

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 008 686**

51 Int. Cl.:

C12Q 1/6841 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.03.2022** **PCT/US2022/020985**

87 Fecha y número de publicación internacional: **22.09.2022** **WO22198068**

96 Fecha de presentación y número de la solicitud europea: **18.03.2022** **E 22715466 (3)**

97 Fecha y número de publicación de la concesión europea: **15.01.2025** **EP 4301870**

54 Título: **Captura en formato múltiple de la expresión génica y proteica de una muestra biológica**

30 Prioridad:

18.03.2021 US 202163162870 P
23.06.2021 US 202163214058 P
17.09.2021 US 202163245697 P
05.10.2021 US 202163252335 P
21.10.2021 US 202163270230 P
18.02.2022 US 202263311703 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
24.03.2025

73 Titular/es:

10X GENOMICS, INC. (100.00%)
6230 Stoneridge Mall Road
Pleasanton, CA 94588-3260, US

72 Inventor/es:

CHEW, JENNIFER;
SHUGA, JOSEPH FRANCIS;
ROELLI, PATRICK;
CHEUNG, DENISE;
HERMES, ALEXANDER;
YIN, YIFENG;
ANAPARTHY, NAISHITHA;
HATORI, RYO;
STOECKIUS, MARLON;
SPALINSKAS, RAPOLAS;
KATSORI, ANNA-MARIA;
UYTINGCO, CEDRIC;
TURKEKUL, MESRUH;
SUKOVICH, DAVID;
GALONSKA, CHRISTINA;
JUREK, ALEKSANDRA y
BLOJU, OCTAVIAN MARIAN

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 3 008 686 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

Captura en formato múltiple de la expresión génica y proteica de una muestra biológica

5 Campo de la invención

En la presente descripción se proporcionan métodos, composiciones y kits para preparar muestras biológicas para el análisis proteómico y de expresión génica espacial en formato múltiple, tal como determinar una localización de un analito de ácido nucleico y un analito de proteína en una muestra biológica.

10

Antecedentes

Las células dentro de un tejido de un sujeto tienen diferencias en la morfología y/o función celular debido a niveles variados de analito (por ejemplo, expresión génica y/o de proteínas) dentro de las diferentes células. La posición específica de una célula dentro de un tejido (por ejemplo, la posición de la célula con relación a las células vecinas o la posición de la célula con relación al microentorno del tejido) puede afectar, por ejemplo, la morfología, la diferenciación, el destino, la viabilidad, la proliferación, el comportamiento, y la señalización y el intercambio de la célula con otras células en el tejido.

15

20

La heterogeneidad espacial se ha estudiado previamente mediante el uso de técnicas que solo proporcionan datos para un pequeño puñado de analitos en el contexto de un tejido intacto o una porción de un tejido, o proporcionan datos sustanciales del analito para el tejido disociado (es decir, células individuales), pero no logran proporcionar información sobre la posición de la célula individual en una muestra biológica madre (por ejemplo, muestra de tejido).

25

Además, la detección en formato múltiple de diferentes analitos (por ejemplo, ácido nucleico, proteína, etc.) en la misma muestra biológica, mientras se conserva la información espacial sigue siendo un desafío en el campo. Los métodos actuales incluyen la detección espacial amplia del transcriptoma o varios métodos de detección de proteínas, sin embargo, aún se necesitan métodos que logren ambos dentro del contexto espacial de una muestra biológica.

30

Resumen

La invención se expone en el conjunto de reivindicaciones adjuntas.

35

La detección en formato múltiple de diferentes analitos (por ejemplo, ácido nucleico, proteína, etc.) en la misma muestra biológica, mientras se conserva la información espacial, sigue siendo un desafío en el campo. Como se describió anteriormente, los métodos actuales incluyen la detección espacial amplia del transcriptoma o varios métodos de detección de proteínas, incluida la inmunofluorescencia, sin embargo, aún se necesitan métodos que detecten tanto la expresión génica espacial como la expresión proteica espaciales dentro del contexto espacial de una muestra biológica simultáneamente.

40

En la presente descripción se proporcionan métodos para determinar la localización espacial de un ácido nucleico y una proteína a partir de una muestra biológica que incluye: a) proporcionar una matriz espacial que incluye una primera y segunda pluralidad de sondas de captura donde cada pluralidad incluye un código de barras espacial y un dominio de captura, b) poner en contacto la matriz espacial con una muestra biológica, c) poner en contacto la muestra biológica con (i) una pluralidad de agentes de captura de analito, donde un agente de captura de analito incluye un resto de unión al analito y un oligonucleótido que incluye un código de barras de un resto de unión al analito y una secuencia de captura de analito, donde la secuencia de captura de analito incluye una secuencia complementaria a una segunda pluralidad de dominios de captura, y (ii) una pluralidad de sondas de ligazón con plantilla, donde una de las sondas de ligazón con plantilla incluye una secuencia complementaria a una primera pluralidad de dominios de captura, d) unir el resto de unión al analito del agente de captura de analito a una proteína objetivo, e) hibridar las sondas de ligazón con plantilla a un ácido nucleico objetivo y ligar las sondas para producir productos de ligazón, f) hibridar los productos de ligazón a la primera pluralidad de dominios de captura y las secuencias de captura de analito de los agentes de captura de analito unidos a la segunda pluralidad de dominios de captura en la matriz espacial, y g) determinar la secuencia o una porción de esta de un producto de ligazón capturado, o un complemento de esta, y la secuencia del código de barras espacial de la sonda de captura, o un complemento de esta, que se asocia con el producto de ligazón, y la secuencia del resto de unión al analito de la sonda de captura unida, o un complemento de esta, del agente de captura de analito unido, para determinar de esta manera la localización espacial de un ácido nucleico y la proteína de la muestra biológica.

55

60

En algunas modalidades, los dominios de captura de la primera pluralidad de sondas de captura se definen como secuencias de captura no homopoliméricas o secuencias homopoliméricas. En algunas modalidades, los dominios de captura de la segunda pluralidad de sondas de captura se definen como secuencias no homopoliméricas o secuencias de captura homopoliméricas. En algunas modalidades, los dominios de captura de la primera pluralidad de sondas de captura son diferentes de los dominios de captura de la segunda pluralidad de sondas de captura. En algunas modalidades, la secuencia homopolimérica incluye una secuencia de poliT y la secuencia no homopolimérica incluye

65

una secuencia fija o una secuencia redundante. En algunas modalidades, la secuencia fija incluye al menos una secuencia seleccionada de SEQ ID NO: 1 a SEQ ID NO: 11.

En algunas modalidades, el ácido nucleico es un ARN o ADN. En algunas modalidades, el ARN es un ARNm.

En algunas modalidades, la muestra biológica es una muestra de tejido. En algunas modalidades, la muestra de tejido es una muestra de tejido recién congelada o una muestra de tejido fijada, donde la muestra de tejido fija es una muestra de tejido fijada con formaldehído, una muestra de tejido fijada con acetona, una muestra de tejido con paraformaldehído o una muestra de tejido fijada con metanol. En algunas modalidades, la muestra biológica es una sección de tejido, donde la sección de tejido es una sección de tejido recién congelada o una sección fijada, y opcionalmente, donde la sección de tejido fijada es una sección de tejido embebida en parafina fijada con formalina, una sección de tejido fijada con acetona, una sección de tejido con paraformaldehído o una sección de tejido fijada con metanol. En algunas modalidades, la muestra de tejido se deriva de una muestra de biopsia o un embrión de roedor completo.

En algunas modalidades, antes de la etapa (b) la muestra biológica se desparafina y desreticula. En algunas modalidades, la desreticulación incluye el uso de un tampón. En algunas modalidades, el tampón incluye tampón Tris-EDTA a un pH de aproximadamente 8 a aproximadamente 10 y una temperatura de aproximadamente 60 °C a aproximadamente 80 °C. En algunas modalidades, el tampón incluye tampón citrato a un pH de aproximadamente 5 a aproximadamente 7 y una temperatura de aproximadamente 70 °C a aproximadamente 100 °C.

En algunas modalidades, la hibridación de los productos de ligazón con plantilla y las secuencias de captura de analito de los agentes de captura de analito unidos incluye la permeabilización de la muestra biológica.

En algunas modalidades, la secuencia de captura de analito del oligonucleótido se bloquea antes de unirse a la proteína objetivo. En algunas modalidades, el oligonucleótido del agente de captura de analito se bloquea mediante una sonda de bloqueo. En algunas modalidades, la sonda de bloqueo se elimina antes de hibridar la secuencia de captura de analito del oligonucleótido de la secuencia de captura de analito con el dominio de captura de la sonda de captura.

En algunas modalidades, la determinación en la etapa (g) incluye: a) extender los productos de ligazón capturados y los oligonucleótidos capturados de los agentes de captura de analito, donde los productos de extensión incluyen el código de barras espacial o un complemento de este, b) liberar los productos de extensión, o los complementos de estos, de la matriz espacial, c) producir una biblioteca a partir de los productos de extensión liberados o los complementos de estos, y d) secuenciar la biblioteca.

En algunas modalidades, antes de la etapa (c) el método incluye preamplificar los productos de extensión, o complementos de estos.

En algunas modalidades, el complemento del oligonucleótido del agente de captura de analito incluye un código de barras del resto de unión al analito específico para la porción de unión al analito del agente de captura de analito.

En algunas modalidades, la primera y segunda pluralidad de sondas de captura incluyen un dominio de escisión, uno o más dominios funcionales, un identificador molecular único y sus combinaciones.

En algunas modalidades, el método incluye la obtención de imágenes de la muestra biológica. En algunas modalidades, la obtención de imágenes incluye una o más de microscopía de expansión, microscopía de campo claro, microscopía de campo oscuro, microscopía de contraste de fases, microscopía electrónica, microscopía de fluorescencia, microscopía de reflexión, microscopía de interferencia y microscopía confocal.

En algunas modalidades, el método incluye tefir la muestra biológica. En algunas modalidades, la tinción incluye hematoxilina y eosina. En algunas modalidades, la tinción incluye el uso de una etiqueta detectable seleccionada del grupo que consiste en un radioisótopo, un fluoróforo, un compuesto quimioluminiscente, un compuesto bioluminiscente o una de sus combinaciones.

En algunas modalidades, la matriz espacial incluye una o más series de dilución de proteínas.

En la presente descripción se proporcionan además matrices espaciales que incluyen: a) una pluralidad de sondas de captura que incluyen códigos de barras espaciales y una primera pluralidad de dominios de captura hibridados con una pluralidad de productos de ligazón con plantilla, y b) una pluralidad de sondas de captura que incluyen códigos de barras espaciales y una segunda pluralidad de dominios de captura hibridados con una pluralidad de oligonucleótidos de los agentes de captura de analitos, donde los oligonucleótidos incluyen una secuencia de captura de analito y un código de barras de la fracción de unión al analito.

En algunas modalidades, las sondas de captura incluyen dominios de escisión, identificadores moleculares únicos, una o más secuencias funcionales, o una de sus combinaciones.

En algunas modalidades, la primera pluralidad de dominios de captura son secuencias homopoliméricas o secuencias no homopoliméricas definidas. En algunas modalidades, la segunda pluralidad de dominios de captura son secuencias homopoliméricas o secuencias no homopoliméricas definidas. En algunas modalidades, la primera pluralidad de dominios de captura son secuencias de poli(T). En algunas modalidades, la matriz espacial incluye una o más series de dilución de proteínas.

En la presente descripción se proporcionan además kits que incluyen: a) una matriz espacial que incluye una pluralidad de sondas de captura, donde las sondas de captura incluyen códigos de barras espaciales y donde la pluralidad de sondas de captura incluye una primera pluralidad de primeros dominios de captura y una segunda pluralidad de segundos dominios de captura, b) uno o más agentes de captura de analito, c) uno o más pares de sondas de ligazón con plantilla con ácido nucleico, y d) una o más enzimas y tampones para la práctica de cualquiera de los métodos descritos en la presente descripción.

También se proporcionan en la presente descripción métodos para determinar la localización espacial de un ácido nucleico y una proteína en una muestra biológica de enfermedad que incluye: a) proporcionar una matriz espacial que incluye una primera y una segunda pluralidad de sondas de captura donde cada pluralidad incluye un código de barras espacial y un dominio de captura, b) poner en contacto la muestra biológica de enfermedad con la matriz espacial, c) poner en contacto la muestra biológica de enfermedad con: (i) una pluralidad de agentes de captura de analito, donde un agente de captura de analito incluye un resto de unión al analito y un oligonucleótido que incluye un código de barras de un resto de unión al analito y una secuencia de captura de analito, donde la secuencia de captura de analito incluye una secuencia complementaria a una segunda pluralidad de dominios de captura, y (ii) una pluralidad de sondas de ligazón con plantilla, donde una de las sondas de ligazón con plantilla incluye una secuencia complementaria a una primera pluralidad de dominios de captura, d) unir el resto de unión al analito del agente de captura de analito a una proteína objetivo, e) hibridar las sondas de ligazón con plantilla a un ARN objetivo y ligar las sondas para producir productos de ligazón con plantilla, f) hibridar los productos de ligazón con plantilla a la primera pluralidad de dominios de captura y las secuencias de captura de analito de los agentes de captura de analito unidos a la segunda pluralidad de dominios de captura en la matriz espacial, y g) determinar la secuencia o una porción de esta de un producto de ligazón capturado, o un complemento, y la secuencia del código de barras espacial de la sonda de captura, o un complemento de esta, que se asocia con el producto de ligazón, y la secuencia del resto de unión al analito de la sonda de captura, o un complemento de esta, del agente de captura de analito unido, para determinar de esta manera la localización espacial de un ácido nucleico y la proteína en la muestra biológica de enfermedad.

En algunas modalidades, la muestra biológica de enfermedad es una muestra biológica cancerosa. En algunas modalidades, la muestra biológica cancerosa es una muestra biológica de cáncer de ovario, una muestra biológica de cáncer de mama, una muestra biológica de cáncer de pulmón o un melanoma. En algunas modalidades, la muestra de cáncer de mama es cáncer de mama triple positivo o una muestra de carcinoma ductal invasivo.

Cuando los valores se describen en términos de intervalos, debe entenderse que la descripción incluye la descripción de todos los posibles subintervalos dentro de tales intervalos, así como también valores numéricos específicos que caen dentro de tales intervalos independientemente de si se indica expresamente un valor numérico específico o un subintervalo específico.

El término “cada uno”, cuando se usa en referencia a una recopilación de artículos, pretende identificar un artículo individual en la recopilación, pero no necesariamente se refiere a cada artículo en la recopilación, a menos que se indique expresamente de cualquier otra manera, o a menos que el contexto del uso indique claramente de cualquier otra manera.

En la presente descripción se describen diversas modalidades de las características de esta descripción. Sin embargo, debe entenderse que tales modalidades se proporcionan meramente a manera de ejemplo, y pueden ocurrir numerosas variaciones, cambios, y sustituciones para los expertos en la técnica.

Descripción de los dibujos

Los siguientes dibujos ilustran ciertas modalidades de las características y ventajas de esta descripción. Estas modalidades no pretenden limitar el alcance de las reivindicaciones adjuntas de ninguna manera. Los símbolos de referencia similares en los dibujos indican elementos similares.

La Figura 1 es un diagrama esquemático que muestra un ejemplo de una sonda de captura con código de barras, como se describe en la presente descripción.

La Figura 2 es un diagrama esquemático de un agente de captura de analito ilustrativo.

La Figura 3 es un diagrama esquemático que representa una interacción ilustrativa entre una sonda de captura inmovilizada en la característica 324 y un agente de captura de analito 326.

Las Figuras 4A-B muestra datos de agrupamiento ilustrativos para secciones de ganglios linfáticos humanos FFPE mediante el uso de un panel de prueba de conjugado anticuerpo-oligonucleótido en formato de ocho. La Figura 4A (ARN) muestra el agrupamiento con respecto a la expresión génica del panel de prueba y la Figura 4B (proteína) muestra el agrupamiento con respecto a la expresión proteica del panel de prueba.

La Figura 5 muestra datos de agrupamiento ilustrativos de genes espaciales (superior central) y expresión proteica (superior derecha) en secciones de tejido de amígdala FFPE superpuestas sobre la imagen de H&E (superior izquierda). Las figuras inferiores muestran la visualización de los marcadores de proteínas PD-1, Ki67 y CD8A que etiquetan las células T foliculares, los folículos individuales y las células T supresoras/citotóxicas, respectivamente.

La Figura 6 muestra una sección de tejido de amígdala FFPE teñida con H&E ilustrativa (parte superior izquierda), datos de expresión génica espacial (parte superior central) y datos de expresión proteica espaciales (parte superior izquierda). La Figura 6 también muestra una vista ampliada de una sección del tejido de las amígdalas FFPE teñido con H&E (parte inferior izquierda) y el agrupamiento de la expresión génica y la expresión proteica de CD8a (parte inferior del medio) y CD4 (parte inferior derecha) superpuestos en la sección de la sección de tejido de las amígdalas FFPE de H&E.

La Figura 7 muestra la expresión espacial ilustrativa de genes y proteínas a partir de un subconjunto de datos de expresión de genes y proteínas de tejido de amígdala. Se muestra un subconjunto de dieciocho mapas de calor de expresión génica y proteica objetivos diferentes.

La Figura 8 muestra vistas ampliadas de datos de mapas de calor ilustrativos para la sección de tejido de amígdala humana FFPE mediante el uso de un panel de prueba de conjugado anticuerpo-oligonucleótido en formato de 12 de varios genes que se muestran en la Figura 7.

Las Figuras 9A-C muestran la expresión espacial ilustrativa de genes y proteínas en secciones de tejido de cáncer de mama triple positivo FFPE. La Figura 9A muestra una sección de tejido de cáncer de mama triple positivo FFPE teñida con H&E. La Figura 9B muestra los grupos representativos de expresión génica y la Figura 9C muestra los grupos representativos de expresión proteica. Los insertos mostrados en las Figuras 9B y 9C muestran gráficos tSNE representativos.

La Figura 10 muestra una vista ampliada de la expresión de ARN y proteína de Her2 y Vimentina en secciones de tejido de carcinoma ductal invasivo FFPE. Las secciones adyacentes se tiñeron por inmunofluorescencia para Her2 (arriba) y Vimentina (abajo) demostrando una señal fuerte dentro de la región del cuadro discontinuo de la muestra de biopsia.

Las Figuras 11A-C muestran la expresión espacial ilustrativa de genes y proteínas en secciones de tejido de carcinoma ductal invasivo FFPE. La Figura 11B muestra los grupos representativos de expresión génica y Figura 11C muestra los grupos representativos de expresión proteica superpuestos en la tinción H&E (Figura 11A) secciones de tejido de carcinoma ductal invasivo.

La Figura 12 muestra la expresión espacial ilustrativa de genes y proteínas a partir de un subconjunto de los datos de expresión génica y proteica de las Figuras 11B-C. Los biomarcadores se usaron para identificar dos regiones de interés en la muestra de tejido de carcinoma ductal invasivo.

La Figuras 13 muestra la expresión espacial ilustrativa de genes y proteínas en la sección de tejido de carcinoma ductal invasivo en FFPE superpuesta a la imagen de H&E del gen ACTA2 y Epcam y sus proteínas correspondientes.

La Figura 14A muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica del ARN del receptor de la proteína tirosina fosfatasa de tipo C (Ptprc) mediante el uso de sondas ligadas en tejido de bazo de ratón FFPE bajo condiciones de configuración de intercalamiento.

La Figura 14B muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica del ARN de Ptprc mediante el uso de sondas ligadas en tejido de bazo FFPE bajo condiciones de control de configuración sin intercalamiento.

La Figura 14C muestra una imagen ilustrativa de información espacialmente resuelta de la expresión proteica de CD45R mediante el uso de agentes de captura de analitos en tejido de bazo de ratón FFPE bajo condiciones de configuración de intercalamiento.

La Figura 14D muestra una imagen ilustrativa de información resuelta espacialmente de la expresión proteica de CD45R mediante el uso de agentes de captura de analitos en tejido de bazo de ratón FFPE bajo condiciones de control de configuración sin intercalamiento.

La Figura 15A muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica del ARN de tirosina hidroxilasa (Th) mediante el uso de sondas ligadas en tejido cerebral de ratón FFPE bajo condiciones de configuración de intercalamiento.

La Figura 15B muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica de Th ARN mediante el uso de sondas ligadas en tejido cerebral de ratón FFPE bajo condiciones de control de configuración sin intercalamiento.

La Figura 15C muestra una imagen ilustrativa de información espacialmente resuelta de la expresión proteica de la proteína tirosina hidroxilasa (TH) mediante el uso de agentes de captura de analitos en tejido cerebral de ratón FFPE bajo condiciones de configuración de intercalamiento.

La Figura 15D muestra una imagen ilustrativa de información espacialmente resuelta de la expresión proteica de TH mediante el uso de agentes de captura de analitos en tejido cerebral de ratón FFPE bajo condiciones de control de configuración sin intercalamiento.

La Figura 16A muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica de ARN de Ter119 mediante el uso de sondas ligadas en tejido de torso de embrión de ratón FFPE en condiciones de configuración de intercalamiento.

La Figura 16B muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica de ARN de Ter119 mediante el uso de sondas ligadas en tejido de torso de embrión de ratón FFPE en condiciones de configuración sin intercalamiento.

La Figura 16C muestra una imagen ilustrativa de información espacialmente resuelta de la expresión proteica de Ter119 mediante el uso de un agente de captura de analito en tejido de torso de embriones de ratón FFPE en condiciones de configuración de intercalamiento.

La Figura 16D muestra una imagen ilustrativa de información resuelta espacialmente de la expresión proteica de Ter119 mediante el uso de agentes de captura de analitos en tejido de torso de embriones de ratón FFPE bajo condiciones de configuración sin intercalamiento.

La Figura 17A muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica de ARN de Ter119 mediante el uso de sondas ligadas en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración de intercalamiento.

La Figura 17B muestra una imagen ilustrativa de información espacialmente resuelta de la expresión génica de ARN de Ter119 mediante el uso de sondas ligadas en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración sin intercalamiento.

La Figura 17C muestra una imagen ilustrativa de información espacialmente resuelta de la expresión proteica de Ter119 mediante el uso de agentes de captura de analitos en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración de intercalamiento.

La Figura 17D muestra una imagen ilustrativa de información espacialmente resuelta de la expresión proteica de Ter119 mediante el uso de un agente de captura de analito en la cabeza del embrión de ratón FFPE y el tejido del torso superior mediante el uso de condiciones de configuración sin intercalamiento.

Las Figuras 18A, 19A, 20A, 21A, y 22A muestran imágenes ilustrativas de información espacialmente resuelta de la expresión génica de ARN de Mpped1 (Figura 18A), Tnnc1 (Figura 19A), Fgf15 (Figura 20A), Epyc (Figura 21A), y Serpina1e (Figura 22A) mediante el uso de sondas ligadas en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración de intercalamiento.

Las Figuras 18B, 19B, 20B, 21B, y 22B muestran imágenes ilustrativas de información espacialmente resuelta de la expresión génica de ARN de Mpped1 (Figura 18B), Tnnc1 (Figura 19B), Fgf15 (Figura 20B), Epyc (Figura 21B), y Serpina1e (Figura 22B) mediante el uso de sondas ligadas en tejido de torso de embrión de ratón FFPE bajo condiciones de configuración de intercalamiento.

Las Figuras 18C, 19C, 20C, 21C, y 22C muestran imágenes ilustrativas de información espacialmente resuelta de la expresión génica de ARN de Mpped1 (Figura 18C), Tnnc1 (Figura 19C), Fgf15 (Figura 20C), Epyc (Figura 21C), y Serpina1e (Figura 22C) mediante el uso de sondas ligadas en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración sin intercalamiento.

Las Figuras 18D, 19D, 20D, 21D, y 22D muestran imágenes ilustrativas de información espacialmente resuelta de la expresión génica de ARN de Mpped1 (Figura 18D), Tnnc1 (Figura 19D), Fgf15 (Figura 20D), Epyc (Figura 21D), y Serpina1e (Figura 22D) mediante el uso de sondas ligadas en tejido de torso de embrión de ratón FFPE bajo condiciones de configuración sin intercalamiento.

Las Figuras 18E, 19E, 20E, 21E, y 22E muestran imágenes ilustrativas de información espacialmente resuelta de la expresión proteica de Mpped1 (Figura 18E), Tnnc1 (Figura 19E), Fgf15 (Figura 20E), Epyc (Figura 21E), y Serpina1e (Figura 22E) mediante el uso de agentes de captura de analito en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración de intercalamiento.

Las Figuras 18F, 19F, 20F, 21F, y 22F muestran imágenes ilustrativas de información espacialmente resuelta de la expresión proteica de Mpped1 (Figura 18F), Tnnc1 (Figura 19F), Fgf15 (Figura 20F), Epyc (Figura 21F), y Serpina1e (Figura 22F) mediante el uso de agentes de captura de analitos en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración sin intercalamiento.

Las Figuras 18G, 19G, 20G, 21G, y 22G muestran imágenes ilustrativas de información espacialmente resuelta de la expresión proteica de Mpped1 (Figura 18G), Tnnc1 (Figura 19G), Fgf15 (Figura 20G), Epyc (Figura 21G), y Serpina1e (Figura 22G) mediante el uso de agentes de captura de analito en tejido de torso de embriones de ratón FFPE bajo condiciones de configuración sin intercalamiento.

Las Figuras 18H, 19H, 20H, 21H, y 22H muestran imágenes ilustrativas de información espacialmente resuelta de la expresión génica de ARN de Mpped1 (Figura 18H), Tnnc1 (Figura 19H), Fgf15 (Figura 20H), Epyc (Figura 21H), y Serpina1e (Figura 22H) mediante el uso de sondas ligadas en la cabeza de embriones de ratón FFPE y el tejido del torso superior en condiciones de configuración sin intercalamiento en el que solo se detectó ARN.

Las Figuras 18I, 19I, 20I, 21I, y 22I muestran imágenes ilustrativas de información espacialmente resuelta de la expresión génica de ARN de Mpped1 (Figura 18I), Tnnc1 (Figura 19I), Fgf15 (Figura 20I), Epyc (Figura 21I), y Serpina1e (Figura 22I) mediante el uso de sondas ligadas en tejido de torso de embrión de ratón FFPE bajo condiciones de configuración sin intercalamiento en el que solo se detectó ARN.

Las Figuras 23A-E muestran la infiltración espacial de células inmunitarias en una sección de tejido de cáncer de mama se correlaciona con las anotaciones del patólogo. La Figura 23A muestra una sección de tejido de cáncer de mama humano FFPE teñida con H&E. La Figura 23B muestra las anotaciones del patólogo. La Figura 23C muestra la expresión proteica espacial de CD3 (por ejemplo, un marcador para los linfocitos T); la Figura 23D muestra la expresión proteica espacial de CD8A (por ejemplo, un marcador para linfocitos T citotóxicos); y la Figura 23E muestra la expresión proteica espacial para HLA-DR (por ejemplo, un marcador para la activación de linfocitos T).

Las Figuras 24A-D muestran la infiltración espacial de células inmunitarias en una sección de tejido FFPE de cáncer de ovario se correlaciona con las anotaciones del patólogo. La Figura 24A muestra una anotación del patólogo para el carcinoma invasivo y las células inmunitarias. La Figura 24B muestra la expresión proteica espacial de CD20 (por ejemplo, un marcador para las células B); la Figura 24C muestra la expresión proteica espacial de CD68 (por ejemplo, un marcador para monocitos); y la Figura 24D muestra la expresión proteica espacial de CD8A (por ejemplo, un marcador para linfocitos T citotóxicos).

Las Figuras 25A-D muestran la infiltración diferencial de linfocitos T citotóxicos espaciales dentro de diferentes regiones de una sección de tejido de FFPE de cáncer de ovario. La Figura 25A muestra la expresión proteica espacial de CD8A (por ejemplo, un marcador para linfocitos T citotóxicos). La Figura 25B muestra áreas altamente infiltradas ("calientes", CD3) y no filtradas ("frías", CD8) de expresión proteica en la sección de tejido FFPE mostrada en la Figura 25A. La Figura 25C muestra la expresión proteica espaciales de HLA-G y Figura 25D muestra la expresión espacial de proteínas de IMPG2.

La Figura 26 muestra una matriz espacial ilustrativa con múltiples series de dilución de proteínas que se ubican en la matriz espacial.

Las Figuras 27A-D muestra un gen espacial ilustrativo (Figura 27B) y la expresión proteica (Figura 27C) en la expresión en el tejido FFPE de cáncer de pulmón. La Figura 27A muestra un tejido de cáncer de pulmón FFPE teñido con H&E y la Figura 27D muestra un gráfico de UMI espacial de proteínas HLA-DR.

Las Figuras 28A-D muestra la expresión espacial ilustrativa de genes (Figura 28B) y proteínas (Figura 28C) en tejido FFPE de melanoma. La Figura 28A muestra un tejido canceroso de melanoma teñido con H&E y la Figura 28D muestra un gráfico de UMI espacial de proteínas HLA-DR.

Las Figuras 29A-C muestra la tinción por inmunofluorescencia de anticuerpos contra vimentina ilustrativa y tinción con DAPI (Figura 29A), la expresión espacial de proteínas ilustrativa (Figura 29B), y la expresión proteica espacial ilustrativa de Vimentina (Figura 29C) en una sección de tejido de cáncer de mama ductal invasivo FFPE de grado II.

Descripción detallada

La presente descripción presenta métodos, composiciones y kits para el análisis espacial de muestras biológicas. Más específicamente, la presente descripción presenta métodos, composiciones y kits tanto para la expresión génica espacial como para la expresión proteica espacial en una muestra biológica.

Las metodologías y composiciones de análisis espacial descritas en la presente descripción pueden proporcionar una gran cantidad de analito y/o datos de expresión para una variedad de analitos dentro de una muestra biológica con alta resolución espacial, mientras se retiene el contexto espacial nativo. Los métodos y composiciones de análisis espacial pueden incluir, por ejemplo, usar una sonda de captura que incluye un código de barras espacial (por ejemplo, una secuencia de ácido nucleico que proporciona información sobre la localización o posición de un analito dentro de una célula o una muestra de tejido (por ejemplo, una célula de mamífero o una muestra de tejido de mamífero) y un dominio de captura que es capaz de unirse a un analito (por ejemplo, una proteína y/o un ácido nucleico) producido por y/o presente en una célula. Los métodos y composiciones de análisis espacial también pueden incluir usar una sonda de captura que tiene un dominio de captura que captura un agente intermedio para la detección indirecta de un analito. Por ejemplo, el agente intermedio puede incluir una secuencia de ácido nucleico (por ejemplo, un código de barras) asociada con el agente intermedio. Por lo tanto, la detección del agente intermedio es indicativa del analito en la muestra de células o tejido.

Los aspectos no limitativos de las metodologías y composiciones de análisis espacial se describen en las Patentes de Estados Unidos núms. 10,774,374, 10,724,078, 10,480,022, 10,059,990, 10,041,949, 10,002,316, 9,879,313, 9,783,841, 9,727,810, 9,593,365, 8,951,726, 8,604,182, 7,709,198, las publicaciones de solicitud de patente de Estados Unidos núms. 2020/239946, 2020/080136, 2020/0277663, 2020/024641, 2019/330617, 2019/264268, 2020/256867, 2020/224244, 2019/194709, 2019/161796, 2019/085383, 2019/055594, 2018/216161, 2018/051322, 2018/0245142, 2017/241911, 2017/089811, 2017/067096, 2017/029875, 2017/0016053, 2016/108458, 2015/000854, 2013/171621, WO 2018/091676, WO 2020/176788, Rodriques y otros, Science 363(6434):1463-1467, 2019; Lee y otros, Nat. Protoc. 10(3):442-458, 2015; Trejo y otros, PLoS ONE 14(2):e0212031, 2019; Chen y otros, Science 348(6233):aaa6090, 2015; Gao y otros, BMC Biol. 15:50, 2017; y Gupta y otros, Nature Biotechnol. 36:1197-1202, 2018; la Guía del usuario de los kits de reactivos de expresión génica espacial de Visium (por ejemplo, Rev C, con fecha de junio de 2020) y/o la Guía del usuario de los kits de reactivos de optimización de tejidos espaciales de Visium (por ejemplo, Rev C, con fecha de julio de 2020), ambas disponibles en el sitio web de documentación 10x Genomics Support y puede usarse en la presente descripción en cualquier combinación. Vickovic y otros (bioRxiv, DOI 10.1101/2020.10.14.338418, 2020), describe una plataforma de alto rendimiento completamente automatizada para la transcriptómica combinada y la proteómica basada en anticuerpos resuelta espacialmente. En la presente descripción se describen aspectos adicionales no limitantes de las metodologías y composiciones de análisis espacial.

Alguna terminología general que puede usarse en esta descripción puede encontrarse en la Sección (I)(b) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. Típicamente, un "código de barras" es una etiqueta o identificador que transmite o es capaz de transmitir información (por ejemplo, información sobre un analito en una muestra, una perla y/o una sonda de captura). Un código de barras puede ser parte de un analito o independiente de un analito. Un código de barras puede unirse a un analito. Un código de barras particular puede ser único con relación a otros códigos de barras. A los efectos de esta descripción, un "analito" puede incluir cualquier sustancia, estructura, resto o componente biológico a analizar. El término "diana" puede referirse de manera similar a un analito de interés.

Los analitos pueden clasificarse en términos generales en uno de dos grupos: analitos de ácidos nucleicos y analitos no ácidos nucleicos. Los ejemplos de analitos no ácidos nucleicos incluyen, pero sin limitarse a, lípidos, carbohidratos, péptidos, proteínas, glicoproteínas (unidas a N o unidas a O), lipoproteínas, fosfoproteínas, variantes específicas de

proteínas fosforiladas o acetiladas, variantes de amidación de proteínas, variantes de hidroxilación de proteínas, variantes de metilación de proteínas, variantes de ubiquitinación de proteínas, variantes de sulfatación de proteínas, proteínas virales (por ejemplo, cápside viral, envoltura viral, cubierta viral, accesorio viral, glicoproteínas virales, pico viral, etc.), proteínas extracelulares e intracelulares, anticuerpos y fragmentos de unión a antígeno. En algunas modalidades, el(los) analito(s) puede(n) localizarse en localización(ones) subcelular(es), incluidas, por ejemplo, orgánulos, por ejemplo, mitocondrias, aparato de Golgi, retículo endoplásmico, cloroplastos, vesículas endocíticas, vesículas exocíticas, vacuolas, lisosomas, etc. En algunas modalidades, el(los) analito(s) puede(n) ser péptidos o proteínas, incluidos sin limitación, anticuerpos y enzimas. Pueden encontrarse ejemplos adicionales de analitos en la Sección (I)(c) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. En algunas modalidades, un analito puede detectarse indirectamente, tal como mediante la detección de un agente intermedio, por ejemplo, un producto de ligazón o un agente de captura de analito (por ejemplo, un anticuerpo conjugado con oligonucleótido), tales como los descritos en la presente descripción.

Una "muestra biológica" se obtiene típicamente del sujeto para su análisis mediante el uso de cualquiera de una variedad de técnicas que incluyen, pero sin limitarse a, biopsia, cirugía y microscopía de captura láser (LCM), y generalmente incluye células y/u otro material biológico del sujeto. En algunas modalidades, una muestra biológica puede ser una sección de tejido. En algunas modalidades, una muestra biológica puede ser una muestra biológica fijada y/o teñida (por ejemplo, una sección de tejido fijada y/o teñida). Los ejemplos no limitantes de tinciones incluyen tinciones histológicas (por ejemplo, hematoxilina y/o eosina) y tinciones inmunológicas (por ejemplo, tinciones fluorescentes). En algunas modalidades, pueden obtenerse imágenes de una muestra biológica (por ejemplo, una muestra biológica fijada y/o teñida). Las muestras biológicas también se describen en la Sección (I)(d) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

En algunas modalidades, una muestra biológica se permeabiliza con uno o más reactivos de permeabilización. Por ejemplo, la permeabilización de una muestra biológica puede facilitar la captura del analito. Las condiciones y agentes de permeabilización ilustrativos se describen en la Sección (I)(d)(ii)(13) o la Sección de modalidades ilustrativas del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

Los métodos de análisis espaciales basados en matrices implican la transferencia de uno o más analitos de una muestra biológica a una matriz de elementos en un sustrato, donde cada elemento está asociado con una localización espacial única en la matriz. El análisis subsecuente de los analitos transferidos incluye determinar la identidad de los analitos y la localización espacial de los analitos dentro de la muestra biológica. La localización espacial de un analito dentro de la muestra biológica se determina basándose en el elemento al que está unido el analito (por ejemplo, directa o indirectamente) en la matriz, y la localización espacial relativa del elemento dentro de la matriz.

Una "sonda de captura" se refiere a cualquier molécula capaz de capturar (directa o indirectamente) y/o etiquetar un analito (por ejemplo, un analito de interés) en una muestra biológica. En algunas modalidades, la sonda de captura es un ácido nucleico o un polipéptido. En algunas modalidades, la sonda de captura incluye un código de barras (por ejemplo, un código de barras espacial y/o un identificador molecular único (UMI)) y un dominio de captura. En algunas modalidades, una sonda de captura puede incluir un dominio de escisión y/o un dominio funcional (por ejemplo, un sitio de unión a cebador, tal como para secuenciación de próxima generación (NGS)). Ver, por ejemplo, la Sección (II)(b) (por ejemplo, las subsecciones (i)-(vi)) del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. La generación de sondas de captura puede lograrse mediante cualquier método apropiado, incluidos los descritos en la Sección (II)(d)(ii) del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

La Figura 1 es un diagrama esquemático que muestra una sonda de captura ilustrativa, como se describe en la presente descripción. Como se muestra, la sonda de captura 102 está opcionalmente acoplada a un elemento 101 por un dominio de escisión 103, tal como un enlazador disulfuro. La sonda de captura puede incluir una secuencia funcional 104 que sea útil para su procesamiento subsecuente. La secuencia funcional 104 puede incluir toda o una parte de la secuencia de unión de células de flujo específica del secuenciador (por ejemplo, una secuencia P5 o P7), toda o una parte de una secuencia de cebador de secuenciación (por ejemplo, un sitio de unión al cebador R1, un sitio de unión al cebador R2), o sus combinaciones. La sonda de captura también puede incluir un código de barras espacial 105. La sonda de captura también puede incluir una secuencia de identificador molecular único (UMI) 106. Mientras que la Figura 1 muestra el código de barras espacial 105 como ubicado aguas arriba (5') de la secuencia UMI 106, debe entenderse que las sondas de captura en donde la secuencia UMI 106 está ubicada aguas arriba (5') del código de barras espacial 105 también son adecuadas para su uso en cualquiera de los métodos descritos en la presente descripción. La sonda de captura también puede incluir un dominio de captura 107 para facilitar la captura de un analito objetivo. En algunas modalidades, la sonda de captura comprende una o más secuencias funcionales adicionales que se pueden ubicar, por ejemplo, entre el código de barras espacial 105 y la secuencia de UMI 106, entre la secuencia de UMI 106 y el dominio de captura 107, o siguiendo el dominio de captura 107. El dominio de captura puede tener una secuencia complementaria a una secuencia de un analito de ácido nucleico. El dominio de captura puede tener una secuencia complementaria a una sonda conectada descrita en la presente descripción. El dominio de captura puede tener una secuencia complementaria a una secuencia de control de captura presente en un agente de captura de analito. El dominio de captura puede tener una secuencia complementaria a un oligonucleótido de férula. Tal oligonucleótido de férula, además de tener una secuencia complementaria a un dominio de captura de una sonda de

captura, puede tener una secuencia de un analito de ácido nucleico, una secuencia complementaria a una porción de una sonda conectada descrita y/o una secuencia de control de captura descrita en la presente descripción.

Las secuencias funcionales generalmente pueden seleccionarse para compatibilidad con cualquiera de una variedad de sistemas de secuenciación diferentes, por ejemplo, Ion Torrent Proton o PGM, instrumentos de secuenciación Illumina, PacBio, Oxford Nanopore, etc., y los requisitos de los mismos. En algunas modalidades, las secuencias funcionales pueden seleccionarse para compatibilidad con sistemas de secuenciación no comercializados. Ejemplos de tales sistemas y técnicas de secuenciación, para los cuales pueden usarse secuencias funcionales adecuadas, incluyen (pero no se limitan a) secuenciación Ion Torrent Proton o PGM, secuenciación Illumina, secuenciación PacBio SMRT y secuenciación Oxford Nanopore. Además, en algunas modalidades, las secuencias funcionales pueden seleccionarse para que sean compatibles con otros sistemas de secuenciación, incluidos sistemas de secuenciación no comercializados.

En algunas modalidades, el código de barras espacial 105 y secuencias funcionales 104 es común a todas las sondas unidas a un elemento determinado. En algunas modalidades, la secuencia UMI 106 de una sonda de captura unida a un elemento dado es diferente de la secuencia UMI de una sonda de captura diferente unida al elemento dado.

En algunas modalidades, puede detectarse más de un tipo de analito (por ejemplo, ácidos nucleicos y proteínas) de una muestra biológica (por ejemplo, simultánea o secuencialmente) mediante el uso de cualquier técnica en formato múltiple apropiada, tales como las descritas en la Sección (IV) del documento WO 2020/176788 y/o publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

En algunas modalidades, la detección de uno o más analitos (por ejemplo, analitos proteicos) puede realizarse mediante el uso de uno o más agentes de captura de analito. Como se usa en la presente, un “agente de captura de analito” se refiere a un agente que interactúa con un analito (por ejemplo, un analito en una muestra biológica) y con una sonda de captura (por ejemplo, una sonda de captura unida a un sustrato o una característica) para identificar el analito. En algunas modalidades, el agente de captura de analito incluye: (i) un resto de unión al analito (por ejemplo, que se une a un analito), por ejemplo, un anticuerpo o fragmento de unión a antígeno del mismo; (ii) código de barras del resto de unión al analito; y (iii) una secuencia de captura de analito. En algunas modalidades, el agente de captura de analito incluye un dominio de código de barras del agente de captura que se conjuga o se une de cualquier otra manera a la porción de unión al analito. En algunas modalidades, el dominio de código de barras del agente de captura se une covalentemente a la porción de unión al analito. En algunas modalidades, un dominio de código de barras del agente de captura es una secuencia de ácido nucleico. En algunas modalidades, un dominio de código de barras del agente de captura incluye un código de barras del resto de unión al analito y una secuencia de captura de analito.

En algunas modalidades, los agentes de captura de analito son capaces de unirse a analitos presentes dentro de una célula. En algunas modalidades, los agentes de captura de analito son capaces de unirse a analitos de la superficie celular que pueden incluir, sin limitación, un receptor, un antígeno, una proteína de superficie, una proteína transmembrana, un conglomerado de proteínas de diferenciación, una proteína canal, una proteína bomba, una proteína portadora, un fosfolípido, una glucoproteína, un glucolípido, un complejo de proteína de interacción célula-célula, un complejo presentador de antígenos, un complejo mayor de histocompatibilidad, un receptor de linfocitos T manipulado genéticamente, un receptor de linfocitos T, un receptor de linfocitos B, un receptor de antígeno quimérico, una proteína de la matriz extracelular, una modificación postraducciona (por ejemplo, fosforilación, glucosilación, ubiquitinación, nitrosilación, metilación, acetilación o lipidación) del estado de una proteína de la superficie celular, una unión en hendidura, y una unión adherente. En algunas modalidades, los agentes de captura de analito son capaces de unirse a analitos de la superficie celular que se modifican postraduccionalmente. En tales modalidades, los agentes de captura de analito pueden ser específicos para los analitos de la superficie celular en base a un estado dado de modificación postraducciona (por ejemplo, fosforilación, glucosilación, ubiquitinación, nitrosilación, metilación, acetilación o lipidación), de manera que un perfil del analito de la superficie celular puede incluir información de modificación postraducciona de uno o más analitos.

Como se usa en la presente, la expresión “resto de unión al analito” se refiere a una molécula o resto capaz de unirse a un constituyente macromolecular (por ejemplo, un analito, por ejemplo, un analito biológico). En algunas modalidades de cualquiera de los métodos de perfilamiento espacial descritos en la presente descripción, el resto de unión al analito del agente de captura de analito que se une a un analito biológico puede incluir, pero no se limita a, un anticuerpo, o un fragmento de unión al epítipo del mismo, una molécula de unión al receptor de superficie celular, un ligando del receptor, una molécula pequeña, un anticuerpo biespecífico, un activador de linfocitos T biespecífico, un activador de receptores de células T, un activador de receptores de células B, un pro cuerpo, un aptámero, un monocuerpo, un afímero, una darpina, y un andamio de proteínas, o cualquiera de sus combinaciones. La porción de unión al analito puede unirse al constituyente macromolecular (por ejemplo, analito) con alta afinidad y/o con alta especificidad. La porción de unión al analito puede incluir una secuencia de nucleótidos (por ejemplo, un oligonucleótido), que puede corresponder a al menos una porción o a la totalidad del resto de unión al analito. La porción de unión al analito puede incluir un polipéptido y/o un aptámero (por ejemplo, un polipéptido y/o un aptámero que se une a una molécula objetivo específica, por ejemplo, un analito). La porción de unión al analito puede incluir un anticuerpo o fragmento de anticuerpo (por ejemplo, un fragmento de unión a antígeno) que se une a un analito específico (por ejemplo, un polipéptido).

En algunas modalidades, una porción de unión al analito de un agente de captura de un analito incluye uno o más anticuerpos o fragmentos de unión a antígeno de estos. Los anticuerpos o fragmentos de unión a antígeno que incluyen la porción de unión al analito pueden unirse específicamente a un analito objetivo. En algunas modalidades, el analito es una proteína (por ejemplo, una proteína sobre una superficie de la muestra biológica (por ejemplo, una célula) o una proteína intracelular). En algunas modalidades, una pluralidad de agentes de captura de analito que comprenden una pluralidad de porciones de unión al analito se une a una pluralidad de analitos presentes en una muestra biológica. En algunas modalidades, la pluralidad de analitos incluye una única especie del analito (por ejemplo, una única especie de polipéptido). En algunas modalidades en las que la pluralidad de analitos incluye una única especie del analito, las porciones de unión al analito de la pluralidad de agentes de captura de analito son los mismos. En algunas modalidades en las que la pluralidad de analitos incluye una única especie del analito, las porciones de unión al analito de la pluralidad de agentes de captura de analito son diferentes (por ejemplo, los miembros de la pluralidad de agentes de captura de analito pueden tener dos o más especies de porciones de unión al analito, en donde cada una de las dos o más especies de porciones de unión al analito se unen a una única especie del analito, por ejemplo, en sitios de unión diferentes). En algunas modalidades, la pluralidad de analitos incluye múltiples especies diferentes de analitos (por ejemplo, múltiples especies diferentes de polipéptidos).

Como se usa en la presente, el término “código de barras del resto de unión al analito” se refiere a un código de barras que se asocia con, o de cualquier otra manera, identifica la porción de unión al analito. En algunos casos, un código de barras del resto de unión al analito (o una porción del mismo) puede eliminarse (por ejemplo, escindirse) del agente de captura de analito. En algunas modalidades, al identificar un resto de unión al analito y su código de barras de resto de unión al analito asociado, también se puede identificar el analito al que se une el resto de unión al analito. Un código de barras del resto de unión al analito puede ser una secuencia de ácido nucleico de una longitud y/o secuencia dadas que se asocia con la porción de unión al analito. Un código de barras del resto de unión al analito puede incluir generalmente cualquiera de la variedad de aspectos de códigos de barras descritos en la presente descripción. Por ejemplo, un agente de captura de analito que es específico para un tipo de analito puede tener acoplado a este un primer dominio de código de barras del agente de captura (por ejemplo, que incluye un primer código de barras del resto de unión al analito), mientras que un agente de captura de analito que es específico para un analito diferente puede tener un dominio de código de barras del agente de captura diferente (por ejemplo, que incluye un segundo código de barras del resto de unión al analito) que se acopla a este. En algunos aspectos, tal dominio de código de barras del agente de captura puede incluir un código de barras del resto de unión al analito que permite la identificación del resto de unión al analito al que se acopla el dominio de código de barras del agente de captura. La selección del dominio de código de barras del agente de captura puede permitir una diversidad significativa en términos de secuencia, mientras que también se puede unir fácilmente a la mayoría de los restos de unión al analito (por ejemplo, anticuerpos o aptámeros) así como también detectarse fácilmente, (por ejemplo, mediante el uso de tecnologías de secuenciación o de matriz). Puede encontrarse una descripción adicional de los agentes de captura de analito en la Sección (II)(b)(ix) del documento WO 2020/176788 y/o en la Sección (II)(b)(viii) de la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

En algunas modalidades, el dominio de código de barras del agente de captura de un agente de captura de analito incluye una secuencia de captura de analito. Como se usa en la presente, el término “secuencia de captura de analito” se refiere a una región o resto configurado para hibridarse, unirse, acoplarse o interactuar de cualquier otra manera con un dominio de captura de una sonda de captura. En algunas modalidades, una secuencia de captura de analito incluye una secuencia de ácido nucleico que es complementaria o sustancialmente complementaria al dominio de captura de una sonda de captura de manera que la secuencia de captura de analito se hibrida con el dominio de captura de la sonda de captura. En algunas modalidades, una secuencia de captura de analito comprende una secuencia de ácido nucleico poli(A) que se hibrida con un dominio de captura que comprende una secuencia de ácido nucleico poli(T). En algunas modalidades, una secuencia de captura de analito comprende una secuencia de ácido nucleico poli(T) que se hibrida con un dominio de captura que comprende una secuencia de ácido nucleico poli(A). En algunas modalidades, una secuencia de captura de analito comprende una secuencia de ácido nucleico que no es homopolimérica que se hibrida con un dominio de captura que comprende una secuencia de ácido nucleico que no es homopolimérica que es complementaria (o sustancialmente complementaria) a la secuencia de ácido nucleico que no es homopolimérica de la región de captura del analito.

Figura 2 es un diagrama esquemático de un agente de captura de analito ilustrativo 202 que comprende una fracción de unión al analito 204 y un dominio de código de barras del agente de captura 208. Un resto de unión al analito 204 es una molécula capaz de unirse a un analito 206 e interactuar con una sonda de captura con código de barras espacial. El resto de unión al analito puede unirse al analito 206 con alta afinidad y/o con alta especificidad. El agente de captura de analito puede incluir un dominio de código de barras del agente de captura 208, una secuencia de nucleótidos (por ejemplo, un oligonucleótido), que puede hibridarse con al menos una porción o la totalidad de un dominio de captura de una sonda de captura. El resto de unión al analito 204 puede incluir un polipéptido y/o un aptámero (por ejemplo, una molécula de oligonucleótido o péptido que se une a un analito objetivo específico). El resto de unión al analito 204 puede incluir un anticuerpo o fragmento de anticuerpo (por ejemplo, un fragmento de unión al antígeno).

La Figura 3 es un diagrama esquemático que representa una interacción ilustrativa entre una sonda de captura inmovilizada en la característica 324 y un agente de captura de analito 326. La sonda de captura inmovilizada en

características 324 puede incluir un código de barras espacial 308 así como también una o más secuencias funcionales 306 y 310, como se describe en otra parte en la presente descripción. La sonda de captura también puede incluir un dominio de captura 312 que es capaz de unirse a un agente de captura de analito 326. El agente de captura de analito 326 puede incluir una secuencia funcional 318, dominio de código de barras del agente de captura 316, y una secuencia de captura de analito 314 que es capaz de unirse al dominio de captura 312 de la sonda de captura 324. El agente de captura de analito también puede incluir un enlazador 320 que permite el dominio de código de barras del agente de captura 316 para acoplarse a la porción de unión al analito 322.

Existen al menos dos métodos para asociar un código de barras espacial con una o más células vecinas, de manera que el código de barras espacial identifique la una o más células, y/o el contenido de la una o más células, como asociado con una localización espacial particular. Un método es promover analitos o sustitutos de analitos (por ejemplo, agentes intermedios) fuera de una célula y hacia una matriz con código de barras espacial (por ejemplo, incluidas sondas de captura con código de barras espaciales). Otro método consiste en escindir sondas de captura con códigos de barras espaciales de una matriz y promover las sondas de captura con códigos de barras espaciales hacia y/o dentro o sobre la muestra biológica.

En algunos casos, las sondas de captura pueden configurarse para cebar, replicar y, en consecuencia, producir productos de extensión opcionalmente con códigos de barra a partir de una plantilla (por ejemplo, una plantilla de ADN o ARN, tal como un analito o un agente intermedio (por ejemplo, un producto de ligazón o un agente de captura de analito), o una porción del mismo), o derivados de los mismos (ver, por ejemplo, la Sección (II)(b)(vii) del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663 con respecto a sondas de captura extendidas). En algunos casos, las sondas de captura pueden configurarse para formar productos de ligazón con una plantilla (por ejemplo, una plantilla de ADN o ARN, tal como un analito o un agente intermedio, o una porción del mismo), para crear de esta manera productos de ligazón que sirven como sustitutos de una plantilla.

Como se usa en la presente, una "sonda de captura extendida" se refiere a una sonda de captura que tiene nucleótidos adicionales añadidos al extremo (por ejemplo, extremo 3' o 5') de la sonda de captura, para extender de esta manera la longitud total de la sonda de captura. Por ejemplo, un "extremo 3' extendido" indica que se añadieron nucleótidos adicionales al nucleótido más cerca de 3' de la sonda de captura para extender la longitud de la sonda de captura, por ejemplo, mediante reacciones de polimerización usadas para extender moléculas de ácido nucleico, incluida la polimerización con plantilla catalizada por una polimerasa (por ejemplo, una ADN polimerasa o una transcriptasa inversa). En algunas modalidades, extender la sonda de captura incluye añadir a un extremo 3' de una sonda de captura una secuencia de ácido nucleico que es complementaria a una secuencia de ácido nucleico de un analito o agente intermedio unido específicamente al dominio de captura de la sonda de captura. En algunas modalidades, la sonda de captura se extiende mediante el uso de transcripción inversa. En algunas modalidades, la sonda de captura se extiende mediante el uso de una o más ADN polimerasas. Las sondas de captura extendidas incluyen la secuencia de la sonda de captura y la secuencia del código de barras espacial de la sonda de captura.

En algunas modalidades, las sondas de captura extendidas se amplifican (por ejemplo, en solución a granel o en la matriz) para producir cantidades que son suficientes para el análisis aguas abajo, por ejemplo, mediante secuenciación de ADN. En algunas modalidades, las sondas de captura extendidas (por ejemplo, moléculas de ADN) actúan como plantillas para una reacción de amplificación (por ejemplo, una reacción en cadena de la polimerasa).

Variantes adicionales de métodos de análisis espacial, que incluyen, en algunas modalidades, una etapa de obtención de imágenes, se describen en la Sección (II)(a) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. Análisis de analitos capturados (y/o agentes intermedios o porciones de los mismos), por ejemplo, incluyendo extracción de muestras, extensión de sondas de captura, secuenciación (por ejemplo, de una sonda de captura extendida escindida y/o una molécula de ADN complementaria a una sonda de captura extendida), la secuenciación en la matriz (por ejemplo, mediante el uso de, por ejemplo, enfoques de hibridación in situ o ligazón in situ), análisis temporal y/o captura de proximidad, se describe en la Sección (II)(g) del documento WO 2020/176788 y/o publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. Algunas mediciones de control de calidad se describen en la Sección (II)(h) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

La información espacial puede proporcionar información de importancia biológica y/o médica. Por ejemplo, los métodos y composiciones descritos en la presente descripción pueden permitir: la identificación de uno o más biomarcadores (por ejemplo, de diagnóstico, pronóstico y/o para la determinación de la eficacia de un tratamiento) de una enfermedad o trastorno; identificación de un candidato a objetivo farmacológica para el tratamiento de una enfermedad o trastorno; identificación (por ejemplo, diagnóstico) de un sujeto que tiene una enfermedad o trastorno; identificación de la etapa y/o pronóstico de una enfermedad o trastorno en un sujeto; identificación de un sujeto que tiene una mayor posibilidad de desarrollar una enfermedad o trastorno; seguimiento de la progresión de una enfermedad o trastorno en un sujeto; determinación de la eficacia de un tratamiento de una enfermedad o trastorno en un sujeto; identificación de una subpoblación de pacientes para la cual un tratamiento es efectivo para una enfermedad o trastorno; modificación de un tratamiento de un sujeto con una enfermedad o trastorno; selección de un sujeto para la participación en un ensayo clínico; y/o selección de un tratamiento para un sujeto con una enfermedad o trastorno. Los métodos ilustrativos para identificar información espacial de importancia biológica y/o médica se pueden encontrar en la publicación de solicitud

de patente de Estados Unidos núm. 2021/0140982A1, la solicitud de patente de Estados Unidos núm. 2021/0198741A1 y/o la solicitud de patente de Estados Unidos núm. 2021/0199660.

La información espacial puede proporcionar información de importancia biológica. Por ejemplo, los métodos y composiciones descritos en la presente descripción pueden permitir: identificación de perfiles de expresión del transcriptoma y/o proteoma (por ejemplo, en tejido sano y/o enfermo); identificación de múltiples tipos de analitos muy cercanos (por ejemplo, análisis del vecino más cercano); determinación de genes y/o proteínas regulados positivamente y/o negativamente en tejido enfermo; caracterización de microambientes tumorales; caracterización de respuestas inmunitarias tumorales; caracterización de tipos celulares y su colocalización en tejido; e identificación de variantes genéticas dentro de los tejidos (por ejemplo, basadas en los perfiles de expresión de genes y/o proteínas asociados con biomarcadores de enfermedades o trastornos específicos).

Típicamente, para los métodos basados en matrices espaciales, un sustrato funciona como un soporte para la unión directa o indirecta de sondas de captura a elementos de la matriz. Un "elemento" es una entidad que actúa como soporte o repositorio para varias entidades moleculares usadas en el análisis espacial. En algunas modalidades, algunos o todos los elementos de una matriz están funcionalizadas para la captura del analito. Se describen sustratos ilustrativos en la Sección (II)(c) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. Pueden encontrarse características y atributos geométricos ilustrativos de una matriz en las Secciones (II)(d)(i), (II)(d)(iii) y (II)(d)(iv) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

Generalmente, los analitos y/o agentes intermedios (o porciones de los mismos) pueden capturarse cuando se pone en contacto una muestra biológica con un sustrato que incluye sondas de captura (por ejemplo, un sustrato con sondas de captura incrustadas, manchadas, impresas, fabricadas sobre el sustrato, o un sustrato con elementos (por ejemplo, perlas, pocillos) que comprende sondas de captura). Como se usa en la presente descripción, "contactar", "contactado" y/o "que entra en contacto", una muestra biológica con un sustrato se refiere a cualquier contacto (por ejemplo, directo o indirecto) de manera que las sondas de captura pueden interactuar (por ejemplo, unirse covalente o no covalentemente (por ejemplo, hibridar)) con analitos de la muestra biológica. La captura puede lograrse activamente (por ejemplo, mediante el uso de electroforesis) o pasivamente (por ejemplo, mediante el uso de difusión). La captura de analitos se describe con más detalle en la Sección (II)(e) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

En algunos casos, el análisis espacial puede realizarse mediante la unión y/o introducción de una molécula (por ejemplo, un péptido, un lípido o una molécula de ácido nucleico) que tiene un código de barras (por ejemplo, un código de barras espacial) a una muestra biológica (por ejemplo, a una célula en una muestra biológica). En algunas modalidades, una pluralidad de moléculas (por ejemplo, una pluralidad de moléculas de ácido nucleico) que tienen una pluralidad de códigos de barras (por ejemplo, una pluralidad de códigos de barras espaciales) se introducen en una muestra biológica (por ejemplo, en una pluralidad de células en una muestra biológica) para usar en análisis espacial. En algunas modalidades, después de unir y/o introducir una molécula que tiene un código de barras a una muestra biológica, la muestra biológica puede separarse físicamente (por ejemplo, disociarse) en células individuales o grupos de células para su análisis. Algunos de estos métodos de análisis espacial se describen en la Sección (III) del documento WO 2020/176788 y/o en la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663.

En algunos casos, el análisis espacial puede realizarse mediante la detección de múltiples oligonucleótidos que se hibridan con un analito. En algunos casos, por ejemplo, el análisis espacial puede realizarse mediante el uso de ligazón con plantilla de ARN (RTL). Los métodos de RTL se han descrito anteriormente. Ver, por ejemplo, Credle y otros, *Nucleic Acids Res.* 21 de agosto de 2017; 45(14):e128. Típicamente, RTL incluye la hibridación de dos oligonucleótidos con secuencias adyacentes en un analito (por ejemplo, una molécula de ARN, tal como una molécula de ARNm). En algunos casos, los oligonucleótidos son moléculas de ADN. En algunos casos, uno de los oligonucleótidos incluye al menos dos bases de ácido ribonucleico en el extremo 3' y/o el otro oligonucleótido incluye un nucleótido fosforilado en el extremo 5'. En algunos casos, uno de los dos oligonucleótidos incluye un dominio de captura (por ejemplo, una secuencia poli(A), una secuencia no homopolimérica). Después de la hibridación con el analito, una ligasa (por ejemplo, la ligasa SplintR) liga los dos oligonucleótidos juntos, creando un producto de ligazón. En algunos casos, los dos oligonucleótidos se hibridan con secuencias que no son adyacentes entre sí. Por ejemplo, la hibridación de los dos oligonucleótidos crea un espacio entre los oligonucleótidos hibridados. En algunos casos, una polimerasa (por ejemplo, una ADN polimerasa) puede extender uno de los oligonucleótidos antes de la ligazón. Después de la ligazón, el producto de ligazón se libera del analito. En algunos casos, el producto de ligazón se libera mediante el uso de una endonucleasa (por ejemplo, ARNasa H). El producto de ligazón liberado puede capturarse a continuación mediante sondas de captura (por ejemplo, en lugar de la captura directa de un analito) en una matriz, opcionalmente amplificarse y secuenciarse, para determinar de este modo la localización y opcionalmente la abundancia del analito en la muestra biológica.

Durante el análisis de información espacial, se obtiene información de secuencia para un código de barras espacial asociado con un analito, y la información de secuencia puede usarse para proporcionar información sobre la distribución espacial del analito en la muestra biológica. Pueden usarse varios métodos para obtener la información espacial. En algunas modalidades, sondas de captura específicas y los analitos que capturan están asociados con

ubicaciones específicas en una matriz de elementos en un sustrato. Por ejemplo, pueden asociarse códigos de barras espaciales específicos con ubicaciones de matriz específicas antes de la fabricación de la matriz, y las secuencias de los códigos de barras espaciales pueden almacenarse (por ejemplo, en una base de datos) junto con información de localización de matriz específica, de modo que cada código de barras espacial se asigne de forma única a una localización particular de la matriz.

Alternativamente, pueden depositarse códigos de barras espaciales específicos en ubicaciones predeterminadas en una matriz de elementos durante la fabricación, de manera que en cada localización, solo esté presente un tipo de código de barras espacial, de modo que los códigos de barras espaciales estén asociados de forma única con un único elemento de la matriz. Cuando sea necesario, las matrices pueden decodificarse mediante el uso de cualquiera de los métodos descritos en la presente descripción de modo que los códigos de barras espaciales se asocien de forma única con las ubicaciones de los elementos de la matriz, y este mapeo puede almacenarse como se describió anteriormente.

Cuando se obtiene información de secuencia para sondas y/o analitos de captura durante el análisis de la información espacial, las ubicaciones de las sondas de captura y/o analitos pueden determinarse al hacer referencia a la información almacenada que asocia de forma única cada código de barras espacial con una localización de elementos de la matriz. De esta manera, las sondas de captura específicas y los analitos capturados se asocian con ubicaciones específicas en la matriz de elementos. Cada localización de elemento de matriz representa una posición con relación a un punto de referencia de coordenadas (por ejemplo, una localización de matriz, un marcador fiducial) para la matriz. En consecuencia, cada localización de elemento tiene una "dirección" o localización en el espacio de coordenadas de la matriz.

Algunos flujos de trabajo ilustrativos de análisis espacial se describen en la sección modalidades ilustrativas del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. Ver, por ejemplo, la modalidad ilustrativa que comienza con "En algunos ejemplos no limitantes de los flujos de trabajo descritos en la presente descripción, la muestra puede sumergirse..." del documento WO 2020/176788 y/o la publicación de solicitud de patente de Estados Unidos núm. 2020/0277663. Ver también, por ejemplo, la Guía del usuario de los kits de reactivos de expresión genética espacial de Visium (por ejemplo, Rev C, con fecha de junio de 2020) y/o la Guía del usuario de los kits de reactivos de optimización de tejidos espaciales de Visium (por ejemplo, Rev C, con fecha de julio de 2020).

En algunas modalidades, el análisis espacial puede realizarse mediante el uso de hardware y/o software dedicado, tal como cualquiera de los sistemas descritos en las Secciones (II)(e)(ii) y/o (V) de los documentos WO 2020/176788 y/o publicación de solicitud de patente de Estados Unidos núm. 2020/0277663, o cualquiera de uno o más de los dispositivos o métodos descritos en las secciones *Control Slide for Imaging, Methods of Using Control Slides and Substrates for Systems of Using Control Slides and Substrates for Imaging, y/o Sample and Array Alignment Devices and Methods, Informational labels* del documento WO 2020/123320.

Los sistemas adecuados para realizar análisis espacial pueden incluir componentes tales como una cámara (por ejemplo, una celda de flujo o una cámara que pueda sellarse, hermética) para contener una muestra biológica. La muestra biológica puede montarse, por ejemplo, en un soporte para muestras biológicas. Pueden conectarse una o más cámaras de fluido a la cámara y/o al soporte para muestras a través de conductos de fluido, y los fluidos pueden suministrarse a la cámara y/o al soporte para muestras a través de bombas fluidicas, fuentes de vacío u otros dispositivos acoplados a los conductos de fluido que crean un gradiente de presión para impulsar el flujo de fluido. También puede conectarse una o más válvulas a conductos de fluido para regular el flujo de reactivos desde los depósitos hasta la cámara y/o el soporte para muestras.

Los sistemas pueden incluir opcionalmente una unidad de control que incluye uno o más procesadores electrónicos, una interfaz de entrada, una interfaz de salida (tal como una pantalla) y una unidad de almacenamiento (por ejemplo, un medio de almacenamiento de estado sólido tal como, pero sin limitarse a, un medio de almacenamiento magnético, óptico u otro de estado sólido, persistente, grabable y/o regrabable). Opcionalmente, la unidad de control puede conectarse a uno o más dispositivos remotos a través de una red. La unidad de control (y componentes de la misma) generalmente puede realizar cualquiera de las etapas y funciones descritas en la presente descripción. Cuando el sistema está conectado a un dispositivo remoto, el dispositivo (o dispositivos) remotos puede realizar cualquiera de las etapas o características descritas en la presente descripción. Los sistemas pueden incluir opcionalmente uno o más detectores (por ejemplo, CCD, CMOS) usados para capturar imágenes. Los sistemas también pueden incluir opcionalmente una o más fuentes de luz (por ejemplo, basadas en LED, basadas en diodos, láseres) para iluminar una muestra, un sustrato con elementos, analitos de una muestra biológica capturada en un sustrato y diversos medios de control y calibración.

Los sistemas pueden incluir opcionalmente instrucciones de software codificadas y/o implementadas en uno o más medios de almacenamiento tangibles y componentes de hardware tales como circuitos integrados de aplicaciones específicas. Las instrucciones de software, cuando son ejecutadas por una unidad de control (y en particular, un procesador electrónico) o un circuito integrado, pueden hacer que la unidad de control, circuito integrado u otro

componente que ejecuta las instrucciones de software realice cualquiera de las etapas o funciones del método descritos en la presente descripción.

En algunos casos, los sistemas descritos en la presente descripción pueden detectar (por ejemplo, registrar una imagen) la muestra biológica en la matriz. Los métodos ilustrativos para detectar la muestra biológica en una matriz se describen en el documento WO 2021/102003 y/o la solicitud de patente de Estados Unidos núm. US 2021/0150707 A1.

Antes de transferir analitos desde la muestra biológica hasta la matriz de elementos en el sustrato, la muestra biológica puede alinearse con la matriz. La alineación de una muestra biológica y una matriz de elementos que incluyen sondas de captura pueden facilitar el análisis espacial, que puede usarse para detectar diferencias en la presencia y/o nivel de analito dentro de diferentes posiciones en la muestra biológica, por ejemplo, para generar un mapa tridimensional de la presencia y/o nivel del analito. Los métodos ilustrativos para generar un mapa bidimensional y/o tridimensional de la presencia y/o nivel del analito se describen en la solicitud PCT núm. 2020/053655 y los métodos de análisis espacial se describen generalmente en el documento WO 2021/102039 y/o la solicitud de patente de Estados Unidos núm. US 2021/0155982 A1.

En algunos casos, un mapa de presencia y/o nivel de analito puede alinearse con una imagen de una muestra biológica mediante el uso de uno o más marcadores de referencia, por ejemplo, objetos colocados en el campo de visión de un sistema de procesamiento de imágenes que aparecen en la imagen producida, como se describe en el *Atributos del sustrato* Sección, *Control Slide for Imaging* Sección del documento WO 2020/123320, WO 2021/102005 y/o solicitud de patente de Estados Unidos núm. US 2021/0158522 A1. Los marcadores fiduciales pueden usarse como punto de referencia o escala de medición para la alineación (por ejemplo, para alinear una muestra y una matriz, para alinear dos sustratos, para determinar la localización de una muestra o matriz en un sustrato con relación a un marcador fiducial) y/o para mediciones cuantitativas de tamaños y/o distancias.

Análisis en formato múltiple de la expresión génica y proteica

La comprensión tanto de la expresión génica como proteica en sistemas biológicos puede ser útil para obtener información sobre los tejidos normales, en desarrollo y enfermos. Mientras que la secuenciación de ARN unicelular (scRNA-seq) hace posible obtener mediciones de expresión génica de alta resolución, la técnica requiere que las células se disocien, lo que hace que se pierda información anatómica y organizativa. De manera similar, se conocen numerosas técnicas de detección de proteínas y pueden proporcionar información espacial de proteínas en una muestra biológica, sin embargo, aún se necesitan métodos para detectar simultáneamente los niveles de expresión génica (por ejemplo, ARNm) de la proteína, o incluso todo el transcriptoma.

Por tanto, en la presente descripción se describen enfoques de "multiómicas" que pueden proporcionar un complemento poderoso a las metodologías tradicionales, lo que permite una mayor comprensión de la heterogeneidad y organización celular dentro de las muestras biológicas. La combinación de detección de proteínas mediante el uso de agentes de captura de analitos en una matriz espacial permite el examen simultáneo de proteínas y expresión génica a partir de la misma muestra biológica (por ejemplo, sección de tejido). Por ejemplo, una matriz que comprende las sondas de captura (por ejemplo, cualquiera de las sondas de captura descritas en la presente descripción) se puede poner en contacto con una muestra biológica, una pluralidad de sondas de ligazón con plantilla y una pluralidad de agentes de captura de analitos que dan como resultado un análisis simultáneo de expresión génica y proteica.

En algunas modalidades, la pluralidad de sondas de ligazón con plantilla incluye un par de sondas para un ácido nucleico objetivo (por ejemplo, ADN, ARN). Las sondas son complementarias a porciones del ácido nucleico objetivo, sin embargo, cuando ambas sondas se hibridan con el ácido nucleico objetivo, existe un espacio entre las dos sondas. En algunas modalidades, el espacio se liga, generando de esta manera un producto de ligazón con plantilla (por ejemplo, un producto de ligazón con plantilla de ADN o ARN). En algunas modalidades, una de las parejas de sondas incluye una secuencia de flanqueo complementaria a un dominio de captura de la matriz. En algunas modalidades, la secuencia complementaria al dominio de captura del producto de ligazón con plantilla se hibrida con el dominio de captura de la sonda de captura.

En algunas modalidades, los agentes de captura de analito, como se describe en la presente descripción, también pueden ponerse en contacto con la muestra biológica. En algunas modalidades, los agentes de captura de analito se ponen en contacto con la muestra biológica antes de que la muestra biológica se ponga en contacto con una matriz. En algunas modalidades, los agentes de captura de analito se ponen en contacto con la muestra biológica después de que la muestra biológica se pone en contacto con la matriz. En algunas modalidades, el resto de unión al analito del agente de captura de analito interactúa (por ejemplo, se une) con un analito (por ejemplo, proteína) en una muestra biológica. En algunas modalidades, el resto de unión al analito es un anticuerpo o fragmento de unión al antígeno.

Los agentes de captura de analitos también pueden incluir un oligonucleótido conjugado que puede comprender uno o más dominios. Por ejemplo, el oligonucleótido conjugado puede incluir un código de barras de la fracción de unión al analito y una secuencia de captura de analito. En algunas modalidades, el código de barras del resto de unión al analito, o un complemento de esta, se refiere a (por ejemplo, identifica) un código de barras que se asocia con o de

cualquier otra manera identifica la porción de unión al analito. En algunas modalidades, el oligonucleótido conjugado puede incluir una secuencia de captura de analito. En algunas modalidades, la secuencia de captura de analito es capaz de interactuar con (por ejemplo, hibridarse) con un dominio de captura de una sonda de captura en un sustrato.

En algunas modalidades, se permite que las sondas de ligazón con plantilla se unan al ácido nucleico objetivo antes de que los agentes de captura de analito se suministren a la muestra biológica. En algunas modalidades, las sondas de ligazón con plantilla se pueden ligar juntas antes, simultáneamente o después de que se suministren los agentes de captura de analito a la muestra biológica. En algunas modalidades, los agentes de captura de analito se suministran a la muestra biológica y se permite que el resto de unión al analito se una al analito objetivo (por ejemplo, proteína) antes de que se suministren las sondas de ligazón con plantilla. En algunas modalidades, los agentes de captura de analitos se suministran a la muestra biológica y la secuencia de captura de analitos se bloquea (por ejemplo, bloqueada por cualquiera de los métodos descritos en la presente descripción). En algunas modalidades, la secuencia de captura de analito de los agentes de captura de analito no se bloquea (por ejemplo, no se bloquea por ninguno de los métodos descritos en la presente descripción) antes, simultáneamente, o después de que las sondas de ligazón con plantilla (por ejemplo, las sondas de ligazón con plantilla de ARN) se suministren y/o antes, simultáneamente, o después de que las sondas de ligazón con plantilla se ligan entre sí.

Por lo tanto, en la presente descripción se proporcionan métodos para determinar la localización espacial de un ácido nucleico y una proteína a partir de una muestra biológica que incluye: a) proporcionar una matriz espacial que incluye una primera y segunda pluralidad de sondas de captura donde cada pluralidad incluye un código de barras espacial y un dominio de captura, b) poner en contacto la matriz espacial con una muestra biológica, c) poner en contacto la muestra biológica con (i) una pluralidad de agentes de captura de analito, donde un agente de captura de analito incluye un resto de unión al analito y un oligonucleótido que incluye un código de barras de un resto de unión al analito y una secuencia de captura de analito, donde la secuencia de captura de analito incluye una secuencia complementaria a una segunda pluralidad de dominios de captura, y (ii) una pluralidad de sondas de ligazón con plantilla, donde una de las sondas de ligazón con plantilla incluye una secuencia complementaria a una primera pluralidad de dominios de captura, d) unir el resto de unión al analito del agente de captura de analito a una proteína objetivo, e) hibridar las sondas de ligazón con plantilla a un ácido nucleico objetivo y ligar las sondas para producir productos de ligazón, f) hibridar los productos de ligazón a la primera pluralidad de dominios de captura y las secuencias de captura de analito de los agentes de captura de analito unidos a la segunda pluralidad de dominios de captura en la matriz espacial, y g) determinar la secuencia o una porción de esta de un producto de ligazón capturado, o un complemento de esta, y la secuencia del código de barras espacial de la sonda de captura, o un complemento de esta, que se asocia con el producto de ligazón, y la secuencia del resto de unión al analito de la sonda de captura, o un complemento de esta, del agente de captura de analito unido, para determinar de esta manera la localización espacial de un ácido nucleico y la proteína de la muestra biológica.

También se proporcionan en la presente descripción métodos para determinar la localización espacial de un ácido nucleico y una proteína en una muestra biológica de enfermedad que incluye: a) proporcionar una matriz espacial que incluye una primera y una segunda pluralidad de sondas de captura donde cada pluralidad incluye un código de barras espacial y un dominio de captura, b) poner en contacto la muestra biológica de enfermedad con la matriz espacial, c) poner en contacto la muestra biológica de enfermedad con: (i) una pluralidad de agentes de captura de analito, donde un agente de captura de analito incluye un resto de unión al analito y un oligonucleótido que incluye un código de barras de un resto de unión al analito y una secuencia de captura de analito, donde la secuencia de captura de analito incluye una secuencia complementaria a una segunda pluralidad de dominios de captura, y (ii) una pluralidad de sondas de ligazón con plantilla, donde una de las sondas de ligazón con plantilla incluye una secuencia complementaria a una primera pluralidad de dominios de captura, d) unir el resto de unión al analito del agente de captura de analito a una proteína objetivo, e) hibridar las sondas de ligazón con plantilla a un ARN objetivo y ligar las sondas para producir productos de ligazón con plantilla, f) hibridar los productos de ligazón con plantilla a la primera pluralidad de dominios de captura y las secuencias de captura de analito de los agentes de captura de analito unidos a la segunda pluralidad de dominios de captura en la matriz espacial, y g) determinar la secuencia o una porción de esta de un producto de ligazón capturado, o un complemento, y la secuencia del código de barras espacial de la sonda de captura, o un complemento de esta, que se asocia con el producto de ligazón, y la secuencia del resto de unión al analito de la sonda de captura, o un complemento de esta, del agente de captura de analito unido, para determinar de esta manera la localización espacial de un ácido nucleico y la proteína en la muestra biológica de enfermedad.

En algunas modalidades, el ácido nucleico es ARN. En algunas modalidades, el ARN es ARNm. En algunas modalidades, el ácido nucleico es ADN.

En algunas modalidades, la muestra biológica de enfermedad es una muestra biológica cancerosa. En algunas modalidades, la muestra biológica cancerosa es una muestra biológica de cáncer de ovario o una muestra biológica de cáncer de mama. En algunas modalidades, la muestra de cáncer de mama es cáncer de mama triple positivo. En algunas modalidades, el cáncer de mama es cáncer de mama de células ductales invasivas. En algunas modalidades, el carcinoma ductal invasivo es un carcinoma ductal invasivo de grado II. En algunas modalidades, el carcinoma ductal invasivo es un carcinoma ductal invasivo de grado III. En algunas modalidades, el cáncer de mama es un carcinoma lobular invasivo. En algunas modalidades, la muestra biológica cancerosa es cáncer de pulmón. En algunas modalidades, la muestra biológica cancerosa es melanoma. En algunas modalidades, la muestra biológica cancerosa

es cáncer de colon. En algunas modalidades, la muestra biológica cancerosa es glioblastoma. En algunas modalidades, la muestra biológica cancerosa es cáncer de próstata.

En algunas modalidades, las sondas de captura incluyen identificadores moleculares únicos, secuencias funcionales o sus combinaciones.

En algunas modalidades, la primera pluralidad de dominios de captura son secuencias homopoliméricas. En algunas modalidades, la primera pluralidad de dominios de captura comprende secuencias de poli(T). En algunas modalidades, la primera pluralidad de dominios de captura se definen como secuencias no homopoliméricas. En algunas modalidades, la primera pluralidad de dominios de captura incluye una secuencia redundante. En algunas modalidades, la primera pluralidad de dominios de captura incluye una secuencia fija. Por ejemplo, la primera pluralidad de dominios de captura puede comprender una de las SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, o SEQ ID NO: 11.

En algunas modalidades, la segunda pluralidad de dominios de captura son secuencias homopoliméricas. En algunas modalidades, la segunda pluralidad de dominios de captura se definen como secuencias no homopoliméricas. En algunas modalidades, la segunda pluralidad de dominios de captura comprende secuencias de poli(T). En algunas modalidades, la segunda pluralidad de dominios de captura incluye una secuencia redundante. En algunas modalidades, la segunda pluralidad de dominios de captura incluye una secuencia fija. Por ejemplo, la segunda pluralidad de dominios de captura puede comprender una de las SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, o SEQ ID NO: 11.

En algunas modalidades, la primera pluralidad de dominios de captura y la segunda pluralidad de dominios de captura incluyen la misma secuencia. En algunas modalidades, la primera pluralidad de dominios de captura y la segunda pluralidad de dominios de captura incluyen diferentes secuencias.

Generalmente, los métodos de la presente descripción pueden usarse con cualquier muestra biológica (por ejemplo, cualquier muestra biológica descrita en la presente descripción). En algunas modalidades, la muestra biológica es una sección de tejido. En algunas modalidades, la muestra biológica es una muestra de tejido. En algunas modalidades, la muestra biológica es una muestra biológica recién congelada. En algunas modalidades, la muestra biológica es una muestra biológica fijada (por ejemplo, embebida en parafina fijada con formalina (FFPE), paraformaldehído, acetona o metanol). En algunas modalidades, la muestra biológica es una muestra FFPE. En algunas modalidades, la muestra biológica es una sección de tejido FFPE. En algunas modalidades, la muestra de tejido es una muestra de tumor. En alguna modalidad, la sección de tejido es una sección de tejido tumoral. En algunas modalidades, la sección de tejido tumoral es una sección de tejido tumoral fija (por ejemplo, una sección de tejido tumoral fijada en parafina fijada con formalina). En algunas modalidades, la muestra de tumor comprende uno o más tumores cancerosos. Se conocen numerosos tipos de cáncer en la técnica. En algunas modalidades, la muestra de tejido se deriva de una muestra de biopsia. En algunas modalidades, la muestra de tejido se deriva de un embrión de roedor completo. En algunas modalidades, el tejido se selecciona de, pero no se limita a, tejido cerebral, tejido de mama, tejido de colon, tejido de corazón, tejido pulmonar, tejido de bazo, tejido de testículos, tejido de amígdala inflamada, tejido de cuello uterino y tejido de ganglios linfáticos.

En algunas modalidades, una muestra FFPE se desparafina y desreticular antes de suministrar una pluralidad de sondas de ligazón con plantilla (por ejemplo, sondas de ligazón con plantilla de ARN) y agentes de captura de analitos. En algunas modalidades, una muestra biológica FFPE se desparafina y descruza antes de la etapa (b). Por ejemplo, el material de incrustación en parafina se puede eliminar (por ejemplo, desparafinado) de la muestra biológica (por ejemplo, sección de tejido) al incubar la muestra biológica en un solvente apropiado (por ejemplo, xileno), seguido de una serie de enjuagues (por ejemplo, etanol de diversas concentraciones) y rehidratación en agua. En algunas modalidades, la muestra biológica se puede secar después de la desparafinización. En algunas modalidades, después de la etapa de secar la muestra biológica, la muestra biológica se puede teñir (por ejemplo, tinción H&E, cualquiera de la variedad de tinciones descritas en la presente descripción).

En algunas modalidades, el método incluye teñir la muestra biológica. En algunas modalidades, la tinción incluye el uso de hematoxilina y eosina. En algunas modalidades, una muestra biológica se puede teñir mediante el uso de cualquier número de colorantes biológicos, que incluyen, entre otros, naranja de acridina, marrón de Bismarck, carmín, azul coomassie, violeta de cresilo, DAPI, eosina, etidio bromuro, ácido fucsina, hematoxilina, tinciones de Hoechst, yodo, verde de metilo, azul de metileno, rojo neutro, azul de Nilo, rojo de Nilo, tetróxido de osmio, yoduro de propidio, rodamina o safranina.

La muestra biológica se puede teñir mediante el uso de técnicas de tinción conocidas, que incluyen Can-Grunwald, Giemsa, hematoxilina y eosina (H&E), Jenner, Leishman, tricrómico de Masson, Papanicolaou, Romanowsky, plata, Sudán, Wright y/o técnicas de tinción de ácido periódico de Schiff (PAS). La tinción con PAS se realiza típicamente después de la fijación en formalina o acetona.

En algunas modalidades, la tinción incluye el uso de una etiqueta detectable seleccionada del grupo que consiste en un radioisótopo, un fluoróforo, un compuesto quimioluminiscente, un compuesto bioluminiscente o una de sus combinaciones.

En algunas modalidades, se obtienen imágenes de la muestra biológica después de teñir la muestra biológica. En algunas modalidades, se obtienen imágenes de la muestra biológica antes de teñir la muestra biológica. En algunas modalidades, la muestra biológica se visualiza o se obtienen imágenes de esta mediante el uso de microscopía de campo claro. En algunas modalidades, la muestra biológica se visualiza o se obtienen imágenes de esta mediante el uso de microscopía de fluorescencia. Se conocen en la técnica métodos adicionales de visualización y obtención de imágenes. Los ejemplos no limitativos de visualización y obtención de imágenes incluyen microscopía de expansión, microscopía de campo claro, microscopía de campo oscuro, microscopía de contraste de fases, microscopía electrónica, microscopía de fluorescencia, microscopía de reflexión, microscopía de interferencia y microscopía confocal. En algunas modalidades, la muestra se tiñe y se obtienen imágenes de esta antes de añadir el primer y/o segundo cebador a la muestra biológica en la matriz.

Después de que una muestra biológica fijada (por ejemplo, FFPE, PFA, acetona, metanol) se haya desparafinado, la muestra biológica fijada (por ejemplo, FFPE, PFA) se puede procesar adicionalmente. Por ejemplo, las muestras biológicas fijadas (por ejemplo, FFPE, PFA) se pueden tratar para eliminar las reticulaciones (por ejemplo, las reticulaciones inducidas por formaldehído (por ejemplo, la desreticulación)). En algunas modalidades, la desreticulación de las reticulaciones (por ejemplo, reticulaciones inducidas por formaldehído) en la muestra biológica fijada (por ejemplo, FFPE, PFA) puede incluir tratar la muestra con calor. En algunas modalidades, la desreticulación de las reticulaciones inducidas por formaldehído puede incluir realizar una reacción química. En algunas modalidades, la desreticulación de las reticulaciones inducidas por formaldehído, puede incluir tratar la muestra con un reactivo de permeabilización. En algunas modalidades, la desreticulación de las reticulaciones inducidas por formaldehído puede incluir calor, una reacción química y/o reactivos de permeabilización.

En algunas modalidades, la desreticulación de las reticulaciones (por ejemplo, reticulaciones inducidas por formaldehído) se pueden realizar en presencia de un tampón. En algunas modalidades, el tampón es tampón Tris-EDTA (TE) (por ejemplo, tampón TE para muestras biológicas FFPE). En algunas modalidades, el tampón es tampón de citrato (por ejemplo, tampón de citrato para muestras biológicas FFPE). En algunas modalidades, el tampón es tampón Tris-HCl (por ejemplo, tampón Tris-HCl para muestras biológicas fijadas con PFA). En algunas modalidades, el tampón (por ejemplo, tampón TE, tampón Tris-HCl) tiene un pH de aproximadamente 5,0 a aproximadamente 10,0 y una temperatura de entre aproximadamente 60 °C a aproximadamente 100 °C.

En algunas modalidades, la muestra biológica se permeabiliza (por ejemplo, se permeabiliza por cualquiera de los métodos descritos en la presente descripción). En algunas modalidades, la permeabilización es una permeabilización enzimática. En algunas modalidades, la permeabilización es una permeabilización química. En algunas modalidades, la muestra biológica se permeabiliza antes de suministrar las sondas de ligazón con plantilla de ARN y los agentes de captura de analitos a la muestra biológica. En algunas modalidades, la muestra biológica se permeabiliza al mismo tiempo que las sondas de ligazón con plantilla de ARN y los agentes de captura de analitos se suministran a la muestra biológica. En algunas modalidades, la muestra biológica se permeabiliza después de que las sondas de ligazón con plantilla de ARN y los agentes de captura de analitos se suministran a la muestra biológica. En algunas modalidades, la hibridación de los productos de ligazón con plantilla por ARN a los segundos dominios de captura y las secuencias de captura de analito de los agentes de captura de analito unidos a los primeros dominios de captura comprende además la permeabilización de la muestra biológica.

En algunas modalidades, la muestra biológica se permeabiliza de aproximadamente 30 a aproximadamente 120 minutos, de aproximadamente 40 a aproximadamente 110 minutos, de aproximadamente 50 a aproximadamente 100 minutos, de aproximadamente 60 a aproximadamente 90 minutos, o de aproximadamente 70 a 80 minutos. En algunas modalidades, las muestras biológicas se permeabilizan durante aproximadamente 30, aproximadamente 35, aproximadamente 40, aproximadamente 45, aproximadamente 50, aproximadamente 55, aproximadamente 60, aproximadamente 65, aproximadamente 70, aproximadamente 75, aproximadamente 80, aproximadamente 90, aproximadamente 95, aproximadamente 100, aproximadamente 105, aproximadamente 110, aproximadamente 115, o aproximadamente 130 minutos.

En algunas modalidades, el tampón de permeabilización comprende urea. En algunas modalidades, la urea está en una concentración de aproximadamente 0,5 M a 3,0 M. En algunas modalidades, la concentración de la urea es aproximadamente 0,5, 1,0, 1,5, 2,0, 2,5, o aproximadamente 3,0 M. En algunas modalidades, el tampón de permeabilización incluye un detergente. En algunas modalidades, el detergente es sarcosilo. En algunas modalidades, el sarcosilo está presente en aproximadamente 2 % a aproximadamente 10 % (v/v). En algunas modalidades, el sarcosilo está presente en aproximadamente 3 %, 4 %, 5 %, 6 %, 7 %, 8 % o 9 % (v/v). En algunas modalidades, el tampón de permeabilización comprende polietilenglicol (PEG). En algunas modalidades, el PEG es de aproximadamente PEG 2K a aproximadamente PEG 16K. En algunas modalidades, el PEG es PEG 2K, 3K, 4K, 5K, 6K, 7K, 8K, 9K, 10K, 11K, 12K, 13K, 14K, 15K o 16K. En algunas modalidades, el PEG está presente en una concentración de aproximadamente 2 % a 25 %, de aproximadamente 4 % a aproximadamente 23 %, de aproximadamente 6 % a aproximadamente 21 %, o de aproximadamente 8 % a aproximadamente 20 % (v/v).

En algunas modalidades, el método incluye una etapa de permeabilizar la muestra biológica (por ejemplo, una sección de tejido). Por ejemplo, la muestra biológica puede permeabilizarse para facilitar la transferencia de los productos extendidos a las sondas de captura en la matriz. En algunas modalidades, la permeabilización incluye el uso de un solvente orgánico (por ejemplo, acetona, etanol y metanol), un detergente (por ejemplo, saponina, Triton X-100™, Tween-20™ o dodecil sulfato de sodio (SDS)) y una enzima (una endopeptidasa, una exopeptidasa, una proteasa) o sus combinaciones. En algunas modalidades, la permeabilización incluye el uso de un endopeptidasa, una proteasa, SDS, éter de polietilenglicol terc-octilfenilo, polisorbato 80 y polisorbato 20, solución de sal de N-lauroilsarcosina de sodio, saponina, Tritón X-100™, Tween-20™, o sus combinaciones. En algunas modalidades, la endopeptidasa es pepsina. En algunas modalidades, la endopeptidasa es la proteinasa K. Se describen métodos adicionales para la permeabilización de muestras, por ejemplo, en Jamur y otros, *Method Mol. Biol.* 588:63-66, 2010.

Los métodos proporcionados en la presente descripción también pueden incluir la tinción de anticuerpos. En alguna modalidad, la tinción de anticuerpos incluye el uso de un tampón de tinción de anticuerpos. En algunas modalidades, el tampón de tinción de anticuerpos (por ejemplo, un tampón a base de PBS) incluye un detergente (por ejemplo, Tween-20, SDS, sarcosilo). En algunas modalidades, el tampón de tinción de anticuerpos incluye un suero, tal como por ejemplo, un suero de cabra. En algunas modalidades, el suero de cabra es de aproximadamente 1 % a aproximadamente 10 % (v/v), de aproximadamente 2 % a aproximadamente 9 % (v/v), de aproximadamente 3 % a aproximadamente 8 % (v/v), o de aproximadamente 4 % a aproximadamente 7 % (v/v). En algunas modalidades, el tampón de tinción de anticuerpos incluye sulfato de dextrano. En algunas modalidades, el sulfato de dextrano está a una concentración de aproximadamente 1 mg/ml a aproximadamente 20 mg/ml, de aproximadamente 5 mg/ml a aproximadamente 15 mg/ml, o de aproximadamente 8 mg/ml a aproximadamente 12 mg/ml.

Los métodos proporcionados en la presente descripción también pueden utilizar sondas de bloqueo para bloquear la unión no específica (por ejemplo, hibridación) de la secuencia de captura de analito y el dominio de captura de una sonda de captura en una matriz. En algunas modalidades, después del contacto entre la muestra biológica y la matriz, la muestra biológica entra en contacto con una pluralidad de agentes de captura de analitos, donde un agente de captura de analito incluye una secuencia de captura de analito que se bloquea de manera reversible con una sonda de bloqueo. En algunas modalidades, la secuencia de captura de analito se bloquea de manera reversible con más de una sonda de bloqueo (por ejemplo, 2, 3, 4 o más sondas de bloqueo). En algunas modalidades, el agente de captura de analito se bloquea antes de que se una el analito objetivo (por ejemplo, una proteína objetivo).

En algunas modalidades, el oligonucleótido del agente de captura de analito (por ejemplo, la secuencia de captura de analito) se bloquea mediante una sonda de bloqueo. En algunas modalidades, las sondas de bloqueo se hibridan con la secuencia de captura de analito de los agentes de captura de analito antes de introducir los agentes de captura de analito en una muestra biológica. En algunas modalidades, las sondas de bloqueo se hibridan a la secuencia de captura de analito de los agentes de captura de analito después de introducir los agentes de captura de analito en la muestra biológica. En tales modalidades, el dominio de captura también puede bloquearse para evitar la unión no específica y/o para controlar el tiempo de unión, entre la secuencia de captura de analito y el dominio de captura. En algunas modalidades, las sondas de bloqueo se pueden introducir de forma alternativa o adicional durante la tinción de la muestra biológica. En algunas modalidades, la secuencia de captura de analito se bloquea antes de unirse al dominio de captura, donde la sonda de bloqueo incluye una secuencia complementaria o sustancialmente complementaria a la secuencia de captura de analito.

En algunas modalidades, la secuencia de captura de analito se bloquea con una sonda de bloqueo. En algunas modalidades, la secuencia de captura de analito se bloquea con dos sondas de bloqueo. En algunas modalidades, la secuencia de captura de analito se bloquea con más de dos sondas de bloqueo (por ejemplo, 3, 4, 5 o más sondas de bloqueo). En algunas modalidades, se usa una sonda de bloqueo para bloquear el extremo 3' libre de la secuencia de captura de analito. En algunas modalidades, se usa una sonda de bloqueo para bloquear el extremo 5' de la secuencia de captura de analito. En algunas modalidades, se usan dos sondas de bloqueo para bloquear ambos extremos 5' y 3' de la secuencia de captura de analito. En algunas modalidades, tanto la secuencia de captura de analito como el dominio de la sonda de captura se bloquean.

En algunas modalidades, las sondas de bloqueo pueden diferir en longitud y/o complejidad. En algunas modalidades, la sonda de bloqueo puede incluir una secuencia de nucleótidos de aproximadamente 8 a aproximadamente 24 nucleótidos de longitud (por ejemplo, de aproximadamente 8 a aproximadamente 22, de aproximadamente 8 a aproximadamente 20, de aproximadamente 8 a aproximadamente 18, de aproximadamente 8 a aproximadamente 16, de aproximadamente 8 a aproximadamente 14, de aproximadamente 8 a aproximadamente 12, de aproximadamente 8 a aproximadamente 10, de aproximadamente 10 a aproximadamente 24, de aproximadamente 10 a aproximadamente 22, de aproximadamente 10 a aproximadamente 20, de aproximadamente 10 a aproximadamente 18, de aproximadamente 10 a aproximadamente 16, de aproximadamente 10 a aproximadamente 14, de aproximadamente 10 a aproximadamente 12, de aproximadamente 12 a aproximadamente 24, de aproximadamente 12 a aproximadamente 22, de aproximadamente 12 a aproximadamente 20, de aproximadamente 12 a aproximadamente 18, de aproximadamente 12 a aproximadamente 16, de aproximadamente 12 a aproximadamente 14, de aproximadamente 14 a aproximadamente 24, de aproximadamente 14 a aproximadamente 22, de aproximadamente 14 a aproximadamente 20, de aproximadamente 14 a aproximadamente 18, de aproximadamente 14 a aproximadamente 16, de aproximadamente 16 a aproximadamente 24, de aproximadamente 16 a

aproximadamente 22, de aproximadamente 16 a aproximadamente 20, de aproximadamente 16 a aproximadamente 18, de aproximadamente 18 a aproximadamente 24, de aproximadamente 18 a aproximadamente 22, de aproximadamente 18 a aproximadamente 20, de aproximadamente 20 a aproximadamente 24, de aproximadamente 20 a aproximadamente 22 o de aproximadamente 22 a aproximadamente 24 nucleótidos de longitud).

En algunas modalidades, la sonda de bloqueo se elimina antes de hibridar la secuencia de captura de analito del oligonucleótido de la secuencia de captura de analito con el primer dominio de captura. Por ejemplo, una vez que la sonda de bloqueo se libera de la secuencia de captura de analito, la secuencia de captura de analito puede unirse al primer dominio de captura en la matriz. En algunas modalidades, la secuencia de captura de analito de bloqueo reduce la tinción de fondo no específica. En algunas modalidades, bloquear la secuencia de captura de analito permite controlar cuándo permitir la unión de la secuencia de captura de analito al dominio de captura de una sonda de captura durante un flujo de trabajo espacial, controlando así el tiempo de captura de la secuencia de captura de analito en la matriz. En algunas modalidades, las sondas de bloqueo se unen de manera reversible, de manera que las sondas de bloqueo se pueden eliminar de la secuencia de captura de analito durante o después del tiempo en que los agentes de captura de analito están en contacto con la muestra biológica. En algunas modalidades, la sonda de bloqueo se puede eliminar con el tratamiento con ARNasa (por ejemplo, el tratamiento con ARNasa H). En algunas modalidades, las sondas de bloqueo se eliminan al aumentar la temperatura (por ejemplo, calentando) la muestra biológica. En algunas modalidades, las sondas de bloqueo se eliminan enzimáticamente (por ejemplo, escindidas). En algunas modalidades, las sondas de bloqueo se eliminan por una enzima USUARIO. En algunas modalidades, las sondas de bloqueo se eliminan por una endonucleasa. En algunas modalidades, la endonucleasa es endonucleasa IV. En algunas modalidades, la endonucleasa es endonucleasa V.

En algunas modalidades, la determinación en la etapa (g) incluye a) extender los productos de ligazón capturados y los oligonucleótidos capturados de los agentes de captura de analito, en donde los productos de extensión comprenden el código de barras espacial o un complemento de este, b) liberar los productos de extensión, o los complementos de estos, de la matriz espacial, c) producir una biblioteca a partir de los productos de extensión liberados o los complementos de estos, y d) secuenciar la biblioteca. En algunas modalidades, la extensión (por ejemplo, la extensión de los productos de ligazón de ácido nucleico que se capturan y los oligonucleótidos que se capturan de los agentes de captura de analito y/o la extensión de la pluralidad de sondas de captura) se realiza con una polimerasa (por ejemplo, cualquier polimerasa adecuada, por ejemplo, la polimerasa T4).

En algunas modalidades, los productos de extensión liberados se pueden preparar para aplicaciones posteriores, tales como la generación de una biblioteca de secuenciación y la secuenciación de próxima generación. La producción de bibliotecas de secuenciación se conoce en la técnica. Por ejemplo, los productos de extensión liberados se pueden purificar y recolectar para las etapas de amplificación corriente abajo. Los productos de extensión liberados pueden amplificarse mediante el uso de PCR, donde los sitios de unión de cebadores flanquean el código de barras espacial y el producto de ligazón o el código de barras de la fracción de unión al analito, o los complementos de estos, generando una biblioteca asociada con un código de barras espacial particular. En algunas modalidades, la preparación de la biblioteca puede cuantificarse y/o controlarse con calidad para verificar el éxito de las etapas de preparación de la biblioteca. Los amplicones de la biblioteca se secuencian y analizan para decodificar la información espacial y el código de barras del producto de ligazón o del resto de la fracción de unión al analito, o complementos de estos.

Alternativa o adicionalmente, los amplicones pueden fragmentarse enzimáticamente y/o seleccionarse por tamaño para proporcionar el tamaño de amplicón deseado. En algunas modalidades, cuando se utiliza una metodología de preparación de la biblioteca de Illumina®, por ejemplo, P5 y P7, se pueden agregar secuencias a los amplicones lo que permite de esta manera la captura de la preparación de la biblioteca en una célula de flujo de secuenciación (por ejemplo, en instrumentos de secuenciación de Illumina). Adicionalmente, se pueden agregar secuencias de i7 e i5 como índices de muestra si se van a agrupar y secuenciar juntas varias bibliotecas. Además, las secuencias de lectura 1 y lectura 2 se pueden añadir a la preparación de la biblioteca para propósitos de secuenciación. Las secuencias antes mencionadas pueden añadirse a una muestra de preparación de la biblioteca, por ejemplo, a través de reparación de extremos, adición de cola de poliA, ligazón de adaptador y/o PCR. Los fragmentos de ADNc pueden entonces secuenciarse mediante el uso, por ejemplo, de la secuenciación de extremos emparejados mediante el uso de TruSeq Lectura 1 y TruSeq Lectura 2 como sitios de cebador de secuenciación, aunque se conocen otros métodos en la técnica.

En algunas modalidades, la determinación en la etapa (g) puede incluir una etapa de preamplificación. Por ejemplo, una hebra complementaria a los productos de ligazón de ARN extendido y/o el producto de extensión de los oligonucleótidos capturados de los agentes de captura de analito la etapa puede generarse y además incluir una etapa de preamplificación de los productos de extensión o complementos de estos (por ejemplo, productos extendidos) antes de la producción de la biblioteca (por ejemplo, producción de la biblioteca RTL; oligonucleótido capturado de la producción del agente de captura de analito).

Composiciones

En la presente descripción se proporcionan además matrices espaciales, que incluyen las matrices espaciales descritas en los métodos en la presente descripción, que incluyen una serie de diluciones de estándares de proteínas directamente en la matriz. En general, la cuantificación de proteínas con anticuerpos puede ser difícil debido a la afinidad variable que tienen los anticuerpos por sus objetivos de proteínas. Para cuantificar con precisión la abundancia de proteínas con anticuerpos, se pueden aplicar curvas estándar con la proteína de interés (por ejemplo, similar a un ensayo ELISA) a una matriz espacial en paralelo con el análisis proteómico espacial. En algunas modalidades, se coloca un estándar de proteína en la matriz (por ejemplo, en la parte superior de las características de la matriz). En algunas modalidades, se coloca más de un estándar de proteína en la matriz (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, o más estándares de proteína se colocan en la matriz). La lectura de la serie de estándares de proteínas internos permite la cuantificación de las proteínas de interés en paralelo a partir de la muestra biológica (por ejemplo, una sección de tejido) directamente en la matriz, lo que puede conducir a una mayor precisión al determinar la concentración de proteínas. Figura 26 muestra un sustrato ilustrativo con una configuración ilustrativa que incluye una serie de dilución para dos proteínas diferentes que se detectan en los márgenes de una matriz espacial. Figura 26 muestra una serie de diluciones para dos proteínas de interés, sin embargo, como se describe en la presente descripción, se pueden observar más de dos series de diluciones de proteínas en la matriz espacial.

En la presente descripción se proporcionan composiciones tales como matrices espaciales que incluyen a) una pluralidad de sondas de captura que incluyen códigos de barras espaciales y una primera pluralidad de dominios de captura hibridados con una pluralidad de productos de ligazón con plantilla, y b) una pluralidad de sondas de captura que comprenden códigos de barras espaciales y una segunda pluralidad de dominios de captura hibridados con una pluralidad de oligonucleótidos de los agentes de captura de analitos, en donde los oligonucleótidos comprenden una secuencia de captura de analito y un código de barras del resto de unión al analito. En algunas composiciones, las secuencias de captura de analito de los oligonucleótidos se hibridan con la segunda pluralidad de dominios de captura.

En algunas composiciones, las sondas de captura incluyen dominios de escisión, identificadores moleculares únicos, secuencias funcionales o sus combinaciones. En algunas composiciones, la primera pluralidad de dominios de captura son secuencias homopoliméricas. En algunas composiciones, la primera pluralidad de dominios de captura comprende secuencias de poli(T). En algunas composiciones, la primera pluralidad de dominios de captura se define como secuencias no homopoliméricas. En algunas composiciones, la primera pluralidad de dominios de captura comprende una secuencia redundante. En algunas composiciones, la primera pluralidad de dominios de captura comprende una secuencia fija. Por ejemplo, la primera pluralidad de dominios de captura puede comprender una de las SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, o SEQ ID NO: 11.

En algunas composiciones, la segunda pluralidad de dominios de captura son secuencias homopoliméricas. En algunas composiciones, la segunda pluralidad de dominios de captura se definen secuencias no homopoliméricas. En algunas composiciones, la segunda pluralidad de dominios de captura comprende secuencias de poli(T). En algunas composiciones, la segunda pluralidad de dominios de captura comprende una secuencia redundante. En algunas composiciones, la segunda pluralidad de dominios de captura comprende una secuencia fija. Por ejemplo, la segunda pluralidad de dominios de captura puede comprender una de las SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, o SEQ ID NO: 11.

En algunas composiciones, la matriz espacial comprende una o más series de dilución de proteínas.

Kits

En la presente descripción también se proporcionan kits que incluyen a) una matriz espacial que comprende una pluralidad de sondas de captura, donde las sondas de captura incluyen códigos de barras espaciales y en donde la pluralidad de sondas de captura comprende una primera pluralidad de primeros dominios de captura y una segunda pluralidad de segundos dominios de captura, b) uno o más agentes de captura de analitos, c) uno o más pares de sondas de ligazón con plantilla de ARN, y d) una o más enzimas y tampones para la práctica de cualquiera de los métodos descritos en la presente descripción. En algunos kits, una o más enzimas incluyen polimerasas, RNasas, DNasas, proteasas, lipasas o sus combinaciones.

Listado de secuencias

SEQ ID NO: 1 – Dominio de captura 22 mer
TTGCTAGGACCGGCCTTAAAGC

SEQ ID NO: 2 – Dominio de captura 21 mer
TTGCTAGGACCGGCCTTAAAG

SEQ ID NO: 3 – Dominio de captura 20 mer
TTGCTAGGACCGGCCTTAAA

SEQ ID NO: 4 – Dominio de captura 19 mer
TTGCTAGGACCGGCCTTAA

SEQ ID NO: 5 – Dominio de captura 18 mer

TTGCTAGGACCGGCCTTA

SEQ ID NO: 6 – Dominio de captura 17 mer

TTGCTAGGACCGGCCTT

SEQ ID NO: 7 – Dominio de captura 16 mer

5 TTGCTAGGACCGGCCT

SEQ ID NO: 8 – Dominio de captura 15 mer

TTGCTAGGACCGGCC

SEQ ID NO: 9 – Dominio de captura 14 mer

TTGCTAGGACCGGC

10 SEQ ID NO: 10 – Dominio de captura 13 mer

TTGCTAGGACCGG

SEQ ID NO: 11 – Dominio de captura 12 mer

TTGCTAGGACCG

15 Ejemplos

Ejemplo 1. Proteómica y expresión génica espacial en ganglios linfáticos humanos FFPE

20 Se realizaron experimentos para determinar si los agentes de captura de analito podrían proporcionar un análisis de la expresión proteica simultáneamente con la expresión génica asociada. En estos experimentos, el agente de captura de analito incluye un anticuerpo como el resto de unión al analito y la secuencia de captura de analito que comprende una secuencia de código de barras que identifica el anticuerpo así como también la secuencia de captura que es complementaria al dominio de captura de la sonda de captura asociada en la matriz.

25 Se permitió que las sondas de ligazón con plantilla se hibridaran con sus ARNm objetivo y los conjugados de anticuerpo-oligonucleótido se incubaron con las muestras en donde se permitió que los anticuerpos se unieran a sus objetivos de proteínas (como se describe en el documento PCT/US2020/66720). Brevemente, los tejidos de ganglios linfáticos humanos FFPE se seccionaron, se montaron en portaobjetos de matriz espacial y se desparafinaron mediante el uso de una serie de lavados con xileno y etanol antes de la obtención de imágenes de campo claro. Los
30 tejidos se lavaron y desenlazararon al incubar los tejidos en tampón TE (pH 9,0) durante 1 hora a 70°C. Después de la desreticulación de los tejidos, las sondas de ligazón con plantilla dirigidas se añadieron a los tejidos y la hibridación de la sonda se ejecutó durante la noche a 50 °C. Después de la hibridación, las sondas se ligaron juntas a 37 °C durante 1 hora. Después de la hibridación de la sonda de ligazón con plantilla, se añadieron los agentes de captura de analito a los tejidos, que se incubaron en un tampón de tinción de anticuerpos con los tejidos durante toda la noche a temperatura ambiente. Las muestras de tejido se lavaron cuatro veces con el tampón de tinción de anticuerpos sin anticuerpos.

Los tejidos se permeabilizaron y se permitió que los productos de sonda ligada o los productos de ligazón migraran para la hibridación a los dominios de captura de las sondas de captura en la superficie de la matriz espacial. Los
40 oligonucleótidos de los agentes de captura de analito también migraron en paralelo y se capturaron a través de sus secuencias de captura que se hibridaron para capturar los dominios de captura de la sonda de en la matriz espacial que son complementarios a las secuencias de captura. Como tal, los productos de ligazón que representan los ARNm objetivo y los oligonucleótidos de los agentes de captura de analito que representan la unión de los anticuerpos a las proteínas objetivo se capturaron simultáneamente en la superficie de la matriz. Para permitir la liberación y captura de la sonda y el oligonucleótido, los tejidos se incubaron con ARNasa H y un tampón asociado durante 30 minutos a 37 °C, los tejidos se permeabilizaron mediante el uso de una proteasa durante 40 minutos adicionales, seguido de un lavado para eliminar las enzimas de los tejidos.

Los productos de ligazón capturados y los oligonucleótidos del agente de unión al analito se extendieron para crear
50 productos extendidos de las moléculas capturadas, incluido el código de barras espacial o un complemento del mismo, el código de barras del resto del agente de unión al analito si está presente y otras secuencias funcionales de la sonda de captura. Se realizaron preparaciones de bibliotecas, las bibliotecas se secuenciaron en un instrumento de secuenciación Illumina y las localizaciones espaciales se determinaron mediante el uso de Space Ranger y Loupe Browser (10X Genomics). Las secuencias de anticuerpos (por ejemplo, el complemento del oligonucleótido capturado de los agentes de unión al analito) se amplificaron con Truseq_pR1 y Truseq_pR2. Para la localización de proteínas, se usaron secuencias relacionadas con el código de barras del resto de unión al analito para determinar la abundancia y la localización de la proteína marcada por los agentes de unión al analito. Los patrones de expresión espacial se determinaron mediante el uso del software de análisis de datos SpaceRanger y el software de visualización del navegador Loupe (10X Genomics).

60 Las Figuras 4A-B demuestra los resultados de un experimento donde se combinaron los agentes de captura de analito para determinar si se podían identificar múltiples objetivos simultáneamente en un tejido. En este experimento, se conjugaron ocho anticuerpos diferentes a oligonucleótidos que comprenden códigos de la fracción de unión al analito y secuencias de captura que se dirigen a ocho proteínas: NCAM1, Ki67, CD8A, PDL1, CD20, CD11B, CD45RA y
65 CD66B. La Figura 4A (ARN) muestra el agrupamiento de la expresión génica espacial de la proteína asociada en la muestra de tejido de ganglio linfático, mientras que la Figura 4B (proteína) muestra el agrupamiento de la expresión

proteica espacial de los ocho objetivos. Este experimento demuestra que es factible analizar en formato múltiple diferentes agentes de captura de analitos para identificar simultáneamente los patrones espaciales de expresión génica y proteica de los objetivos en formato múltiple en una muestra de tejido.

Resultados experimentales adicionales del experimento en formato múltiple que se muestra en la Figura 4A incluyó el direccionamiento a CD20 y CD20 ARNm cuando solo se dirige a una proteína con un agente de captura de analito. Estos resultados (datos no mostrados) muestran una muestra de tejido de ganglio linfático donde un anticuerpo dirigido a CD20 se conjugó con un oligonucleótido para su captura subsecuente en una matriz espacial. La ARNm de la sección de tejido se capturó y extendió con éxito, proporcionando de esta manera un análisis de expresión génica espacial de CD20 (datos no mostrados). Simultáneamente, la matriz de proteínas presentada espacialmente demuestra que el anticuerpo CD20 del agente de captura de analito fue capaz de unirse a la proteína CD20 seguido de la captura de oligonucleótidos en la matriz espacial, la extensión y la determinación espacial de la localización de la proteína CD20 a través del código de barras de la parte de unión al analito (datos no mostrados). Los patrones de expresión génica y de expresión proteica para CD20 se superponen y se localizan en los linfocitos B en los ganglios linfáticos (datos no mostrados). Como tal, los métodos demuestran la utilidad de los métodos y agentes de unión al analito que se dirigen a una proteína de interés y detectar simultáneamente la expresión génica espacial y la expresión proteica en una muestra de tejido.

Se generaron datos adicionales con el objetivo de CD8A (por ejemplo, mediante el uso de un anticuerpo CD8A) similar a CD20 descrito anteriormente. Los patrones de expresión génica espacial y de expresión proteica fueron consistentes con el direccionamiento de células T positivas para CD8 citotóxicas en el tejido de ganglios linfáticos (datos no mostrados), lo que demuestra nuevamente la utilidad de los métodos y agentes de unión al analito que se dirigen a proteínas específicas de interés para determinar simultáneamente la expresión espacial de genes y proteínas de un objetivo dado.

Ejemplo 2. Proteómica y expresión génica espacial en tejido de amígdala humana FFPE

Como se describió anteriormente, se realizaron experimentos para determinar si los agentes de captura de analitos podrían proporcionar un análisis de expresión proteica simultáneamente con la expresión génica asociada en el tejido de las amígdalas humanas FFPE.

Brevemente, los tejidos de amígdala humana FFPE se seccionaron, se montaron en portaobjetos de matriz espacial y se desparafinaron mediante el uso de una serie de lavados de xileno y etanol antes de la tinción H&E y la obtención de imágenes en campo claro. Los tejidos se lavaron y se desenlizaron al incubar los tejidos en tampón de citrato (pH 6,0) durante 1 hora a 90 °C. Después de desenlazar los tejidos, las sondas de ligazón dirigidas se aplicaron a los tejidos y la hibridación de la sonda se ejecutó durante la noche a 50 °C. Después de la hibridación, las sondas se ligaron juntas a 37 °C durante 1 hora para generar productos de ligazón. Después de la hibridación de la sonda de ligazón, los agentes de captura de analito se aplicaron a los tejidos, que se incubaron en un tampón de tinción de anticuerpos (tampón a base de PBS, suero de cabra al 5 %, ADN de esperma de salmón y sulfato de dextrano) con los tejidos durante toda la noche a 4°C. Las muestras de tejido se lavaron cuatro veces con tampón de tinción de anticuerpos sin anticuerpos.

Los tejidos se permeabilizaron y los productos de ligazón se dejaron migrar para su captura por hibridación a los dominios de captura de las sondas de captura en la superficie de la matriz espacial. Los oligonucleótidos de los agentes de captura de analito complementarios a las secuencias de captura alternativas del segundo conjunto de sondas de captura en la matriz también se capturaron por hibridación. Como tal, ambos productos de ligazón que representan el ARNm de la proteína objetivo y el oligonucleótido del agente de captura de analito que representa la unión del anticuerpo a la proteína objetivo se capturaron simultáneamente en la superficie de la matriz. Para permitir la liberación y captura de la sonda y el oligonucleótido, los tejidos se incubaron con RNasa H y un tampón asociado, y polietilenglicol (PEG) durante 30 minutos a 37 °C. Los tejidos se permeabilizaron mediante el uso de un tampón de permeabilización que comprende una proteasa (por ejemplo, proteinasa K), PEG, urea 1 M, durante 60 minutos adicionales, seguido de un lavado para eliminar las enzimas de los tejidos.

Los productos de ligazón capturados y los oligonucleótidos del agente de unión al analito se extendieron para incluir el código de barras espacial o un complemento del mismo, el código de barras del resto del agente de unión al analito o un complemento del mismo si está presente y otras secuencias funcionales de la sonda de captura. Adicionalmente, dichos productos se preamplificaron antes de la preparación de la biblioteca.

Las preparaciones de la biblioteca se realizaron a partir de los productos extendidos, las bibliotecas se secuenciaron en un instrumento de secuenciación Illumina y las localizaciones espaciales se determinaron y visualizaron. Las secuencias de anticuerpos (por ejemplo, el complemento del oligonucleótido capturado de los agentes de unión al analito) se amplificaron con Truseq_pR1 y Truseq_pR2. Para la localización de proteínas, se usaron secuencias relacionadas con el código de barras del resto de unión al analito para determinar la abundancia y la localización de la proteína marcada por los agentes de unión al analito. Los patrones de expresión espacial se determinaron mediante el uso del software de análisis de datos SpaceRanger y el software de visualización del navegador Loupe (10X Genomics).

La Figura 5 muestra datos de agrupamiento ilustrativos de genes espaciales (superior central) y expresión proteica (superior derecha) en secciones de tejido de amígdala FFPE superpuestas sobre la imagen de H&E (superior izquierda). La Figura 5 (abajo) muestra la visualización de los marcadores de proteínas PD-1, Ki67 y CD8A que etiquetan las células T foliculares, los folículos individuales y las células T supresoras/citotóxicas, respectivamente. Los datos demuestran la utilidad de los agentes de captura de analitos para dirigirse a una proteína de interés y detectar simultáneamente la expresión génica espacial en muestras de tejido de las amígdalas FFPE. Además, los datos demuestran la especificidad de los marcadores de proteínas (por ejemplo, anticuerpos) para identificar diferentes tipos de células dentro de una sección de tejido heterogénea y detectar simultáneamente la expresión génica espacial.

En otro experimento realizado en secciones de tejido de amígdala FFPE, se realizó el análisis espacial ilustrativo de la expresión génica y proteica de un conjugado de anticuerpo-oligonucleótido en formato de 20 dirigido a proteínas de interés y sondas con plantilla de ARN dirigidas a 18 000 objetivos de ARNm. La sección de tejido de amígdala humana FFPE se tiñó con H&E que identificó dónde se pueden ver los folículos que contienen células inmunitarias en maduración y la capa epitelial (datos no mostrados). Los datos de los grupos representativos de expresión génica y los datos de los grupos representativos de expresión proteica se alinean con las estructuras macroscópicas visibles en la imagen de H&E mostrada (datos no mostrados).

Los datos demuestran la viabilidad de usar en formato múltiple diferentes agentes de captura de analitos para identificar simultáneamente los patrones espaciales de expresión génica y proteica de los objetivos en formato múltiple en una muestra de tejido.

La Figura 6 muestra una sección de tejido de amígdala FFPE teñida con H&E ilustrativa y datos de expresión génica y espacial (arriba). La Figura 6 también muestra una vista ampliada de la sección de tejido de las amígdalas FFPE teñida con H&E y el agrupamiento de la expresión génica y la expresión proteica de CD8a y CD4 superpuestas en la sección de tejido de las amígdalas FFPE con H&E (parte inferior). La visualización de los marcadores CD8a y CD4 coincide con los folículos mostrados en la tinción H&E donde CD8a se localiza en el borde de los folículos y CD4 se localiza fuera de los folículos.

Se realizó una agrupamiento ilustrativa adicional de datos de expresión espacial de genes y proteínas en diferentes secciones de tejido de las amígdalas FFPE. El agrupamiento de la expresión génica espacial y el agrupamiento de proteínas espaciales se realizaron en la muestra de tejido de amígdala FFPE que también se tiñó con H&E (datos no mostrados). Los resultados demuestran un experimento en el que se combinaron los agentes de captura de analito para determinar si se podían identificar múltiples objetivos simultáneamente en el tejido de las amígdalas FFPE. En este experimento, se conjugaron 21 anticuerpos diferentes con oligonucleótidos que comprenden códigos de barras del resto de unión al analito y secuencias de captura (por ejemplo, secuencias de captura de analito) que se dirigen a 21 proteínas (18 se muestran en la Figura 7): Ki67, NCAM1, BCL, CD138, CD21, CD268, CD11b, LAG3, CD4, CD20, PD-L1, CD45RA, CD68, PanCK, CD66b, PCNA, CD235a /b y CD79a.

La Figura 7 muestra un panel de 18 de las 21 proteínas dirigidas, que incluye específicamente los genes: Ki67, NCAM1, BCL, CD138, CD21, CD268, CD11b, LAG3, CD4, CD20, PD-L1, CD45RA, CD68, PanCK, CD66b, PCNA, CD235a /b y CD79a. En el panel, la fila superior muestra la expresión proteica y el panel inferior muestra la expresión de ARN. Vistas ampliadas de la expresión espacial ilustrativa de genes y proteínas a partir de un subconjunto de siete mapas de calor de expresión génica y proteica dirigidos diferentes, que incluyen específicamente: PanCK, PD-L1, Ki67, CD4, CD68, CD20 y CD11b. Expresión génica (ARN) y expresión proteica para Ki67, NCAM1, BCL, CD138, CD11b, LAG, CD4, CD20, PD-L1, PanCK, CD66b y CD79a como se muestra en la Figura 7 superposición, demostrando la utilidad de los agentes de captura de analitos para dirigirse a una proteína de interés y detectar simultáneamente la expresión génica espacial en muestras de tejido de las amígdalas FFPE (datos no mostrados).

La Figura 8 muestra vistas ampliadas de datos de agrupamiento ilustrativos para el tejido de amígdala humana FFPE mediante el uso de un panel de prueba de conjugado anticuerpo-oligonucleótidos en formato de 12 de nueve genes que se muestran en la Figura 7, que incluyen específicamente: CD4, PD-L1, CD20, CD11b, BCL-2, CD21, CD791, CD45RA, PCNA, Ki67, PanCK y CD235a/b.

Colectivamente, los datos demuestran la utilidad de los métodos y los agentes de unión al analito que se dirigen a proteínas específicas de interés para determinar simultáneamente la expresión espacial de genes y proteínas de una objetivo dada.

Ejemplo 3. Proteómica y expresión génica espacial en tejido de cáncer de mama FFPE

Tejido de cáncer de mama triple positivo humano

Como se describió anteriormente, se realizaron experimentos para determinar si los agentes de captura de analitos podrían proporcionar un análisis de expresión proteica simultáneamente con la expresión génica asociada en el tejido de cáncer de mama humano triple positivo FFPE.

Brevemente, los tejidos de cáncer de mama triple positivo humano FFPE se seccionaron, se montaron en portaobjetos de matriz espacial y se desparafinaron mediante el uso de una serie de lavados de xileno y etanol antes de secarse a temperatura ambiente. A continuación, los portaobjetos se calentaron a 37 °C durante 15 minutos, seguido de una serie de lavados con etanol (100 %, 96 %, 96 % y 70 % de etanol). A continuación, los tejidos se tiñeron con H&E y se obtuvieron imágenes de campo claro. Alternativamente, los tejidos se pueden teñir (por ejemplo, teñidos por inmunofluorescencia) en lugar de la tinción H&E.

Los tejidos se lavaron y desreticularon al incubar los tejidos en tampón Tris-EDTA (TE) (pH 9,0) durante 1 hora a 95 °C seguido de una serie de lavados con HCl 0,1 N. Después de la desreticulación de los tejidos, las sondas de ligazón dirigidas se añadieron a los tejidos y la hibridación de la sonda se ejecutó durante la noche a 50 °C. Los tejidos se lavaron en un tampón de poshibridación que incluye 3x SSC, ARNt de levadura de Baker y agua sin nucleasa y seguido de un lavado con tampón 2x SSC. Después de la hibridación, las sondas se ligaron juntas a 37 °C durante 1 hora. Después de la hibridación de la sonda, los tejidos se incubaron en un tampón de bloqueo de anticuerpos (tampón a base de PBS (pH 7,4), suero de cabra, ADN de esperma de salmón, Tween-20, un inhibidor de RNasa y sulfato de dextrano) a temperatura ambiente durante 60 minutos. El tampón de bloqueo se eliminó de los tejidos y los tejidos se incubaron durante toda la noche a 4 °C con los agentes de captura de analito en una mezcla de tinción de anticuerpos (tampón a base de PBS (pH 7,4), suero de cabra al 5 %, ADN de esperma de salmón de 0,1 µg/µl, Tween-20 al 0,1 %, inhibidor de RNasa 1 U/µl, oligonucleótidos de bloqueo, agentes de captura de analito (por ejemplo, anticuerpos con un oligonucleótido conjugado) y 10 mg/ml de sulfato de dextrano)). Las muestras de tejido se lavaron varias veces con tampón de tinción de anticuerpos sin anticuerpos.

Los tejidos se permeabilizaron y las sondas ligadas se liberaron para su captura mediante la hibridación en los dominios de captura de las sondas de captura en la superficie de la matriz espacial. Los oligonucleótidos de los agentes de captura de analito complementarios a las secuencias de captura alternativas del segundo conjunto de sondas de captura en la matriz también se capturaron por hibridación. Como tal, ambos productos de ligazón que representan el ARNm objetivo y el oligonucleótido del agente de captura de analito que representa la unión del anticuerpo a la proteína objetivo se capturaron simultáneamente en la superficie de la matriz. Para permitir la liberación y captura de la sonda y el oligonucleótido, los tejidos se incubaron con una RNasa (por ejemplo, RNasa H), un tampón asociado y polietilenglicol (PEG) durante 30 minutos a 37 °C. Los tejidos se permeabilizaron mediante el uso de un tampón de permeabilización que comprende una proteasa (por ejemplo, Proteinasa K), PEG, urea 3 M, durante 60 minutos adicionales, seguido de un lavado para eliminar las enzimas de los tejidos. Después de la permeabilización, los tejidos se lavaron en 2x SSC varias veces.

Los productos de ligazón que se capturan y los oligonucleótidos del agente de unión al analito se extienden para crear productos extendidos de las moléculas que se capturan, incluido el complemento del código de barras espacial, el código de barras del resto del agente de unión al analito si está presente y otras secuencias funcionales de la sonda de captura. Adicionalmente, dichos productos se preamplificaron antes de la preparación de la biblioteca.

Las preparaciones de la biblioteca se realizaron a partir de los productos extendidos, secuenciados en un instrumento de secuenciación Illumina, y las localizaciones espaciales se determinaron y visualizaron. Las secuencias de anticuerpos (por ejemplo, el complemento del oligonucleótido capturado de los agentes de unión al analito) se amplificaron con Truseq_pR1 y Truseq_pR2. Para la localización de proteínas, se usaron secuencias relacionadas con el código de barras del resto de unión al analito para determinar la abundancia y la localización de la proteína marcada por los agentes de unión al analito. Los patrones de expresión espacial se determinaron mediante el uso del software de análisis de datos SpaceRanger y el software de visualización del navegador Loupe (10X Genomics).

Las Figuras 9A-C demostrar los resultados de un experimento donde los agentes de captura de analito se combinaron para determinar si las objetivos podían identificarse simultáneamente en el tejido de cáncer de mama ductal invasivo triple positivo de FFPE de grado II. En este experimento, se conjugaron 11 anticuerpos diferentes a oligonucleótidos que comprenden códigos de barras del resto de unión al analito y secuencias de captura (por ejemplo, secuencias de captura de analito) que se dirigen a once proteínas: Her2, EpCAM, PanCK, N-Cadherina, PCNA, AlfaSMA, Vimentina, CD8a, CD4, CD68, CD20 y HLA-DR. La Figura 9B (RNA) muestra el agrupamiento de la expresión génica espacial, mientras que la Figura 9C (proteína) muestra el agrupamiento de la proteína asociada en la muestra de tejido de cáncer de mama triple positivo FFPE. Los datos demuestran la viabilidad de usar en formato múltiple diferentes agentes de captura de analitos para identificar simultáneamente los patrones de grupos de expresión espacial de genes y proteínas de los objetivos en formato múltiple en muestras de tejido de cáncer de mama triple positivo.

Un panel de 11 proteínas objetivo que incluye específicamente los genes: Se generó Her2, EpCAM, PanCK, N-Cadherina, PCNA, AlfaSMA, Vimentina, CD8a, CD4, CD68, CD20 y HLA-DR donde para cada expresión de ARN y la expresión proteica objetivo se generó (datos no mostrados). También se generó una expresión espacial ilustrativa de genes y proteínas para KRT10, KRT18 y PanCK Ab y el agrupamiento imparcial de la expresión génica y proteica superpuestas en la imagen de H&E demuestra patrones similares (datos no mostrados). Se muestra una vista ampliada de la expresión de ARN de Her2 y Vimentina y la expresión proteica en la Figura 10. Las secciones adyacentes se tiñeron por inmunofluorescencia para HER2 (arriba) y Vimentina (abajo) demostrando una señal fuerte dentro de la región del cuadro discontinuo de la muestra de biopsia. La tinción de anticuerpos para dos biomarcadores, HER2 y Vimentina, a partir de secciones de tejido adyacentes se correlaciona con tumores en el tejido de cáncer de

mama triple positivo. La expresión génica (ARN) y la expresión proteica para Her2, EpCAM, PanCK, N-Cadherina, PCNA, AlfaSMA, Vimentina, CD8a, CD4, CD68, CD20 y HLA-DR también se superponen (datos no mostrados) demostrando la utilidad de los agentes de captura de analitos para dirigirse a una proteína de interés y detectar simultáneamente la expresión génica espacial en muestras de tejido de cáncer de mama triple positivo.

Como se describió anteriormente, se realizaron experimentos para determinar si los agentes de captura de analitos podrían proporcionar un análisis de expresión proteica simultáneamente con la expresión génica asociada en el tejido de cáncer de mama de carcinoma ductal invasivo FFPE.

Las muestras se prepararon como se describió anteriormente en el ejemplo de cáncer de mama triple positivo FFPE. Las Figuras 11A-C muestran de tejido de carcinoma ductal invasivo humano teñido con H&E (Figura 11A). Las Figuras 11B y 11C muestran el agrupamiento de la expresión génica y el agrupamiento de la expresión proteica, respectivamente, superpuestas sobre la imagen de H&E con patrones similares. La Figura 12 muestra un examen adicional de diferentes biomarcadores: CD11b, PD-L1 y alfa-SMA exhiben variación regional dentro de la muestra de tejido. Estos biomarcadores de proteínas se usaron para definir la región de interés (por ejemplo, mostrada como región 1 y región 2). Las regiones de interés identificadas de patrones de expresión génica específicos se identificaron con los 50 principales genes altamente expresados. Los datos demuestran la viabilidad de usar en formato múltiple diferentes agentes de captura de analitos para identificar simultáneamente los patrones de agrupamiento espacial de expresión génica y proteica de los objetivos de en formato múltiple en muestras de tejido de carcinoma ductal invasivo.

Figura 13 muestra la expresión espacial ilustrativa de genes y proteínas para Acta2 (arriba) y Epcam (abajo). El agrupamiento no sesgado de la expresión génica y proteica superpuesto sobre una imagen de H&E de la misma muestra de tejido demuestra patrones similares. También se generaron gráficos que muestran la correlación entre la expresión génica de Acta2 y la expresión proteica de Alfa-SMA (codificada por el gen Acta2) y la expresión génica del Epcam y la expresión proteica de EPCAM (codificada por el gen EPCAM) (datos no mostrados). Se realizó un experimento similar como se describe en Figura 13 en un tejido de cáncer de mama de carcinoma ductal invasivo FFPE diferente y la expresión espacial ilustrativa de genes y proteínas para Acta2 y, por separado, HLA-DRA (gen) y HLA-DR (proteína) se correlacionaron entre sí (datos no mostrados).

Colectivamente, los datos demuestran la viabilidad de usar en formato múltiple diferentes agentes de captura de analitos para identificar simultáneamente la expresión espacial de genes y proteínas en muestras de tejido de carcinoma ductal invasivo FFPE triple positivo.

Ejemplo 4. Proteómica y expresión génica espacial en secciones de bazo, cerebro, cabeza y torso de ratón FFPE en una muestra de embrión de ratón

Se realizaron experimentos para determinar si los agentes de captura de analito podrían proporcionar un análisis de la expresión proteica simultáneamente con la expresión génica asociada en tejidos de ratón FFPE, incluidos los embriones de ratón completos o porciones de los mismos. Los siguientes experimentos probaron la expresión espacial de genes y proteínas en formatos de intercalamiento y sin intercalamiento. En algunos ejemplos, la alineación de un primer sustrato con una muestra biológica y un segundo sustrato con una matriz espacial sobre el mismo se facilita mediante un proceso de intercalamiento. En consecuencia, en la presente descripción se describen métodos intercalar juntos el primer sustrato con una muestra biológica con un segundo sustrato que comprende una matriz con una pluralidad de sondas de captura, donde la sonda de captura incluye un código de barras espacial y un dominio de captura.

En un ejemplo no limitante, las muestras de bazo de ratón FFPE, las muestras de ratón FFPE, las muestras de torso de embriones de ratón FFPE, las muestras de cabeza de embriones de ratón FFPE se colocaron en portaobjetos estándar (para condiciones de intercalamiento) o portaobjetos de expresión espacial (GEx) (como condiciones sin intercalamiento). Las diapositivas de GEx incluyen una matriz de sondas de captura con código de barras espacial. Brevemente, los tejidos se seccionaron y montaron en portaobjetos y se secaron durante toda la noche en un deshidratador. Al día siguiente, los tejidos se calentaron a 60 °C, seguido de desparafinado y rehidratación. Los tejidos se tiñeron con H&E y se obtuvieron imágenes de campo claro. Los tejidos se desmancharon con HCl y se desenlazaran durante 1 hora en tampón de citrato (pH 6,0) a 95 °C. Después del desenlazado, los tejidos se incubaron durante toda la noche con conjuntos de sondas del transcriptoma completo de ratón (ligazón con plantilla de ARN) a 50 °C. Al día siguiente, los tejidos se lavaron para eliminar las sondas no hibridadas, luego se trataron con ligasa para enlazar las sondas RTL. Después de otra etapa de lavado, los tejidos se bloquearon con el tampón de bloqueo de anticuerpos. Los tejidos se incubaron durante toda la noche con una biblioteca de anticuerpos conjugados (por ejemplo, una biblioteca que comprende una pluralidad de agentes de captura de analitos, cada uno de los cuales comprende un anticuerpo específico de antígeno conjugado con un oligonucleótido). Al día siguiente, los tejidos se sometieron a condiciones de intercalamiento o sin intercalamiento como sigue.

Los tejidos colocados en portaobjetos estándar para las condiciones de intercalamiento se lavaron con PBS-T, se sometieron a una tinción de eosina y se lavaron con SSC. Los tejidos se sometieron a condiciones de intercalamiento. Brevemente, los portaobjetos de tejido se montaron en un instrumento de intercalamiento junto con un portaobjetos GEx y una solución de reactivo que incluye una RNasa y una Proteinasa K. Tras el cierre en el instrumento, las

secciones de tejido se permeabilizaron durante 30 minutos, lo que permite que los productos de ligazón y los oligonucleótidos de los agentes de captura de analito migren al portaobjetos GEx para su captura por las sondas de captura. Después de la permeabilización y captura, se retiraron los portaobjetos de GEx del instrumento.

- 5 Los tejidos colocados en portaobjetos GEx para condiciones sin intercalamiento se lavaron con PBS-T y SSC. Los tejidos se sometieron a una etapa de liberación de la sonda de 30 minutos con una RNasa, seguido de permeabilización con un tampón de permeabilización que incluye la proteasa K. En consecuencia, los productos de ligazón y los agentes de captura de analito se capturaron mediante las sondas de captura en el portaobjetos GEx.
- 10 Independientemente de las condiciones, los portaobjetos de GEx se lavaron dos veces con 2x SSC, y se sometieron a la extensión de la sonda, la desnaturalización y la preamplificación seguidas de la amplificación y el secuenciación de las bibliotecas de ligazón con plantilla y el agente de captura de analito.
- 15 Después de la secuenciación, se evaluó la calidad, la sensibilidad y la detección bajo cada condición (condiciones de intercalamiento y sin intercalamiento). Como se muestra en la Tabla 1, la calidad, sensibilidad y detección de las transcripciones detectadas globalmente (es decir, ARNm) y proteínas fueron comparables entre las condiciones de intercalamiento y sin intercalamiento.

Tabla 1.

	Métrica	Condiciones de intercalamiento	Condiciones sin intercalamiento
Calidad de ligazón con plantilla	Códigos de barras válidos	99,00 %	98,90 %
	Lecturas de fracciones en el objetivo	85,60 %	82,10 %
	Lecturas de fracciones utilizables	79,80 %	78,70 %
	Lecturas de fracciones en puntos	81,60 %	81,00 %
	Lecturas de fracciones no mapeadas	0,90 %	1,00 %
Sensibilidad de ligazón con plantilla	Mediana de genes (20K prps)	4856	4705
	Mediana de UMI (20K prps)	16 966	15 156
Detección de proteínas mediante el uso de agentes de captura de analito	Lecturas de fracciones utilizables	75,20 %	67,10 %
	Lecturas de fracciones en el punto	78,70 %	70,10 %
	Fracción desconocida	3,70 %	3,60 %
	Mediana de UMI por punto (5K lecturas utilizables por punto)	4632	4114
	Correlación del agente de captura de analito/ligazón con plantilla seleccionado	0,77	0,76

- 20 Se generaron imágenes para evaluar la superposición de la expresión génica y los perfiles de proteínas génicas en el tejido del bazo de ratón y el tejido cerebral de ratón para los biomarcadores individuales. Como se muestra en la Figura 14A (condiciones de intercalamiento) y la Figura 14B (condiciones sin intercalamiento), se detectó la expresión de ARNm del receptor de fosfatasa de tirosina de tipo C (Ptpcr; por ejemplo, Ensembl: ENSMUSG00000026395) en una muestra de tejido del bazo. Además, el producto génico (por ejemplo, la proteína) CD45R también se detectó tanto en condiciones de intercalamiento (Figura 14C) y en condiciones sin intercalamiento (Figura 14D). CD45R es el nombre de la proteína de Ptpcr, y se determinó si había superposición de la expresión de ARNm de Ptpcr y la expresión proteica de CD45R. Como se muestra en las Figuras 14A-14D, tanto las condiciones de intercalamiento (77 % de correlación) como sin intercalamiento (76 % de correlación), respectivamente, demuestran la superposición de transcripción y proteína, lo que indica que la detección de transcripción (es decir, ARNm) y proteína (1) se identificó en áreas similares de los tejidos y (2) fue comparable en las condiciones de intercalamiento y sin intercalamiento en muestras de bazo de ratón.

- 35 La calidad, sensibilidad y detección bajo cada condición (condiciones de intercalamiento y sin intercalamiento) se evaluaron de manera global en muestras de cerebro de ratón. Como se muestra en la Tabla 2, la calidad, sensibilidad y detección de transcripciones detectadas globalmente (es decir, ARNm) y proteínas fueron comparables entre las condiciones de intercalamiento y sin intercalamiento en muestras de cerebro de ratón.

Tabla 2.

	Métrica	Condiciones de intercalamiento	Condiciones sin intercalamiento
Calidad de ligazón con plantilla	Códigos de barras válidos	99,00 %	98,90 %
	Lecturas de fracciones en el objetivo	92,00 %	87,30 %
	Lecturas de fracciones utilizables	88,80 %	72,30 %
	Lecturas de fracciones en puntos	91,80 %	79,10 %
	Lecturas de fracciones no mapeadas	1,10 %	1,90 %

Sensibilidad de ligazón con plantilla	Mediana de genes (10K prps)	4405	3185
	Mediana de UMI (10K prps)	9386	5553
Detección de proteínas mediante el uso de agentes de captura de analito	Lecturas de fracciones utilizables	80,70 %	62,30 %
	Lecturas de fracciones en el punto	84,40 %	65,10 %
	Fracción desconocida	3,70 %	3,70 %
	Mediana de UMI por punto (5K lecturas utilizables por punto)	3232	3221
	Correlación del agente de captura de analito/ligazón con plantilla seleccionado	0,63	0,63

De manera similar a las imágenes de muestras de bazo de ratón descritas anteriormente, se evaluaron la expresión génica individual y los productos de proteínas en muestras de cerebro de ratón. Como se muestra en la Figura 15A (condiciones de intercalamiento) y la Figura 15B (condiciones sin intercalamiento), se detectó la expresión de ARNm de tirosina hidroxilasa (Th; por ejemplo, Ensembl: ENSMUSG00000000214) en el cerebro. Además, el producto génico (por ejemplo, proteína) TH también se detectó en condiciones de intercalamiento (Figura 15C) y condiciones sin intercalamiento (Figura 15D). Dado que TH es la proteína producida por Th, se determinó si había superposición de la expresión de ARNm de Th y la expresión proteica de TH. Como se muestra en las Figuras 15A-15D, tanto las condiciones de intercalamiento (63 % de correlación) como las condiciones sin intercalamiento (63 % de correlación), respectivamente, vieron la superposición de transcripción y proteína, lo que demuestra que la detección de transcripción (es decir, ARNm) y proteína fue comparable en las condiciones de intercalamiento y sin intercalamiento en muestras de cerebro de ratón.

Los experimentos que usan los mismos métodos (es decir, probar las condiciones de intercalamiento contra las condiciones sin intercalamiento mientras se detecta tanto el ARN como la proteína) se realizaron en secciones de la cabeza y el torso del embrión de ratón completo. Además de las condiciones en las que se detectaron tanto el ARN como la proteína, se incluyó una tercera condición como control. Esta tercera condición (Condición 3 en las Tablas 3 y 4) detectó la presencia y abundancia de solo ARN. En cada condición, el ARN se detectó mediante el uso de la ligazón con plantilla como se describió anteriormente. Para las condiciones 1 y 2 como se muestra en las Tablas 3 y 4 más abajo, la proteína también se detectó mediante el uso de métodos de agente de captura de analitos como se describió anteriormente. La calidad, sensibilidad y detección bajo cada condición (condiciones de intercalamiento frente a sin intercalamiento) se evaluaron en muestras de cabeza y torso de embriones de ratón. Como se muestra en las Tablas 3 y 4, la calidad, sensibilidad y detección de los transcritos detectados globalmente (es decir, ARNm) y proteínas fueron comparables entre las condiciones de intercalamiento y sin intercalamiento en muestras de torso y cabeza/torso superior de embriones de ratón. Además, la calidad, sensibilidad y detección de las transcripciones detectadas globalmente (es decir, ARNm) fue aproximadamente la misma entre las Condiciones 1 y 3, lo que demuestra que tanto la captura de proteínas como los métodos de intercalamiento no interfirieron con la captura de ARN mediante el uso de métodos de ligazón con plantilla.

Tabla 3. Datos de la muestra de torso de embrión de ratón completo

	Métrica	Condición 1: Condiciones de intercalamiento que detectan ARN y proteína	Condición 2: Condiciones sin intercalamiento que detectan ARN y proteína	Condición 3: Condiciones sin intercalamiento que detectan ARN
Calidad de ligazón con plantilla	Códigos de barras válidos	98,9 %	99,0 %	98,5 %
	Lecturas de fracciones en el objetivo	88,5 %	88,5 %	87,9 %
	Lecturas de fracciones utilizables	81,5 %	77,7 %	84,8 %
	Lecturas de fracciones en puntos	84,0 %	79,9 %	88,4 %
	Lecturas de fracciones no mapeadas	1,0 %	1,0 %	1,4 %
Sensibilidad de ligazón con plantilla	Mediana de genes (10K prps)	4144	4356	3562
	Mediana de UMI (10K prps)	8767	8176	6121
Detección de proteínas mediante el uso de agentes de captura de analito	Lecturas de fracciones utilizables	77,3 %	72,6 %	-
	Lecturas de fracciones en el punto	80,8 %	75,9 %	-
	Fracción desconocida	3,7 %	3,7 %	-
	Mediana de UMI por punto (5K lecturas utilizables por punto)	3358	3264	-
	Correlación del agente de captura de analito/ligazón con plantilla seleccionado	0,90	0,80	-

Tabla 4. Datos de la Muestra de Cabeza/Parte Superior del Torso del Embrión de Ratón Completo

	Métrica	Condición 1: Condiciones de intercalamiento que detectan ARN y proteína	Condición 2: Condiciones sin intercalamiento que detectan ARN y proteína	Condición 3: Condiciones sin intercalamiento que detectan ARN
Calidad de ligazón con plantilla	Códigos de barras válidos	99,0 %	99,0 %	98,5 %
	Lecturas de fracciones en el objetivo	89,1 %	89,4 %	89,0 %
	Lecturas de fracciones utilizables	86,7 %	82,9 %	82,2 %
	Lecturas de fracciones en puntos	89,5 %	85,2 %	85,2 %
	Lecturas de fracciones no mapeadas	1,0 %	1,0 %	1,5 %
Sensibilidad de ligazón con plantilla	Mediana de genes (10K prps)	4400	4708	3562
	Mediana de UMI (10K prps)	9077	8431	6571
Detección de proteínas mediante el uso de agentes de captura de analito	Lecturas de fracciones utilizables	84,3 %	77,9 %	-
	Lecturas de fracciones en el punto	88,2 %	81,5 %	-
	Fracción desconocida	3,7 %	3,8 %	-
	Mediana de UMI por punto (5K lecturas utilizables por punto)	3358	3349	-
	Correlación del agente de captura de analito/ligazón con plantilla seleccionado	0,72	0,59	-

Además de realizar un análisis de expresión global en cada grupo, se analizaron los objetivos individuales para la localización y abundancia de la expresión génica espacial (por ejemplo, ARNm) y la expresión proteica espaciales de objetivos individuales en la muestra de torso y cabeza/torso superior del embrión de ratón. Las Figuras 16A y 16C muestran la detección de ARNm (Figura 16A) y proteína (Figura 16C) del antígeno 76 de linfocitos (Ter119) (por ejemplo, NCBI Gene ID: 104231), de muestras de torso de embriones de ratón en la Condición 1 (es decir, en condiciones de intercalamiento). Las Figuras 16B y 16D muestran la detección de ARNm (Figura 16B) y proteína (Figura 16D) de Ter119 de muestras de torso de embriones de ratón en la Condición 2 (es decir, en condiciones sin intercalamiento). Las Figuras 17A y 17C muestran la detección de ARNm (Figura 17A) y proteína (Figura 17C) de Ter119 de muestras de cabeza de embriones de ratón en la Condición 1 (es decir, en condiciones de intercalamiento). Las Figuras 17B y 17D muestran la detección de ARNm (Figura 17B) y proteína (Figura 17D) de Ter119 de muestras de cabeza de embriones de ratón en la Condición 2 (es decir, en condiciones sin intercalamiento). Como se muestra en las Figuras 17A-17D, el ARNm y la proteína de Ter119 se detectaron fácilmente con una considerable superposición de ARNm y detección de proteína tanto en condiciones de intercalamiento como en condiciones sin intercalamiento, lo que demuestra la adaptabilidad y la reproducibilidad de los métodos independientemente de la condición.

Se analizaron biomarcadores únicos adicionales en las muestras de cabeza y torso del embrión de ratón. Como se muestra en las Figuras 18A-18I, se analizó el ARNm y la proteína del dominio de metalofosfoesterasa que contiene 1 (Mpped1; Ensembl: ENSMUSG000000041708). Mpped1 se detectó fácilmente en la región cerebral de muestras de cabeza de embriones de ratón. La Figura 18A muestra la detección de ARNA de Mpped1 en una muestra de cabeza de embrión de ratón mediante el uso de condiciones de intercalamiento (Condición 1 de las Tablas 3 y 4). La Figura 18E muestra la detección de la proteína Mpped1 en una muestra de cabeza de embrión de ratón mediante el uso de condiciones de intercalamiento (Condición 1 de las Tablas 3 y 4). Sin embargo, el ARN de Mpped1 no se detectó fácilmente en una muestra de torso de embrión de ratón (Figura 18B). De manera consistente con estas observaciones, Mpped1 tiene actividad de metalofosfoesterasa, que podría tener un papel en el desarrollo cerebral, como se esperaba que estuviera presente en la muestra de la cabeza del embrión y no en la muestra del torso del embrión. De acuerdo con los datos de las condiciones de intercalamiento, se detectó el ARNm y la proteína de Mpped1 en condiciones sin intercalamiento en la cabeza (Figura 18C (ARNm) y la Figura 18F (proteína)) pero no en el torso (Figura 18D (ARNm) y Figura 18G (proteína)). De manera similar, en condiciones sin intercalamiento en las que solo se detectó ARN, se detectó ARN Mpped1 en la cabeza (Figura 18H) pero no en el torso (Figura 18I). Como tal, independientemente de las condiciones, tanto la expresión génica como la proteica, hasta el nivel de un solo biomarcador, pueden detectarse simultáneamente en la misma muestra.

Se examinaron cuatro biomarcadores adicionales: troponina C1, tipo cardíaco y esquelético lento (Tnnc1; por ejemplo, Ensembl: ENSMUSG000000091898); factor de crecimiento de fibroblastos 15 (Fgf15; por ejemplo, Ensembl: ENSMUSG000000031073); epifcano (Epyc, por ejemplo, Ensembl: ENSMUSG000000019936); e inhibidor de peptidasa de serina (o cisteína), clado A, miembro 1E (Serpina1e; por ejemplo, Ensembl: ENSMUSG000000072849) en muestras de embriones de ratón de cabeza/cuerpo superior bajo las Condiciones 1, 2 y 3 de las Tablas 3 y 4. Ver las Figuras 19A-22I. Tnnc1 está involucrado en la regulación de la contracción muscular y su expresión se detectó fácilmente en

todas las muestras bajo cada condición. Ver las Figuras 19A-19I. La función de Fgf15 en la neurogénesis retiniana y como un factor de determinación del destino celular. De hecho, se encontró su expresión en el ojo del embrión en cada muestra de cabeza, pero no se detectó en el torso. Ver las Figuras 20A-20I. Funciones de Epyc en la formación ósea y en la estructura del cartílago y se detectó en cada muestra (cabeza y torso). Ver las Figuras 21A-21I. Finalmente, la Serpina1e es activa en el hígado ya que funciona en la producción de la proteína alfa-1 antitripsina. Como se muestra en las Figuras 22A-22I, Serpina1e se detectó en el torso de embriones de ratón pero no se detectó en las muestras de cabeza/torso superior. Consistente entre cada una de estas imágenes de biomarcadores, los métodos sin intercalamiento detectaron fácilmente los biomarcadores de ARNm y proteínas individuales en comparación con los métodos de intercalamiento.

Como tal, mediante el uso de condiciones de intercalamiento o sin intercalamiento, tanto la expresión génica como la proteica, hasta el nivel de un solo biomarcador, se pueden detectar simultáneamente en múltiples tipos de tejidos mediante el uso de los métodos descritos en la presente descripción.

Ejemplo 5. Proteómica espacial, expresión génica espacial e identificación de células inmunitarias en secciones de tejido canceroso FFPE

Se realizaron experimentos para determinar si los agentes de captura de analito podrían proporcionar un análisis espacial de proteínas y expresión génica en secciones de tejido canceroso FFPE. Adicionalmente, se realizaron experimentos para identificar varios tipos de células inmunitarias en secciones de tejido FFPE de cáncer de mama y secciones de tejido FFPE de cáncer de ovario. Las secciones de tejido se prepararon y el análisis se realizó mediante los métodos descritos en el Ejemplo 4.

Cáncer de mama de Carcinoma Ductal Invasivo

Como se muestra en las Figuras 23A-E se usó un panel de anticuerpos 25-plex para estudiar el microambiente tumoral y mostrar la infiltración espacial de células inmunitarias en una sección de tejido FFPE de cáncer de mama. Los datos mostrados en las Figuras 23A-E son de muestras extraídas de secciones de tejido más grandes que demostraron una expresión espacial consistente de genes y proteínas espaciales como se muestra en las Figuras 23C-E y descrito en la presente descripción (datos no mostrados). La Figura 23A muestra una sección de tejido de cáncer de mama de tipo de cáncer de mama ductal invasivo de grado III teñida con H&E y la Figura 23B muestra la imagen de H&E de la Figura 23A anotado por un patólogo. Las anotaciones del patólogo se delinean en la imagen y corresponden a Vasos Sanguíneos, DCIS (carcinoma ductal in situ), Células Inmunitarias, Carcinoma Invasivo, Necrosis o Glándula Normal. El patólogo identificó el carcinoma ductal in situ, así como también áreas de tejido con infiltrados de células inmunitarias.

La agrupamiento de la expresión génica espacial y de proteínas espaciales superpuesta a la imagen teñida con H&E mostrada en la Figura 23A demostraron patrones de expresión similares que delinean el tumor y la región estromal (datos no mostrados). Más específicamente, se usaron marcadores inmunitarios de proteínas para identificar células inmunitarias infiltrantes. Por ejemplo, la Figura 23C muestra la expresión proteica espacial de CD3 (por ejemplo, un marcador para células T) y la Figura 23D muestra la expresión proteica espacial de CD8A (por ejemplo, un marcador para linfocitos T citotóxicos). Además, dentro de un infiltrado, se distinguió la orientación espacial de diferentes tipos de células inmunitarias. Un ejemplo de esto se muestra en la Figura 23E donde se observó una acumulación de linfocitos T citotóxicos que expresan el isotipo de antígeno de leucocitos humanos – DR (HLA-DR) (por ejemplo, un marcador de activación de linfocitos T).

Adicionalmente, se examinó la expresión espacial de genes y proteínas de genes y proteínas adicionales a partir de la misma sección de tejido mostrada en la Figura 23A. Más específicamente, el gen ACTA2 y su proteína correspondiente, la actina de músculo liso alfa (SMA), demostraron patrones de expresión similares con correlaciones de recuento de UMI de proteína génica positiva (datos no mostrados). Un examen adicional del gen de la citoqueratina, KRT18, también demostró patrones similares a un anticuerpo PanCK y nuevamente se observaron correlaciones positivas de UMI de gen-proteína (datos no mostrados).

En una sección de tejido de cáncer de mama ductal invasivo FFPE de grado II diferente, la sección de tejido se teñió con H&E y la expresión proteica espacial de PanCK y HLA-DR dentro de la sección de tejido. La selección manual de las regiones positivas para HLA-DR y PanCK se realizó con el 10x Loupe Browser que muestra las regiones contrastantes dentro de la sección de tejido. El análisis de expresión diferencial local de ambas regiones generó los 50 principales genes asociados con la expresión en las regiones seleccionadas de HLA-DR o PanCK identificadas y se muestran en la Tabla 5.

Tabla 5.

ID del marcador	Nombre del marcador	Promedio de HLA-DR	Factor de cambio de Log2 de HLA-DR	Valor de P de HLA-DR	Promedio de PanCK	Factor de cambio de Log2 de PanCK	Valor de P de PanCK
ENSG00000106483	SFRP4	11,59067363	2,58736286	3,34E-49	1,926061057	-2,58736286	3,34E-49
ENSG00000168685	IL7R	3,519297315	2,433061874	1,02E-40	0,650696303	-2,433061874	1,02E-40

ES 3 008 686 T3

ENSG00000011465	DCN	29,40019273	2,218895222	4,74E-39	6,307770284	-2,218895222	4,74E-39
ENSG00000091986	CCDC80	3,432380045	2,310230989	7,36E-39	0,691066033	-2,310230989	7,36E-39
ENSG00000139329	LUM	33,17683334	2,178290672	9,94E-38	7,321262975	-2,178290672	9,94E-38
ENSG00000197614	MFAP5	3,117091908	2,247513206	2,12E-36	0,655476929	-2,247513206	2,12E-36
ENSG00000211772	TRBC2	4,483908589	2,20449644	3,84E-36	0,971529419	-2,20449644	3,84E-36
ENSG00000087245	MMP2	9,819947279	2,17666454	3,84E-36	2,169341797	-2,17666454	3,84E-36
ENSG00000182326	C1S	7,999797384	2,155136429	4,02E-36	1,793797074	-2,155136429	4,02E-36
ENSG00000108821	COL1A1	150,9804112	2,114722952	4,02E-36	34,81889197	-2,114722952	4,02E-36
ENSG00000136235	GNPMB	5,620650143	2,161589748	2,39E-34	1,254648708	-2,161589748	2,39E-34
ENSG00000090382	LYZ	5,075286879	2,204681169	3,37E-34	1,099543957	-2,204681169	3,37E-34
ENSG00000064205	CCN5	3,718695759	2,138373605	4,86E-34	0,843514881	-2,138373605	4,86E-34
ENSG00000211751	TRBC1	4,436189304	2,185129984	7,54E-34	0,974185322	-2,185129984	7,54E-34
ENSG00000164692	COL1A2	97,7921546	2,041795601	8,18E-34	23,72199689	-2,041795601	8,18E-34
ENSG00000277734	TRAC	4,685011293	2,099070424	4,58E-33	1,092107428	-2,099070424	4,58E-33
ENSG00000211592	IGKC	190,0999999	2,203176006	9,62E-33	41,23342956	-2,203176006	9,62E-33
ENSG00000271503	CCL5	4,115788386	2,085695324	1,34E-32	0,968342335	-2,085695324	1,34E-32
ENSG00000158747	NBL1	8,572428812	2,007423552	1,09E-31	2,129503248	-2,007423552	1,09E-31
ENSG00000106624	AEBP1	24,52600855	1,986108768	1,64E-31	6,183474011	-1,986108768	1,64E-31
ENSG00000169442	CD52	2,54616474	2,103282373	1,85E-31	0,59173525	-2,103282373	1,85E-31
ENSG00000168542	COL3A1	55,19587513	1,945997239	8,85E-31	14,30841332	-1,945997239	8,85E-31
ENSG00000163520	FBLN2	4,983256828	1,994115903	8,85E-31	1,249336902	-1,994115903	8,85E-31
ENSG00000172724	CCL19	4,214635477	2,569106092	1,13E-30	0,709126175	-2,569106092	1,13E-30
ENSG00000145423	SFRP2	12,30135013	1,975672819	1,40E-30	3,123873435	-1,975672819	1,40E-30
ENSG00000140937	CDH11	6,69433407	1,973568341	1,57E-30	1,702434001	-1,973568341	1,57E-30
ENSG00000106565	TMEM176B	2,784761169	2,020982234	4,69E-30	0,685223046	-2,020982234	4,69E-30
ENSG00000172061	LRRC15	6,252930678	1,96380266	5,20E-30	1,600978496	-1,96380266	5,20E-30
ENSG00000166741	NNMT	3,541452698	1,979047111	1,90E-29	0,897164127	-1,979047111	1,90E-29
ENSG00000184347	SLIT3	2,425162266	2,012530816	3,18E-29	0,600234141	-2,012530816	3,18E-29
ENSG00000107562	CXCL12	6,52390805	1,925167509	5,56E-29	1,715713517	-1,925167509	5,56E-29
ENSG00000149131	SERPING1	7,503857666	1,912104472	5,56E-29	1,991396278	-1,912104472	5,56E-29
ENSG00000123500	COL10A1	6,305762744	1,924239545	7,60E-29	1,659408368	-1,924239545	7,60E-29
ENSG00000140285	FGF7	1,174235279	2,165189314	8,22E-29	0,261340883	-2,165189314	8,22E-29
ENSG00000115594	IL1R1	2,104761348	2,001816916	1,23E-28	0,524806488	-2,001816916	1,23E-28
ENSG00000159403	C1R	10,59027289	1,895587338	2,40E-28	2,842878868	-1,895587338	2,40E-28
ENSG00000165507	DEPP1	5,92741698	2,006365507	2,68E-28	1,473495138	-2,006365507	2,68E-28
ENSG00000142871	CCN1	11,35378146	1,8940456	3,90E-28	3,051101685	-1,8940456	3,90E-28
ENSG00000143196	DPT	2,212129741	1,984447325	3,90E-28	0,558270869	-1,984447325	3,90E-28
ENSG00000213886	UBD	1,508270278	2,211600224	3,90E-28	0,325082561	-2,211600224	3,90E-28
ENSG00000186340	THBS2	6,307467004	1,90036637	3,90E-28	1,687560943	-1,90036637	3,90E-28
ENSG00000060718	COL11A1	5,138344506	1,953199093	1,11E-27	1,325295736	-1,953199093	1,11E-27
ENSG00000197747	S100A10	6,643206264	1,86886665	1,72E-27	1,816637842	-1,86886665	1,72E-27
ENSG00000132386	SERPINF1	11,63498439	1,853278452	1,80E-27	3,216298869	-1,853278452	1,80E-27
ENSG00000118523	CCN2	16,70174997	1,860743031	3,07E-27	4,593119128	-1,860743031	3,07E-27
ENSG00000127083	OMD	1,624159972	2,000597227	4,93E-27	0,40529084	-2,000597227	4,93E-27
ENSG00000084636	COL16A1	3,253432724	1,891287213	6,63E-27	0,875916901	-1,891287213	6,63E-27
ENSG00000180447	GAS1	2,45413469	1,901572191	2,20E-26	0,65600811	-1,901572191	2,20E-26
ENSG00000163430	FSTL1	10,87829286	1,80177238	6,32E-26	3,116436906	-1,80177238	6,32E-26
ENSG00000227507	LTB	2,731929102	1,888847173	8,98E-26	0,736747569	-1,888847173	8,98E-26

Los datos demuestran que la expresión proteica espacial puede identificar la infiltración de células inmunitarias en secciones de tejido FFPE de cáncer de mama y, además, la expresión proteica espacial se correlaciona con las anotaciones de un patólogo que demuestra la utilidad de los métodos descritos en la presente descripción para identificar células inmunitarias dentro de un microambiente tumoral que puede usarse como una herramienta de diagnóstico.

Carcinoma de ovario

Como se muestra en las Figuras 24A-D se usó un panel de anticuerpos en formato de 25 para estudiar una sección de tejido de cáncer de ovario FFPE (por ejemplo, carcinoma) y mostrar la infiltración espacial de células inmunitarias.

El panel de anticuerpos incluyó anticuerpos para marcadores intracelulares y extracelulares. La Figura 24A muestra una anotación del patólogo para el carcinoma invasivo y las células inmunitarias en una sección de tejido de cáncer de ovario FFPE teñida con H&E. Las anotaciones del patólogo se delinean en la imagen y corresponden a Vasos Sanguíneos, DCIS (carcinoma ductal in situ), Células Inmunitarias, Carcinoma Invasivo o Necrosis.

El panel de anticuerpos en formato de 25 incluyó anticuerpos para marcadores inmunitarios de proteínas que confirmaron la presencia de células inmunitarias dentro de la sección de tejido de FFPE de cáncer de ovario. Además, el panel de anticuerpos distinguió (por ejemplo, subtipificó) las células inmunitarias en base a sus marcadores de superficie característicos. Por ejemplo, la Figura 24B muestra la expresión proteica espacial de CD20 (por ejemplo, un marcador para las células B) y la Figura 24C muestra la expresión proteica espacial de CD68 (por ejemplo, un marcador para monocitos). La Figura 24D muestra la expresión proteica espacial de CD8A (por ejemplo, un marcador para linfocitos T citotóxicos) y una pequeña región del carcinoma de ovario incluye la infiltración de linfocitos T citotóxicos (flecha) mientras que el carcinoma más grande no.

Las Figuras 25A-D muestran la infiltración diferencial de linfocitos T citotóxicos espaciales dentro de diferentes regiones de una sección de tejido de FFPE de cáncer de ovario. Los perfiles de expresión génica de las regiones de carcinoma grande y pequeño se compararon para encontrar diferencias que pueden correlacionarse con la variación de la infiltración de células inmunitarias observada en los datos de expresión proteica. La Figura 25A muestra la expresión proteica espacial de CD8A (por ejemplo, un marcador para linfocitos T citotóxicos) y la Figura 25B muestra áreas altamente infiltradas ("calientes") y no infiltradas ("frías") de la sección de tejido de FFPE de cáncer de ovario. El mapeo de infiltración se basa en la detección de proteínas de la expresión proteica CD3 para la infiltración de linfocitos T citotóxicos "activos" o alta, y la expresión proteica CD8 para la infiltración de linfocitos T citotóxicos "fríos" o mínima a ninguna. Las Figuras 25C y 25D muestran la expresión génica espacial de los genes de respuesta inmunitaria. Por ejemplo, la Figura 25C muestra la expresión génica espacial del antígeno de histocompatibilidad de clase I HLA, cadena alfa G (HLA-G; también conocido como antígeno de leucocitos humanos G) en el pequeño carcinoma que se correlaciona con los datos de expresión proteica y la Figura 25D muestra la expresión génica espacial en el carcinoma grande de proteoglicano de matriz interfectorreceptora 2 (IMPG2), que se sabe que está involucrado en el crecimiento del tumor, que también se correlaciona con los datos de expresión proteica. Los datos demuestran que la expresión proteica espacial puede identificar la infiltración de células inmunitarias en secciones de tejidos FFPE de cáncer de ovario y, además, la expresión proteica espacial se correlaciona con las anotaciones de un patólogo y con los datos de expresión génica que demuestran la utilidad de los métodos descritos en la presente descripción para identificar células inmunitarias dentro de un microambiente tumoral. Los métodos descritos en la presente descripción también pueden usarse como una herramienta de diagnóstico y/o método de cribado. Adicionalmente, la expresión génica espacial que se correlaciona con diferentes carcinomas ováricos dentro de una sección de tejido es distinguible por los métodos descritos en la presente descripción.

Ejemplo 6. Proteómica espacial y expresión génica espacial en secciones de tejido canceroso FFPE

Se realizaron experimentos para determinar si los agentes de captura de analito podrían proporcionar un análisis espacial de proteínas y expresión génica en secciones de tejido canceroso FFPE. Las secciones de tejido se prepararon y el análisis se realizó mediante los métodos descritos en el Ejemplo 4.

Cáncer de pulmón

Como se muestra en las Figuras 27A-D análisis espacial de la expresión génica y proteica en una sección de tejido de cáncer de pulmón FFPE. La Figura 27A muestra una sección de tejido de cáncer de pulmón FFPE teñida con H&E y las Figuras 27B y 27C, muestran el agrupamiento de la expresión génica espacial (Figura 27B) y agrupamiento de la expresión proteica espacial (Figura 27C), respectivamente. La Figura 27D muestra un gráfico de UMI espacial de proteínas HLA-DR. Como se describe en la presente descripción, HLA-DR es un marcador de activación de linfocitos T y la Figura 27D

Los datos demuestran que la expresión génica espacial y la expresión proteica espaciales se correlacionan entre sí, lo que demuestra la utilidad de los métodos descritos en la presente descripción para identificar la expresión espacial de genes y proteínas en secciones de tejido FFPE de cáncer de pulmón que pueden usarse como una herramienta de diagnóstico.

Melanoma

Como se muestra en las Figuras 28A-D análisis espacial de la expresión génica y proteica en una sección de tejido de melanoma FFPE. La Figura 28A muestra una sección de tejido de cáncer de pulmón FFPE teñida con H&E y las Figuras 28B y 28C, muestran el agrupamiento de la expresión génica espacial (Figura 28B) y agrupamiento de la expresión proteica espacial (Figura 28C), respectivamente. La Figura 28D muestra el gráfico de UMI espacial de la proteína HLA-DR. Como se describe en la presente descripción, HLA-DR es un marcador de activación de linfocitos T y la Figura 28D

Los datos demuestran que la expresión génica espacial y la expresión proteica espaciales se correlacionan entre sí, lo que demuestra la utilidad de los métodos descritos en la presente descripción para identificar la expresión espacial de genes y proteínas en secciones de tejido FFPE de melanoma que pueden usarse como una herramienta de diagnóstico.

5 Otras tejidos analizados
Además de las diversas secciones de tejido FFPE analizadas como se describe en la presente descripción, se analizaron otros tipos de tejido, que incluyen: tejido cerebral sano, tejido de carcinoma lobular invasivo de cáncer de mama, tejido de mama sano, tejido de cáncer de colon, tejido de colon sano, tejido de glioblastoma, tejido de corazón,
10 tejido pulmonar sano, tejido de cáncer de próstata, tejido de bazo sano, tejido de testículos, tejido de amígdala inflamada, tejido de cuello uterino y tejido de ganglios linfáticos (no se muestran los datos).

Ejemplo 7. La expresión proteica espacial correlaciona con la tinción de inmunofluorescencia

15 Las Figuras 29A-C muestran la tinción por inmunofluorescencia de anticuerpos contra vimentina (VIM) (Figura 29A) y tinción con DAPI en una sección de tejido de cáncer de mama de grado II de carcinoma ductal invasivo FFPE. La Figura 29B muestra la expresión proteica espaciales superpuesta sobre la imagen fluorescente en la Figura 29A. La Figura 29C muestra la expresión proteica espacial con un agente de captura de analito específico de VIM en la misma sección de tejido. Los datos demuestran una distribución espacial similar entre la tinción por inmunofluorescencia de anticuerpos contra vimentina (Figura 29A) y la expresión proteica espaciales con un agente de captura de analito
20 específico de VIM (Figura 29C) y demuestra la utilidad de los métodos descritos en la presente descripción para recapitular la expresión proteica espacial con agentes de captura de analito en relación con la tinción de anticuerpos inmunofluorescentes.

25 Otras modalidades

Debe entenderse que, aunque la invención se ha descrito junto con la descripción detallada de la misma, la descripción anterior pretende ilustrar y no limitar el alcance de la invención, que se define por el alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un método para determinar la localización espacial de un ácido nucleico y una proteína en una muestra biológica que comprende:
 - 5 a) proporcionar una matriz espacial que comprende una primera y segunda pluralidad de sondas de captura, en donde cada una de la primera y segunda pluralidad de sondas de captura comprende un código de barras espacial y un dominio de captura,
 - b) poner en contacto la matriz espacial con la muestra biológica,
 - 10 c) poner en contacto la muestra biológica con
 - (i) una pluralidad de agentes de captura de analito, en donde un agente de captura de analito de la pluralidad de agentes de captura de analito comprende un resto de unión a proteína y un oligonucleótido que comprende un código de barras de resto de unión a proteína y una secuencia de captura de analito, en donde la secuencia de captura de analito comprende una secuencia complementaria al dominio de captura de la segunda pluralidad de sondas de captura, y
 - 15 (ii) una pluralidad de sondas de ligazón con plantilla, en donde una de las sondas de ligazón con plantilla comprende una secuencia complementaria al dominio de captura de la primera pluralidad de sondas de captura,
 - 20 d) unir el resto de unión a proteína del agente de captura de analito a una proteína objetivo,
 - e) hibridar la pluralidad de sondas de ligazón con plantilla a un ácido nucleico objetivo y ligar las sondas de ligazón con plantilla hibridadas para producir un producto de ligazón con plantilla,
 - f) hibridar: (i) el producto de ligazón con plantilla al dominio de captura de la primera pluralidad de sondas de captura y (ii) la secuencia de captura de analito del agente de captura de analito unido al dominio de captura de la segunda pluralidad de sondas de captura en la matriz espacial, y
 - 25 g) determinar: (i) la secuencia o una porción de esta del producto de ligazón con plantilla capturado, o un complemento de esta, y la secuencia del código de barras espacial de la primera pluralidad de sondas de captura, o un complemento de esta, que se asocia con el producto de ligazón con plantilla, y (ii) la secuencia del código de barras del resto de unión a proteína, o un complemento de esta, del agente de captura de analito unido, y la secuencia del código de barras espacial de la segunda pluralidad de sondas de captura, o un complemento de esta, que se asocia con el agente de captura de analito, para determinar de esta manera la localización espacial del ácido nucleico y la proteína en la muestra biológica.
 - 30
- 35 2. El método de la reivindicación 1, en donde los dominios de captura de la primera pluralidad de sondas de captura son secuencias de captura no homopoliméricas o secuencias homopoliméricas definidas, opcionalmente en donde los dominios de captura de la segunda pluralidad de sondas de captura son secuencias de captura homopoliméricas o secuencias no homopoliméricas definidas, en donde opcionalmente los dominios de captura de la primera pluralidad de sondas de captura son diferentes de los dominios de captura de la segunda pluralidad de sondas de captura.
- 40 3. El método de la reivindicación 2, en donde las secuencias homopoliméricas comprenden secuencias de poliT y las secuencias no homopoliméricas comprenden secuencias fijas o secuencias redundantes, preferentemente en donde las secuencias fijas comprenden al menos una secuencia seleccionada de SEQ ID NO: 1 a SEQ ID NO: 11.
- 45 4. El método de una cualquiera de las reivindicaciones 1-3, en donde el ácido nucleico es un ARN o ADN, preferentemente en donde el ARN es un ARNm.
- 50 5. El método de una cualquiera de las reivindicaciones 1-4, en donde la muestra biológica es una muestra de tejido, preferentemente en donde la muestra de tejido es una muestra de tejido recién congelada o una muestra de tejido fijada, en donde la muestra de tejido fijada es una muestra de tejido fijada con formaldehído, una muestra de tejido fijada con acetona, una muestra de tejido con paraformaldehído o una muestra de tejido fijada con metanol.
- 55 6. El método de una cualquiera de las reivindicaciones 1-4, en donde la muestra biológica es una sección de tejido, en donde la sección de tejido es una sección de tejido recién congelada o una sección de tejido fijada, y opcionalmente, en donde la sección de tejido fijada es una sección de tejido embebida en parafina fijada con formalina, una sección de tejido fijada con acetona, una sección de tejido con paraformaldehído o una sección de tejido fijada con metanol.
- 60 7. El método de la reivindicación 5, en donde la muestra de tejido se deriva de una muestra de biopsia o un embrión de roedor completo.
- 65 8. El método de una cualquiera de las reivindicaciones 1-7, en donde antes de (b) la muestra biológica se desparafina y desreticula, preferentemente en donde la desreticulación comprende el uso de un tampón, opcionalmente en donde:

- (i) el tampón comprende tampón Tris-EDTA a un pH de aproximadamente 8 a aproximadamente 10 y una temperatura de aproximadamente 60 °C a aproximadamente 80 °C; o
- (ii) el tampón comprende tampón de citrato a un pH de aproximadamente 5 a aproximadamente 7 y una temperatura de aproximadamente 70 °C a aproximadamente 100 °C.

9. El método de la reivindicación 1, en donde la hibridación del producto de ligazón con plantilla y las secuencias de captura de analito de los agentes de captura de analito unidos en (f) comprende permeabilizar la muestra biológica.

10. El método de la reivindicación 1, en donde la secuencia de captura de analito del oligonucleótido se bloquea antes de la unión a la proteína objetivo, preferentemente en donde la secuencia de captura de analito se bloquea mediante una sonda de bloqueo, preferentemente en donde la sonda de bloqueo se elimina antes de hibridar la secuencia de captura de analito del agente de captura de analito unido al dominio de captura de la segunda pluralidad de sondas de captura en (f).

11. El método de la reivindicación 1, en donde la determinación en (g) comprende además:

- a) extender el producto de ligazón con plantilla capturado y el oligonucleótido capturado de los agentes de captura de analito, en donde los productos de extensión comprenden el código de barras espacial de la primera y segunda pluralidad de sondas de captura o un complemento de estas, respectivamente,
- b) liberar los productos de extensión, o complementos de estos, de la matriz espacial,
- c) producir una biblioteca a partir de los productos de extensión liberados o complementos de estos, y
- d) secuenciar la biblioteca,

opcionalmente en donde:

- (i) antes de (c) el método comprende además preamplificar los productos de extensión, o complementos de estos; o
- (ii) el complemento del oligonucleótido capturado del agente de captura de analito comprende el código de barras del resto de unión a proteína específico para el resto de unión a proteína del agente de captura de analito.

12. El método de una cualquiera de las reivindicaciones 1-11, en donde cada una de la primera y segunda pluralidad de sondas de captura comprende además un dominio de escisión, uno o más dominios funcionales, un identificador molecular único y una de sus combinaciones.

13. El método de una cualquiera de las reivindicaciones 1-12, en donde el método comprende además la obtención de imágenes de la muestra biológica, preferentemente en donde la obtención de imágenes comprende uno o más de microscopía de expansión, microscopía de campo claro, microscopía de campo oscuro, microscopía de contraste de fases, microscopía electrónica, microscopía de fluorescencia, microscopía de reflexión, microscopía de interferencia y microscopía confocal.

14. El método de una cualquiera de las reivindicaciones 1-13, en donde el método comprende además teñir la muestra biológica, preferentemente en donde:

- (i) la tinción comprende hematoxilina y eosina; o
- (ii) la tinción comprende el uso de un radioisótopo, un fluoróforo, un compuesto quimioluminiscente, un compuesto bioluminiscente o una de sus combinaciones.

15. El método de una cualquiera de las reivindicaciones 1-14, en donde la matriz espacial comprende una o más series de dilución de proteínas.

16. Una matriz espacial que comprende:

- a) una pluralidad de sondas de captura que comprenden códigos de barras espaciales y una primera pluralidad de dominios de captura hibridados con una pluralidad de productos de ligazón con plantilla producidos por la ligazón de sondas de ligazón con plantilla, y
- b) una pluralidad de sondas de captura que comprenden códigos de barras espaciales y una segunda pluralidad de dominios de captura hibridados con una pluralidad de oligonucleótidos de los agentes de captura de analito, en donde la pluralidad de oligonucleótidos comprende una secuencia de captura de analito y un código de barras del resto de unión a proteína,

opcionalmente en donde:

- i) las sondas de captura comprenden además dominios de escisión, identificadores moleculares únicos, una o más secuencias funcionales, o una de sus combinaciones;

- ii) la primera pluralidad de dominios de captura son secuencias homopoliméricas o secuencias no homopoliméricas definidas;
- iii) la segunda pluralidad de dominios de captura son secuencias homopoliméricas o secuencias no homopoliméricas definidas; o
- iv) la primera pluralidad de dominios de captura son secuencias de poli(T);

opcionalmente en donde la matriz espacial comprende una o más series de dilución de proteínas.

17. Un método para determinar la localización espacial de un ácido nucleico y una proteína en una muestra biológica de enfermedad que comprende:

a) proporcionar una matriz espacial que comprende una primera y una segunda pluralidad de sondas de captura en donde cada una de la primera y segunda pluralidad de sondas de captura comprende un código de barras espacial y un dominio de captura,

b) poner en contacto la muestra biológica de enfermedad con la matriz espacial,

c) poner en contacto la muestra biológica de enfermedad con:

- (i) una pluralidad de agentes de captura de analito, en donde un agente de captura de analito de la pluralidad de agentes de captura de analito comprende un resto de unión a proteína y un oligonucleótido que comprende un código de barras de resto de unión a proteína y una secuencia de captura de analito, en donde la secuencia de captura de analito comprende una secuencia complementaria al dominio de captura de la segunda pluralidad de sondas de captura, y
- (ii) una pluralidad de sondas de ligazón con plantilla, en donde una de las sondas de ligazón con plantilla comprende una secuencia complementaria al dominio de captura de la primera pluralidad de sondas de captura,

d) unir el resto de unión a proteína del agente de captura de analito a una proteína objetivo,

e) hibridar la pluralidad de sondas de ligazón con plantilla a un ácido nucleico objetivo y ligar las sondas de ligazón con plantilla hibridadas para producir un producto de ligazón con plantilla,

f) hibridar: (i) el producto de ligazón con plantilla al dominio de captura de la primera pluralidad de sondas de captura y (ii) las secuencias de captura de analito de los agentes de captura de analito unidos al dominio de captura de la segunda pluralidad de sondas de captura en la matriz espacial, y

g) determinar: (i) la secuencia o una porción de esta del producto de ligazón con plantilla capturado, o un complemento de esta, y la secuencia del código de barras espacial de la primera pluralidad de sondas de captura, o un complemento de esta, que se asocia con el producto de ligazón con plantilla, y (ii) la secuencia del código de barras del resto de unión a proteína, o un complemento de esta, del agente de captura de analito unido, y la secuencia del código de barras espacial de la segunda pluralidad de las sondas de captura, o un complemento de esta, que se asocia con el agente de captura de analito unido, para determinar de esta manera la localización espacial del ácido nucleico y la proteína en la muestra biológica de enfermedad,

opcionalmente en donde la muestra biológica de enfermedad es una muestra biológica cancerosa, preferentemente en donde la muestra biológica cancerosa es una muestra biológica de cáncer de ovario, una muestra biológica de cáncer de mama, una muestra biológica de cáncer de pulmón o un melanoma, preferentemente en donde la muestra biológica de cáncer de mama es una muestra de cáncer de mama triple positivo o un carcinoma ductal invasivo.

DIBUJOS

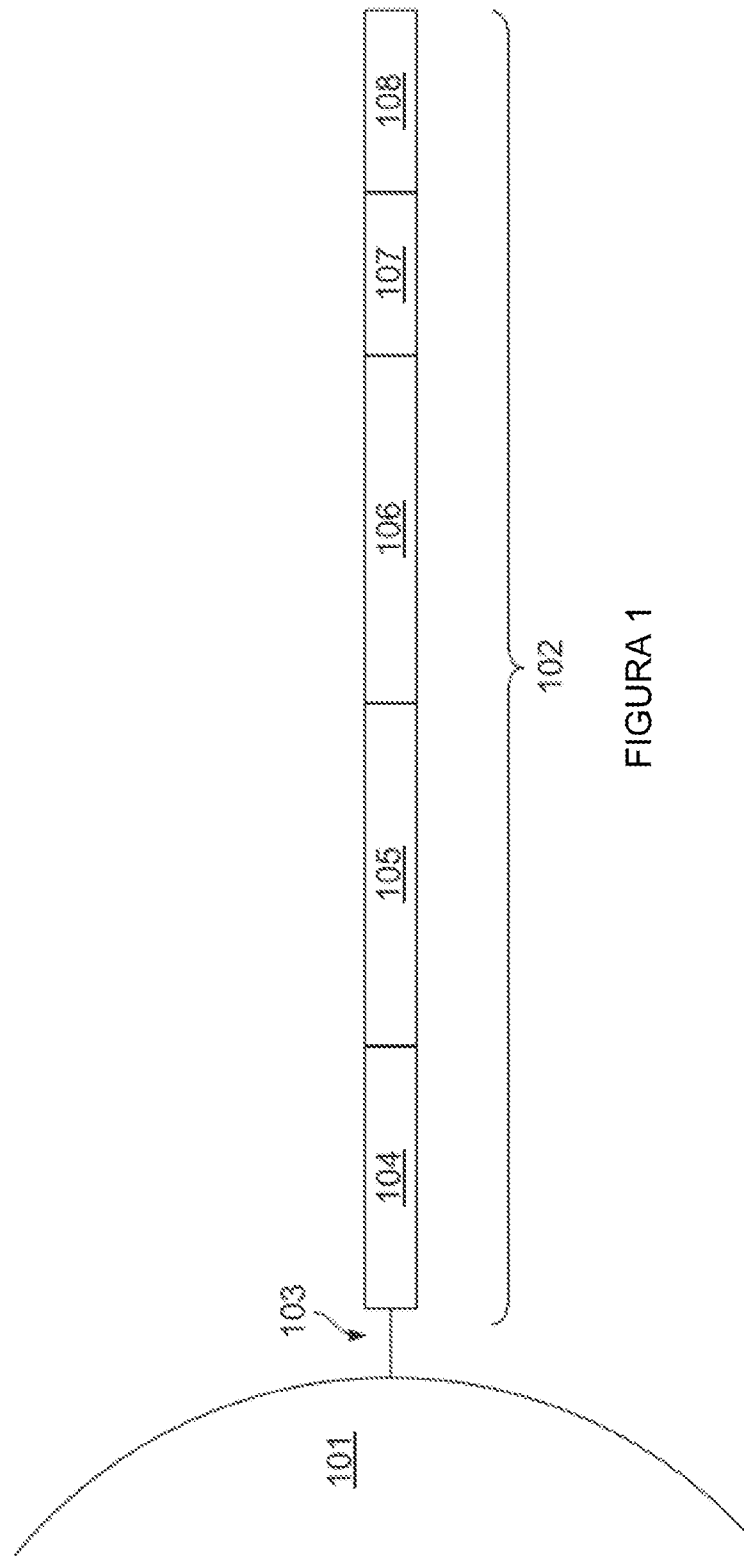


FIGURA 1

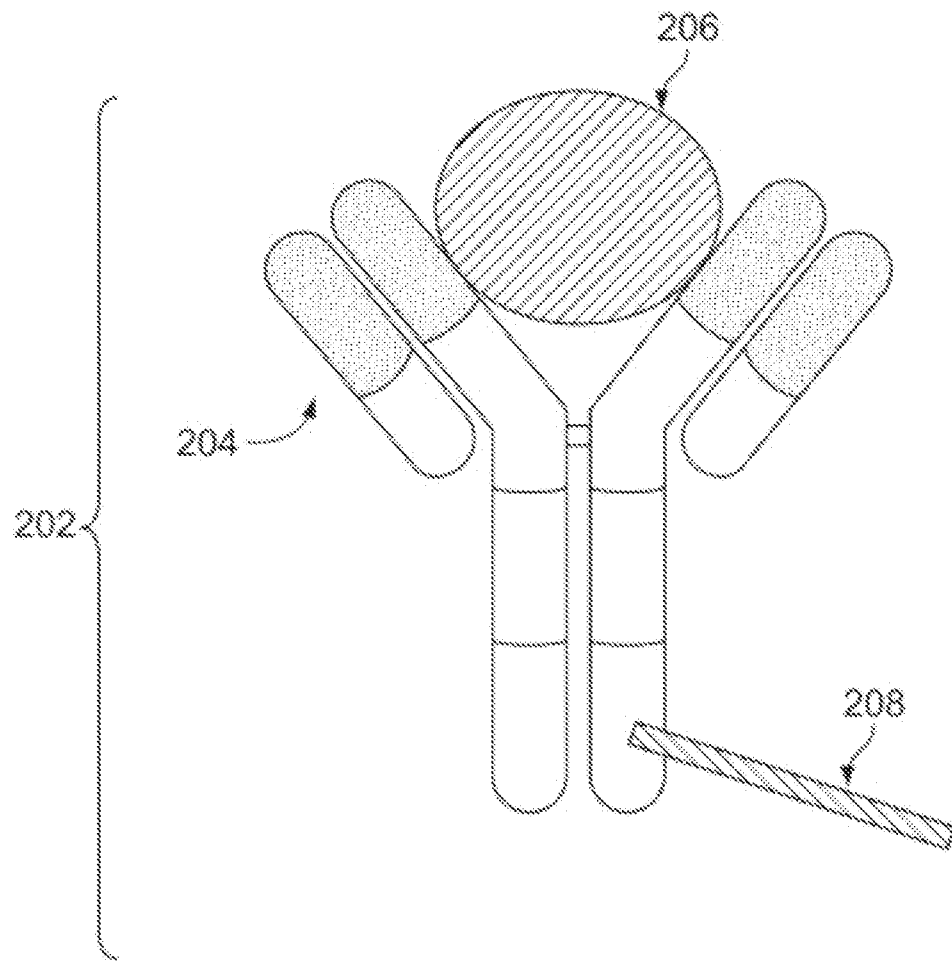


FIGURA 2

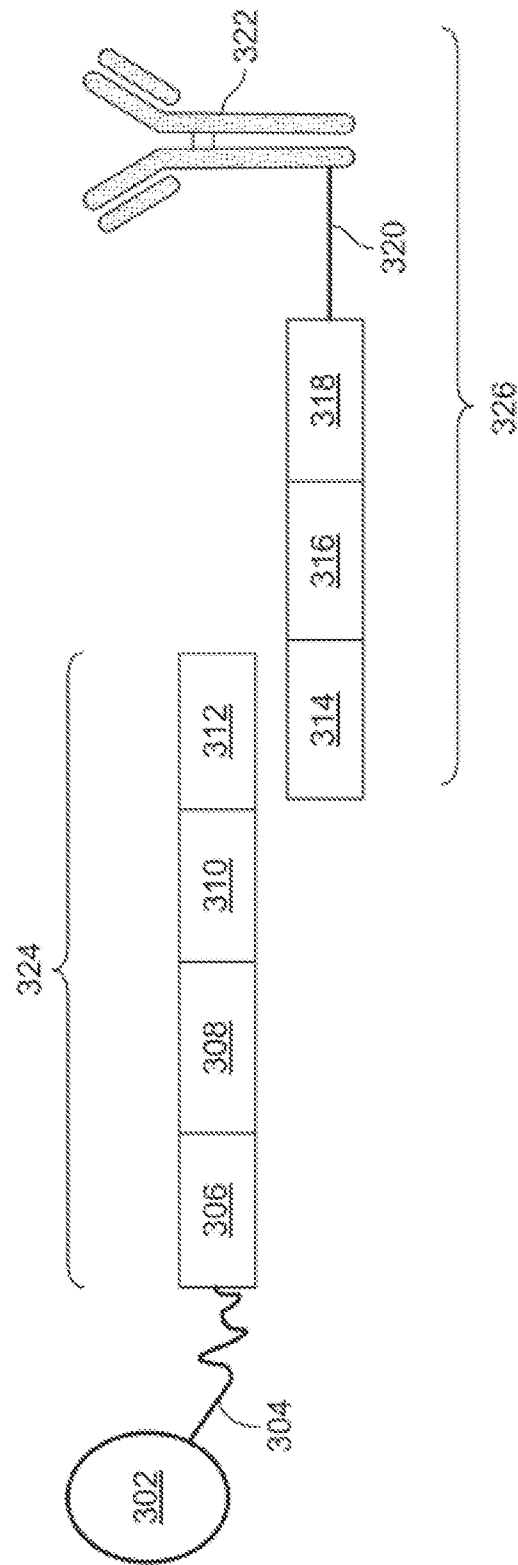


FIGURA 3

PROTEÍNA

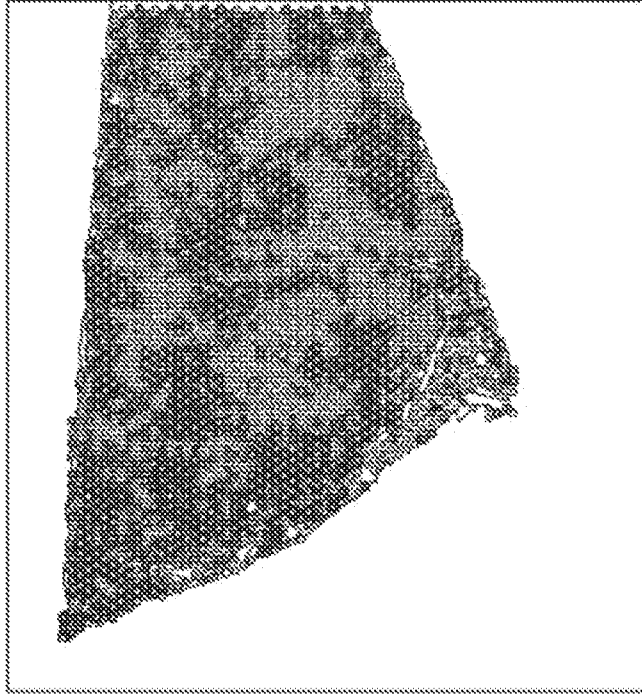


FIGURA 4B

ARN

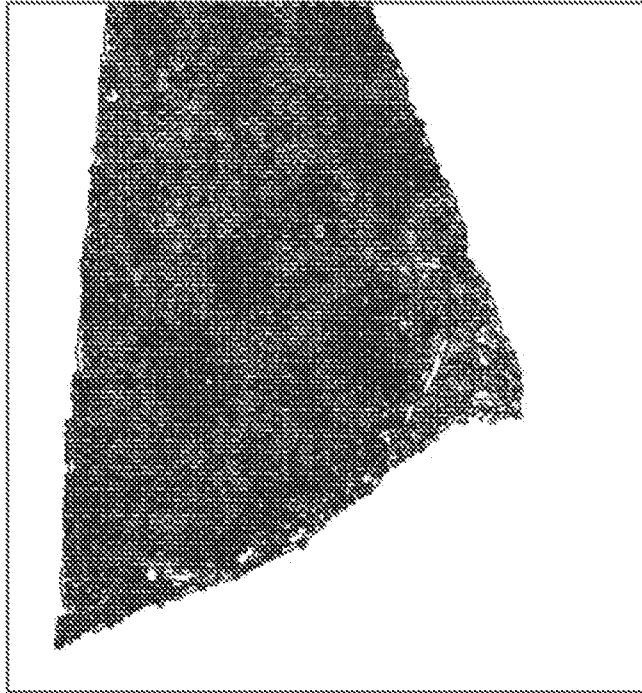


FIGURA 4A

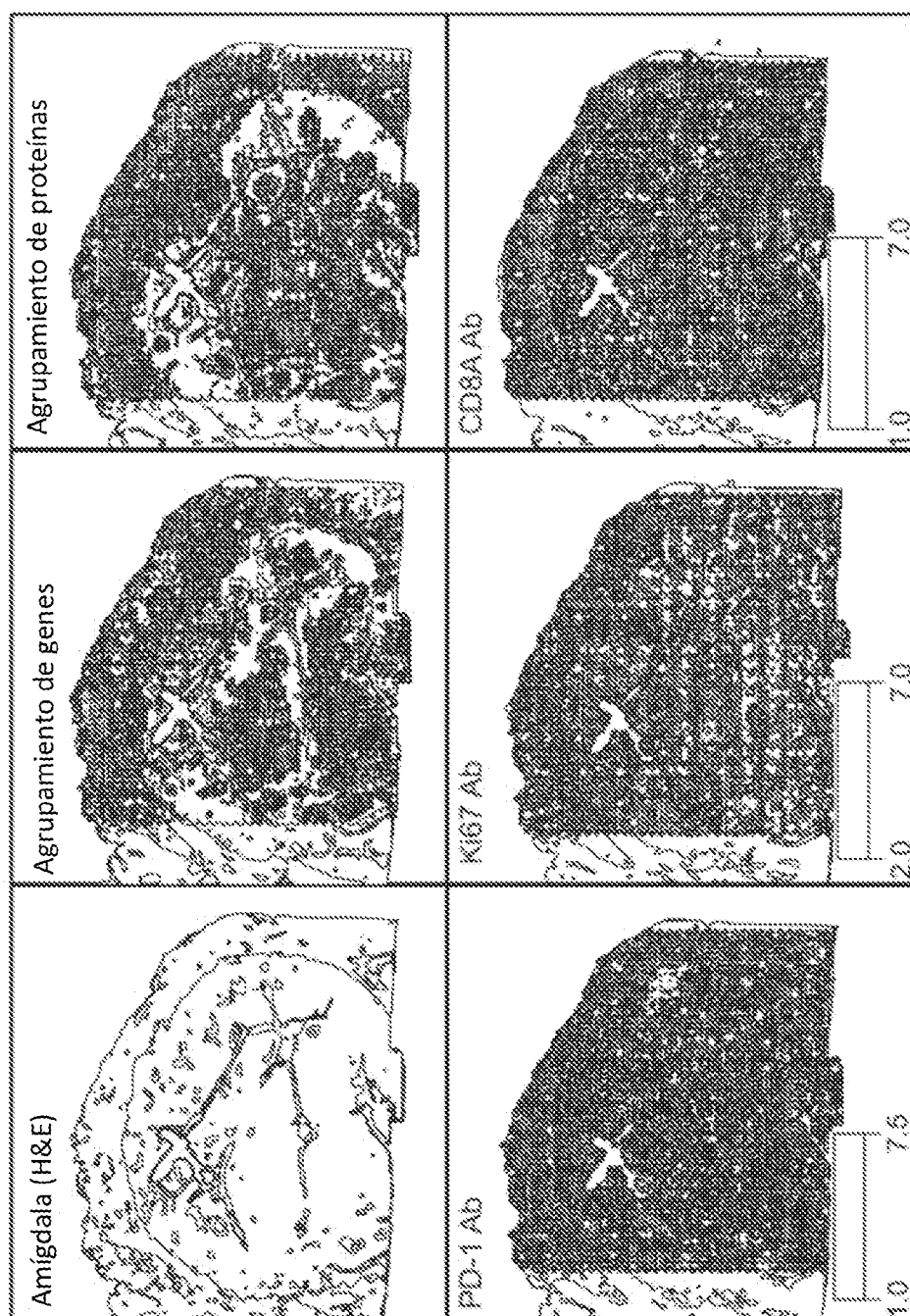


FIGURA 5



FIGURA 6

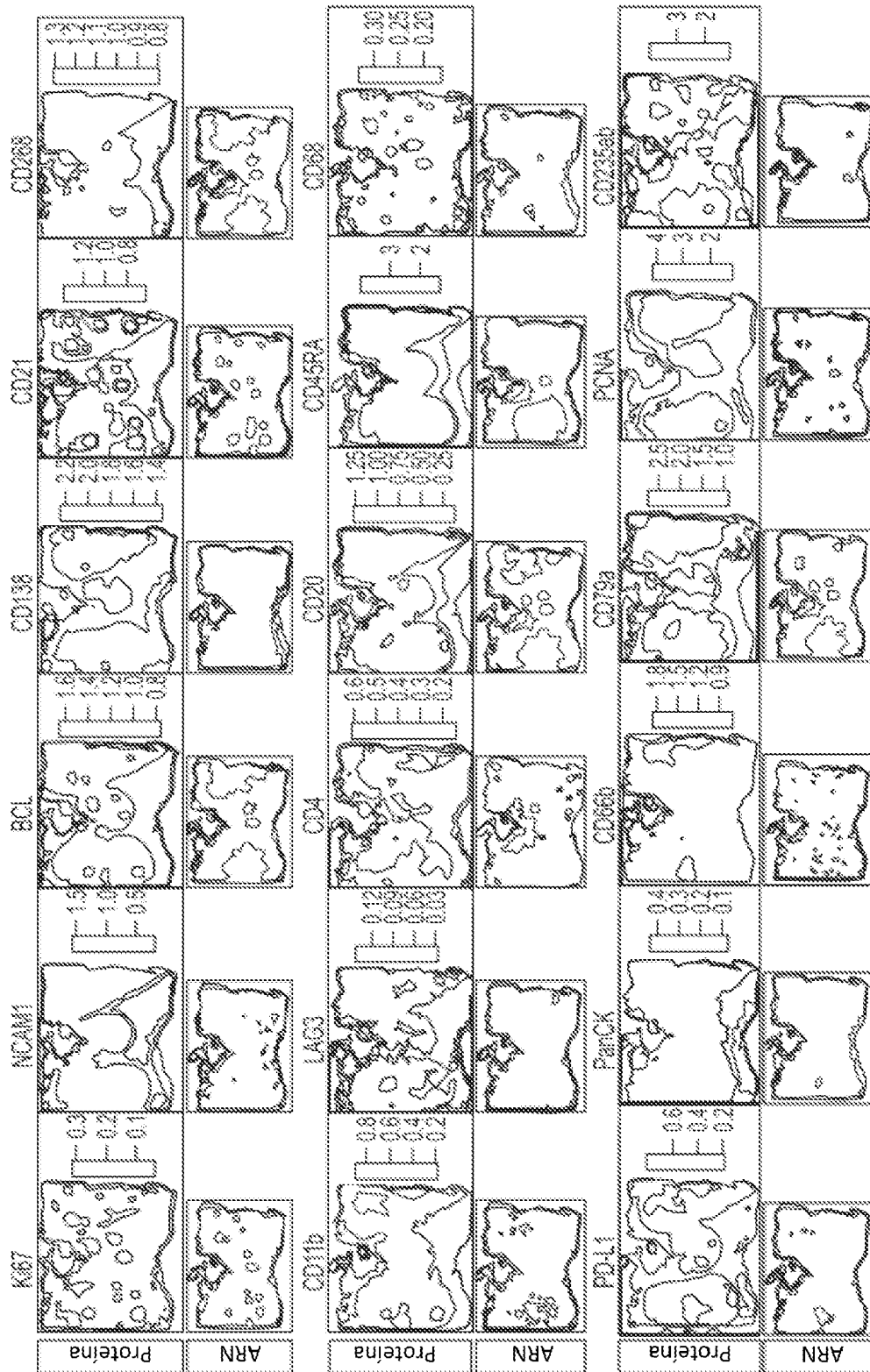


FIGURA 7

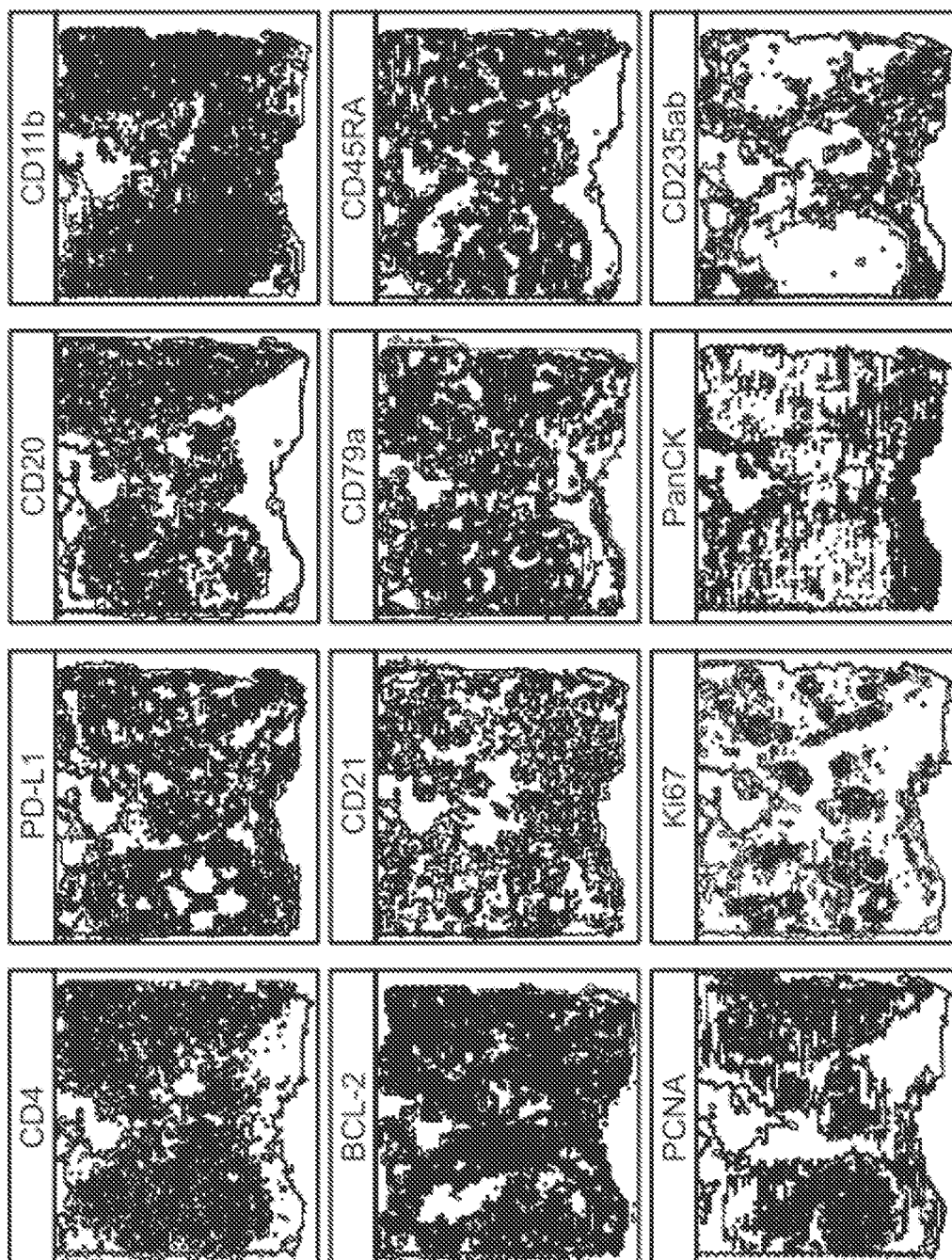


FIGURA 8

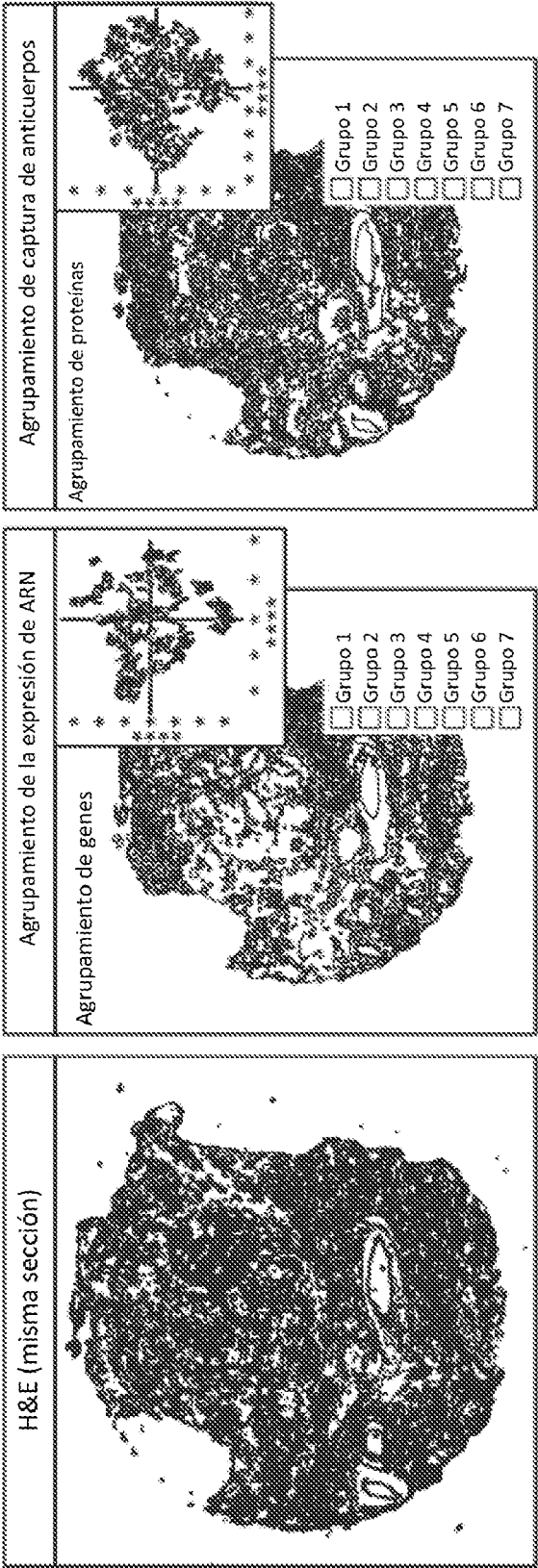


FIGURA 9A

FIGURA 9B

FIGURA 9C

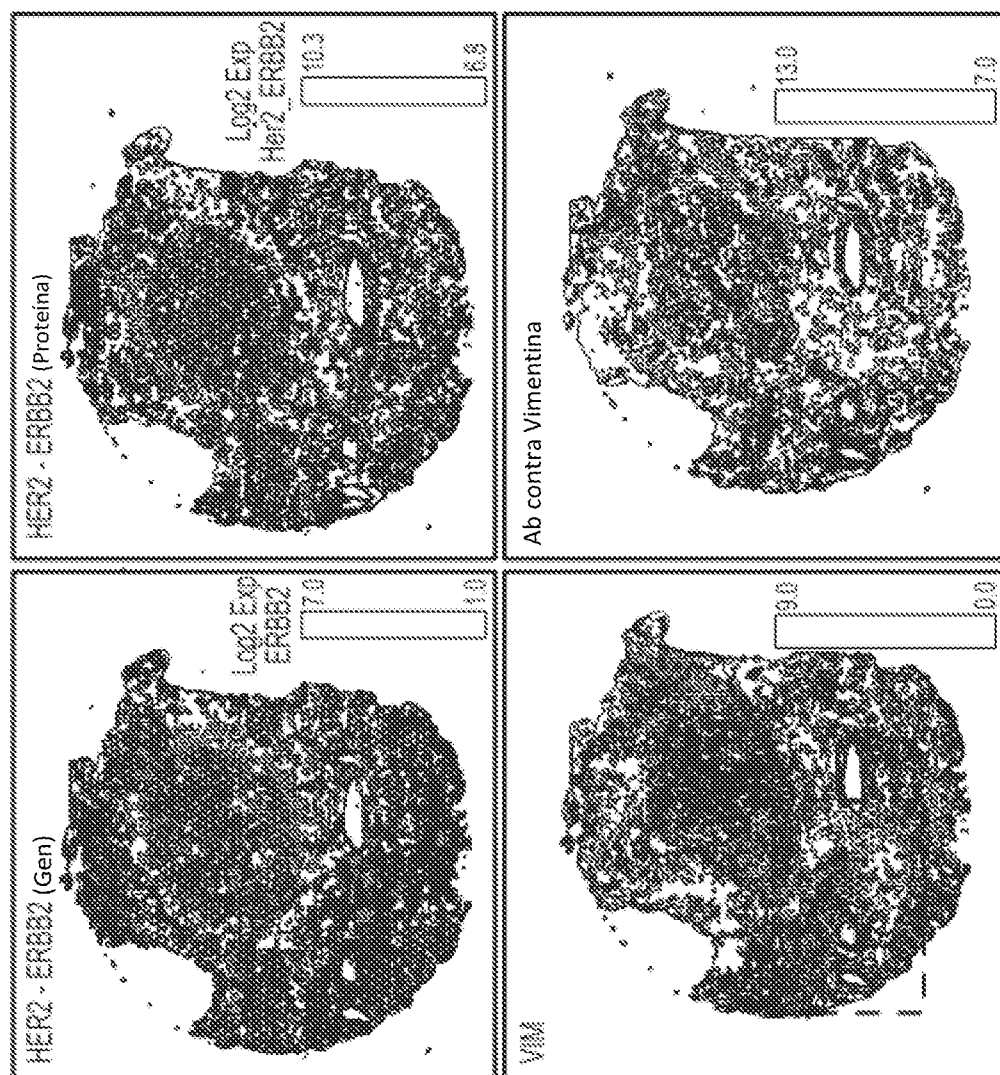
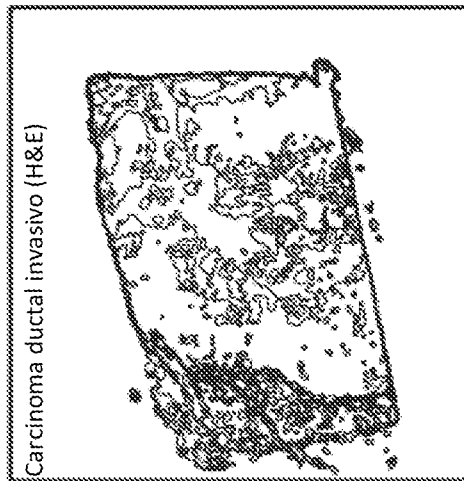
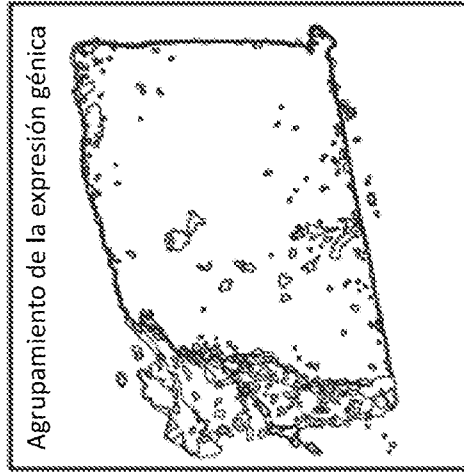


FIGURA 10



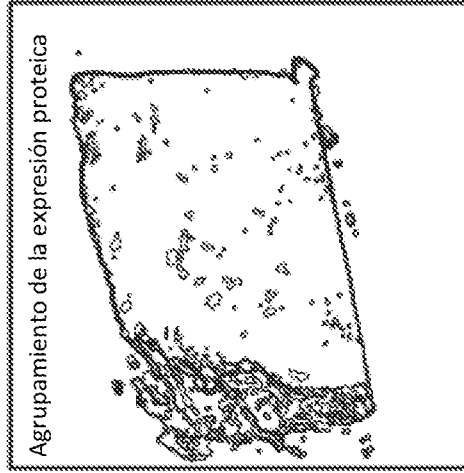
Carcinoma ductal invasivo (H&E)

FIGURA 11A



Agrupamiento de la expresión génica

FIGURA 11B



Agrupamiento de la expresión proteica

FIGURA 11C

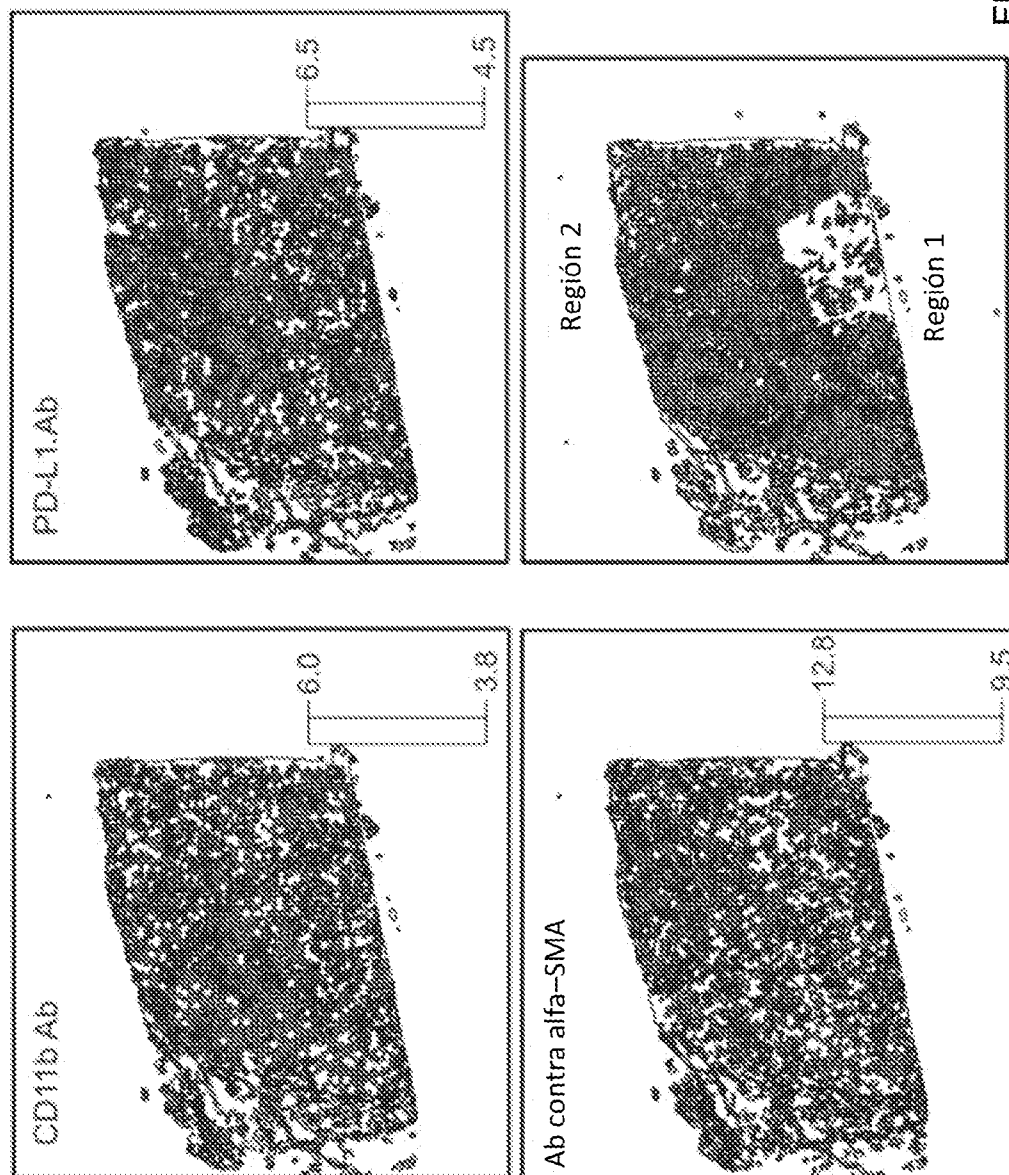


FIGURA 12

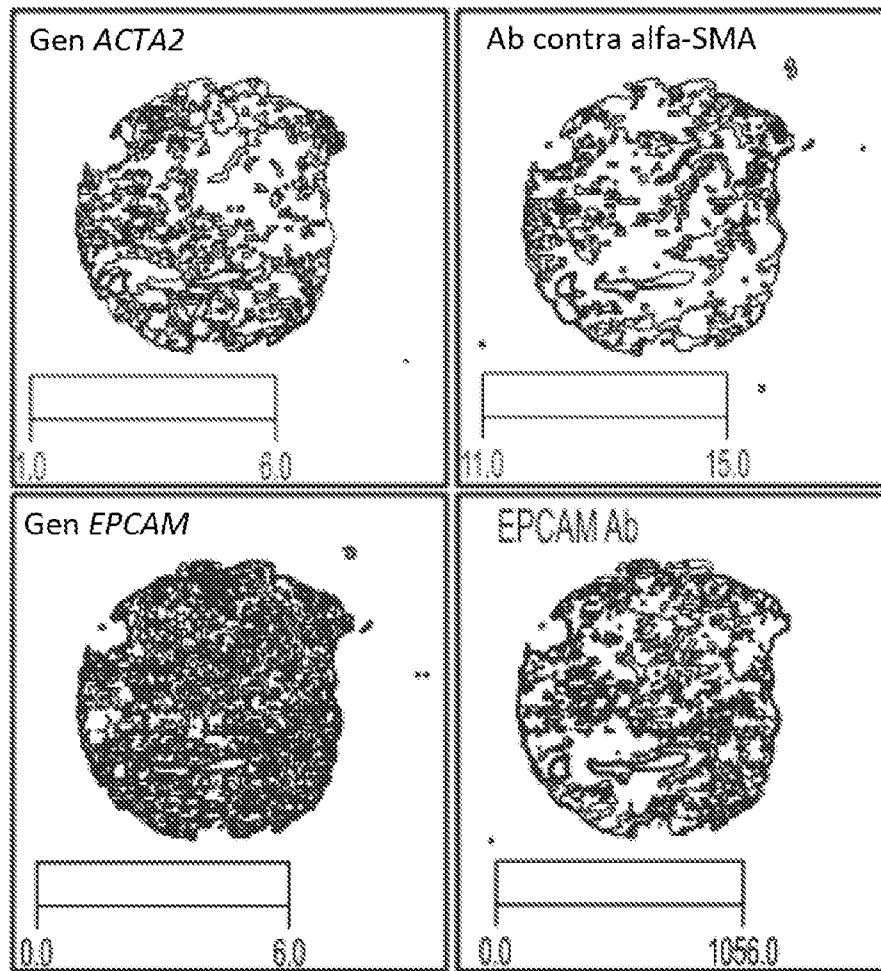


FIGURA 13

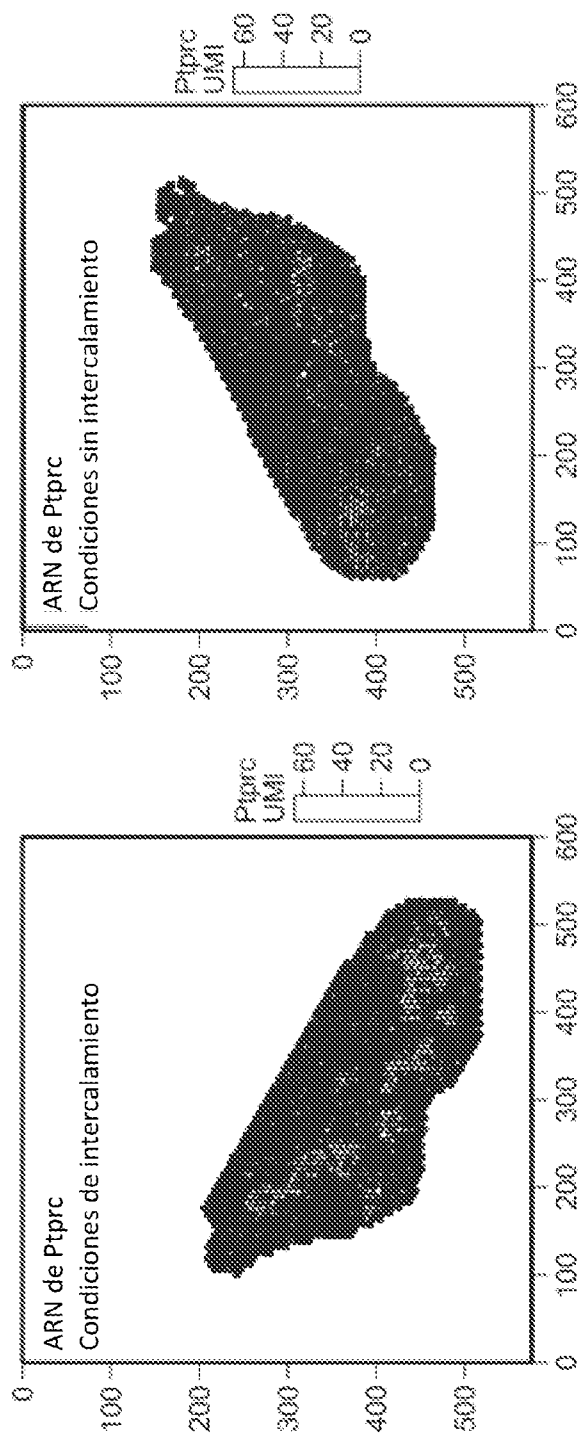


FIGURA 14A

FIGURA 14B

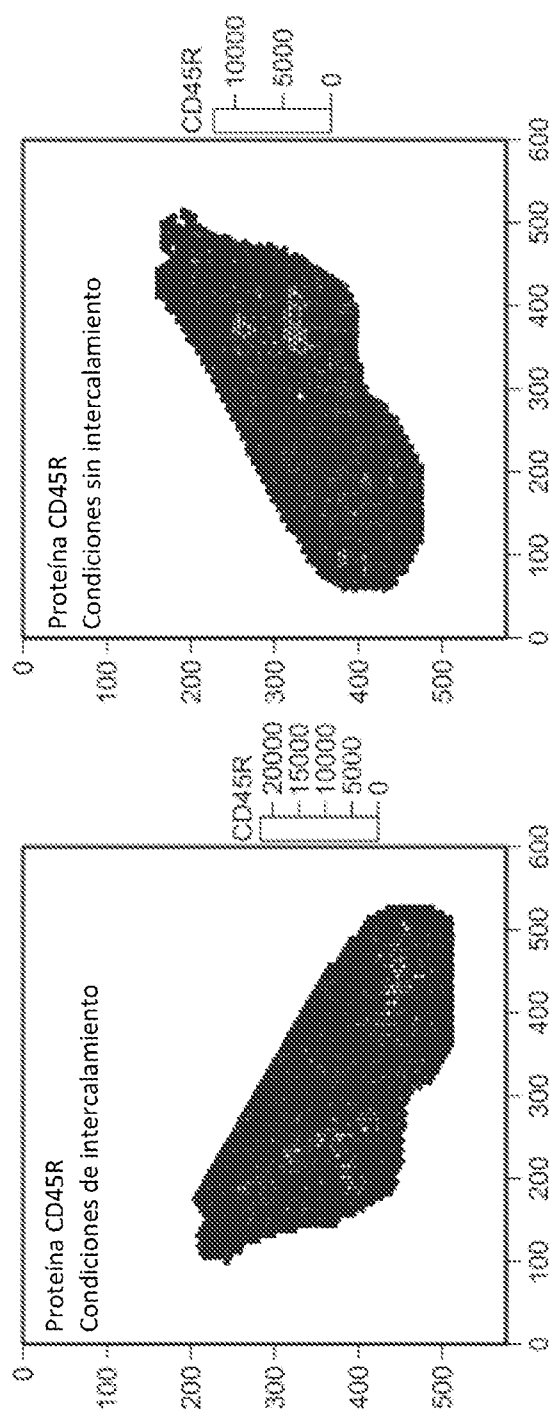


FIGURA 14C

FIGURA 14D

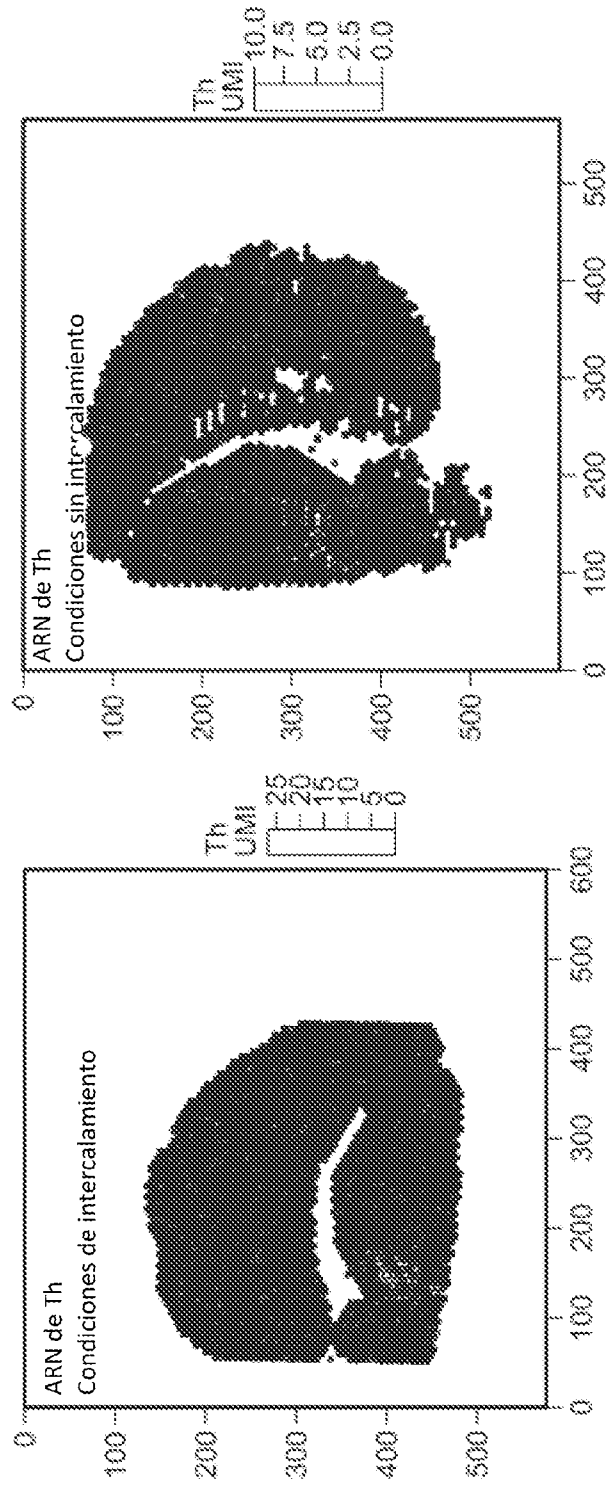


FIGURA 15A

FIGURA 15B

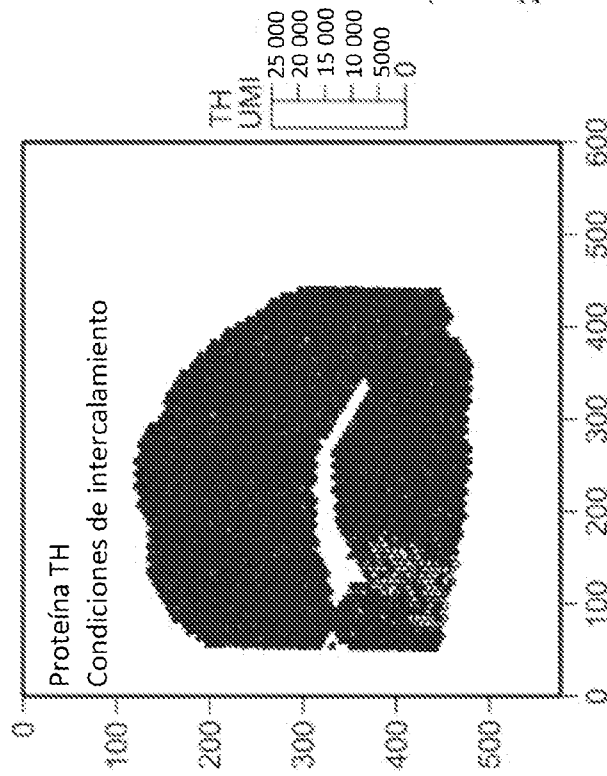


FIGURA 15C

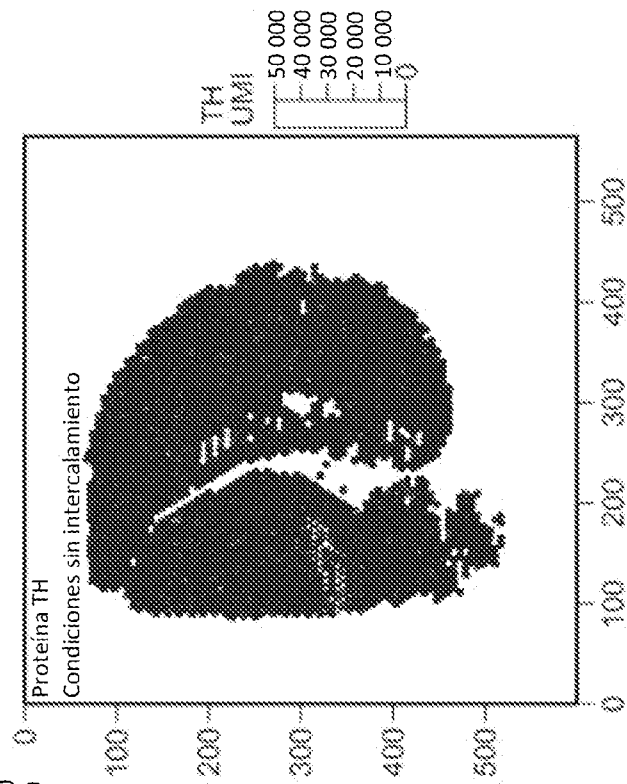


FIGURA 15D

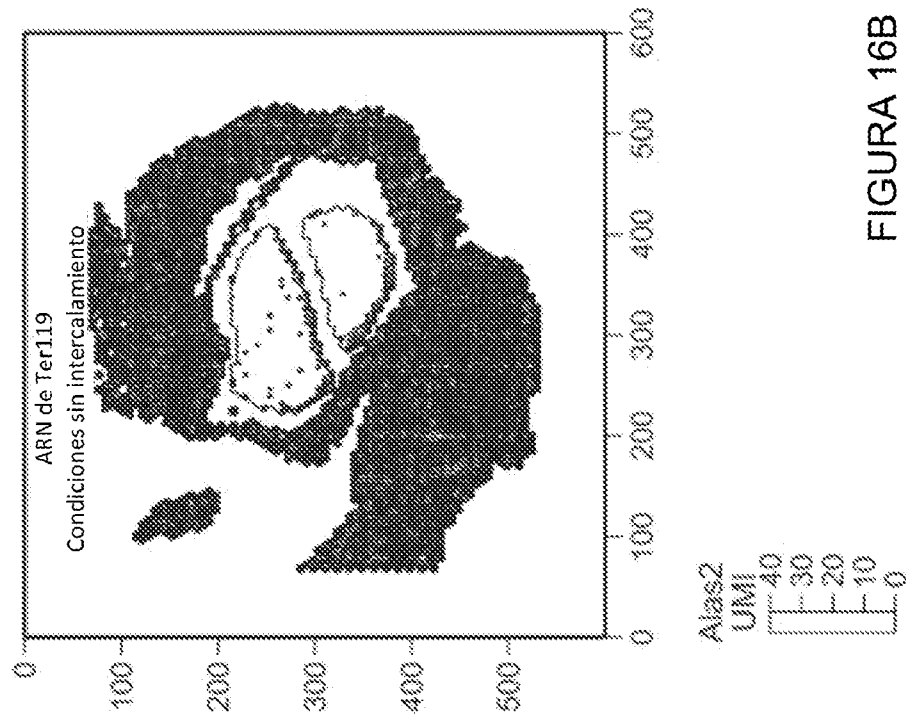


FIGURA 16A

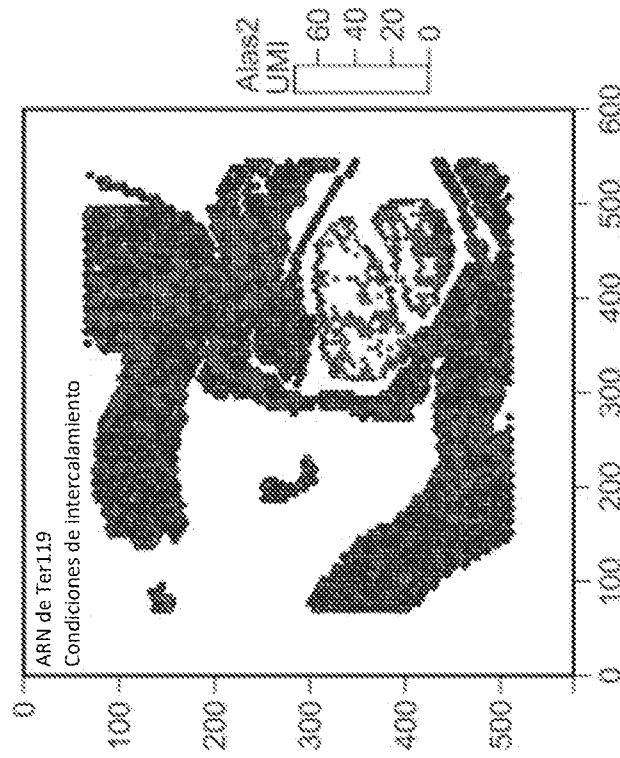


FIGURA 16B

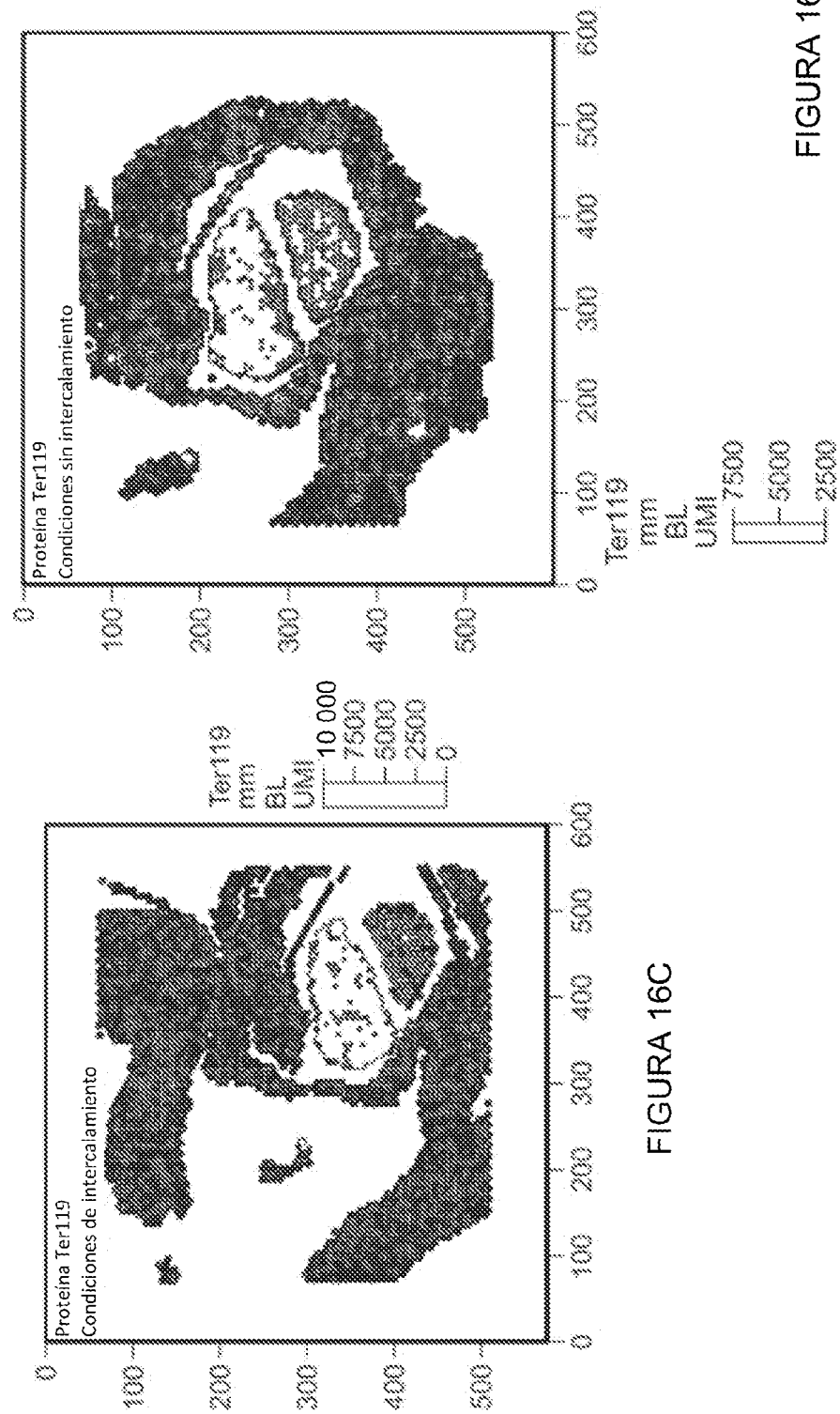


FIGURA 16C

FIGURA 16D

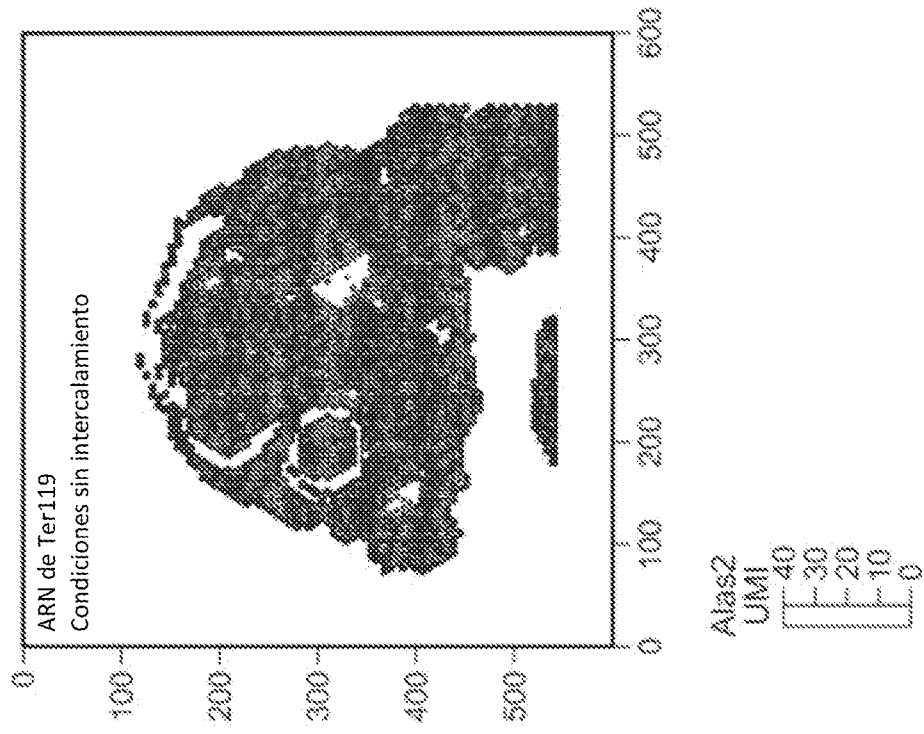


FIGURA 17A

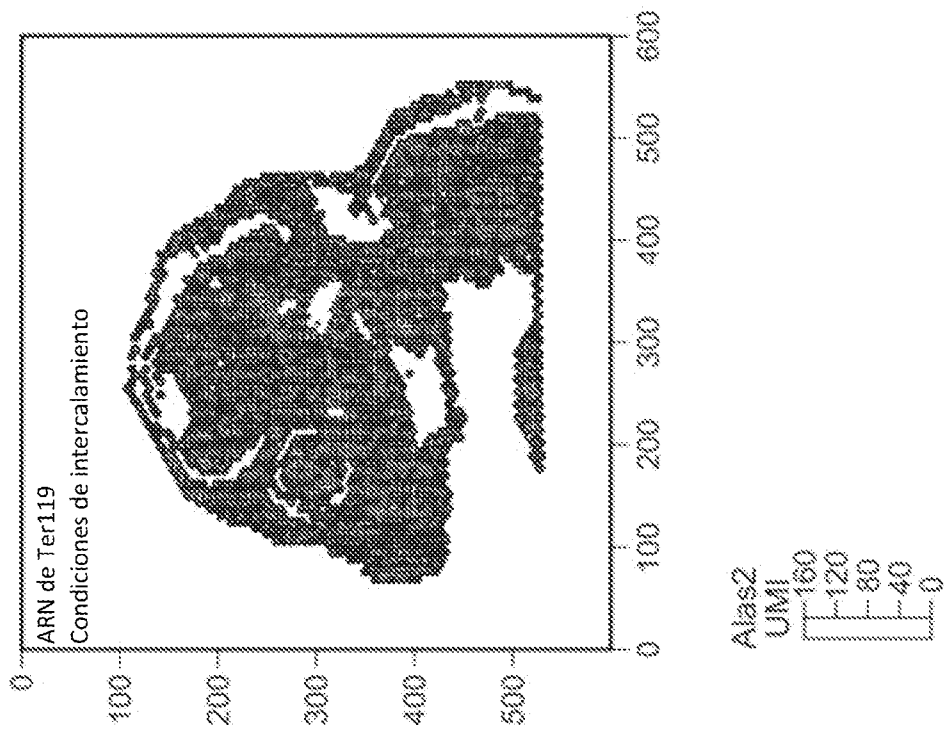


FIGURA 17B

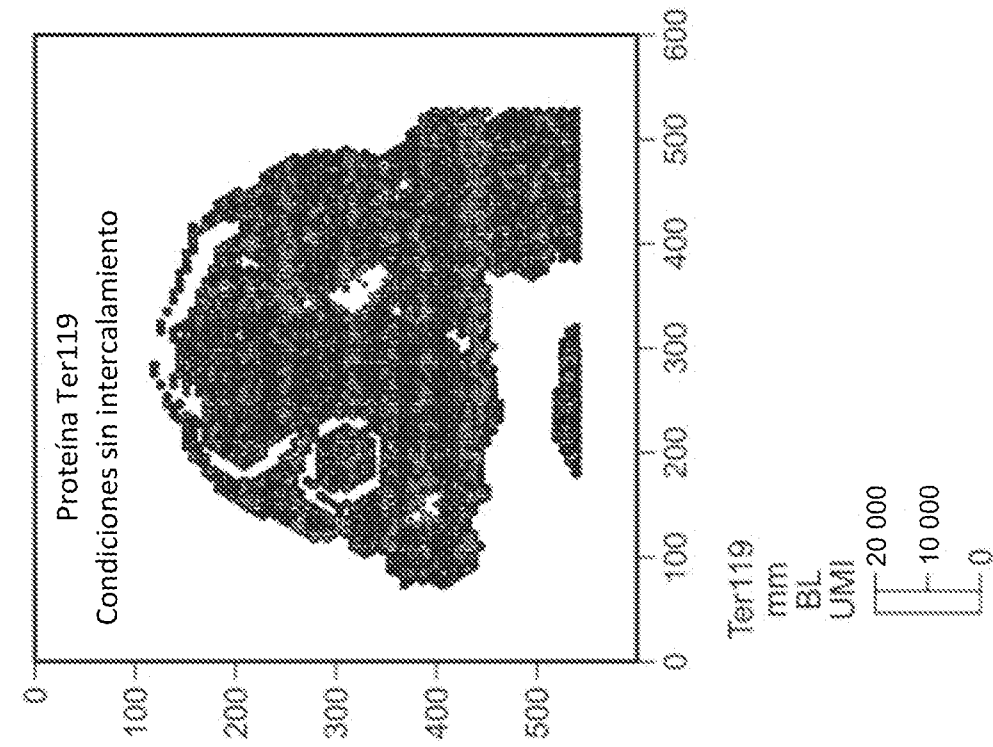


FIGURA 17D

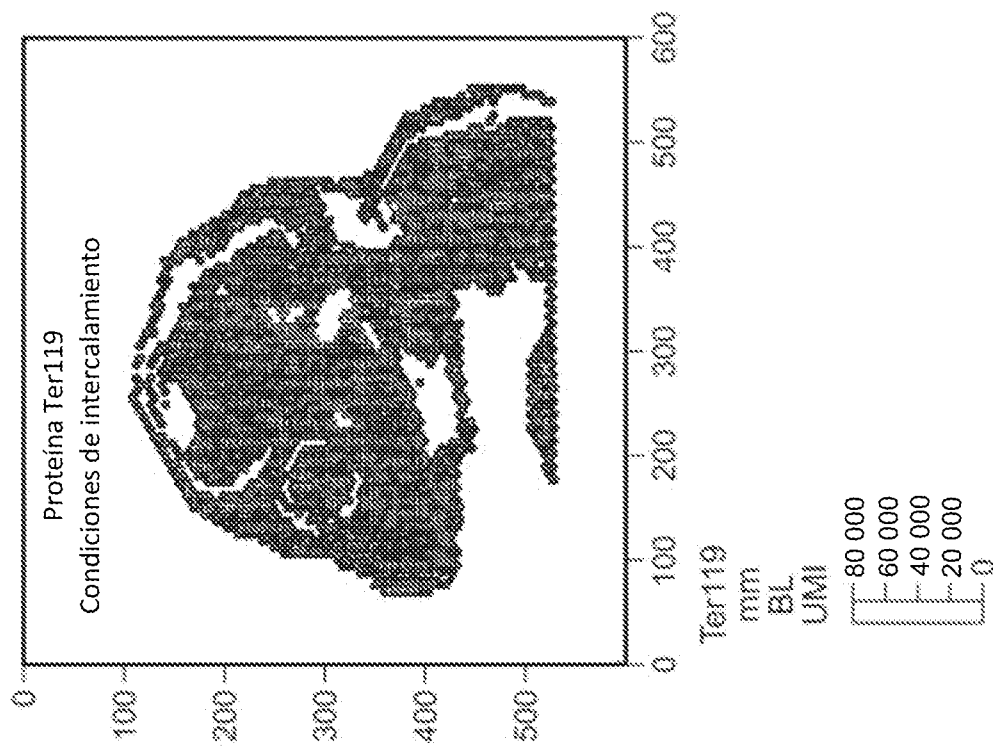
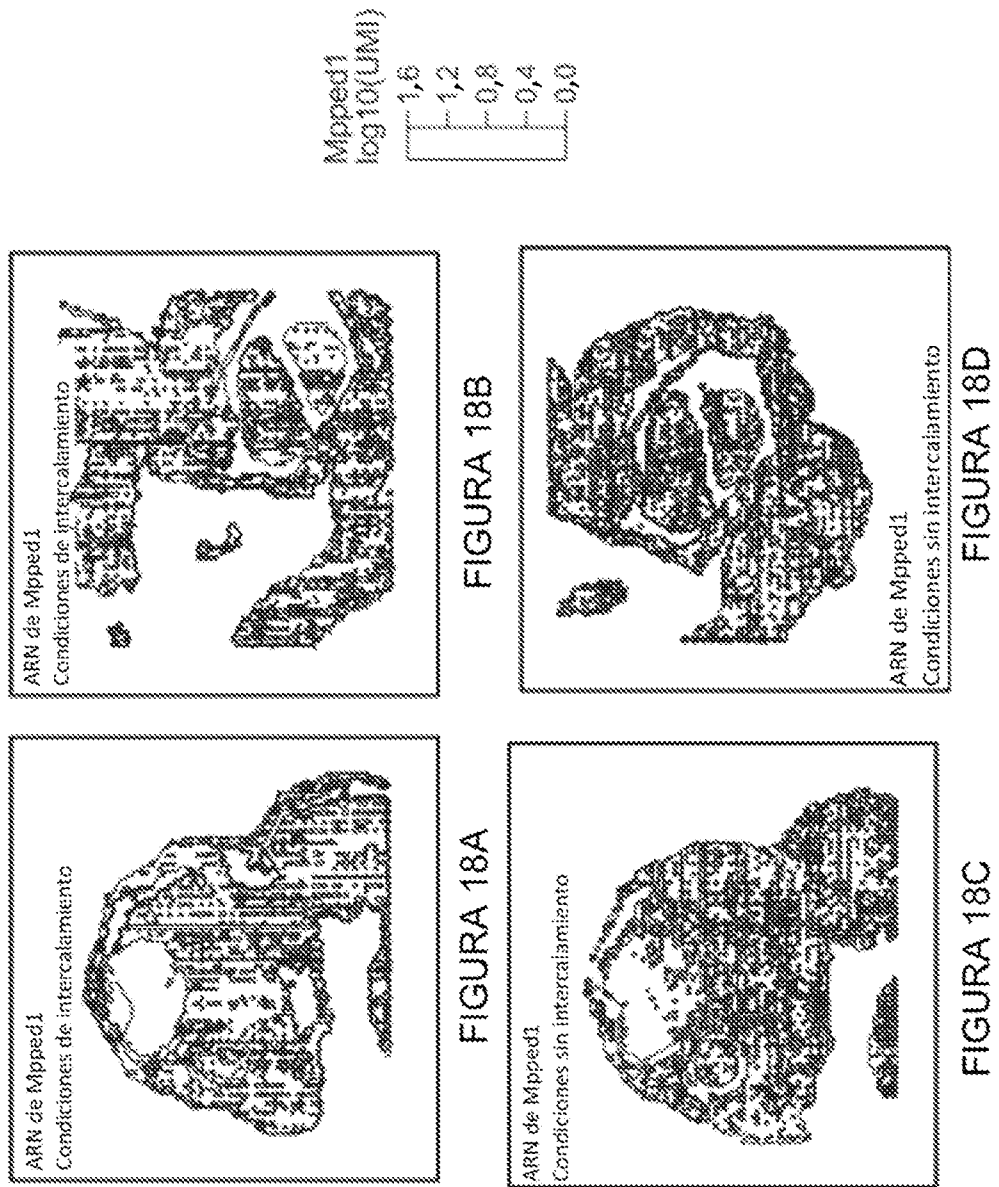


FIGURA 17C



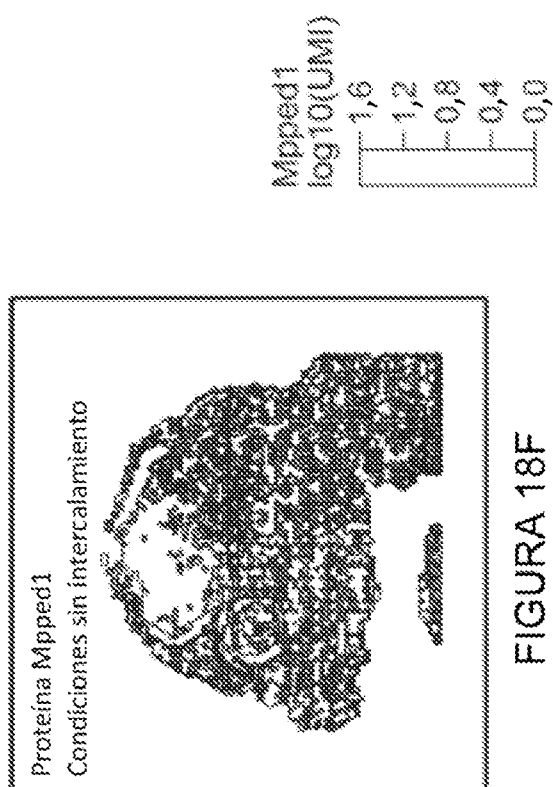


FIGURA 18F

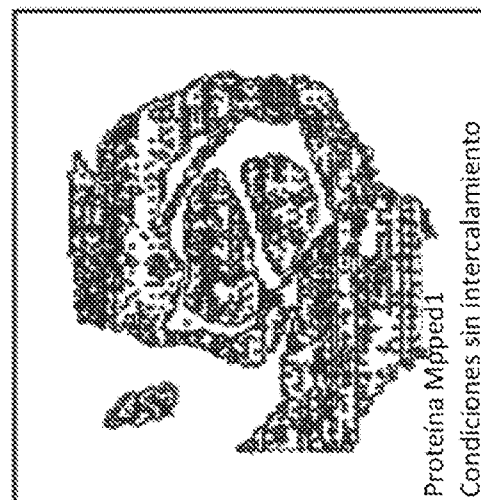


FIGURA 18G

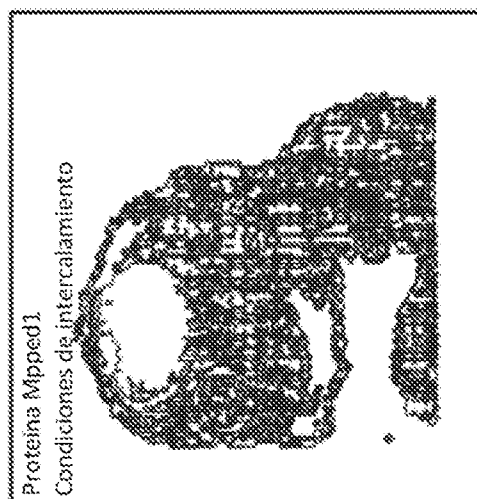


FIGURA 18E

ARNm de Mpped1 Condiciones sin
intercalamiento (detección de ARN
solamente)

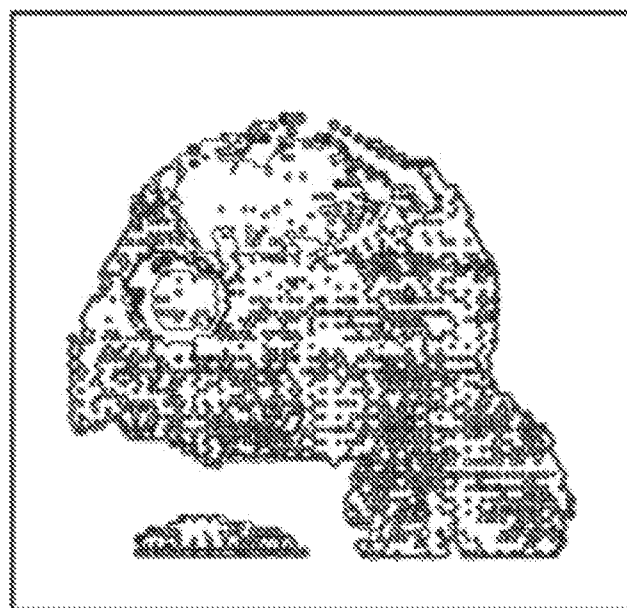


FIGURA 18H

ARNm de Mpped1 Condiciones sin
intercalamiento (detección de ARN
solamente)

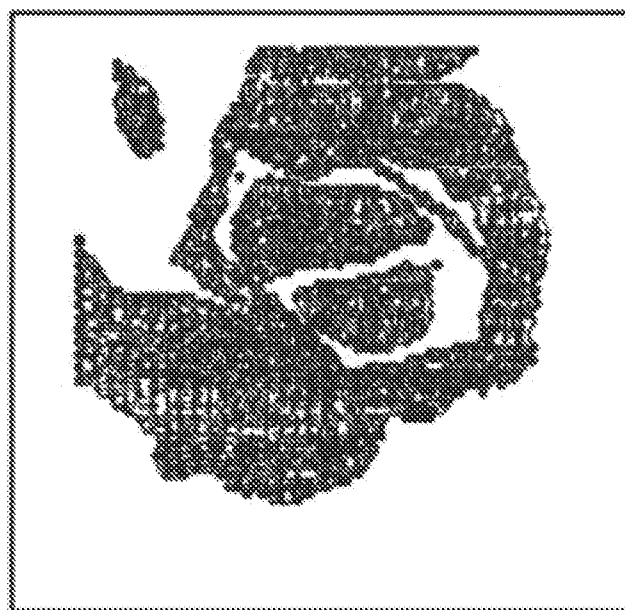
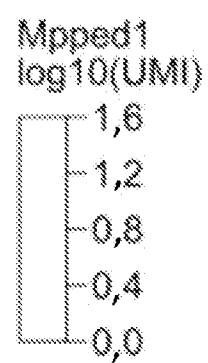
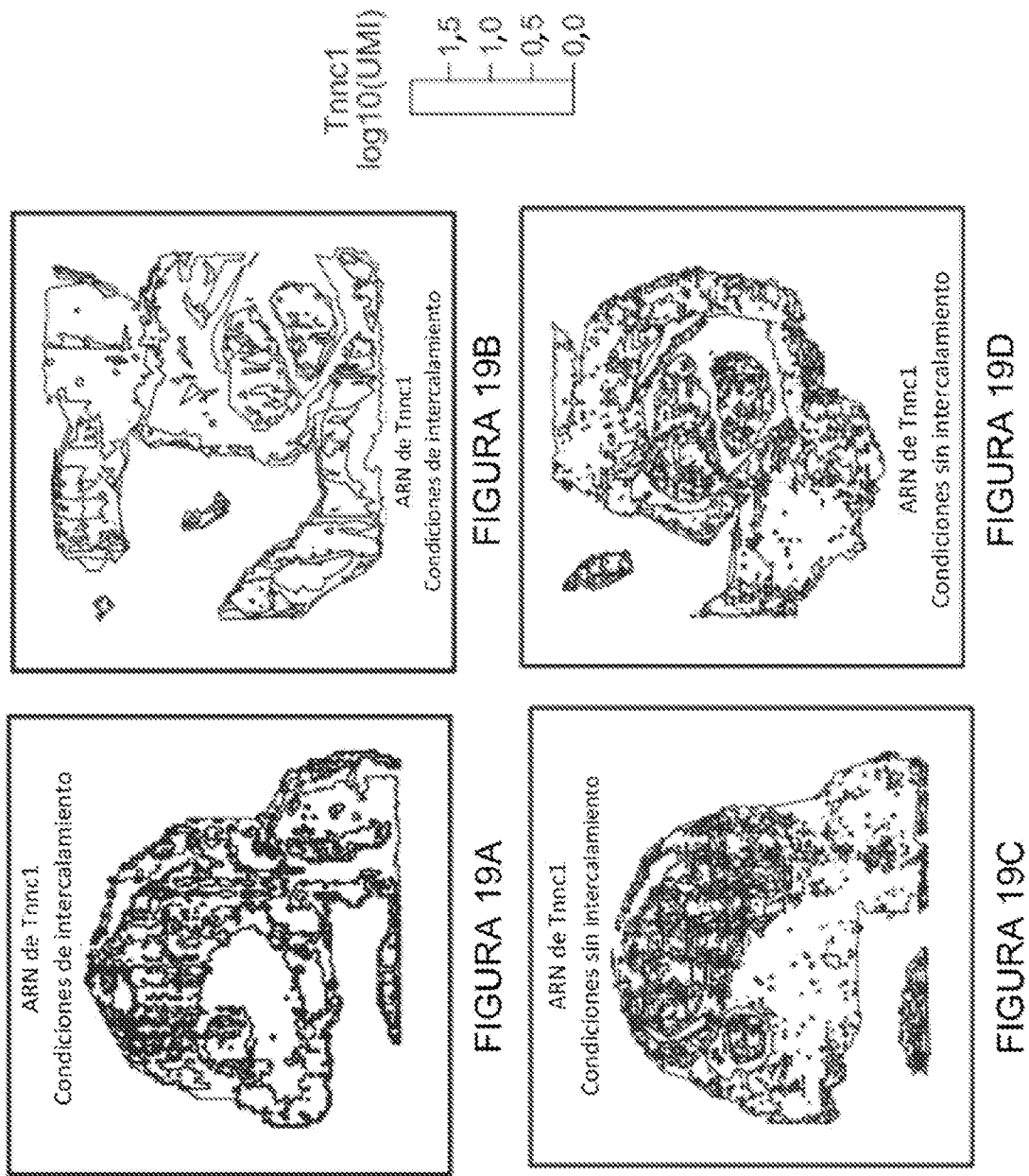
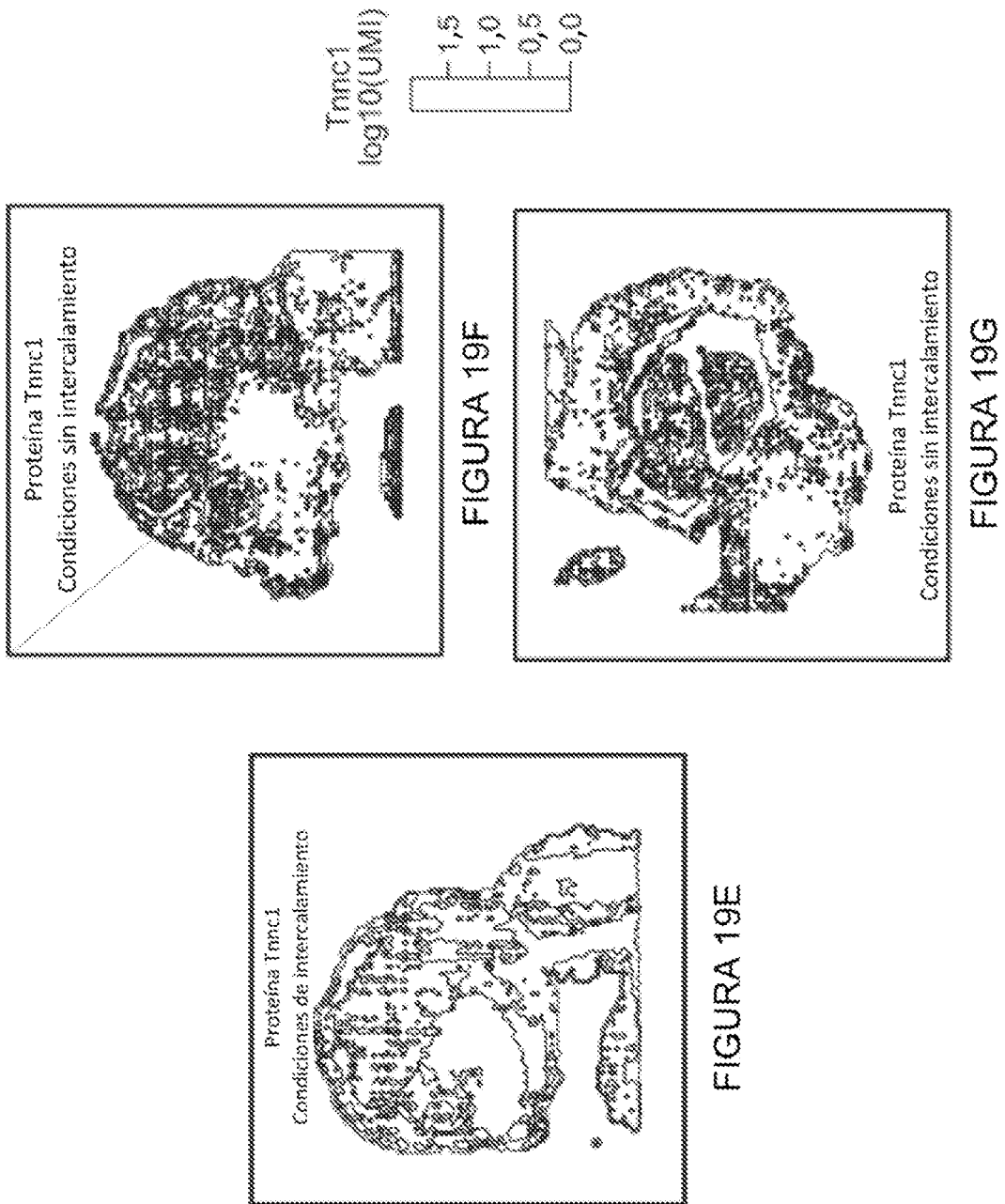


FIGURA 18I







(Detección de ARN solamente) ARN de
Tnnc1 Condiciones sin intercalamiento

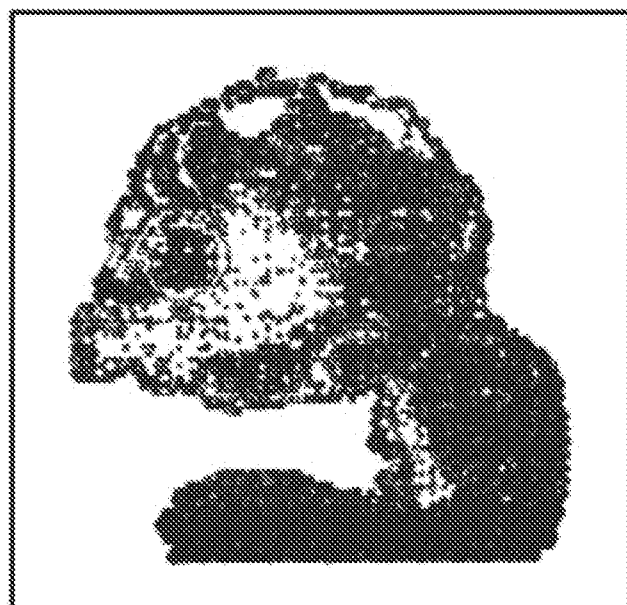
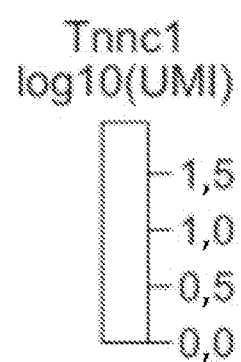


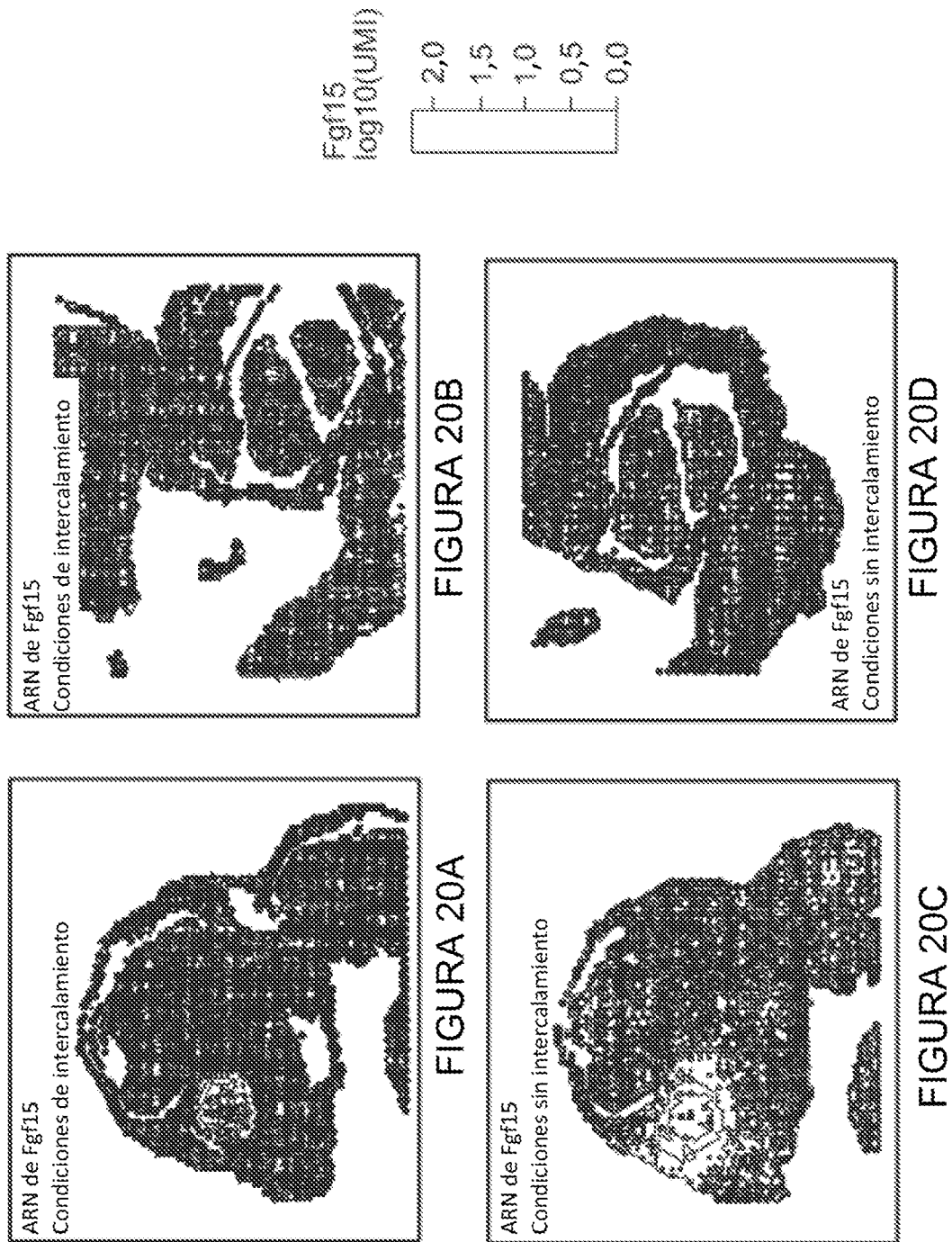
FIGURA 19H

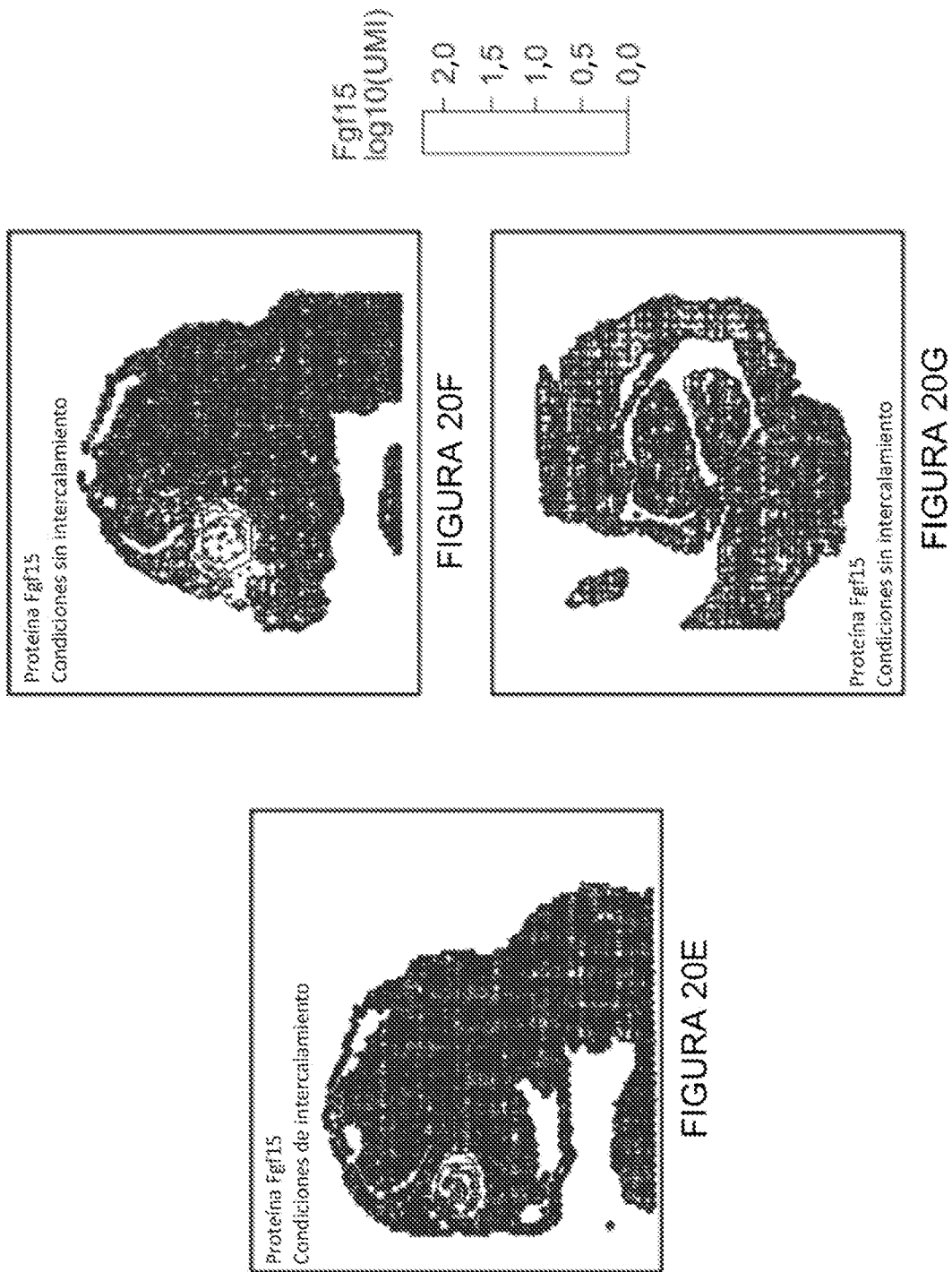


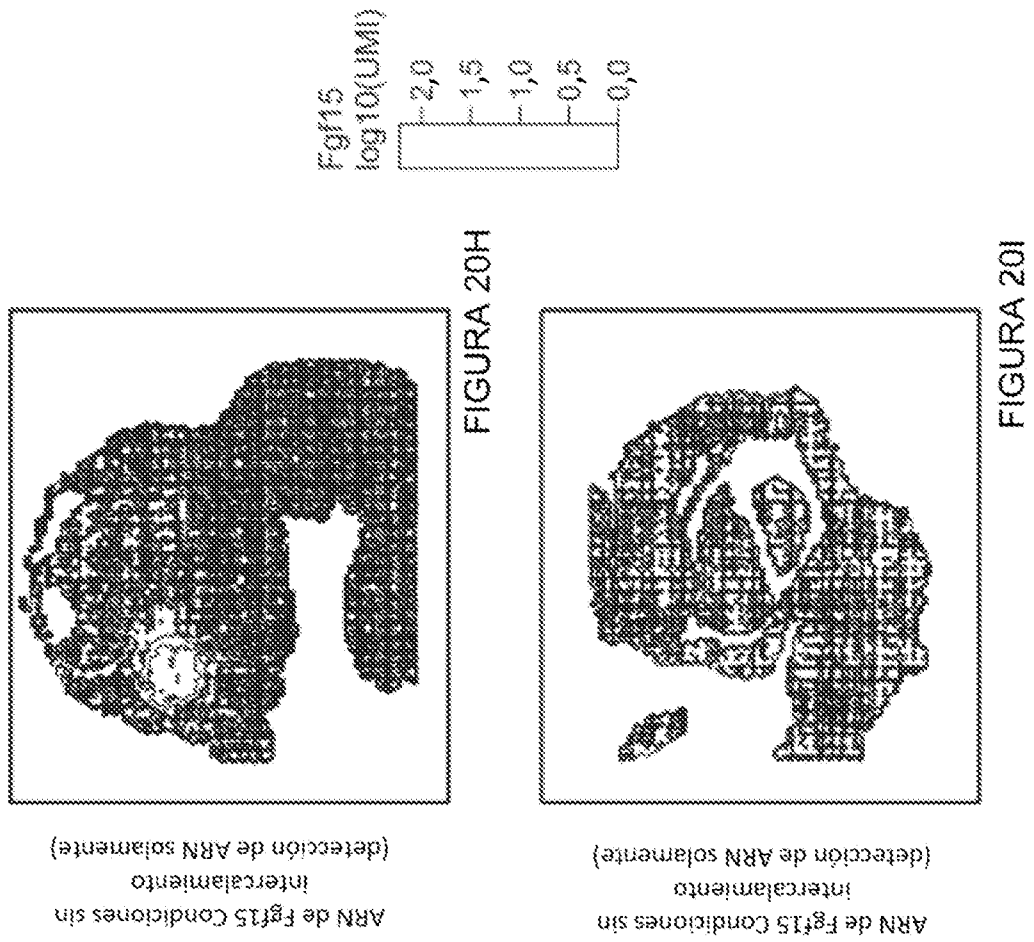
(Detección de ARN solamente) ARN de
Tnnc1 Condiciones sin intercalamiento

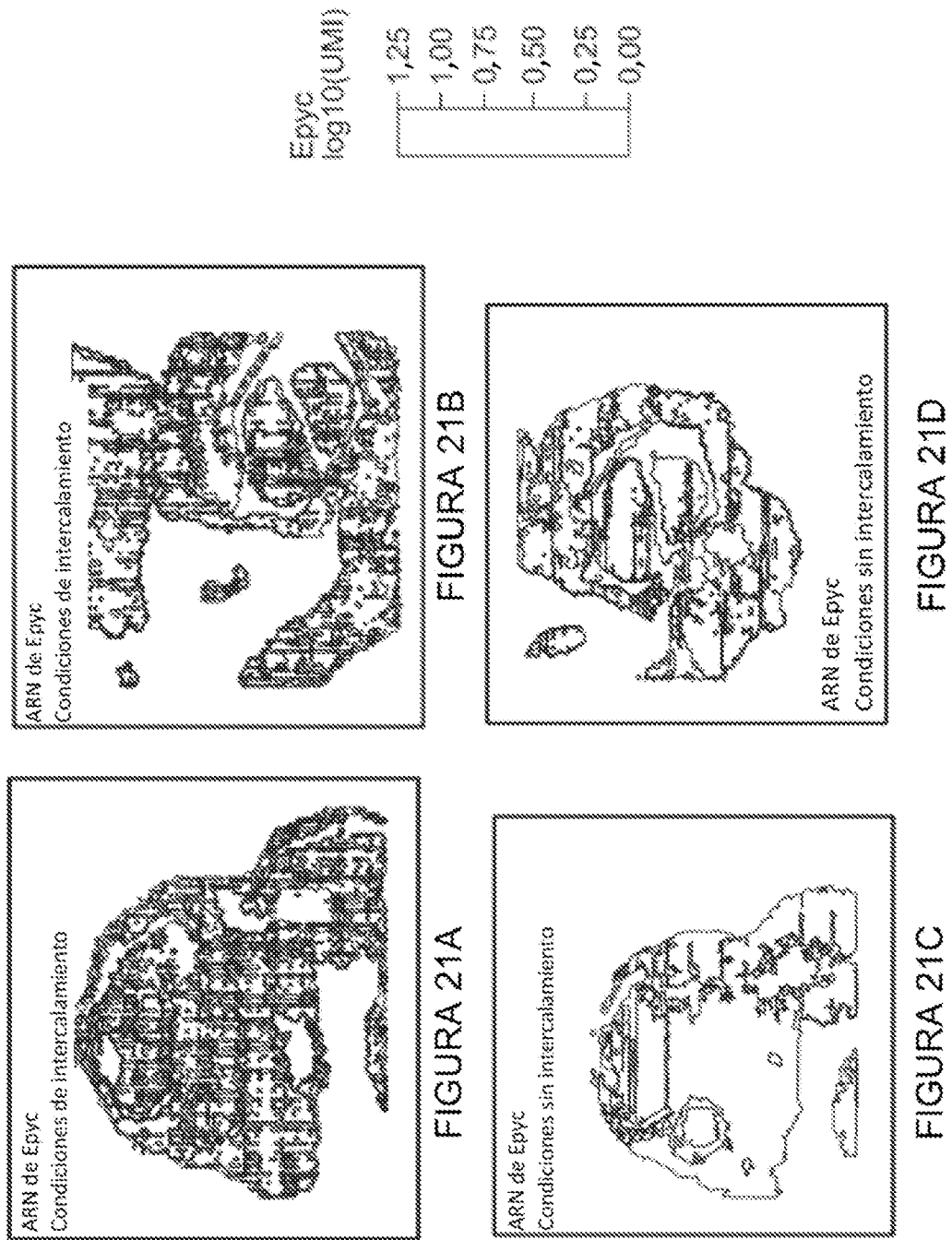


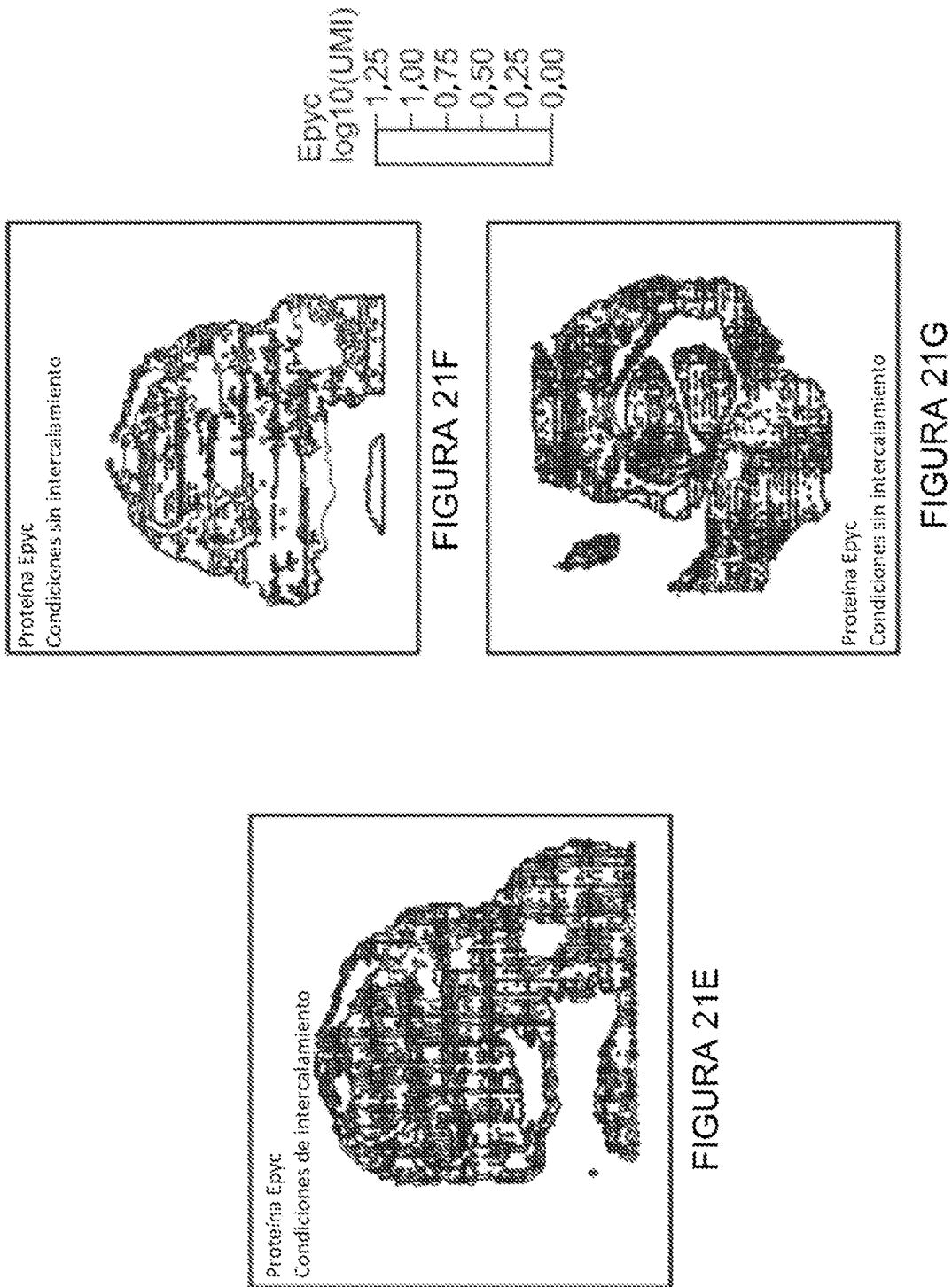
FIGURA 19I

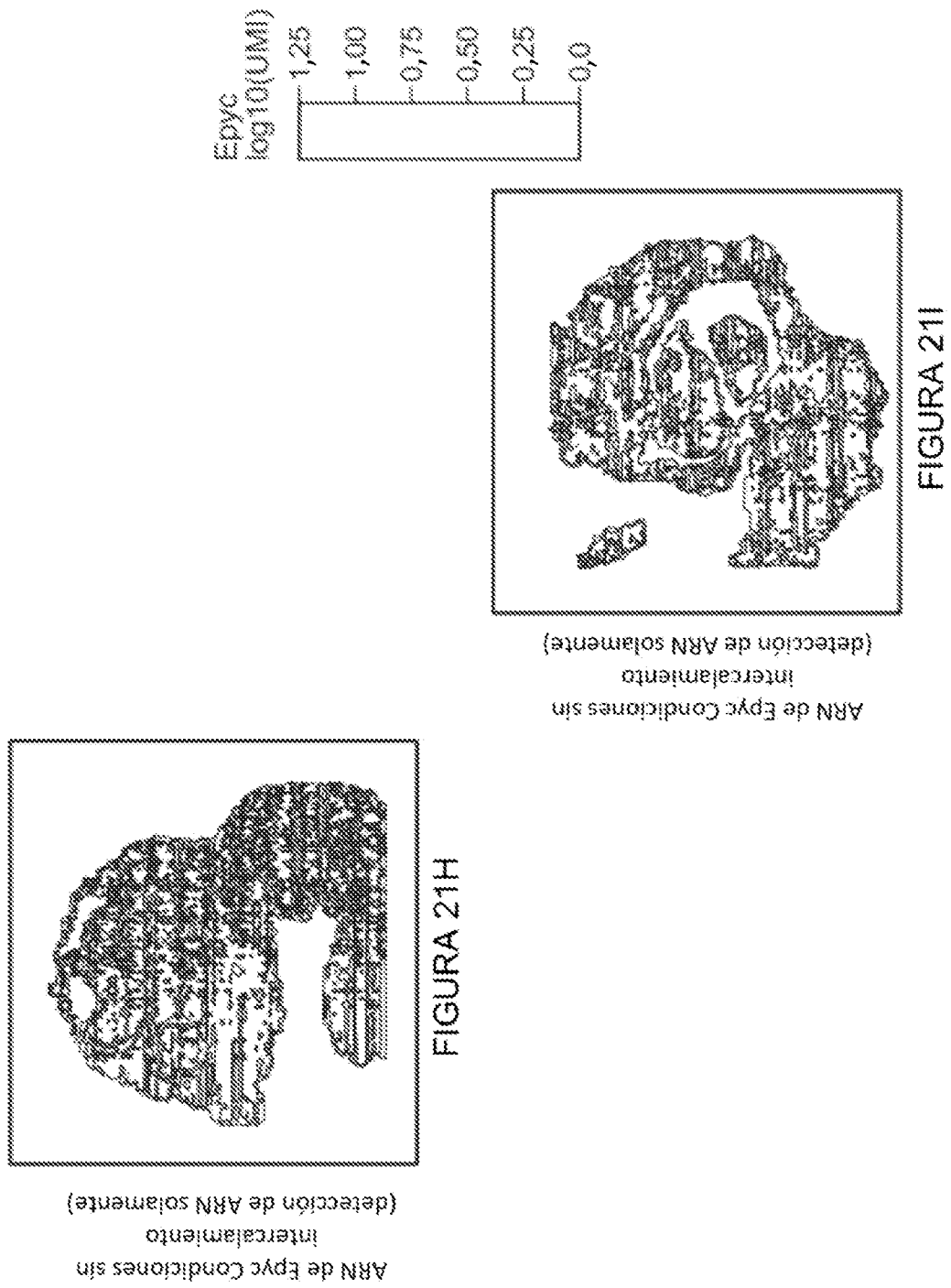












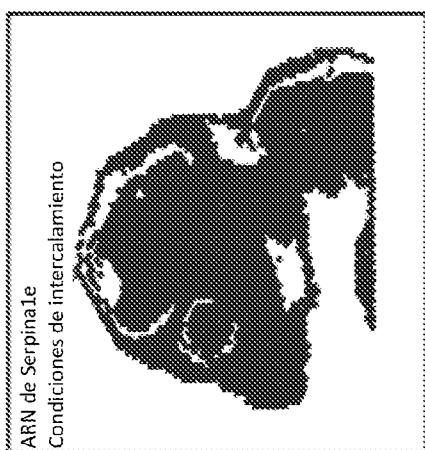


FIGURA 22A

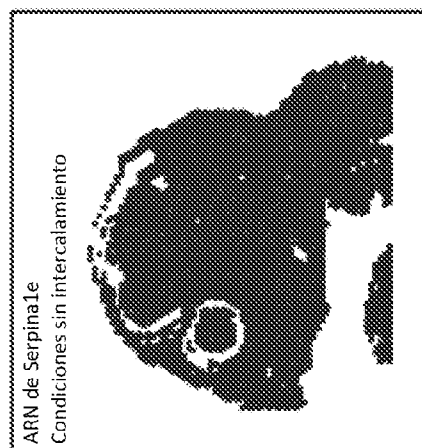


FIGURA 22C

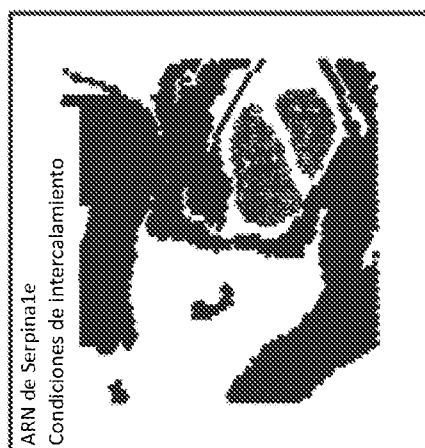


FIGURA 22B

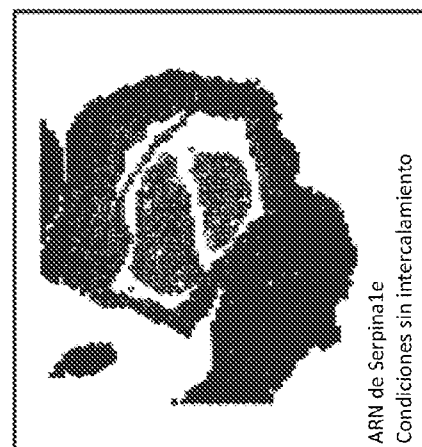


FIGURA 22D

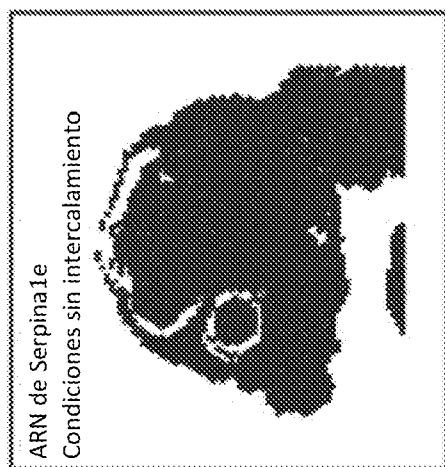


FIGURA 22F

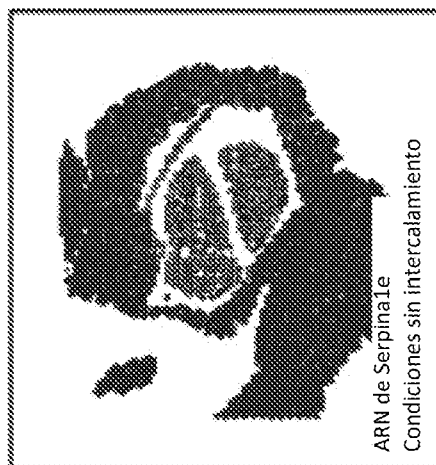


FIGURA 22G

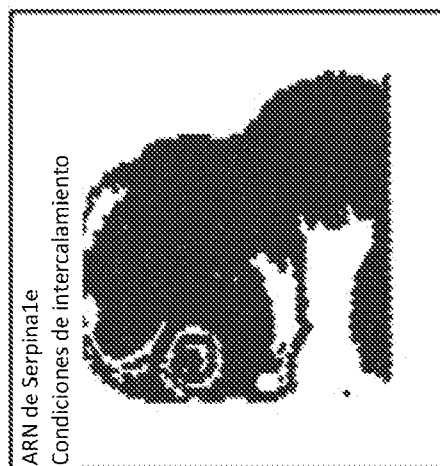


FIGURA 22E

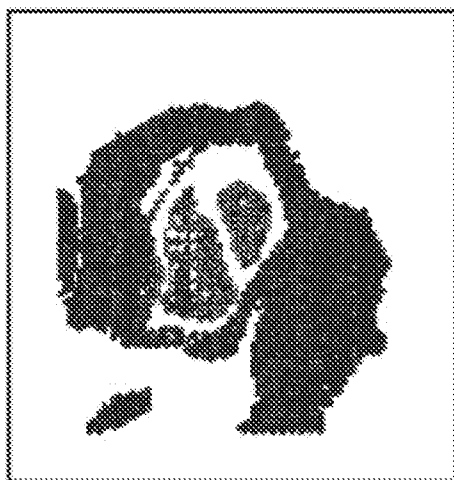


FIGURA 22I

ARN de SerpinA1 Condiciones sin
intercalamiento
(detección de ARN solamente)

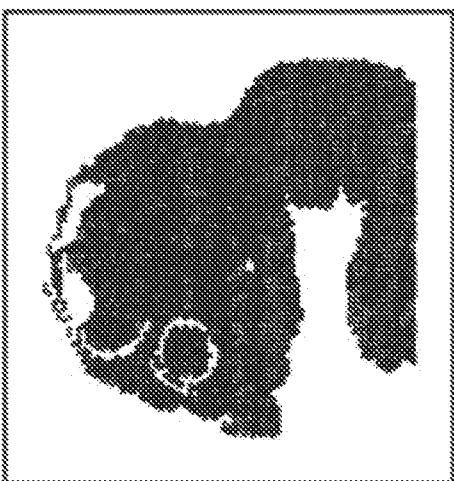


FIGURA 22H

ARN de SerpinA1 Condiciones sin
intercalamiento
(detección de ARN solamente)

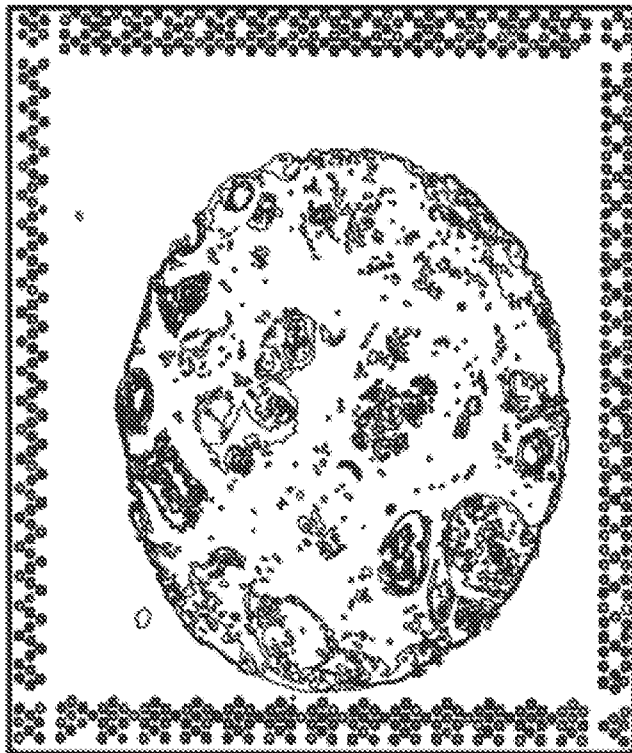


FIGURA 23A

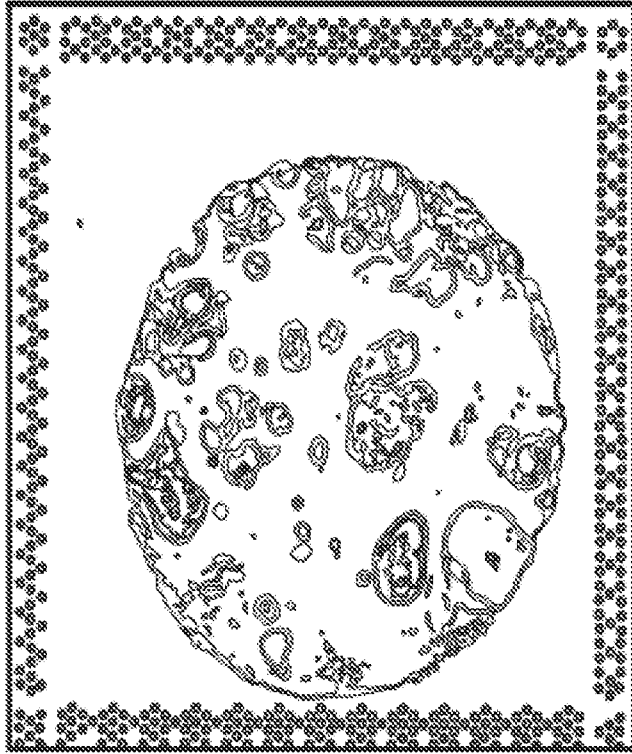


FIGURA 23B

Anotación del patólogo

- | | |
|---|---|
| <input type="checkbox"/> Vaso sanguíneo | <input type="checkbox"/> DCIS |
| <input type="checkbox"/> Células inmunitarias | <input type="checkbox"/> Carcinoma invasivo |
| <input type="checkbox"/> Necrosis | <input type="checkbox"/> Glándula normal |

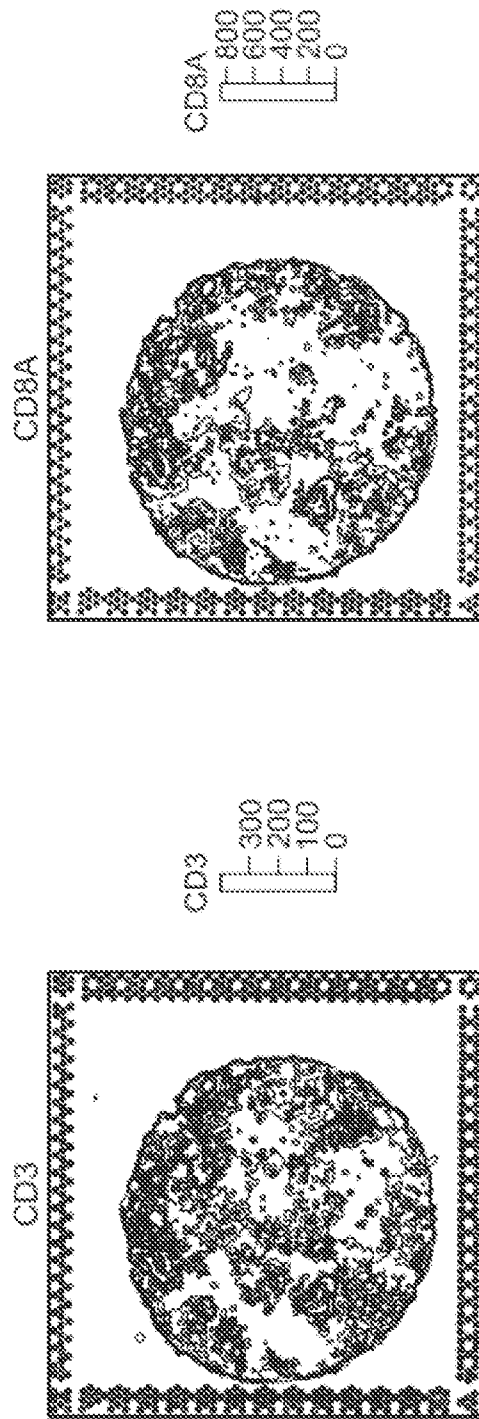


FIGURA 23C

FIGURA 23D

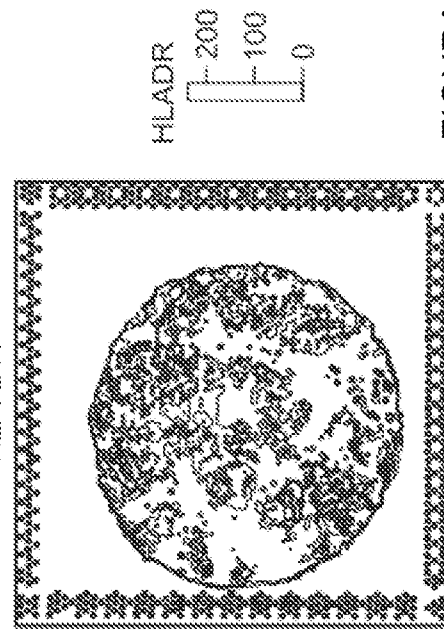


FIGURA 23E

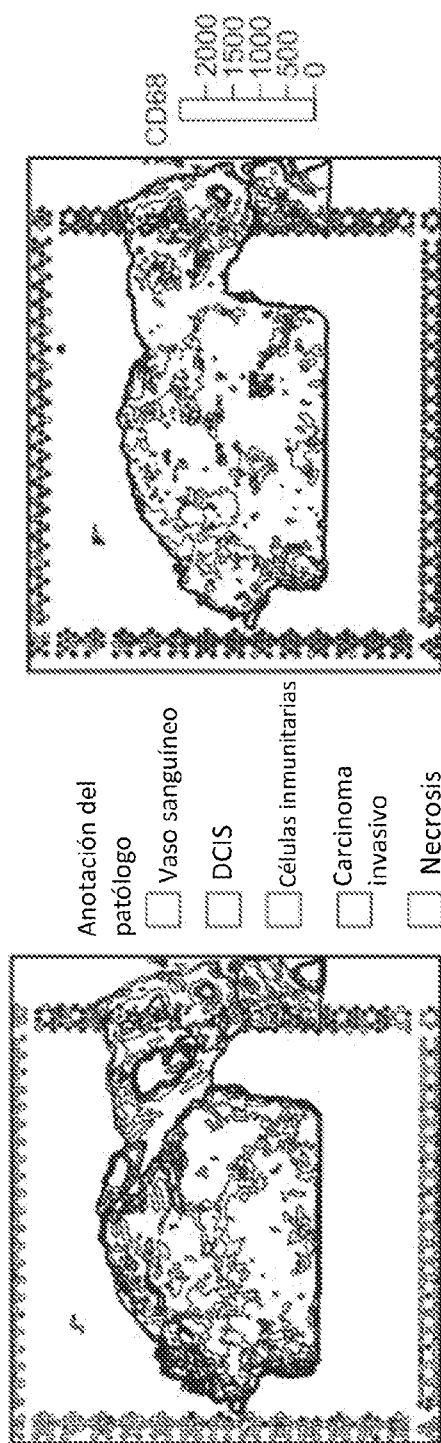


FIGURA 24A

FIGURA 24C

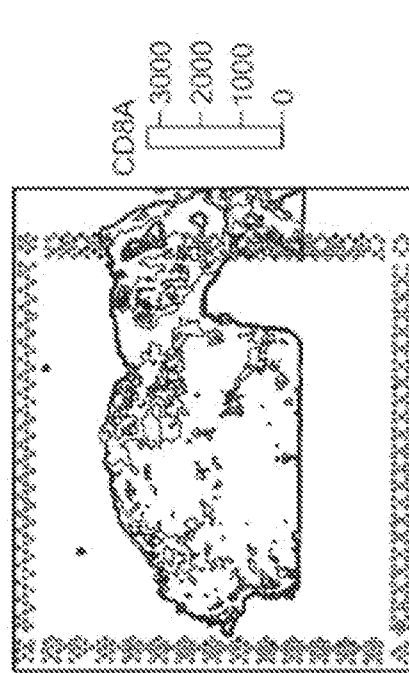
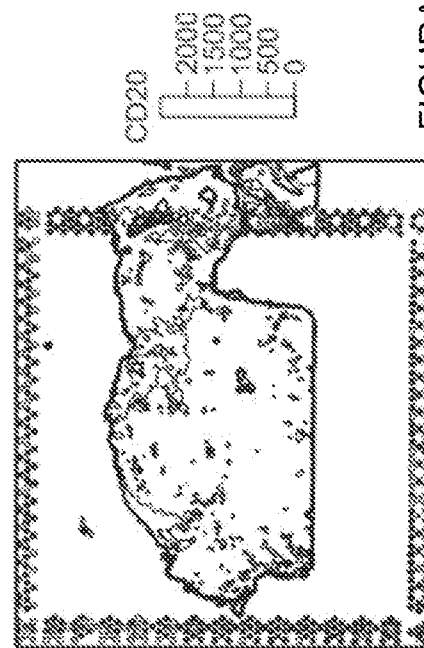


FIGURA 24B

FIGURA 24D



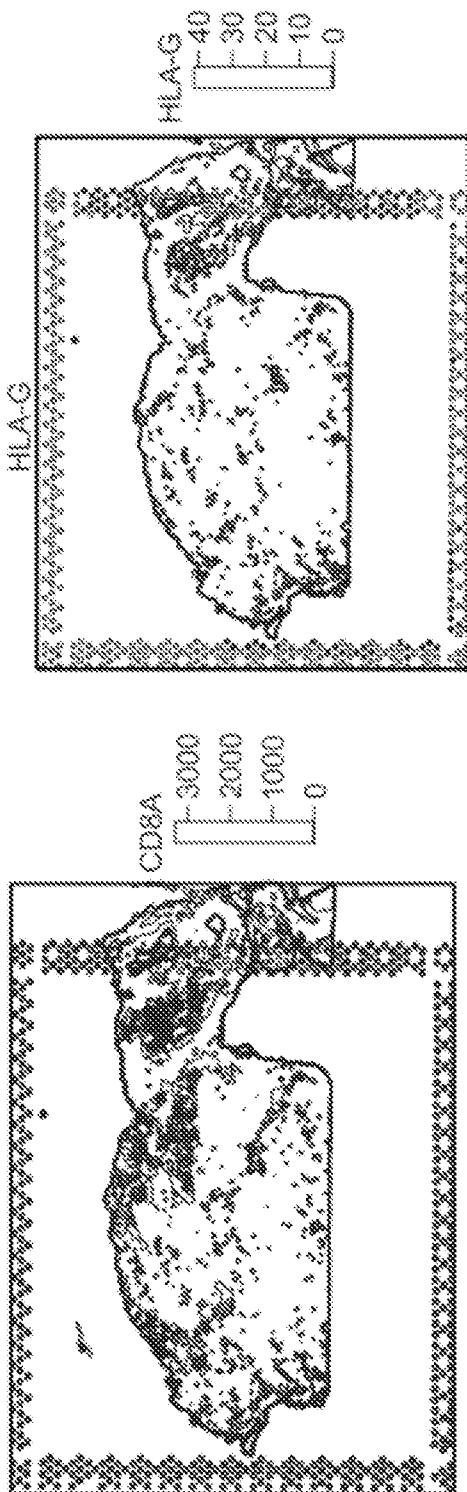


FIGURA 25A

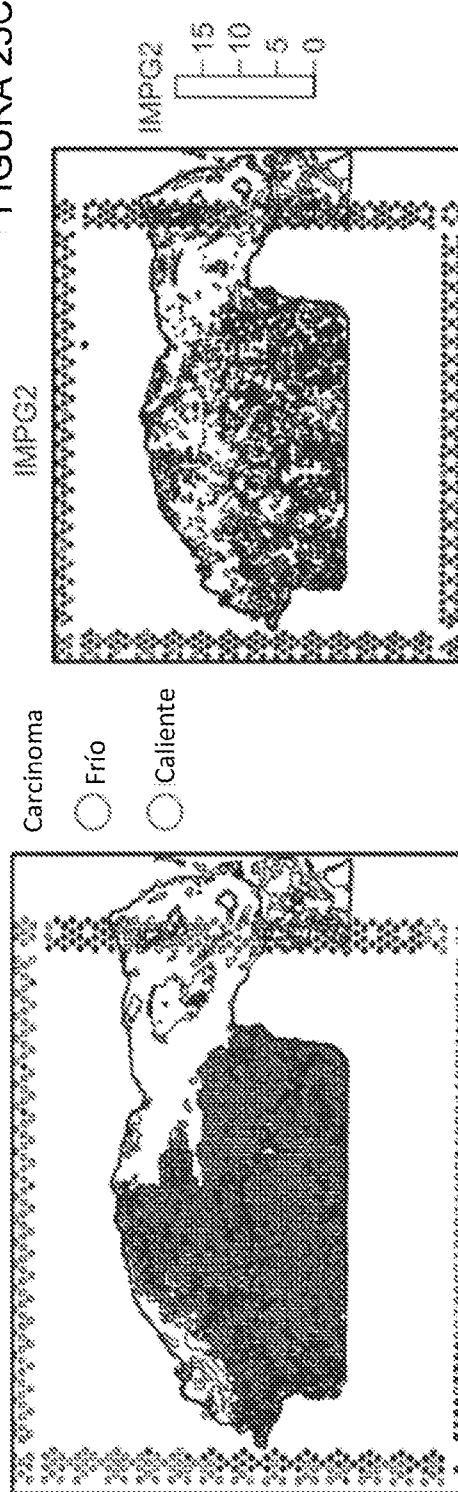


FIGURA 25B

FIGURA 25C

FIGURA 25D

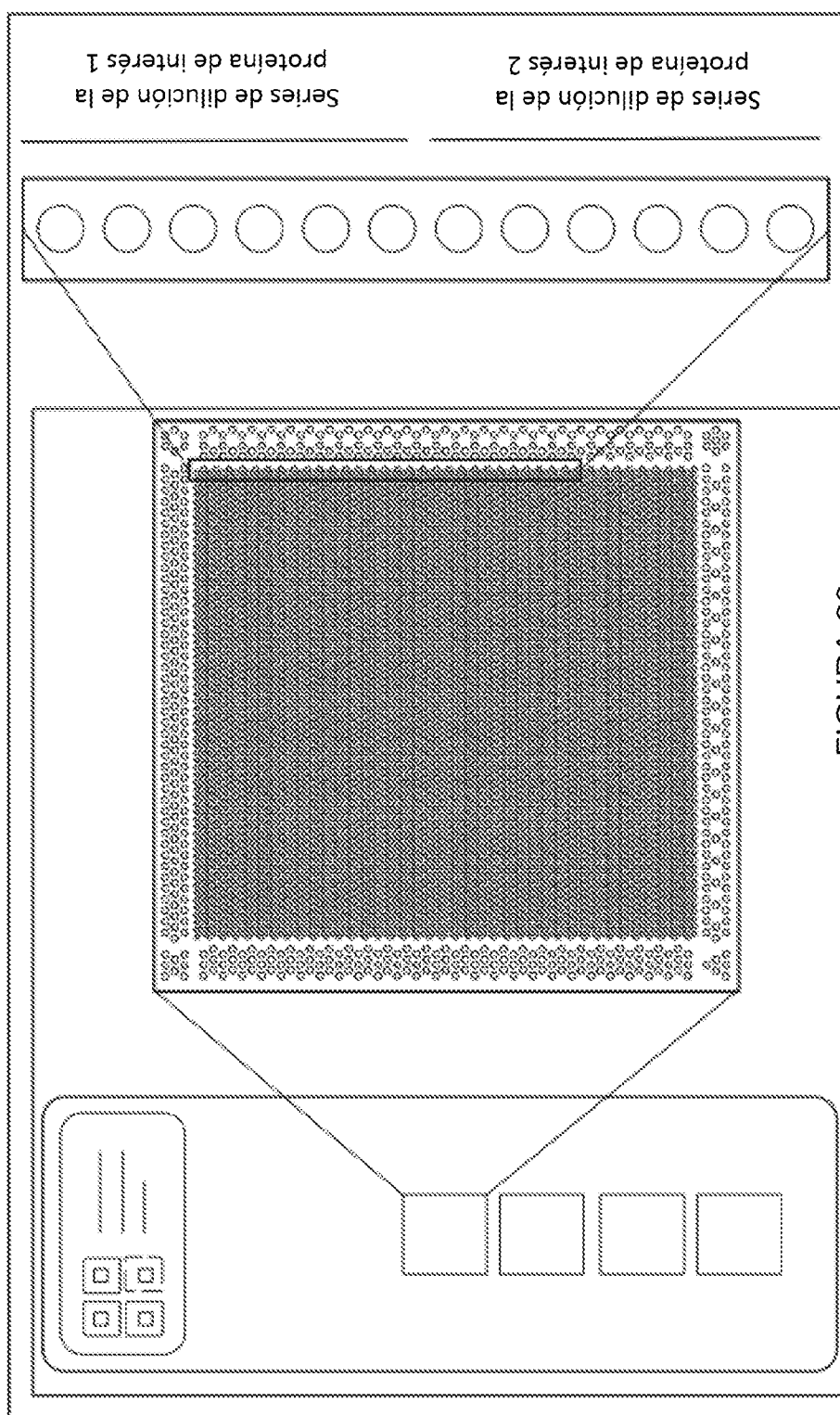


FIGURA 26

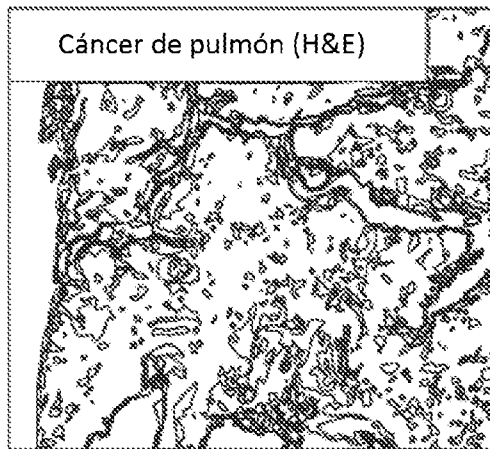


FIGURA 27A

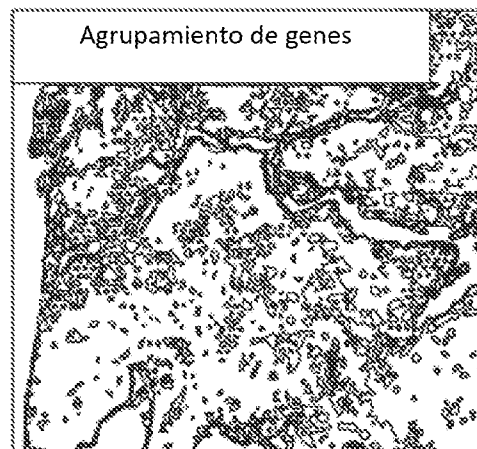


FIGURA 27B

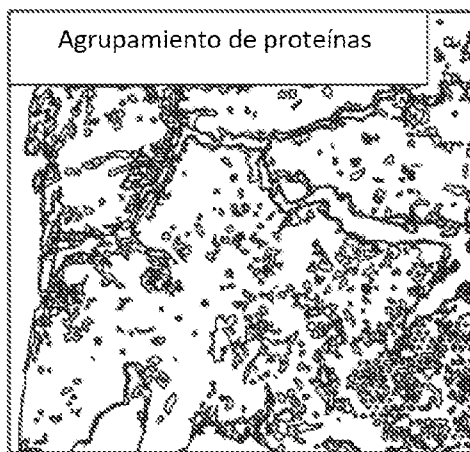


FIGURA 27C

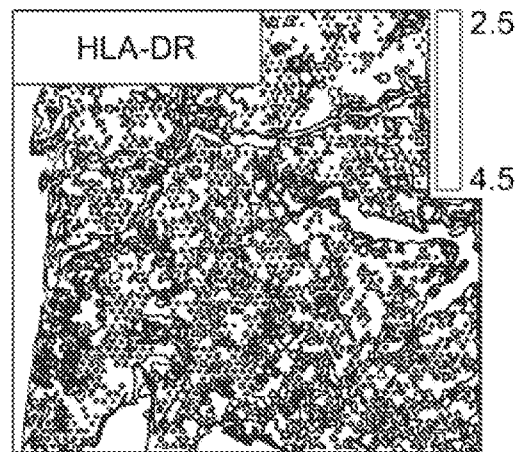


FIGURA 27D

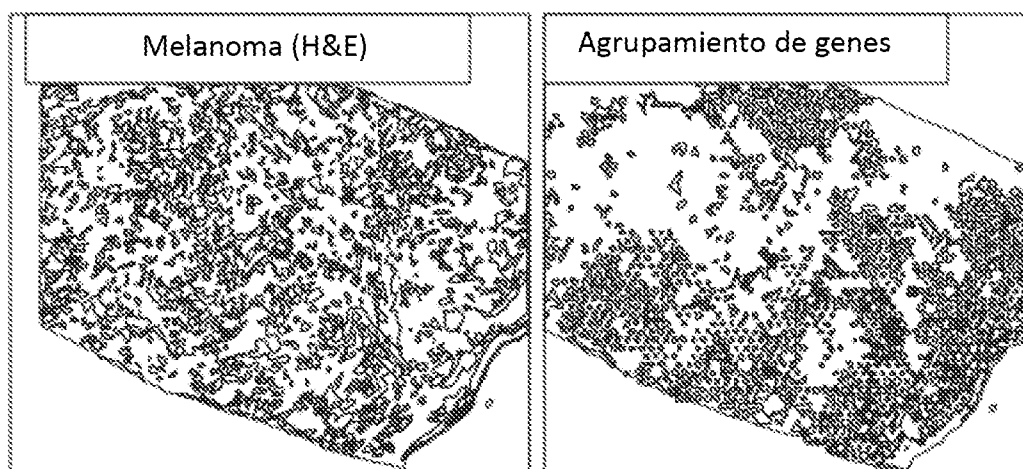


FIGURA 28A

FIGURA 28B

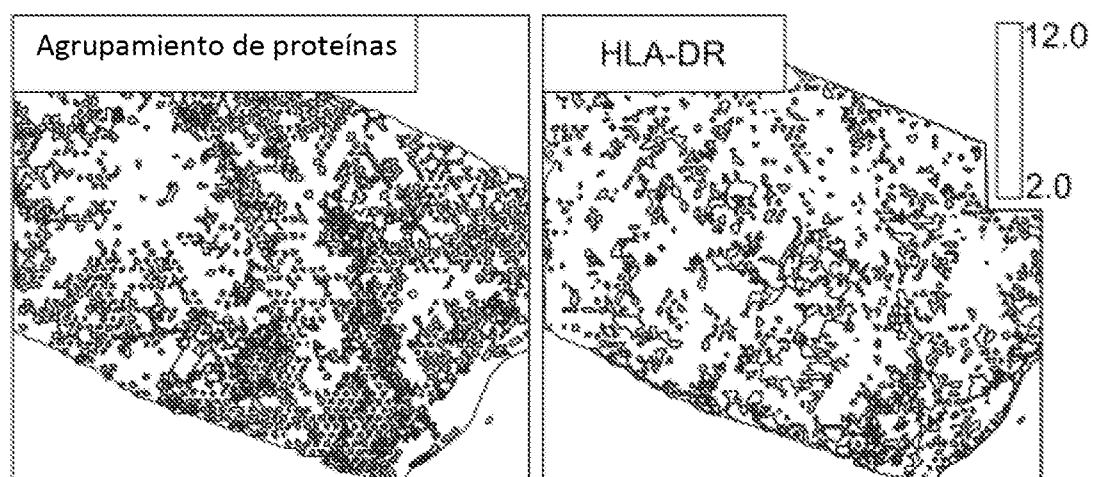


FIGURA 28C

FIGURA 28D

