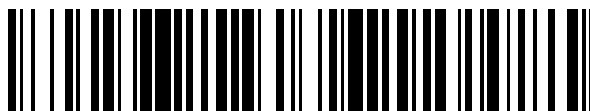


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 374 686**

51 Int. Cl.:
G06K 9/00 (2006.01)
G06T 7/00 (2006.01)
G01N 21/00 (2006.01)
G06F 19/00 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **08754418 .5**
96 Fecha de presentación: **14.05.2008**
97 Número de publicación de la solicitud: **2156370**
97 Fecha de publicación de la solicitud: **24.02.2010**

54 Título: **SEPARACIÓN EN COMPARTIMENTOS POR CARACTERIZACIÓN DE PÍXEL USANDO AGRUPAMIENTO DE DATOS DE IMÁGENES.**

30 Prioridad:
14.05.2007 US 917853 P

45 Fecha de publicación de la mención BOPI:
21.02.2012

45 Fecha de la publicación del folleto de la patente:
21.02.2012

73 Titular/es:
HISTORX, INC.
300 GEORGE STREET
NEW HAVEN, CT 06511, US

72 Inventor/es:
CHRISTIANSEN, Jason, H.;
PINARD, Robert;
GUSTAVSON, Mark;
BOURKE, Brian;
REILLY, Dylan, M. y
TEDESCHI, Gregory, R.

74 Agente: **Tomas Gil, Tesifonte Enrique**

ES 2 374 686 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

Separación en compartimentos por caracterización de píxel usando agrupamiento de datos de imágenes

5 Aplicación relacionada

[0001] La presente aplicación reivindica el beneficio de la solicitud provisional estadounidense con número de serie: 60/917,853, solicitado el 14 de mayo de 2007.

10 Antecedentes de la invención

[0002] El análisis cuantitativo automatizado de la expresión de biomarcador en las secciones de tejido o microarrays tisulares presenta diferentes desafíos, incluyendo heterogeneidad de secciones de tejido, localización subcelular de coloración y la presencia de señal de fondo. Por ejemplo, dependiendo del tipo de tumor o sección del tejido que está siendo analizado, el área de interés puede representar casi la muestra entera o sólo un pequeño porcentaje. Por ejemplo, un carcinoma pancreático o carcinoma lobular del pecho con respuesta desmoplástica sustancial puede mostrar tejido estromal que representa un porcentaje grande del área total. Si el objetivo del ensayo es determinar la expresión de la célula epitelial de un marcador dado, un protocolo debe usarse que valore sólo esa región. El protocolo no sólo debe ser capaz de seleccionar la región de interés, sino también normalizarla, de modo que el nivel de expresión leído de cualquier área dada se pueda comparar con la de otras áreas. La localización subcelular presenta desafíos similares. Sistemas automatizados y métodos para analizar rápidamente secciones de tejido, incluyendo microarrays tisulares, que permiten la identificación y la localización de marcadores biológicos identificados dentro de compartimentos subcelulares en tejidos y otras células con muestras, se necesitan. Un primer documento de la técnica anterior es la publicación de la patente internacional WO2006/083969 que divulga un método para la discriminación de características morfológicas presentes en una muestra biológica de interés usando metodologías basadas en fluorescencia, mediante la substracción de una máscara de imagen nuclear de una máscara de imagen de campo luminoso. Otro documento de la técnica anterior es la publicación de la patente internacional WO2006/133325 que divulga un método para análisis de tejido desde el margen quirúrgico de tumor resecado, y el uso de esta información para predecir recurrencia, supervivencia y eficacia del tratamiento en pacientes con cáncer.

[0003] Determinados métodos (incluyendo microscopia confocal convolución/deconvolución) se ha usado para cuantificar la expresión de proteínas a nivel celular (o subcelular) dentro de un único campo de alta potencia. Estos métodos, no obstante, son computacionalmente técnicas intensivas y laboriosas que operan imágenes múltiples en serie. Como resultado, el estándar actual para análisis de inmunohistología es análisis patológico convencional basado y clasificación de la muestra según escala.

[0004] Sistemas automatizados para análisis histológico de secciones de tejido incluyen frecuentemente métodos que o tienen 1) un operador que examina una imagen de un campo de visión de un tejido teñido y parámetros de ajuste para condiciones de análisis óptimas, o 2) ajustes constantes que tratan un conjunto de datos entero de la misma manera, pero un operador sigue siendo requerido para hacer llamadas de decisión en el ajuste de los parámetros iniciales, es decir, umbrales. Ambos métodos padecen al menos la desventaja de que los datos no están siendo tratados por un único método constante que sea completamente objetivo. Estas decisiones pueden influir en la salida del sistema y afectar a la calidad de los datos. Ellos también añaden una capa extra de complejidad del sistema en el que los métodos de análisis se pueden ajustar a experimentos individuales o muestras individuales y ningún método universal se usa.

50 Resumen de la invención

[0005] La presente invención se refiere en general a métodos de detección y cuantificación de expresión de proteína y de identificación de compartimentos biológicos definidos por marcador. Es un objeto de la presente invención proporcionar métodos de compartimentos de definición en los que la expresión de biomarcador se localiza y cuantifica en tejidos y células con muestras, que requieren intervención mínima de usuario y proporcionan compartimento óptimo, incluyendo resolución de compartimento subcelular.

[0006] En una forma de realización, la presente invención se refiere a un método para la definición de un primer compartimento biológico definido por marcador en relación a un segundo compartimento biológico definido por marcador presente en una muestra biológica de interés, comprendiendo comparación de la intensidad en cada uno de los lugares de píxel en una primera imagen de alta resolución del primer compartimento biológico definido por marcador con la intensidad en cada uno de los lugares de píxel correspondiente a una segunda imagen de alta resolución del segundo compartimento biológico definido por marcador, en el que la primera imagen de alta resolución se preparó usando un primer agente de imágenes que es específico para el primer compartimento biológico definido por marcador, y en el que la segunda imagen de alta resolución se preparó usando un segundo agente de imágenes que es específico para el segundo compartimento biológico definido por marcador, en el que

diferencias en la intensidad del píxel definen el primer compartimento biológico definido por marcador en relación al segundo compartimento biológico definido por marcador. En una forma de realización particular, el método es automatizado, por ejemplo, en el que el método se implementa por un ordenador. En una forma de realización particular, los píxeles de las dos imágenes de alta resolución se representan en un gráfico, en los que los ejes del gráfico comprenden la intensidad del primer agente de imágenes y la intensidad del segundo agente de imágenes. En formas de realización particulares, los métodos de la presente invención comprenden opcionalmente i) asignación de píxeles a una agrupación caracterizada por el hecho de intensidad alta del primer agente de imágenes e intensidad baja del segundo agente de imágenes en el primer compartimento; ii) asignación de píxeles a una agrupación caracterizada por el hecho de intensidad alta del segundo agente de imágenes e intensidad baja del primer agente de imágenes en el segundo compartimento; y iii) asignación de píxeles a una agrupación caracterizada por el hecho de intensidad baja del primer agente de imágenes e intensidad baja del segundo agente de imágenes en fondo y eliminación de estos píxeles de análisis posterior. En una forma de realización particular, cualquiera de los pasos de asignación son realizados usando un algoritmo de agrupamiento, por ejemplo, una k significa método de agrupación para determinar un miembro de agrupación por cada píxel. En una forma de realización particular, los métodos de la invención comprenden opcionalmente la asignación de píxeles restantes intensidad de primer agente de imágenes e intensidad de segundo agente de imágenes para el primer compartimento o el segundo compartimento se basan en probabilidad. En una forma de realización particular, los métodos de la invención comprenden opcionalmente la asignación de esos píxeles restantes con intensidad de primer agente de imágenes e intensidad de segundo agente de imágenes ni para el primer compartimento ni para el segundo compartimento. En una forma de realización particular, el compartimento biológico se selecciona del grupo consistente en: un tipo celular, compartimento subcelular, un compartimento tisular y un compartimento celular o tisular localizado. En una forma de realización particular, el compartimento biológico es un compartimento subcelular seleccionado del grupo consistente en: núcleo celular, citoplasma, membrana nuclear, membrana celular, mitocondria, retículo endoplasmático, peroxisoma y lisosoma. En una forma de realización particular, el compartimento biológico es un compartimento tisular seleccionado del grupo consistente en: epitelio, estroma, mesotelio. En una forma de realización particular, la muestra es una muestra de tejido, preparación celular o fracción subcelular. En una forma de realización particular, los métodos de la invención comprenden opcionalmente la definición de una máscara definida por la intensidad de píxel del primer y/o segundo agente de imágenes y definición de asignación de compartimento para sólo esos píxeles en la máscara. En una forma de realización particular, los métodos de la invención comprenden opcionalmente la incubación de la muestra con un primer agente de imágenes que específicamente etiqueta el primer compartimento biológico definido por marcador, un segundo agente de imágenes que específicamente etiqueta un segundo compartimento biológico definido por marcador.

[0007] En una forma de realización, la presente invención se refiere a un método implementado en ordenador para la definición de un primer compartimento biológico definido por marcador en relación a un segundo compartimento biológico definido por marcador presente en una muestra biológica comprendiendo: a) incubación de la muestra con un primer agente de imágenes que específicamente etiqueta el primer compartimento definido por marcador, y un segundo agente de imágenes que específicamente etiqueta el segundo compartimento definido por marcador; b) obtención de una primera imagen de alta resolución de la primera muestra etiquetada de agente de imágenes, y una segunda imagen de alta resolución de la segunda muestra etiquetada de agente de imágenes; c) determinación de intensidad de un primer y un segundo agente de imágenes en cada lugar de píxel correspondiente en la primera y la segunda imagen; d) realización de un análisis de agrupamiento en cada píxel basado en la intensidad de cada píxel del primer y el segundo agente de imágenes en cada uno del píxel para calcular agrupaciones; e) asignación de esos píxeles en la agrupación caracterizada por intensidad alta de píxel del primer de agente de imágenes e intensidad baja de píxel de agente de imágenes en el primer compartimento; f) asignación de esos píxeles en la agrupación caracterizada por intensidad alta de píxel de segundo agente de imágenes e intensidad baja de píxel del primer agente de imágenes al segundo compartimento; y g) asignación de esos píxeles en la agrupación caracterizada por intensidad baja de píxel del primer agente de imágenes e intensidad baja de píxel del segundo agente de imágenes al fondo y eliminación de estos píxeles de análisis posterior, definiendo así un primer compartimento subcelular definido por marcador en relación a un segundo compartimento subcelular definido por marcador.

[0008] En una forma de realización, la presente invención se refiere a un método implementado en ordenador para la localización y cuantificación de un biomarcador particular dentro de un primer compartimento biológico definido por marcador en relación a un segundo compartimento biológico definido por marcador presente en una muestra biológica comprendiendo: a) incubación de la muestra de tejido con un primer agente de imágenes que específicamente etiqueta el primer compartimento definido por marcador, un segundo agente de imágenes que específicamente etiqueta un segundo compartimento subcelular definido por marcador, y un tercer agente de imágenes que específicamente etiqueta el biomarcador; b) obtención de una primera imagen de alta resolución de la primera muestra de agente de imágenes marcada, una segunda imagen de alta resolución de la segunda muestra de agente de imágenes etiquetada y una tercera imagen de alta resolución de la tercera muestra de agente de imágenes marcada; c) determinación de la intensidad de píxel del primer agente de imágenes y del segundo agente de imágenes en cada uno de los lugares de píxel en la primera y la segunda imagen; d) realización de un análisis de agrupamiento en los píxeles para la asignación de píxeles al primer compartimento definido por marcador o el segundo compartimento definido con marcador; y e) análisis en la tercera imagen los lugares de píxel asignados a los compartimentos para así identificar esos lugares de píxel con un valor de intensidad indicativo del tercer agente

de imágenes, y determinación del valor de intensidad total del tercer agente de imágenes en los lugares de píxel asignados a cada uno de los compartimentos primero y segundo, con el fin de localizar y cuantificar el biomarcador en el primer compartimento en relación al segundo compartimento. En una forma de realización particular, las imágenes de alta resolución son obtenidas usando un microscopio óptico vertical o invertido. En una forma de realización particular, el análisis de agrupación se realiza usando agrupación de k medias reiterativas en la intensidad del primer y el segundo píxel en cada uno de los lugares de píxel para calcular tres centroides usando distancias de algoritmo de Euclides o de logaritmo de verosimilitud. En una forma de realización particular, los métodos de la presente invención comprenden opcionalmente i) representación en el gráfico de los lugares de píxel y los centroides calculados en los que los ejes del gráfico comprenden la intensidad del primer agente de imágenes y la intensidad del segundo agente de imágenes en los lugares de píxel para el primer compartimento y el segundo compartimento; ii) conexión de los centroides para la definición de un triángulo; iii) asignación de esos lugares de píxel con una intensidad que no está en la área del triángulo: (1) al primero compartimento si la intensidad de píxel es sustancialmente indicativa del primer agente de imágenes; (2) al segundo compartimento si la intensidad de píxel es sustancialmente indicativa del segundo agente de imágenes, o (3) para cualquier compartimento si la intensidad de píxel es sustancialmente indicativa al fondo; y iv) asignación de esos lugares de píxel dentro del área del triángulo del primero compartimento o el segundo compartimento basada en un valor correspondiente a la probabilidad de que el píxel se origine desde el primer o el segundo compartimento. En una forma de realización particular, el biomarcador se selecciona del grupo consistente en: una proteína, un péptido, un ácido nucleico, un lípido y un carbohidrato. En una forma de realización particular, cada uno de los agentes de imágenes primero, segundo y tercero comprenden un fluoróforo. En una forma de realización particular, la cantidad del biomarcador presente en el primer o el segundo compartimento comprende la suma de los valores de intensidad del tercer agente de imágenes en los lugares de píxel en el compartimento y la división de la suma por el número de píxeles en el compartimento. En una forma de realización particular, un lugar de píxel no asignado al primer o el segundo compartimento se asigna a un tercer compartimento. En una forma de realización particular, la muestra es una muestra de tejido con un espesor de aproximadamente cinco micras. En una forma de realización particular, el primer compartimento es una membrana celular y el segundo compartimento es un núcleo celular. En una forma de realización particular, la muestra biológica es una sección de tejido fijado. En una forma de realización particular, el primer o el segundo agente de imágenes reacciona con un marcador que se selecciona del grupo consistente en: citoqueratina, beta catenina, alfa catenina y vimentina. En una forma de realización particular, al menos uno del primero, el segundo o el tercer agente de imágenes comprenden un fluoróforo seleccionado del grupo consistente en: 4'.6-diamidino-2-fenilindol (DAPI), Cy3, Cy-5-tiramida y colorantes flúor de Alexa. En una forma de realización particular, el biomarcador se selecciona del grupo consistente en: Her-2/neu, receptor de estrógeno, receptor de progesterona, receptor factor de crecimiento epidérmico, fosfatasa y homólogo de tensina (PTEN) y grupo de complementación cruz de reparación por escisión 1 (ERCC1). En una forma de realización particular, una máscara se aplica a las imágenes de alta resolución primera, segunda y tercera, y sólo píxeles en la máscara se analizan.

[0009] En una forma de realización, la presente invención se refiere a un medio informático legible comprendiendo las instrucciones informáticas legibles almacenadas en este para la ejecución por un procesador de la realización de un método para determinar una dilución óptima de un reactivo para el uso en un inmunoensayo cuantitativo que incluye las etapas de: recepción de una pluralidad de conjuntos de dilución, cada conjunto de dilución con un valor de dilución diferente respectivo y comprendido una pluralidad respectiva de valores de intensidad de coloración de inmunoensayo; determinación de cada uno de la pluralidad de conjuntos de dilución un intervalo respectivo dinámico métrico relacionado con la pluralidad respectiva de valores de intensidad de coloración de inmunoensayo; e identificación del conjunto de dilución con el intervalo numéricamente más grande dinámico métrico, el valor de dilución del conjunto de dilución identificada representando un nivel de dilución óptima del reactivo para el uso en el inmunoensayo cuantitativo.

[0010] En una forma de realización, la presente invención se refiere a una señal electromagnética que lleva unas instrucciones legibles por ordenador para la determinación de una dilución de un reactivo óptima para el uso en un inmunoensayo cuantitativo incluyendo las etapas de: recepción de una pluralidad de conjuntos de dilución, cada conjunto de dilución con un valor de dilución diferente respectivo y comprendiendo una pluralidad respectiva de valores de intensidad de coloración de inmunoensayo; determinación de cada uno de la pluralidad de conjuntos de dilución un intervalo respectivo dinámico métrico relacionado con la pluralidad respectiva de valores de intensidad de coloración de inmunoensayo; e identificación del conjunto de dilución con el intervalo dinámico métrico numéricamente mayor, el valor de dilución del conjunto de dilución identificada representando un nivel de dilución óptima del reactivo para el uso en el inmunoensayo cuantitativo.

[0011] En una forma de realización, la presente invención se refiere a un método implementado en ordenador para la definición de un primer compartimento subcelular definido por marcador en relación a un segundo compartimento subcelular definido por marcador presente en células individuales de interés contenidas en una muestra de tejido comprendiendo: a) incubación de la muestra de tejido con una primera mancha que específicamente etiqueta el primer compartimento subcelular definido por marcador, una segunda mancha que específicamente etiqueta un segundo compartimento subcelular definido por marcador, b) obtención de una imagen de alta resolución de cada primera y segunda mancha en la muestra de tejido usando un microscopio para la obtención de: i) una primera imagen del primer compartimento subcelular definido por marcador; ii) una segunda imagen del segundo compartimento subcelular definido por marcador; y c) determinación de la primera y segunda intensidad de mancha

en cada uno de los lugares de píxel en la primera y la segunda imagen; d) representación en el gráfico de los píxeles, en el que los ejes del gráfico comprenden la intensidad de la primera mancha y la intensidad de la segunda mancha; e) realización de agrupación de k medias reiterativas en la intensidad de la primera y la segunda mancha en cada uno de los lugares de píxel para calcular tres agrupaciones; f) asignación de esos píxeles en la agrupación caracterizada por el hecho de intensidad alta de la primera mancha e intensidad baja de la segunda mancha al primer compartimento; g) asignación de esos píxeles en la agrupación caracterizada por el hecho de intensidad alta de la segunda mancha y intensidad baja de la primera mancha al segundo compartimento; h) asignación de esos píxeles en la agrupación caracterizada por el hecho de intensidad baja de la primera mancha e intensidad baja de la segunda al fondo y eliminación de esos píxeles de análisis posterior; i) asignación de esos píxeles con intensidad de la primera mancha e intensidad de la segunda mancha para el primer compartimento o el segundo compartimento basado en la probabilidad definiendo así un primer compartimento subcelular definido por marcador en relación a un segundo compartimento subcelular definido por marcador.

[0012] En una forma de realización, la presente invención se refiere a un método implementado en ordenador para la definición de un primer compartimento subcelular definido por marcador en relación a un segundo compartimento subcelular definido por marcador presente en células individuales de interés contenidas en una muestra de tejido comprendiendo: a) incubación de la muestra de tejido con una primera mancha que específicamente etiqueta el primer compartimento subcelular definido por marcador, una segunda mancha que específicamente etiqueta un segundo compartimento subcelular definido por marcador, b) obtención de una imagen de alta resolución de cada una de la primera y la segunda mancha en la muestra de tejido usando un microscopio para la obtención de: i) una primera imagen del primer compartimento subcelular definido por marcador; ii) una segunda imagen del segundo compartimento subcelular definido por marcador; y c) determinación de la intensidad de la primera y segunda mancha en cada uno de los lugares de píxel en la primera y la segunda imagen; d) mostrando en el gráfico los píxeles, en el que los ejes del gráfico comprenden la intensidad de la primera mancha y la intensidad de la segunda mancha; e) realización de agrupación de k medias reiterativas en la intensidad de la primera y la segunda mancha en cada uno de los lugares de píxel para calcular tres agrupaciones; f) asignación de esos píxeles en la agrupación caracterizada por el hecho de que intensidad alta de la primera mancha y intensidad baja de la segunda mancha en el primer compartimento; g) asignación de esos píxeles en la agrupación caracterizada por el hecho de que intensidad alta de la segunda mancha y bajo intensidad de la primera mancha en el segundo compartimento; h) asignación de esos píxeles en la agrupación caracterizada por el hecho de que intensidad baja de la primera mancha e intensidad baja de la segunda mancha al fondo y eliminación de estos píxeles de análisis posterior; i) asignación de esos píxeles con intensidad de la primera mancha e intensidad de la segunda para ni el primer compartimento ni el segundo compartimento; definiendo así un primer compartimento subcelular definido por marcador en relación a un segundo compartimento subcelular definido por marcador.

[0013] En una forma de realización, la presente invención se refiere a un método implementado en ordenador para localizar y cuantificar un biomarcador particular dentro de un primer compartimento subcelular definido por marcador en relación a un segundo compartimento subcelular definido por marcador presente en células individuales de interés contenidas en una muestra de tejido comprendiendo: a) incubación de la muestra de tejido con una primera mancha que específicamente etiqueta el primer compartimento subcelular definido por marcador, una segunda mancha que específicamente etiqueta un segundo compartimento subcelular definido por marcador, y una tercera mancha que específicamente etiqueta el biomarcador; b) obtención de una imagen de alta resolución de cada una de la primera, la segunda y la tercera mancha en la muestra de tejido usando un microscopio óptico vertical o invertido para la obtención de: i) una primera imagen del primer compartimento subcelular definido por marcador; ii) una segunda imagen del segundo compartimento subcelular definido por marcador; y iii) una tercera imagen del biomarcador, c) determinación de la intensidad de la primera y la segunda mancha en cada uno de los lugares de píxel en la primera y la segunda imagen; d) realización de agrupación de k medias reiterativas en la intensidad de la primera y la segunda mancha en cada uno de los lugares de píxel para asignar píxeles para el primer compartimento subcelular definido por marcador del segundo compartimento subcelular definido por marcador; e) análisis en la tercera imagen en los lugares de píxel asignados al primer compartimento subcelular, el segundo compartimento subcelular, o ambos compartimentos en la fase (f) y paso (g) identificación de esos lugares de píxel con un valor de intensidad indicativo de la tercera mancha, y determinación del valor de intensidad total de la tercera mancha en los lugares de píxel asignados a cada uno del primer y segundo compartimento subcelular; con el fin de localizar y cuantificar el biomarcador en el primer compartimento subcelular en relación al segundo compartimento subcelular.

[0014] En una forma de realización, la presente invención se refiere a un método implementado en ordenador para localizar y cuantificar un biomarcador particular dentro de un primer compartimento subcelular definido por marcador en relación a un segundo compartimento subcelular definido por marcador presente en células individuales de interés contenidas en una muestra de tejido comprendiendo: a) incubación de la muestra de tejido con una primera mancha que específicamente etiqueta el primer compartimento subcelular definido por marcador, una segunda mancha que específicamente etiqueta un segundo compartimento subcelular definido por marcador, y una tercera mancha que específicamente etiqueta el biomarcador; b) obtención de una imagen de alta resolución de cada una de la primera, la segunda y la tercera mancha en la muestra de tejido usando un microscopio óptico vertical o invertido para la obtención de: i) una primera imagen del primer compartimento subcelular definido por marcador; ii) una segunda imagen del segundo compartimento subcelular definido por marcador; y iii) una tercera imagen del biomarcador, c) determinación de la intensidad de la primera y la segunda mancha en cada uno de los lugares de

píxel en la primera y la segunda imagen; d) realización de agrupamiento de k medias reiterativas en la intensidad de la primera y la mancha en cada uno de los lugares de píxel para calcular tres centroides usando distancias de algoritmo de Euclides o logaritmo de verosimilitud; e) representación en el gráfico de los lugares de píxel y los centroides calculados donde los ejes del gráfico comprenden la intensidad de la primera mancha y la intensidad de la segunda mancha en los lugares de píxel para el primer compartimento y el segundo compartimento. f) conexión de los centroides para la definición de un triángulo, g) asignación de esos lugares de píxel con una intensidad que no se encuentra en el área del triángulo: (1) al primer compartimento si la intensidad de píxel es sustancialmente indicativa de la primera mancha; (2) al segundo compartimento si la intensidad de píxel es sustancialmente indicativa de la segunda mancha, o (3) para cualquier compartimento si la intensidad de píxel es sustancialmente indicativa de fondo; h) asignación de esos lugares de píxel dentro del área del triángulo del primer compartimento o el segundo compartimento basado sobre un valor correspondiente a la probabilidad de que el píxel se origine desde el primer o el segundo compartimento; i) análisis en la tercera imagen de los lugares de píxel asignado al primer compartimento subcelular, el segundo compartimento subcelular, para indentificar esos lugares de píxel con un valor de intensidad indicativo de la tercera mancha, y determinación del valor de intensidad total de la tercera mancha en los lugares de píxel asignados a cada uno del primer y segundo compartimento subcelular; con el fin de localizar y cuantificar el biomarcador en el primer compartimento subcelular en relación al segundo compartimento subcelular.

[0015] Otras características, objetos y ventajas de la invención se desprenderán de la siguiente figura, descripción detallada y reivindicaciones.

Breve descripción de los dibujos

[0016] La FIG. 1 es esquemática del agrupamiento usado para la determinación de tres centroides en los datos.

[0017] La FIG. 2 muestra una línea celular teñida con DAPI (núcleos), anti-citoqueratina (Cy3), y anti-integrina alfa-V (Cy5).

[0018] La FIG. 3 es un gráfico de dispersión que muestra DAPI (normal) frente a Cy3 (normal).

[0019] La FIG.4 presenta datos que muestran agrupamiento usando distancia de logaritmo de verosimilitud (Auto).

[0020] La FIG. 5 presenta datos que muestran agrupamiento de porcentaje de DAPI: distancia de logaritmo de verosimilitud (Fuerza 3). Ambos producen agrupamiento equivalente, uso de porcentaje Cy3, tomando la agrupación superior como píxeles positivos Cy3-predominante y en la parte inferior agrupamiento como píxeles positivos DAPI-predominante.

[0021] La FIG. 6 muestra porcentaje Cy3 "agrupación de señalización" (basado en valores normalizados) logaritmo de verosimilitud (Estandarizado)-Fuerza 3.

[0022] La FIG. 7 muestra coloración de tipo de célula Her2 Spot 17.

[0023] La FIG. 8 muestra coloración de tipo de célula p53 Spot 2.

[0024] Figuras de 9A a 9C son gráficos que muestran análisis de agrupamiento. FIG. 9A) descripción de modelo de método C-AQUA que muestra asignación de píxel específico: fondo (caja entre puntos A y C), 100% citoplasma/Cy3 (caja entre el borde izquierdo y C; eje Y; 0% nuclear/DAPI), 0-100% citoplasma/Cy3 (triángulo ABC; 0% nuclear/DAPI), 0-100% nuclear/DAPI (triángulo ABD, 0% citoplasma/Cy3), y 100% nuclear/DAPI (caja inferior; eje X; 0% citoplasma/Cy3). FIG. 9B) 2 _ 2 diagrama de dispersión que muestra Cy3 (Y) y Dapi (X) intensidades de píxel graficado uno contra el otro con centroides indicados (B, fondo; C, citoplasma; N, nuclear). Esta imagen pasó la validación en la que ambos centroides de compartimento fueron mayores que 1 desviación típica fuera del centroide inferior. FIG. 9C) 2 X 2 diagrama de dispersión de un punto de tejido diferente que muestra intensidades de píxel indicadas y centroides. Esta imagen falló la validación debido a distancia insuficiente (<1 desviación típica) entre el centroide citoplásmico (C) y centroide de fondo (B).

[0025] Las FIG. 10A y 10B son gráficos que muestran comparaciones entre análisis AQUA® y C-AQUA. FIG. 10A) análisis de regresión con R de Pearson indicado y valores Rho de Spearman entre puntuaciones AQUA® generadas por dos operadores altamente entrenados usando algoritmos de análisis tradicional AQUA®. FIG. 10B) análisis de regresión con R de Pearson indicado y valores Rho de Spearman entre puntuaciones AQUA® generadas por dos operadores altamente entrenados usando algoritmos C-AQUA en los mismos datos establecidos como en A.

[0026] Las FIG. 11A y 11B son gráficos que muestran comparaciones entre análisis AQUA® y C-AQUA. FIG. 11A) análisis de regresión lineal para tamaño de compartimento nuclear entre análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con Pearson R indicado y valores rho de Spearman. FIG. 11B) análisis de regresión lineal para tamaño de compartimento citoplásmico entre análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con Pearson R indicado y valores rho de Spearman.

[0027] FIG. de 12A a 12C son gráficos que muestran comparaciones entre análisis AQUA® y C-AQUA. FIG. 12A) análisis de regresión lineal para puntuaciones ER AQUA® entre análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con Pearson R indicado y valores Rho de Spearman. FIG. 12B) análisis de regresión lineal para puntuaciones PR AQUA® entre análisis AQUA® (eje Y) y análisis C-AQUA (eje x) con Pearson R indicado y valores Rho de Spearman. FIG. 12C) análisis de regresión lineal para puntuaciones Hcr2 AQUA® entre análisis AQUA® (eje Y) y análisis C-AQUA (eje x) con Pearson R indicado y valores Rho de Spearman.

[0028] FIG. de 13A a 13C son gráficos que muestran comparaciones entre análisis AQUA® y C-AQUA. Comparaciones de resultados de supervivencia para FIG. 13A) ER, FIG. 13B) PR y FIG. 13C) Her2 expresión para análisis AQUA® (izquierda) y análisis C-AQUA (derecha) que muestra consecuencias de supervivencia similares basadas en atribución de punto de corte como determinado por distancia de logaritmo de verosimilitud de agrupamiento no supervisado [FIG. 13A) AQUA® tradicional: 11,4% de reducción en la supervivencia total (rango logarítmico $p = 0,018$) de 80,9% (ER alto) a 69,5% (ER bajo); C-AQUA: 13,8% de reducción en la supervivencia total (rango logarítmico $p = 0,005$) de 81,6% (ER alto) a 67,8% (ER bajo); FIG. 13B) AQUA® tradicional: 12,4% de reducción en la supervivencia total (rango logarítmico $p = 0,021$) de 84,2% (PR alto) a 71,8% (PR bajo); C-AQUA: 14,5% de reducción en la supervivencia total (rango logarítmico $p = 0,001$) de 83,3% (PR alto) a 68,8% (PR bajo); y FIG. 13C) AQUA® tradicional: 18,5% de reducción total en la supervivencia total (rango logarítmico $p = 0,022$) de 77,1% (Her2 bajo) a 73,8% (Her2 medio) a 58,6% (Her2 alto); C-AQUA: 24,2% de reducción total en la supervivencia total (rango logarítmico $p = 0,002$) de 77,8% (Her2 bajo) a 73,8% (Her2 medio) a 53,6% (Her2 alto)].

[0029] Las figuras 14A y 14B muestran comparación de puntuación de expresión PTEN AQUA® (regresión lineal) como determinado por análisis AQUA® y C-AQUA. FIG. 14A) análisis de regresión lineal para expresión PTEN nuclear como determinado por análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con R de Pearson indicado. FIG. 14B) análisis de regresión lineal para expresión PTEN citoplasmático como determinado por análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con valores R de Pearson indicados.

[0030] FIG. 15 puntuaciones de expresión PTEN AQUA® citoplasmático.

[0031] La FIG. 16 muestra la correlación de puntuaciones PTEN AQUA® derivadas por ambos métodos para resultados en pacientes se muestran en curvas Kaplan Meier. Puntuaciones PTEN AQUA® fueron significativamente correlacionadas con la supervivencia del paciente. Baja expresión PTEN se asoció a resultados pobres en comparación con alta expresión PTEN.

[0032] Las figuras 17A y 17B muestran una comparación de puntuaciones de expresión ERCC1 AQUA® (regresión lineal) como determinado por análisis AQUA® y C-AQUA. FIG. 14A) análisis de regresión lineal para expresión ERCC 1 nuclear como determinado por análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con R de Pearson indicado. FIG. 14B) análisis de regresión lineal para expresión ERCC1 citoplasmática como determinada por análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con valores de R de Pearson indicados.

[0033] FIG. 18 puntuaciones de expresión de ERCC1 AQUA® citoplasmática.

[0034] La FIG. 19 muestra la correlación de puntuaciones ERCC1 AQUA® derivadas por ambos métodos para resultados en pacientes se muestran en curvas Kaplan Meier. Puntuaciones ERCC1 AQUA® se correlacionaron significativamente con la supervivencia del paciente. Expresión ERCC1 baja se asoció a resultados pobres en comparación con expresión ERCC1 alta.

Descripción detallada

[0035] Se debe apreciar que ciertos aspectos, modos, formas de realización, variaciones y características de la invención se describen abajo a varios niveles de detalle para la proporción de una comprensión sustancial de la presente invención.

[0036] A menos que se defina de otra manera, todos los términos científicos y técnicos usados aquí tienen generalmente el mismo significado como entendido comúnmente por alguien técnico en la materia a la que esta invención pertenece. Como se usa en esta especificación y en las reivindicaciones anexas, las formas singulares "un" y "el" incluyen referentes plurales a menos que el contenido claramente lo dicte de otra manera. Por ejemplo, en referencia a "una célula" incluye una combinación de dos o más células y similares. Generalmente, la nomenclatura usada aquí y los procedimientos de laboratorio en la biología celular, inmunohistoquímica y formación de imágenes (p. ej., células y tejido) descritos abajo son aquellos bien conocidos y empleados de forma común en la técnica. Técnicas estándar, o modificaciones de estas se usan para análisis inmunohistoquímicos. Todas las referencias citadas aquí se incorporan aquí para referencia en su totalidad y para todos los efectos en la misma extensión como si cada publicación individual, patente, o solicitud de patente fue específicamente e individualmente incorporada por referencia en su totalidad a todos los efectos.

[0037] La tecnología de microarray tisular ofrece la oportunidad para análisis de rendimiento alto de muestras de tejido (Konen, J. et al., *Nat. Med.* 4:844-7 (1998); Kallioniemi, O. P. et al., *Hum. Mol. Genet.* 10:657-62 (2001); Rimm, D. L. et al., *Cancer J.* 7:24-31 (2001)). Por ejemplo, la capacidad para realizar rápidamente estudios a gran escala usando microarrays tisulares puede proporcionar información decisiva para la identificación y validación de dianas farmacológicas/marcadores pronósticos (p. ej. receptor de estrógeno (ER) y HER2/neu) y candidatos terapéuticos.

[0038] Muchos marcadores biológicos muestran una distribución paramétrica (normal "bell-shaped") y en consecuencia se analizan mejor en una escala continua (p. ej., 0 a 1000). Desafortunadamente, la observación manual tiende a ser nominal (p. ej. 1+, 2+, 3+), principalmente debido a que el ojo humano es incapaz de distinguir de forma fiable diferencias sutiles en la intensidad de la coloración. Diferentes métodos se han desarrollado para traducir observaciones nominales manuales en una escala continua. La más importante entre estos es la puntuación H en la que el porcentaje de células coloreadas positivamente (0 a 100) se multiplica por la intensidad de coloración (p. ej. 0 a 3) para hacer una escala continua teóricamente (0 a 300). No obstante, la incapacidad para detectar diferencias sutiles en la intensidad de coloración, particularmente en los extremos inferiores y superiores de la escala, al igual que la tendencia para puntuaciones redondas (por ejemplo, 50% a 3+ para una puntuación de 150, frente a 47% a 3+ para una puntuación de 141), limita la eficacia de la puntuación H.

[0039] En algunos aspectos, la presente invención proporciona métodos mejorados para la cuantificación y la localización de un objetivo particular en componentes celulares definidos. Los presentes inventores han descubierto un método para lograrlo que tiene la ventaja de ser completamente objetivo y minimiza la intervención de operador o la toma de una decisión. El método realiza un agrupamiento en los datos de intensidad para cada compartimento celular adquirido. Este agrupamiento permite la eliminación de fondo, la asignación de píxeles específicos a un compartimento dado y la asignación probabilística de píxeles para cada compartimento en el que puede haber superposición de señales. Una vez que los píxeles se asignan a cada compartimento (o se descartan en caso de ruido) las señales de objetivo asociadas pueden medirse, por ejemplo, sumarse y calcular una puntuación.

[0040] La invención proporciona métodos para la asignación de píxel objetivo para compartimentos específicos. La asignación se determina preferentemente sobre una base de imagen-a-imagen, mejor que criterio universal de ajuste. Además, la asignación de píxeles (p. ej., píxeles de CY3/citoqueratina al citoplasma) es también una función de otras imágenes de compartimento, de manera que la consideración se da en el estado de píxeles en otras imágenes de compartimento. En una forma de realización, una imagen es de una primera mancha que específicamente etiqueta un primer compartimento (p. ej., una imagen de CY3/citoqueratina, representando el compartimento citoplasmático) y una segunda imagen es de una segunda mancha que específicamente etiqueta un segundo compartimento (p. ej., imagen DAPI, representando el compartimento nuclear) y las asignaciones de píxeles se basan en cuatro criterios:

- 1.) Intensidad baja en ambas primera y segunda imagen (p. ej., DAPI y Cy3): FONDO: ELIMINACIÓN
- 2.) Intensidad alta de la segunda mancha (por ejemplo; DAPI) en relación a la intensidad de la primera mancha (Cy3): SEGUNDO COMPARTIMENTO (p. ej., NUCLEAR)
- 3.) Intensidad alta de la primera mancha (por ejemplo; Cy3) en relación a la intensidad de la segunda mancha (p. ej., DAPI): primer compartimento (por ejemplo; citoplasmático)
- 4.) Intensidad alta de segunda y primera mancha (p. ej., DAPI y Cy3): INDETERMINADO: ELIMINACIÓN

[0041] El agrupamiento es una función matemática algorítmica por la que los centroides dentro de conjuntos de datos se definen por distancias relativas de cada punto de datos al otro, como determinado, por ejemplo, por distancia de algoritmo de Euclides o distancia de logaritmo de verosimilitud. Sin querer estar atado a la teoría, se cree que intensidades de píxel de agrupamiento de al menos dos imágenes (es decir, DAPI y Cy3), podrían suponer centroides que definen píxeles como se describe, al menos, por el criterio anterior. Porque el agrupamiento es objetivo y se puede realizar individualmente en cada imagen, el agrupamiento se descubrió para proporcionar una asignación de píxeles fiable para compartimentos, independiente de la intervención de operador.

[0042] En otra forma de realización, los píxeles con señal indicativa de la primera y segunda mancha se asignan a compartimentos por el siguiente método. Cada píxel en imágenes adquiridas tiene tres atributos, intensidad de la aportación del marcador de compartimento A, aportación de intensidad de marcador de compartimento B y una aportación de intensidad del objetivo o biomarcador de interés. Estas intensidades se miden en sus canales de fluorescencia respectivos por la configuración experimental. Para evitar desvío experimental, la intensidad del objetivo no se manipula en este presente método. Así, los datos para los dos atributos de compartimento se puede ilustrar en un gráfico bidimensional esquemáticamente mostrado en la FIG. 1. La extensión típica de los datos se representa por el triángulo rectángulo discontinuo.

[0043] Los píxeles con una inclinación fuerte hacia cualquiera de los ejes se pueden asignar a ese compartimento (p. ej., píxeles en regiones A y B podrían asignarse absolutamente a los compartimentos A y B, respectivamente). Los

píxeles cerca del origen representan intensidades bajas para ambos canales y se pueden descartar como fondo con píxeles atípicos que tienen intensidad alta pero valores similares, mostrados en la región D. Los píxeles que permanecen en la región etiquetado A/B pueden asignarse después a cada compartimento basado en probabilidad. Esta asignación permite señal de objetivo en aquellos píxeles que han de ser distribuidos a través de ambos compartimentos basados en la caracterización de probabilidad.

[0044] Para definir las regiones anteriormente descritas, por ejemplo, para cada imagen, el agrupamiento se utiliza para determinar tres centroides en los datos (mostrados como C1, C2 y C3). Este método está completamente automatizado y no requiere ninguna decisión de operadores para proceder. El análisis se realiza mediante la realización de agrupamiento de k medias en tres centroides usando distancias euclídeas. Una vez estos puntos son determinados, las regiones ilustradas en la FIG. 1 se generan usando estos puntos. Los datos se analizan luego de la siguiente manera: (i) fondo y píxeles atípicos se descartan de cálculo posterior. Un píxel se define como fondo si su distancia al origen es inferior a dos veces la del centroide de fondo (C2) distancia al origen. Un píxel se define como atípico si su intensidad excede el valor definido por la línea o plano definido por los centroides extremos (p. ej., C1 y C3 en la FIG. 1 región D); (ii) Píxeles en regiones A y B se asignan exclusivamente a los dos compartimentos; (iii) Píxeles en la región triangular A/B se asignan luego a un valor de probabilidad que les permite ser distribuido esencialmente en múltiples compartimentos. Este valor de probabilidad puede calcularse basado en la distancia desde las dos regiones A y B, o usando una función de forma que también asigna una probabilidad de cada píxel con una aportación de la región de fondo examinando cada distancia del píxel desde los tres vértices dados por los centroides; (iv) Con todos píxeles asignados, las puntuaciones de objetivo asociadas se pueden sumar para cada compartimento y una puntuación calculada usando métodos estándar:

$$\frac{\sum_i^{\# \text{ Píxeles}} \text{Int}_i * P_i}{\sum_i^{\# \text{ Píxeles}} P_i}$$

donde *Int* es la intensidad del píxel, *P* es la probabilidad del píxel que es asignada a un compartimento particular (que varía de 0 a 1).

Métodos generales

[0045] En general, lo que aquí se describe es una serie de técnicas que pueden usarse para análisis rápido automatizado de célula con muestras, incluyendo tejidos y microarrays tisulares. Aunque estas técnicas se basan en otras y se describen como un proceso compacto, cada técnica tiene una aplicabilidad amplia y se pueden utilizar individualmente o en otras combinaciones que las descritas abajo.

[0046] En una forma de realización particular, los métodos de la invención se usan preferentemente con análisis AQUA®, las características se describen en la patente US 7,219,016, que se incorpora para referencia en su totalidad.

[0047] En un disposición experimental AQUA® típica, muestras de tejido se manchan con marcadores que definen, por ejemplo, los compartimentos subcelulares de interés y el objetivo específico (u objetivos) que se estudian. Asignación local basada en píxel para compartimentación de expresión (PLACE) es el algoritmo de clave que funciona para segmentar eficazmente píxeles de imagen de segmento con motivo de compartimentación de expresión. Una paso decisivo en este algoritmo es el ajuste de los umbrales de intensidad que se utilizan para delinear el fondo o los píxeles no-específicos de píxeles específicos de señal. Imágenes que han sido "enmascaradas" de esta manera se combinan posteriormente en una forma mutuamente exclusiva de manera que píxeles sobre los umbrales se asignan a compartimentos subcelulares específicos. Una vez que los píxeles se han asignado a cada compartimento, la señal para el biomarcador objetivo puede después calcularse según el promedio sobre todos los píxeles asignados a un compartimento dado, que es la puntuación AQUA® para esta muestra.

[0048] Por ejemplo, en una muestra de tumor epitelial, dos manchas pueden utilizarse para diferenciar la región tumoral y los compartimentos subcelulares incorporados: DAPI (4'-6-diamidino-2-fenilindol: una etiqueta específica de coloración nuclear/ADN bicatenario) y citoqueratina (un biomarcador epitelial específico etiquetado para lectura fluorescente). Estas imágenes se umbralizan individualmente para eliminar la señal no-específica, luego se combinan para producir una imagen que representa píxeles que no son sólo epitelial específicos, sino que también representan citoplasma y píxeles específicos nucleares. Intensidades de píxel de un objetivo específico que ha sido etiquetado para lectura en un tercer canal fluorescente puede cuantificarse posteriormente dentro de esta imagen "PLACEd".

[0049] Sería ventajoso, específicamente para operación clínica, mejorar el algoritmo de análisis de puntuación AQUA®, de manera que la segmentación de imagen se automatiza completamente, eliminando así el paso de

umbral definido por el usuario. Este mejoraría el sistema en diferentes maneras: primero, debido al tiempo de operador asociado con la definición de un ajuste de umbral optimizado, la eficiencia del sistema aumentaría inmensamente. Segundo, debido a la naturaleza subjetiva de los umbrales de ajuste, incluso por operadores con experiencia, la variabilidad operador-a-operador se podría eliminar. Tercero, para fines clínicos y/o eficiencia de investigación de laboratorio y control de la calidad, un método constante de umbrales de ajuste debe aplicarse para todas las imágenes específicas de canal adquiridas a través de cohorte TMA o una sección de tejido entera. El desarrollo de un método automatizado tipo PLACE permitiría la segmentación de imagen que debe optimizarse en una base imagen-por-imagen. Y, finalmente, el método descrito aquí implica examen de imágenes de compartimento simultáneamente, así los umbrales se fijan en el contexto de datos de píxel para todos los marcadores de compartimento.

[0050] La presente invención se puede utilizar para localizar y cuantificar un biomarcador dentro de cualquier imagen de célula con muestra imprimible, incluyendo, pero no limitado a, biopsias de tejido y célula con muestras fluidas, como, por ejemplo, sangre, orina, líquido cefalorraquídeo, saliva, linfa, líquido pleural, líquido peritoneal y líquido pericárdico y para el análisis de microarrays tisulares.

[0051] Cualquier dispositivo de formación de imágenes óptico o no óptico puede usarse, como, por ejemplo, microscopios ópticos verticales o invertidos, microscopios confocales de escaneo, cámaras, microscopios electrónicos de escaneo o efecto túnel, microscopios de sonda de escaneo y detectores infrarrojos de imágenes, etc.

[0052] En las formas de realización anteriormente descritas, el ordenador puede incluir hardware, software, o una combinación de ambos para controlar los otros componentes del sistema y para analizar las imágenes. El análisis anteriormente descrito se implementa en programas informáticos con uso de técnicas de programación estándar. Estos programas se diseñan para ejecutarse en ordenadores programables comprendiendo cada uno un procesador, un sistema de almacenamiento de datos (incluyendo memoria y/o elementos de almacenamiento), al menos un dispositivo de entrada, al menos un dispositivo de salida, como una pantalla o una impresora. El código de programa se aplica para datos de entrada (p. ej., imágenes pegadas o pilas de imágenes) para realizar las funciones descritas aquí y generar información (p. ej., localización de señal), que se aplica para uno o más dispositivos de salida. Cada programa de ordenador se puede implementar en un lenguaje de programación procesal de alto nivel o lenguaje de programación orientada a objetos, o un lenguaje ensamblador o lenguaje de máquina. Cada programa de ordenador se puede almacenar en un medio de almacenamiento legible por ordenador (p. ej., CD ROM o disquete magnético) que, cuando es leído por un ordenador, puede provocar que el procesador en el ordenador realice el análisis descrito aquí.

[0053] Lo siguiente proporciona una descripción detallada de una forma de realización específica de la preparación y el análisis de microarrays tisulares según métodos descritos aquí, aunque pasos similares podrían realizarse con respecto a cualquier célula con muestra. Un microarray tisular incluye muestras múltiples de histospots obtenidos a partir de histocores introducidos típicamente en un bloque fino de parafina a intervalos regulares, formando una serie de filas y columnas. Histospots (secciones finas de histocores) pueden tener sustancialmente forma tipo disco y típicamente un espesor de aproximadamente cinco micras y un diámetro de aproximadamente 0,6 milímetros. Típicamente los centros de los histospots se distancian sobre unas décimas de un milímetro aparte en los bloques de parafina. Secciones del histospots se pueden montar en un portaobjetos del microscopio. Un microarray tisular puede incluir cualquier número de histospots, típicamente en el orden de varios cientos a unos pocos miles.

[0054] Una estación de microscopía óptica puede utilizarse para obtener una imagen apropiada del tejido. Una estación de microscopía incluye un microscopio óptico para la formación de imágenes del tejido y un ordenador para el análisis de las imágenes. Un microscopio óptico incluye una montura, aloja una fuente de luz, una fase de muestra, una lente del objetivo y una cámara CCD. Un software capturador se utiliza para adquirir las imágenes a través de cámara CCD.

[0055] Un microscopio óptico también incluye diferentes filtros de luz para proporcionar los espectros de iluminación apropiados para microscopía estándar o fluorescente. Por ejemplo, para microscopía fluorescente los filtros pueden estar en las ruedas de filtro y un alojamiento, que alberga una serie de filtros dicróicos. Los filtros en la rueda permiten selección de los espectros de iluminación apropiados. Los filtros en la rueda alteran (filtro) la luz transmitida para el aislamiento de firmas espectrales en la microscopía fluorescente. Una fase de muestra sostiene y posiciona apropiadamente el portaobjetos del microscopio con la muestra de tejido o microarray tisular. Una fase de muestra puede traducirse linealmente traducida en las direcciones "x", "y" y "z" (los ejes se muestran). Una fase de muestra incluye motores para permitir la traducción automatizada. Un ordenador controla la traducción de fase de muestra por control de servo de los motores.

[0056] Un microarray tisular se puede representar de la siguiente manera: un usuario posiciona el microarray en una fase de muestra. El usuario ajusta la fase de muestra de modo que el primer histospot (por ejemplo; superior-izquierda) es al central del campo de visión y focalizado en por la cámara CCD. La lente del objetivo debería ajustarse a la resolución apropiada, por ejemplo, un histospot 0,6 milímetro se puede ver a una 10X ampliación.

Los histospots generalmente corresponden a áreas de intensidad de luz más alta que la de la parafina circundante, como se evaluó a través de varios medios incluyendo señales derivadas de la dispersión de luz visible de tejidos

- teñidos, autofluorescencia de tejido o de una marca fluorescente. Un ordenador puede adquirir una imagen de baja resolución (p. ej. 64 píxel X 64 píxel con resolución de 16 bit) usando software (Softworx 2.5, Applied Precision, Issaquah, WA) y una plataforma de formación de imágenes (p. ej., Deltavision). Un ordenador traduce automáticamente la fase de muestra por una cantidad aproximadamente igual a un campo de visión. El ordenador luego adquiere una segunda imagen de baja resolución. Este proceso se repite hasta que el ordenador ha adquirido imágenes de la muestra de tejido entera o microarray. Con uso de software disponible comercialmente, el ordenador luego genera una imagen compuesta del microarray tisular entero por unión de la secuencia de imágenes como mosaico.
- [0057] Marcadores biológicos, que se pueden detectar conforme a la presente invención incluyen, pero de forma no limitativa, cualquiera de los siguientes: ácidos nucleicos, proteínas, péptidos, lípidos, carbohidratos u otros componentes de una célula. Marcadores determinados son característicos de células particulares, mientras otros marcadores han sido identificados mientras son asociados a una enfermedad particular o condición. Ejemplos de marcadores pronósticos conocidos incluyen marcadores enzimáticos tales como, por ejemplo, galactosil transferasa II, enolasa específica de neurona, protón ATPasa-2 y fosfatasa ácida. Hormona o marcadores receptores de hormonas incluyen gonadotropina humana coriónica (HCG), corticotropina, antígeno carcinoembrionario (CEA), antígeno prostático específico (PSA), receptor de estrógeno, receptor de progesterona, receptor de andrógeno, receptor de complemento gC1q-R/p33, receptor IL-2, receptor de neurotrofina p75, receptor de PTH, receptor de hormonas tiroideas y receptor de insulina.
- [0058] Marcadores linfoides incluyen alfa-1-antiquimotripsina, alfa-1-antitripsina, marcador celular B, bcl-2, bcl-6, antígeno de linfocito B 36 kD, BM1 (marcador mielóide), BM2 (marcador mielóide), galectin-3, granzima B, antígeno HLA clase I, antígeno HLA clase II (DP), antígeno HLA clase II (DQ), antígeno HLA clase II (DR), defensinas neutrofilicas humanas, inmunoglobulina A, inmunoglobulina D, inmunoglobulina G, inmunoglobulina M, kappa de cadena ligera, lambda de cadena ligera, antígeno de linfocito/histocito, marcador macrófago, muramidasa (lisozima), quinasa de linfoma anaplástico p80, marcador de plasmocito, inhibidor de proteasa de leucocito secretor, receptor de antígeno de célula T (JOVI 1), receptor de antígeno de célula T (JOVI 3), desoxinucleotidil transferasa terminal, marcador no-agrupado de células B.
- [0059] Marcadores tumorales incluyen alfa fetoproteína, apolipoproteína D, BAG-1 (proteína RAP46), CA19-9 (sialil lewisa), CA50 (antígeno mucínico asociado a carcinoma), CA125 (antígeno de cáncer de ovario), CA242 (antígeno mucínico asociado a tumor), cromogranina A, clusterina (apolipoproteína J), antígeno de membrana epitelial, antígeno relacionado epitelial, antígeno específico epitelial, receptor factor de crecimiento epidérmico, receptor de estrógeno, proteína 15 fluida de enfermedad quística gruesa, antígeno específico de hepatocito, HER2, heregulina, mucina gástrica humana, glóbulo de grasa de leche humana, MAGE-1, metaloproteinasas matriciales, melan A, marcador de melanoma (HMB45), mesotelina, metalotioneína, factor de transcripción de microftalmia (MITF), glicoproteína de núcleo Muc-1. Glicoproteína Muc-1, glicoproteína Muc-2, glicoproteína Muc-5AC, glicoproteína Muc-6, mieloperoxidasa, Myf-3 (marcador de rabdomiosarcoma), Myf-4 (marcador de rabdomiosarcoma), MyoD 1 (marcador de rabdomiosarcoma), mioglobina, proteína nm23, fosfatasa alcalina de placenta, transtiretina, receptor de progesterona, antígeno específico de próstata, fosfatasa ácida prostática, péptido inhibina prostática, PTEN, marcador de carcinoma de célula renal, antígeno mucinoso del intestino delgado, tetranectina, factor-1 de transcripción de tiroides, inhibidor de tejido de metaloproteinasas matricial 1, inhibidor de tejido de metaloproteinasas matricial 2, tirosinasa, proteína-1 relacionada de tirosinasa, vellosidades, factor von Willebrand, CD34, CD34, clase II, CD51 Ab-1; CD63; CD69, Chk1, Chk2, claspin C-met, COX6C, CREB, ciclina D1, citoqueratina, citoqueratina 8, DAPI, desmina, DHP (1-6 difenil-1,3,5-hexatrieno), E-cadherina, EEA1, EGFR, EGFRvIII, EMA (antígeno de membrana epitelial), ER, ERB3, ERCC1, ERK, e-selectina, FAK, fibronectina, FOXP3, Gamma-H2AX, GB3, GFAP, giantina, GM130, Golgin 97, GRB2, GRP78BiP, GSK3 Beta, HER-2, histona 3, histona 3 K14-Ace [anti-acetil-histona H3 (Lys 14)], histona 3 K18-Ace [histona h3-acetil Lys 18], Histona 3 K27-TriMe, [histona H3 (trimetil K27)], histona 3 K4-diMe [anti-dimetil-histona H3 (Lys 4)], histona 3 K9- ECA [acetil-histona H3 (Lys 9)], histona 3 K9-triMe [histona 3-tri metil Lys 9], histona 3 S10-Fos [anti-fosfo histona H3 (Ser 10)], Marcador de mitosis], histona 4, histona H2A.X S139-Fos [fosfo histona H2A.X (Ser139)anticuerpo], histona H2B, histona H3 DiMetil K4, histona H4 TriMetil K20-Chip grad, HSP70, uroquinasa, VEGF R1, ICAM-1, IGF-1; IGF-1R, IGF-1 receptor beta, IGF-II, IGF-IIR, IKB-alfa IKKE, IL6, IL8, integrina alfa V beta 3, integrina alfa V beta 6, integrina alfa V / CD51, integrina B5, integrina B6, integrina B8, integrina beta 1(CD 29), integrina beta 3, integrina beta 5 integrina B6, IRS-1, Jagged 1, anti-proteína quinasa C Beta2, LAMP-1, cadena ligera Ab-4 (cóctel), cadena ligera lambda, cadena ligera kappa, M6P, Mach 2, MAPKAPK-2, MEK 1, MEK 1/2 (Ps222), MEK 2, MEK1/2 (47E6), péptido de bloqueo MEK1/2, MET/HGFR, MGMT, antígeno mitocondrial, Mitotracker verde FM, MMP-2, MMP9, E-cadherina, mTOR, ATPasa, N-cadherina, nefrina, NFkB, NFkB p105/p50, NF-KB P65, notch 1, notch 2, notch 3, oxOhos complejo IV, p130Cas, p38 MAPK, p44/42 MAPK anticuerpo, P504S, P53; P70; P70 S6K, pan-cadherina, paxilina, P-cadherina, PDI, pEGFR, fosfo AKT, fosfo CREB, fosfo receptor EGF, fosfo GSK3 Beta, fosfo H3, fosfo HSP-70, fosfo MAPKAPK- 2, fosfo MEK1/2, MAP quinasa fosfo p38, fosfo p44/42 MAPK, fosfo p53, fosfo PKC, proteína ribosómica fosfo S6, fosfo Src, fosfo-Akt, fosfo-Bad, fosfo-IKB-a, fosfo-mTOR, fosfo-NF-kappaB p65; Phospho-p38; fosfo-p44/42 MAPK, fosfo-p70 S6 quinasa, fosfo-Rb, fosfo-Smad2, PIM1, PIM2, PKC β , podocalixina, PR, PTEN, R1, Rb 4H1, R-cadherina, ribonucleótido reductasa, RRM1, RRM11, SLC7A5, NDRG, HTF9C, HTF9C, CEACAM, p33, S6 proteína ribosómica, Src, survivina, sinaptopodina, sindecan 4, Talin, tensina, timidilato sintasa, Tuberlin, VCAM-1, VEGF, vimentina, aglutinina, YES, ZAP-70 y ZEB.

5 [0060] Marcadores asociados al ciclo celular incluyen factor activador de proteasa apoptótica 1, bcl-w, bcl-x, bromodesoxiuridina, CAK (cdk-activación quinasa), proteína de susceptibilidad a apoptosis celular (CAS), caspasa 2, caspasa 8; CPP32 (caspasa-3), CPP32 (caspasa-3), quinasas dependientes de ciclina, ciclina A, ciclina B1, ciclina D1, ciclina D2, ciclina D3, ciclina E, ciclina G, factor de fragmentación de ADN (N-terminal), Fas (CD95), proteína de dominio de muerte asociada a Fas, Fas ligando, Fen-1; IPO-38, Mcl-1, proteínas de mantenimiento de minicromosoma, proteína de reparación de desajuste (MSH2), poli (ADP-ribosa) polimerasa, antígeno nuclear de proliferación celular, proteína p16, p27 proteína, p34cdc2; proteína p57 (Kip2), proteína p105, Stat 1 alfa, topoisomerasa I, topoisomerasa II alfa, topoisomerasa III alfa, topoisomerasa II beta.

10 [0061] Tejido neuronal y marcadores de tumor incluyen alfa B cristalina, alfa-internexina, alfa-sinucleína, proteína amiloide precursora, beta-amiloide, calbindina, colina acetiltransferasa, transportador 1 de aminoácido excitatorio; GAP43, proteína ácida fibrilar glial, receptor de glutamato 2, proteína básica de mielina, receptor de factor de crecimiento de nervio (gp75), marcador de neuroblastoma, neurofilamento 68 kD, neurofilamento 160 kD, neurofilamento 200 kD, enolasa específica de neurona, receptor alfa4 de acetilcolina nicotínica, receptor beta2 de acetilcolina nicotínica, periférica, producto genético de proteína 9; proteína S-100, serotonina, SNAP-25, sinapsina I, sinaptofisina, tau, triptófano hidroxilasa, tirosina hidroxilasa, ubiquitina.

15 [0062] Marcadores de diferenciación de agrupación incluyen CD1a, CD1b, CD1c, CD1d, CD1e, CD2, CD3delta, CD3epsilon, CD3gamma, CD4, CD5, CD6, CD7, CD8alfa, CD8beta, CD9; CD10; CD11a, CD11b, CD11c, CDw12; CD13; CD14; CD15; CD15s, CD16a, CD16b, CDw17; CD18; CD19; CD20; CD21; CD22; CD23; CD24; CD25; CD26; CD27; CD28; CD29; CD30; CD31; CD32; CD33; CD34; CD35; CD36; CD37; CD38; CD39; CD40; CD41; CD42a, CD42b, CD42c, CD42d, CD43; CD44; CD44R, CD45; CD46; CD47; CD48; CD49a, CD49b, CD49c, CD49d, CD49e, CD49f, CD50; CD51; CD52; CD53; CD54; CD55; CD56; CD57; CD58; CD59; CDw60; CD61; CD62E, CD62L, CD62P, CD63; CD64; CD65; CD65s, CD66a, CD66b, CD66c, CD66d, CD66e, CD66f, CD68; CD69; CD70; CD71; CD72; CD73; CD74; CDw75; CDw76; CD77; CD79a, CD79b, CD80; CD81; CD82; CD83; CD84; CD85; CD86; CD87; CD88; CD89; CD90; CD91; CDw92; CDw93; CD94; CD95; CD96; CD97; CD98; CD99; CD100; CD101, CD 102, CD 103; CD104; CD105; CD106; CD107a, CD107b, CDw108; CD109, CD 114, CD 115; CD116, CD 117; CDw119; CD120a,CD120b, CD121a, CDw121b, CD122; CD123; CD124; CDw125; CD126; CD127; CDw128a, CDw128b, CD130; CDw131; CD132; CD134; CD135; CDw136; CDw137; CD138; CD139; CD140a, CD140b, CD141; CD142; CD143; CD144; CDw145; CD146; CD147; CD148; CDw149; CDw150; CD151; CD152; CD153; CD154; CD155; CD156; CD157; CD158a, CD158b, CD161; CD162; CD163; CD164; CD165; CD166, y TCR-zeta.

20 [0063] Otros marcadores celulares incluyen proteína-F centrómero (CENP-F), giantin, involucrina, lamin A&C [XB 10], LAP- 70, mucina, proteínas complejas de poro nuclear, p180 proteína de cuerpo lamelar, ran, r, catepsina D, proteína Ps2, Her2-neu, P53; S100, antígeno marcador epitelial (EMA), TdT, MB2, MB3, PCNA, y Ki67.

25 [0064] Células con muestras se pueden manchar usando colorantes o manchas, histoquímicos, o inmunohistoquímicos que reaccionan directamente con los marcadores biológicos específicos o con varios tipos de células o compartimentos subcelulares. No todas las manchas son compatibles. Por lo tanto, el tipo de manchas empleado y su secuencia de aplicación debería ser bien considerada, pero puede ser fácilmente determinado por alguien instruido en la técnica. Estos histoquímicos pueden ser cromóforos detectables por microscopía de transmisión o fluoróforos detectables por microscopía de fluorescencia. En general, célula con muestras se pueden incubar con una solución comprendiendo al menos un histoquímico, que reacciona directamente con o se unen con grupos químicos del objetivo. Algunos histoquímicos deben ser co-incubados con un mordiente o metal para permitir la coloración. Una célula con muestra se puede incubar con una mezcla de al menos un histoquímico que mancha un componente de interés y otro histoquímico que actúa como una contratinción y se une a una región exterior del componente de interés. Alternativamente, mezclas de sondas múltiples se pueden utilizar en la coloración y proporcionan una vía de identificación para las posiciones de sondas específicas.

30 [0065] La siguiente lista no limitativa proporciona cromóforos ejemplares que se pueden utilizar como agentes de imágenes histológicas (manchas o contratinciones) y sus células diana, compartimentos subcelulares, o componentes celulares: eosina (componentes alcalinos celulares, citoplasma), hematoxilina (ácidos nucleicos), naranja G (glóbulos rojos, páncreas y células de glándula pituitaria), verde claro SF (colágeno), Romanowsky-Giemsa (morfología de célula total), May-Grunwald (glóbulos rojos), contratinción azul (Trevigen), verde etilo (CAS) (amiloide), Feulgen-Naftol amarillo S (ADN), Giemsa (diferencialmente manchas de varios compartimentos celulares), verde metilo (amiloide), pironina (ácidos nucleicos), amarillo de naftol (glóbulos rojos), rojo neutro (núcleos), tinción de Papanicolaou (que incluye típicamente una mezcla de hematoxilina, eosina Y, naranja G y mezcla marrón de Bismarck (morfología de célula total), contratinción roja B (Trevigen), contratinción roja C (Trevigen), rojo de Sirius (amiloide), reactivo de Feulgen (pararosanilina) (ADN), galocianina-cromo alúmina (ADN), galocianina-cromo alúmina y amarillo de naftol amarillo S (ADN), verde metilo pironina Y (ADN), reactivo de tionina-Feulgen (ADN), naranja de acridina (ADN), azul de metileno (ARN y ADN), azul de toluidina (ARN y ADN), azul de alciano (carbohidratos), rojo de rutenio (carbohidratos), negro de Sudán (lípidos), Sudán IV (lípidos), aceite rojo-O (lípidos), tinción tricrómica de Van Gieson (mezcla de fucsina ácida y de ácido pícrico) (células musculares), tinción tricrómica de Masson (hematoxilina, fucsina ácida y mezcla de verde claro) (mancha colágeno, citoplasma, nucléolos diferentemente), aldehído fucsina (fibras de elastina) y tinción de Weigert (diferencia fibras reticulares y de

colágeno). Una lista completa de tales tinciones, su descripción y uso general se proporciona en R. D. Lillie, "*Conn's Biological Stains*", 8th ed., Williams y Wilkins Company, Baltimore, Md. (1969). Mordientes adecuados y composiciones del precedente son bien conocidos por alguien instruido en la técnica.

5 [0066] La siguiente lista no limitativa proporciona tinciones fluorescentes histológicas ejemplares y sus células diana, compartimentos subcelulares, o componentes celulares si aplicable: 4',6-diamidino-2-fenilindol (DAPI) (ácidos nucleicos), eosina (componentes alcalinos celulares, citoplasma), Hoechst 33258 y Hoechst 33342 (dos bisbenzimidias) (ácidos nucleicos), yoduro de propidio (ácidos nucleicos), espectro naranja (nucleico ácidos). Espectro verde (ácidos nucleicos), quinacrina (ácidos nucleicos), fluoresceína-faloidina (fibras de actina).
10 Cromomicina A 3 (ácidos nucleicos), reacción de acriflavina-Feulgen (ácido nucleico), reacción de auramina O-Feulgen (ácidos nucleicos), bromuro de etidio (ácidos nucleicos). Tinción de Nissl (neuronas), fluoróforos de ADN de alta afinidad alta como POPO, BOBO, YOYO y TOTO y otros, y proteína verde fluorescente fundida a proteína de enlace de ADN, como histonas, ACMA, quinacrina y naranja de acridina.

15 [0067] Una amplia variedad de sondas fluorescentes específicas de orgánulo patentadas están disponibles comercialmente, e incluyen sondas específicas de mitocondria (colorantes MitoFluor y MitoTracker), retículo endoplasmático (ER) y sondas de Golgi (ER-Tracker y varios conjugados de ceramida) y sondas lisosómicas (colorantes LysoTracker). Estas sondas, al igual que muchos histoquímicos fluorescentes no patentados, están disponibles y se describen extensivamente en *Handbook of Fluorescent Probes and Research Products* 8.sup.th Ed. (2001), disponible de *Molecular Probes*, Eugene, OR.

20 [0068] Cada célula con muestra se puede co-incubar con sustratos apropiados para una enzima que es un componente celular de interés y reactivos apropiados de los que los coloreados de rendimiento precipitan a los sitios de actividad enzimática. Estas tinciones histoquímicas enzimáticas son específicas para la enzima objetivo particular. La coloración con manchas enzimáticas histoquímicas se puede utilizar para definir un componente subcelular o un tipo particular de célula. Alternativamente, tinciones histoquímicas enzimáticas se pueden utilizar diagnósticamente para cuantificar la cantidad de actividad enzimática en células. Una amplia variedad de sustratos enzimáticos y ensayos de detección se conocen y se describen en la técnica.

25 [0069] Las fosfatasa ácidas se pueden detectar a través de diferentes métodos. En el método de Gomori para fosfatasa ácida, una preparación celular se incuba con glicerofosfato y nitrato de plomo. La enzima libera fosfato, que se combina con plomo para producir fosfato de plomo, un precipitado incoloro. El tejido se sumerge luego en una solución de sulfuro de amonio, que reacciona con fosfato de plomo para formar sulfuro de plomo, un precipitado negro. Alternativamente, células se pueden incubar con una solución comprendiendo pararosanilina-HCl, nitrito de sodio, naftol ASB1 fosfato (sustrato) y tampón acetato veronal. Este método produce un precipitado rojo en las áreas de actividad de la fosfatasa ácida. Debido a el contenido característico de la fosfatasa ácida, los lisosomas se pueden distinguir de otros gránulos citoplasmáticos y orgánulos a través del uso de este ensayo.

30 [0070] Las deshidrogenasas se pueden localizar por incubación de células con un sustrato apropiado para las especies de deshidrogenasa y tetrazol. Los iones de hidrógeno de transferencias enzimáticas del sustrato para tetrazol, reduciendo el tetrazol a formazán, un precipitado oscuro. Por ejemplo, NADH-deshidrogenasa es un componente del complejo I de la cadena respiratoria y se localiza predominantemente en la mitocondria.

35 [0071] Otras enzimas bien conocidas en las técnicas de coloración han sido desarrolladas, y sus localizaciones celulares primarias o actividades, incluyen, pero de forma no limitativa, lo siguiente: ATPasas (fibras musculares), succinato deshidrogenasas (mitocondria), citocromo c oxidasas (mitocondria), fosforilasas (mitocondria), fosfofructoquinasas (mitocondria), acetilcolinesterasas (células nerviosas), lactasas (intestino delgado), leucina aminopeptidasas (células hepáticas), mioadenilato desaminasas (células musculares), diaforasas NADH (eritrocitos) y sacarosas (intestino delgado).

40 [0072] La inmunohistoquímica se encuentra entre las técnicas histoquímicas más específicas y sensibles. Cada histospot se puede combinar con una composición de unión etiquetada comprendiendo específicamente una sonda de unión. Varias etiquetas se pueden emplear, tales como fluoróforos, o enzimas que producen un producto que absorbe luz o fluorescencias. Una amplia variedad de etiquetas se sabe que proveen de señales fuertes en relación con un único evento de unión. Sondas múltiples usadas en la coloración se pueden marcar con más de un marcador fluorescente distinguible. Estas diferencias de color proporcionan una vía para identificar las posiciones de sondas específicas. El método de preparación de conjugados de fluoróforos y proteínas, como anticuerpos, se describe extensivamente en la literatura y no requiere ejemplificación aquí.

45 [0073] Aunque hay al menos 120.000 anticuerpos disponibles comercialmente, anticuerpos primarios ejemplares, que se conocen por enlazar específicamente componentes celulares y emplearse actualmente como componentes en tinciones inmunohistoquímicas usadas para la investigación y, en casos limitados, para el diagnóstico de varias enfermedades, incluyen, por ejemplo, anticuerpo receptor de anti-estrógeno (cáncer de mama), anticuerpo receptor de anti-progesterona (cáncer de mama), anticuerpos anti-p53 (cánceres múltiples), anticuerpos anti-Her-2/neu (cánceres múltiples), anticuerpos anti-EGFR (factor de crecimiento epidérmico, cánceres múltiples), anticuerpo de anti-catepsina D (mama y otros cánceres), anticuerpo anti-Bcl-2 (células apoptóticas), anticuerpo anti-E-cadherina,

anticuerpo anti-CA125 (de ovario y otros cánceres), anticuerpo anti-CA15-3 (cáncer de mama), anticuerpo anti-CA19-9 (cáncer de colon), anticuerpo anti-c-erbB-2, anticuerpo anti-P-glicoproteína (MDR, resistencia multifármaco), anticuerpo anti-CEA (antígeno carcinoembrionario), proteína anti-retinoblastoma (Rb) anticuerpo, anticuerpo de oncoproteína anti-ras (p21), anticuerpo anti-Lewis X (también llamado CD15), anticuerpo anti-Ki-67 (proliferación celular), anticuerpo anti-PCNA (cánceres múltiples), anticuerpo anti-CD3 (células T), anticuerpo anti-CD4 (auxiliar de células T), anticuerpo anti-CD5 (células T), anticuerpo anti-CD7 (timocitos, células T inmaduras, células asesinas NK), anticuerpo anti-CD8 (supresor de células T), anticuerpo anti-CD9/p24 (ALL), anticuerpo anti-CD10 (también llamado CALLA) (leucemia linfoblástica aguda común), anticuerpo anti-CD11c (monocitos, granulocitos, AML), anticuerpo anti-Cd 13 (células mielomonocíticas, AML), anticuerpo anti-Cd 14 (monocitos maduros, granulocitos), anticuerpo anti-CD15 (enfermedad de Hodgkin), anticuerpo anti-Cd 19 (células B), anti-CD20 de anticuerpos (B células), anti-CD22 de anticuerpos (B células), anti-CD23 anticuerpo (células B activadas, CLL), anticuerpo anti-CD30 (células T y B activadas, enfermedad de Hodgkin), anticuerpo anti-CD31 (marcador de angiogénesis), anticuerpo anti-CD33 (células mieloides, AML), anticuerpo anti-CD34 (células madre endoteliales, tumores estromales), anticuerpo anti-CD35 (células dendríticas), anticuerpo anti-CD38 (células plasmáticas, células T y B activadas y células mieloides), anticuerpo anti-CD41 (plaquetas; megacariocitos), anticuerpo anti-LCA/CD45 (antígeno de leucocito común), anticuerpo anti-CD45RO (auxiliar, inductor de células T), anticuerpo anti-CD45RA (células B), anticuerpo anti-CD39; CD100, anticuerpo anti-CD95/Fas (apoptosis), anticuerpo anti-CD99 (marcador de sarcoma de Ewings, producto genético MIC2), anticuerpo anti-CD106 (VCAM-1; células endoteliales activadas), anticuerpo anti-ubiquitina (enfermedad de Alzheimer), anticuerpo anti-CD71 (receptor de transferrina), anticuerpo anti-c-myc (oncoproteína y un hapteno), anticuerpo anti-citoqueratinas (receptor de transferrina), anticuerpo anti-vimentinas (células endoteliales) (células B y T), anticuerpo anti-HPV proteínas (virus del papiloma humano), anticuerpo anti-kappa cadenas ligeras (célula B), anticuerpo anti-lambda de cadenas ligeras (célula B), anticuerpo anti-melanosomas (HMB45) (melanoma), antígeno específico anticuerpos anti-próstata (PSA) (cáncer de próstata), anticuerpo anti-S-100 (melanoma, células salivares, células gliales), antígeno de anticuerpo anti-tau (enfermedad de Alzheimer), anticuerpo antifibrina (células epiteliales), anticuerpo anti-queratinas, anticuerpo anti-citoqueratina (tumor), anticuerpo anti-alfa-catenina (membrana celular), anticuerpo anti-Tn-antígeno (carcinoma de colon, adenocarcinomas y cáncer pancreático); anticuerpo anti-1,8-ANS (1-anilino naftaleno-8-ácido sulfónico) anticuerpo anti-C4; anticuerpo grado CASP anti-2C4; α anticuerpo anti-2C4 CASP; anticuerpo anti-HER-2; anticuerpo cristalino anti-alfa B; anticuerpo anti-alfa galactosidasa; anticuerpo anti-alfa-catenina; anticuerpo anti-humano VEGF R1 (Flt-1); anticuerpo anti-integrina B5 ; anticuerpo anti-integrina beta 6; anticuerpo anti-fosfo-SRC; anticuerpo anti-Bak; anticuerpo anti-BCL-2; anticuerpo anti-BCL-6; anticuerpo anti-beta catenina; anticuerpo anti-beta catenina; anticuerpo anti-integrina alfa V beta 3; anticuerpo anti-C ErbB-2 Ab-12; anticuerpo anti-calnexina; anticuerpo anti-calreticulina; anticuerpo anti-calreticulina; anticuerpo anti-CAM5.2 (bajo peso molecular de anti-citoqueratina); anticuerpo anti-cardiotina (R2G); anticuerpo anti-catepsina D; anticuerpo policlonal de pollo para galactosidasa alfa; anticuerpo anti-c-Met; anticuerpo anti-CREB; anticuerpo anti-COX6C; anticuerpo anti-ciclina D1 Ab-4; anticuerpo anti-citoqueratina; anticuerpo anti-desmina; anticuerpo anti-DHP (1-6 difenil-1,3,5-hexatrieno); antígeno biotina de cabra anti-pollo DSB-X; anticuerpo de anti-E-cadherina; anticuerpo anti-EEA1; anticuerpo anti-EGFR; anticuerpo anti-EMA (antígeno de membrana epitelial); anticuerpo anti-ER (receptor de estrógeno); anticuerpo anti-ERB3; anticuerpo anti-ERCC1 ERK (Pan ERK); anticuerpo anti-E-selectina; anticuerpo anti-FAK; anticuerpo anti-fibronectina; anticuerpo FITC de cabra anti-ratón IgM; anticuerpo anti-FOXP3; anticuerpo anti-GB3; anticuerpo anti-GFAP (proteína ácida fibrilar glial); anticuerpo anti-giantina; anticuerpo anti-GM130; anticuerpo anti-cabra a h Met; anticuerpo anti-Golgin 97; anticuerpo anti- GRB2; anticuerpo anti-GRP78BiP; anticuerpo anti-GSK-3Beta; anticuerpo anti-hepatocito; anticuerpo anti-HER-2; anticuerpo anti- HER-3; anticuerpo anti-histona 3; anticuerpo anti-histona 4; anticuerpo anti-histona H2A X; anticuerpo anti-histona H2B; anticuerpo anti-HSP70; anticuerpo anti-ICAM-1; anticuerpo anti-IGF-1; receptor de anticuerpo anti-IGF-1; receptor beta de anticuerpo anti- IGF-1; anticuerpo anti-IGF-II; anticuerpo anti-IKB-alfa; anticuerpo anti-IL6; anticuerpo anti-IL8; anticuerpo anti-integrina beta 3; anticuerpo anti-integrina beta 5; anticuerpo anti-integrina b8; anticuerpo anti-Jagged 1; anticuerpo anti-proteína quinasa C Beta2; anticuerpo anti- LAMP-1; anticuerpo anti-M6P (receptor manosa 6-fosfato); anticuerpo anti-MAPKAPK-2; anticuerpo anti-MEK 1; anticuerpo anti-MEK 2; anticuerpo de antígeno anti-mitocondrial; anticuerpo marcador anti-mitocondrial; anticuerpo anti-Mitotracker verde FM; anticuerpo anti-MMP-2; anticuerpo anti-MMP9; anticuerpo anti-Na⁺/K ATPasa; anticuerpo anti-Na⁺/K ATPasa alfa 1; anticuerpo anti-Na⁺/K ATPasa alfa 3; anticuerpo anti-N-cadherina; anticuerpo anti-nefrina; anticuerpo anti-NF-KB p50; anticuerpo anti-NF- KB P65; anticuerpo anti-Notch 1; anticuerpo conjugado anti-OxPhos complejo IV - Alexa488; anticuerpo anti-p130Cas; anticuerpo anti-P38 MAPK; anticuerpo anti-p44/42 MAPK; anticuerpo anti-P504S clon 13H4; anticuerpo anti-P53; anticuerpo anti-P70 S6K; anticuerpo de péptido de bloqueo de fosfo quinasa anti-P70; anticuerpo anti-Pan cadherina; anticuerpo anti-paxilina; anticuerpo anti-P cadherina; anticuerpo anti-PDI; anticuerpo anti-fosfo AKT; anticuerpo anti-fosfo CREB; anticuerpo anti-fosfo GSK-3-beta; anticuerpo anti-fosfo GSK-3 beta; anticuerpo anti-fosfo H3; anticuerpo anti-fosfo MAPKAPK-2; anticuerpo anti-fosfo MEK; anticuerpo anti-fosfo p44/42 MAPK; anticuerpo anti-fosfo p53; anticuerpo anti-fosfo-NF-kb p65; anticuerpo anti-fosfo-p70 S6 quinasa; anticuerpo anti-fosfo PKC (Pan); anticuerpo anti-fosfo S6 proteína ribosómica; anticuerpo anti-fosfo Src; anticuerpo anti-fosfo-Bad; anticuerpo anti-fosfo-HSP27; anticuerpo anti-fosfo-IKB-a; anticuerpo anti-fosfo-p44/42 MAPK; anticuerpo anti-fosfo-p70 S6 quinasa; anticuerpo anti-fosfo-Rb (Ser807/811) (retinoblastoma); anticuerpo anti-fosfo HSP-7; anticuerpo anti-fosfo-p38; anticuerpo anti-Pim-1; anticuerpo anti-Pim-2; β anticuerpo anti-PKC; β 11 anticuerpo anti-PKC; anticuerpo anti-podocalixina; anticuerpo anti-PR; anticuerpo anti-PTEN; anticuerpo anti-R1; anticuerpo anti-Rb 4H1 (retinoblastoma); anticuerpo anti-R-cadherina; anticuerpo anti-RRM1; anticuerpo anti-S6 proteína ribosómica; anticuerpo anti-S-100; anticuerpo anti-sinaptopodina; anticuerpo anti-sinaptopodina; anticuerpo anti-Sindecin 4; anticuerpo anti-Talin; anticuerpo anti-

Tensin; anticuerpo anti-Tuberlin; anticuerpo anti-uroquinasa; anticuerpo anti-VCAM-1; anticuerpo anti-VEGF; anticuerpo anti-vimentina; anticuerpo anti-ZAP-70; y anti-ZEB.

5 [0074] Fluoróforos que se pueden conjugar a un anticuerpo primario incluyen, pero de forma no limitativa, fluoresceína, rodamina, rojo de Texas, Cy2, Cy3, Cry5, Rojo de VECTOR, ELF.TM. (fluorescencia etiquetada de enzima), Cy, Cy0.5, Cy1, Cry1.5, Cy3, Cy3.5, Cy5, Cy7, FluorX, calceína, calceína-AM, CRYPTOFLUOR.TM.'S, naranja (42 kDa), tangerina (35 kDa), oro (31 kDa), rojo (42 kDa), carmesí (40 kDa), BHMP, BHDMAP, Br-Oregón, amarillo Lucifer, familia de tinte Alexa, N-[6-(7-nitrobenc-2-oxa-1, 3-diazol-4-il)amino]caproil] (NBD), BODIPi.TM., boro dipirrometeno difluoruro, verde de Oregón, MITOTRACKER.TM. Rojo, DiOC.sub.7 (3), DiIC.sub. 18, ficoeritrina, fibobiliproteínas BPE (240 kDa) RPE (240 kDa), CPC (264 kDa) APC (104 kDa), espectro azul, espectro Aqua, espectro verde, espectro oro, espectro naranja, espectro rojo, NADH, NADPH, FAD, colorantes infrarrojo (IR), GDP-ribose cíclica (cGDPR), blanco de calcoflúor, lisamina, umbeliferona, tirosina y triptófano. Una amplia variedad de otra sondas fluorescentes están disponibles de y/o extensivamente se describen en *Handbook of Fluorescent Probes and Research Products* 8.sup.th Ed. (2001), disponible de *Molecular Probes, Eugene*, OR, o, al igual que muchos otros fabricantes.

10 [0075] Además, la amplificación de la señal puede conseguirse usando combinaciones de elementos de enlace específicos, como anticuerpos y anti-anticuerpos, donde los anti-anticuerpos se unen con una región conservada de la sonda de anticuerpo objetivo, particularmente donde los anticuerpos son de distintas especies. Alternativamente, pares receptor-ligantes de enlace específicos, como biotina-estreptavidina, pueden utilizarse, donde el anticuerpo primario se conjuga a un elemento del par y el otro elemento se marca con una sonda detectable. Así, uno construye eficazmente un sándwich de miembros de enlace, donde el primer elemento de enlace se une al componente celular y sirve para proveer una unión secundaria, donde el elemento de la unión secundaria puede incluir o no una etiqueta, que puede además proveer una unión terciaria donde el elemento de la unión terciario proporciona una etiqueta.

20 [0076] Los anticuerpos secundarios, avidina, estreptavidina o biotina, se marcan independientemente con una fracción detectable, que puede ser una enzima que dirige una reacción colorimétrica de un sustrato con un producto de reacción de color sustancialmente no-soluble, un tinte fluorescente (mancha), un tinte luminiscente o un tinte no fluorescente. Ejemplos en lo referente a cada una de estas opciones se enumeran más adelante.

25 [0077] En principio, cualquier enzima que (i) se puede conjugar o unir indirectamente a (p. ej., por medio de avidina conjugada, estreptavidina, biotina, anticuerpo secundario) un anticuerpo primario, y (II) usa un sustrato soluble para proporcionar un producto insoluble (precipitado) podría usarse.

30 [0078] La enzima empleada puede ser, por ejemplo, fosfatasa alcalina, peroxidasa de rábano, beta galactosidasa y/o glucosa-oxidasa; y el sustrato puede respectivamente ser una fosfatasa alcalina, peroxidasa de rábano, beta-galactosidasa o sustrato de glucosa-oxidasa.

35 [0079] Sustratos de fosfatasa alcalina (AP) incluyen, pero de forma no limitativa, sustrato AP-azul (precipitado azul, Zymed catálogo p. 61); sustrato AP-naranja (naranja, precipitado, Zymed), sustrato AP-rojo (rojo, precipitado rojo, Zymed), 5-bromo, 4-cloro, 3-indolifosfato (BCIP sustrato, precipitado turquesa), 5-bromo, 4-cloro, fosfato/nitroazul 3-indolil tetrazolio / iodonitrotetrazolio (BCIP/int sustrato, precipitado amarillo-marrón, Biomeda) 5-bromo, 4-cloro, 3-indolifosfato/nitroazul tetrazolio (BCIP/NBT sustrato, azul/morado) 5-bromo, 4-cloro, fosfato/nitroazul 3-indolil tetrazolio / iodonitrotetrazolio (BCIP/NBT/INT, precipitado marrón, DAKO, rojo rápido (rojo), de color magenta fos (color magenta), AS-BI-fosfato de naftol (NABP)/rápido rojo TR (rojo), AS-BI-fosfato de naftol (NABP)/nuevo fucsina (rojo), AS-MX-fosfato de naftol (NAMP)/nuevo fucsina (rojo), fucsina nueva AP sustrato (rojo), fosfato de p-nitrofenilo (PNPP, amarillo soluble en agua), VECTORTM negro (negro), VECTOR.TM. Azul (azul), VECTOR.TM. Rojo (rojo), Rojo de vega (color rojo frambuesa).

40 [0080] Sustratos de peroxidasa de rábano (HRP, a veces abreviado PO) incluyen, pero de forma no limitativa, 2,2' azino-di-3-etilbenc-tiazolin sulfonato (ABTS, verde, soluble en agua), aminoetil carbazol, 3-amino, 9-etilcarbazol AEC (3A9EC, rojo). Alfa-naftol pironina (rojo), 4-cloro-1-naftol (4C1N, azul, azul-negro), 3,3'-diaminobenzidina tetrahidrocloruro (DAB, marrón) orto-dianisidina (verde), o-fenileno diamina (OPD, marrón, soluble en agua), TACS azul (azul), TACS rojo (rojo), 3,3'.5,5'tetrametilbenzidina (TMB, verde/azul o verde) real BLUE.TM. (azul), VECTOR.TM. VIP (morado), VECTOR.TM. SG (enrarecido azul-gris), y Zymed azul HRP sustrato (azul vivo).

45 [0081] Sustratos de glucosa-oxidasa (ir), incluyen, pero de forma no limitativa, nitroazul tetrazolio (NBT, precipitado morado), tetranitroazul tetrazolio (TNBT, precipitado negro), 2-(4-iodofenil)-5-(4-nitrofenil)- 3-feniltetrazolio cloruro (INT, rojo o precipitado naranja), tetrazolio azul (azul), violeta de nitrotetrazolio (violeta), y 3-(4 2,5-difeniltetrazolio bromuro (MTT, morado). Todos los sustratos de tetrazolio requieren glucosa como un co-sustrato. La glucosa se oxida y la sal de tetrazolio se reduce y forma un formazán insoluble que forma el precipitado de color.

50 [0082] Sustratos de beta-galactosidasa, incluyen, pero de forma no limitativa, 5-bromo-4-cloro-3-indoil beta-D-galactopiranosido, precipitado (X-gal azul). Los precipitados asociados a cada uno de los sustratos enumerados tienen firmas espectrales detectables únicas (componentes).

[0083] La enzima puede dirigirse también para catalizar una reacción de luminiscencia de un sustrato, como, pero no limitado a, luciferasa y aequorina, con un producto de reacción sustancialmente no-soluble capaz de luminiscencia o de dirigir una segunda reacción de un segundo sustrato, como, pero no limitado a, luciferina y ATP o coelenterazina y Ca^{2+} , con un producto de luminiscencia.

[0084] Las siguientes referencias, que se incorporan aquí en su totalidad, proporcionan ejemplos adicionales: J. M. Elias (1990) *Immunohistopathology: A practical approach to diagnosis*. ASCP Press (American Society of Clinical Pathologists), Chicago ; J. F. McGinty. F. E. Bloom (1983) *Double immunostaining reveals distinctions among opioidpeptidergic neurons in the medial basal hypothalamus*. *Brain Res.* 278: 145-153 ; and T. Jowett (1997) *Tissue In situ Hybridization: Methods in Animal Development*. John Wiley & Sons, Inc., New York ; *J Histochem Cytochem* 1997 December 45(12): 1629-1641.

[0085] Preparaciones celulares se pueden someter a hibridación *in situ* (ISH). En general, una sonda de secuencia de ácidos nucleicos se sintetiza y se marca o bien con una sonda fluorescente o con un elemento de un par ligando-receptor, como biotina/avidina, etiquetado con una fracción detectable. Sondas ejemplares y fracciones se describen en la sección precedente. La sonda de secuencia es complementaria a una secuencia de nucleótidos objetivo en la célula. Cada célula o compartimento celular con la secuencia de nucleótidos objetivo puede unirse a la sonda marcada. Sondas usadas en el análisis pueden ser o bien oligonucleótidos de ADN o de ARN o polinucleótidos y pueden contener no sólo nucleótidos de origen natural, sino sus análogos, como dioxigenina dCTP, biotina dCTP 7-azaguanosina, azidotimidina, inosina o uridina. Otras sondas útiles incluyen sondas peptídicas y sus análogos, ADN de gen ramificado, peptidomiméticos, ácidos nucleicos peptídicos y/o anticuerpos. Las sondas deberían tener complementariedad suficiente a la secuencia de ácidos nucleicos objetivo de interés, de modo que enlace estable y específico ocurre entre la secuencia de ácidos nucleicos objetivo y la sonda. El grado de homología requerido para hibridación estable varía con la astringencia de la hibridación. Metodologías convencionales para ISH, hibridación y selección de sonda se describen en Leitch, et al. *In Situ Hybridization: a practical guide*, Oxford BIOS Scientific Publishers, *Microscopy Handbooks* v. 27 (1994); and Sambrook, J., Fritsch, E. F., Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (1989).

[0086] La presente invención se ilustra además por los siguientes ejemplos, que no deberían ser interpretados como limitación de ninguna manera. El contenido de todas las referencias citadas se incorporan por la presente expresamente por referencia.

Ejemplificación

Ejemplo 1: Metodología-Uso de herramientas disponibles públicamente

[0087] Para empezar a dirigir la posibilidad de realización de algoritmos de agrupamiento en los datos de imagen, paquete de software estadístico SPSS (SPSS, Inc., Chicago, IL) se aplicó en estudios con intensidades de píxel representantes de ficheros de datos para cada píxel (DAPI, Cy3, y Cy5) de una imagen de alta resolución seleccionada (p. ej., 2048 X 2048 píxeles). La FIG. 2 muestra las imágenes que se usaron en el primer análisis, representando un control de línea celular teñido con DAPI (núcleos), anti-citoqueratina (Cy3) e anti-integrina alfa-V (Cy5). Para estas imágenes, cada 64° píxel se emitió al fichero de datos. La FIG. 3 muestra un gráfico de dispersión de intensidades de píxel normalizadas para DAPI y Cy3 (normalizado en una escala 0-1 por división por intensidad de píxel máxima). Valores de píxel de agrupamiento normalizado [usando algoritmo de agrupación en dos etapas: distancia de logaritmo de verosimilitud; límite de agrupación = 15] resultaron en dos agrupaciones (tabla 1).

Tabla 1: Distribución de agrupación

	N	% del combinado	% del total
Agrupación 1	58719	89,6	89,6
Agrupación 2	6816	10,4	10,4
Combinado	65535	100,0	100,0
Total	65535		100,0

[0088] La agrupación 1 fue la agrupación de valor más bajo y contuvo ~90% de los píxeles para ambos Cy3 y DAPI. La visualización de las asignaciones de agrupación (FIG. 4) revela que la agrupación 1 son píxeles que representan valores bajos en ambos DAPI y Cy3 con píxeles representantes de agrupación 2 con valor en ambas imágenes. La agrupación 1 así define píxeles que se ajustan al criterio 1 de arriba. Estos son píxeles de fondo que no tienen valor ni en Cy3 ni en DAPI. Por lo tanto, esta agrupación se puede quitar del análisis.

[0089] Para diferenciar los tres criterios posteriores, un métrico denominado "Porcentaje Cy3" fue desarrollado/definido por lo siguiente:

$$\frac{\text{Intensidad de píxel Cy3 (Normalizada)}}{(\text{Intensidad de píxel Cy3 (NORM)} + \text{Intensidad de píxel DAPI (NORM)})}$$

[0090] Este produce un métrico para la intensidad de píxel relativa entre Cy3 y DAPI. Esto podría también realizarse usando DAPI como numerador, donde el método dio resultados equivalentes (FIG. 5). Agrupamiento en este método indicará:

- Alta intensidad DAPI en relación a intensidad Cy3: bajo porcentaje de agrupación Cy3
- Alta intensidad Cy3 en relación a intensidad DAPI: alto porcentaje de agrupación Cy3
- Alta intensidad DAPI y Cy3: porcentaje medio de agrupación Cy3

[0091] La realización de este agrupamiento [usando algoritmo de agrupación en dos etapas: distancia de logaritmo de verosimilitud; límite de agrupación = 3], 3 agrupaciones (tabla 2 y FIG. 6) se observó que representaba, basado en relaciones de ejes, los tres criterios anteriores.

Tabla 2: Distribución de agrupación

	N	% del combinado	% del total
Agrupación 1	1689	25,1	25,1
Agrupación 2	1645	24,5	24,5
Agrupación 3	3388	50,4	50,4
Combinado	6722	100,0	100,0
Total	6722		100,0

[0092] Tratamiento de la "agrupación DAPI" (agrupación 1) y "agrupación Cy3" (agrupación 3) separadamente para calcular una puntuación AQUA® objetivo [sumar intensidades de píxel Cy5 en cada agrupación, dividir por el número total de píxeles en la agrupación, multiplicar por una constante, 100, 000] puntuaciones AQUA® dadas que se ajustan con la biología prevista del objetivo en que la expresión Cy3 se observó como mayor que la expresión DAPI (la integrina se asocia predominantemente a la membrana/citoplasma) como se muestra en la tabla 3.

Como prueba adicional del concepto, señal alta de DAPI en píxeles DAPI frente a píxeles Cy3 (tabla 4) y mayor señal Cy3 en píxeles Cy3 frente a píxeles DAPI (tabla 5) se observaron por estos métodos.

Resultados similares se observaron cuando los distancias algoritmos de Euclides se usaron en lugar de logaritmo de verosimilitud (tablas 6-8).

Tabla 3: Puntuaciones AQUA® resultantes, agrupamiento de logaritmo de verosimilitud

	Píxeles Cy3	Píxeles DAPI	Proporción relevante
AltP	3102	976	3,18
Cy3%-LL	3630	1983	1,83
**Potencia Cy5 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

Tabla 4: Puntuaciones AQUA® resultantes, agrupamiento de logaritmo de verosimilitud

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA
Cy3%-LL	115	291	2,6
**Potencia DAPI sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

| superior)); dividido por el número total de píxeles * 100000. |

Tabla 5: Puntuaciones AQUA® resultantes, agrupamiento de logaritmo de verosimilitud

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA
Cy3%-LL	4014	767	5,28
** Potencia Cy3 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000			

Tabla 6: Puntuaciones AQUA® resultantes, agrupamiento de distancia euclídea

	Píxeles Cy3	Píxeles DAPI	Proporción relevante
AltP	3102	976	3,18
Cy3%-LL	3560	2396	1,49
** Potencia Cy5 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

5

Tabla 7: Puntuaciones AQUA® resultantes, agrupamiento de distancia euclídea

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA
Cy3%-LL	321	535	1,6
** Potencia DAPI sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000			

Tabla 8: Resultados AQUA® resultantes, agrupamiento de distancia euclídea

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA
Cy3%-LL	7551	1163	6,5
** Potencia Cy3 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

10 **Ejemplo 2**

[0093] Datos para cada imagen de píxel se obtuvo y analizó para las imágenes presentadas en la FIG. 7 (Her2 epitelio de cáncer de mama teñido).

15 Primer agrupamiento de paso se realizó como antes para eliminar píxeles de fondo (tabla 9) seguido de agrupamiento del porcentaje métrico Cy3 (tabla 10).

Tabla 9: Distribución de agrupación

	N	% del combinado	% del total
Agrupación 1	3145854	75,0	75,0
Agrupación 2	1048450	25,0	25,0
Combinado	4194304	100,0	100,0
Total	4193404		100,0

20

Tabla 10: Distribución de agrupación

	N	% del combinado	% del total
Agrupación 1	236673	22,6	22,6
Agrupación 2	296558	28,3	28,3
Agrupación 3	515219	49,1	49,1
Combinado	1048450	100,0	100,0
Total	1048450		100,0

5 [0094] Las puntuaciones AQUA resultantes se ajustan a lo esperado en que aumentó la expresión Her2 en Cy3 en relación a DAPI, aumentó DAPI en DAPI en relación a Cy3, y aumentó Cy3 en Cy3 en relación a DAPI (tablas 11-13). Además, el método de agrupamiento excedió el rendimiento del método actual de AQUA® como una proporción alta Cy3/DAPI para Her2 se observó (ver FIG. 7). Her2 es una proteína predominantemente citoplasmática/membranosa.

Tabla 11. Puntuaciones Cy5 AQUA® resultantes

	Píxeles Cy3	Píxeles DAPI	Proporción relevante
AltP	617	346	1,78
Píxel #	304138	292536	
Cy3%-LL	1133	290	3,9
	515219	236673	
**Potencia Cy5 sumada en cada agrupación (agrupación DAPI cluster (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

10

Tabla 12. Puntuaciones DAPI AQUA® resultantes

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA
Cy3%-LL	309	547	1,8
**Potencia DAPI sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000			

Tabla 13. Puntuaciones Cy3 AQUA® resultantes

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA
Cy3%-LL	309	547	1,8
**Potencia DAPI sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000			

15 **Ejemplo 3**

[0095] Datos para cada imagen de píxel se obtuvieron y analizaron para las imágenes presentadas en la FIG. 8 (epitelio de cáncer cervical teñido p53).

20 Primer agrupamiento de paso se realizó como antes para eliminar píxeles de fondo seguidos de agrupamiento del porcentaje métrico Cy3 (tabla 14).

Tabla 14: Distribución de agrupación

	N	% del combinado	% del total
Agrupación 1	1926335	45,9	45,9
Agrupación 2	776690	18,5	18,5
Agrupación 3	1491279	35,6	35,6
Combinado	4194304	100,0	100,0
Total	4194304		100,0

[0096] La nota del primer agrupamiento de paso originó 3 agrupaciones.

No obstante, la agrupación de fondo (agrupación 1) es equivalente a la agrupación de fondo obtenida cuando dos agrupaciones fueron "forzadas" (tabla 15).

5

Tabla 15: Distribución de agrupación

	N	% del combinado	% del total
Agrupación 1	1932895	46,1	46,1
Agrupación 2	2261409	53,9	53,9
Combinado	4194304	100,0	100,0
Total	4194304		100,0

[0097] Las puntuaciones de AQUA® resultantes se ajustan a lo esperado en que aumentó la expresión p53 en DAPI en relación a Cy3, aumentó DAPI en DAPI en relación a Cy3, y aumentó Cy3 en Cy3 en relación a DAPI (tablas 16-18).

10

Tabla 16. Puntuaciones Cy5 AQUA® resultantes

	Píxeles Cy3	Píxeles DAPI	Proporción relevante
AltP	92	440	4,7
Cuenta de píxeles:	985051	455948	
Cy3%-LL	190	735	3,9
Píxeles	1484438	277449	
**Potencia Cy5 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

15

Tabla 17. Puntuaciones Cy3 AQUA® resultantes

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA
Cy3%-LL	1753	477	3,7
**Potencia DAPI sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

Tabla 18. Puntuaciones DAPI AQUA® resultantes

	Cy3	DAPI	Proporción relevante
AltP	NA	NA	NA

Cy3%-LL	376	1793	4,8
**Potencia Cy3 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

[0098] Píxeles de fondo en la imagen de destino pueden también ser un problema. Para dirigir este problema, agrupamiento [usando algoritmo de agrupación en dos etapas: distancia de logaritmo de verosimilitud; límite de agrupación = 15] se realizó en valores de píxel Cy5 para eliminar fondo (tabla 19).

5

Tabla 19: Distribución de agrupación

	N	% del combinado	% del total
Agrupación 1	1879937	82,9	82,9
Agrupación 2	313604	13,8	13,8
Agrupación 3	74428	3,3	3,3
Combinado	2267969	100,0	100,0
Total	2267969		100,0

[0099] Dejando caer la agrupación inferior mientras el fondo mejora la proporción métrica de área (Citoplasma: Núcleo para Her2 y Núcleo: Citoplasma para p53) como se muestra en las tablas 20 y 21 (compare las filas 2 y 3).

10

Tabla 20. P53: COMPARTIMENTO: Puntuaciones Cy5 AQUA® resultantes

	Píxeles Cy3	Píxeles DAPI	Proporción relevante
AltP	53	440	4,7
Cuneta de píxeles	985051	455948	
Cy3%-LL-All	190	735	3,9
Píxeles objetivo	1484438	277449	
Cy3%-LL-Top	34	612	18
Píxeles Comp objetivo			
Cy3 %-LL-Top	797	1405	1,76
Píxeles Comp objetivo (sólo objetivo)			
**Potencia Cy5 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

Tabla 21. Her2: Puntuaciones Cy5 de AQUA® resultantes

	Píxeles Cy3	Píxeles DAPI	Proporción relevante
AltP	617	346	1,78
Píxel #	304138	292536	
Cy3%-LL	1133	290	3,9
	515219	236673	
Cy3%-LL-Top	612	145	4,2
Píxeles Comp objetivo			
Cy3%-LL-Top	2379	2171	1,10

Píxeles Comp objetivo (sólo objetivo)			
**Potencia Cy5 sumada en cada agrupación (agrupación DAPI (agrupación inferior); agrupación Cy3 (agrupación superior)); dividido por el número total de píxeles * 100000.			

Ejemplo 4: Construcción de microarrays tisulares y métodos de coloración inmunohistoquímicos para receptor de estrógeno (ER) y HER2/neu y para análisis de beta-catenina nuclear asociada

5 [0100] Diseño de microarray tisular: muestras fijadas en formalina introducidas en parafina de 345 casos de carcinoma de mama positiva invasivo con nódulos positivos se identificaron. Áreas de carcinoma invasivo, lejos de lesiones *in situ* y epitelio normal, se identificaron y tres núcleos de 0,6 cm por biopsia de punción se tomaron de áreas separadas. Cada núcleo se dispuso en un bloque receptor separado y secciones gruesas de cinco micras se cortaron y procesaron tal y como anteriormente describió (Konenen, J. et al., *Nat. Med.*, 4:844-7, 1987). De forma similar, 310 casos de carcinoma de colon se obtuvieron y dispusieron, tal y como anteriormente describió (Chung, G. et al., *Clin. Cancer Res.* (In Press)). Inmunohistoquímica: portaobjetos de microarray tisular revestido con parafina pre-cortado se desparafinaron y antígeno recuperado por presión de cocción (Kato, A. K. et al., *Biotech. Histochem.*, F2:291-8, 1997). Los portaobjetos se tiñeron con anticuerpos para uno de tres antígenos objetivo: monoclonal anti-E.R. (ratón, Dako Corporation, Carpinteria, CA), policlonal anti-HER2/neu (conejo, Dako Corp.), monoclonal (clon de ratón 14, BD Transduction Labs, San Diego, CA) anti-beta-catenina, o policlonal anti-beta-catenina de conejo. Primarios se incubaron durante toda la noche a 40C. Un anticuerpo de cabra antiratón correspondiente o secundario de anticonejo conjugado a una peroxidasa de rábano acompañada de esqueleto de polímero de dextrano se aplicó luego durante 1 hr (Envision, DAKO Corp.). Antígenos objetivo fueron bien visualizados con un cromógeno de luz visible (diaminobenzidina, Dako) para análisis visual, o un cromógeno fluorescente (Cy-5-tiramida, NEN Life Science Products, Boston, MA). Portaobjetos designados para análisis automatizado se contratiñeron con DAPI para visualización de núcleos, y o bien policlonal anticitoqueratina de conejo (Zymed, So. San Francisco, CA) o anti-alfa-catenina de conejo para distinguir entre células tumorales y estroma al igual que para visualizar la membrana celular. En muchos casos, imágenes sustraídas exponencialmente de histospots teñidos con anti-citoqueratina proporcionaron una etiqueta aceptable para la membrana celular debido a la unión submembranosa de citoqueratina en células tumorales. Estos anticuerpos se visualizaron usando o bien Cy3- o Alexa 488-conjugado de cabra anti-ratón o anticuerpos secundarios de anticonejo (Amersham, Piscataway, NJ and Molecular Probes, Eugene, OR). Portaobjetos designados para inspección visual se contratiñeron con hidróxido de amonio hematoxilina acidificada. Examen manual de microarrays para niveles de E.R., HER2/neu y beta-catenina se han descrito previamente (Snead, D. R. et al., *Histopathology*, 23:233-8, 1993).

Ejemplo 5: Análisis de agrupamiento de AQUA®

35 [0101] La plataforma de análisis cuantitativo automatizada (plataforma AQUA®) es una plataforma de análisis de imagen automatizada basada en fluorescencia usada para el objetivo y la cuantificación reproducible de expresión de proteína en compartimentos celulares y subcelulares específicos que usan el algoritmo de asignación local basada en píxel para la compartimentación de expresión (PLACE). Inherente a PLACE es un paso definido por usuario por el cual umbrales de intensidad de píxel específico deben establecerse manualmente para diferenciar fondo de píxeles específicos de señal dentro de imágenes de compartimento múltiples. Para reducir el tiempo del operador, eliminar la variabilidad operador-a-operador y obtener separación de píxel objetiva y óptima para cada imagen, un algoritmo de agrupamiento dicotómico no supervisado basado en píxel (matemáticas basados en agrupación de k medias) permite para el objetivo y diferenciación automatizada de señal de fondo al igual que diferenciación de píxeles de compartimento específico (p. ej., DAPI frente a Cy3) en una base imagen-por-imagen, se describe aquí. Este nuevo algoritmo se evaluó por cuantificación de receptor de estrógeno específico de compartimento (ER), receptor de progesterona (PR), expresión de Her2 en gran cohorte (n = 682) de pacientes con cáncer de mama con un alto grado de correlación (R = 0,992, 0,987 y 0,990 respectivamente) con análisis AQUA® convencional usando ajustes de umbral manuales como determinado por un operador experimentado. Puntuaciones de expresión obtenidas por AQUA de agrupamiento (c-AQUA) mantuvo relaciones cuantitativas equivalentes como se muestra por agrupamiento de datos comparables y consecuencias de supervivencia asociadas. Aunque cualquier sistema es conveniente para los métodos de la invención, este nuevo algoritmo de agrupamiento realza la eficiencia y objetividad de la presente plataforma de AQUA®.

Métodos

55 [0102] Cohorte: una gran cohorte de cáncer de mama en formato de microarray tisular (TMA) se empleó en estos estudios para probar algoritmos C-AQUA. Esta cohorte de Yale Tissue Microarray Facility (YTMA49) se ha descrito en detalle previamente (Dolled-Filhart, M. et al., *Clin. Cancer Res.*, 12:6459-68, 2006). Brevemente, la cohorte de

mama (n = 652) de carcinoma ductal invasivo recogido en serie del departamento de Patología de la universidad de Yale de 1961 a 1983. También, en el array hay una selección de tejido normal y controles de línea celular. El tiempo de seguimiento es de 12,8 años con una edad media de diagnóstico de 58,1 años. Esta cohorte contiene aproximadamente muestras de medio nódulo positivo y de medio nódulo negativo.

[0103] Tinción de inmunofluorescencia: YTMA49 fue teñido usando un protocolo de inmunofluorescencia indirecto. En resumen, portaobjetos de microarray tisular revestidos con parafina pre-cortada se desparafinaron y antígeno recuperado por recuperación de epítipo inducida por calor en 10 mM Tris (pH 9,0). Usando un autoteñidor (LabVision, Fremont, CA), los portaobjetos se pre-incubaron con tiradores de fondo ((BioCare Medical, Concord, CA). Portaobjetos se incubaron luego con anticuerpos primarios contra ER (Dako, Carpintería, California), clon 1D5, 1:200 dilución), PR (Dako (Carpintería, CA), clon monoclonal de ratón PgR636, 1:1000 dilución), o Her2 (Dako (Carpintería, CA), policlonal de conejo, 1:8000 dilución) y pan-citoqueratina (policlonal de conejo, 1:200 dilución, DAKO, Carpintería, CA) diluido en DaVinci verde (BBioCare Medical, Concord, CA) durante 1 hora a RT. Los portaobjetos se lavaron 3 X 5 min con 1X TBS con Tween-20 0,05%. Anticuerpos secundarios correspondientes se diluyeron en Da Vinci verdes y se incubaron durante 30 minutos a temperatura ambiente. Estos incluyen tanto anticuerpos conjugados directamente a un fluoróforo para anti-citoqueratina (Alexa 555-conjugado cabra de anticonejo; 1:100, Molecular Probes, Eugene, OR), y/o conjugado a una peroxidasa de rábano (HRP) para ER, PR, y Her2 (Dako, Carpintería, CA), Envision de anti-ratón o anti-conejo (Dako, Carpintería, CA)). Portaobjetos se lavaron otra vez 3 X 5 min con TBS con Tween-20 0,05%. Portaobjetos se incubaron con un cromógeno fluorescente (cy-5-tiramida, NEN Life Science Products, Boston, MA), que, como DAB, se activan por HRP y producen la deposición de numerosos colorantes Cy-5 asociados de manera covalente inmediatamente adyacentes al anticuerpo conjugado con HRP secundario. Cy-5 (rojo) se usó porque su valor máximo de emisión está bien fuera del espectro verde-naranja de auto-fluorescencia de tejido. Los portaobjetos para análisis automatizado se incubaron deslizados con DAPI *anti-fade* con un medio de montura (ProLong Gold, Molecular Probes, Eugene, OR).

[0104] Adquisición de imagen: captura de imagen automatizada se realizó por el HistoRx PM-2000™, que ha sido descrito previamente en detalle (Camp, R. et al., *Nat. Med.*, 8:1323-1327, 2002 ; Giltneane, J. & Rimm, D., *Nat. Clin. Pract. Oncol.*, 1:104-11, 2004 ; Cregger, M. et al., *Arch. Pathol. Lab. Med.*, 130:1026-30, 2006). Imágenes digitales de alta resolución, 8 bit (dando como resultado 256 valores de intensidad específicos por píxel de una imagen adquirida) de la coloración de citoqueratina visualizadas con Cy3, DAPI y coloración de objetivo con Cy5 se capturaron y guardaron para cada histospot en el array. Píxeles se escribieron en ficheros de imagen como una función de potencia (potencia (P) = ((Intensidad de píxel/256)/tiempo de exposición) para ayudar a compensar variaciones experimentales en la intensidad de coloración. Imágenes en foco y fuera de foco se tomaron para cada canal para uso futuro con el script de AQUA® tradicional y programa de validación.

[0105] Análisis AQUA® tradicional: análisis AQUA® se realizó. En resumen, una máscara específica de tumor se genera umbralizando manualmente la imagen de un marcador (citoqueratina) que diferencia tumor de estroma circundante y/o leucocitos. Este crea una máscara binaria (cada píxel es o bien 'on' o ' off'). Los niveles de umbral se verificaron y ajustaron si fue necesario, por control de spot de una pequeña muestra de imágenes e imágenes restantes se enmascararon automáticamente usando el único valor de umbral determinado. Todas las manipulaciones de imagen posteriores implicaron sólo información de imagen de la área enmascarada. Después, dos imágenes (uno en-foco, uno fuera-de-foco, tomada 6 µm más-profundo en la muestra) se toman de los identificativos específicos de compartimento y el marcador de objetivo. Un porcentaje de la imagen de fuera-de-foco se sustrae de la imagen en-foco, basado en un análisis píxel-por-píxel de las dos imágenes usando un algoritmo llamado RESA (algoritmo de sustracción exponencial rápida). El algoritmo de RESA mejora la interfaz entre áreas de coloración de intensidad mayor y áreas adyacentes de coloración de intensidad inferior, permitiendo una asignación de píxeles a fondo y compartimentos adyacentes más fácil. Finalmente, el algoritmo PLACE asigna cada píxel en la imagen a un compartimento subcelular específico. Píxeles que no se pueden asignar con precisión a un compartimento dentro de un grado de confianza definido por usuario (100% en este caso) se descartan. Por ejemplo, píxeles en los que las intensidades de píxel citoplasmáticas y nucleares son demasiado similares para ser asignados con precisión se niegan (normalmente comprendiendo <8% de los píxeles totales). Una vez que cada píxel se asigna a un compartimento subcelular (o excluido como se ha descrito anteriormente), la señal en cada ubicación se suma. Estos datos se guardan y pueden expresarse posteriormente o bien como un porcentaje de señal total o como la intensidad de señal media por área de compartimento. Imágenes se validaron según lo siguiente: 1) >2% área tumoral cubierta, 2) Imágenes en el fondo 10% de DAPI y/o Cy3 intensidad total quitado, 3) proporción de puntuación DAPI AQUA® (DAPI medido en núcleo/DAPI medido en el citoplasma) > 1,5.

[0106] Algoritmo de agrupamiento de AQUA® (c-AQUA): máscaras tumorales se aplicaron a las imágenes para excluir cualquier región de no tejido o no tumor y considerar sólo tejido tumoral para análisis (como en el experimento tradicional anteriormente descrito, no obstante, un conjunto fijo de parámetros se usan para todos los experimentos). Este también mejora la sensibilidad y eficiencia computacional del método eliminando un gran número de píxeles no contribuyentes (por ejemplo, en una imagen de alta resolución de un histospot 0,6 mm, tomado a 20X potencia objetiva, <50% de los píxeles representarán tejido). La generación de la máscara tumoral se realiza como se ha descrito anteriormente y usando valores que han sido definidos por examen de varias muestras diferentes.

[0107] La segmentación de la imagen por agrupamiento se realizó usando agrupamiento de k medias basado en distancias euclídeas (Jain, A. et al., *ACM Computing Surveys*, 31:264-323, 1999). Primero, a todos los píxeles se le asignaron características basadas en potencia (ver adquisición de imagen) proporcionado por imágenes de compartimento y se pueden representar como coordenadas (PDAPI, Pcy3). Como resultado de esto, los píxeles podrían presentarse en un diagrama de dispersión 2-D de intensidades de compartimento (figs. 9A-C). El modelo usado para realizar la segmentación de imagen afirma que los píxeles se separarán en dos clasificaciones: 1) aquellos que tienen señal baja en todos los compartimentos evaluados (es decir, fondo), y 2) píxeles con la propiedad de que una etiqueta de compartimento muestra coloración mayor que la de otros (p. ej., mayor intensidad Cy3 que DAPI). Por los datos presentados aquí, para dos compartimentos subcelulares, esto supondría la necesidad de identificar tres centroides de datos. La selección de posiciones de valor inicial es importante porque puede impactar en cuánto tardará el algoritmo de k medias en converger en una solución y prevenir 'intercambio' de centros, que supondría asignación incorrecta. Por el modelo descrito aquí, la agrupación de antecedentes se inicia al origen mientras la citoqueratina y los centros nucleares se inician para sus valores máximos respectivos y cero (p. ej., para el marcador DAPI, el valor inicial es (PDAPI(máx), 0). Los píxeles se asignan luego a cada agrupación basada en distancia euclídea. Valores de centroide de agrupación se calculan luego y miembro de agrupación de píxel es reevaluado. El método se ejecuta iterativamente y termina, después hay convergencia (ningún miembro cambia) o 30 iteraciones.

[0108] Una vez que los centroides de agrupación se han definido, control de error ocurre para detectar condiciones que pueden suponer resultados erróneos. La primera condición es si no hay suficiente señal en bien uno de los compartimentos, que supondrá una segregación del compartimento basado en ruido de fondo. Para detectar esto, un método se implementa que comparará cada compartimento central a la agrupación de fondo central. Si un compartimento central está dentro de un estándar del miembro de primer plano a la agrupación de fondo, el punto de datos será etiquetado y excluido del análisis. El segundo control de error deriva del hecho de que el algoritmo también detecta la cantidad de área que está produciendo una señal viable. Debería esta área ser demasiado pequeña para obtener un buen tamaño de la muestra, el punto debería ser etiquetado y omitido del análisis al estar subrepresentado. Por el presente sistema, que está equipado con un chip 2048 X 2048 CCD para la adquisición de imágenes, el número de píxeles proporcionado debe ser mayor de 210.000 píxeles (5%). La FIG. 9B muestra una imagen que pasa todas las especificaciones de control de calidad, mientras que la FIG. 9C muestra una imagen que falla debido a fallo de distancia de agrupación en el canal Cy3.

[0109] Del examen de los diagramas de dispersión en las figuras 9B y 9C, se puede observar que habrá generalmente píxeles que tienen intensidades superiores a fondo, pero que tienen aportación de intensidad similar para cada canal. Así, una vez que la convergencia se alcanza, un método geométrico se usa luego para además definir la certeza de que un píxel es un elemento de cualquier agrupación. Cada píxel es caracterizado basado en su ubicación en la agrupación y la proximidad a otras agrupaciones. Si el de píxeles de Cy3 y DAPI es menor que B, entonces hay cero certeza en ambos compartimentos y el valor de píxel se fija a cero en ambos compartimentos (FIG. 9A). Este representa fondo en la imagen. Si Cy3 es mayor que B y DAPI es inferior que B, luego hay un 100% de probabilidad para citoplasma y un 0% de probabilidad para nuclear (FIG. 9A). Recíprocamente, si DAPI es mayor que B y Cy3 es inferior que B, luego hay un 100% de probabilidad para nuclear y un 0% de probabilidad para citoplasma (FIG. 9A). Para valores en la región central del diagrama de dispersión que no se asignan definitivamente para cualquier compartimento, una región de función de probabilidad se define por los triángulos ABC y ABD. En estas regiones, los píxeles se asignan para cualquier Cy3 (triángulo ABC) o DAPI (triángulo ABD) exclusivamente. No obstante, su aportación al cálculo total se modifica por su ubicación en los triángulos. A los píxeles en el triángulo ABC se les asigna una probabilidad basada en su proximidad a los vértices. Probabilidades para píxeles en las regiones triangulares se calculan por medio de una función continua y bien definida que varía de 0 a 1 (100% de probabilidad). Mientras un píxel se aproxima a C, el valor se aproxima al 100%, mientras un píxel se aproxima a los vértices A o B (o el segmento de línea conectado a A y B) el valor se aproxima a 0. El triángulo ABD sigue la misma lógica, con valores aproximándose al 100% mientras los píxeles se aproximan al vértice D.

Resultados

[0110] Comparación de algoritmos PLACE para asignación de compartimento: análisis AQUA® y análisis C-AQUA se realizaron en el mismo conjunto de imágenes validadas y adquiridas (n = 388 fuera de un total posible de 652). Aunque los resultados son similares, un operador se requirió para determinar la configuración y los niveles de umbral para generar imágenes, mientras que las imágenes asociadas generadas con C-AQUA se generaron automáticamente, en una forma no supervisada.

[0111] Dos operadores configuraron un experimento AQUA® tradicional y C-AQUA en los mismos datos establecidos.

La configuración para el experimento AQUA® tradicional tardó un promedio de 20 minutos, mientras que el tiempo medio de configuración para C-AQUA fue menos de 2 minutos y no requirió la intervención subjetiva de un operador. Análisis de regresión entre dos operadores para los dos métodos se muestra en la FIG. 10. Aunque altamente correlativo (FIG. 10A; R de Pearson = 0,992, $p < 0.001$ R de Spearman = 0,989, $p < 0,001$), puntuaciones de AQUA® resultante de análisis AQUA® fue, sin embargo, diferente entre operadores, mientras que puntuaciones de

AQUA® generadas con C-AQUA fueron idénticas (FIG. 10B; R de Pearson = 1,000, $p = 0$ R de Spearman = 1,000, $p = 0$).

[0112] Análisis de regresión lineal se realizó para examinar comparaciones totales para todas las imágenes. La comparación de tamaño de compartimento nuclear mostró una correlación altamente significativa entre análisis AQUA® convencional y C-AQUA (FIG. 10A; R de Pearson = 0,779 ($p < 0,001$); R de Spearman = 0,793 ($p < 0,001$)). El tamaño de compartimento citoplasmático se correlacionó también significativamente (FIG. 10B; R de Pearson = 0,923 R de Spearman = 0,914 ($p < 0,001$)). Estos datos no sólo demuestran la capacidad de C-AQUA para establecer compartimentos celulares comparables al de los algoritmos de AQUA®/PLACE convencional, sino que también demuestran que compartimentación C-AQUA no es absolutamente equivalente. Esto se debe al hecho de que la compartimentación se optimiza para cada imagen, en lugar de umbralización universalmente aplicada a través de todas las imágenes como con los algoritmos PLACE de análisis AQUA® convencional.

[0113] La comparación de algoritmos de posición para la expresión de cálculos de puntuación: para confirmar que C-AQUA produce puntuaciones AQUA® equivalentes, análisis AQUA® convencional y análisis C-AQUA se realizaron para tres marcadores biológicos comunes de cáncer de mama; receptor de estrógeno (ER), receptor de progesterona (PR), y Her2. El análisis se realizó en una gran cohorte de cáncer de mama ($n = 607$) de muestras de pacientes con cáncer de mama en el formato TMA. Puntuaciones AQUA® para ambos análisis AQUA® convencional y análisis C-AQUA en compartimentos celulares pertinentes produjeron resultados altamente correlativos en ambos valor y análisis de orden de rango [FIG. 11A (ER, núcleo): R de Pearson = 0,992 ($p < 0,001$) y R de Spearman = 0,993 ($p < 0,001$); FIG. 11B (PR, núcleo): R de Pearson = 0,987 ($p < 0,001$) y R de Spearman = 0,962 ($p < 0,001$); FIG. 11C (Her2, citoplasma/membrana): R de Pearson = 0,990 ($p < 0,001$) y R de Spearman = 0,976 ($p < 0,001$)].

[0114] Comparación de algoritmos PLACE para los resultados de supervivencia: aunque correlación altamente significativa entre puntuaciones AQUA® obtenidas a través de análisis AQUA® convencional y C-AQUA se observó, es importante para demostrar que relaciones de datos equivalentes se mantienen de tal manera que puntos de corte comparables respecto a los resultados (por ejemplo; supervivencia) pueden obtenerse. Para probar esta distancia de logaritmo de verosimilitud no supervisada, agrupamiento se realizó para cada conjunto de puntuaciones de AQUA®. Para ER, dos agrupaciones se identificaron para ambos puntuaciones AQUA® convencional y C-AQUA con 95% de coincidencia total. ER expresión en el cáncer de mama es predictivo de mejor supervivencia. Análisis de supervivencia Kaplan-Meier demostró puntuaciones análisis de agrupaciones de AQUA® obtenidas para ambos análisis AQUA® tradicional y análisis C-AQUA produjeron resultados de supervivencia equivalentes que producen que expresión ER alta predice significativamente un aumento en la supervivencia específica a enfermedad en cinco años [FIG. 12A; análisis AQUA® tradicional: 11,4% de reducción en la supervivencia total (rango logarítmico $p = 0,018$) de 80,9% (ER alto) a 69,5% (ER bajo); C-AQUA: 13,8% de reducción en la supervivencia total (rango logarítmico $p = 0,005$) de 81,6% (ER alto) a 67,8% (ER bajo)].

[0115] Para PR, dos agrupaciones se identificaron para ambos análisis AQUA® convencional y puntuaciones de C-AQUA con 83% de coincidencia total. Expresión PR en el cáncer de mama es también predictivo de mejor supervivencia. Análisis de supervivencia Kaplan-Meier demostró puntuaciones de agrupaciones de AQUA® obtenidas para ambos análisis AQUA® tradicional y análisis C-AQUA produjeron resultados de supervivencia equivalentes en los que expresión alta PR predice significativamente un aumento en la supervivencia específica de enfermedad en cinco años [FIG. 12B; AQUA® tradicional: 12,4% de reducción en la supervivencia total (rango logarítmico $p = 0,021$) de 84,2% (PR alto) a 71,8% (PR bajo); C-AQUA: 14,5% de reducción en la supervivencia total (rango logarítmico $p = 0,001$) de 83,3% (PR alto) a 68,8% (PR bajo)].

[0116] Para Her2, tres agrupaciones se identificaron para ambos análisis AQUA® convencional y puntuaciones C-AQUA con 94% de coincidencia total. Expresión Her2 en el cáncer de mama es predictivo de supervivencia disminuida. Análisis de supervivencia Kaplan-Meier demostró agrupaciones de puntuaciones AQUA® obtenidas para ambos AQUA® tradicional y resultados C-AQUA produjeron resultados de supervivencia equivalentes en los que expresión alta Her2 predice significativamente supervivencia específica de enfermedad disminuida en cinco años [FIG. 12C; AQUA® tradicional: 18,5% de reducción total en la supervivencia total (rango logarítmico $p = 0,022$) de 77,1% (Her2 bajo) a 73,8% (Her2 medio) a 58,6% (Her2 alto); C-AQUA: 24,2% de reducción total en la supervivencia total (rango logarítmico $p = 0,002$) de 77,8% (Her2 bajo) a 73,8% (Her2 medio) a 53,6% (Her2 alto)].

Discusión

[0117] El uso de análisis de imagen avanzado está siendo adoptado rápidamente para facilitar el análisis de muestras en los laboratorios de patología. La automatización asociada, cuantificación y métodos analíticos más objetivos están proveyendo a los patólogos de acceso a cantidades superiores y mejoradas de información. El sistema AQUA® es una plataforma inmunohistoquímica cuantitativa y robusta (IHC), es ahora un sistema de investigación para caracterización de biomarcador y descubrimiento (Berger, A. et al., *Cancer Res.*, 64:8767-8772, 2004; Camp, R. et al., *Cancer Res.*, 63:1445-1448, 2003; Dolled-Filhart, M. et al., *Clin. Cancer Res.*, 9:594-600, 2003; McCabe, A. et al., *J. Natl. Cancer Inst.*, 97:1808-15, 2005). A diferencia del tradicional IHC, el sistema AQUA®

es objetivo y produce estrictamente datos cuantitativos *in situ* de expresión de proteína en una escala continua en vez de datos subjetivos categóricos. El sistema AQUA® prefiere tomar ventaja de la potencia multiplexada de fluorescencia usando marcadores múltiples para diferenciar molecularmente compartimentos celulares y subcelulares en los que cuantificación simultánea de biomarcadores de interés en los tipos de célula específicos y compartimentos subcelulares pueden realizarse.

[0118] El agrupamiento es un método matemático por el cual datos son segregados basados en las relaciones de varias propiedades inherentes a cada medición (Miller, D. et al., *Front. Biosci.*, 13:677-90, 2008), en este ejemplo, las intensidades de mediciones fluorescentes para píxeles dentro de una imagen. El agrupamiento se puede aplicar a imágenes múltiples de un único campo de visión usando intensidades de píxel para adscribir centroides específicos a fondo y señal o incluso niveles de señalización diferentes. La aplicación de estos tipos de algoritmos de agrupamiento para imágenes individuales fluorescentes (p. ej., DAPI o Cy3) permiten segmentación automatizada de fondo desde señal específica para compartimentos celulares individuales, así como los umbrales definidos por usuario cumplidos.

[0119] Un algoritmo de segmentación específica que produce agrupaciones como se ha descrito anteriormente se describe aquí, permitiendo así a un algoritmo automatizado tipo PLACE que retira variabilidad operador-a-operador y optimiza compartimentación de expresión en una base imagen-por-imagen. C-AQUA muestra un alto grado de correlación con análisis AQUA® tradicional como realizado por un operador experimentado.

[0120] La segmentación de imagen de manera que la expresión de proteína se puede cuantificar en compartimentos subcelulares y celulares específicos es un avance sobre otros métodos de cuantificación de imágenes manual en los que estos niveles de expresión se deben identificar manualmente y la expresión etiquetada categóricamente por ojo. Es también un avance sobre PLACE que requiere entrada de operador, facilitado por mejora de imagen (RESA), para definir umbrales de intensidad de píxel específicos para separar señal no específica y fondo de señal específica en cada imagen de compartimento.

[0121] Aunque otras plataformas existen para realización digital de análisis patológico, el sistema cuantitativo AQUA® proporciona ventajas. El criterio de evaluación, no obstante, en el análisis AQUA®, que es segmentación de imagen de no sólo señal específica de fondo, sino también dos o más señales independientes de imágenes independientes, presentaron un único desafío. Se supuso que intensidades de píxel de dos o más imágenes podrían ser agrupadas en paralelo, así no sólo señal de fondo común de eliminación de todas imágenes consultadas sino diferenciación, con un alto grado de confianza, señales específicas de imágenes múltiples que permiten la compartimentación estricta de expresión para píxeles de objetivo específico. Las características provechosas de las imágenes de la capacidad del algoritmo PLACE para segmentar imágenes y compartimentar expresión se mantienen por lo tanto, pero se mejora por medio de un método automatizado para compartimentación y generación de una puntuación AQUA®.

[0122] Hay diferentes ventajas clave para compartimentación por el método C-AQUA. Primero, el tiempo de entrada del operador se reduce significativamente. Con C-AQUA, la necesidad de optimizar umbrales se elimina, reduciendo así el tiempo de preparación a 1-2 minutos. Segundo, un operador no es siempre capaz de analizar con precisión una imagen usando métodos visuales manuales. Aunque operadores entrenados típicamente determinen umbrales equivalentes aproximadamente, cambios en ajustes de umbrales llevan a diferencias en la compartimentación, que pueden llevar potencialmente a diferencia en las puntuaciones AQUA® (FIG. 10; las comparaciones se muestran en la FIG. 13). Dada que la exactitud y la precisión son de suma importancia, especialmente en un ajuste clínico, la mejora de un sistema reproducible y robusto ya a un nivel de automatización completa, y eliminando así todas las fuentes de variabilidad, es de extrema importancia. Tercero, umbrales definidos por usuario deben aplicarse igualmente a través de todas las imágenes dentro de un conjunto que va a ser analizado, mientras que C-AQUA proporciona la oportunidad para optimizar compartimentación en una base de imagen por imagen.

[0123] Es posible aplicar C-AQUA a más de dos imágenes permitiendo la compartimentación óptima y automatizada de tres o más compartimentos molecularmente-definidos. Por ejemplo, la asignación de píxeles para nuclear (DAPI), citoplasma (citoqueratina) y también membrana (pan-cadherina) podría conseguirse con el mismo nivel de exactitud y eficiencia.

Ejemplo 6. Agrupamiento de análisis AQUA® de expresión PTEN

[0124] Un glioblastoma (GBM) TMA (110 GBM muestras de pacientes a 2X redundancia; tiempo de seguimiento medio: 13,2 meses) se manchó para detección de PTEN (clon de ratón monoclonal 138G6, CST #9559) con compartimentos nucleares y no nucleares generalmente como se ha descrito anteriormente, excepto que S 100 se usó como marcador de compartimento no nuclear. Imágenes adquiridas como se ha descrito anteriormente se analizaron usando análisis AQUA® tradicional y análisis de agrupamiento de AQUA®.

[0125] La comparación de puntuaciones PTEN AQUA® (regresión lineal) entre análisis AQUA® y C-AQUA se muestra en la FIG. 14. FIG. 14A) análisis de regresión lineal para expresión PTEN nuclear como determinado por

análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con R de Pearson indicado. FIG. 14B) análisis de regresión lineal para expresión de PTEN citoplasmático como determinado por análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con valores R de Pearson indicados.

5 [0126] Puntuaciones de AQUA® de expresión de PTEN citoplasmático se describen además en la FIG. 15. La correlación de puntuaciones PTEN AQUA® derivadas por ambos métodos para resultados en paciente se muestran en curvas Kaplan Meier en la FIG. 16. Puntuaciones PTEN AQUA® se correlacionaron significativamente con la supervivencia del paciente. Expresión baja de PTEN se asoció a resultados pobres en comparación con expresión alta de PTEN.

10

Ejemplo 7. Análisis de agrupamiento de AQUA® de expresión de ERCC1

15 [0127] Un cáncer pulmonar TMA [insertar detalles TMA] se mancharon para detección de ERCC1 con compartimentos nucleares y no nucleares como se ha descrito anteriormente. Imágenes adquiridas como se ha descrito anteriormente se analizaron usando análisis AQUA® tradicional y análisis de agrupamiento de AQUA®.

20 [0128] La comparación de la puntuación ERCC1 AQUA® (regresión lineal) entre análisis AQUA® y C-AQUA se muestra en la FIG. 17. FIG. 17A) análisis de regresión lineal para expresión de ERCC1 nuclear como determinado por análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con R de Pearson indicado. FIG. 14B) análisis de regresión lineal para expresión de ERCC1 citoplasmático como determinado por análisis AQUA® (eje Y) y análisis C-AQUA (eje X) con valores R de Pearson indicados.

25 [0129] Puntuaciones de AQUA® de expresión de ERCC 1 citoplasmático se describen además en la FIG. 18. La correlación de puntuaciones ERCC1 AQUA® derivadas por ambos métodos para resultados en pacientes se muestran en curvas Kaplan Meier en la FIG. 19. Puntuaciones ERCC1 AQUA® se correlacionaron significativamente con la supervivencia del paciente. Expresión baja de ERCC1 se asoció a resultados relativamente pobres en comparación con expresión alta de ERCC1.

30

Equivalentes

35 [0130] Métodos funcionalmente equivalentes y aparatos dentro del campo de la invención, además de aquellos enumerados aquí, serán evidentes para los expertos en la técnica de las descripciones precedentes. Estas modificaciones y variaciones están destinadas a entrar dentro del campo de las reivindicaciones anexas.

REIVINDICACIONES

1. Método para la identificación de compartimentos biológicos definidos por marcadores por medio de la definición de un primer compartimento biológico definido por marcador en relación a un segundo compartimento biológico definido por marcador presente en una muestra biológica de interés, comprendiendo la comparación de la intensidad en cada uno de los lugares de píxel en una primera imagen de alta resolución del primer compartimento biológico definido por marcador con la intensidad en cada uno de los lugares de píxel correspondientes de una segunda imagen de alta resolución del segundo compartimento biológico definido por marcador, **caracterizado por el hecho de que** la primera imagen de alta resolución ha sido preparada usando un primer agente de imágenes que es específico para el primer compartimento biológico definido por marcador, y donde la segunda imagen de alta resolución ha sido preparada usando un segundo agente de imágenes que es específico para el segundo compartimento biológico definido por marcador, donde diferencias en la intensidad de píxel definen el primer compartimento biológico definido por marcador en relación al segundo compartimento biológico definido por marcador.
2. Método según la reivindicación 1, donde los píxeles de las dos imágenes de alta resolución se representan en un gráfico, donde los ejes del gráfico comprenden la intensidad del primer agente de imágenes y la intensidad del segundo agente de imágenes.
3. Método según la reivindicación 1, comprendiendo además:
- i) asignación de píxeles a una agrupación **caracterizada por el hecho de** intensidad alta de primer agente de imágenes e intensidad baja de segundo agente de imágenes en el primer compartimento;
 - ii) asignación de píxeles a una agrupación **caracterizada por el hecho de** intensidad alta de segundo agente de imágenes e intensidad baja de primer agente de imágenes en el segundo compartimento; y
 - iii) asignación de píxeles a una agrupación **caracterizada por el hecho de** intensidad baja de primer agente de imágenes e intensidad baja de segundo agente de imágenes en el fondo y eliminación de estos píxeles de análisis posterior.
4. Método según la reivindicación 3, donde cualquiera de los pasos de asignación se realiza usando un algoritmo de agrupamiento.
5. Método según la reivindicación 4, donde cualquiera de los pasos de asignación se realiza usando método de agrupamiento de las k medias para determinar un miembro de agrupación para cada píxel.
6. Método según la reivindicación 3, comprendiendo además asignación de píxeles restantes con intensidad del primer agente de imágenes e intensidad del segundo agente de imágenes para el primer compartimento o el segundo compartimento basado en probabilidad.
7. Método según la reivindicación 3, comprendiendo además asignación de los píxeles restantes con intensidad de primer agente de imágenes e intensidad de segundo agente de imágenes ni para el primer compartimento ni para el segundo compartimento.
8. Método según la reivindicación 1, comprendiendo además definición de una máscara definida por la intensidad de píxel del primer y/o segundo agente de imágenes y definición de asignación del compartimento sólo para los píxeles en la máscara.
9. Método según la reivindicación 1, comprendiendo además incubación de la muestra con un primer agente de imágenes que específicamente etiqueta el primer compartimento biológico definido por marcador, un segundo agente de imágenes que específicamente etiqueta un segundo compartimento biológico definido por marcador.
10. Método según la reivindicación 1, donde un biomarcador particular en el primer compartimento biológico definido por marcador en relación al segundo compartimento biológico definido por marcador se localiza y se cuantifica por:
- a) incubación de un tercer agente de imágenes que específicamente etiqueta el biomarcador;
 - b) obtención de una tercera imagen de alta resolución de la muestra etiquetada del tercer agente de imágenes;
 - d) realización de un análisis de agrupamiento en los píxeles para asignar píxeles al primer compartimento definido con marcador o el segundo compartimento definido con marcador; y
 - e) análisis en la tercera imagen de los lugares de píxel asignados a los compartimentos con el fin de identificar los lugares de píxel con un valor de intensidad indicativo del tercer agente de imágenes, y

determinación del valor de la intensidad total de la tercera imagen en los lugares de píxel asignados a cada uno de los primeros y segundos compartimentos, con el fin de localizar y cuantificar el biomarcador en el primer compartimento en relación al segundo compartimento.

- 5 11. Método según la reivindicación 10, donde el análisis de agrupación se realiza usando agrupamiento de k medias reiterativas en la primera y la segunda intensidad de píxel en cada uno de los lugares de píxel para calcular tres centroides usando distancias euclídeas o distancias de logaritmo de verosimilitud.
- 10 12. Método según la reivindicación 11, comprendiendo además:
- 15 i) representación en el gráfico de los lugares de píxel y los centroides calculados donde los del gráfico comprenden la intensidad del primer agente de imágenes y la intensidad del segundo agente de imágenes en lugares de píxel para el primer compartimento y el segundo compartimento;
- 20 ii) conexión de los centroides para definición de un triángulo;
- 25 iii) asignación de los lugares de píxel con una intensidad que no se encuentra en el área del triángulo: (1) al primer compartimento si la intensidad de píxel es sustancialmente indicativa del primer agente de imágenes; (2) al segundo compartimento si la intensidad de píxel es sustancialmente indicativa del segundo agente de imágenes, o (3) para cualquier compartimento si la intensidad de píxel es sustancialmente indicativa de fondo; y
- iv) asignación de los lugares de píxel en el área del triángulo en el primer compartimento o el segundo compartimento basada en un valor correspondiente a la probabilidad de que el píxel se origine desde el primer o el segundo compartimento. 14.

FIG. 1

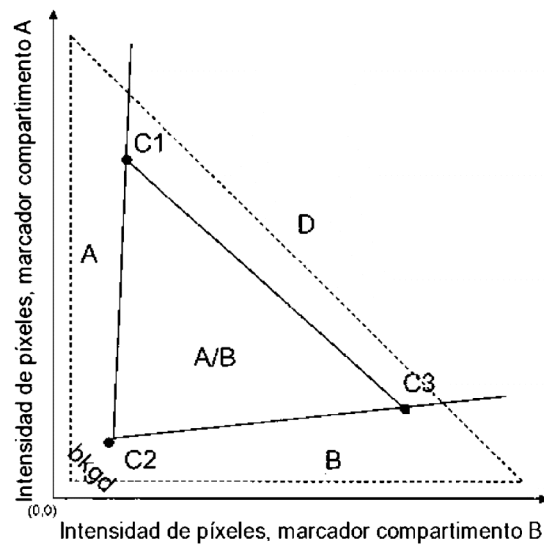


FIG. 2

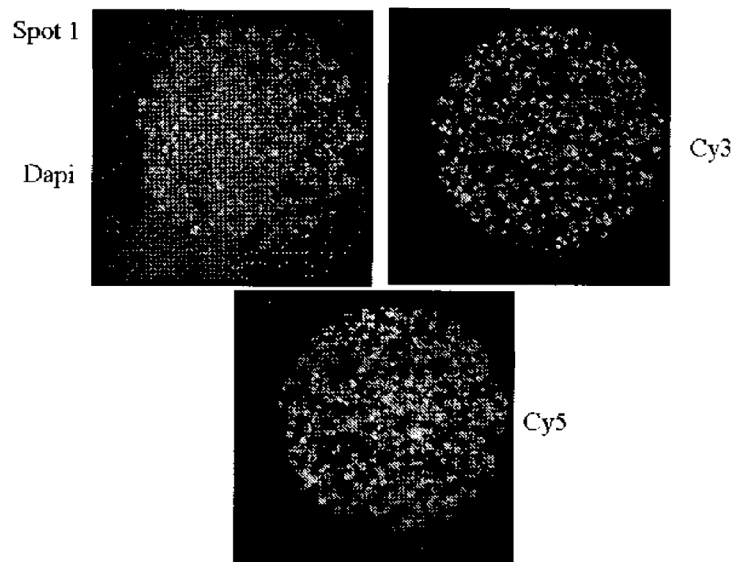


FIG. 3

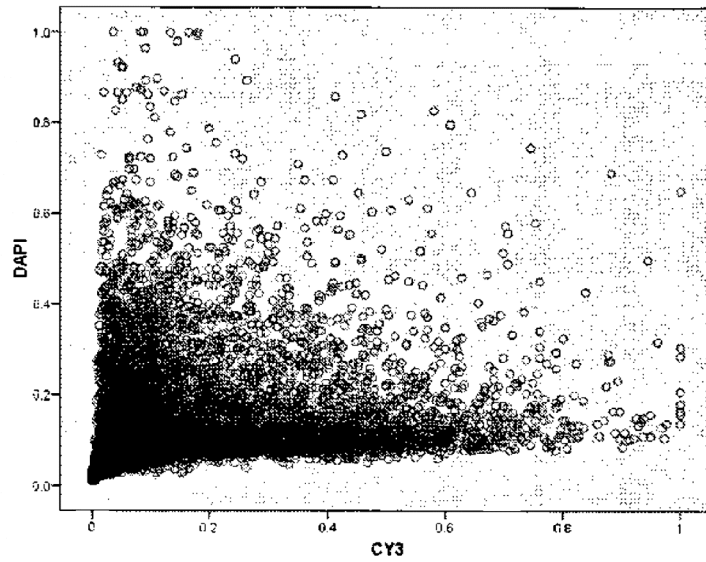


FIG. 4

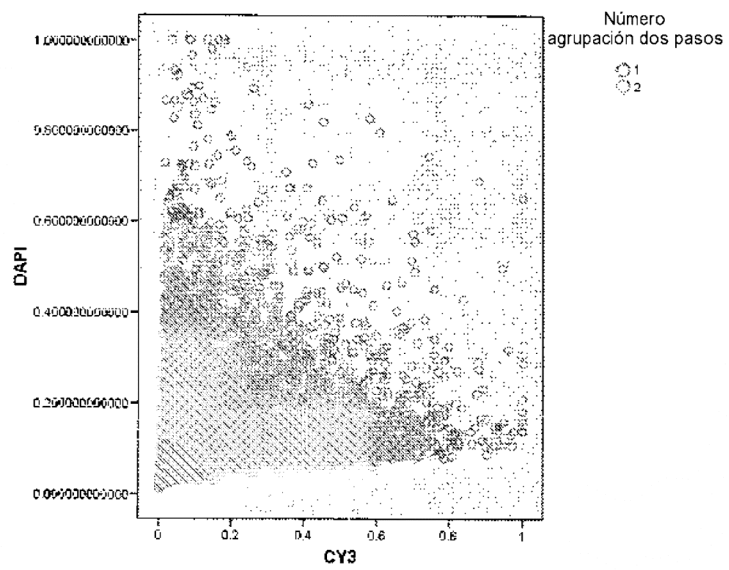


FIG. 5

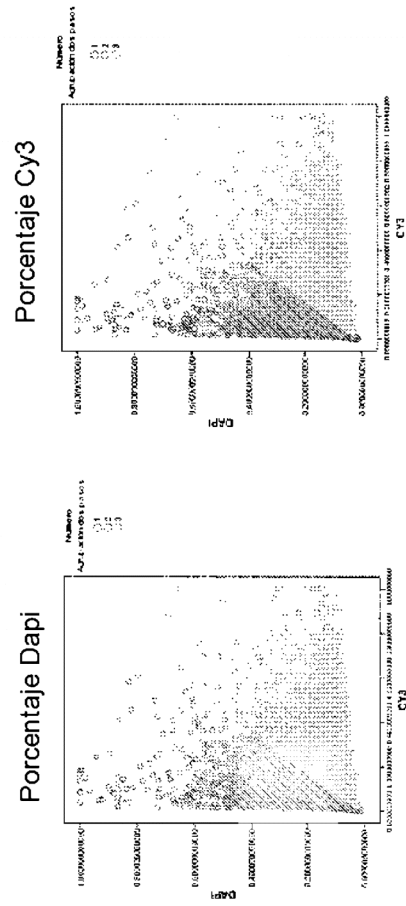


FIG. 6

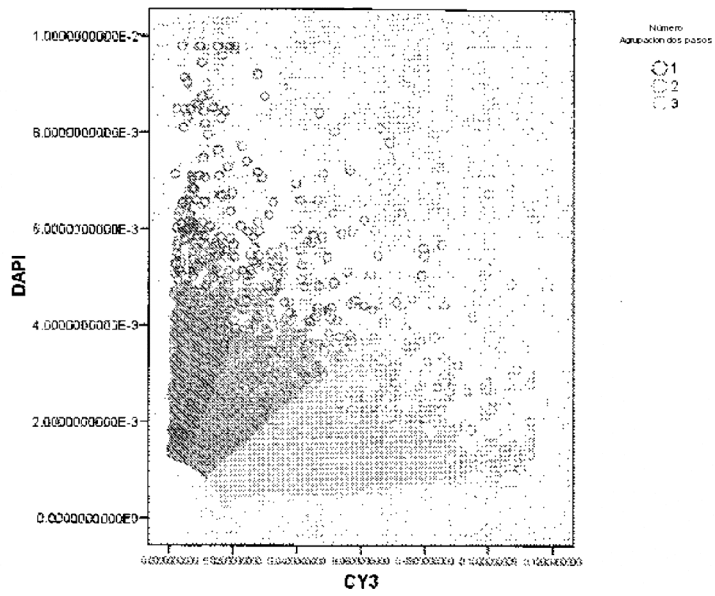


FIG. 7

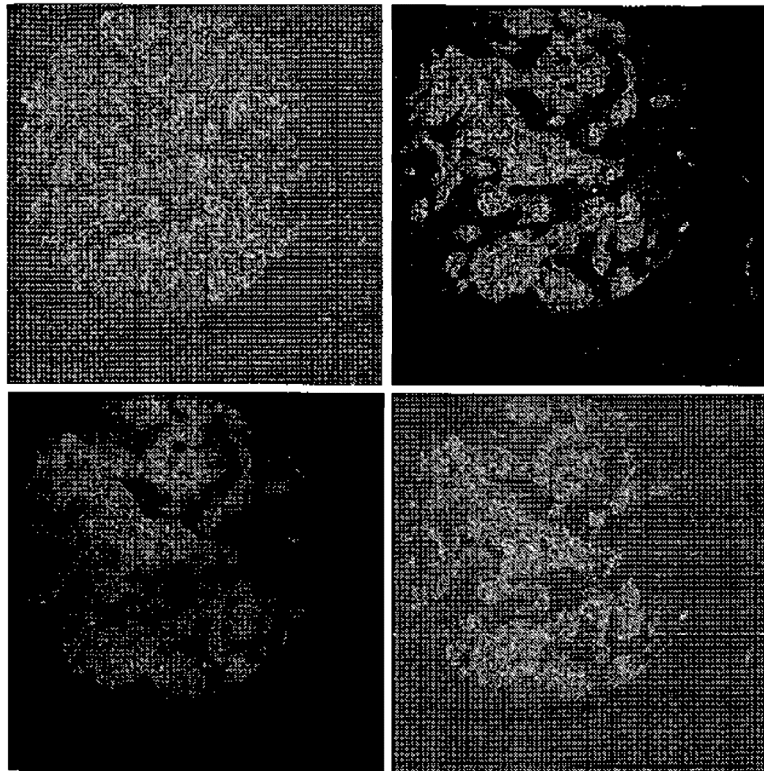


FIG. 8

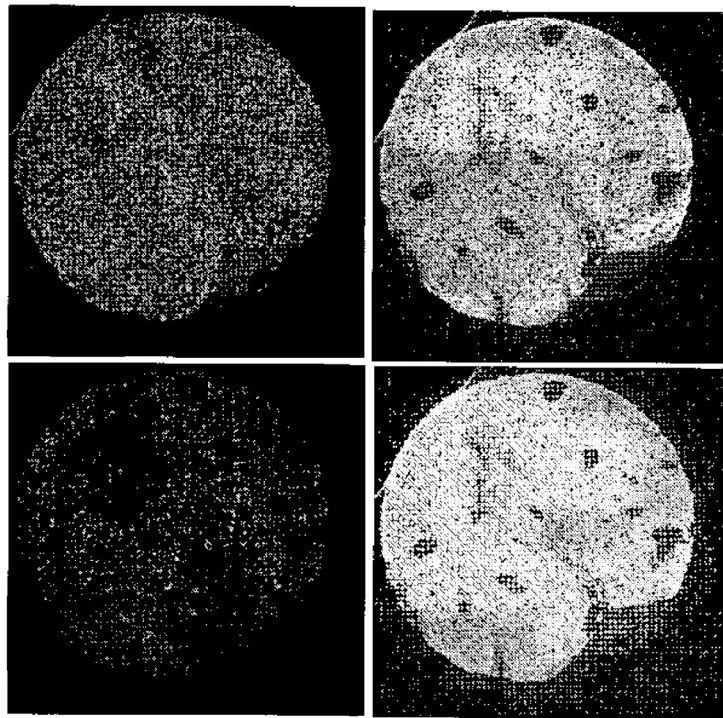
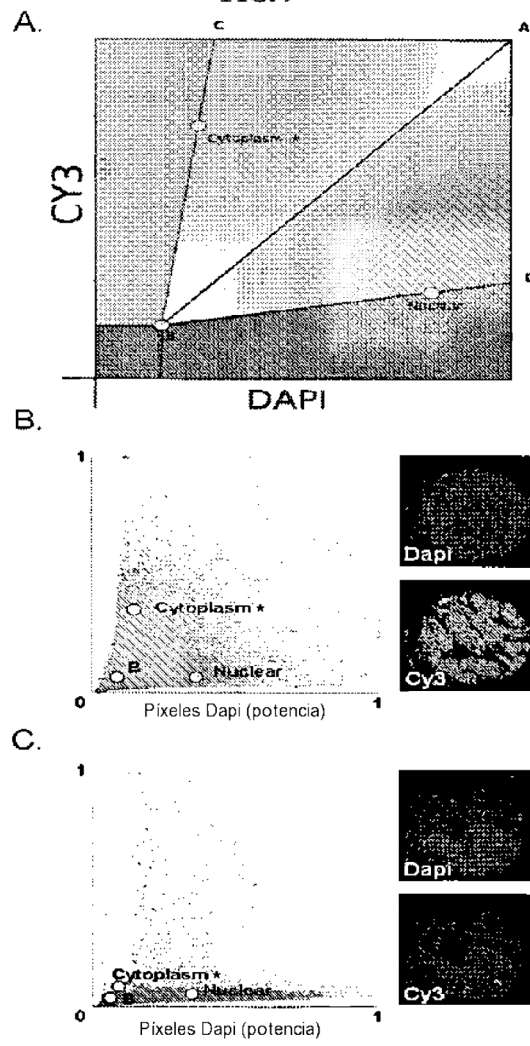


FIG. 9



* Citoplasma

FIG. 10

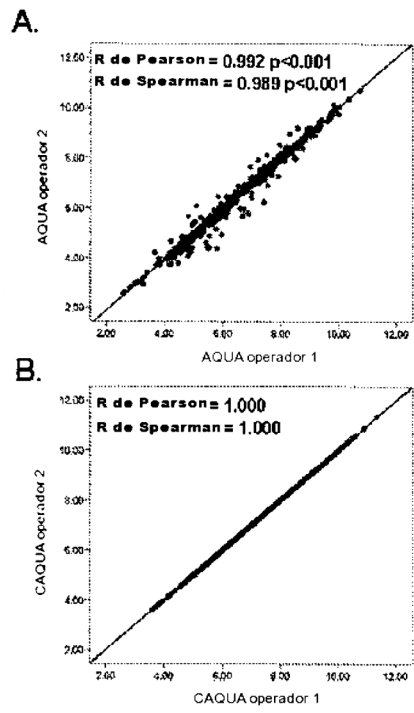


FIG. 11

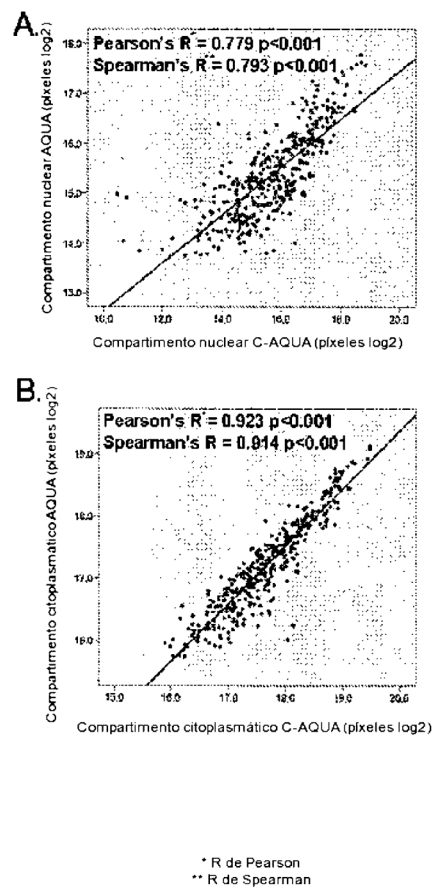


FIG. 12

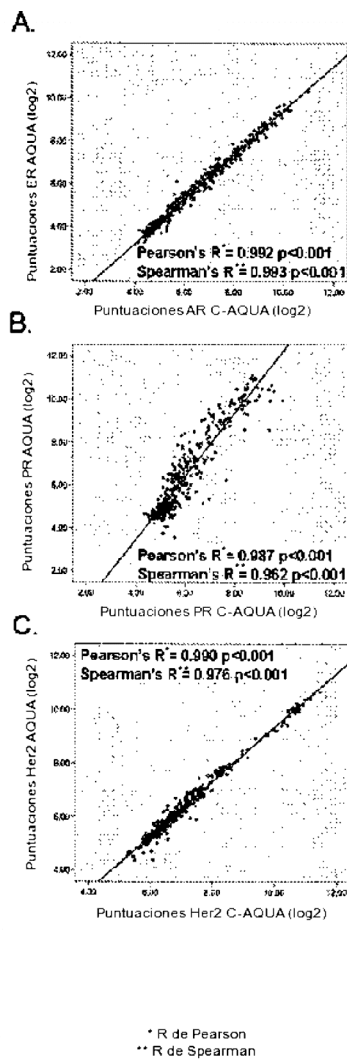
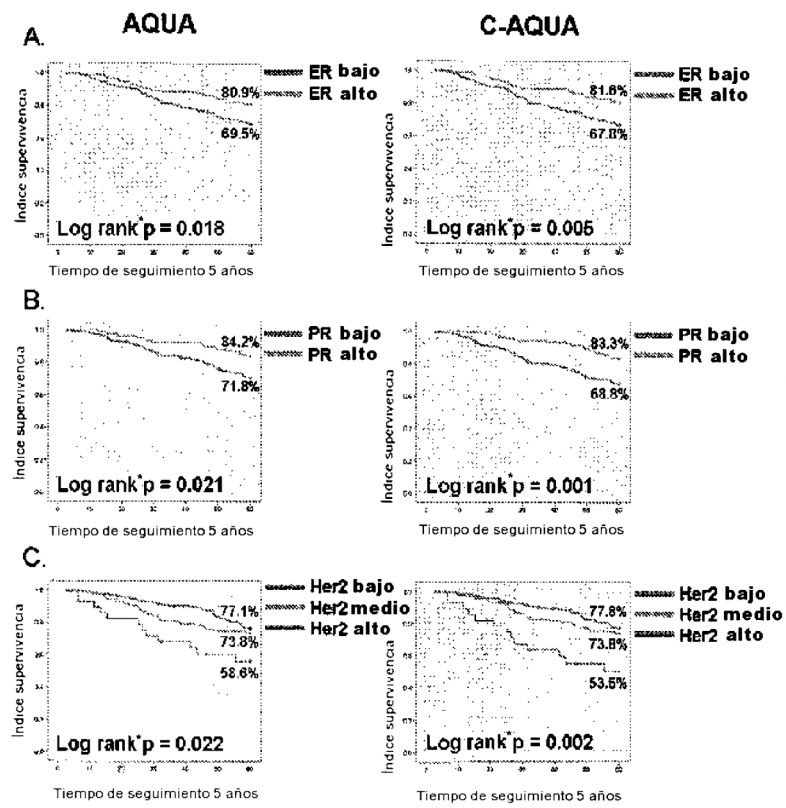


FIG. 13



* Rango logarítmico

FIG. 14

Nuclear

Citoplasmático

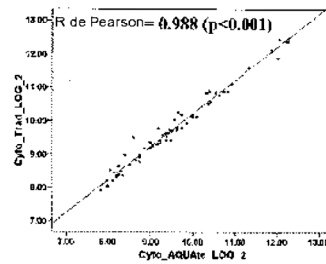
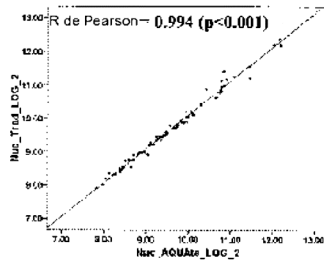
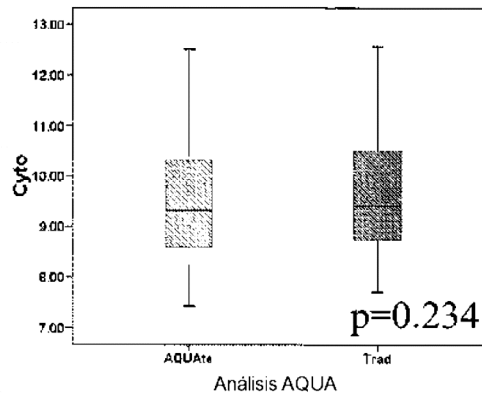


FIG. 15



Estadística

		Cyto_Avg_ Trad	Cyto_Avg_ AQUAte	Cyto_Avg_ Trad LOG	Cyto_Avg_ AQUAte LOG
N	Valido	98	98	98	98
	Faltante	59	59	59	59
Media		1134.87154	997.880689	9.6461	9.4500
Mediana		685.16835	641.032300	9.4203	9.3241
Desviación estándar		1131.341508	1037.3721254	1.15022	1.14747
Diferencia		1279933.607	1076.140.927	1.323	1.317
Oblicuidad		2.292	2.504	.560	.651
Error estándar de oblicuidad		.244	.244	.244	.244
Rango		5883.409	5689.1072	4.86	5.09
Mínimo		209.955	172.1173	7.71	7.43
Máximo		6093.364	5861.2245	12.57	12.52

FIG. 16

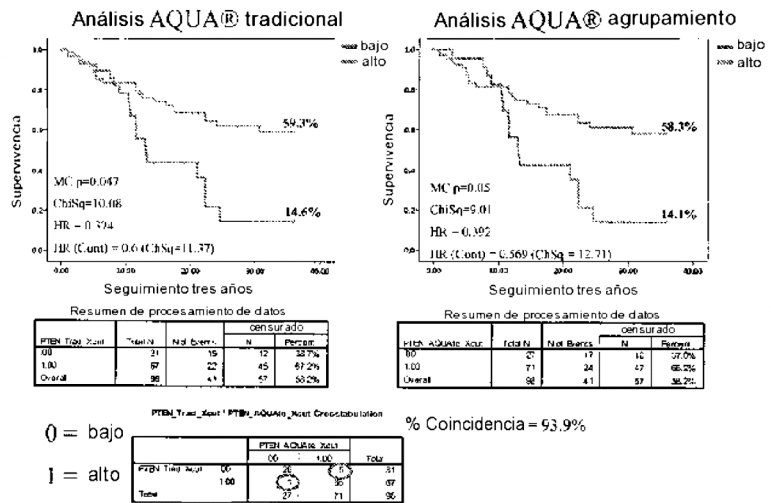


FIG. 17

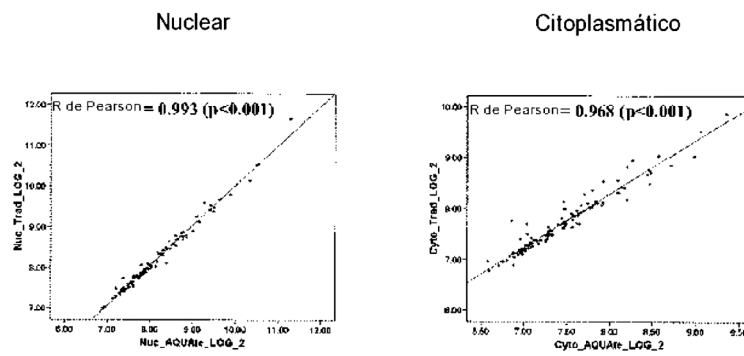
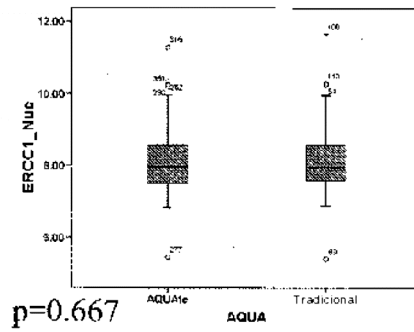


FIG. 18



Estadística

		Nuc_AqL Trad	Nuc_AqL AQUA	Nuc_AqL Trad LOG	Nuc_AqL AQUA LOG
N	Valido	154	154	154	154
	Faltante	54	54	54	54
Media		334,9513	327,2717	8,1237	8,0645
Mediana		248,0382	250,3658	7,9512	7,9678
Desviación estándar		307,55615	274,63550	,78534	,80915
Diferencia		94590,732	75424,658	,817	,655
Oblicuidad		6,154	4,200	,918	,825
Error estándar de oblicuidad		,195	,195	,195	,195
Rango		3204,59	2453,33	8,25	6,83
Mínimo		42,56	43,90	5,41	5,46
Máximo		3247,15	2497,23	11,66	11,29

FIG. 19

