

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 608 956**

51 Int. Cl.:

**G01N 33/543** (2006.01)

**G01N 33/574** (2006.01)

**G01N 33/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.06.2011 PCT/GB2011/000865**

87 Fecha y número de publicación internacional: **15.12.2011 WO11154698**

96 Fecha de presentación y número de la solicitud europea: **09.06.2011 E 11736127 (9)**

97 Fecha y número de publicación de la concesión europea: **05.10.2016 EP 2580592**

54 Título: **Método, serie y uso de los mismos**

30 Prioridad:

**11.06.2010 GB 201009798**

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

**17.04.2017**

73 Titular/es:

**IMMUNOVIA AB (100.0%)  
Medicon Village  
223 81 Lund, SE**

72 Inventor/es:

**BORREBAECK, CARL ARNE KRISTER y  
WINGREN, CHRISTER LARS BERTIL**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

ES 2 608 956 T3

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

**DESCRIPCIÓN**

Método, serie y uso de los mismos

**5 Campo de la invención**

La presente invención se refiere a métodos para el pronóstico de cáncer de mama y biomarcadores y series para el uso de los mismos.

**10 Antecedentes**

El cáncer de mama es uno de los tipos de tumor más común en el mundo occidental, afectando a más de una de cada diez mujeres (1). Una cuarta parte de todos los diagnósticos malignos son para cáncer de mama y desde la final de la década de 1990 la incidencia en todo el mundo, especialmente para el grupo de edad de 45 a 70, ha estado aumentando, aunque ha habido una ligera disminución durante los pasados años. Una explicación para la disminución es que las actividades de detección prematura (mamografía) y las terapias endocrinas y citostáticas han mejorado la supervivencia.

A pesar de esta disminución, el cáncer de mama sigue matando a más de cien mujeres cada día en los Estados Unidos. Además, a pesar de las mejoras en la detección prematura y la comprensión de la base molecular del cáncer de mama, aproximadamente el 30 % de todos los pacientes con cáncer de mama primario desarrollarán recidiva distante de la enfermedad (es decir, metástasis) (2).

Una vez se desarrolla enfermedad metastásica, la posibilidad de cura es muy limitada o inexistente. Para ofrecer mejores tratamientos con una eficacia aumentada y menor toxicidad, deben seleccionarse terapias basándose en las características clínicas y moleculares del tumor. Actualmente, existen posibilidades solamente limitadas para predecir el resultado de la enfermedad usando predictores clínicos tradicionales, tales como el grado histológico y el estado del ganglio linfático (3, 4).

Recientemente, los estudios genómicos han abierto la posibilidad de identificar a pacientes con mal pronóstico respecto a la supervivencia sin recidiva, usando análisis de expresión de una multitud de diferentes genes (5). El documento WO2009/114862 describe un método para controlar la eficacia del tratamiento para un cáncer de mama triple negativo, donde se miden dos o más biomarcadores en dos puntos temporales diferentes, comparados, y se deduce la eficacia del tratamiento a partir de la comparación.

Sin embargo, a pesar de los tremendos esfuerzos, actualmente no es posible realizar evaluaciones de riesgo para metástasis de cáncer de mama usando un ensayo simple de sangre, aunque los recientes desarrollos en la proteómica de afinidad han hecho avanzar el campo de los biomarcadores de cáncer (7-9). El suero es una fuente particularmente valiosa, ya que no solamente es útil para la detección inicial de la enfermedad, sino que también podría usarse para control continuo y análisis de la eficacia de la terapia, que está en contraste con los predictores basados en mediciones realizadas en el momento de la cirugía.

Por tanto, existe una necesidad de métodos mejorados para el pronóstico de cáncer de mama y, en particular, el riesgo de desarrollar recidiva distante de la enfermedad (es decir, metástasis).

Frente a este fondo, los presentes inventores ahora han desarrollado un enfoque proteómico basado en suero para pronosticar el cáncer de mama e identificar una serie de biomarcadores de suero para determinar el riesgo de desarrollar recidiva distante de la enfermedad (es decir, metástasis).

**50 Sumario de la invención**

Por consiguiente, en un primer aspecto, la invención proporciona un método para el pronóstico de cáncer de mama en un sujeto, que comprende las etapas de:

- (a) proporcionar una primera muestra del proteoma del sujeto;
- (b) medir en la primera muestra del proteoma la cantidad de dos o más biomarcadores seleccionados del grupo de biomarcadores enumerados en la Tabla 1;
- (c) proporcionar una muestra adicional del proteoma (por ejemplo, una segunda) del sujeto;
- (d) medir en la muestra adicional del proteoma la cantidad de los dos o más biomarcadores medidos en la etapa (b); y (e) determinar la diferencia entre la cantidad de los dos o más biomarcadores en la primera muestra del proteoma y la muestra adicional del proteoma;

donde la primera muestra del proteoma y la muestra adicional del proteoma son representativas de la composición del proteoma del sujeto en diferentes días,

y donde la diferencia entre la cantidad de los dos o más biomarcadores en la primera muestra del proteoma y la muestra adicional del proteoma (por ejemplo, segunda) es indicativa del riesgo de recidiva y/o de metástasis distante

- de cáncer de mama en el sujeto;  
donde la primera muestra del proteoma del sujeto es representativa de la composición del proteoma del sujeto en cuatro semanas antes o después de la resección del tumor del individuo;  
donde la muestra adicional del proteoma del individuo proporcionada en la etapa (c) es representativa de la  
5 composición del proteoma del sujeto en un punto temporal en 3 a 9 meses desde la resección del tumor del individuo; y  
donde la etapa (b) comprende medir la apolipoproteína A4 (APOA4) y/o la subunidad beta de la ATP sintasa, mitocondrial (ATP5B).
- 10 Por "el pronóstico de cáncer de mama" se entiende determinar el riesgo de recidiva de cáncer de mama en el sujeto. Dicho riesgo puede expresarse como la probabilidad de recidiva del cáncer de mama en un periodo de tiempo específico, por ejemplo, en uno, dos, tres, cuatro, cinco o diez años desde la fecha en que se hace el pronóstico.
- 15 Se apreciará que el sujeto que se está ensayando en el método de la invención normalmente es un ser humano.
- Sin embargo, los métodos también pueden usarse para el pronóstico de un animal tal como un mamífero doméstico o de granja (por ejemplo, un caballo, un cerdo, una vaca, una oveja, un perro o un gato).
- 20 Las etapas (a) y (c) del método de la invención comprenden proporcionar una primera muestra del proteoma y una muestra adicional del proteoma del sujeto a ensayar.
- 25 Por "muestra del proteoma" se entiende una muestra de las proteínas expresadas por las células del sujeto a ensayar. La muestra del proteoma también puede incluir otras moléculas biológicas, o componentes o fragmentos de las mismas, cuya medición puede proporcionar información útil en el pronóstico de cáncer de mama. Por ejemplo, la muestra del proteoma puede incluir restos de proteína o de carbohidrato, o componentes antigénicos o fragmentos de los mismos.
- En una realización, la muestra del proteoma es una muestra soluble del proteoma.
- 30 Preferiblemente, las muestras del proteoma proporcionadas en las etapas (a) y (c) son muestras de sangre. En una realización, la muestra o muestras de sangre están sin fraccionar (es decir, la sangre no se ha separado en sus partes componentes por centrifugación o por cualquier otro medio). En otra realización, la muestra o muestras de sangre están fraccionadas.
- 35 Por tanto, la muestra o muestras del proteoma pueden ser suero o plasma o pueden obtenerse de suero o plasma.
- Preferiblemente, la muestra del proteoma proporcionada en las etapas (a) y/o (c) son muestras de suero o de plasma, derivadas de muestra de sangre tomadas del sujeto a ensayar.
- 40 Las muestras del proteoma proporcionadas en las etapas (a) y (c) pueden prepararse usando métodos bien conocidos en la técnica. Los expertos en la materia apreciarán que la muestra puede estar en un estado nativo o en un formato digerido, dependiendo del método usado para detectar las proteínas en la misma.
- 45 En una realización, una o más de las muestras recogidas de un individuo son muestras de sangre y el resto son muestras de suero. Sin embargo, se prefiere que las muestras recogidas del individuo sean todas muestras de sangre o todas muestras de suero.
- 50 Las muestras del proteoma proporcionadas en las etapas (a) y (c) son representativas de la composición del proteoma del sujeto. Por "representativas de la composición del proteoma del sujeto" se entiende que la muestra del proteoma refleja la composición real del proteoma del sujeto en el momento en que se proporcionó la muestra. La muestra del proteoma puede ser, por ejemplo, una muestra de fluido o tejido tomada directamente del sujeto o un derivado procesado de dicha muestra.
- 55 Las muestras del proteoma proporcionadas en las etapas (a) y (c) se recogen del sujeto en diferentes días y, por tanto, son representativas de la composición del proteoma del sujeto en diferentes días. Por "representativas de la composición del proteoma del sujeto en diferentes días" se entiende que cada muestra adicional o posterior es representativa de la composición del proteoma del sujeto en un punto temporal posterior que la muestra o muestras previas del proteoma proporcionadas.
- 60 En una realización, las etapas (c) y (d) se repiten para proporcionar una o más muestras adicionales del proteoma del sujeto y para medir en las mismas la cantidad de los dos o más biomarcadores medidos en la etapa (b), donde la una o más muestras adicionales del proteoma son representativas de la composición del proteoma del sujeto en diferentes días a partir de la primera muestra del proteoma y de la muestra adicional del proteoma.
- 65 Por tanto, puede proporcionarse una pluralidad de muestras adicionales del proteoma del sujeto, por ejemplo, una tercera, cuarta, quinta, sexta (y así sucesivamente) muestras del proteoma. Se apreciará que la cantidad de

muestras del proteoma proporcionadas del sujeto puede ser suficiente para proporcionar y/o confirmar y/o controlar el pronóstico de cáncer de mama.

5 En una realización, la etapa (e) comprende determinar el cambio log en la cantidad de biomarcador entre la primera muestra del proteoma y la muestra o muestras adicionales del proteoma. Puede determinarse el cambio log en presencia y/o cantidad de dos o más biomarcadores entre las dos o más muestras del proteoma. Esta diferencia en la cantidad de biomarcador en el tiempo se conoce como "velocidad de biomarcador" y puede calcularse usando métodos bien conocidos en la técnica (por ejemplo, véase Makarov et al., 2009, Ann. Rev. Med., 60:139-151).

10 Por tanto, la velocidad de biomarcador refleja la tasa a la cual cambia la concentración de un biomarcador (es decir, la cantidad) en el tiempo en lugar de su concentración/cantidad absoluta o estática.

15 Las etapas de (d) y (b) del método de la invención comprenden medir en las muestras del proteoma la cantidad de dos o más biomarcadores seleccionados del grupo de biomarcadores enumerados en la Tabla 1, que comprende medir la apolipoproteína A4 (APOA4) y/o la subunidad beta de la ATP sintasa, mitocondrial (ATP5B).

20 Por "biomarcador" se entiende una molécula biológica de origen natural, o componente o fragmento de la misma, cuya medición puede proporcionar información útil en el pronóstico de cáncer de mama. Por ejemplo, el biomarcador puede ser una proteína de origen natural o un resto de carbohidrato, o un componente antigénico o un fragmento del mismo.

Por consiguiente, puede medirse la cantidad de apolipoproteína A4 (APOA4). Puede medirse la cantidad de la subunidad beta de la ATP sintasa, mitocondrial (ATP5B).

25 Puede medirse la cantidad de tirosina quinasa de Bruton (BTK). Puede medirse la cantidad de inhibidor de la esterasa del complemento-1 (inhibidor de esterasa de C1). Puede medirse la cantidad del factor B del complemento (factor B). Puede medirse la cantidad de CD40. Puede medirse la cantidad de factor de transcripción de homeodominio CHX10 (CHX10). Puede medirse la cantidad de interleuquina-1 $\alpha$  (IL-1 $\alpha$ ). Puede medirse la cantidad de interleuquina-5 (IL-5). Puede medirse la cantidad de interleuquina-6 (IL-6). Puede medirse la cantidad de interleuquina-7 (IL-7). Puede medirse la cantidad de interleuquina-9 (IL-9). Puede medirse la cantidad de interleuquina-12 (IL-12). Puede medirse la cantidad de interleuquina-13 (IL-13). Puede medirse la cantidad de interleuquina-18 (IL-18). Puede medirse la cantidad del miembro 9 de la familia del dominio TBC1 (KIAA0882). Puede medirse la cantidad de Lewis<sup>x</sup>/CD15. Puede medirse la cantidad de proteína de unión a oxisterol tipo 3 (OSBPL3). Puede medirse la cantidad de proteína quimiotáctica de monocitos-1 (MCP-1). Puede medirse la cantidad de Sialil Lewis<sup>x</sup>. Puede medirse la cantidad de factor-1 $\beta$  de necrosis tumoral (TNF- $\beta$ ).

30 Puede medirse la cantidad de integrina alfa-10. Puede medirse la cantidad de  $\beta$ -galactosidasa. Puede medirse la cantidad de ligando CD40. Puede medirse la cantidad de eotaxina. Puede medirse la cantidad de receptor de péptido-1 tipo glucagón (GLP-1 R). Puede medirse la cantidad de interleuquina-1 $\beta$ . Puede medirse la cantidad de interleuquina-3. Puede medirse la cantidad de interleuquina-8. Puede medirse la cantidad de interleuquina-10. Puede medirse la cantidad de interleuquina-16. Puede medirse la cantidad de MCP-4. Puede medirse la cantidad de miosina-2 (MYOM2). Puede medirse la cantidad de Rantes. Puede medirse la cantidad de factor- $\beta$  de crecimiento transformante (TGF-p1). Puede medirse la cantidad de regulador de transcritos sin sentido 3B (UPF3B). Puede medirse la cantidad de interleuquina-4. Puede medirse la cantidad de interleuquina-1ra.

45 En una realización, la etapa (b) y la etapa (d) comprenden medir la cantidad de cada uno de los biomarcadores enumerados en la Tabla 1(A).

50 En otra realización, la etapa (b) y la etapa (d) comprenden medir la cantidad de uno o más de biomarcadores seleccionados del grupo enumerado en la Tabla 1(B), por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 o los 19 biomarcadores seleccionados del grupo enumerado en la Tabla 1(B).

Por tanto, en una realización de los métodos de la invención, se mide la cantidad de todos los biomarcadores enumerados en la Tabla 1(A) y 1(B) en la etapa (b) y en la etapa (d).

55 En una realización adicional, la etapa (b) y la etapa (d) comprenden medir la cantidad de uno o más de los biomarcadores adicionalmente o como alternativa de la Tabla 1(C), por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o los 15 biomarcadores adicionalmente o como alternativa de la Tabla 1(C).

60 Por tanto, en una realización de los métodos de la invención, se mide la cantidad de todos los biomarcadores enumerados en la Tabla 1 en la etapa (b) y en la etapa (d).

65 Normalmente, la etapa (b) y la etapa (d) se realizan usando dos o más primeros agentes de unión capaces cada uno de unirse a un biomarcador medido en la etapa (d). Normalmente, cada primer agente de unión será capaz de unirse específicamente a un biomarcador diferente medido en la etapa (b). Sin embargo, los expertos en la materia

apreciarán que puede haber alguna duplicación en el biomarcador al que se une el primer agente de unión. Por ejemplo, dos o más de los primeros agentes de unión pueden unirse a diferentes epítomos en el mismo biomarcador (dicha duplicación de antígenos diana puede servir como control interno útil para los métodos de la invención, asegurando que las mediciones del biomarcador son precisas).

5 Preferiblemente, la etapa (b) y la etapa (d) se realizan usando dos o más agentes de unión capaces de unirse a un biomarcador medido en la etapa (b). Cuando se mide un biomarcador enumerado en la Tabla 1 en la etapa (b) y (d), se prefiere que dichas mediciones se realicen usando un agente de unión definido en la sección de Ejemplos (a continuación). Sin embargo, se apreciará que puede usarse igualmente cualquier otro agente de unión con especificidad por el biomarcador relevante.

10 En la etapa (a), la primera muestra del proteoma del sujeto es representativa de la composición del proteoma del sujeto en 4 semanas antes o después de la resección del tumor del individuo. Normalmente, la primera muestra del proteoma del individuo es representativa de la composición del proteoma del sujeto 4 semanas antes de la resección del tumor del individuo, por ejemplo, 3 semanas, 2 semanas, 1 semana, 6 días, 5 días, 4 días, 3 días, 2 días, 1 día o 0 días antes de la resección del tumor del individuo. Preferiblemente, la primera muestra del proteoma del individuo es representativa de la composición del proteoma del sujeto en 1 semana antes de la resección del tumor del individuo. Más preferiblemente, la primera muestra del proteoma del individuo es representativa de la composición del proteoma del sujeto en el mismo día que la resección del tumor del individuo.

20 Por "resección del tumor" se entiende retirada quirúrgica de la totalidad o parte de un tumor. Sin embargo, también se incluye cualquier otro enfoque principal de tratamiento. Dependiendo de la fase y tipo del cáncer de mama, puede no usarse resección del tumor como enfoque de tratamiento principal o puede ser secundario a otro enfoque terapéutico tal como terapia de radiación (es decir, radioterapia), crioterapia, terapia fotodinámica, quimioterapia, terapia hormonal e inmunoterapia. En dichos casos, puede usarse el inicio, la conclusión o un periodo intermedio entre el inicio y la conclusión de la terapia no quirúrgica principal como punto de referencia para recoger muestras del proteoma de un individuo.

25 Por tanto, por "antes o después de la resección del tumor del individuo" también incluye antes o después del inicio, la conclusión o un periodo intermedio entre el inicio y la conclusión del enfoque de tratamiento principal como un punto de referencia posible para recoger muestras de suero o plasma de un individuo.

30 La muestra adicional (por ejemplo, segunda) del proteoma del individuo proporcionada en la etapa (c) es representativa de la composición del proteoma del sujeto en un punto temporal en 3 a 9 meses desde la resección del tumor del individuo, por ejemplo, en 4 a 8 meses, en 5 a 7 meses o a los 6 meses desde la resección del tumor.

35 Por tanto, la muestra adicional del proteoma del individuo puede ser representativa de la composición del proteoma del sujeto en 40 semanas desde la resección del tumor del individuo, por ejemplo, en 39 semanas, 38 semanas, 37 semanas, 36 semanas, 35 semanas, 34 semanas, 33 semanas, 32 semanas, 31 semanas, 30 semanas, 29 semanas, 28 semanas, 27 semanas, 26 semanas, 25 semanas, 24 semanas, 23 semanas, 22 semanas, 21 semanas, 20 semanas, 19 semanas, 18 semanas, 17 semanas, 16 semanas, 15 semanas, 14 semanas, 13 semanas, 12 semanas, 11 semanas, 10 semanas, 9 semanas, 8 semanas, 7 semanas, 6 semanas, 5 semanas, 4 semanas, 3 semanas, 2 semanas, o 1 semana después de la resección del tumor del individuo. Normalmente, sin embargo, la muestra adicional del proteoma es representativa de la composición del proteoma del sujeto en 6 meses desde la resección del tumor del individuo.

40 Cuando se repiten las etapas (c) y (d), la una o más muestras adicionales del proteoma (es decir, tercera, cuarta, quinta, sexta muestras, y así sucesivamente) son representativas de la composición del proteoma del sujeto en un aniversario regular de la resección del tumor del sujeto. Por tanto, las muestras adicionales pueden ser representativas de la composición del proteoma del sujeto a intervalos de tiempos de cada mes o más, por ejemplo, cada dos meses, tres meses, cuatro meses, cinco meses, seis meses, siete meses, ocho meses, nueve meses, diez meses, once meses, un año, dos años, tres años, cinco años o más después de la resección del tumor. Normalmente, sin embargo, la uno o más muestras adicionales del proteoma son representativas de la composición del proteoma del sujeto en 8 semanas antes o después de un aniversario anual de la resección del tumor, por ejemplo, en aproximadamente 7 semanas, 6 semanas, 4 semanas, 3 semanas, 2 semanas o 1 semana antes o después de un aniversario anual de la resección del tumor del individuo.

50 Por tanto, se prefiere que una o más muestras adicionales del proteoma sean representativas de la composición del proteoma del sujeto en o aproximadamente en un aniversario anual de la resección del tumor del individuo.

60 Sin embargo, se apreciará que la muestra o muestras adicionales del proteoma pueden ser representativas de la composición del proteoma del sujeto al menos 1 año después de la resección del tumor en el individuo, por ejemplo, al menos 2 años, 3 años, 4 años, 5 años, 6 años, 7 años, 8 años, 9 años o al menos 10 años después de la resección del tumor en el individuo.

65

En una realización, la resección del tumor es una resección de tumor de cáncer de mama primario. Por "cáncer de mama primario" se entiende que el tumor está localizado en el órgano o tejido donde comenzó el cáncer.

5 En una realización alternativa, la resección del tumor es una resección del tumor de cáncer de mama secundario. Por "cáncer de mama secundario" se entiende que el tumor está localizado en otro órgano o tejido de donde comenzó el cáncer (es decir, un tumor metastásico).

10 En una realización adicional, al sujeto se le tiene que administrar o se le ha administrado quimioterapia y/o radioterapia.

15 En una realización, las etapas (b) y (c) comprenden control continuo del individuo. Por "control continuo" se incluye la realización rutinaria de mediciones en una base ininterrumpida, iterativa (es decir, mediciones repetidas a cortos intervalos de tiempo regulares). También se incluye la realización de mediciones en una base ininterrumpida, no iterativa (es decir, mediciones a tiempo real). El control continuo puede realizarse de forma manual, pero se realiza preferiblemente por un sistema automatizado (por ejemplo, basado en ordenador).

20 En una realización, el riesgo de recidiva de cáncer de mama en un individuo se determina como una probabilidad numérica de recidiva de cáncer de mama, por ejemplo, como un porcentaje de probabilidad de recidiva de cáncer de mama en un periodo de tiempo dado.

Sin embargo, en una realización alternativa, el riesgo de cáncer de mama en un individuo se determina como un alto riesgo o un bajo riesgo. Por tanto, el individuo a ensayar se colocará en un grupo de "alto riesgo" o de "bajo riesgo" con respecto a la recidiva de cáncer de mama (es decir, metástasis).

25 Por "alto riesgo" se entiende que el individuo tiene al menos una probabilidad del 72 % de recidiva del cáncer de mama en 3 años desde la resección del tumor del individuo (en particular, desde la resección de tumor primario del individuo). Por ejemplo, el individuo puede tener una probabilidad del 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % o un 100 % de recidiva del cáncer de mama en 3 años desde la resección del tumor del individuo. Por "bajo riesgo" se entiende que el individuo tiene como mucho una probabilidad del 15 % de recidiva del cáncer de mama en 3 años desde la resección del tumor del individuo (en particular, desde la resección del tumor primario del individuo). Por ejemplo, el individuo puede tener una probabilidad del 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 %, 1 %, 0,5 %, 0,25 %, 0,1 %, 0,01 % o un 0 % de recidiva del cáncer de mama en 3 años desde la resección del tumor del individuo.

35 Generalmente, el riesgo de recidiva de cáncer de mama en un individuo se determina con una AUC ROC de al menos 0,55, por ejemplo, con una AUC ROC de al menos 0,60, 0,65, 0,70, 0,75, 0,80, 0,85, 0,90, 0,95, 0,96, 0,97, 0,98 o con una AUC ROC de al menos 0,99. Preferiblemente, el riesgo de recidiva de cáncer de mama en un individuo se determina con una AUC ROC de al menos 0,85.

40 Normalmente, el riesgo de recidiva de cáncer de mama en un individuo se determina usando una máquina de soporte vectorial (SVM), tal como las disponibles en <http://cran.r-project.org/web/packages/e1071/index.html> (por ejemplo, e1071 1.5-24). Sin embargo, también puede usarse cualquier otro medio adecuado.

45 Las máquinas de soporte vectorial (SVM) son un conjunto de métodos de aprendizaje supervisado relacionados usados para la clasificación y regresión. Dado un conjunto de ejemplos de instrucción, cada uno marcado como perteneciente a uno de dos categorías, un algoritmo de instrucción SVM construye un modelo que predice si un nuevo ejemplo cae en una categoría o en otra. De forma intuitiva, un modelo SVM es una representación de los ejemplos como puntos en el espacio, mapeados de modo que los ejemplos de las categorías diferentes están divididos por un claro hueco que es lo más ancho posible. Los nuevos ejemplos después se mapean en el mismo espacio y se predice que pertenecen a una categoría basándose en qué sitio del hueco caen.

50 De forma más formal, una máquina de soporte vectorial construye un hiperplano o conjunto de hiperplanos en un espacio de alta dimensión o dimensión infinita, que puede usarse para la clasificación, regresión y para otras tareas. De forma intuitiva, se consigue una buena separación por el hiperplano que tiene la distancia más larga hasta los puntos de datos de instrucción más cercanos de cualquier clase (el llamado margen funcional), ya que en general, cuanto mayor sea el margen menor será el error de generalización del clasificador. Para más información sobre SVM, véase, por ejemplo, Burges, 1998, Data Mining and Knowledge Discovery, 2:121-167.

60 En una realización de la invención, la SVM se "instruye" antes de realizar los métodos de la invención usando muestras del proteoma de los sujetos asignados a grupos de pacientes conocidos (concretamente, aquellos pacientes en que el cáncer de mama no sufrió recidiva durante un periodo de tiempo dado y aquellos pacientes en que el cáncer de mama sufrió recidiva sobre el periodo de tiempo dado). Ejecutando dichas muestras de instrucción, la SVM es capaz de aprender los perfiles de biomarcadores que están asociados con un riesgo de recidiva de cáncer de mama. Una vez completado el proceso de instrucción, la SVM entonces tiene la capacidad de si una muestra del proteoma ensayada es o no de un sujeto que está en riesgo de recidiva de cáncer de mama.

Sin embargo, este procedimiento de instrucción puede evitarse programando la SVM con los parámetros de instrucción necesarios. Por ejemplo, el riesgo de recidiva de cáncer de mama en un sujeto puede determinarse de acuerdo con los parámetros SVM conocidos detallados en la Tabla 2, basándose en la medición de todos los biomarcadores enumerados en la Tabla 1(A) y 1(B).

Los expertos en la materia apreciarán que los parámetros SVM adecuados pueden determinarse por cualquier combinación de los biomarcadores enumerados en la Tabla 1 instruyendo a una máquina SVM con la selección apropiada de datos (es decir, mediciones de biomarcadores en muestras del proteoma de grupos de pacientes conocidos).

En una realización, el riesgo de recidiva de cáncer de mama se determina en combinación con al menos otro método de determinación del riesgo de recidiva del cáncer de mama tal como en combinación con uno o más perfiles de expresión de ácidos nucleicos asociados con el riesgo de recidiva de cáncer de mama. Por ejemplo, puede usarse el perfil de expresión de ácidos nucleicos definido en van't Veer et al., 2002, *Nature*, 415:530-536 y/o en Paik et al., 2004, *N. Engl. J. Med.*, 351:2817-2826.

Como alternativa de forma adicional, el riesgo de recidiva de cáncer de mama puede determinarse en combinación con marcadores clínicos convencionales. Por "marcadores clínicos convencionales" se entiende la estratificación de los pacientes en diferentes grupos de pronóstico basándose en características clínicas tales como estado menopáusico, estado de receptores de estrógenos, estado de progesterona, tipo ductal/lobular, tamaño del tumor, fase/estado de los ganglios linfáticos y grado histológico (en oposición a la expresión de firmas complejas de expresión). Un ejemplo común de un índice pronóstico usando marcadores clínicos convencionales es el índice pronóstico de Nottingham (NPI) que se obtuvo de un estudio de múltiples variables retrospectivo, es decir (en algún grado) capaz de predecir la supervivencia en pacientes con cáncer de mama. Para más información, véase, por ejemplo, Eden et al., 2004, *European Journal of Cancer*, 40(12):1837-1841.

En una realización adicional, el dos o más agentes de unión capaces de unirse a un biomarcador definido en la Tabla 1 comprenden o consisten en un anticuerpo o un fragmento de unión a antígeno del mismo. Preferiblemente, el anticuerpo o fragmento de unión a antígeno del mismo es un anticuerpo monoclonal o un fragmento del mismo. Más preferiblemente, el anticuerpo o fragmento a antígeno del mismo es un anticuerpo recombinante o fragmento de unión a antígeno del mismo.

El término "anticuerpo" incluye cualquier anticuerpo sintético, anticuerpo recombinante o híbrido de anticuerpo, tal como, aunque sin limitación, una molécula de anticuerpo de cadena sencilla producida por presentación en fagos de regiones variables y/o constantes de cadena ligera y/o pesada de inmunoglobulina, u otras moléculas inmunointeractivas capaces de unirse a un antígeno en un formato de inmunoensayo que es conocido para los expertos en la materia.

También se incluye el uso de agentes de unión tipo anticuerpo, tales como aficuerpos y aptámeros.

Los métodos para obtener anticuerpos, o fragmentos de unión a antígenos derivados de los mismos, con especificidad para un antígeno dado (tal como los biomarcadores de la Tabla 1) son bien conocidos en la técnica. Se encuentra una revisión general de las técnicas implicadas en la síntesis de fragmentos de anticuerpo que retienen sus sitios de unión específica en Winter y Milstein (1991) *Nature* 349, 293-299.

Adicionalmente, o como alternativa, una o más de las primeras moléculas de unión pueden ser un aptámero (véase Collett et al., 2005, *Methods* 37:4-15).

Convenientemente, se explora una biblioteca de anticuerpos (tal como la biblioteca n-CoDeR de Bioinvent International AB; véase Carlsson y Söderlind, 2001, *Expert Rev Mol Diagn.* 1:102-8.) para identificar anticuerpos, o fragmentos de unión a antígeno derivados de los mismos, con especificidades de unión deseadas. Por tanto, pueden usarse bibliotecas moleculares tales como bibliotecas de anticuerpos (Clackson et al., 1991, *Nature* 352, 624-628; Marks et al., 1991, *J Mol Biol* 222(3): 581-97), bibliotecas de péptidos (Smith, 1985, *Science* 228(4705): 1315-7), bibliotecas de ADNc expresados (Santi et al., (2000) *J Mol Biol* 296(2): 497-508), bibliotecas sobre otras estructuras diferentes a la región flanqueante de anticuerpos tales como aficuerpos (Gunneriusson et al., 1999, *Appl Environ Microbiol* 65(9): 4134-40) o bibliotecas basadas en aptámeros (Kenan et al., 1999, *Methods Mol Biol* 118, 217-31) como fuente a partir de la cual se seleccionan moléculas de unión que son específicas para un motivo dado para su uso en los métodos de la invención.

Las bibliotecas moleculares pueden expresarse *in vivo* en células procariotas (Clackson et al., 1991, *op. cit.*; Marks et al., 1991, *op. cit.*) o en células eucariotas (Kieck et al., 1999, *Proc Natl Acad Sci USA*, 96(10):5651-6) o pueden expresarse *in vitro* sin implicación de células (Hanes y Pluckthun, 1997, *Proc Natl Acad Sci USA* 94(10):4937-42; He y Taussig, 1997, *Nucleic Acids Res* 25(24):5132-4; Nemoto et al., 1997, *FEBS Lett*, 414(2):405-8).

En casos donde se usan bibliotecas basadas en proteínas, los genes que codifican las bibliotecas de moléculas potenciales de unión a menudo se empaquetan en virus y la molécula potencial de unión se presenta en la superficie

del virus (Clackson et al., 1991, *supra*; Marks et al., 1991, *supra*; Smith, 1985, *supra*).

Quizá el sistema de presentación más habitualmente usado es bacteriófago filamentosos que presenta fragmentos de anticuerpo en sus superficies, expresándose los fragmentos de anticuerpo como una fusión a la proteína minoritaria de envuelta del bacteriófago (Clackson et al., 1991, *supra*; Marks et al., 1991, *supra*). Sin embargo, otros sistemas adecuados para la presentación incluyen el uso de otros virus (documento EP 39578), bacterias (Gunneriusson et al., 1999, *supra*; Daugherty et al., 1998, *Protein Eng* 11(9):825-32; Daugherty et al., 1999, *Protein Eng* 12(7):613-21), y levaduras (Shusta et al., 1999, *J Mol Biol* 292(5):949-56).

Además, se han desarrollado sistemas de expresión que utilizan la unión del producto polipeptídico a su ARNm codificante en los llamados sistemas de presentación de ribosoma (Hanes y Pluckthun, 1997, *supra*; He y Taussig, 1997, *supra*; Nemoto et al., 1997, *supra*), o como alternativa la unión del producto polipeptídico al ADN codificante (véase la patente de Estados Unidos n.º 5.856.090 y el documento WO 98/37186).

Los dominios pesado variable ( $V_H$ ) y ligero variable ( $V_L$ ) del anticuerpo están implicados en el reconocimiento del antígeno, un hecho reconocido por primera vez por los primeros experimentos de digestión con proteasa. Se encontró confirmación adicional por "humanización" de anticuerpos de roedor. Los dominios variables de origen de roedor pueden fusionarse con dominios constantes de origen humano de modo que el anticuerpo resultante retenga la especificidad antigénica del anticuerpo precursor de roedor (Morrison et al., (1984) *Proc. Natl. Acad. Sci. USA* 81,6851-6855).

La especificidad antigénica se confiere por dominios variables y es independiente de los dominios constantes como se sabe por los experimentos que implican la expresión bacteriana de fragmentos de anticuerpo, conteniendo todos uno o más dominios variables. Estas moléculas incluyen moléculas tipo Fab (Better et al., (1988) *Science* 240, 1041); moléculas Fv (Skerra et al., (1988) *Science* 240, 1038); moléculas Fv de cadena sencilla (scFv) donde los dominios asociados  $V_H$  y  $V_L$  se unen mediante un oligopéptido flexible (Bird et al., (1988) *Science* 242, 423; Huston et al., (1988) *Proc. Natl. Acad. Sci. USA* 85, 5879) y anticuerpos de un único dominio (dAb) que comprenden dominios V aislados (Ward et al., (1989) *Nature* 341,544). Se encuentra una revisión general de las técnicas implicadas en la síntesis de fragmentos de anticuerpo que retienen sus sitios de unión específica en Winter y Milstein (1991) *Nature* 349, 293-299.

Por tanto, el anticuerpo o el fragmento de unión a antígeno puede seleccionarse del grupo que consiste en anticuerpos intactos, fragmentos Fv (por ejemplo, Fv de cadena sencilla [scFv] y Fv unido por disulfuro), fragmentos tipo Fab (por ejemplo, fragmentos Fab, fragmentos Fab' y fragmentos F(ab)<sub>2</sub>), dominios variables individuales (por ejemplo, dominios  $V_H$  y  $V_L$ ) y anticuerpos de dominio (dAb, incluyendo formatos individuales y dobles [es decir, dAb-enlazador-dAb]).

Por "moléculas scFv" se entiende moléculas donde los dominios asociados  $V_H$  y  $V_L$  están unidos mediante un oligopéptido flexible.

Las ventajas de usar fragmentos de anticuerpo, en lugar de anticuerpos completos, son diversas. El tamaño más pequeño de los fragmentos puede conducir a propiedades farmacológicas mejoradas, tales como mejor penetración de tejido sólido. Las funciones efectoras de los anticuerpos completos, tales como unión al complemento, se eliminan. Los fragmentos de anticuerpo Fab, Fv, scFv y dAb pueden expresarse todos y secretarse en *E. coli*, permitiendo de ese modo la fácil producción de grandes cantidades de dichos fragmentos.

Los anticuerpos completos, y los fragmentos F(ab)<sub>2</sub> son "bivalentes". Por "bivalente" se entiende que dichos anticuerpos y fragmentos F(ab)<sub>2</sub> tiene dos sitios de combinación con antígeno. En contraste, los fragmentos Fab, Fv, scFv y dAb son monovalentes, teniendo solamente un sitio de combinación con antígeno.

Los anticuerpos pueden ser monoclonales o policlonales. Los anticuerpos monoclonales adecuados pueden prepararse por técnicas conocidas, por ejemplo, las descritas en "Monoclonal Antibodies: A manual of techniques", H Zola (CRC Press, 1988) y en "Monoclonal Hybridoma Antibodies: Techniques and applications", J G R Hurrell (CRC Press, 1982), ambas cuales se incorporan en este documento por referencia.

Cuando se seleccionan moléculas potenciales de unión de las bibliotecas, habitualmente se emplean uno o más péptidos de selección que tienen motivos definidos. Los restos de aminoácido que proporcionan estructura, disminuyendo la flexibilidad en el péptido o cadenas laterales cargadas, polares o hidrófobas que permiten la interacción con la molécula de unión, pueden usarse en el diseño de motivos para péptidos de selección. Por ejemplo:

- (i) La prolina puede estabilizar una estructura peptídica ya que su cadena lateral se une tanto al carbono alfa como al nitrógeno;
- (ii) La fenilalanina, tirosina y triptófano tienen cadenas laterales aromáticas y son muy hidrófobos, mientras que la leucina y la isoleucina tienen cadenas laterales alifáticas y también son hidrófobos;
- (iii) La lisina, arginina e histidina tienen cadenas laterales básicas y estarán cargados positivamente a pH neutro,

mientras que el aspartato y el glutamato tienen cadenas laterales ácidas y estarán cargados negativamente a pH neutro;

(iv) La asparagina y la glutamina son neutros a pH neutro pero contienen un grupo amida que puede participar en enlaces de hidrogeno;

5 (v) Las cadenas laterales de serina, treonina y tirosina contienen grupos hidroxilo, que pueden participar en enlaces de hidrógeno.

Normalmente, la selección de moléculas de unión puede implicar el uso de tecnologías de serie y sistemas para analizar la unión a manchas correspondientes a tipos de moléculas de unión.

10 Por tanto, se apreciará que el anticuerpo o fragmento de unión a antígeno puede seleccionarse del grupo que consiste en anticuerpos intactos, fragmentos Fv (por ejemplo, Fv de cadena sencilla o Fv unido por disulfuro), fragmentos tipo Fab (por ejemplo, fragmentos Fab, fragmentos Fab' y fragmentos F(ab)<sub>2</sub>, dominios variables individuales (por ejemplo, dominios V<sub>H</sub> y V<sub>L</sub>) y anticuerpos de dominio (por ejemplo, dAb, incluyendo formatos individuales y dobles (es decir, dAb-enlazador-dAb)).

15 Como alternativa, un primer agente de unión puede comprender o consistir en agentes de unión tipo anticuerpo, por ejemplo, anticuerpos o aptámeros.

20 En una realización, los dos o más biomarcadores medidos en la etapa (b), preferiblemente todas las moléculas orgánicas de la muestra se marcan con una molécula detectable.

25 Por un "resto detectable" se incluye un resto que permite determinar su presencia y/o cantidad relativa y/o localización (por ejemplo, la localización en una serie), directa o indirectamente. Los restos detectables adecuados son bien conocidos en la técnica. Preferiblemente, el resto detectable se selecciona del grupo que consiste en un resto fluorescente; un resto luminiscente; un resto quimioluminiscente; un resto radiactivo; un resto enzimático, un resto de ligando o un resto de unión a ligando.

30 Por ejemplo, el resto detectable puede ser un resto fluorescente y/o luminiscente y/o quimioluminiscente que, cuando se expone a condiciones específicas, puede detectarse. Dicho resto fluorescente puede tener que exponerse a radiación (es decir, luz) a una longitud e intensidad de onda específicas para causar la excitación del resto fluorescente, posibilitando de ese modo que emita fluorescencia detectable a una longitud de onda específica que puede detectarse.

35 Como alternativa, el resto detectable puede ser una enzima que es capaz de convertir un sustrato (preferiblemente indetectable) en un producto detectable que puede visualizarse y/o detectarse. Ejemplos de enzimas adecuadas se analizan en más detalle a continuación en relación con, por ejemplo, ensayos ELISA.

40 En una realización adicional, el resto detectable puede ser un átomo radiactivo que es útil en la formación de imágenes. Los átomos radiactivos adecuados incluyen <sup>99m</sup>Tc y <sup>231</sup>I para estudios de centelleografía. Otros restos fácilmente detectables incluyen, por ejemplo, marcadores de spin para imágenes de resonancia magnética (MRI) tales como <sup>123</sup>I de nuevo, <sup>131</sup>I, <sup>111</sup>In, <sup>19</sup>F, <sup>13</sup>C, <sup>15</sup>N, <sup>17</sup>O, gadolinio, manganeso o hierro. Claramente, el agente de detectar (tal como, por ejemplo, la una o más proteínas en la muestra de ensayo y/o en la muestra de control descritas en este documento y/o una molécula de anticuerpo para su uso en la detección de una proteína seleccionada) debe tener suficiente de los isótopos atómicos apropiados para que el resto detectable sea fácilmente detectable.

45 Los radiomarcadores u otros marcadores pueden incorporarse en las proteínas presentes en las muestras de los métodos de la invención y/o en los agentes de unión de la invención de modos conocidos. Por ejemplo, si el agente de unión es un polipéptido puede biosintetizarse o puede sintetizarse por síntesis química de aminoácidos usando precursores adecuados de aminoácidos que implican, por ejemplo, flúor-19 en lugar de hidrogeno. Marcadores tales como <sup>99m</sup>Tc, <sup>123</sup>I, <sup>186</sup>Rh, <sup>188</sup>Rh y <sup>111</sup>In pueden unirse, por ejemplo, a través de restos de cisteína en el resto de unión. El itrio-90 puede unirse mediante un resto de lisina. El método IODOGEN (Fraker et al., (1978) Biochem. Biophys. Res. Comm. 80, 49-57) puede usarse para incorporar <sup>123</sup>I. La referencia ("Monoclonal Antibodies in Immunoscintigraphy", J-F Chatal, CRC Press, 1989) describe otros métodos en detalle. Los métodos para conjugar otros restos detectables (tales como restos enzimáticos, fluorescentes, luminiscentes, quimioluminiscentes o radiactivos) con proteínas son bien conocidos en la técnica.

50 Los expertos en la materia apreciarán que las proteínas de la muestra o muestra a ensayar pueden marcarse con un resto que ayude de forma indirecta a la determinación de la presencia, cantidad y/o localización de dichas proteínas. Por tanto, el resto puede constituir un componente de un resto detectable de múltiples componentes. Por ejemplo, las proteínas en la muestra o muestras a ensayar pueden marcarse con biotina, que permite su posterior detección usando estreptavidina fusionada o unida de otro modo a un marcador detectable.

65 En una realización adicional la etapa (b) y/o la etapa (d) se realizan usando una serie.

- Las series per se son bien conocidas en la técnica. Normalmente están formadas de una estructura lineal o bidimensional que tiene regiones ("manchas") espaciadas (es decir, concretas), teniendo cada una un área finita, formada sobre la superficie de un soporte sólido. Una serie también puede ser una estructura de perla donde cada perla puede identificarse por un código molecular o un código de color o identificarse en un flujo continuo. También puede realizarse el análisis de forma secuencial donde la muestra se pasa sobre una serie de manchas que adsorben cada una clase de molécula de la solución. El soporte sólido es normalmente vidrio o un polímero, siendo los polímeros más comúnmente usados celulosa, poli(acrilamida), nailon, poliestireno, cloruro de polivinilo o polipropileno. Los soportes sólidos pueden estar en forma de tubos, perlas, discos, chips de silicio, microplacas, membrana de difluoruro de polivinilideno (PVDF), membrana de nitrocelulosa, membrana de nailon, otra membrana porosa, membrana no porosa (por ejemplo, plástico, polímero, metacrilato, silicio, entre otros, una pluralidad de pernos poliméricos, o una pluralidad de pocillos de microtitulación, o cualquier otra superficie adecuada para inmovilizar proteínas, polinucleótidos y otras moléculas adecuadas y/o para realizar un inmunoensayo. Los procesos de unión son bien conocidos en la técnica y generalmente consisten en reticular, unir de forma covalente o adsorber de forma física una molécula proteica, polinucleótido o similar al soporte sólido. Como alternativa, puede emplearse el acoplamiento de afinidad de las sondas a través de marcas de afinidad o de construcciones similares. Usando técnicas bien conocidas, tales como impresión por contacto o sin contacto, cobertura o fotolitografía, puede definirse la localización de cada mancha. Para revisiones, véase Jenkins, R.E., Pennington, S.R. (2001, *Proteomics*, 2,13-29) y Lal et al., (2002, *Drug Discov Today* 15;7(18 Supl.):S143-9).
- La serie puede ser una serie basada en perlas. Normalmente, la serie es una serie basada en superficie, por ejemplo, una macroserie, microserie o nanoserie, preferiblemente una serie de anticuerpos o una microserie de anticuerpos.
- Normalmente, la serie es una microserie. Por "microserie" se incluye el significado de una serie de regiones que tienen una densidad de regiones concretas de al menos aproximadamente 100/cm<sup>2</sup>, y preferiblemente de al menos aproximadamente 1000/cm<sup>2</sup>. Las regiones en una microserie tienen dimensiones típicas, por ejemplo, diámetro, en el intervalo entre aproximadamente 10-250 nm, y están separadas de otras regiones en la serie por aproximadamente la misma distancia. La serie puede ser como alternativa una macroserie o una nanoserie.
- Una vez se han identificado y aislado las moléculas de unión adecuadas (analizadas anteriormente), los expertos en la materia pueden fabricar una serie usando métodos bien conocidos en la técnica de biología molecular; véanse los ejemplos a continuación.
- En una realización adicional más, la etapa (b) y/o la etapa (d) se realizan usando una serie que comprende un segundo agente de unión capaz de unirse a los dos o más primeros agentes de unión, teniendo el segundo agente de unión un resto detectable. Los segundos agentes de unión adecuados se describen en detalle anteriormente en relación a los primeros agentes de unión.
- Por tanto, las proteínas de interés en la muestra a ensayar primero pueden aislarse y/o inmovilizarse el primer agente de unión, después de lo cual puede determinarse la presencia y/o cantidad relativa de dichas proteínas usando un segundo agente de unión.
- Preferiblemente, el segundo agente de unión es un anticuerpo o un fragmento de unión a antígeno del mismo. De forma conveniente, el anticuerpo o fragmento del mismo se selecciona del grupo que consiste en: scFv; Fab; un dominio de unión de una molécula de inmunoglobulina. Los anticuerpos y fragmentos adecuados, y métodos para preparar los mismos, se describen en detalle anteriormente.
- Como alternativa, el segundo agente de unión puede ser un agente de unión de tipo anticuerpo, tal como un anticuerpo o aptámero.
- Como alternativa, cuando el resto detectable sobre la proteína en la muestra a ensayar comprende o consiste en un miembro de un par de unión específica (por ejemplo, biotina), el segundo agente de unión puede comprender o consistir en el miembro complementario del par de unión específica (por ejemplo, estreptavidina).
- Cuando se usa un ensayo de detección, se prefiere que el resto detectable se seleccione del grupo que consiste en: un resto fluorescente; un resto luminiscente; un resto quimioluminiscente; un resto radiactivo; un resto enzimático, un resto de ligando o un resto de unión a ligando. Ejemplos de restos detectables adecuados para su uso en los métodos de la invención se describen anteriormente.
- Los ensayos preferidos para detectar proteínas en suero o plasma incluyen ensayos inmunoabsorbentes ligados a enzimas (ELISA), radioinmunoensayos (RIA), ensayos inmunoradiométricos (IRMA) y ensayos inmunoenzimáticos (IEMA), incluyendo ensayos de tipo sándwich usando anticuerpos monoclonales y/o policlonales. Los ensayos de tipo sándwich a modo de ejemplo se describen por David et al., en las patentes de Estados Unidos n.º 4.376.110 y 4.486.530, incorporadas por la presente por referencia. En la tinción con anticuerpos de células sobre portaobjetos, puede usarse en métodos bien conocidos en ensayos de diagnóstico de laboratorio de citología, que son bien conocidos por los expertos en la materia.

- 5 Por tanto, en una realización, el ensayo es un ELISA (ensayo inmunoabsorbente ligado a enzimas) que normalmente implica el uso de enzimas que dan un producto de reacción coloreado, habitualmente en ensayos en fase sólida. Enzimas tales como peroxidasa de rábano rústico y fosfatasa se han empleado ampliamente. Un modo de amplificar la reacción de fosfatasa es usar NADP como sustrato para generar NAD que ahora actúa como co-enzima para un segundo sistema enzimático. La pirofosfatasa de *Escherichia coli* proporciona un buen conjugado porque la enzima no está presente en los tejidos, es estable y da un buen color de reacción. También pueden usarse sistemas quimioluminiscentes basados en enzimas tales como luciferasa.
- 10 La conjugación con la vitamina biotina se usa frecuentemente ya que esta puede detectarse fácilmente por su reacción con avidina unida a enzima o estreptavidina a que se une con gran especificidad y afinidad.
- 15 En una realización alternativa, el ensayo usado para la detección de proteínas es convenientemente un ensayo fluorométrico. Por tanto, el resto detectable del segundo agente de unión puede ser un resto fluorescente, tal como un fluoróforo *Alexa* (por ejemplo, *Alexa-647*).
- Preferiblemente, el anticuerpo o fragmento de unión a antígeno del mismo es un anticuerpo recombinante o fragmento de unión a antígeno del mismo.
- 20 En una realización, el método del primer aspecto de la invención es un método *in vitro*. En una realización alternativa, el método del primer aspecto de la invención es un método *in vivo*.
- 25 También se describe en este documento una serie para su uso en un método de acuerdo con el primer aspecto de la invención, comprendiendo la serie uno o más primeros agentes de unión como se define anteriormente. Preferiblemente, la serie comprende uno o más agentes de unión capaces de unirse a un biomarcador de la Tabla 1. Más preferiblemente, la serie comprende agentes de unión capaces de unirse a un biomarcador de la Tabla 1(A), mucho más preferiblemente, a todos los biomarcadores de la Tabla 1(A).
- 30 La serie puede comprender de forma adicional o alternativa uno o más agentes de unión capaces de unirse a un biomarcador de la Tabla 1(B), por ejemplo, a 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 o a 19 biomarcadores de la Tabla 1(B). Por tanto, se prefiere que la serie comprenda agentes de unión capaces de unirse a cada uno de los biomarcadores definidos en la Tabla 1(A) y en la Tabla 1(B).
- 35 Además, la serie puede comprender agentes de unión capaces de unirse a uno o más de los biomarcadores definidos en la Tabla 1(C), por ejemplo, a 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o a 15 biomarcadores de la Tabla 1(C). Preferiblemente, la serie comprende agentes de unión capaces de unirse a cada uno de los biomarcadores definidos en la Tabla 1(C).
- 40 Los expertos en la materia apreciarán que la serie puede comprender una pluralidad de diferentes agentes de unión capaces de unirse al mismo biomarcador seleccionado de los grupos anteriores de biomarcadores.
- 45 Los agentes de unión pueden comprender o consistir en un anticuerpo o fragmento de unión a antígeno del mismo, tal como un anticuerpo monoclonal o un fragmento del mismo. El anticuerpo o fragmento de unión a antígeno del mismo puede ser un anticuerpo recombinante o un fragmento de unión a antígeno del mismo.
- 50 El anticuerpo o el fragmento de unión a antígeno puede seleccionarse del grupo que consiste en anticuerpos intactos, fragmentos Fv (por ejemplo, Fv de cadena sencilla y Fv unido por disulfuro), fragmentos tipo Fab (por ejemplo, fragmentos Fab, fragmentos Fab' y fragmentos F(ab)<sub>2</sub>), dominios variables individuales (por ejemplo, dominios V<sub>H</sub> y V<sub>L</sub>) y anticuerpos de dominio (por ejemplo, dAb, incluyendo formatos individuales y dobles, es decir, dAb-enlazador-dAb). Preferiblemente, el anticuerpo o el fragmento de unión a antígeno es un Fv de cadena sencilla (scFv).
- 55 Como alternativa, el agente de unión puede comprender o consistir en un agente de unión de tipo anticuerpo, por ejemplo, un anticuerpo o un aptámero.
- Los agentes de unión pueden inmovilizarse. Sin embargo, como alternativa pueden estar sin fijar.
- 60 La serie puede comprender o consistir en al menos 2 a 10 agentes diferentes de unión (por ejemplo, diferentes anticuerpos) para cada uno de los biomarcadores enumerados en la Tabla 1(A) y/o 1(B) y/o 1(C), por ejemplo al menos de 2-3, 2-4, 2-5, 2-6, 2-7, 2-8, o de 2-9 diferentes especies de agente de unión para cada biomarcador. Por tanto, la serie puede comprender o consistir en al menos 3 diferentes agentes de unión para cada biomarcador.
- Las series adecuadas para su uso en los métodos de la invención se analizan en detalle anteriormente.
- 65 También se describe en este documento el uso *in vitro* de un biomarcador seleccionado del grupo de biomarcadores de la Tabla 1(A) como marcador pronóstico para determinar el riesgo de recidiva y/o metástasis de cáncer de mama

en un sujeto. El biomarcador o biomarcadores pueden ser para su uso en combinación con uno o más biomarcadores adicionales seleccionados del grupo de biomarcadores de la Tabla 1(B) y/o 1(C). El uso puede ser en combinación con todos los biomarcadores de (i) la Tabla 1(A) y la Tabla 1(B); (ii) la Tabla 1(A) y la Tabla 1(C); (iii) la Tabla 1(B) y la Tabla 1(C); (iv) la Tabla 1(A), la Tabla 1(B) y la Tabla 1(C), o por ejemplo, 1 o 2 de los biomarcadores de la Tabla 1(A), 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 o 19 biomarcadores de la Tabla 1(B) y/o 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 biomarcadores de la Tabla 1(C).

Convenientemente, el uso comprende o consiste en 3 o más biomarcadores seleccionados del grupo definido en la Tabla 1, por ejemplo, al menos 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37 o 38 biomarcadores seleccionados del grupo definido en la Tabla 1. Se prefiere que el uso comprenda o consista en biomarcadores que comprenden el grupo de apolipoproteína A4 (APOA4), subunidad beta de ATP sintasa, mitocondrial (ATPSB), tirosina quinasa de Bruton (BTK), inhibidor de la esterasa del complemento-1 (inhibidor de esterasa de C1), factor B del complemento (factor B), CD40, factor de transcripción de homeodominio CHX10 (CHX10), interleuquina-1 $\alpha$  (IL-1 $\alpha$ ), interleuquina-5 (IL-5), interleuquina-6 (IL-6), interleuquina-7 (IL-7), interleuquina-9 (IL-9), interleuquina-12 (IL-12), interleuquina-13 (IL-13), interleuquina-18 (IL-18), miembro 9 de la familia del dominio TBC1 (KIAA0882), Lewis<sup>x</sup>/CD15, proteína de unión a oxisterol tipo 3 (OSBPL3), proteína quimiotáctica de monocitos-1 (MCP-1), Sialil Lewis<sup>x</sup>, factor-1 $\beta$  de necrosis tumoral (TNF-  $\beta$ ).

También se describe en este documento un kit para determinar el riesgo de recidiva de cáncer de mama en un individuo, que comprende o consiste en:

- (i) uno o más agentes de unión capaces de unirse a un biomarcador enumerado en la Tabla 1(A) y/o 1(B) y/o 1(C); y
- (ii) instrucciones para realizar el método de la invención.

El kit puede comprender adicionalmente uno o más segundos agentes de unión como se define anteriormente. Preferiblemente el kit también comprende o consiste en:

- (iii) una serie de acuerdo con el tercer aspecto de la invención; y
- (iv) instrucciones para realizar el método de la invención.

El kit puede comprender agentes de unión capaces de unirse a cada uno de los biomarcadores definidos en la Tabla 1(A) y/o 1(B) y puede comprender adicionalmente agentes de unión capaces de unirse a uno o más de los biomarcadores definidos en la Tabla 1(C).

Como alternativa, el kit puede comprender una pluralidad de diferentes agentes de unión capaces de unirse al mismo biomarcador seleccionado del grupo de biomarcadores de la Tabla 1(A) y/o 1(B) y/o 1(C).

Los agentes de unión pueden comprender o consistir en un anticuerpo o fragmento de unión a antígeno del mismo tal como un anticuerpo monoclonal o un fragmento del mismo. Preferiblemente, el anticuerpo o el fragmento de unión a antígeno del mismo es un anticuerpo recombinante o un fragmento de unión a antígeno del mismo.

Ahora se describirán ejemplos no limitantes preferidos que encarnan ciertos aspectos de la invención, con referencia a las siguientes figuras:

**Figura 1:** Análisis SVM para la predicción de cáncer de mama metastásico. (A) Se compararon los niveles de analito para muestras recogidas a los 3-6 meses después de la cirugía con las muestras tomadas en el momento de la cirugía. Las diferencias en todos los analitos identificados en las series se suministraron a una SVM que, usando un procedimiento de validación cruzada de una exclusión, se instruyó para clasificar al paciente como uno que desarrollará cáncer metastásico o uno que no. El análisis produjo una AUC ROC de 0,88. En el mapa térmico se muestran todos los analitos que presentaban un valor p de Wilcoxon de  $< 0,05$ , indicando el rojo aumento y el verde disminución, respecto a la velocidad de biomarcador. (B) El mismo análisis SVM, pero usando la velocidad de analito comparando muestras tomadas en el momento de la cirugía con muestras tomadas 12 meses después de la cirugía. Este análisis produjo una AUC ROC de 0,75.

**Figura 2:** Análisis de la tasa de falso descubrimiento y velocidad de analito. (A) Se identificó la cantidad de pacientes con mayor señal de biomarcador para la segunda señal para cada anticuerpo. Los anticuerpos se clasificaron en consecuencia y se representaron (rojo), con el anticuerpo con señal creciente en la menor cantidad de pacientes a la izquierda. La primera y la segunda muestra entonces se recombinaron aleatoriamente, tratando la segunda muestra como la primera y viceversa para un conjunto aleatorio de pacientes. La cantidad de pacientes con mayor señal en la "segunda" muestra se identificó de nuevo, para conseguir una estimación de la tasa de falsos positivos (FPR) del análisis. Este procedimiento se repitió 10.000 veces. El color azul describe la frecuencia con que cierta cantidad de pacientes con señal creciente se clasificaban para cada anticuerpo durante la permutación del conjunto de datos recombinados aleatoriamente; el blanco indicaba no existente y el azul oscuro el más común. El valor promedio para cada clasificación de anticuerpo se representa como una línea

amarilla. A partir de este análisis, es evidente que la verdadera combinación de la primera y la segunda muestra da un resultado que diverge claramente de la línea amarilla que describe la FPR, no solamente en las colas del diagrama, sino incluso más en el centro, lo que indica una regulación positiva general. Los diagramas muestran la velocidad para (B) Lewis x; (C) IL-16; y (D) CD40, durante los primeros 3-6 meses después de la cirugía, respectivamente. La velocidad de la señal para estos análisis era claramente positiva para pacientes que finalmente desarrollaban cáncer de mama metastásico.

**Figura 3:** Prevalidación de la clasificación de pacientes en alto o bajo riesgo para recidiva de cáncer de mama. Se instruyó una SVM sobre el conjunto de datos de descubrimiento y se ensayó sobre el conjunto de datos de prevalidación, usando (A) datos de serie de anticuerpos obtenidos de la firma de 21 biomarcadores generada, usando eliminación regresiva, como se describe en los materiales y métodos suplementarios. La AUC ROC para esta clasificación fue de 0,85; (B) se instruyó una SVM usando datos clínicos convencionales (amarillo) que dan una AUC ROC = 0,66, mientras que la combinación de la firma de 21 biomarcadores y los parámetros clínicos convencionales (rojo) produjo una AUC ROC = 0,90, que demuestra el valor añadido en los datos de la serie.

**Figura 4:** Se descubrió que la quimioterapia no era un factor desconcertante para nuestra firma de biomarcadores candidatos. Los valores de decisión de la SVM para todos los pacientes se representan y los pacientes que reciben quimioterapia durante los primeros 3-6 meses se indican por flechas. Cuando se dividen los pacientes en grupos de riesgo basados en la decisión de la SVM, el 72 % de los pacientes en el grupo de alto riesgo (rojo) después desarrollaron metástasis, y solamente el 15 % en el grupo de bajo riesgo (azul). Cuando se correlaciona el resultado con los pacientes a los que se les dio tratamiento adyuvante (flechas negras), fue evidente que este factor no desviaba el análisis, con 5/18 y 7/20 en el grupo de alto y bajo riesgo, respectivamente.

**Figura 5:** Eliminación regresiva de variables. Para reducir la cantidad de anticuerpos usados para la clasificación, se combinó el procedimiento anterior de una exclusión con un proceso de eliminación regresiva para los anticuerpos. Este proceso también produjo una clasificación de los anticuerpos que permite asignar clasificaciones bajas a anticuerpos aleatoriamente correlacionados.

## Ejemplos

### Introducción

Para definir biomarcadores predictivos de suero asociados con recidiva del tumor en cáncer de mama, se elaboró la hipótesis de que los patrones descodificantes de proteínas de suero inmunorreguladoras de baja abundancia podrían revelar importante información acerca del riesgo de recidiva. Por consiguiente, utilizando la tecnología de serie de anticuerpos recombinantes del estado de la técnica capaz de analizar grandes cantidades de análisis proteicos de baja abundancia, usando solamente cantidades minoritarias de suero no fraccionado (10), se exploraron muestras de pacientes recogidas sobre un periodo de tres años. Las muestras se recogieron de pacientes con cáncer de mama antes de la resección del tumor primario y después cada 6-12 meses, produciendo hasta 5 muestras/paciente. Analizando la velocidad, es decir, el cambio en el tiempo, de los marcadores, pudo extraerse información indicativa del riesgo de los pacientes de desarrollar metástasis.

Este estudio demuestra, por primera vez, que una muestra simple de sangre alberga información predictiva de recidiva de tumor sistémico en pacientes con cáncer de mama, que podría permitir nuevas posibilidades en la selección de la terapia. Además, en términos de rendimiento, logra un valor añadido sustancial para los predictores clínicos derivados en el momento de la cirugía primaria.

### Materiales y métodos

#### Muestras y análisis de serie

Se recogieron muestras de suero de dos cohortes independientes de pacientes, indicadas como la cohorte de descubrimiento y la cohorte de prevalidación (Tabla 3). En la cohorte de descubrimiento, se recogieron muestras de 38 pacientes diagnosticados con cáncer de mama primario. Se recogieron consentimientos informados por escrito durante la visita preoperatoria en el departamento de cirugía (Lund University Hospital, Lund), cuando también se recogieron las muestras de suero. Se registraron la hora y la fecha en que se extrajeron las muestras de sangre. El suero se almacenó a -80 °C, se marcó con códigos de serie para posibilitar los análisis con ocultación. Para la mayoría de los pacientes, la visita preoperatoria tuvo lugar menos de una semana antes de la cirugía. Se extrajo sangre una segunda vez en el primer seguimiento (3-6 meses) y después aproximadamente cada 12 meses durante tres años. Este estudio estaba aprobado por el comité ético regional (Lund, Suecia). En la cohorte de prevalidación, se extrajeron muestras de suero, como se describe para la cohorte de descubrimiento, de 26 pacientes nuevos e independientes con cáncer de mama. Los pacientes que no desarrollaron recidiva distante tuvieron un seguimiento de hasta 7 años (Tabla 3).

La plataforma de microserie de anticuerpos recombinantes contenía 135 anticuerpos contra 65 antígenos diferentes, es decir, para asegurar la calidad se utilizaron 2-5 clones diferentes de anticuerpo contra la mayoría de antígenos

(Tabla 4). El análisis de microserie, incluyendo la preparación de muestras, la producción de anticuerpos, la fabricación y normalización de las series, se describe en detalle (10, 11).

5 Para ser capaces de comparar la predicción de recidiva distante, usando parámetros clínicos convencionales, se asignó cada uno de los parámetros disponibles a valores numéricos, donde el estado premenopáusico/posmenopáusico, ER y PgR-negativo/positivo, tipo ductal/lobular y el estado del ganglio linfático se establecieron a -1 o 1, respectivamente. Además, el grado histológico I, II o III se estableció a -1, 0 o 1, mientras que el tamaño del tumor se graduó de forma continua de -1 a 1. Una máquina de soporte vectorial (SVM) entonces se instruyó sobre el conjunto de datos de descubrimiento, usando los datos clínicos convencionales, o una combinación de los datos clínicos y de microserie.

#### Análisis de datos usando la máquina de soporte vectorial

15 Se empleó una SVM para clasificar las muestras como pertenecientes a uno de dos grupos definidos, usando un núcleo lineal con el coste de restricciones establecido a 1. No se hicieron intentos por ajustarlo para evitar el riesgo de sobreajuste. Las diferencias log entre las señales para muestras recogidas en el momento de la operación y entre tras a seis meses después se calcularon para cada paciente, y se usaron para instruir y ensayar el clasificador SVM, usando validación cruzada de una exclusión. Esta parte de instrucción, usando la cohorte de descubrimiento, incluyó la creación de un subpanel de anticuerpos seleccionando anticuerpos que, en el conjunto de instrucción, presentaban la mayor potencia de discriminación combinada. Esta selección de anticuerpos se hizo usando una estrategia de eliminación regresiva de validación cruzada. Usando este enfoque, se compiló una lista de 21 anticuerpos con los valores más altos (Tabla 1(A) y (B)) y se instruyó un modelo SVM final, ahora llamado congelado. Los parámetros SVM derivados de la instrucción se muestran en la Tabla 2.

25 Durante la prevalidación, se analizaron y ensayaron muestras de suero de 26 pacientes nuevos e independientes con cáncer de mama (cohorte de prevalidación), usando las velocidades de análisis en el clasificador SVM previamente congelado.

#### Preparación de muestras

30 Las muestras de suero se biotinilaron, usando un protocolo optimizado previamente para marcar el proteoma de suero (10). Todas las muestras de suero se marcaron, usando EZ-Link Sulfo-NHS-LC-Biotina (Pierce, Rockford, IL, EE. UU.), como se ha descrito previamente (10). En resumen, se centrifugaron alícuotas de suero de 50 µl y se diluyeron 1:45 en PBS, produciendo una concentración final de proteína de aproximadamente 2 mg/ml. Después se añadió Sulfo-NHS-biotina hasta una concentración final de 10 mM y las muestras se incubaron en hielo durante 2 h. Se retiró la biotina no conjugada por diálisis frente a PBS a 4 °C durante 72 h. Finalmente, las muestras se separaron en alícuotas y se almacenaron a -20 °C antes de su uso.

#### Producción y purificación de fragmentos de anticuerpo recombinante

40 Se seleccionaron ciento treinta y cinco fragmentos de anticuerpo scFv recombinante humano contra 65 antígenos, principalmente contra factores inmunorreguladores, usando la biblioteca de presentación en fagos n-CoDeR (25) (Tabla 4). Los criterios de selección fueron rigurosos, usando procedimientos operativos convencionales, para asegurar la correcta especificidad del anticuerpo. Todas las sondas scFv se produjeron en 100 ml de cultivos de *E. coli*, y se purificaron a partir de sobrenadantes de expresión o preparaciones periplásmicas, usando cromatografía de afinidad sobre Ni-NTA agarosa (Qiagen, Hilden, Alemania). Se realizó elución, usando imidazol 250 mM, seguido de diálisis extensiva con PBS. La concentración de proteína se determinó midiendo la absorbancia a 280 nm, y el scFv purificado se almacenó a 4 °C hasta su uso adicional.

#### Fabricación y procesamiento de microseries de anticuerpos

La producción y manipulación de las microseries de anticuerpos se realizaron de acuerdo con una configuración previamente optimizada (10, 25). En resumen, se fabricaron microseries de scFv, usando una impresora sin contacto (Biochip Arrayer1, Perkin Elmer Life & Analytical Sciences, Wellesley, MA, EE. UU.), que deposita aproximadamente 330 pl/gota, usando tecnología piezoeléctrica. Los anticuerpos scFv se dispusieron en serie depositando 2 gotas en cada posición, donde se permitió que la primera gota se secase antes de distribuir la segunda gota. El soporte sólido era los portaobjetos de microserie MaxiSorp poliméricos negros (NUNC A/S, Roskilde, Dinamarca) y cada anticuerpo se dispuso en serie en ocho réplicas para asegurar una estadística adecuada. Para ayudar en la alineación en rejilla durante la posterior cuantificación, se manchó una fila que contenía estreptavidina conjugada con Alexa647 (2 µg/ml) como la fila superior en las ocho subseries que constituyen la serie. Los portaobjetos se bloquearon con leche desnatada en polvo al 5 % (p/v) (Semper AB, Sundbyberg, Suecia) en PBS durante una noche y después se colocaron en una Protein Array Workstation (PAW) (Perkin Elmer Life & Analytical Sciences) donde se lavaron durante cuatro minutos a 60 µl/min con Tween-20 al 0,05 % en PBS (PBS-T). Después de ello, se inyectaron 75 µl de la muestra y se dejó que se agitaran sobre la serie cada 15 segundos, durante 60 minutos. Después de otro lavado de cuatro minutos, las series se incubaron durante 60 minutos con 350 µl de 1 µg/ml de estreptavidina conjugada con Alexa-647 en PBS, con leche desnatada en polvo al 1 % (p/v) y Tween 20 al 1 %. Finalmente,

después de una última etapa de lavado, las series se secaron bajo una corriente de gas nitrógeno y se exploraron con un escáner de microserie confocal (ScanArray Express, Perkin Elmer Life & Analytical Sciences) a resolución de 5  $\mu\text{m}$ , usando tres ajustes diferentes del escáner.

5 Se usó el software ScanArray Express versión 4.0 (Perkin Elmer Life & Analytical Sciences) para cuantificar la intensidad de cada mancha, usando el método de círculo fijo. El fondo local se sustrajo, y para compensar los posibles defectos locales, se excluyeron automáticamente las dos réplicas mayores e inferiores. En todos los análisis de datos adicionales, cada dato puntual representa el valor medio de las cuatro manchas replicadas restantes. Para analitos de proteína que presentan señales saturadas, se usaron los valores de ajustes de escáner inferiores.

#### Normalización de datos de microserie

15 Se realizó la normalización entre chips del conjunto de datos, usando un enfoque de normalización semiglobal, similar al método de normalización usado para microseries de ADN. Para encontrar un factor de escala (7, 11, 27), se calculó el coeficiente de variación (CV) para cada anticuerpo, y el 15 % de los anticuerpos que presentaban los valores CV más bajos sobre todas las muestras se identificaron, correspondientes a 20 analitos. El factor de normalización  $N_i$  se calculó por la fórmula  $N_i = S_i/\mu$ , donde  $S_i$  es la suma de las intensidades de señal de los 20 analitos para cada muestra, y  $\mu$  es el promedio de  $S_i$  de todas las muestras. Para normalizar las muestras, todas las intensidades de anticuerpo en una muestra se dividieron con su factor de normalización  $N_i$ .

#### Análisis de datos

##### *Calibración del clasificador y ensayo independiente*

25 Para reducir las variaciones entre muestras, se usaron diferencias entre las expresiones de proteína a  $t = 0$  y  $t = 3-6$  meses, como variables independientes para un clasificador SVM. Se instruyeron y se ensayaron las SVM con un procedimiento de una exclusión, usando  $N = 38$  muestras e inicialmente todo  $M = 135$  variables. Como la cantidad de anticuerpos excedía la cantidad de muestras, se necesitó eliminar aquellos anticuerpos con bajo impacto sobre las predicciones para evitar el ajuste al ruido debido a correlaciones aleatorias. Esto se hizo para cada una de las muestras de una exclusión usando un procedimiento de eliminación regresiva. El rendimiento se midió con el área de ROC (características operativas del receptor). Las listas de anticuerpos clasificados resultantes de cada modelo SVM después se fusionaron en una firma consenso. La última después se usó con una SVM congelada sobre un conjunto independiente de ensayo. A continuación, se describen las diferentes etapas en algún detalle y se resume el procedimiento completo en la Figura 5.

##### *Procedimiento de una exclusión*

40 El principio para el procedimiento es, para  $N$  muestras que pertenecen a dos clases, instruir y ensayar una SVM usando validación cruzada de una exclusión, es decir, instruir la SVM sobre todas las muestras salvo una, y ensayar el modelo resultante usando la muestra excluida. La muestra de ensayo se asigna a un valor de decisión usando el modelo SVM instruido, y se pone de nuevo en el conjunto de instrucción, después de lo cual se excluye la siguiente muestra a usarse como muestra de ensayo. El procedimiento se repite hasta que cada muestra se ha asignado a un valor de decisión una vez usando todas las demás muestras como conjunto de instrucción. Los valores de decisión para todas las muestras se usan para crear una curva ROC, y se calcula el área bajo la curva. El área obtenida sirve como estimación del área esperada para este conjunto de muestras.

##### *Eliminación regresiva de variables*

50 Para reducir la cantidad de anticuerpos usados para la clasificación, se combinó el procedimiento anterior de una exclusión con un proceso de eliminación regresiva para los anticuerpos. Este proceso también producirá una clasificación de los anticuerpos con el propósito de asignar bajas clasificaciones a anticuerpos correlacionados aleatoriamente.

55 El proceso se describe en la Figura 5, a la que aquí se hace referencia como etapa 1-7. Partiendo con  $M$  anticuerpos, se crean  $M$  conjuntos de datos donde cada conjunto de datos tiene un anticuerpo remplazado con un valor constante, que es el valor promedio de ese anticuerpo para todas las muestras (**etapa 1**). Para evaluar cada importancia del anticuerpo para la clasificación en el conjunto de datos en curso, se hace un procedimiento de una exclusión de la SVM (como se ha descrito anteriormente) para cada uno de los  $M$  conjuntos de datos. Posteriormente, se crean  $M$  curvas ROC usando el resultado de SVM, y se identifica el conjunto de datos del que el área ROC generada tenía la disminución más pequeña (podría ser negativa) en comparación con el original. El conjunto de anticuerpos con un valor constante en ese conjunto de datos se identifica y se elimina (**etapa 2**). Los conjuntos de datos ahora contienen  $M-1$  anticuerpos y, por lo tanto, se crean  $M-1$  nuevos conjuntos de datos, teniendo todos uno de los restantes anticuerpos remplazado con su promedio. El procedimiento de ensayo de una exclusión junto con la evaluación de la SVM usando anticuerpos constantes se repite, eliminando, en efecto, el siguiente anticuerpo que porta la mínima información. El procedimiento se continúa hasta que queda solamente un

anticuerpo, produciendo un orden de clasificación para cada importancia del anticuerpo en la clasificación de las muestras en ese momento en el conjunto de datos (**etapa 3**). Esta información entonces puede usarse para construir subpaneles de anticuerpos de cualquier longitud deseada (**etapa 4**). Para evaluar la potencia de predicción de dicho panel de anticuerpos, no es posible ensayar un nuevo modelo usando las mismas muestras en un modo imparcial. Además, los anticuerpos correlacionados aleatorios aún pueden obtener una alta clasificación. Por lo tanto, se añade un ciclo adicional de una exclusión (**etapa 5a**) donde una muestra se retira antes del inicio del procedimiento de eliminación regresiva que, por tanto, puede usarse como una muestra de ensayo (**etapa 5b**). Este ciclo más externo se itera excluyendo cada muestra como un conjunto de ensayo una vez, y las muestras restantes producen una lista de clasificación de anticuerpos. Esta lista de clasificación entonces se usa para escoger un subpanel de anticuerpos con la clasificación más alta e instruir un único modelo SVM, que se ensaya sobre la muestra de ensayo. El resultado del proceso más externo de una exclusión es una lista de valores de decisión para todas las muestras, para cualquier longitud dada de subpanel. El área ROC correspondiente sirve como un ensayo del rendimiento usando subpaneles de anticuerpos del tamaño dado, que pueden usarse para estimar la cantidad de anticuerpo necesaria para hacer una clasificación adecuada en el conjunto de datos.

#### *Firma de anticuerpo consenso*

El procedimiento de eliminación regresiva produce la misma cantidad de listas de clasificación de anticuerpos que la cantidad de muestras. Para producir un único orden de clasificación, se concatena la información de cada ejecución en una lista consenso asignando cada anticuerpo a un valor (etapa 6) basado en su supervivencia promedio en las rondas de eliminación, donde el anticuerpo con la mayor supervivencia promedio se clasifica como el más importante. Finalmente, se instruye una nueva SVM, usando los 21 anticuerpos con los mejores valores, y se ensaya sobre el nuevo conjunto de datos independiente (**etapa 7**).

## **Resultados**

### Forma de biomarcadores de suero asociada con recidiva del tumor

En un intento por identificar una firma de biomarcadores de suero que fuera predictiva de recidiva distante en cáncer de mama, se recogieron muestras de pacientes con cáncer de mama primario sobre un periodo de tres años. Encontrar un clasificador potente para enfermedad metastásica comparando los niveles absolutos de proteína en el suero (analito) de las muestras recogidas en el momento de la cirugía primaria no fue satisfactorio, como fue evidente por un área bajo la curva (AUC) ROC de 0,54 (datos no mostrados). Por consiguiente, en lugar de comparar los niveles absolutos de los analitos, se analizó la velocidad, que se definió como el cambio log para cada analito entre la primera muestra de suero extraída de forma preoperatoria y cada muestra consecutiva de cada paciente sobre el periodo de tres años. Después, se aportó a la SVM la dirección, es decir, regulación positiva o negativa y la magnitud del cambio sobre el tiempo. El clasificador permitió una estratificación de los pacientes en riesgo alto frente a bajo para recidiva del tumor con una AUC ROC de 0,88 (Figura 1A), usando muestras de suero recogidas 3-6 meses después de la cirugía. Por tanto, este enfoque basado en la velocidad de biomarcador, demostró que había suficiente información en el conjunto de datos para identificar una forma de biomarcadores candidatos que permitiera una clasificación de los pacientes y para identificar un grupo de alto riesgo para desarrollaron una recidiva de tumor distante en tres años.

Adoptando la misma estrategia, pero en su lugar usando las muestras recogidas a los 12 meses después de la cirugía, se encontró un resultado cualitativo similar, pero con una potencia de predicción disminuida, como se ilustra por una AUC ROC de 0,75 (Figura 1 B). Esta tendencia podría no seguirse más, usando muestras recogidas a los 24 y 36 meses, ya que la cohorte de pacientes diagnosticados sin enfermedad metastásica era demasiado pequeña para un análisis estadísticamente relevante.

Analizando todos los analitos (Figura 2A), fue evidente que las velocidades crecientes de los analitos para pacientes metastásicos durante este periodo eraban claramente sobre-representadas, mientras que lo opuesto es cierto para pacientes no metastásicos. Cuando se analizó la firma de biomarcadores predictivos en más detalle, se demostró que Lewis x ( $p=0,0005$ ), IL-16 ( $p=0,002$ ) y CD40 ( $p=0,0003$ ) diferían al máximo entre pacientes con cáncer de mama recidivante frente a no recidivante, durante los primeros 6 meses después de la cirugía, como se determina por ensayo de clasificación de Wilcoxon. En la Figura 2B-D, se muestra la dinámica para estos tres analitos, donde fue evidente que un aumento significativo durante estos primeros meses después de la cirugía era más frecuente para pacientes que desarrollan una recidiva del tumor distante en un punto temporal posterior.

### Prevalidación de la firma de biomarcadores derivada de la cohorte de descubrimiento

Para ensayar la potencia de la clasificación derivada de la cohorte de descubrimiento de 38 pacientes, primero se condensó la cantidad total de analitos hasta los 21 biomarcadores no redundantes que contribuyen al máximo a la clasificación de recidiva metastásica, usando una estrategia de eliminación regresiva. La segunda cohorte independiente de pacientes, indicada como la cohorte de prevalidación, comprendía otros 26 pacientes, de los cuales el 50 % desarrolló una recidiva de cáncer de mama distante en el periodo de tres años que duró el estudio. Se analizaron dos muestras de suero de cada uno de los 26 nuevos pacientes, la primera recogida en el momento

de la resección del tumor inicial y la segunda después de 3 a 6 meses, siguiendo el mismo procedimiento que para la cohorte de descubrimiento. Por consiguiente, se procesaron 52 muestras con esta plataforma de microserie de anticuerpos, como se ha descrito anteriormente, y se determinó la velocidad de cada biomarcador.

5 El clasificador, que consiste en los 21 biomarcadores de máxima contribución derivados de la cohorte de descubrimiento, permitió una estratificación de los pacientes en riesgo alto frente a bajo para recidiva del tumor con una AUC ROC de 0,85 en la cohorte de prevalidación independiente (Figura 3). Es de observar que las firmas de biomarcadores asociadas a cáncer previamente definidas solapaban hasta el 30 % con los actuales 21 biomarcadores (7, 9, 11). Por consiguiente, esta firma predictiva no reflejaba una respuesta inflamatoria general (14).

10

El efecto de la quimioterapia adyuvante sobre la clasificación en grupos de alto o bajo riesgo

15 Para ensayar si la clasificación, presentada en la Figura 1, era dependiente de la terapia recibida por los pacientes en la cohorte de descubrimiento, se analizó el efecto de la quimioterapia adyuvante sobre la clasificación en alto o bajo riesgo de recidiva de cáncer de mama. En la Figura 4, los valores de decisión de la SVM para todos los pacientes y los pacientes que reciben quimioterapia, durante los primeros 3-6 meses, se indican por flechas. No pudo detectarse estratificación de los pacientes individuales que reciben quimioterapia. Se obtuvieron resultados similares cuando se analizaron los pacientes que reciben tratamiento endocrino adyuvante (datos no mostrados). Por tanto, la clasificación en grupos de alto y bajo riesgo para recidiva del tumor distante no estaba desviada por una terapia adyuvante particular. De forma interesante, como unos pocos pacientes que recibieron quimioterapia aún después de 6 meses se clasificaban como pertenecientes al grupo de alto riesgo, no pudo demostrarse ningún efecto beneficioso de esa terapia particular. Por consiguiente, basándose en el perfil molecular derivado de nuestro análisis de microserie, estos pacientes podrían seleccionarse para otro régimen de tratamiento.

25

Firma de biomarcadores moleculares frente a parámetros convencionales de diagnóstico

30 La potencia del diagnóstico molecular a veces se ha cuestionado, en particular en relación con los parámetros clínicos convencionales (12), tales como estado del ganglio linfático, tamaño del tumor, grado histológico y estado del receptor de estrógenos (ER) y receptor de progesterona (PgR). Por lo tanto, es esencial comparar nuestro predictor de suero con el rendimiento de predictores basados en dichos marcadores clínicos. Utilizando la combinación de parámetros convencionales, se compararon con nuestra firma de biomarcadores de suero, que presentaban una AUC ROC de 0,85 (Figura 3A). Se instruyeron dos modelos SVM adicionales sobre el conjunto de datos de descubrimiento, usando los datos clínicos y una combinación de los datos clínicos y de microserie, respectivamente.

35

40 Los modelos se ensayaron, usando los pacientes de la cohorte de prevalidación, y el resultado se presentó como curvas ROC. La AUC ROC para el predictor, usando los parámetros clínicos convencionales, fue de 0,66. Por consiguiente, se consiguió un poder predictivo significativamente mejorado cuando se usaba la firma molecular, basada en velocidades de analito. Es de observar que cuando se combinaban los parámetros clínicos convencionales con nuestra firma de 21 biomarcadores y se usaban ambos como un enfoque para la clasificación del riesgo, se obtenía una AUC ROC de 0,90 (Figura 3B). Por consiguiente, los dos conjuntos diferentes de variables no contenían información solapante con respecto al resultado clínico, como es evidente por un análisis de correlación de Pearson, que mostró que las velocidades de analito tenían correlaciones muy débiles (< 0,35) con los marcadores clínicos convencionales. Esto apoya el hecho de que nuestro enfoque de suero de proteínas contenía información única no presente en marcadores convencionales, que estaba en contraste con las correlaciones observadas entre los marcadores clínicos y el perfilado de microserie génica (12).

50

55 Para comparar adicionalmente nuestros datos con las herramientas disponibles de toma de decisiones, se investigó Adjuvant online ([www.adjuvantonline.com](http://www.adjuvantonline.com)), que estima el riesgo de recidiva del tumor en cáncer de mama. La edad del paciente, el estado ER, el grado del tumor, el tamaño del tumor y la cantidad de ganglios positivos se introdujeron online y se registró el riesgo calculado para recidiva en diez años para cada paciente. Los pacientes entonces se clasificaron de acuerdo con su riesgo estimado y, usando el resultado verdadero para los pacientes en relación a la recidiva distante, se generó una ROC que presentaba una AUC de 0,60. Esto está en concordancia con la AUC de 0,66 dada por nuestro análisis, basado en parámetros clínicos convencionales, aunque el tiempo de seguimiento sin enfermedades fue de siete años en nuestra cohorte de prevalidación.

**Discusión**

60 Los pacientes con cáncer de mama se tratan con terapia regional local o por la adición de quimioterapia sistémica, terapia hormonal o biológica, produciendo un resultado clínico mejorado, pero también un tratamiento excesivo considerable, con efectos secundarios para los pacientes y costes para los profesionales de asistencia sanitaria. Por consiguiente, se necesitan parámetros de pronóstico, para estratificar a los pacientes en diferentes grupos de riesgo, que ayudaría a tomar decisiones racionales de tratamiento. Además, serían muy deseables parámetros predictivos para controlar la progresión de la enfermedad y la eficacia de la terapia, ya que permitirían la selección basada en el paciente de las terapias adyuvantes y la evitación del tratamiento excesivo. Actualmente, pacientes con la misma fase de la enfermedad pueden tener respuestas completamente diferentes a la terapia y de supervivencia global, lo

65

que apunta al hecho de que los parámetros tradicionales no logran clasificar de forma precisa los pacientes con cáncer de mama, de acuerdo con su necesidad clínica.

5 Los progresos recientes en el perfilado de la expresión génica han producido, sin embargo, resultados prometedores para predecir el resultado clínico de cáncer de mama (5, 6), predecir el riesgo de desarrollar recidiva local después de cirugía de conservación de la mama (14), así como para un diagnóstico más prematuro de cáncer de mama (15). Como la predicción del resultado clínico de cáncer de mama, basado en perfiles de expresión génica, requiere la presencia del tumor, los enfoques basados en genes no permiten el control continuo de la enfermedad y la evaluación de la eficacia del tratamiento después de la resección del tumor inicial. Para dichos propósitos, la elección preferida de muestra sería suero, que debido a las limitaciones técnicas no ha sido posible analizar usando enfoques proteómicos tradicionales (16).

15 Hemos diseñado una plataforma de microserie de anticuerpos recombinantes (21-23), que presenta alta sensibilidad y reproducibilidad y que está centrada en proteínas del suero asociadas con el sistema inmunitario. Este enfoque de proteómica de afinidad ya ha permitido recientemente la identificación de varias firmas de biomarcadores de suero que distinguen entre diferentes indicaciones de cáncer e individuos sanos (7, 8, 11), lo que demuestra la potencia de la plataforma.

20 En un intento por investigar si la información almacenada de forma transitoria en el suero podría descifrarse y ayudar en el tratamiento del cáncer de mama, se seleccionó un total de 64 pacientes con cáncer de mama primario y un total de unos pocos cientos de muestras de suero para perfiles moleculares asociados con el riesgo de desarrollar recidiva de tumor distante. Comparando las muestras recogidas sobre un periodo de tres años, no se pudieron descifrar los patrones moleculares en el proteoma de suero, que usa niveles estáticos de analito, que indicaban que las señales eran vectores de variabilidad demasiado débiles, paralelos en lugar de ortogonales, o inexistentes.

25 Sin embargo, cuando se analizaba el cambio en la intensidad del analito sobre el tiempo, se pudo clasificar a los pacientes en un grupo de alto y bajo riesgo para el desarrollo de cáncer de mama metastásico, dentro del periodo del estudio. Estaba claro que durante los primeros 3-6 meses, los biomarcadores discriminatorios aumentaban en intensidad analítica si un paciente pertenecía al grupo de alto riesgo y viceversa para los pacientes de bajo riesgo. Los estudios proteómicos clínicos más previos están limitados por el hecho de que se usa solamente una cohorte de pacientes para descubrir biomarcadores potenciales, y para evadir este obstáculo, se usó una segunda cohorte independiente para ensayar la capacidad de clasificar por riesgo a los pacientes. Para este propósito, primero se condensó la cantidad de analitos usados para prevalidar nuestra firma de biomarcadores candidatos, usando la estrategia de eliminación regresiva imparcial. Esto produjo la forma más discriminatoria, que no habría sido el caso si se seleccionaran solamente biomarcadores con los valores p más bajos.

35 Después de la condensación hasta 21 marcadores, se analizó la cohorte de ensayo y se pudieron clasificar los pacientes en alto y bajo riesgo de desarrollo de cáncer de mama metastásico con una AUC ROC de 0,85. Se esperaba esta disminución del poder predictivo, ya que los biomarcadores siempre funcionan mejor en los conjuntos de datos de los que se obtienen. No obstante, una AUC ROC de 0,85 es la primera demostración de que el contenido de información albergado en una muestra de suero podría descodificarse, usando una microserie de anticuerpos, que allana el camino para el desarrollo de un enfoque más personalizado para el tratamiento de pacientes con cáncer de mama. Es de observar que la forma no parecía verse afectada por la terapia recibida por los pacientes, por ejemplo, la quimioterapia no influía en la estratificación de los pacientes. Esta observación (Figura 4) también demostraba la necesidad de una predicción precisa de cáncer de mama metastásico, ya que varios pacientes en el grupo de alto riesgo podrían haberse beneficiado de otro tratamiento adyuvante, igual que algunos de los pacientes en el grupo de bajo riesgo podrían haber estado tratados excesivamente.

40 Para obtener un impacto clínico, los biomarcadores basados en suero tienen que funcionar mejor que los marcadores de pronóstico tradicionales, tales como clínicos (edad), histopatológicos (grado histológico, estado del ganglio linfático, tamaño del tumor) y estado del receptor de hormonas (ER, PgR) (14). Los predictores existentes, precalibrados basados en dichos marcadores no pueden usarse de forma directa en este caso, ya que normalmente asumen tiempo de recidiva más largos que los implicados en nuestro estudio y estarían desfavorecidos en una comparación. Por lo tanto, se evaluó el poder predictivo de los parámetros tradicionales por instrucción y ensayo con ocultación, usando una SVM con estos marcadores como valores de entrada. Se demostró que incluso su potencia combinada, representada por una AUC ROC de 0,66, era significativamente menor que nuestra firma prevalidada de biomarcadores de suero. Esto de soporte a nuestra hipótesis de que había suficiente información albergada en el suero que podría descifrarse, usando una tecnología adecuada, para la predicción de la enfermedad.

50 De forma interesante, cuando se combinaban los parámetros clínicos tradicionales con nuestra firma de 21 biomarcadores, el poder predictivo aumentaba, produciendo una AUC ROC incluso más mejorada de 0,90, que sugiere fuertemente que los marcadores de suero proporcionaban valor añadido. El presente análisis se basaba en 65 analitos distintos (Tabla 4) y, inspirado por el presente estudio, diseñaremos un chip ampliado de cáncer de mama capaz de analizar más proteínas del suero en cohortes de pacientes más grandes, para refinar adicionalmente la firma de biomarcadores y, por consiguiente, aumentar si poder predictivo.

La firma de biomarcadores que podría clasificar los pacientes en bajo o alto riesgo de desarrollar recidiva distante en cáncer de mama podría usarse, en un entorno clínico, junto con una firma de biomarcadores asociada con cáncer de mama metastásico ya diagnosticado (11). La firma (11) se basaba en ocho biomarcadores y ha demostrado dar una AUC ROC de 0,85 en una cohorte de ensayo independiente del presente estudio (datos no mostrados). Por consiguientes, una aplicación clínica potenciar podría ser que después de realizar una evaluación del riesgo sobre el paciente en la primera visita (3-6 meses) después de la operación, podría usarse la forma asociada con cáncer de mama metastásico establecido para hacer el seguimiento de los pacientes durante varios años y controlar sus perfiles moleculares del suero que funcionan como una señal prematura de aviso de recidiva del tumor. Esto tiene que validarse, por supuesto, en estudios prospectivos más grandes, pero indican un modo potencial directo hacia un enfoque más personalizado en terapia contra el cáncer de mama.

En conclusión, se ha demostrado que un simple ensayo de sangre alberga suficiente información para permitir una evaluación del riesgo de la probabilidad de desarrollar cáncer de mama metastásico, después de la operación primaria, usando proteómica de afinidad.

## Referencias

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005;55:74-108.
2. Gonzalez-Angulo AM, Morales-Vasquez F, Hortobagyi GN. Overview of resistance to systemic therapy in patients with breast cancer. *Adv Exp Med Biol.* 2007; 608:1-22.
3. Eifel P, Axelson JA, Costa J et al. National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, Noviembre 1-3, 2000. *J Natl Cancer Inst.* 2001; 93:979-89.
4. Goldhirsch A, Ingle JN, Gelber RD. et al. Thresholds for therapies: highlights from the St Gallen international expert consensus on the primary therapy of early breast cancer 2009. *Annals Oncol.* 2009; 20:1319-1329.
5. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415:530-536.
6. Paik S, Shak S, Tang G et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* 2004; 351:2817-2826.
7. Ingvarsson J, Wingren C, Carlsson A, et al. Detection of pancreatic cancer using antibody microarray-based serum protein profiling. *Proteomics* 2008; 8:2211-2219.
8. Sanchez-Carbayo M, Socci ND, Lozano JJ, Haab BB, Cordon-Cardo C, Profiling bladder cancer using targeted antibody arrays. *Am. J. Pathol.* 2006; 168:93-103.
9. Ellmark P, Ingvarsson J, Carlsson A, et al. Identification of protein expression signatures associated with *Helicobacter pylori* infection and gastric adenocarcinoma using recombinant antibody microarrays. *Mol. Cell. Proteomics* 2006; 5:1638-1646.
10. Wingren C, Ingvarsson J, Dextrin L, Szul D y Borrebaeck CAK. Design of recombinant antibody microarrays for complex proteome analysis: Choice of sample labeling-tag and solid support. *Proteomics* 2007; 7:3055-3065.
11. Carlsson A, Wingren C, Ingvarsson J, et al. Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays. *Eur. J. Cancer* 2008; 44:472-480.
12. Edén P, Ritz C, Rose C, Fernö M, Peterson C. "Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. *Eur J Cancer.* 2004; 40:1837-41.
13. Nimeus-Malmström E, Krogh M, Malmström P, et al. Gene expression profiling in primary breast cancer distinguishes patients developing local recurrence after breast-conservation surgery, with or without postoperative radiotherapy. *Breast Cancer Res.* 2008; 10:1-11.
14. Chelinska M, Kowalewska M. Systemic inflammation as a confounding factor in cancer biomarker discovery and validation. *Nature* 2010; 10;2.
15. Sharma P, Sahni NS, Tibshirani R, et al. Early detection of breast cancer based on gene-expression patterns in peripheral blood. *Breast Cancer Res.* 2005; 7:634-44.
16. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003; 13:198-207.
17. Ekins RP. Ligand assays: from electrophoresis to miniaturized microarrays. *Clin. Chem.* 1998; 44:2015-2030.
18. Zhu H, Snyder M. Protein chip technology. *Curr Opin Chem Biol* 2003; 7:55-63.
19. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* 2002; 1:845-867.
20. Borrebaeck CAK, Wingren C. Design of high-density antibody microarrays for disease proteomics: Key technological issues. *J Proteomics* 2009; 72:928-935.
21. Steinhauer C, Wingren C, Malmberg Hager AC y Borrebaeck CAK. Single framework recombinant antibody fragments designed for protein chip applications. *Biotechniques* 2002; Supl.: High-Throughput Proteomics: Protein Arrays: 38-45.
22. Wingren C, Steinhauer C, Ingvarsson J, et al. Microarrays based on affinity-tagged scFv antibodies: sensitive detection of analyte in complex proteomes. *Proteomics* 2005; 5:1281-1291.
23. Borrebaeck CAK y Wingren C. High-throughput proteomics using antibody microarrays: an update. *Expert Rev Mol Diagn.* 2007; 7:673-686.
24. Söderlind E, Strandberg L, Jirholt P, et al. Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat Biotechnol* 2000; 18:852-856.
25. Ingvarsson J, Larsson A, Sjöholm AG, et al. Design of recombinant antibody microarrays for serum protein profiling: targeting of complement proteins. *J. Proteome Res.* 2007; 6:3527-36.

26. Hamelinck D, Zhou H, Li L, et al. Optimized normalization for antibody microarrays and application to serum-protein profiling. *Mol Cell Proteomics* 2005; 4:773-84.

Tabla 1

| <b>(A) Biomarcadores centrales</b>  |   |
|---|---|
| <b>Nombre del biomarcador</b>   | <b>Secuencia(s) a modo de ejemplo</b>     |
| Apolipoproteína A4 (APOA4)  | P06727                                    |
| Subunidad beta de la ATP sintasa, mitocondrial (ATP5B)                      | P06576                                    |
| <b>(B) Firma de biomarcadores preferidos</b>                                |   |
| <b>Nombre del biomarcador</b>   | <b>Secuencia(s) a modo de ejemplo</b>     |
| Tirosina quinasa de Bruton (BTK)  | Q06187                                    |
| Inhibidor de la esterasa del complemento-1 (inhibidor de la esterasa de C1) | P05155                                    |
| Factor B del complemento (Factor B)   | P00751                                    |
| CD40  | Q6P2H9                                    |
| Factor de transcripción de homeodominio CHX10 (CHX10)                       | P05155                                    |
| Interleuquina-1 $\alpha$ (IL-1 $\alpha$ )                                   | P01583                                    |
| Interleuquina-5 (IL-5)  | BC066282, CH471062, P05113                |
| Interleuquina-6 (IL-6)  | P05231                                    |
| Interleuquina-7 (IL-7)  | AK226000, AB102893, AB102885, P13232      |
| Interleuquina-9 (IL-9)  | P15248                                    |
| Interleuquina-12 (IL-12)  | O60595                                    |
| Interleuquina-13 (IL-13)  | P35225                                    |
| Interleuquina-18 (IL-18)  | Q14116                                    |
| Miembro 9 de la familia del dominio TBC1 (KIAA0882)                         | Q6ZT07                                    |
| Lewis <sup>x</sup> /CD15  | [No aplicable]                            |
| Proteína de unión a oxisterol tipo 3 (OSBPL3)                               | C9JEF2, C9JZ19, Q9H4L5, C9J8P4            |
| Proteína-1 quimiotáctica de monocitos (MCP-1)                               | P13500                                    |
| Sialil Lewis <sup>x</sup>   | Estructura de carbohidrato [No aplicable] |
| Factor de necrosis tumoral-1 $\beta$ (TNF- $\beta$ )                        | P01374                                    |
| <b>(C) Biomarcadores adicionales opcionales</b>                             |   |
| <b>Nombre del biomarcador</b>   | <b>Secuencia(s) a modo de ejemplo</b>     |
| Integrina alfa-10   | Hs.158237                                 |
| $\beta$ -galactosidasa  | P16278                                    |
| Ligando de CD40   | P29965                                    |
| Eotaxina  | P51671                                    |
| Receptor del péptido-1 tipo glucagón (GLP-1 R)                              | P43220                                    |
| Interleuquina-1 $\beta$   | P01584                                    |
| Interleuquina-3   | P08700                                    |
| Interleuquina-8   | CR623827, CR623683, DQ893727,             |
| <b>Nombre del biomarcador</b>   | <b>Secuencia(s) a modo de ejemplo</b>     |
|   | DQ890564, P10145                          |
| Interleuquina-10  | P22301                                    |
| Interleuquina-16  | Q05BE6, Q8IUU6, B5TY35                    |
| MCP-4   | Q99616                                    |
| Miomesina-2 (MYOM2)   | P54296                                    |
| Rantes  | P13501                                    |
| Factor- $\beta$ de crecimiento transformante (TGF- $\beta$ 1)               | P01137                                    |
| Regulador de transcritos sin sentido 3B (UPF3B)                             | QOVAK7                                    |
| Interleuquina-4   | P05112                                    |
| Interleuquina-1ra   | P18510                                    |

**Tabla 2**

Los siguientes parámetros se obtuvieron usando la SVM e1071 1.5-24, disponible en <http://cran.r-project.org/web/packages/e1071/index.html>.

|    |  |                                |                                       |                                   |                           |                                |
|----|--|--------------------------------|---------------------------------------|-----------------------------------|---------------------------|--------------------------------|
| 5  | \$tipo<br>[1] 0  |                                |                                       |                                   |                           |                                |
| 10 | \$núcleo<br>[1] 0  |                                |                                       |                                   |                           |                                |
|    | \$coste<br>[1] 1   |                                |                                       |                                   |                           |                                |
| 15 | \$grado<br>[1] 3   |                                |                                       |                                   |                           |                                |
| 20 | \$gamma<br>[1] 0,04761905  |                                |                                       |                                   |                           |                                |
|    | \$coef0<br>[1] 0   |                                |                                       |                                   |                           |                                |
| 25 | \$nu<br>[1] 0,5  |                                |                                       |                                   |                           |                                |
|    | \$épsilon<br>[1] 0,1   |                                |                                       |                                   |                           |                                |
| 30 | \$esparcido<br>[1] FALSO   |                                |                                       |                                   |                           |                                |
|    | \$ponderado  |                                |                                       |                                   |                           |                                |
| 35 | [1] CIERTO CIERTO<br>CIERTO CIERTO CIERTO CIERTO CIERTO CIERTO CIERTO CIERTO CIERTO |                                |                                       |                                   |                           |                                |
|    | \$x.escala   |                                |                                       |                                   |                           |                                |
|    | \$x.escala\$'ponderado:centro'   |                                |                                       |                                   |                           |                                |
|    | EG22.260123.   | IL.7..1.<br>APOA4<br>0,6093909 | TNF.b..2.<br><br>0,4888461            | EG7.141932.CHX10<br><br>0,5194064 | BTK<br><br>0,3885171      | CD40..4.<br><br>0,4091274      |
|    | 0,5693457<br>EG23.252478.<br>IL.12..3.   | KIAA0882<br><br>0,4535247      | EG25.252336.anti.MCL<br><br>0,5956616 | EG9.350100.ATP5B<br><br>0,5668468 | EI<br><br>0,5194224       | Factor.B..2..<br><br>0,5307535 |
|    | 0,4051029<br>IL.9..2.  | IL.13..3.<br><br>0,5284938     | IL.18..1.<br><br>0,4075885            | IL.1a..3.<br><br>0,5941146        | IL.5..2.<br><br>0,5428361 | IL.6..1.<br><br>0,3743462      |
|    | 0,5567893  | Lewisx..1.<br>0,5204801        | MCP.1..3.<br>0,5608207                | Sialil.x<br>0,5048063             |                           |                                |
| 40 | \$x.escala\$"ponderado:escala'   |                                |                                       |                                   |                           |                                |
|    | EG22.260123.   | IL.7..1.<br>APOA4<br>0,2234510 | TNF.b..2.<br><br>0,2340201            | EG7.141932.CHXIO<br><br>0,2253392 | BTK<br><br>0,1617694      | CD40..4.<br><br>0,2003829      |
|    | 0,1947099<br>EG23.252478.<br>IL.12..3.   | KIAA0882<br><br>0,1673467      | EG25.252336.anti.MCL<br><br>0,2073111 | EG9.350100.ATP5B<br><br>0,2686860 | EI<br><br>0,2275948       | Factor.B..2..<br><br>0,2338916 |
|    | 0,1468782<br>IL.9..2.  | IL.13..3.<br><br>0,2020372     | IL.18..1.<br><br>0,1715512            | IL.1a..3.<br><br>0,2541235        | IL.5..2.<br><br>0,2015772 | IL.6..1.<br><br>0,2359194      |

ES 2 608 956 T3

0,2176524

Lewisx..1.  
0,2520771

MCP.1..3.  
0,2209286

Sialil.x  
0,2607917

\$y.escala  
NULO

5 \$nclases  
[1] 2

10 \$niveles  
[1] "MetDistante" "NoMetDistante" "reposito"

\$tot.nSV  
[1] 16

15 \$nSV  
[1] 5 11

\$marcadores  
[1] 1 2

20 \$SV

|               | IL.7..1.    | TNF.b..2.   | EG7.141932. | BTK        | CD40..4.    | EG22.260123. | EG23.252478. |
|---------------|-------------|-------------|-------------|------------|-------------|--------------|--------------|
| EG25.252336.a |             |             | CHX10       |            |             | APOA4        | KIAA0882     |
| nti.MCL       |             |             |             |            |             |              |              |
| BC104         | -1,72878005 | 0,16404151  | 0,08194673  | -1,4821892 | -0,20945935 | -0,008551152 | 0,56474259   |
| 0,647552034   |             |             |             |            |             |              |              |
| BC152         | -1,18673489 | 0,16789721  | -0,37353361 | -0,2558364 | -0,07045099 | -0,300217895 | -0,13423148  |
| -0,255316821  |             |             |             |            |             |              |              |
| BC156         | 1,07470984  | -1,25840379 | 0,14648518  | 0,5124307  | 0,43604106  | 0,701639774  | 0,30184949   |
| 0,112261016   |             |             |             |            |             |              |              |
| BC224         | -0,38641301 | -0,38465450 | 0,58904238  | -0,2198718 | 0,18262343  | 0,971031555  | 0,51977281   |
| 0,556924433   |             |             |             |            |             |              |              |
| BC70          | 0,49494080  | 0,02968846  | -0,02105569 | 1,4284150  | 0,55805233  | 1,001445091  | -0,17663963  |
| 0,309008339   |             |             |             |            |             |              |              |
| BC119         | 0,68089956  | -0,26970554 | 0,01709544  | 1,2594195  | 0,80095567  | 0,985860689  | 0,32059607   |
| 0,962229843   |             |             |             |            |             |              |              |
| BC123         | -1,47972460 | -0,64487597 | 0,23352593  | -2,4016723 | -0,30192272 | -0,613369998 | 1,31842626   |
| 0,791827496   |             |             |             |            |             |              |              |
| BC162         | 0,08652534  | -0,49958930 | -0,62779888 | 0,4662603  | -1,62702240 | -1,265192129 | -0,63413299  |
| -1,026124999  |             |             |             |            |             |              |              |
| BC212         | -0,25019319 | 0,62731854  | -0,15142958 | -1,3452997 | -2,04172864 | -0,862262747 | 0,12307725   |
| -0,006164769  |             |             |             |            |             |              |              |
| BC41          | -0,56280296 | -0,20014615 | 0,81993713  | -0,2446241 | 0,40022261  | 1,011699296  | 1,81890868   |
| 0,941110371   |             |             |             |            |             |              |              |
| BC44          | 0,34946805  | -0,35003605 | -0,28882815 | -0,2513542 | 0,43765530  | -0,729949938 | -0,31092196  |
| -0,205794722  |             |             |             |            |             |              |              |
| BC57          | 1,07732967  | 1,42323505  | 2,12950113  | -0,4789719 | 0,67120469  | 1,438276723  | 3,26552710   |
| 1,931485894   |             |             |             |            |             |              |              |
| BC82          | -0,85194455 | -0,02634785 | -0,07671850 | -1,0847651 | 0,29761712  | -0,037555372 | 0,35122821   |
| -0,190928380  |             |             |             |            |             |              |              |
| BC87          | -0,09104134 | 0,55117224  | -0,21032290 | 0,3012652  | 0,24433613  | -0,386117835 | -0,71641769  |
| -0,673912564  |             |             |             |            |             |              |              |
| BC92          | 0,61584785  | 1,38011151  | 0,70604393  | 0,1980872  | 0,37732185  | 0,401803985  | 0,01251850   |
| 0,340860075   |             |             |             |            |             |              |              |
| BC98          | -0,09870669 | -2,08890693 | 0,53803970  | -0,2163412 | -0,79686632 | 0,438660744  | 0,50131918   |
| -0,334405791  |             |             |             |            |             |              |              |

EG9.350100.ATP5B EI Factor.B..2. IL.12..3. IL.13..3. IL.18..1. IL.1a..3. IL.5..2.

| IL.6..1.    | IL.9..2.     |             |            |            |             |             |            |             |
|-------------|--------------|-------------|------------|------------|-------------|-------------|------------|-------------|
| BC104       | -0,06841983  | -0,79806974 | -0,4818731 | -0,4980696 | -1,00268354 | -0,24927827 | -0,3977183 | -0,84185368 |
| -0,03695358 | -0,884569546 |             |            |            |             |             |            |             |
| BC152       | -0,81984741  | -0,46271285 | 0,4618282  | -0,3297756 | -0,27851745 | 0,11636339  | 0,4584601  | -0,95665226 |
| -1,58675456 | -0,859445462 |             |            |            |             |             |            |             |

# ES 2 608 956 T3

```

BC156      1,20529203  0,67424741 -0,8271355 0,6113120 0,25845095 -0,04614928 -0,1349133 0,48169777
-1,36041255 0,394145570
BC224      1,61211669  0,06826851  0,1333717 -0,4364411 -0,02215751 -2,37589954 0,3223370 -0,06663768
-1,35422913 -1,295940601
BC70       0,52641522  1,03788753  1,6460879 0,8074381 0,58898984 0,39387896 1,1476412 0,30869312
0,60644973 0,881224796
BC119      0,60826765  0,62258224 -2,1781676 0,8169730 2,33375910 0,27336511 1,0704331 0,21547230
-1,07250229 2,004479796
BC123      -0,24189341 -2,28222440 1,2468878 -2,7580871 -1,54337127 -0,71185770 -1,1497185 -1,75367819
-1,10930970 -2,558158402
BC162      -1,61374915 -0,73716455 -0,1773728 0,4033345 -0,86244597 -0,55776651 -0,9937919 -0,96602749
0,69835443 -0,250338305
BC212      -0,02665158 -1,43219060 0,1276030 -0,3183847 -0,48332555 -0,04380831 -2,1500522 -0,43269681
0,03655087 -1,532758825
BC41       0,68983913  1,46753304  0,3906515 -0,3938307 -0,44744993 3,45326238 0,2856195 0,77145772
0,60966231 -0,262739745
BC44       -0,85059772 -0,41608430 -0,9024877 -0,3314940 -0,02539801 -0,48721493 0,4488824 1,27502167
0,63585407 0,510734350
BC57       0,94863487  -1,55045128 -0,7237189 -0,2999112 -1,32472302 -0,15814942 -0,7699348 -1,08134602
0,41453433 0,007046511
BC82       0,44597738  -0,23940059 0,2982282 -0,7716297 0,09473559 0,31271803 0,8128847 -0,40799788
-0,48090892 -0,275352403
BC87       -0,43323779 0,94550916 0,8373675 1,0839010 -0,30577379 -0,39521361 0,7705388 -0,26220576
0,96962268 0,729620447
BC92       0,69062228  -1,09554932 0,1361003 -0,3563486 0,08877688 -0,08982865 -0,1900476 -0,31901322
0,64425431 -0,230280344
BC98       0,86837302  -0,21731264 0,2780535 -1,1415435 2,18218964 -0,11678834 -0,0551481 0,07821912
1,12818573 0,156899742

```

|       | Lewisx..1.  | MCP.1.,.3.    | Sialil.x    |
|-------|-------------|---------------|-------------|
| BC104 | 0,04913151  | -1,1699038164 | 0,48580074  |
| BC152 | 0,19352752  | -0,7047500356 | 0,40039174  |
| BC156 | -0,72699766 | 0,1762927414  | -0,78994670 |
| BC224 | -2,06476589 | 1,1772446443  | -0,17406871 |
| BC70  | 0,45931367  | 1,7611954453  | -1,03576811 |
| BC119 | -1,73469677 | 1,4619597768  | 1,21056825  |
| BC123 | -0,92230494 | -1,8412777307 | 1,58508822  |
| BC162 | 0,07699985  | 0,0406312609  | 1,57473207  |
| BC212 | 0,19392949  | -0,3019735097 | 0,81326702  |
| BC41  | -0,19053962 | 0,2703781740  | -0,06397621 |
| BC44  | 1,01036207  | -0,7046768943 | 0,51387072  |
| BC57  | -0,32801357 | 0,6285212119  | -0,28667403 |
| BC82  | -0,25212082 | -0,3259224211 | -0,18688331 |
| BC87  | -0,32483706 | 0,8884765739  | -1,77813625 |
| BC92  | -0,10882989 | 0,0009727735  | -0,49132772 |
| BC98  | -1,38143674 | -0,3223663269 | 0,66550080  |

\$índice

[1] 1 5 6 15 16 19 20 23 25 27 28 30 35 36 37 38

5

\$rho

[1] 0,2847347

\$sondacom

[1] CIERTO

10

\$sondaA

[1] -1,094657

15

\$sondaB

[1] 0,2128045

\$sigma

[1] 0

20

\$coefs

[1,] 0,79306050  
 [2,] 0,81101350  
 [3,] 0,50522912  
 5 [4,] 0,08584558  
 [5,] 0,62204473  
 [6,] -0,28084545  
 [7,] -0,43343460  
 10 [8,] -0,08939337  
 [9,] -0,04883494  
 [10,] -0,28094053  
 [11,] -0,19803954  
 [12,] -0,01242534  
 [13,] -0,40895362  
 15 [14,] -0,87795820  
 [15,] -0,14985990  
 [16,] -0,03650794

\$na.acción  
 20 NULO

| \$ajustado    | BC104         | BC140         | BC146         | BC149         | BC152         | BC156         | BC163         | BC171         | BC178         |
|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| BC181         | BC193         |               |               |               |               |               |               |               |               |
| MetDistante   |
| MetDistante   | MetDistante   |               |               |               |               |               |               |               |               |
| BC198         | BC208         | BC221         | BC224         | BC70          | BC108         | BC115         | BC119         | BC123         |               |
| BC128         | BC135         |               |               |               |               |               |               |               |               |
| MetDistante   | MetDistante   | MetDistante   | MetDistante   | MetDistante   | NoMetDistante | NoMetDistante | NoMetDistante | NoMetDistante | NoMetDistante |
| NoMetDistante | NoMetDistante |               |               |               |               |               |               |               |               |
| BC162         | BC185         | BC212         | BC217         | BC41          | BC44          | BC48          | BC57          | BC62          |               |
| BC66          | BC74          |               |               |               |               |               |               |               |               |
| NoMetDistante |
| NoMetDistante | NoMetDistante |               |               |               |               |               |               |               |               |
| BC78          | BC82          | BC87          | BC92          | BC98          |               |               |               |               |               |
| NoMetDistante | NoMetDistante | NoMetDistante | NoMetDistante | NoMetDistante |               |               |               |               |               |

Niveles: MetDistante NoMetDistante reposo

**Tabla 3.** Demográfica de pacientes y parámetros clínicos

| Grupo  | Cohorte de descubrimiento |           | Cohorte de prevalidación |           |
|--|---------------------------|-----------|--------------------------|-----------|
|  | Met                       | No Met    | Met                      | No Met    |
| Número de pacientes                                  | 16                        | 22        | 13                       | 13        |
| Seguimiento sin enfermedad (3/5/7 años) <sup>a</sup> | -                         | 4/14/4    | -                        | -/12/1    |
| Edad   | 54 (14) <sup>b</sup>      | 52 (12)   | 59 (13)                  | 60(9)     |
| Muestra 2, tiempo de recogida después de OP (meses)  | 4,4 (1,1)                 | 4,5 (1,1) | 4,9 (2,1)                | 4,0 (1,0) |
| Muestra 3, tiempo de recogida después de OP (meses)  | 13 (1,0)                  | 13 (1,2)  | -                        | -         |
| Tiempo hasta recidiva (meses)                        | 21 (10)                   | -         | 19 (13)                  | -         |
| Tamaño del tumor (mm)                                | 22 (10)                   | 21 (11)   | 38 (23)                  | 16 (11)   |
| Pre/posmenopáusico                                   | 5/11                      | 10/12     | 9/4                      | 11/2      |
| ER+/-  | 13/2 <sup>c</sup>         | 18/4      | 9/4                      | 11/1      |
| PgR+/-   | 9/6                       | 13/9      | 7/6                      | 7/5       |
| Ganglio linfático +/-                                | 12/4                      | 10/11     | 6/5                      | 3/9       |
| Grado I/M/IN   | 1/9/5                     | 3/13/5    | 0/9/4                    | 5/6/1     |
| Ductal/lobular                                       | 16/1 <sup>d</sup>         | 20/2      | 11/2                     | 11/2      |
| Radiación/sin radiación                              | 11/5                      | 14/8      | 11/2                     | 9/4       |
| Terapia adyuvante: hormonal (tamoxifeno)             | 12/4                      | 16/6      | 6/7                      | 7/6       |
| Terapia adyuvante: inhibidor de aromatasa            | 6/10                      | 3/19      | 4/9                      | 4/9       |

<sup>a</sup>Tiempo de seguimiento sin enfermedad, para paciente no metastásicos.

<sup>b</sup>Los valores en paréntesis son desviaciones típicas.

<sup>c</sup>En casos donde la suma es menor que la cantidad en el grupo, los datos de los pacientes están ausentes.

<sup>d</sup>En el caso de tipo de tumor ductal/lobular, los pacientes pueden tener ambos, produciendo una suma mayor que la cantidad de pacientes en el grupo.

**Tabla 4.** Sumario de biomarcadores analizados por la microserie

| <i>Antígeno (n.º de clones)</i>     | <i>Antígeno (n.º de clones)</i> |
|-------------------------------------|---------------------------------|
| Alfa-10 (1)                         | IL-8 (3)                        |
| Alfa-11 (1)                         | IL-9 (3)*                       |
| Angiomotina (2)                     | IL-10 (3)**                     |
| APOA4 (3)*                          | IL-11 (3)                       |
| ATPSB (3)*                          | IL-12 (4)*                      |
| β-galactosidasa (1)                 | IL-13 (2)*/**                   |
| BTK (1)*                            | IL-16 (2)                       |
| C1q (1)**                           | IL-18 (3)*                      |
| C1s (1)                             | INF-γ (2)                       |
| Inhibidor de la esterasa de C1 (1)* | JAK3 (1)                        |
| C3 (2)**                            | KIAA0882 (3)*                   |
| C4 (1)                              | LDL (2)                         |
| C5 (2)                              | Leptina (1)                     |
| CD40 (4)*                           | Lewis <sup>x</sup> (2)*         |
| Ligando de CD40 (1)                 | Lewis <sup>y</sup> (1)          |
| CHX10 (3)*                          | LUM (1)                         |
| Digoxina (1)                        | OSBPL3 (2)*                     |
| DUSP9 (1)                           | MCP-1 (3)**/**                  |
| Eotaxina (3)                        | MCP-3 (1)                       |
| Factor B (2)*/**                    | MCP-4 (2)                       |
| GLP-1 (1)                           | MYOM2 (2)                       |
| GLP-1 R (1)                         | Procatepsina (1)                |
| GM-CSF (3)                          | Properdina (1)                  |
| IL-1α (3)*                          | PSA (1)                         |
| IL-1β (3)                           | Rantes (1)                      |
| IL-1ra (3)                          | Sialil Lewis <sup>x</sup> (1)*  |
| IL-2 (3)                            | TGF-β1 (2)                      |
| IL-3 (3)                            | Péptido TM (1)                  |
| IL-4 (3)**                          | TNF-α (2)                       |
| IL-5 (3)*/**                        | TNF-β (4)*                      |
| IL-6 (4)*/**                        | UPF3B (3)                       |
| IL-7 (2)*                           | VEGF (4)                        |

\* Incluido en la firma de 21 biomarcadores

\*\* La especificidad del anticuerpo contra estos antígenos se validó adicionalmente, por espectrometría de masas, series de proteínas, o ELISA.

REIVINDICACIONES

1. Un método para el pronóstico de cáncer de mama en un sujeto, que comprende las etapas de:

- 5 (a) proporcionar una primera muestra del proteoma del sujeto;
- (b) medir en la primera muestra del proteoma la cantidad de dos o más biomarcadores seleccionados del grupo de biomarcadores de la Tabla 1;
- (c) proporcionar una muestra adicional del proteoma del sujeto;
- 10 (d) medir en la muestra adicional del proteoma la cantidad de los dos o más biomarcadores medidos en la etapa (b);
- (e) determinar la diferencia entre la cantidad de los dos o más biomarcadores en la primera muestra del proteoma y la muestra adicional del proteoma;

15 en el que la primera muestra del proteoma y la muestra adicional del proteoma son representativas de la composición del proteoma del sujeto en diferentes días,  
 en el que la diferencia entre la cantidad de los dos o más biomarcadores en la primera muestra del proteoma y la adicional es indicativa del riesgo de recidiva y/o metástasis distante de cáncer de mama en el sujeto,  
 en el que la primera muestra del proteoma del sujeto es representativa de la composición del proteoma del sujeto dentro de las 4 semanas antes o después de la resección del tumor del individuo;

20 en el que la muestra adicional del proteoma del individuo proporcionada en la etapa (c) es representativa de la composición del proteoma del sujeto en un punto temporal dentro de los 3 a 9 meses desde la resección del tumor del individuo; y  
 en el que la etapa (b) comprende medir la apolipoproteína A4 (APOA4) y/o la subunidad beta de la ATP sintasa, mitocondrial (ATP5B).

25 2. El método de acuerdo con la reivindicación 1, en el que se repiten las etapas (c) y (d) para proporcionar dos o más muestras adicionales del proteoma del sujeto y para medir en las mismas la cantidad de los dos o más biomarcadores medidos en la etapa (b), en donde las dos o más muestras adicionales del proteoma son representativas de la composición del proteoma del sujeto en diferentes días de la primera y la segunda muestras del proteoma.

30 3. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que la etapa (b) comprende medir la cantidad de 1 o más biomarcadores de los biomarcadores enumerados en la Tabla 1(B), por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 o 19 biomarcadores enumerados en la Tabla 1(B).

35 4. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que la etapa (b) comprende medir la cantidad de 1 o más biomarcadores de los biomarcadores enumerados en la Tabla 1(C), por ejemplo, al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o los 15 biomarcadores enumerados en la Tabla 1(C).

40 5. El método de acuerdo con la reivindicación 4, en el que la etapa (b) comprende medir la cantidad de todos los biomarcadores enumerados en la Tabla 1.

45 6. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que la etapa (b) y/o la etapa (d) se realizan usando dos o más agentes de unión, cada uno capaz de unirse a un biomarcador medido en la etapa (b).

50 7. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que la primera muestra del proteoma, proporcionada en la etapa (a), es representativa de la composición del proteoma del sujeto en un punto temporal dentro de aproximadamente 4 semanas antes o después de la resección de un tumor de cáncer de mama en el sujeto; y/o  
 en el que la muestra adicional del proteoma es representativa de la composición del proteoma del sujeto en un punto temporal dentro de 3 a 6 meses después de la resección del tumor.

55 8. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que se repiten la etapa (c) y la etapa (d) de modo que se proporcionan muestras adicionales del proteoma y se miden los niveles de biomarcadores en las mismas, en donde las muestras adicionales del proteoma son representativas de la composición del proteoma del sujeto en un punto temporal cada 6 a 18 meses después de la resección del tumor, hasta al menos 24 meses.

60 9. El método de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde la precisión predictiva del método, determinada por un valor de AUC ROC, es de al menos 0,70, por ejemplo, al menos 0,75, 0,80 o al menos 0,85.

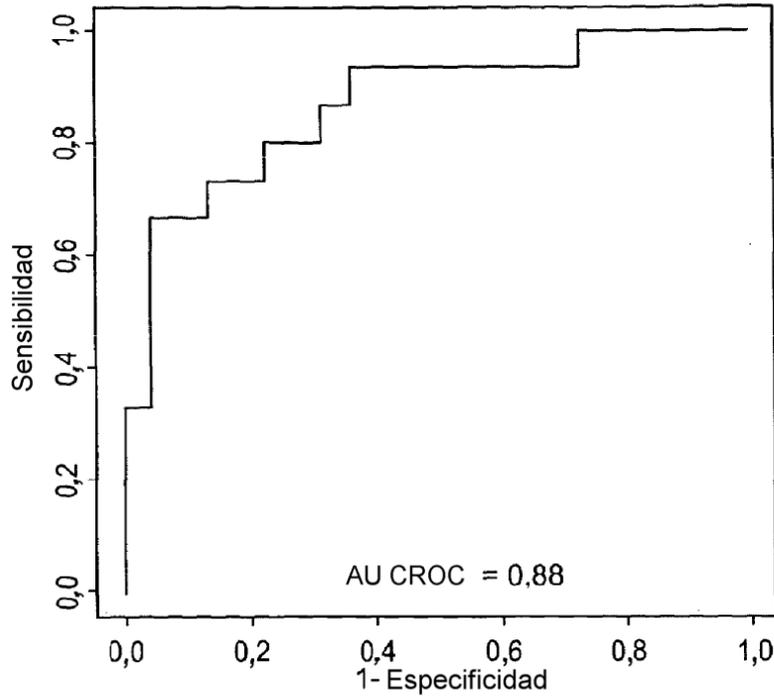
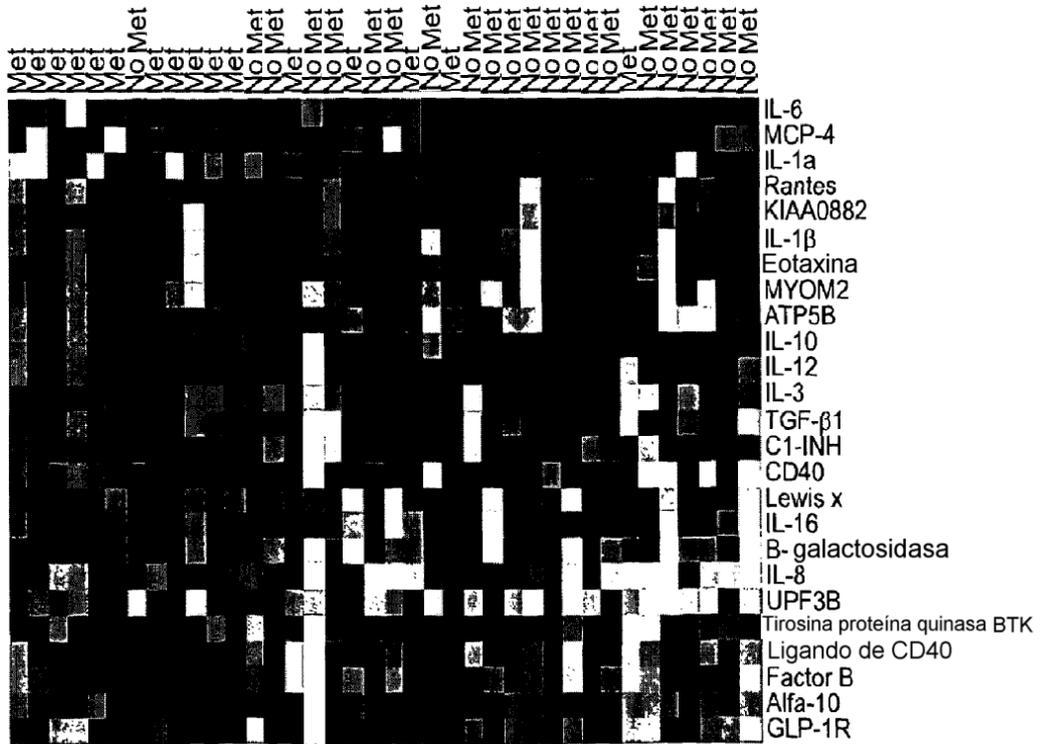


Figura 1A



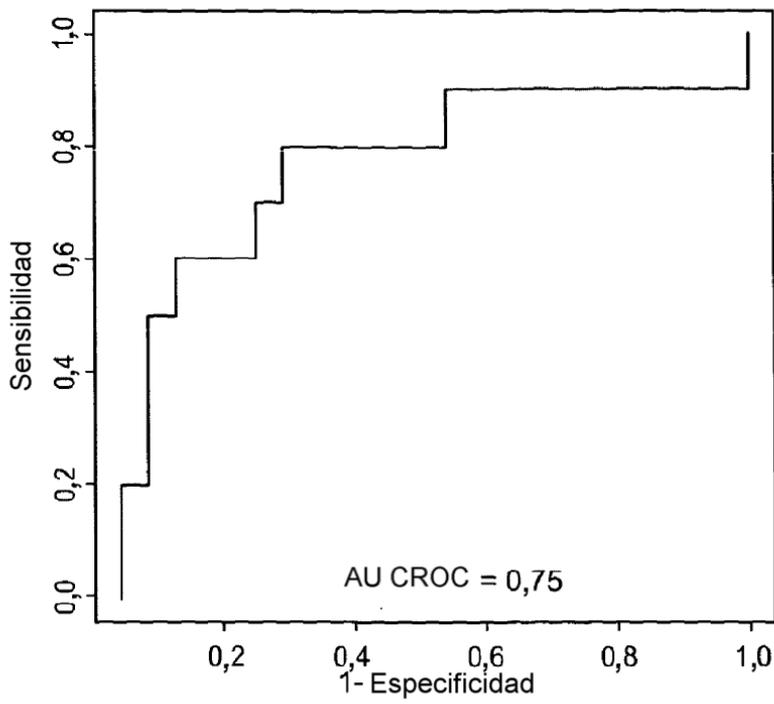
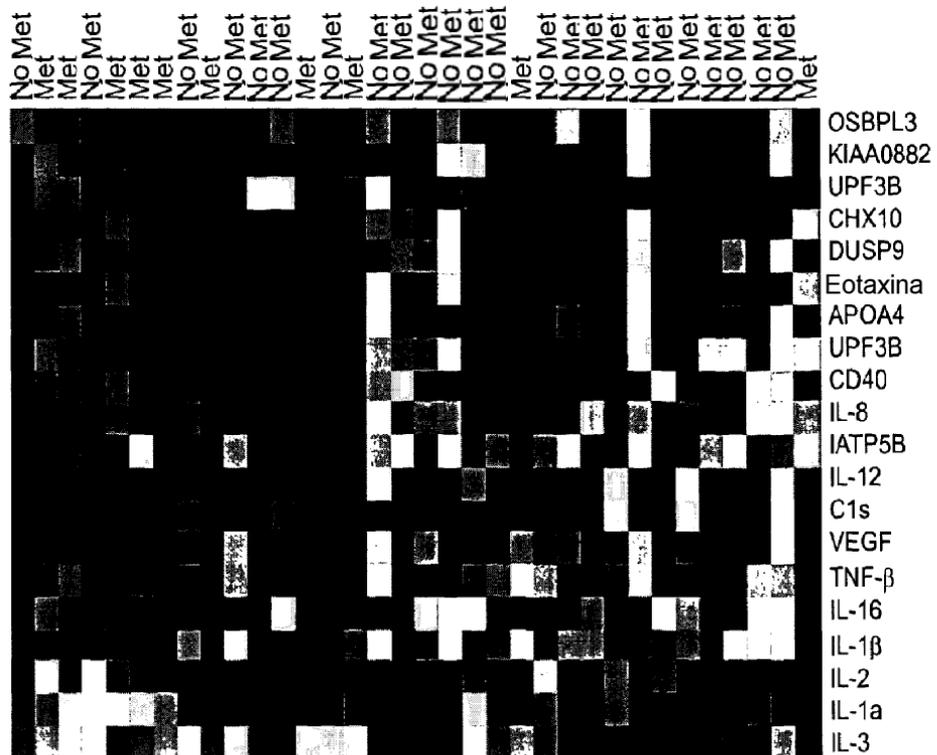


Figura 1B



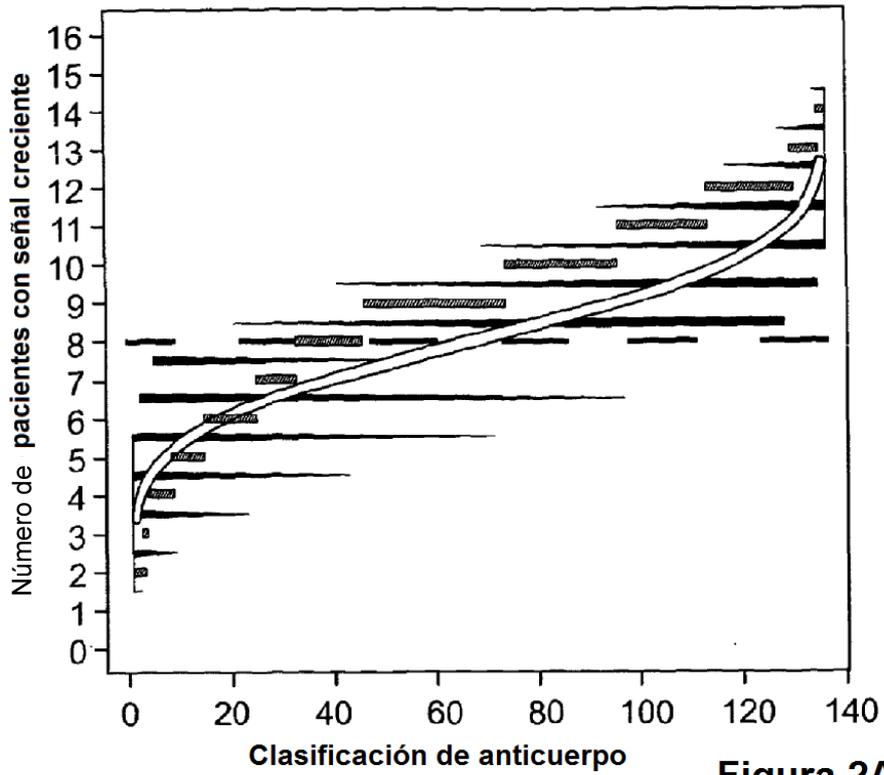


Figura 2A

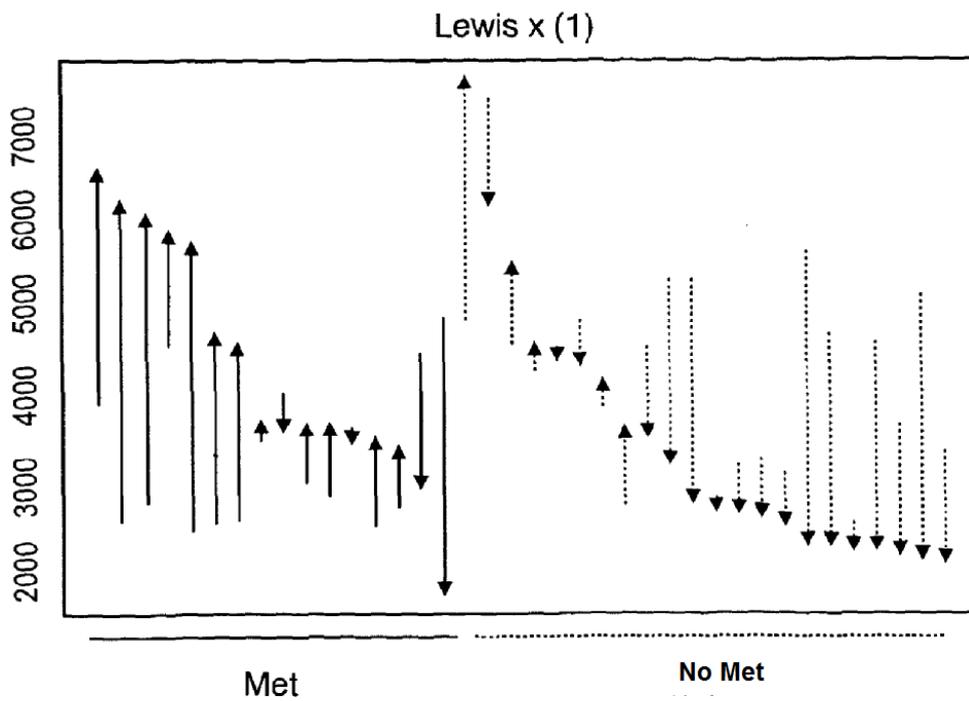


Figura 2B

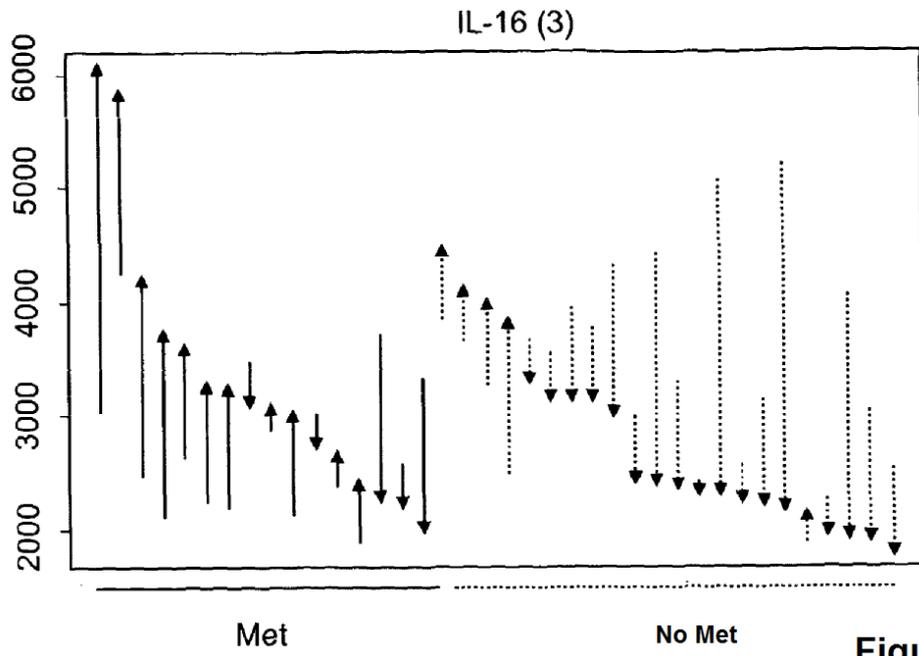


Figura 2C

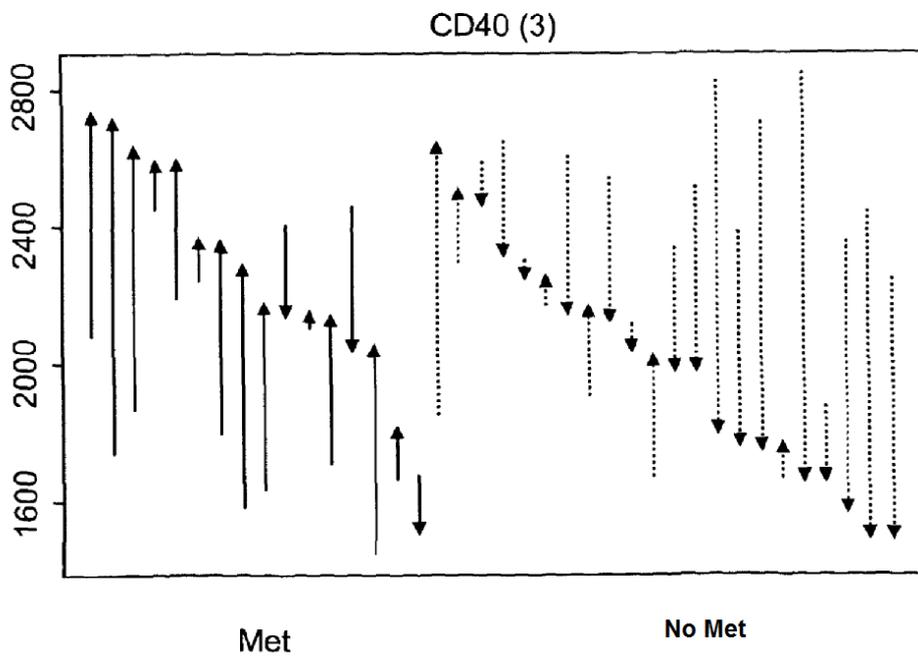


Figura 2D

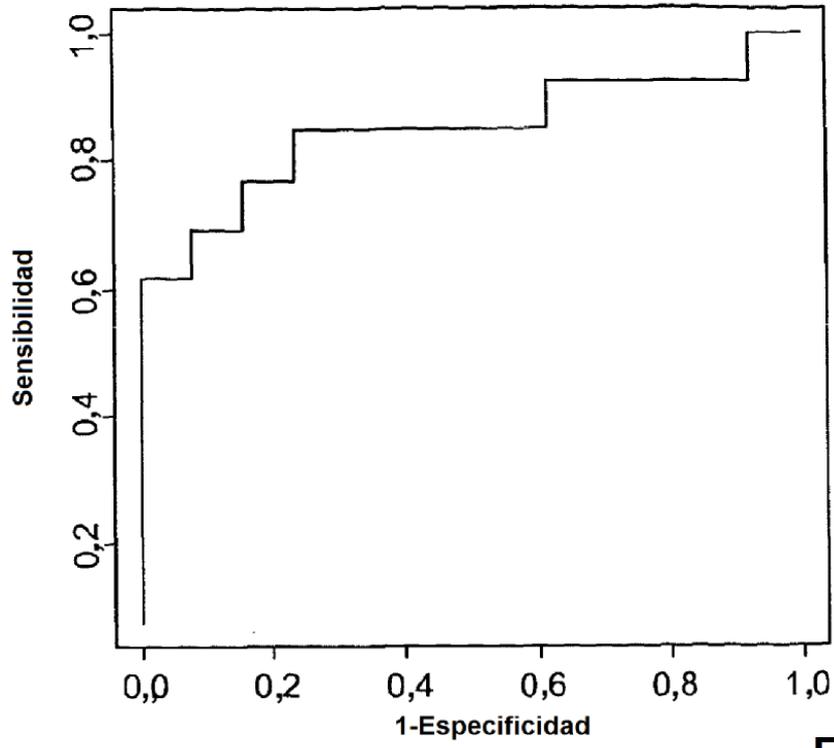


Figura 3A

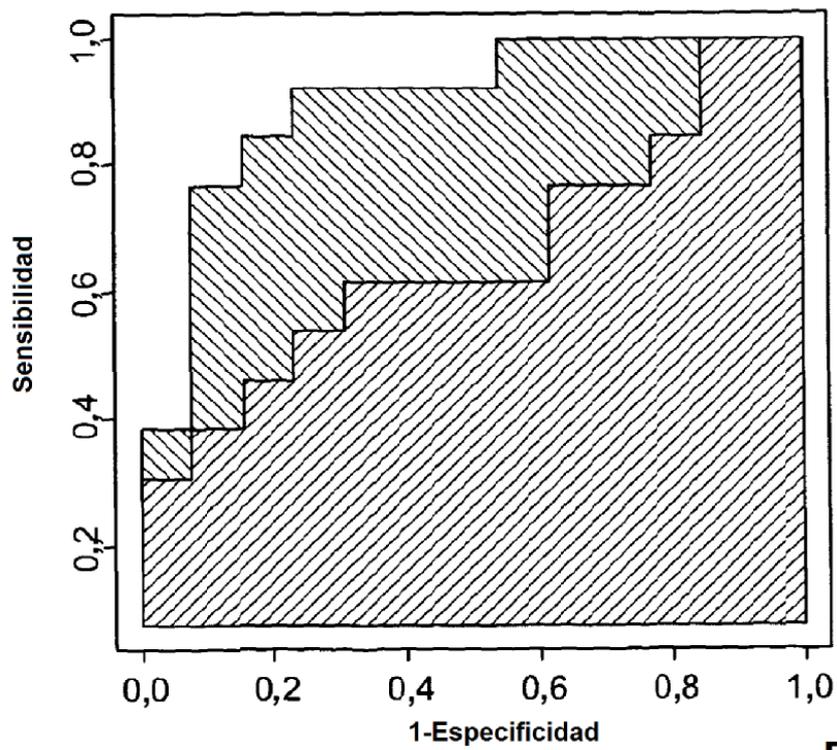


Figura 3B

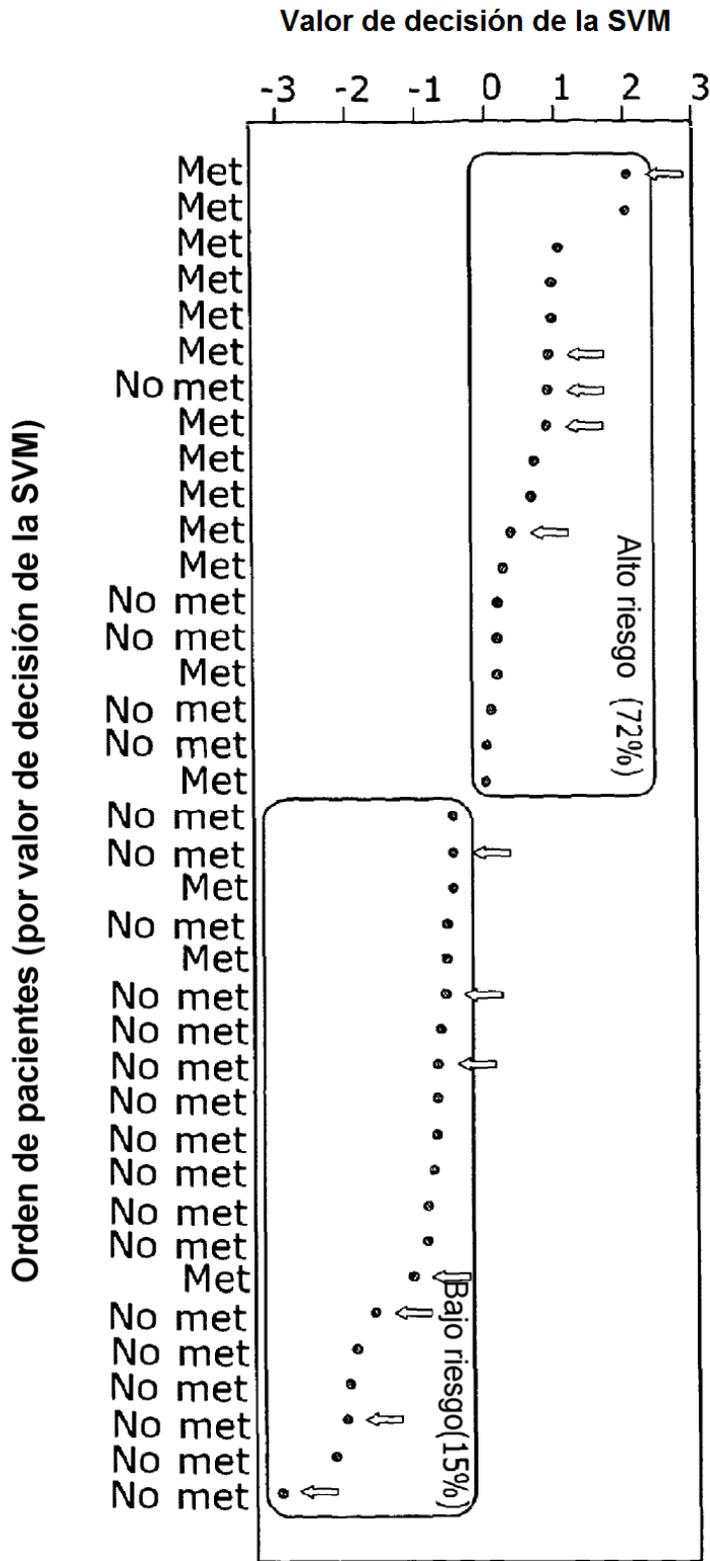


Figura 4

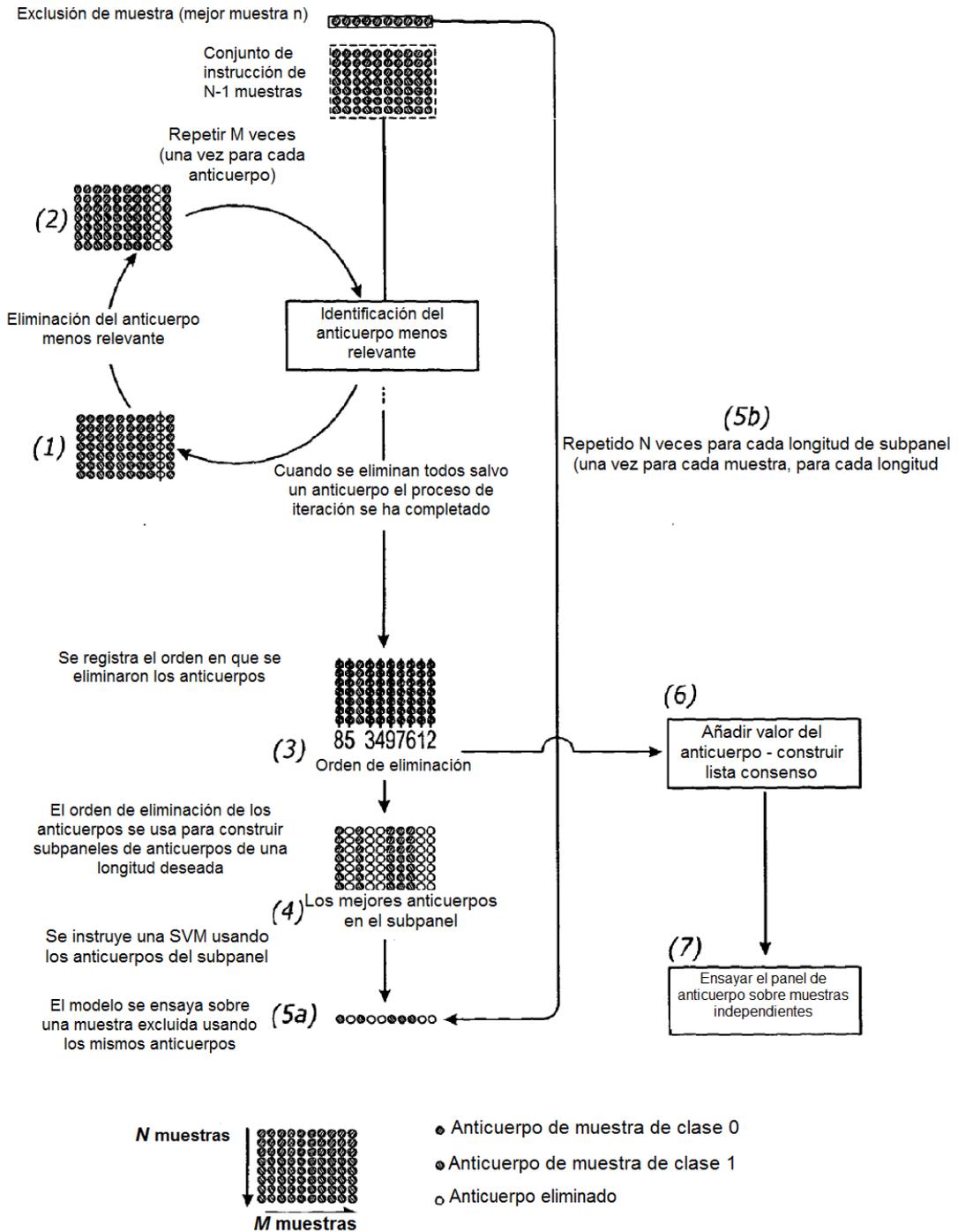


Figura 5