécimen biológico, por ejemplo muestra de tejido, para determinar el gen al cual corresponde. Al determinar qué región (por ejemplo, célula) del espécimen biológico, por ejemplo, muestra de tejido, estaba en contacto con el elemento, es posible determinar qué región del espécimen biológico, por ejemplo, muestra de tejido, expresaba dicho gen (o contenía el gen, por ejemplo, en el caso de la genómica espacial). 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 o genoma espacial del espécimen biológico, por ejemplo, muestra de tejido.

A manera de ejemplo representativo, los datos de secuenciación pueden analizarse para clasificar las secuencias en especies específicas de sonda de captura, es decir, 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 (o grupo de elementos adyacentes), 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 el espécimen biológico, por ejemplo, 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 sondas extendidas, por ejemplo moléculas de ADNc, 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 elemento del cual se liberó la sonda de captura y, por lo tanto, la ubicación o cercanía en el sustrato sólido en el que se capturó la molécula de ácido nucleico, por ejemplo, ARNm, (es decir, su ubicación en el sustrato sólido, por ejemplo, matriz) y el gen expresado. 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 genes o 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 genes o transcritos.

En algunas modalidades puede ser conveniente secuenciar la longitud completa de las sondas extendidas generadas por los métodos de la invención. Como se mencionó anteriormente, el cambio de plantilla es particularmente útil para generar moléculas de ADNc de doble hebra de longitud completa y, en algunas modalidades, puede ser conveniente o ventajoso secuenciar la longitud completa de estas moléculas. Cualquier técnica de secuenciación adecuada puede usarse para secuenciar las moléculas de ADN de longitud completa y más abajo se presentan ejemplos representativos.

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), como se describió anteriormente. 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 marcados con fluorescencia y 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 liberados y una señal electrónica proporcionalmente mayor.

Por lo tanto, está claro que están disponibles nuevos formatos de secuenciación, y con tiempos de funcionamiento más cortos como una de las principales características de esas plataformas, será evidente que otras tecnologías de secuenciación serán útiles en los métodos de la invención.

Una característica esencial de la presente invención, como se describió anteriormente, es una etapa de extensión de las sondas de captura para producir una hebra complementaria de las moléculas de ácidos nucleicos "capturadas" (es decir, las moléculas hibridadas al dominio de captura de 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. Como se mencionó anteriormente, la mezcla de reacción de transcripción inversa comprende además medios para liberar dichas sondas de captura de la superficie del sustrato sólido, por ejemplo, una enzima de escisión, tal como la enzima USER. La mezcla de reacción puede comprender otros componentes, tales como inhibidor(es) de RNasa, inhibidores de la síntesis de ADN dependiente del ADN, por ejemplo, actinomicina D. 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 , usualmente de aproximadamente 20 a 1000 μM . La mezcla de reacción puede incubarse en el sustrato sólido en condiciones adecuadas para la extensión de las sondas de captura y la escisión de dichas sondas del sustrato sólido. Típicamente, la mezcla de reacción puede incubarse sobre el sustrato sólido durante al menos 1 hora, por ejemplo, 1-2 horas o 2, 3, 4 o más horas. Sin embargo, tiempos de incubación más cortos pueden ser suficientes, por ejemplo, aproximadamente 30 minutos, tales como 10-30 minutos, 15-30 minutos, etc. 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 puede ser útil o ventajoso usar dNTP marcados en la reacción de extensión, por ejemplo, reacción de transcripción inversa. Los dNTP marcados se incorporarán en la molécula de ADN sintetizada, por ejemplo, molécula de ADNc, de esta manera marcarán las sondas extendidas. En una modalidad representativa, el dNTP marcado es un dNTP marcado con fluorescencia, por ejemplo, Cy3-dCTP.

Por lo tanto, en algunas modalidades, el método comprende además una etapa de marcaje de las sondas extendidas generadas en la etapa (c). En algunas modalidades, la etapa de marcaje puede ser simultánea con dicha etapa de extensión, es decir, mediante la incorporación de nucleótidos marcados en las sondas extendidas. En algunas modalidades, la etapa de marcaje puede ser subsecuente a la etapa de extensión, por ejemplo, mediante el uso de nucleótidos marcados en las etapas de síntesis y/o amplificación de la segunda hebra, de esta manera se incorporan nucleótidos marcados en la segunda hebra (es decir, la hebra complementaria de las sondas extendidas) y/o amplicones de la misma. Alternativamente, los marcadores (moléculas detectables) pueden incorporarse en el ADNc sintetizado mediante la unión a las moléculas, por ejemplo, por medio 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, hidroxietilbamidina, 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 marca la primera hebra de ADNc, puede ser preferible usar tinciones de ácidos nucleicos capaces de marcar 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.

Otras modalidades de marcaje y detección de moléculas de ADN marcadas en el contexto de métodos transcriptómicos y genómicos espaciales se describen en el documento WO2014/060483.

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 5-120 minutos, preferentemente 5-60, 5-45 o 5-30 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.

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 generadas, por ejemplo, ADNc, por ejemplo, para enriquecer la muestra para obtener una mejor representación de los ácidos nucleicos, por ejemplo transcritos, capturados del espécimen biológico, por ejemplo, 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 sondas extendidas anteriores, por ejemplo moléculas de ADNc, que pueden mezclarse, 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 sondas extendidas, por ejemplo moléculas de 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 de la invención, 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 sondas extendidas, por ejemplo moléculas de ADNc, pueden modificarse mediante la adición de dominios de amplificación a los extremos de las moléculas de ácidos nucleicos, que pueden implicar una reacción de ligazón.

Como se conoce en la técnica, las ligasas catalizan la formación de un enlace fosfodiéster entre los extremos 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 RNasa 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 RNasa 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.

A manera de ejemplo representativo, el método descrito puede comprender las siguientes etapas:

(a) proporcionar un sustrato sólido en el que se inmovilizan múltiples especies de sondas de captura de manera que cada especie ocupe una posición distinta en el sustrato sólido, en donde dichas sondas son para una reacción de extensión del cebador de transcripción inversa (preferentemente cuando dichas sondas se orientan sobre el sustrato sólido de modo que tengan un extremo 3' libre) y en donde cada especie de dicha sonda de captura comprende una molécula de ácido nucleico que comprende:

- (i) un dominio de escisión para liberar la sonda de captura de la superficie del sustrato sólido,
- (ii) un dominio posicional que corresponde a la posición de la sonda de captura en el sustrato sólido, y
- (iii) un dominio de captura;
- (b) poner en contacto dicho sustrato sólido con un espécimen biológico;

5 (c) obtener imágenes del espécimen biológico sobre el sustrato sólido;

(d) liberar dichas sondas de captura de la superficie del sustrato sólido en condiciones que permitan que el ARN del espécimen biológico se hibride con el dominio de captura en dichas sondas de captura y, simultáneamente y/o subsecuentemente, extender dichas sondas de captura mediante el uso del ARN hibridado con las sondas de captura como plantillas de extensión para producir sondas extendidas (moléculas de ADNc) de esta manera se etiqueta espacialmente el ARN del espécimen biológico;

10

(e) realizar la síntesis de ADNc de segunda hebra en las sondas extendidas;

y

(f) analizar las sondas extendidas, por ejemplo, la secuencia de las sondas extendidas (moléculas de ADNc).

A manera de un ejemplo representativo alternativo, el método descrito puede comprender las siguientes etapas:

15 (a) proporcionar un sustrato sólido en el que se inmovilizan múltiples especies de sondas de captura de manera que cada especie ocupe una posición distinta en el sustrato sólido, en donde dichas sondas son para una reacción de extensión del cebador de transcripción inversa (preferentemente cuando dichas sondas se orientan sobre el sustrato sólido de modo que tengan un extremo 3' libre) y en donde cada especie de dicha sonda de captura comprende una molécula de ácido nucleico que comprende:

- 20 (i) un dominio de escisión para liberar la sonda de captura de la superficie del sustrato sólido,
- (ii) un dominio posicional que corresponde a la posición de la sonda de captura en el sustrato sólido, y
- (iii) un dominio de captura;
- (b) poner en contacto dicho sustrato sólido con un espécimen biológico;
- (c) opcionalmente rehidratar la muestra de tejido;

25 (d) liberar dichas sondas de captura de la superficie del sustrato sólido en condiciones que permitan que el ARN del espécimen biológico se hibride con el dominio de captura en dichas sondas de captura y, simultáneamente y/o subsecuentemente, extender dichas sondas de captura mediante el uso del ARN hibridado con las sondas de captura como plantillas de extensión para producir sondas extendidas (moléculas de ADNc) de esta manera se etiqueta espacialmente el ARN del espécimen biológico.

30 (e) obtener imágenes del espécimen biológico sobre el sustrato sólido;

(f) amplificar las sondas extendidas; y

(g) analizar las sondas extendidas amplificadas, por ejemplo, la secuencia de las sondas extendidas amplificadas (moléculas de ADNc).

35 A manera de aún un ejemplo representativo adicional, el método descrito puede comprender las siguientes etapas:

(a) proporcionar un sustrato sólido en el que se inmovilizan múltiples especies de sondas de captura de manera que cada especie ocupe una posición distinta en el sustrato sólido, en donde dichas sondas son para una reacción de extensión del cebador de transcripción inversa (preferentemente cuando dichas sondas se orientan sobre el sustrato sólido de modo que tengan un extremo 3' libre) y en donde cada especie de dicha sonda de captura comprende una molécula de ácido nucleico que comprende:

40

- (i) un dominio de escisión para liberar la sonda de captura de la superficie del sustrato sólido,
- (ii) un dominio posicional que corresponde a la posición de la sonda de captura en el sustrato sólido, y
- (iii) un dominio de captura;
- (b) poner en contacto dicho sustrato sólido con un espécimen biológico;

45 (c) opcionalmente rehidratar la muestra de tejido;

(d) liberar dichas sondas de captura de la superficie del sustrato sólido en condiciones que permitan que el ARN del espécimen biológico se hibride con el dominio de captura en dichas sondas de captura y, simultáneamente y/o subsecuentemente, extender dichas sondas de captura mediante el uso del ARN hibridado con las sondas de captura como plantillas de extensión para producir sondas extendidas (moléculas de ADNc) de esta manera se etiqueta espacialmente el ARN del espécimen biológico;

(e) opcionalmente obtener imágenes del espécimen biológico sobre el sustrato sólido;

(f) realizar la síntesis de ADNc de segunda hebra en las sondas extendidas;

(g) agrupar opcionalmente las sondas extendidas de doble hebra de (f);

(h) amplificar las sondas extendidas de doble hebra de (f) o (g); y

(i) analizar las sondas extendidas amplificadas, por ejemplo, la secuencia de las sondas extendidas amplificadas (moléculas de ADNc).

Las etapas en los métodos descritos anteriormente pueden combinarse de cualquier manera adecuada. Se debe entender que la invención también abarca variaciones de estos métodos, por ejemplo, cuando se realiza la amplificación en el sustrato sólido. También se abarcan métodos que omiten la etapa de obtención de imágenes y métodos que comprenden una etapa de correlación de la información del análisis de secuencias obtenida del análisis de las sondas extendidas (que incluyen sondas extendidas de doble hebra y/o amplicones de las mismas) con una imagen de dicho espécimen biológico.

Como se mencionó anteriormente, aunque la invención se describe principalmente con referencia a la detección o análisis del ARN, y al análisis o detección del transcriptoma, se apreciará que los principios descritos pueden aplicarse de manera análoga a la detección o análisis del ADN en células y a estudios genómicos. Por lo tanto, visto en sentido más general, la invención puede verse como que es generalmente aplicable a la detección de ácidos nucleicos en general y en un aspecto más particular adicional, como que proporciona métodos para el análisis o detección de ADN. La información espacial puede ser valiosa también en un contexto genómico, es decir, la detección y/o análisis de una molécula de ADN con resolución espacial. Esto puede lograrse mediante el etiquetado genómico de acuerdo con la presente invención. Tales métodos de etiquetado espacial, por ejemplo, detección localizada o espacial, pueden ser útiles por ejemplo en el contexto de estudiar variaciones genómicas en diferentes células o regiones de un tejido, por ejemplo, comparar células o tejidos normales y enfermos (por ejemplo, células o tejidos normales frente a tumorales) o en el estudio de los cambios genómicos en la progresión de enfermedades, etc. Por ejemplo, los tejidos tumorales pueden comprender una población heterogénea de células que pueden diferir en las variantes genómicas que contienen (por ejemplo, mutaciones y/u otras aberraciones genéticas, por ejemplo, reordenamientos cromosómicos, amplificaciones/deleciones/inserciones cromosómicas, etc.). La detección de variaciones genómicas, o loci genómicos diferentes, en diferentes células de una manera localizada puede ser útil en un contexto de este tipo, por ejemplo, para estudiar la distribución espacial de las variaciones genómicas. Una utilidad principal de un método de este tipo sería en el análisis tumoral. En el contexto de la presente invención, puede prepararse un sustrato sólido, por ejemplo, matriz, que se diseña, por ejemplo, para capturar el genoma de una célula completa en las sondas de captura a partir de un elemento. Por lo tanto, pueden compararse diferentes células en el espécimen biológico, por ejemplo, muestra de tejido. Por supuesto, el método no se limita a tal diseño y pueden ser posibles otras variaciones, en donde el ADN se detecta de manera localizada y la posición del ADN capturado por las sondas de captura liberadas del sustrato sólido, por ejemplo, matriz, se correlaciona con una posición o ubicación en el espécimen biológico, por ejemplo, muestra de tejido.

La información del análisis de secuencias (por ejemplo, secuenciación) obtenida en la etapa de análisis de las sondas extendidas (que incluyen sondas extendidas de doble hebra y/o amplicones de las mismas) puede usarse para obtener información espacial sobre el ácido nucleico en la muestra. En otras palabras, la información del análisis de secuencias puede proporcionar información sobre la ubicación del ácido nucleico en la muestra. Esta información espacial puede derivarse de la naturaleza de la información del análisis de secuencias obtenida, por ejemplo, de una secuencia determinada o identificada, por ejemplo, puede revelar la presencia de una molécula de ácido nucleico 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 del análisis de secuencias. Sin embargo, como se describió anteriormente, la información espacial puede obtenerse convenientemente mediante la correlación de los datos de análisis de secuencias con una imagen del espécimen biológico, por ejemplo, muestra de tejido, y esto representa una modalidad preferida de la invención.

La reacción de extensión referida en la etapa (c) puede definirse como una reacción de extensión catalizada por polimerasa y actúa para adquirir una hebra complementaria de la "molécula de ácido nucleico capturada", es decir, la molécula de ácido nucleico que se hibrida con el dominio de captura de la sonda de captura, es decir, mediante la síntesis de la hebra complementaria que utiliza la sonda de captura como un cebador y el ácido nucleico capturado como una plantilla. En otras palabras, puede ser cualquier reacción de extensión del cebador llevada a cabo por cualquier enzima polimerasa. Por lo tanto, cuando la molécula de ácido nucleico capturada es ADN, por ejemplo ADN genómico, la polimerasa será una ADN polimerasa.

Por lo tanto, se proporciona además un método que comprende:

(a) proporcionar un sustrato sólido en el que se inmovilizan múltiples especies de sondas de captura de manera que cada especie ocupe una posición distinta en el sustrato sólido, en donde dichas sondas son para una reacción de extensión del cebador (preferentemente cuando dichas sondas se orientan sobre el sustrato sólido de modo que tengan un extremo 3' libre) y en donde cada especie de dicha sonda de captura comprende una molécula de ácido nucleico que comprende:

- (i) un dominio de escisión para liberar la sonda de captura de la superficie del sustrato sólido,
- (ii) un dominio posicional que corresponde a la posición de la sonda de captura en el sustrato sólido, y
- (iii) un dominio de captura;

(b) poner en contacto dicho sustrato sólido con un espécimen biológico; y

(c) fragmentar ADN en dicha muestra de tejido de espécimen biológico, en donde dicha fragmentación se lleva a cabo antes, durante o después de poner en contacto el sustrato sólido con el espécimen biológico en la etapa (b);

(d) liberar dichas sondas de captura de la superficie del sustrato sólido en condiciones que permitan que el ADN del espécimen biológico se hibride con el dominio de captura en dichas sondas de captura y, simultáneamente y/o subsecuentemente, extender dichas sondas de captura mediante el uso del ADN hibridado con las sondas de captura como plantillas de extensión para producir sondas extendidas de esta manera se etiqueta espacialmente el ADN del espécimen biológico;

(e) generar opcionalmente una hebra complementaria de dicho ADN etiquetado y/u opcionalmente amplificar dicho ADN etiquetado;

(f) mezclar opcionalmente las sondas extendidas (ADN etiquetado) de (d) o mezclar las sondas extendidas de doble hebra (ADN etiquetado) y/o amplicones de la etapa (e); y

(g) analizar las sondas extendidas (ADN etiquetado) de (d) o (f) y/o las sondas extendidas de doble hebra (ADN etiquetado) y/o amplicones de la etapa (e) o (f), por ejemplo, la secuencia de las sondas extendidas (ADN etiquetado) y/o sondas extendidas de doble hebra (ADN etiquetado) y/o amplicones.

El método puede incluir además una etapa de:

(h) correlacionar dicha información del análisis de secuencias con una imagen de dicho espécimen biológico, en donde se obtienen imágenes del espécimen biológico antes o después de la etapa (d), preferentemente antes de la etapa (d).

En el contexto de la genómica espacial, donde el ácido nucleico diana es ADN, puede preferirse la inclusión de etapas de obtención de imágenes y correlación de imágenes en algunas circunstancias.

En modalidades en las que se captura el ADN, el ADN puede ser cualquier molécula de ADN que pueda aparecer en una célula. Por lo tanto, puede ser ADN genómico, es decir, nuclear, ADN mitocondrial o ADN plasmídico, por ejemplo, ADN de cloroplasto. En una modalidad preferida, el ADN es ADN genómico.

Se debe entender que cuando la fragmentación se lleva a cabo después del contacto en la etapa (b), es decir después de que el espécimen biológico se coloca sobre el sustrato sólido, por ejemplo, matriz, la fragmentación ocurre antes de que el ADN se hibride con el dominio de captura. En otras palabras, los fragmentos de ADN se hibridan (o más particularmente, se dejan hibridar) con el dominio de captura en dichas sondas de captura.

Como se indicó anteriormente, en algunas modalidades se prefiere que las moléculas de ácidos nucleicos (diana) (particularmente ARN) en el espécimen biológico no se traten o modifiquen para mejorar o facilitar su interacción con las sondas de captura antes de poner en contacto el espécimen biológico con el sustrato sólido, por ejemplo, se ponen en contacto con o se hibridan con ácidos nucleicos que funcionan como un intermediario entre el ácido nucleico diana del espécimen biológico y las sondas de captura, es decir, que funcionan para unir o enlazar indirectamente el ácido nucleico diana del espécimen biológico a la sonda de captura. Sin embargo, ventajosamente, pero no necesariamente, en una modalidad particular de este aspecto de la invención, los fragmentos de ADN del espécimen biológico, por ejemplo muestra de tejido, pueden proporcionarse con un dominio de unión para permitir o facilitar su captura por las sondas de captura. En consecuencia, el dominio de unión es capaz de hibridarse con el dominio de captura de la sonda de captura. Por lo tanto, un dominio de unión de este tipo puede considerarse un complemento del dominio de captura (es decir, puede verse como un dominio de captura complementario), aunque no se requiere complementariedad absoluta entre los dominios de captura y unión, simplemente que el dominio de unión sea suficientemente complementario para permitir que tenga lugar una hibridación productiva, es decir, que los fragmentos de ADN en el espécimen biológico, por ejemplo muestra de tejido, sean capaces de hibridarse con el dominio de captura de las sondas de captura. La provisión de un dominio de unión de este tipo puede garantizar que el ADN en la muestra no se una a las sondas de captura hasta después de la etapa de fragmentación. El dominio de unión puede

proporcionarse a los fragmentos de ADN mediante procedimientos bien conocidos en la técnica, por ejemplo mediante la ligazón de secuencias adaptadoras o enlazadoras que pueden contener el dominio de unión. Por ejemplo, puede usarse una secuencia enlazadora con un extremo sobresaliente. El dominio de unión puede estar presente en la porción de hebra única de un enlazador de este tipo, de manera que después de la ligazón del enlazador a los fragmentos de ADN, la porción de hebra única que contiene el dominio de unión está disponible para la hibridación con el dominio de captura de las sondas de captura. Alternativamente y en una modalidad preferida, el dominio de unión puede introducirse mediante el uso de una enzima transferasa terminal para introducir una cola de polinucleótido, por ejemplo, una cola homopolimérica tal como un dominio poli-A. Esto puede llevarse a cabo mediante el uso de un procedimiento análogo al descrito anteriormente para introducir un segundo dominio de amplificación en el contexto de los métodos de ARN. Por lo tanto, en modalidades ventajosas puede introducirse un dominio de unión común. En otras palabras, un dominio de unión que es común a todos los fragmentos de ADN y que puede usarse para lograr la captura de los fragmentos en el sustrato sólido, por ejemplo, matriz. Como se describe más abajo, en algunas modalidades, el dominio de unión común puede introducirse en el ADN durante (es decir, simultáneamente con) la etapa de fragmentación del ADN, por ejemplo, cuando el ADN se fragmenta mediante el uso de una transposasa (por ejemplo, Tn5) el dominio de unión común puede ligarse simultáneamente a los fragmentos de ADN.

Cuando se lleva a cabo una reacción de adición de cola para introducir un dominio de unión (común), las sondas de captura en el sustrato sólido, por ejemplo, matriz, pueden protegerse de la reacción de adición de cola, es decir, las sondas de captura pueden bloquearse o enmascararse como se describió anteriormente. Esto puede lograrse, por ejemplo, mediante la hibridación de un oligonucleótido de bloqueo a la sonda de captura, por ejemplo, al extremo sobresaliente (por ejemplo, porción de hebra única) de la sonda de captura. Cuando el dominio de captura comprende una secuencia de poli-T, por ejemplo, un oligonucleótido de bloqueo de este tipo puede ser un oligonucleótido de poli-A. El oligonucleótido de bloqueo puede tener un extremo 3' bloqueado (es decir, un extremo incapaz de extenderse, o al que no puede adicionarse cola). Las sondas de captura también pueden protegerse, es decir bloquearse, mediante modificaciones químicas y/o enzimáticas, como se describió en detalle anteriormente.

Por lo tanto, en algunas modalidades de la invención, el método puede comprender una etapa de proporcionar a los fragmentos de ADN un dominio de unión que es capaz de hibridarse con el dominio de captura. Esta etapa puede realizarse después (o al mismo tiempo que) de la etapa de fragmentación del ADN pero antes de la etapa de liberación y extensión de las sondas de captura. En modalidades preferidas, la etapa de proporcionar a los fragmentos de ADN un dominio de unión que es capaz de hibridarse con el dominio de captura se realiza después si el espécimen biológico entra en contacto con el sustrato (por ejemplo, matriz), es decir, la etapa se realiza *in situ* en el sustrato.

En los métodos de detección de ADN establecidos anteriormente, será evidente que la etapa de extensión de la sonda de captura y la etapa opcional de generación de una copia complementaria del ADN etiquetado o de amplificar el ADN etiquetado, puede implicar o requerir el uso de una enzima polimerasa que desplaza la hebra. Por ejemplo, el ADN fragmentado hibridado a la sonda de captura puede ser ADN de doble hebra parcial y, por lo tanto, la extensión de la sonda de captura puede requerir el desplazamiento de una de las hebras de ADN. Las polimerasas que desplazan la hebra adecuadas se describieron anteriormente. En el contexto de producir hebras complementarias de la sonda extendida y/o amplicones de la sonda extendida, el uso de polimerasas de desplazamiento de hebra es para garantizar que el dominio posicional se copie en la copia o amplicón complementario. Este será particularmente el caso en el que la sonda de captura se inmovilice en el sustrato sólido, por ejemplo, una matriz, mediante hibridación con una sonda de superficie.

Por lo tanto, en una modalidad, el método de la invención puede usarse para determinar y/o analizar todo el genoma de una muestra de tejido, por ejemplo, el genoma 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 genoma. Por lo tanto, el método puede implicar determinar y/o analizar una parte o subconjunto del genoma, por ejemplo, un genoma parcial correspondiente a un subconjunto o grupo de genes o de cromosomas, por ejemplo, un conjunto de genes o cromosomas particulares o una región o parte particular del genoma, por ejemplo relacionado con una enfermedad o afección particular, tipo de tejido, etc. Por lo tanto, el método puede usarse para detectar o analizar secuencias genómicas o loci genómicos de tejido tumoral en comparación con tejido normal, o incluso dentro de diferentes tipos de células en una muestra de tejido. La presencia o ausencia, o la distribución o ubicación de diferentes variantes genómicas o loci en diferentes células, grupos de células, tejidos o partes o tipos de tejido pueden examinarse.

Visto desde otro aspecto, las etapas del método expuestas anteriormente pueden verse como que proporcionan un método para obtener información espacial con respecto a los ácidos nucleicos, por ejemplo, secuencias genómicas, variantes o loci de un espécimen biológico, por ejemplo, muestra de tejido. Dicho de otra manera, los métodos de la invención pueden usarse para el marcaje (o etiquetado) de genomas, particularmente genomas individuales o distribuidos espacialmente.

Visto alternativamente, el método de la invención puede verse como un método para la detección espacial de ADN en un espécimen biológico, por ejemplo, muestra de tejido, o un método para detectar ADN con resolución espacial, o para la determinación y/o análisis localizado o espacial de ADN en un espécimen biológico, por ejemplo, muestra de tejido. En particular, el método puede usarse para la detección o determinación localizada o espacial y/o análisis de genes o secuencias genómicas o variantes genómicas o loci (por ejemplo, distribución de variantes genómicas o loci) en un espécimen biológico, por ejemplo, muestra de tejido. La detección/determinación/análisis localizado/espacial

significa que el ADN puede localizarse en su posición o ubicación nativa dentro de una célula o tejido en el espécimen biológico, por ejemplo, muestra de tejido. Por lo tanto, por ejemplo, el ADN puede localizarse en una célula o grupo de células, o tipo de células en la muestra, o en regiones particulares de áreas dentro de un espécimen biológico, por ejemplo, muestra de tejido. Puede determinarse la ubicación o posición nativa del ADN (o en otras palabras, la ubicación o posición del ADN en el espécimen biológico, por ejemplo, muestra de tejido), por ejemplo, una variante genómica o locus.

La etapa de fragmentación de ADN en una muestra de tejido puede llevarse a cabo mediante el uso de cualquier procedimiento deseado conocido en la técnica. Por lo tanto, pueden usarse métodos físicos de fragmentación, por ejemplo, sonicación o tratamiento con ultrasonido. También se conocen métodos químicos. También pueden usarse métodos enzimáticos de fragmentación, por ejemplo con endonucleasas, por ejemplo enzimas de restricción o transposasas, por ejemplo Tn5. Nuevamente, los métodos y enzimas para esto se conocen bien en la técnica. En particular, una transposasa, tal como Tn5, puede dar como resultado ventajosamente la fragmentación simultánea del ADN y la ligazón de adaptadores (por ejemplo, que comprenden dominios de unión comunes) a los extremos 5' de ambas hebras de fragmentos de ADN, lo que puede facilitar la captura del ADN fragmentado por las sondas de captura, es decir, los adaptadores ligados a los extremos 5' del ADN pueden contener un dominio que es capaz de interactuar con (es decir, hibridarse con) el dominio de captura de las sondas de captura. La fragmentación puede realizarse antes durante o después de la preparación del espécimen biológico, por ejemplo, muestra de tejido, para colocarla sobre un sustrato sólido, por ejemplo, matriz, por ejemplo, preparación de una sección de tejido. Convenientemente, la fragmentación puede lograrse en la etapa de fijación del tejido. Por lo tanto, por ejemplo, la fijación con formalina dará como resultado la fragmentación del ADN. Otros fijadores pueden producir resultados similares.

En términos del detalle de preparación y uso del sustrato sólido, por ejemplo, matriz, en estos aspectos de la invención, se entenderá que la descripción y detalle dados anteriormente en el contexto de los métodos de ARN se aplican de manera análoga a los métodos de detección de ADN más establecidos en la presente descripción. Por lo tanto, todos los aspectos y detalles descritos anteriormente se aplican de manera análoga. Por ejemplo, la descripción de cebadores y reacciones de transcriptasa inversa, etc. puede aplicarse de manera análoga a cualquier aspecto de los cebadores de extensión, reacciones de ADN polimerasa, etc. referidos anteriormente. Igualmente, las referencias y a la síntesis de ADNc de la primera y segunda hebra pueden aplicarse de manera análoga a la molécula de ADN etiquetada y su complemento. Pueden usarse métodos de análisis de secuencias como se describió anteriormente.

A manera de ejemplo, el dominio de captura puede ser como se describió anteriormente para las sondas de captura. Un dominio de captura poli-T o que contiene poli-T puede usarse, por ejemplo, cuando los fragmentos de ADN se proporcionan con un dominio de unión que comprende una secuencia de poli-A.

Las sondas de captura/moléculas de ADN etiquetadas (es decir, las sondas extendidas que comprenden complementos de moléculas de ADN capturadas) pueden proporcionarse con dominios de amplificación como se describió anteriormente.

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 un esquema del diseño experimental descrito en el Ejemplo 2.

La Figura 2 muestra una imagen de microscopía de fluorescencia de una matriz en la que se ha sintetizado ADNc (que incorpora nucleótidos marcados con Cy3) de acuerdo con los métodos descritos en el Ejemplo 2. (A) muestra una imagen después de la síntesis de ADNc (C) y (D) o la síntesis y liberación combinada de ADNc de las sondas de captura (E). (B) muestra una imagen después de la liberación de las sondas de captura extendida (C) y (D) o después de un período de incubación adicional (E).

La Figura 3 muestra gráficos de barras que muestran: (A) el número total de transcritos únicos identificados después de las etapas de extensión y liberación separadas (ST1.0) y después de las etapas de extensión y liberación combinadas (ST2.0); y (B) el número total de genes únicos identificados después de las etapas de extensión y liberación separadas (ST1.0) y después de las etapas de extensión y liberación combinadas (ST2.0).

La Figura 4 muestra mapas térmicos que muestran el número de transcritos capturados en cada elemento de secciones de tejido cerebral de ratón mediante el uso de: (A) etapas de extensión y liberación separadas (ST1.0); y (B) etapas de extensión y liberación combinadas (ST2.0).

La Figura 5 muestra una superposición de una imagen de tejido cerebral de ratón y transcritos de Penk capturados en elementos de la matriz en contacto con dicho tejido mediante el uso de: (A) etapas de extensión y liberación separadas (ST1.0); y (B) etapas de extensión y liberación combinadas (ST2.0).

Ejemplo 1

El protocolo descrito más abajo es representativo de los métodos transcripcionales espaciales "estándar" en los que las etapas de síntesis de ADNc y liberación de las sondas extendidas se realizan por separado, lo que se conoce como "ST1.0".

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

El oligonucleótido de captura de ARN (Tabla 1) se imprimió en portaobjetos de vidrio para funcionar como la sonda de captura. La sonda se sintetizó con un enlazador amino de extremo 5' terminal con un espaciador C6. 1007 oligonucleótidos de captura con códigos de barras únicos de 18 mer, un UMI semialeatorizado de 9 mer y una región de captura de poli-20TVN se dispusieron sobre la superficie de los portaobjetos activados Codelink de acuerdo con las instrucciones del fabricante (Surmodics). Los oligonucleótidos se inmovilizaron en elementos de 100 µm con una distancia de centro a centro (paso) de 200 µm entre los puntos. De acuerdo con el fabricante del portaobjetos, se inmovilizaron aproximadamente 200 millones de oligonucleótidos en cada punto. Un área en la esquina superior izquierda de la matriz se dejó en blanco con el propósito de orientación. Se prepararon seis matrices de 6200 x 6600 µm por portaobjeto de vidrio. Después de la impresión, se realizó el bloqueo de la superficie de acuerdo con las instrucciones del fabricante.

Tabla 1

Sonda 1

Amino-C6- GTCCGATATGATTGCCGWSNNWSNNVTTTTTTTTTTTTTTTTTTVN (SEQ ID NO: 3)	UUUUUGACTCGTAATACGACTCACTATAGGGACACGACGCTCTTCCGATCT
--	---

Recolección y preparación de bulbos olfativos

Se sacrificaron ratones adultos C57BL /6 (>2 meses de edad) y se aislaron inmediatamente los bulbos olfativos y se congelaron al instante en isopentano. El tejido se incrustó en OCT frío antes del corte en secciones. El bulbo olfativo se cortó en secciones en un criostato con un grosor de 10 µm. Las secciones se montaron en matrices con códigos de barras espaciales.

Permeabilización y transcripción inversa

Para cada pocillo, correspondiente a cada submatriz con una sección, se añadieron 70 µl de tampón de reacción de exonucleasa I 1x (#B0293S, NEB) con BSA 0,19 µg/µl (#B9000S, NEB) y se incubó a 37 °C durante 30 minutos. Cada pocillo se lavó con 100 µl de SSC 0,1x, se diluyó en agua desionizada a partir de solución madre (#S6639, Sigma-Aldrich). A continuación, se añadieron 70 µl de pepsina al 0,1 % (#P7000-25G, Sigma-Aldrich) disuelta en HCl 0,1 M (#318965-1000ML, Sigma-Aldrich) a cada pocillo y se incubó a 37 °C durante 10 minutos. Cada pocillo se lavó como se describió anteriormente y se añadieron 70 µl de mezcla de transcripción inversa a cada pocillo y se incubó durante toda la noche (ON).

La mezcla de transcripción inversa contenía tampón de primera hebra 1x (#18080-044, Invitrogen), DTT 5 mM (#18080-044, Invitrogen), 500 µM de dATP/dGTP/dTTP, 12,5 µM de dCTP, 25 µM de cianina 3-dCTP (#NEL576001EA, PerkinElmer), BSA 0,19 µg/µl, Actinomicina D 50 ng/µl (#A1410-2MG, Sigma-Aldrich), DMSO 1 % (#472301-500 ML, Sigma-Aldrich), Superscript III 20 U/µl (#18080-04, Invitrogen) y RNaseOUT 2U /µl (#10777-019). La permeabilización se llevó a cabo durante 2, 10 o 30 minutos. La síntesis de ADNc se llevó a cabo durante 1 hora, 3 horas, 6 horas o durante toda la noche.

Para degradar y eliminar tejido de ratón, se añadieron 70 µl de proteinasa K (#19131, Qiagen) y tampón PKD (pH 7,5, de acuerdo con el fabricante, #1034963, Qiagen), a una relación de 1:7, a cada pocillo y se incubó a 56 °C durante 1 hora. Después de la incubación, los portaobjetos se lavaron con SSC 2x con SDS al 0,1 % a 50 °C (#71736-100ML, Sigma-Aldrich) seguido de SSC 0,2x y SSC 0,1x a temperatura ambiente (RT) y por último se secaron por centrifugación.

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.

Liberación de ADNc de la matriz

La matriz se unió a un soporte de portaobjetos Arraylt y una máscara de 16 pocillos (Arraylt Corporation). Se añadió a cada pocillo una mezcla de escisión (50 µl) que contenía tampón de Exo I 1x (New England Biolabs, Ipswich, MA, USA), BSA 1x, agua libre de RNasa/DNasa, 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.

Ejemplo 2

Se diseñó un experimento para determinar los efectos de realizar las etapas de extensión y liberación simultáneamente. En la Figura 1 se muestra un esquema del diseño experimental, que representa seis tratamientos diferentes: C1, C2, D1, D2, E1 y E2.

C1 y C2 se refieren al protocolo "estándar" descrito en el Ejemplo 1 anterior, mediante el uso de la mezcla de síntesis de ADNc expuesta más abajo:

	a) Tampón de la primera hebra	64 µl
	b) DTT	16 µl
5	c) Agua libre de RNasa y DNasa	133 µl
	d) Actinomicina D	32 µl
	e) Mezcla de dNTP	16 µl
	f) Cy3 dCTP	8 µl
	g) BSA	3 µl
10	h) RNasa OUT	16 µl
	i) SuperScript® III	32 µl

La etapa de permeabilización tisular incluyó tratamiento con un tampón de exonucleasa (C1) o tampón de exonucleasa que comprende exonucleasa I.

15 D1 y D2 difieren de C1 y C2 solo con respecto a la etapa de escisión, en donde la mezcla de escisión comprendía la enzima USER en la mezcla de tampón de ADNc. La mezcla de escisión se expone más abajo:

	a) Tampón de la primera hebra	32 µl
	b) DTT	8 µl
	c) agua libre de RNasa y DNasa	70,5 µl
	d) Actinomicina D	16 µl
20	e) dNTP	8 µl
	f) BSA	1,5 µl
	g) Enzima USER	16 µl

25 E1 y E2 difieren de C1 y C2 en la medida en que la mezcla de síntesis de ADNc incluía la enzima de escisión, es decir, la enzima USER. La mezcla de reacción para la síntesis de ADNc y liberación de la sonda combinadas, es decir simultáneas, se expone más abajo:

	a) Tampón de la primera hebra	32 µl
	b) DTT	8 µl
	c) Agua libre de RNasa y DNasa	50,5 µl
	d) Actinomicina D	16 µl
30	e) Mezcla de dNTP	8 µl
	f) Cy3 dCTP	4 µl
	g) BSA	1,5 µl
	h) RNasa OUT	8 µl
	i) SuperScript® III	16 µl
35	j) Enzima USER	16 µl

Los protocolos E1 y E2 no incluían la etapa de eliminación de tejido. El protocolo E también se denomina en la presente descripción ST2.0.

Para cada pocillo, la mezcla de liberación se añadió e incubó a 37 °C durante 1-2 horas.

40 La Figura 2 muestra la señal de Cy3 de la superficie de las matrices, es decir, el marcador fluorescente incorporado en las sondas extendidas, después de la síntesis de ADNc (A) y aproximadamente 3 horas después de la reacción de

escisión. La Figura 2 demuestra que el protocolo C da como resultado la síntesis de ADNc en la superficie de la matriz (A) y la liberación subsecuente de las sondas extendidas de la superficie de la matriz (B), como se esperaba.

El protocolo D muestra que la enzima de escisión (USER) es capaz de funcionar eficientemente en el tampón de transcriptasa inversa, ya que las sondas extendidas se liberan de la superficie de la matriz con eficacia similar a la reacción de escisión en el protocolo C.

El protocolo E muestra que las sondas de captura se escindieron con éxito simultáneamente con la síntesis de ADNc. En este sentido, las sondas extendidas de los protocolos C y E se combinaron y se sometieron a análisis de secuencias para confirmar que el ADNc se sintetizó exitosamente en el protocolo E.

Recolección de la mezcla de sondas extendidas

Se recogieron 65 µl de la mezcla de sondas extendidas de cada uno de los pocillos usados y se colocaron en tubos de PCR de 0,2 ml.

Síntesis de la segunda hebra de ADNc para la preparación de una biblioteca de ADNc para la secuenciación

Se añadieron 5 µl de la mezcla de segunda hebra que contiene tampón de primera hebra 2,7x, 3,7 U/µl de ADN polimerasa I (#18010-017, Invitrogen) y 0,18 U/µl de RNasaH (#18021-014, Invitrogen) y las muestras se incubaron a 16 °C durante 2 horas, se añadieron 5 µl de ADN polimerasa de T4 (#M0203S, NEB) y las muestras se incubaron a 16 °C durante otros 20 minutos. Se añadieron 25 µl de agua desionizada o EDTA 80 mM (#15575-038, Invitrogen) y se purificó mediante el uso de perlas Agencourt RNAClean XP (#A63987, Beckman Coulter) de acuerdo con el protocolo del fabricante y se eluyó en agua desionizada. Se mezclaron 5,6 µl de las muestras con 10,4 µl de mezcla de transcripción in vitro con un contenido final de tampón de reacción de T7 1x (#AM1333, Ambion), 7,5 mM de cada NTP (#AM1333, Ambion), mezcla de enzima T7 1x (#AM1333, Ambion) y SUPERaseIN 1 U/µl (#AM2694, Ambion). Las muestras se incubaron a 37 °C durante 14 horas.

Las muestras se purificaron mediante el uso de perlas Agencourt RNAClean XP de acuerdo con el protocolo del fabricante y se eluyeron en 10 µl de agua desionizada. La cantidad y longitud promedio de fragmento de ARN amplificado (ARNA) se determinó mediante el uso del kit RNA 6000 Pico (#5067-1513, Agilent) con un bioanalizador 2100 (Agilent) de acuerdo con el protocolo del fabricante. La muestra restante y 2,5 µl de adaptador de ligazón se añadieron a una concentración final de 0,71 µM. Las muestras se calentaron a 70 °C durante 2 minutos y después se colocaron en hielo antes de añadir 4,5 µl de la mezcla de ligazón a un contenido final de tampón de reacción de ligasa de ARN de T4 1x (#B0216L, NEB), 20 U/µl de ligasa de ARN de T4 2, truncada (#M0242L, NEB), inhibidor de RNasa 4 U/µl, murino (#M0314L, NEB) y adaptador de ligazón 0,5 µM. Las muestras se incubaron a 25 °C durante 1 hora. Las muestras se purificaron mediante el uso de perlas Agencourt RNAClean XP de acuerdo con el protocolo descrito anteriormente y se mezclaron con 1 µl de cebador RT (IDT) hasta una concentración final de 1,7 µM y 1 µl de dNTP y una concentración final de 0,83 mM de cada dNTP. Las muestras se calentaron a 65 °C durante 5 min y después se colocaron en hielo, se añadieron 8 µl de una mezcla de transcripción inversa a un contenido final de tampón de primera hebra 1x, DTT 0,05 M, 500 µM de cada dNTP, cebador de RT 1 mM, Superscript III 10 U/µl y RNaseOUT 2U/µl. Las muestras se incubaron a 50 °C durante 1 hora antes de colocarlas en hielo y las muestras se purificaron subsecuentemente mediante el uso de perlas Agencourt RNAClean XP de acuerdo con el protocolo descrito anteriormente. Un volumen de reacción total de 10 µl que contenía la mezcla 1xKAPA HiFi HotStart ReadyMix (#KK2601, KAPA Biosystems), EVA green 1x (#31000, Biotium), PCR InPE1.0 0,5 µM (Eurofins), PCR InPE2.0 0,01 µM (Eurofins), PCR Index 0,5 µM (Eurofins) y 2 µl de ADNc purificado se amplificaron mediante qPCR con el siguiente protocolo: 98 °C durante 3 minutos, seguido de ciclo a 98 °C durante 20 segundos, 60 °C durante 30 segundos y 72 °C durante 30 segundos. Después de determinar la cantidad de ciclos necesarios en base a la qPCR, las muestras se amplificaron en un volumen de reacción total de 25 µl. Las bibliotecas se purificaron y las muestras se eluyeron en 20 µl de tampón de elución (#19086, Qiagen) y la longitud promedio de fragmento de las bibliotecas terminadas se determinó mediante el uso del kit de ADN 1000 (#5067-1504, Agilent) con un bioanalizador 2100 de acuerdo con el protocolo del fabricante. La concentración de las bibliotecas terminadas se determinó con el kit de ensayo Qubit dsDNA HS (#Q32854, Life Technologies) de acuerdo con el protocolo del fabricante. Las bibliotecas terminadas se diluyeron a 4 nM y se secuenciaron en la plataforma Illumina NextSeq mediante el uso de secuenciación de extremos pareados de acuerdo con el protocolo del fabricante. Típicamente, 31 o 51 bases se secuenciaron en la lectura uno, y 121 o 101 bases se secuenciaron en la lectura dos. Los oligonucleótidos usados durante la preparación de la biblioteca fueron:

Adaptador de ligazón:

rApp]AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC[ddC] (SEQ ID NO: 4)

Segundo cebador de transcripción inversa:

TGACTGGAGTTCAGACGTGTGCTCTTCCGA (SEQ ID NO: 5)

Cebador de PCR InPE1.0:

ATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT (SEQ ID NO: 6)

Cebador de PCR InPE2.0:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT (SEQ ID NO: 7)

Cebador de PCR Index:

AAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTC (SEQ ID NO: 8)

- 5 Los resultados de la Figura 3 demuestran sorprendentemente que la escisión y extensión simultáneas de las sondas de captura (protocolo E, es decir ST2.0) dio como resultado un aumento en el número total de transcritos (es decir, sondas extendidas) producidos (Figura 3A) y también un aumento en el número de genes únicos detectados, es decir, la escisión y extensión simultáneas de las sondas de captura mejoraron la diversidad de los ácidos nucleicos capturados de la muestra de tejido.
- 10 Ejemplo 3
- El protocolo estándar ST1.0 (es decir, reacciones de extensión y escisión separadas) y el protocolo ST2.0 (es decir, la escisión y extensión simultáneas de las sondas de captura) se realizaron en tejido cerebral de ratón mediante el uso de una matriz con sondas de captura que contenían dominios posicionales. La Figura 4 muestra mapas térmicos de los datos de secuenciación obtenidos para cada elemento en la matriz mediante el uso de ST1.0 (A) y ST2.0 (B).
- 15 La Figura 5 muestra una superposición de los datos de expresión para el gen de la proencefalina (Penk) en una imagen del tejido cerebral de ratón obtenida mediante el uso de ST1.0 (A) y ST2.0 (B).
- Es evidente a partir de los datos de la Figura 4 que la mayoría de las sondas extendidas se correlacionan directamente con la parte de la matriz en contacto con la muestra de tejido. Por lo tanto, los resultados en la Figura 4 indican que los cebadores liberados en el protocolo ST2.0 son capaces de reunirse con, e hibridarse con, transcritos en el tejido.
- 20 Además, los datos en la Figura 4 sugieren que un mayor número de sondas se extienden mediante el uso del protocolo ST2.0 con relación al protocolo ST1.0.
- La Figura 5 demuestra que pueden detectarse transcritos únicos del gen de Penk espacialmente mediante el uso de los protocolos ST1.0 y ST 2.0. Ambos protocolos muestran que los transcritos de Penk se ubican en la estructura central en forma de pera del tejido, lo que demuestra que ambos protocolos facilitan el etiquetado espacial de moléculas de ácidos nucleicos de un espécimen biológico.
- 25

REIVINDICACIONES

1. Un método para etiquetar espacialmente ácidos nucleicos de un espécimen biológico que comprende:
 - (a) proporcionar un sustrato sólido en el que se inmovilizan múltiples especies de sondas de captura de manera que cada especie ocupe una posición distinta sobre el sustrato sólido, en donde cada sonda de captura de las múltiples especies es un cebador para una reacción de extensión y en donde la sonda de captura comprende una molécula de ácido nucleico que comprende:
 - (i) un dominio de escisión para liberar la sonda de captura del sustrato sólido;
 - (ii) un dominio posicional que corresponde a una posición de la sonda de captura en el sustrato sólido;
 - (iii) un dominio de captura
 - (b) poner en contacto el sustrato sólido con el espécimen biológico
 - (c) liberar simultáneamente las múltiples especies de sondas de captura del sustrato sólido en condiciones que permiten que un ácido nucleico del espécimen biológico se hibride con el dominio de captura de la sonda de captura como una plantilla de extensión para producir una sonda de captura extendida, en donde la etapa (c) comprende poner en contacto el sustrato sólido con una mezcla de reacción acuosa que comprende:
 - (i) una transcriptasa inversa capaz de extender la sonda de captura mediante el uso del ácido nucleico hibridado a la sonda de captura como una plantilla de extensión para producir la sonda de captura extendida, en donde la sonda de captura extendida comprende además un saliente de ADN de hebra única;
 - (ii) un oligonucleótido de cambio de plantilla que comprende una secuencia capaz de hibridarse con el saliente de ADN de hebra única y un dominio de amplificación, en donde la transcriptasa inversa es capaz de extender aún más la sonda de captura extendida mediante el uso del oligonucleótido de cambio de plantilla como una plantilla de extensión;
 - (iii) una polimerasa capaz de extender un cebador hibridado al dominio de amplificación incorporado en la sonda de captura extendida como una plantilla de extensión para producir una sonda de captura extendida de doble hebra; y
 - (iv) un medio para liberar las múltiples especies de sondas de captura del sustrato sólido.
2. El método de la reivindicación 1, en donde el medio para liberar las múltiples especies de sondas de captura comprende: una enzima de escisión capaz de escindir específicamente las múltiples especies de sondas de captura dentro del dominio de escisión, opcionalmente, en donde el dominio de escisión comprende una secuencia de poli-U y la enzima de escisión comprende una mezcla de uracil ADN glucosilasa (UDG) y una enzima endonucleasa capaz de reconocer sitios apurínicos/apirimidínicos (AP) de ADNdh, preferentemente una enzima ADN glucosilasa-liasa tal como la endonucleasa VIII, opcionalmente, donde el dominio de escisión comprende una secuencia de reconocimiento de endonucleasa de restricción y la enzima de escisión es una endonucleasa de restricción o en donde el dominio de escisión es un reticulador químico escindible.
3. El método de cualquier reivindicación 1 o 2, en donde la transcriptasa inversa se selecciona de la lista que consiste en transcriptasas inversas de M-MLV, MuLV, AMV y HIV y un derivado o un mutante de las mismas, preferentemente, en donde el derivado es un derivado de secuencia modificada, opcionalmente en donde la transcriptasa inversa comprende una secuencia polipeptídica como se expone en la SEQ ID NO: 2 o una secuencia polipeptídica que tiene al menos 80 % de identidad de secuencia con la misma, opcionalmente, en donde la transcriptasa inversa está codificada por una secuencia de nucleótidos que comprende una secuencia como se expone en la SEQ ID NO: 1 o una secuencia de nucleótidos que tiene al menos 80 % de identidad de secuencia con la misma.
4. El método de cualquiera de las reivindicaciones 1 a 3, en donde el método comprende además una etapa de amplificación de la sonda de captura extendida y/o la sonda de captura extendida de doble hebra.
5. El método de cualquiera de las reivindicaciones 1 a 4, en donde el método comprende además una etapa de análisis de la sonda de captura extendida y/o la sonda de captura extendida de doble hebra y/o un amplicón de la misma, en donde la etapa de análisis de la sonda de captura extendida y/o la sonda de captura extendida de doble hebra y/o el amplicón de la misma comprende analizar las secuencias de la sonda de captura extendida y/o la sonda de captura extendida de doble hebra y/o el amplicón de la misma.
6. El método de la reivindicación 5, en donde el método comprende además una etapa de correlación de la información de análisis, por ejemplo, la información del análisis de secuencias, con una imagen del espécimen biológico.

7. El método de cualquiera de las reivindicaciones 1 a 6, en donde el método comprende además una etapa de marcaje de la sonda de captura extendida generada en la etapa (c) y/o una hebra complementaria de la sonda de captura extendida y/o el amplicón de la misma, en donde la etapa de marcaje puede realizarse simultáneamente con, o subsecuente a, la etapa de extensión.
- 5 8. El método de cualquiera de las reivindicaciones 1 a 7, en donde las múltiples especies de sondas de captura comprenden además un dominio de amplificación.
9. El método de cualquiera de las reivindicaciones 1 a 8, en donde el dominio posicional de las múltiples especies de sondas de captura comprende una secuencia de código de barras.
- 10 10. El método de cualquiera de las reivindicaciones 1 a 9, en donde el dominio de captura comprende un oligonucleótido de ADN poli-T y/o una secuencia oligonucleotídica aleatoria o redundante, opcionalmente en donde el dominio de captura comprende una secuencia específica para un gen o grupo de genes diana particular.
11. El método de cualquiera de las reivindicaciones 1 a 10, en donde el sustrato sólido comprende una matriz, opcionalmente, en donde la matriz es una matriz de perlas.
- 15 12. El método de cualquiera de las reivindicaciones 1 a 11, en donde las múltiples especies de sondas de captura se inmovilizan en el sustrato sólido en una disposición aleatoria, opcionalmente en donde el método comprende además una etapa de realizar una reacción de detección de ácido nucleico en el sustrato sólido para determinar las secuencias de dominio posicional de las múltiples especies de sondas de captura en el sustrato sólido.
- 20 13. El método de cualquiera de las reivindicaciones 1 a 12, en donde el espécimen biológico es una sección de tejido, opcionalmente en donde la sección de tejido se prepara mediante el uso de un tejido fijado, por ejemplo, un tejido fijado con formalina e incrustado en parafina (FFPE), o tejido congelado.
14. El método de cualquiera de las reivindicaciones 1 a 13, que comprende además una etapa de permeabilizar y/o fijar el espécimen biológico después de poner en contacto el espécimen biológico con el sustrato sólido.
- 25 15. El método de cualquiera de las reivindicaciones 1-14, en donde la polimerasa es una ADN polimerasa que desplaza la hebra, opcionalmente, en donde la ADN polimerasa que desplaza la hebra es ADN polimerasa phi29, ADN polimerasa BST o ADN polimerasa Klenow.

Figura 1

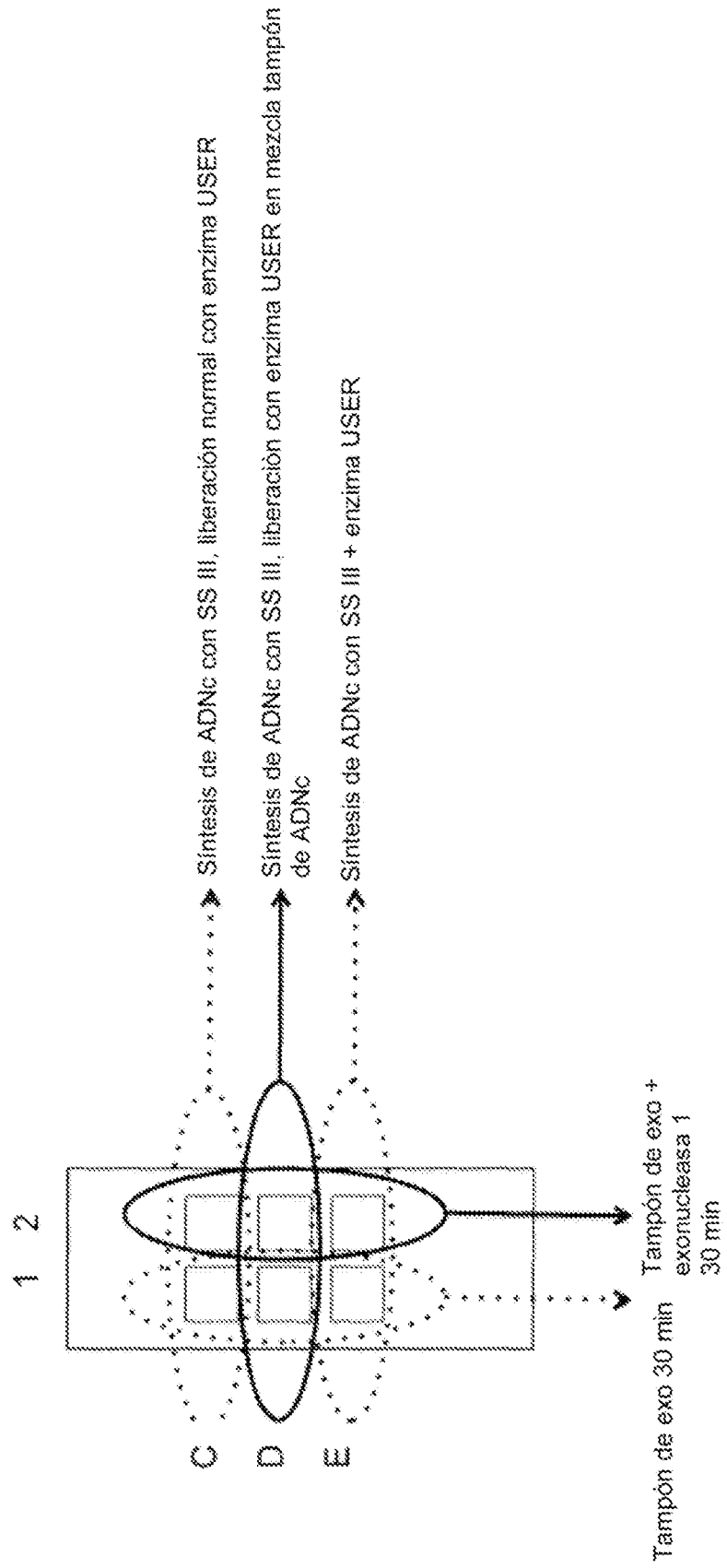


Figura 2

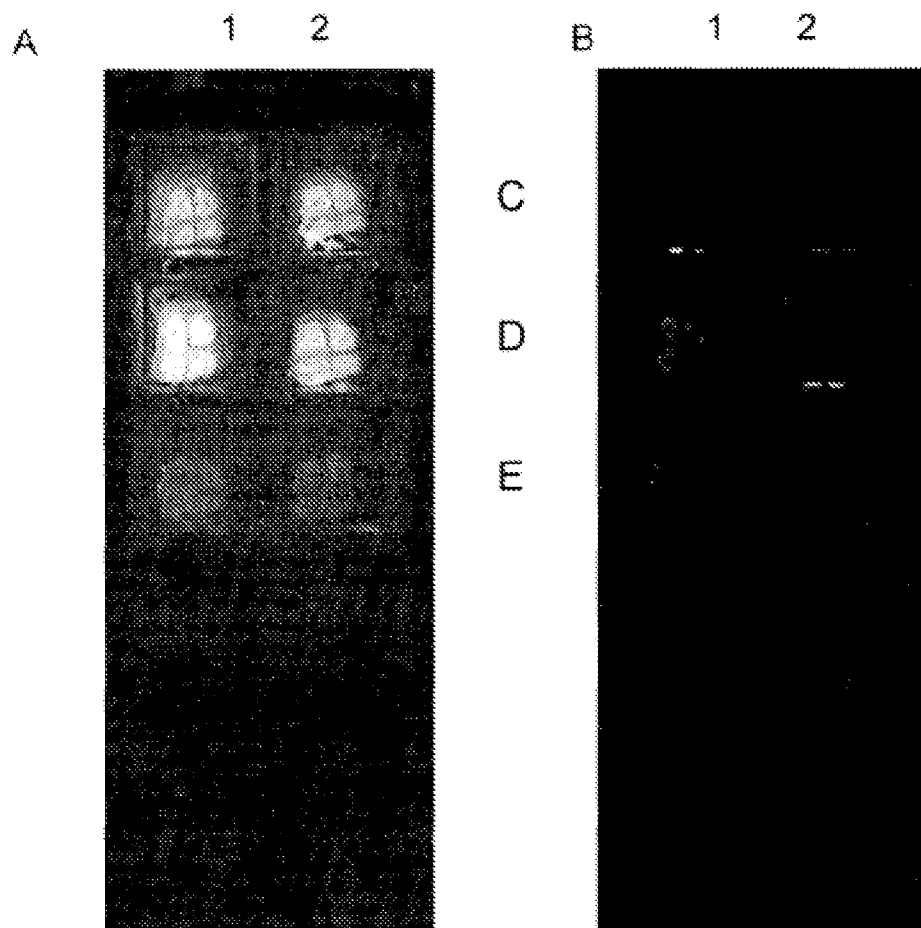


Figura 3

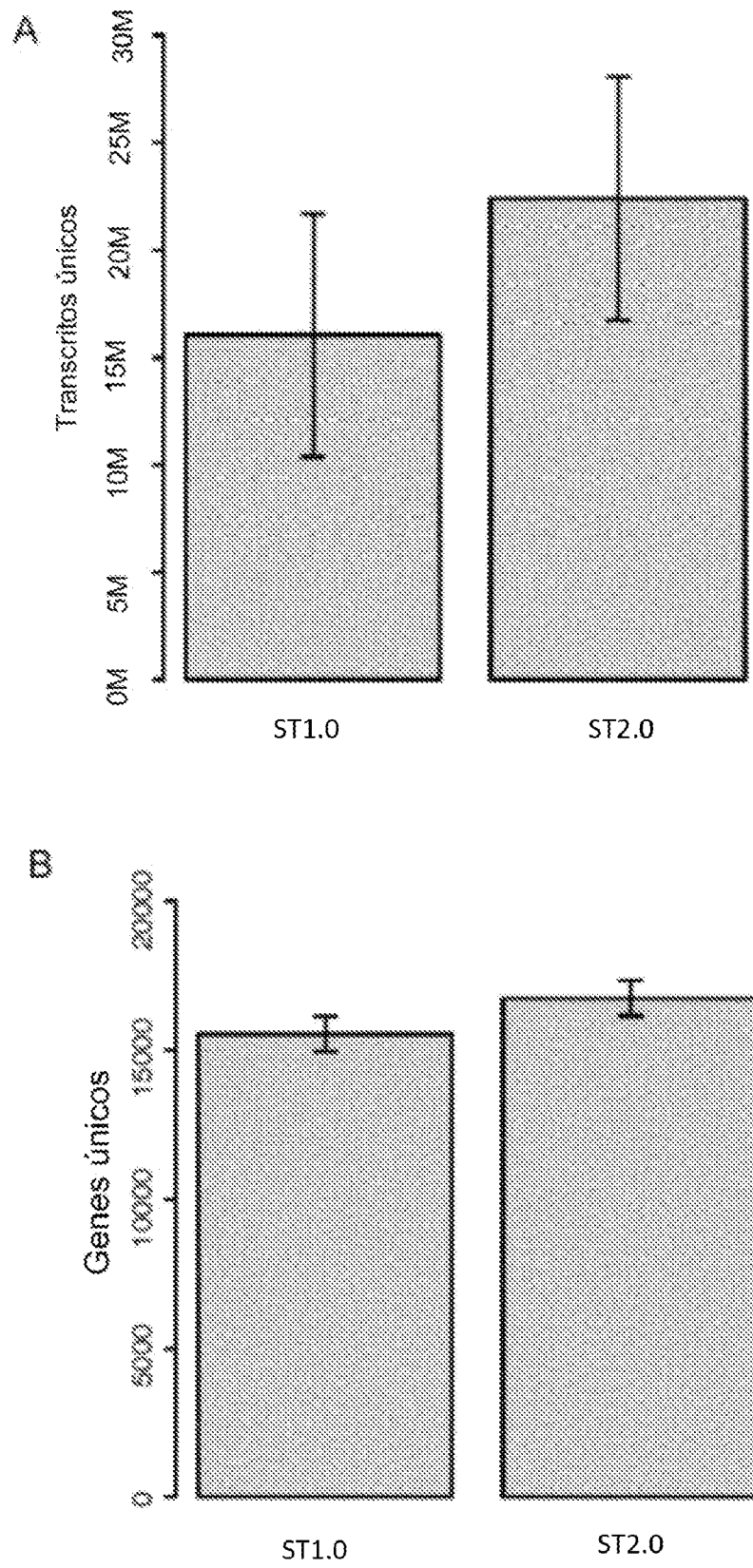


Figura 4

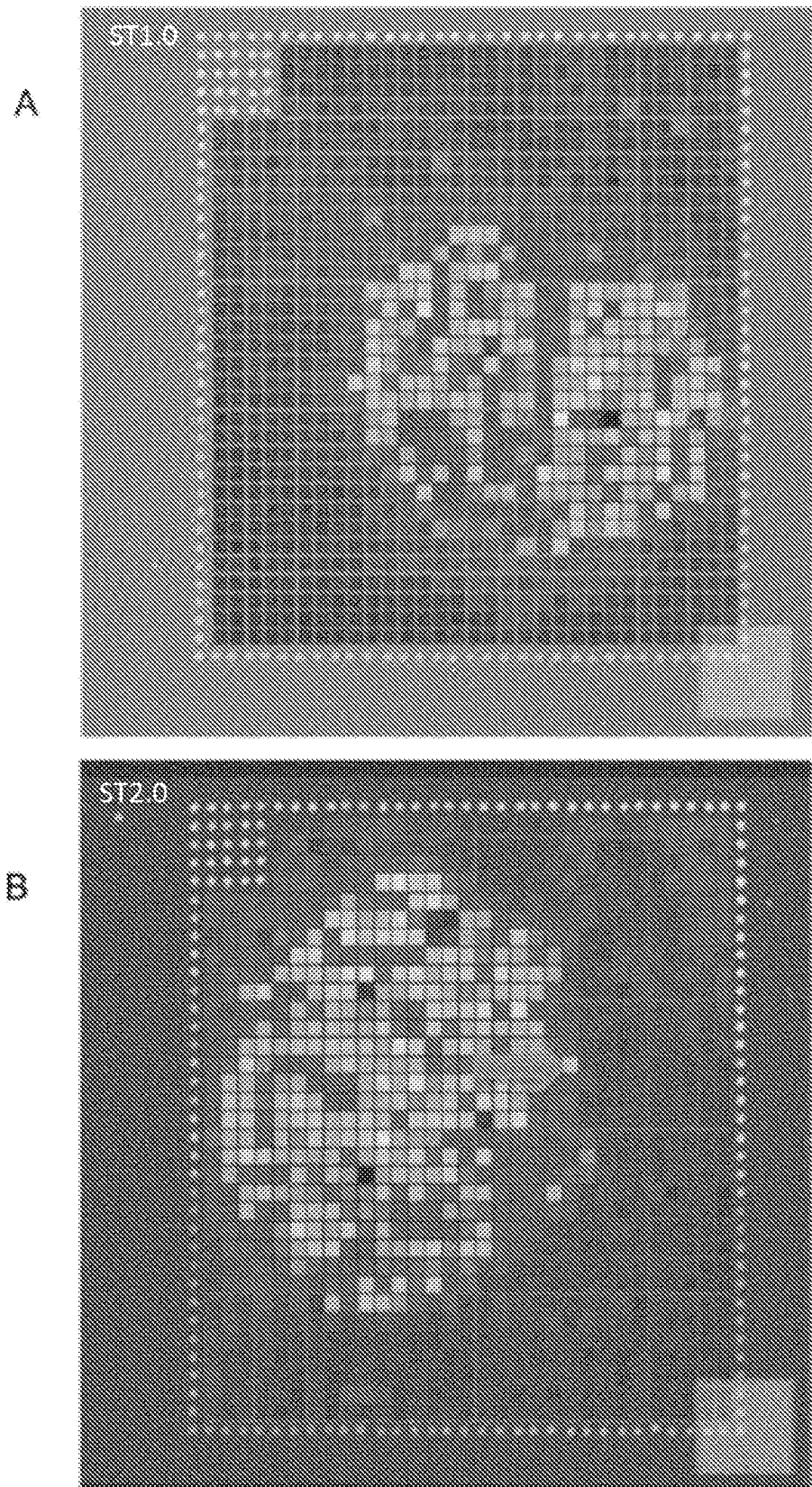
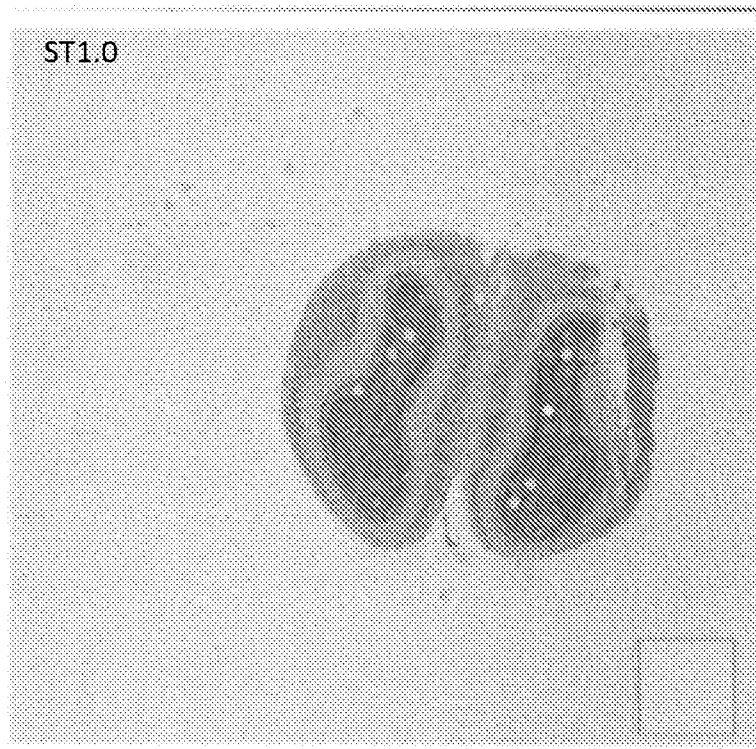


Figura 5

A



B

