



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 346 202**

51 Int. Cl.:
G01N 33/574 (2006.01)
G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03767259 .9**
96 Fecha de presentación : **05.08.2003**
97 Número de publicación de la solicitud: **1540010**
97 Fecha de publicación de la solicitud: **15.06.2005**

54 Título: **Uso de biomarcadores para detectar cáncer.**

30 Prioridad: **06.08.2002 US 401837 P**
21.01.2003 US 441727 P
04.04.2003 US 460342 P

45 Fecha de publicación de la mención BOPI:
13.10.2010

45 Fecha de la publicación del folleto de la patente:
13.10.2010

73 Titular/es: **The Johns Hopkins University**
3400 N. Charles Street
Baltimore, Maryland 21218, US
Ciphergen Biosystems, Inc.

72 Inventor/es: **Chan, Daniel, W.;**
Zhang, Zhen;
Fung, Eric y
Meng, Xiao-Ying

74 Agente: **Carpintero López, Mario**

ES 2 346 202 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

Uso de biomarcadores para detectar cáncer.

5 **Campo de la invención**

La invención proporciona biomarcadores importantes en la detección del cáncer ovárico. Los marcadores se identificaron distinguiendo el perfil de proteína sérica en pacientes con cáncer ovárico respecto de los individuos sanos usando el análisis SELDI. La presente invención se refiere a los biomarcadores para un sistema y el procedimiento en el que se usan los biomarcadores para la calificación del estado del cáncer ovárico. La presente invención también identifica los biomarcadores como proteínas conocidas.

Antecedentes de la invención

El cáncer ovárico está entre las neoplasias ginecológicas más letales en los países desarrollados. Solo en los Estados Unidos, anualmente se diagnostican aproximadamente 23.000 mujeres con la enfermedad y casi 14.000 mujeres mueren debido a este. (Jamal, A., *et al.*, CA Cancer J. Clin, 2002; 52:23-47). A pesar del avance en la terapia del cáncer, la mortalidad del cáncer ovárico ha permanecido prácticamente sin modificaciones durante las últimas dos décadas. (Id.) Debido al gradiente de supervivencia considerable en relación con la etapa en que se diagnostica la enfermedad, la detección precoz es aún el factor más importante para aumentar la supervivencia a largo plazo de las pacientes con cáncer ovárico.

El mal pronóstico del cáncer ovárico diagnosticado en etapas tardías, el costo y riesgo asociado con los procedimientos de diagnóstico confirmatorios, y su relativamente baja prevalencia en la población general en conjunto plantean requerimientos extremadamente estrictos sobre la sensibilidad y especificidad de un ensayo para ser usado en la detección del cáncer ovárico en la población general.

La identificación de los marcadores tumorales adecuados para la detección y el diagnóstico precoz del cáncer mantiene la gran promesa de mejorar la evolución clínica de los pacientes. Es de especial importancia para las pacientes que presentan síntomas vagos o que no los tienen o con tumores que son relativamente inaccesibles en el examen físico. A pesar del esfuerzo considerable dirigido a la detección precoz, no se han desarrollado ensayos de detección de bajo costo (Paley PJ., Curr Opin Oncol, 2001; 13(5):399-402) y las mujeres generalmente se presentan con la enfermedad diseminada en el momento de efectuar el diagnóstico. (Ozols RF, *et al.*, Epithelial ovarian cancer. En: Hoskins WJ, Perez CA, Young RC, editors. Principles and Practice of Gynecologic Oncology. 3rd ed. Philadelphia: Lippincott, Williams y Wilkins; 2000. p.981-1057).

El marcador tumoral mejor caracterizado, CA125, es negativo en aproximadamente 30-40% de los carcinomas de ovarios de etapa I y sus niveles están elevados en una variedad de enfermedades benignas. (Meyer T, *et al.*, Br J Cancer, 2000;82 (9):1535-8; Buamah P., J Surg Oncol, 2000;75(4):264-5; Tuxen MK, *et al.*, Cancer Treat Rev, 1995;21(3):215-45). Su utilización como herramienta de identificación basada en la población para la detección y diagnóstico precoz del cáncer ovárico está dificultada por su baja sensibilidad y especificidad. (MacDonald ND, *et al.*, Eur J Obstet Gynecol Reprod Biol, 1999;82(2):155-7; Jacobs I, *et al.*, Hum Reprod, 1989;4(1):1-12; Shih I-M, *et al.*, Tumor markers in ovarian cancer. En: Diamandis EP, Fritsche, H., Lilja, H., Chan, D.W., a Schwartz, M., editor. Tumor markers physiology, pathobiology, technology and clinical applications. Philadelphia: AACCC Press; en prensa). Si bien más recientemente se ha usado la ecografía pélvica y vaginal para identificar pacientes de alto riesgo, ninguna de estas técnicas tiene la sensibilidad y especificidad suficientes para ser aplicadas en la población general. (MacDonald ND, *et al.*, *supra*). Recientes esfuerzos para usar CA125 en combinación con marcadores tumorales adicionales (Woolas RP XF, *et al.*, JNatl Cancer Inst, 1993;85 (21):1748-51; Woolas RP, *et al.*, Gynecol Oncol, 1995;59 (1):111-6; Zhang Z, *et al.*, Gynecol Oncol, 1999;73(1):56-61; Zhang Z, *et al.*, Use of Multiple Markers to Detect Stage I Epithelial Ovarian Cancer: Neural Network Analysis Improves Performance. American Society of Clinical Oncology 2001; Annual Meeting, Abstract) en un modelo de cáncer de riesgo longitudinal (Skates SJ, *et al.*, Cancer, 1995; 76(10 Suppl):2004-10), y en tándem con ultrasonido como una prueba de segunda línea (Jacobs I DA, *et al.*, Br Med J, 1993;306(6884):1030-34; Menon U TA, *et al.*, British Journal of Obstetrics y Gynecology, 2000;107(2):165-69) han demostrado resultados promisorios para aumentar la especificidad de la prueba global, lo cual es crítico para una enfermedad tal como el cáncer ovárico que posee una prevalencia relativamente baja.

Debido al pésimo pronóstico del cáncer ovárico en etapa tardía, es de consenso general que un médico aceptará una prueba con un valor predictivo positivo mínimo del 10%. (Bast, R.C., *et al.*, Cancer Tratamiento y Research, 2002; 107: 61-97). Si se extiende esto a la población general, una prueba de detección general debería requerir una sensibilidad mayor del 70% y una especificidad del 99,6%. En la actualidad, ninguno de los marcadores serológicos existentes, tales como CA125, CA72-4, o M-CSF, da individualmente tal desempeño. (Bast, R.C., *et al.*, Int J Biol Marcadores, 1998; 13:179-87).

WO 02/37112 describe un procedimiento para el diagnóstico, pronóstico y control del cáncer ovárico en un sujeto mediante la detección de hK10 en una muestra del sujeto, tal como una muestra de suero o un extracto de tejido tumoral. hK10 se puede medir usando un reactivo que detecta o se une a hK10, tal como los anticuerpos específicamente reactivos con hK10 o con una parte de ellos.

Bergman, Ann-Charlotte *et al.*, Identification of gel-separated tumor marker proteins by mass spectrometry Electrophoresis 2000, 21, 679-686 describe que la electroforesis en gel en dos dimensiones con posterior análisis por espectrometría de masa se aplicó para estudiar las diferencias de la expresión de proteína entre los tumores sólidos benignos y malignos de células humanas de mamas, pulmón y ovario. Las identificaciones de proteína en los tumores ováricos malignos incluyen proteína de unión al ácido retinoico II así como galectina-1 y apolipoproteína A1 y anexina IV.

Vlahou, Antonia *et al.*, SELDI protein profiling in ovarian cancer diagnosis, Proceedings of the American Association for Cancer Research, vol. 43, March 2002, 36, sugieren al chip de proteína de SELDI (ionización/desorción láser potenciada por superficie) como una herramienta diagnóstica adyuvante para el cáncer ovárico.

En consecuencia, existe una necesidad crítica de nuevos marcadores que, en forma individual o en combinación con otros marcadores o modalidades diagnósticas, proporcionen la sensibilidad y especificidad requeridas para la detección precoz del cáncer ovárico. (Bast RC *et al.*, Early detection of ovarian cancer: promise and reality. Ovarian cancer: ISIS Medical Media Ltd., Oxford, UK; 2001. en prensa). Sin una prueba de detección aceptable, la detección precoz aún es el factor más crítico para aumentar la supervivencia a largo plazo de los pacientes con cáncer ovárico.

Por consiguiente, es deseable contar con un procedimiento confiable y preciso para determinar el estado del cáncer ovárico en los pacientes, cuyos resultados se pueden usar posteriormente para manejar el tratamiento del sujeto.

Sumario de la invención

La presente invención proporciona procedimientos sensibles y rápidos y kits que son útiles para determinar el estado del cáncer ovárico mediante la medición de estos marcadores. La medición de estos marcadores en las muestras de pacientes proporciona información que los diagnosticadores pueden correlacionar con un probable diagnóstico de cáncer humano o con un diagnóstico negativo (por ejemplo, normal o libre de enfermedad). Los marcadores se caracterizan por peso molecular y/o por sus propias identidades de proteína. Los marcadores se pueden separar de otras proteínas en una muestra por medio de una variedad de técnicas de fraccionamiento, por ejemplo, separación cromatográfica acoplada con espectrometría de masa, captura de proteína usando anticuerpos inmovilizados o por inmunoensayos tradicionales. En realizaciones preferidas, el procedimiento de resolución involucra espectrometría de masa ionización/desorción láser potenciada por superficie ("SELDI"), en el que la superficie de la sonda de espectrometría de masa comprende adsorbentes que se unen con los marcadores.

Más específicamente, se descubrieron tres biomarcadores que luego se identificaron, de acuerdo con los procedimientos descritos en la presente como (1) apolipoproteína A1 (denominada en la presente como "Apo A1"), (2) una forma truncada de transtiretina, (denominada en la presente como "transtiretina Δ N10"), y (3) un fragmento de escisión del inhibidor de la cadena pesada de inter- α -tripsina H4 (denominada en la presente como "fragmento IAIH4").

La presente invención proporciona un procedimiento para calificar el estado del cáncer ovárico de un sujeto que comprende (a) medir, en una muestra de suero del sujeto, la cantidad de Apo A1, transtiretina Δ N10 y fragmento IAIH4 y (b) correlacionar la medición con el estado del cáncer ovárico de acuerdo con la reivindicación 1.

En ciertos procedimientos, la etapa de medición comprende detectar la presencia o ausencia de marcadores en la muestra. En otros procedimientos, la etapa de medición comprende cuantificar la cantidad de marcadores en la muestra.

En el procedimiento de la invención, la etapa de medición puede comprender: proporcionar una muestra de sangre o un derivado de sangre; fraccionar las proteínas en la muestra en una resina de intercambio iónico y recolectar fracciones que contienen Apo A1, transtiretina Δ N10 y fragmento IAIH4; y capturar Apo A1, transtiretina Δ N10 y fragmento IAIH4 de las fracciones en una superficie de un sustrato que comprende reactivo de capturas que unen los biomarcadores proteicos. El derivado de sangre, por ejemplo, es suero o plasma. En realizaciones preferidas, el sustrato es una sonda de SELDI que comprende una superficie de cobre IMAC y en el que los biomarcadores proteicos se detectan por SELDI. En otras realizaciones, el sustrato es una sonda de SELDI que comprende reactivos de afinidad bioespecíficos que unen a Apo A1, transtiretina Δ N10 y fragmento IAIH4 y en el que los biomarcadores proteicos se detectan por SELDI. En otras realizaciones, el sustrato es una placa de microtitulación que comprende reactivos de afinidad bioespecíficos que unen Apo A1, transtiretina Δ N10 y fragmento IAIH4 y los biomarcadores proteicos se detectan por inmunoensayo:

Los procedimientos se pueden usar para manejar el tratamiento del sujeto sobre la base del estado determinado por el procedimiento. Por ejemplo, si el resultado de los procedimientos de la presente invención es no concluyente o existe una razón por la cual sea necesaria la confirmación del estado, el médico puede ordenar más pruebas. Alternativamente, si el estado indica que la cirugía es apropiada, el médico puede programar la cirugía para la paciente. Asimismo, si el resultado de la prueba es positivo, por ejemplo, el estado es cáncer ovárico de etapa tardía o si el estado es de otro modo agudo, no se justificaría ninguna acción adicional. Por otra parte, si los resultados muestran que el tratamiento ha sido exitoso, puede no ser necesario un tratamiento adicional.

Por otra parte, al menos un biomarcador se puede medir nuevamente después del tratamiento del sujeto. En estos casos, la etapa de manejo del tratamiento del sujeto luego se repite y/o altera según el resultado obtenido.

El término “estado del cáncer ovárico” se refiere al estado de la enfermedad de la paciente. Los ejemplos de tipos de estados del cáncer ovárico incluyen, pero sin limitación, el riesgo de cáncer del sujeto, la presencia o ausencia de la enfermedad, el estado de enfermedad de una paciente y la efectividad del tratamiento de la enfermedad. Otros estados y grados de cada estado son conocidos en la técnica.

Los biomarcadores que son útiles en los procedimientos de la presente invención son Apo A1, transtiretina Δ N10 y fragmento IAIH4. En ciertas realizaciones preferidas, el procedimiento además comprende medir al menos un marcador previamente conocido (en la presente denominado como “Marcador 4”) en una muestra del sujeto y correlacionar la medición de al menos un Marcador 4 y la medición de Apo A1, transtiretina Δ N10 y fragmento IAIH4 con el estado del cáncer ovárico. En ciertas realizaciones se mide solo un Marcador 4, además de los marcadores Apo A1, transtiretina Δ N10 y fragmento IAIH4, mientras que en otras realizaciones se mide más de Marcador 4.

El Marcador 4 incluye biomarcadores del cáncer ovárico conocidos y se selecciona de CA125, CA125 II, CA15-3, CA19-9, CA72-4, CA 195, inhibidor de tripsina asociado al tumor (TATI), CEA, fosfatasa alcalina placentaria (PLAP), Sialil TN, galactosiltransferasa, factor estimulante de colonia de macrófago (M-CSF, CSF-1), ácido lisofosfatídico (LPA), componente de 110 kD del dominio extracelular del receptor del factor de crecimiento epidérmico (p110EGFR), calicreínas de tejido, por ejemplo, calicreína 6 y calicreína 10 (NES-1), prostasina, HE4, creatina quinasa B (CKB), LASA, HER-2/neu, péptido de gonadotropina urinaria, Dianon NB 70/K, antígeno del péptido tisular (TPA), osteopontina y haptoglobina, y variantes de proteína (por ejemplo, forma de escisión, isoformas) de los marcadores.

De acuerdo con la reivindicación 1, el procedimiento proporciona la medición de los tres biomarcadores: Apo A1, transtiretina Δ N10 y fragmento IAIH4. En algunas realizaciones, también se mide al menos un marcador conocido, Marcador 4, en una muestra del sujeto y la medición del Marcador 4 y las mediciones de los otros tres biomarcadores (Apo A1, transtiretina Δ N10 y fragmento IAIH4) se correlacionan con el estado del cáncer ovárico. Como se mencionó antes, en ciertas realizaciones, los biomarcadores que se miden comprenden: los tres biomarcadores (Apo A1, transtiretina Δ N10 y fragmento IAIH4) y dos o más marcadores del grupo designado como Marcador 4.

Para los valores de masa de los marcadores descritos en la presente, la precisión de masa del instrumento espectral se considera que está aproximadamente dentro de $\pm 0,15$ por ciento del valor de peso molecular descrito. En forma adicional, para tales variaciones de precisión reconocidas del instrumento, la determinación del espectro de masa puede variar dentro de los límites de resolución de aproximadamente 400 a 1000 m/dm, donde m es masa y dm es el ancho del pico de espectro de masa a 0,5 de altura de pico.

También se considera que dicha precisión de masa y varianzas de resolución y en consecuencia el significado del término “aproximadamente” con respecto a la masa de cada uno de los marcadores descritos en la presente es inclusivo de las variantes de los marcadores que pueden existir debido al sexo, genotipo y/o etnicidad del sujeto y el cáncer u origen particular o estado del mismo.

La precisión de una prueba diagnóstica se caracteriza por una curva de Características Operador-Receptor (“curva ROC”). Una ROC es un gráfico de la tasa de positivos verdaderos contra la tasa de falsos positivos para los diferentes puntos límites posibles de una prueba diagnóstica. Una curva ROC muestra la relación entre sensibilidad y especificidad. Es decir, un incremento de sensibilidad se acompañará con una reducción de la especificidad. Cuanto más cerca sigue la curva el eje izquierdo y posteriormente el borde superior del espacio ROC, más precisa es la prueba. A la inversa, cuanto más cerca la curva va de la diagonal de 45 del gráfico de ROC, menos precisa es la prueba. El área bajo la ROC es una medida de la precisión de la prueba. La precisión de la prueba depende de cuán bien la prueba separa el grupo analizado en aquellos con y sin la enfermedad en cuestión. Un área bajo la curva (denominada como “AUC”) de 1 representa una prueba perfecta, en cambio un área de 0,5 representa una prueba menos útil. En consecuencia, los biomarcadores preferidos y los procedimientos de diagnóstico de la presente invención tienen una AUC mayor de 0,50, las pruebas más preferidas tienen una AUC mayor de 0,60, las pruebas más preferidas tienen una AUC mayor de 0,70.

Los procedimientos preferidos para medir los biomarcadores incluyen usar una matriz de biochip. Las matrices de biochips útiles en la invención incluyen matrices de proteína y ácido nucleico. Uno o más marcadores se capturan en la matriz de biochip y se someten a ionización láser para detectar el peso molecular de los marcadores. El análisis de los marcadores es, por ejemplo, por peso molecular de uno o más marcadores contra un umbral de intensidad que se normaliza contra una corriente iónica total. Preferiblemente, se usa la transformación logarítmica para reducir los intervalos de intensidad del pico para limitar la cantidad de marcadores detectados.

En los procedimientos preferidos de la presente invención, la etapa para correlacionar la medición de los biomarcadores con el estado del cáncer ovárico se realiza con un algoritmo de clasificación del programa de computación. Preferiblemente, los datos se generan en muestras de sujetos inmovilizados sobre una matriz de biochip, al someter dicha matriz de biochip a la ionización láser y detectar la intensidad de señal para la relación masa/carga; y, transformar los datos en forma legible por ordenador; y ejecutar un algoritmo que clasifica los datos de acuerdo con los parámetros de ingreso del usuario, para detectar señales que representan marcadores presentes en pacientes de cáncer ovárico y que faltan en los sujetos controles sin cáncer.

Preferiblemente las superficies del biochip son, por ejemplo, iónicas, aniónicas, compuestas de iones níquel inmovilizado, compuestas de una mezcla de iones positivos y negativos, compuestas de uno o más anticuerpos, ácidos

nucleicos de cadena simple o doble, proteínas, péptidos o fragmentos de estos, sondas de aminoácidos, o bibliotecas de despliegue de fagos.

5 Preferiblemente, se miden uno o más de los marcadores usando espectrometría de masa por desorción/ionización láser, que comprende proveer una sonda adaptada para usar con un espectrómetro de masa que comprende adsorbente unido a este, y poner en contacto la muestra del sujeto con el adsorbente, y; desorber y ionizar el marcador o los marcadores de la sonda y detectar los marcadores desionizados/ionizados con el espectrómetro de masa.

10 Preferiblemente, la espectrometría de masa por desorción/ionización láser comprende: proporcionar un sustrato que comprende un adsorbente unido a este; poner en contacto la muestra del sujeto con el adsorbente; colocar el sustrato en una sonda adaptada para usar con un espectrómetro de masas que comprende un adsorbente unido a este; y, desorber y ionizar el marcador y los marcadores de la sonda y detectar el marcador o los marcadores desorbidos/ionizados con el espectrómetro de masas.

15 El adsorbente puede ser un adsorbente hidrófobo, hidrófilo, iónico o un quelato metálico, tal como níquel o un anticuerpo, oligonucleótido de cadena simple o doble, aminoácido, proteína, péptidos o sus fragmentos.

20 Los procedimientos descritos se pueden realizar en cualquier tipo de muestra de la paciente que pueda estar disponible para dichos procedimientos, por ejemplo, sangre, suero y plasma.

Tal como se reivindica, se mide una pluralidad de biomarcadores en una muestra del sujeto.

25 La pluralidad de los biomarcadores consiste en Apo A1, transtiretina Δ N10 y fragmento IAIH4. La medición de la pluralidad de los biomarcadores también puede incluir medir al menos un Marcador 4, es decir, al menos un biomarcador conocido que se describió anteriormente. Preferiblemente, los biomarcadores proteicos se miden por SELDI o inmunoensayo. La presente invención también proporciona un kit como el reivindicado, es decir, que comprende (a) un reactivo de captura o reactivo de capturas que se unen a Apo A1, transtiretina Δ N10, y fragmento IAIH4; y (b) un recipiente que comprende Apo A1, transtiretina Δ N10 y fragmento IAIH4. Si bien el reactivo de captura puede ser cualquier tipo de reactivo, preferiblemente el reactivo es una sonda de SELDI. El reactivo de captura también se 30 puede unir a otros biomarcadores conocidos, por ejemplo, Marcador 4. El kit también puede comprender un reactivo de captura que se une a CA125.

35 Los kits pueden comprender una solución de lavado que permite la retención en forma selectiva del biomarcador unido al reactivo de captura en comparación con otros biomarcadores después del lavado.

El kit también proporciona un reactivo de captura, que es un anticuerpo, oligonucleótido de cadena simple o doble, aminoácido, proteína, péptido o fragmento de este.

40 La medición de uno o más biomarcadores proteicos usando el kit se realiza mediante espectrometría de masas o inmunoensayos tales como un ELISA.

También se incluyen proteínas purificadas para la detección del cáncer ovárico y/o la generación de los anticuerpos para posteriores ensayos de diagnóstico. Las proteínas purificadas incluyen un péptido purificado de la SEQ ID NO: 1 (fragmento IAIH4). Este péptido purificado también puede comprender una marca detectable.

45 Otros aspectos de la invención se describen *infra*.

Breve descripción de las figuras

50 La Figura 1 muestra una vista del espectro de masas en pseudo-gel de muestras del conjunto de descubrimiento de biomarcadores, que muestra picos ubicados a valores de m/z de 12828 y 28043 (fracción de pH 4, matriz de IMAC-Cu), y a 3272 (fracción pH 9, matriz de IMAC-Cu).

55 Las Figuras 2(A), (B), (C) y (D) muestran una comparación de las curvas de características operador-receptor (ROC) entre CA125 y tres biomarcadores identificados.

Las Figuras 2 (E), (F), (G), y (H) muestran una comparación de las curvas de características operador-receptor (ROC) para CA125 y dos modelos predictivos multivariados.

60 Las Figuras 3 (a), (b), (c), (d), (e), (f), (g), y (h) muestran los gráficos de dispersión que muestran las distribuciones de los tres biomarcadores identificados y CA125 entre los pacientes y controles sanos en el conjunto de descubrimiento del biomarcador y el conjunto de validación independiente (paneles a - h).

65 Las Figuras 3 (i), (j), (k) y (l) muestran los gráficos de dispersión que muestran el resultado de los dos modelos predictivos multivariados entre los pacientes y controles sanos en el conjunto de ensayo (parte del conjunto de descubrimiento del biomarcador) y el conjunto de validación independiente. La Figura 4 muestra el diagrama del algoritmo de clasificación usado para caracterizar los biomarcadores.

Definiciones

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en la presente tienen el significado comúnmente entendido por los expertos en la técnica a la que pertenece esta invención. Las siguientes referencias proporcionan a los expertos una definición general de muchos de los términos usados en esta invención: Singleton *et al.*, Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics 5th Ed., R. Rieger *et al.* (eds.), Springer Verlag (1991); y Hale & Marham, The Harper Collins Dictionary of Biology (1991). Como se usan en la presente, los siguientes términos tienen los significados atribuidos a ellos a menos que se especifique lo contrario.

“Espectrómetro de iones en fase gaseosa” se refiere a un aparato que detecta iones en fase gaseosa. Los espectrómetros de iones en fase gaseosa incluyen una fuente iónica que suministra iones en fase gaseosa. Los espectrómetros de iones en fase gaseosa incluyen, por ejemplo, espectrómetros de masas, espectrómetros de movilidad de iones, y dispositivos de medición de corriente de iones totales. La “espectrometría de iones en fase gaseosa” se refiere al uso de un espectrómetro de iones en fase gaseosa para detectar iones en fase gaseosa.

“Espectrómetro de masas” se refiere al espectrómetro de iones en fase gaseosa que mide un parámetro que se puede traducir en relaciones de masa a carga de iones en fase gaseosa. Los espectrómetros de masas generalmente incluyen una fuente de iones y un analizador de masa. Los ejemplos de espectrómetros de masas son tiempo de vuelo, sector magnético, filtro cuadrupolar, trampa de iones, resonancia de ciclotrón iónico, analizador del sector electrostático e híbridos de estos. “Espectrometría de masas” se refiere al uso de un espectrómetro de masas para detectar iones en fase gaseosa.

“Espectrómetro de masas por desorción láser” se refiere a un espectrómetro de masas que usa energía láser como medio para desorber, volatilizar e ionizar un analito.

“Espectrómetro de masas en tándem” se refiere a cualquier espectrómetro de masas que es capaz de realizar dos etapas sucesivas de discriminación o medición de iones basadas en m/z , que incluye iones en una mezcla iónica. La frase incluye espectrómetros de masas que tienen dos analizadores de masa que son capaces de realizar dos etapas sucesivas de discriminación o medición de iones basadas en m/z , en tándem en espacio. La frase además incluye los espectrómetros de masas que tienen un solo analizador de masa que es capaz de realizar dos etapas sucesivas de discriminación o medición de iones basadas en m/z en tándem en tiempo. En consecuencia, la frase incluye explícitamente espectrómetro de masas Qq-TOF, espectrómetros de masas con trampa de iones, espectrómetros de masas TOF-TOF, transformada de Fourier de resonancia ciclotrón de iones del espectrómetro de masas, espectrómetros de masas sensor electrostático - sector magnético y sus combinaciones.

“Analizador de masa” se refiere a un subensamblaje de un espectrómetro de masas que comprende un medio para medir un parámetro que se puede traducir en relaciones de masa a carga de los iones en fase gaseosa. En un espectrómetro de masas con tiempo de vuelo, el analizador de masa comprende un ensamblaje óptico iónico, un tubo de vuelo y un detector de iones.

“Fuente de iones” se refiere a un subensamblaje de un espectrómetro de iones en fase gaseosa que proporciona iones en fase gaseosa. En una realización, la fuente de iones proporciona iones a través de un proceso de desorción/ionización. Dichas realizaciones generalmente comprenden una interfaz de sonda que capta en posición una sonda en una relación interrogable con una fuente de energía de ionización (por ejemplo, una fuente de desorción/ionización láser) y en comunicación concurrente a presión atmosférica o subatmosférica con un detector de un espectrómetro de iones en fase gaseosa.

Las formas de energía de ionización para desorber/ionizar un analito de una fase sólida incluyen, por ejemplo: (1) energía láser; (2) átomos rápidos (usados en bombardeo de átomos rápidos); (3) partículas de alta energía generadas por medio del decaimiento de radionúclidos (usados en desorción de plasma); y (4) iones primarios que generan iones secundarios (usados en espectrometría de masas de ion secundario). La forma preferida de la energía de ionización para los analitos de fase sólida es un láser (usado en desorción/ionización láser), en particular, láseres de nitrógeno, láseres de Nd-Yag y otras fuentes de láser pulsado. “Fluencia” se refiere a la energía liberada por unidad de área de la imagen interrogada. Una fuente de alta fluencia, tal como un láser, liberará aproximadamente 1 mJ/mm^2 a 50 mJ/mm^2 . Normalmente, una muestra se coloca en la superficie de una sonda, la sonda se capta con la interfaz de la sonda y la superficie de la sonda es pulsada con la energía de ionización. La energía desorbe las moléculas de analito desde la superficie a la fase gaseosa y las ioniza.

Otras formas de energía de ionización para los analitos incluye, por ejemplo: (1) electrones que ionizan los compuestos neutros en fase gaseosa; (2) campo eléctrico fuerte para inducir la ionización de los compuestos neutros de fase gaseosa, fase sólida o fase líquida; y (3) una fuente que aplica una combinación de partículas de ionización o campos eléctricos con agentes químicos neutros para la ionización química de compuestos neutros de fase gaseosa, fase sólida o fase líquida.

“Soporte sólido” se refiere a un material sólido que se puede derivar de o de otro modo unirse a un reactivo de captura. Los ejemplos de soportes sólidos incluyen sondas, placas de microtitulación y resinas cromatográficas.

“Sonda” en el contexto de esta invención se refiere a un dispositivo adaptado para captar una interfaz de la sonda de un espectrómetro de iones en fase gaseosa (por ejemplo, un espectrómetro de masas) y para presentar un analito a la energía de ionización para la ionización y la introducción en un espectrómetro de iones en fase gaseosa, tal como un espectrómetro de masas. Una “sonda” generalmente comprenderá un sustrato sólido (sea flexible o rígido) que comprende una muestra que presenta una superficie en la que se presenta un analito a la fuente de energía de ionización.

“Ionización/desorción láser potenciada por superficie” o “SELDI” se refiere a un procedimiento de espectrometría de iones en fase gaseosa por desorción/ionización (por ejemplo, espectrometría de masas) en el que el analito es capturado en la superficie de una sonda de SELDI que capta la interfaz de la sonda del espectrómetro de iones en fase gaseosa. En “EM por SELDI”, el espectrómetro de iones en fase gaseosa es un espectrómetro de masas. La tecnología SELDI se describe, por ejemplo, en la patente U.S. 5.719.060 (Hutchens y Yip) y la patente U.S. 6.225.047 (Hutchens y Yip).

“Captura de Afinidad Potenciada por Superficie” o “SEAC” es una versión de SELDI que involucra el uso de sondas que comprenden una superficie absorbente (una “sonda SEAC”). “Superficie adsorbente” se refiere a una superficie que se une a un adsorbente (también llamado un “reactivo de captura” o un “reactivo de afinidad”). Un adsorbente es cualquier material capaz de unir un analito (por ejemplo, un polipéptido blanco o ácido nucleico), “adsorbente cromatográfico” se refiere a un material normalmente usado en cromatografía. Los adsorbentes cromatográficos incluyen, por ejemplo, materiales de intercambio iónico, quelantes de metales (por ejemplo, ácido nitriloacético o ácido iminodiacético), quelatos metálicos inmovilizados, adsorbentes de interacción hidrófoba, adsorbentes de interacción hidrófila, colorantes, biomoléculas simples (por ejemplo, nucleótidos, aminoácidos, azúcares simples y ácidos grasos) y adsorbentes de modo mixto (por ejemplo, adsorbentes por atracción hidrófoba/repulsión electrostática). “Adsorbente bioespecífico” se refiere a un adsorbente que comprende una biomolécula, por ejemplo, una molécula de ácido nucleico (por ejemplo, un aptámero), un polipéptido, a polisacárido, un lípido, un esteroide o un conjugado de estos (por ejemplo, una glicoproteína, una lipoproteína, un glicolípido, un ácido nucleico (por ejemplo, conjugado de ADN-proteína). En ciertos casos el adsorbente bioespecífico puede ser una estructura macromolecular tal como un complejo de multiproteínas, una membrana biológica o un virus. Los ejemplos de adsorbentes bioespecíficos son anticuerpos, proteínas del receptor y ácidos nucleicos. Los adsorbentes bioespecíficos normalmente tienen mayor especificidad para un analito blanco que los adsorbentes cromatográficos. Otros ejemplos de adsorbentes para usar en SELDI se pueden hallar en la patente U.S. 6.225.047 (Hutchens y Yip, “Use of retentate chromatography to generate difference maps”, 1 de mayo de 2001).

En algunas realizaciones, una sonda de SEAC se proporciona como una superficie preactivada que se puede modificar para proporcionar un adsorbente de elección. Por ejemplo, se proporcionan ciertas sondas con un resto reactivo que es capaz de unir una molécula biológica mediante un enlace covalente. Los restos epóxido y carbodiimidazol son útiles para unir en forma covalente adsorbentes bioespecíficos tales como anticuerpos o receptores celulares.

“Adsorción” se refiere a la unión no covalente detectable de un analito a un adsorbente o reactivo de captura.

“Desorción Neta Potenciada por Superficie” o “SEND” es una versión de SELDI que involucra el uso de sondas que comprende moléculas absorbedoras de energía unidas químicamente a la superficie de la sonda (“sonda SEND”). Las “moléculas absorbedoras de energía” (“EAM”) se refieren a las moléculas que son capaces de absorber energía de una fuente de desorción/ionización láser y a partir de esta contribuyen a la desorción e ionización de las moléculas de analito en contacto con ella. La frase incluye las moléculas usadas en MALDI, con frecuencia denominadas como “matriz”, y explícitamente incluye los derivados de ácido cinámico, ácido sinapínico (“SPA”) ácido ciano-hidroxicinámico (“CHCA”) y ácido dihidroxibenzoico, ácido ferúlico, derivados de hidroxiaacetofenona, así como otros. También incluye las EAM usadas en SELDI. SEND también se describe en la Patente de Estados Unidos 5.719.060 y Solicitud de patente de Estados Unidos 60/408.255, presentada el 4 de septiembre de 2002 (Kitagawa, “Monomers And Polymers Having Energy Absorbing Moieties Of Use In Desorption/Ionization Of Analytes”).

“Fijación y Liberación Fotolábil Potenciada en Superficie” o “SEPAR” es una versión de SELDI que involucra el uso de sondas unidas a la superficie que puede unir covalentemente un analito y luego liberar el analito mediante la ruptura de un enlace fotolábil del resto después de la exposición a la luz, por ejemplo, luz láser. SEPAR también se describe en la Patente de Estados Unidos 5.719.060.

“Eluyente” o “solución de lavado” se refiere a un agente, normalmente una solución, que se usa para afectar o modificar la adsorción de un analito a una superficie adsorbente y/o eliminar los materiales no unidos de la superficie. Las características de elución de un eluyente pueden depender, por ejemplo, del pH, fuerza iónica, hidrofobicidad, grado de caotropismo, potencia detergente y temperatura.

“Analito” se refiere a cualquier componente de una muestra que se quiere detectar. El término se puede referir a un componente solo o a una pluralidad de componentes en la muestra.

La “complejidad” de una muestra adsorbida en una superficie de adsorción de una sonda de captura de afinidad significa la cantidad de especies de proteína diferentes que están adsorbidas.

“Ligando de unión molecular” y “ligandos de unión específica” se refieren a pares de moléculas, normalmente pares de biomoléculas que exhiben unión específica. Los compañeros de unión molecular incluyen, sin limitación, receptor y ligando, anticuerpo y antígeno, biotina y avidina, y biotina y estreptavidina.

5 “Control” se refiere al registro de cambios de un parámetro de variación continua.

“Biochip” se refiere a un sustrato sólido que tiene una superficie generalmente plana a la que se fija un adsorbente. Con frecuencia, la superficie del biochip comprende una pluralidad de ubicaciones localizables, cada una de estas ubicaciones tiene el adsorbente unido allí. Los biochips se pueden adaptar para captar una interfaz de sonda y, en consecuencia, funcionan como sondas.

15 “Biochip de proteína” se refiere a un biochip adaptado para la captura de los polipéptidos. Muchos biochips de proteína se describen en la técnica. Estos incluyen, por ejemplo, biochip de proteína producidos por CIPHERGEN en Biosystems (Fremont, CA), Packard BioScience Company (Meriden CT), Zyomyx (Hayward, CA) y Phyllos (Lexington, MA). Los ejemplos de tales biochip de proteína se describen en las siguientes patentes o solicitudes de patente: Patente de Estados Unidos 6.225.047 (Hutchens y Yip, “Use of retentate chromatography to generate difference maps” 1 de mayo de 2001; publicación internacional WO 99/51773 (Kuimelis y Wagner, “Addressable protein arrays” 14 de octubre de 1999); Patente de Estados Unidos 6.329.209 (Wagner *et al.*, “Arrays of protein-capture agents and methods of use thereof”, 11 de diciembre de 2001) y Publicación internacional WO 00/56934 (Englert *et al.*, “Continuous porous matrix arrays” 28 de septiembre de 2000).

20 Los biochips de proteína producidos por CIPHERGEN Biosystems comprenden superficies que tienen adsorbentes cromatográficos o bioespecíficos unidos a ubicaciones localizables. Las matrices de CIPHERGEN ProteinChip® incluyen NP20, H4, H50, SAX-2, WCX-2, CM-10, IMAC-3, IMAC-30, LSAX-30, LWCX-30, IMAC-40, PS-10, PS-20 y PG-20. Estos biochip de proteína comprenden un sustrato de aluminio en la forma de una tira. La superficie de la tira está recubierta con dióxido de silicio.

En el caso del biochip NP-20, el óxido de silicio actúa como un adsorbente hidrófilo para capturar proteínas hidrófilas.

30 Los biochips H4, H50, SAX-2, WCX-2, CM-10, IMAC-3, IMAC-30, PS-10 y PS-20 también comprenden un polímero reticulado, funcionalizado en la forma de un hidrogel físicamente unido a la superficie del biochip o unido covalentemente a través de silano a la superficie del biochip. El biochip H4 tiene grupos funcionales isopropilo para la unión hidrófoba. El biochip H50 tiene nonilfenoxi-poli(etilenglicol)metacrilato para la unión hidrófoba. El biochip SAX-2 tiene grupos funcionales amonio cuaternario para el intercambio aniónico. Los biochips WCX-2 y CM-10 tienen grupos funcionales carboxilato para el intercambio catiónico. Los biochips MAC-3 y IMAC-30 tienen grupos funcionales de ácido nitriloacético que adsorben iones de metal de transición, tales como Cu++ y Ni++, por quelación. Estos iones metálicos inmovilizados permiten la adsorción de péptidos y proteínas por unión coordinada. El biochip PS-10 tiene grupos funcionales carboimidazol que pueden reaccionar con grupos de proteínas para formar una unión covalente. El biochip PS-20 tiene grupos funcionales epóxido para la unión covalente con las proteínas. Los biochips de la serie PS son útiles para la unión de adsorbentes bioespecíficos, tales como anticuerpos, receptores, lectinas, heparina Proteína A, biotina/estreptavidina y similares, a las superficies del chip donde actúan para capturar específicamente los analitos de una muestra. El biochip PG-20 es un chip PS-20 al que se une la Proteína G. Los biochips LSAX-30 (intercambio aniónico), LWCX-30 (intercambio catiónico) e IMAC- 40 (quelato metálico) tienen perlas de látex funcionalizadas en sus superficies. Tales biochips también se describen en: WO 00/66265 (Rich *et al.*, “Probes for a Gas Phase Ion Spectrometer”, 9 de noviembre de 2000); WO 00/67293 (Beecher *et al.*, “Muestra Holder with Hydrophobic Coating for Gas Phase Mass Spectrometer”, 9 de noviembre de 2000); Solicitud de Patente de Estados Unidos US20030032043A1 (Pohl y Papanu, “Latex Based Adsorbent Chip”, 16 de julio de 2002) y Solicitud de Patente de Estados Unidos 60/350.110 (Um *et al.*, “Hydrophobic Surface Chip”, 8 de noviembre de 2001).

50 Después de la captura sobre un biochip, los analitos se pueden detectar por una variedad de procedimientos de detección seleccionados de, por ejemplo, un procedimiento de espectrometría de iones en fase gaseosa, un procedimiento óptico, un procedimiento electroquímico, microscopía de fuerza atómica y un procedimiento de radiofrecuencia. Los procedimientos de espectrometría de iones en fase gaseosa se describen en la presente. De particular interés es el uso de espectrometría de masas y, en particular, SELDI. Los procedimientos ópticos incluyen, por ejemplo, detección de fluorescencia, luminescencia, quimioluminiscencia, absorbancia, reflectancia, transmitancia, birrefringencia o índice de refracción (por ejemplo, resonancia del plasmón superficial, elipsometría, un procedimiento de espejo resonante, un procedimiento de acoplador direccional de guía de onda o interferometría). Los procedimientos ópticos incluyen microscopía (confocal y no confocal), procedimientos de imágenes y procedimientos sin imagen. Los inmunoensayos en varios formatos (por ejemplo, ELISA) son procedimientos populares para la detección de analitos capturados en una fase sólida. Los procedimientos electroquímicos incluyen procedimientos voltamétricos y amperométricos. Los procedimientos de radiofrecuencia incluyen espectroscopía de resonancia multipolar.

65 “Marcador” en el contexto de la presente invención se refiere a un polipéptido (de un peso molecular aparente particular), que está presente en forma diferencial en una muestra tomada de pacientes que tienen cáncer humano en comparación con una muestra comparable tomada de sujetos control (por ejemplo, una persona con un diagnóstico negativo o indetectable de cáncer, sujeto normal o sano). El término “biomarcador” se usa en forma indistinta con el término “marcador”.

El término “medición” significa los procedimientos que incluyen detectar la presencia o ausencia de marcadores en la muestra, cuantificar la cantidad de marcador en la muestra, y/o calificar el tipo de biomarcador. La medición se puede lograr por procedimientos conocidos en la técnica y los que se describen adicionalmente en la presente, que incluyen pero sin limitación, SELDI e inmunoensayo. Se puede usar cualquiera de los procedimientos adecuados para

- 5 detectar y medir uno o más de los marcadores descritos en la presente. Estos procedimientos incluyen, sin limitación, a la espectrometría de masas (por ejemplo, espectrometría de masa por desorción/ionización láser), fluorescencia (por ejemplo inmunoensayo sándwich), resonancia de plasmones superficiales, elipsometría y microscopia de fuerza atómica.
- 10 La frase “presente en forma diferencial” se refiere a las diferencias en la cantidad y/o la frecuencia de un marcador presente en una muestra tomada de pacientes que tienen cáncer humano en comparación con un sujeto control. Por ejemplo, el fragmento IAIH4 está presente en un nivel elevado en las muestras de pacientes con cáncer ovárico en comparación con las muestras de sujetos control. En contraste, Apo A1 y transtiretina Δ N10 descritas en la presente se encuentran en un nivel reducido en las muestras de pacientes con cáncer ovárico en comparación con muestras de
- 15 sujetos control. Por otra parte, un marcador puede ser un polipéptido, que se detecta a una mayor frecuencia o a una menor frecuencia en las muestras de pacientes con cáncer humano en comparación con las muestras de sujetos control. Un marcador puede estar presente en forma diferencial en términos de cantidad, frecuencia o ambas.

- Un polipéptido está presente en forma diferencial entre dos muestras si la cantidad del polipéptido en una muestra
- 20 tiene una diferencia estadísticamente significativa de la cantidad del polipéptido en la otra muestra. Por ejemplo, un polipéptido está presente en forma diferencial entre las dos muestras si está presente en al menos aproximadamente 120%, al menos aproximadamente 130%, al menos aproximadamente 150%, al menos aproximadamente 180%, al menos aproximadamente 200%, al menos aproximadamente 300%, al menos aproximadamente 500%, al menos aproximadamente 700%, al menos aproximadamente 900%, o al menos aproximadamente 1000% más que lo que está
- 25 presente en la otra muestra, o si es detectable en una muestra y no detectable en la otra.

- En forma alternativa o adicional, un polipéptido está presente en forma diferencial entre dos conjuntos de muestras si la frecuencia de detección del polipéptido en las muestras de pacientes con cáncer ovárico es superior o inferior en términos estadísticamente significativos que en las muestras control. Por ejemplo, un polipéptido está presente en
- 30 forma diferencial entre los dos conjuntos de muestras si se detecta al menos aproximadamente 120%, al menos aproximadamente 130%, al menos aproximadamente 150%, al menos aproximadamente 180%, al menos aproximadamente 200%, al menos aproximadamente 300%, al menos aproximadamente 500%, al menos aproximadamente 700%, al menos aproximadamente 900%, o al menos aproximadamente 1000% más frecuentemente o menos frecuentemente en un conjunto de muestras respecto del otro conjunto de muestras.

- “Diagnóstico” significa identificar la presencia o naturaleza de una condición patológica, es decir, cáncer ovárico. Los procedimientos diagnósticos difieren en cuanto a la sensibilidad y especificidad. La “sensibilidad” de un ensayo diagnóstico es el porcentaje de enfermedad de los individuos con pruebas positivas (por ciento de “positivos verdaderos”). Los individuos enfermos no detectados por el ensayo son “falsos negativos”. Los sujetos que no están enfermos
- 40 y que tienen resultados negativos en el ensayo se denominan “negativos verdaderos”. La “especificidad” de un ensayo diagnóstico es 1 menos la tasa de falsos positivos, donde la tasa de “falsos positivos” se define como la proporción de sujetos con la enfermedad que tienen ensayos positivos. Si bien un procedimiento de diagnóstico particular no puede proporcionar un diagnóstico definitivo de una afección, es suficiente si el procedimiento proporciona una indicación positiva que colabora en el diagnóstico.

- Una “cantidad de ensayo” de un marcador se refiere a una cantidad de un marcador presente en una muestra analizada. Una cantidad de ensayo puede ser en cantidad absoluta (por ejemplo, $\mu\text{g/ml}$) o una cantidad relativa (por ejemplo, intensidad relativa de señales).

- Una “cantidad diagnóstica” de un marcador se refiere a una cantidad de un marcador en una muestra del sujeto que es compatible con un diagnóstico de cáncer ovárico. Una cantidad diagnóstica puede ser en cantidad absoluta (por ejemplo, $\mu\text{g/ml}$) o una cantidad relativa (por ejemplo, intensidad relativa de señales).

- Una “cantidad de control” de un marcador puede ser cualquier cantidad o un intervalo de cantidad, que se compara con una cantidad de ensayo de un marcador. Por ejemplo, una cantidad control de un marcador puede ser la cantidad de un marcador en una persona sin cáncer ovárico. Una cantidad control puede ser en cantidad absoluta (por ejemplo, $\mu\text{g/ml}$) o una cantidad relativa (por ejemplo, intensidad relativa de señales).

- “Anticuerpo” se refiere a un ligando polipeptídico sustancialmente codificado por un gen de inmunoglobulina o
- 60 genes de inmunoglobulina, o fragmentos de esta, que se une y reconoce específicamente un epitopo (por ejemplo, un antígeno). Los genes de inmunoglobulina reconocidos incluyen los genes de la región constante liviana kappa y lambda, los genes de la región constante pesada alfa, gamma, delta, epsilon y mu y la miríada de genes de la región variable de la inmunoglobulina. Los anticuerpos existen, por ejemplo, como inmunoglobulinas intactas o como numerosos fragmentos bien caracterizados producidos por la digestión con varias peptidasas. Estos incluyen, por
- 65 ejemplo, los fragmentos Fab’ y F(ab)’².

El término “anticuerpo” como se usa en la presente, también incluye los fragmentos de anticuerpo producidos por la modificación de los anticuerpos completos o los sintetizados de novo usando metodologías de ADN recombinante.

También incluyen anticuerpos policlonales, anticuerpos monoclonales, anticuerpos quiméricos, anticuerpos humanizados o anticuerpos de cadena simple. La porción “Fc” de un anticuerpo se refiere a la porción de una cadena pesada de la inmunoglobulina que comprende uno o más dominios de la región constante de la cadena pesada, CH1, CH2 y CH3, pero no incluye la región variable de la cadena pesada.

“Tratamiento del sujeto” se refiere a la conducta del clínico o médico luego de la determinación del estado del cáncer ovárico. Por ejemplo, si el resultado de los procedimientos de la presente invención no son concluyentes o existe una razón de que es necesaria la confirmación de este estado, el médico debe ordenar más pruebas.

Alternativamente, si el estado indica que la cirugía es apropiada, el médico puede programar la cirugía de la paciente. Asimismo, si el estado es negativo, por ejemplo, cáncer ovárico en etapa final o si el estado es agudo, no se justifica una acción adicional.

Por otra parte, si los resultados muestran que el tratamiento ha sido exitoso, puede no ser necesario un tratamiento adicional.

Descripción detallada de la invención

La presente invención proporciona un procedimiento y un kit reivindicado sobre la base de los biomarcadores generados de la comparación de perfiles proteicos de pacientes diagnosticados con cáncer ovárico y de pacientes sin enfermedades neoplásicas, usando el ProteinChip® Biomarker System Ciphergen Biosystems, Inc., Fremont, CA). Estos biomarcadores definidos en las reivindicaciones 1 y 15, junto con otros marcadores de cáncer ovárico, se evaluaron en forma individual y en modelos predictivos multivariados. En particular, se muestra que los biomarcadores definidos en las reivindicaciones 1 y 15, opcionalmente en combinación con otras pruebas diagnósticas, proporcionan un nuevo procedimiento para determinar el estado de cáncer ovárico en un sujeto.

El perfil de proteína de alto rendimiento combinado con el uso efectivo de herramientas bioinformáticas proporciona un abordaje útil para analizar los marcadores del cáncer. En breves palabras, el sistema usado en la presente invención utiliza matrices cromatográficas ProteinChip® para ensayar muestras usando SELDI (Desorción/ionización láser potenciada por superficie). Las proteínas unidas a las matrices se leen en un ProteinChip® Reader, un espectrómetro de masas de tiempo de vuelo.

La presente invención se basa en el descubrimiento de los marcadores de proteína que están presentes en forma diferencial en las muestras de pacientes con cáncer ovárico y los sujetos control, y la aplicación de este descubrimiento en procedimientos y kits para determinar el estado del cáncer ovárico. Estos marcadores de proteína se hallan en las muestras de pacientes con cáncer ovárico en niveles que son diferentes de los niveles de las mujeres en las que el cáncer es indetectable. Por consiguiente, la cantidad de uno o más de los marcadores hallados en una muestra de ensayo en comparación con un control, o la presencia o ausencia de uno o más marcadores en la muestra de ensayo proporciona información útil con respecto al estado del cáncer ovárico de la paciente. En el procedimiento de la reivindicación 1, se usan tres biomarcadores en combinación.

I. Descripción de los biomarcadores

A. Apolipoproteína A1

Un marcador que se usa en los procedimientos de la presente invención incluye apolipoproteína A1, también denominada en la presente como “Apo A1”. Apo A1 es detectable por espectrometría de masas como un pico que tiene m/z de 28043. Las masas de los marcadores descritos en la presente se consideran precisas dentro del 0,15 por ciento del valor especificado determinado por el protocolo de espectroscopia de masas SELDI. Apo A1 se detectó por el fraccionamiento de la sangre de acuerdo con el protocolo, seguido por la aplicación a un chip de IMAC y la detección por SELDI. La proteína purificada se digirió con tripsina y se identificó como apolipoproteína A1. El protocolo para aislar e identificar Apo A1 se expone a continuación en los Ejemplos. Apo A1 se regula por disminución en las pacientes que tienen cáncer ovárico en alguna etapa. En consecuencia, la ausencia de Apo A1, o una disminución estadísticamente significativa en la cantidad de Apo A1, en comparación con un control normal, se puede correlacionar con un estado del cáncer ovárico. Una disminución estadísticamente significativa es la conocida en la técnica, por ejemplo, el valor de p menor de 0,05.

B. Transtiretina ΔN10

Otro marcador que se usa en los procedimientos de la presente invención incluye una forma de pre-albúmina, también denominada en la presente como “transtiretina ΔN10”. La transtiretina ΔN10 es detectable por espectrometría de masas como un pico que tiene m/z de 12870,9. La transtiretina ΔN10 se detectó por el fraccionamiento de la sangre de acuerdo con el protocolo, seguido por la aplicación a un chip de IMAC y la detección por SELDI. Por inmunoprecipitación y espectrometría de masas en tándem, se halló que la proteína purificada es una forma truncada de la pre-albúmina, que carece de los 10 aminoácidos N-terminales (denominada en la presente como “transtireti-

na $\Delta N10$ ”). El protocolo para aislar e identificar la transtiretina $\Delta N10$ se expone a continuación en los Ejemplos. La transtiretina $\Delta N10$ también está regulada por disminución en los pacientes que tienen cáncer ovárico en alguna etapa. En consecuencia, la ausencia de transtiretina $\Delta N10$, o una disminución estadísticamente significativa de la cantidad de transtiretina $\Delta N10$, en comparación con un control normal, se puede correlacionar con un estado del cáncer ovárico.

La invención que se describe en la presente usa transtiretina $\Delta N10$. Sin embargo, la transtiretina nativa (13900 dalton) también es útil en los procedimientos de la invención.

C. Fragmento IAIH4

Otro marcador que se usa en los procedimientos de la presente invención es un fragmento de escisión de la cadena pesada H4 del inhibidor de inter- α -tripsina, también denominado en la presente como “fragmento IAIH4”. El fragmento IAIH4 es detectable por espectrometría de masas como un pico que tiene m/z de 3272. El fragmento IAIH4 se detectó por el fraccionamiento de la sangre de acuerdo con el protocolo, seguido por la aplicación a un chip IMAC y la detección por SELDI. El pico se purificó del suero combinado de pacientes con cáncer ovárico mediante una serie de técnicas de separación cromatográficas. Se determinó que su secuencia es MNFRPGVLSSRQLGLPGPPDVPD-HAAYHPF (SEQ ID NO: 1), un fragmento que abarca los aminoácidos 660-689 de la cadena pesada H4 del inhibidor de inter-alfa-tripsina (ITIH4; PK-120). Este resultado se confirmó por el análisis de los productos de digestión de pepsina del marcador. El fragmento IAIH4 está regulado por aumento en los pacientes que tienen cáncer ovárico en alguna etapa. En consecuencia, la presencia del fragmento IAIH4, o un aumento de la cantidad del fragmento IAIH4, en comparación con un control normal, se puede correlacionar con un estado del cáncer ovárico.

D. Marcadores de cáncer ovárico conocidos

Ciertas realizaciones de la presente invención también usan biomarcadores de cáncer ovárico conocidos en combinación con Apo A1, transtiretina $\Delta N10$ y fragmento IAIH4. El término “Marcador 4” se usa en la presente para referirse a los marcadores de cáncer ovárico conocidos. El Marcador 4 incluye

CA125, CA125 II, CA15-3, CA19-9, CA72-4, CA 195, TATI, CEA, PLAP, Sialil TN, galactosiltransferasa, M-CSF, CSF-1, LPA, p110EGFR, calicreínas tisulares, prostasina, HE4, CKB, LASA, HER-2/neu, péptido de gonadotropina urinaria, Dianon NB 70/K, TPA, osteopontina y haptoglobina, y variantes de proteína (por ejemplo, formas de escisión, isoformas) de los marcadores.

Estos marcadores son útiles para diagnosticar cáncer ovárico sobre la base de sus niveles en sangre, en comparación con los sujetos normales. Por ejemplo, se sabe que CA125 está elevado en la sangre de mujeres con cáncer ovárico. De modo similar, se sabe que CA 19-9, CA 72.4, CA 195, TATI, inhibina y PLAP, y otros, están elevados en la sangre de mujeres con cáncer ovárico. En ciertas realizaciones preferidas de esta invención, al menos un marcador conocido (Marcador 4) se incluye en el procedimiento con los marcadores Apo A1, transtiretina $\Delta N10$ y fragmento IAIH4.

II. Muestras de ensayo

A) Tipos de sujetos

Las muestras se recolectan de los sujetos, por ejemplo, mujeres, que quieren establecer el estado del cáncer ovárico. Los sujetos pueden ser mujeres en las que se ha determinado que tienen un alto riesgo de cáncer ovárico sobre la base de sus antecedentes familiares. Otras pacientes incluyen mujeres que tienen cáncer ovárico y la prueba se usa para determinar la efectividad de la terapia o el tratamiento que ellas están recibiendo. También, las pacientes podrían incluir mujeres sanas que tienen una prueba como parte de un examen de rutina, o para establecer los niveles basales de los biomarcadores. Las muestras se pueden recoger de mujeres que han sido diagnosticadas con cáncer ovárico y que recibieron tratamiento para eliminar el cáncer o que tal vez están en remisión.

B) Tipos de muestra y preparación de la muestra

Los marcadores se pueden medir en diferentes tipos de muestras biológicas. La muestra es preferiblemente una muestra biológica líquida. Los ejemplos de una muestra biológica líquida útil en esta invención incluyen sangre, suero sanguíneo, plasma, secreciones vaginales, orina, lágrimas, saliva, etc. Debido a que todos los marcadores se hallan en suero sanguíneo, el suero sanguíneo es la fuente de muestra para las realizaciones de la invención.

Si se desea, la muestra se puede preparar para aumentar la capacidad de detección de los marcadores. Por ejemplo, para incrementar la capacidad de detección de los marcadores, una muestra de suero sanguíneo del sujeto se puede fraccionar preferiblemente mediante, por ejemplo, cromatografía en agarosa con azul de Cibacron y cromatografía de afinidad de ADN monocatenario, cromatografía de intercambio aniónico, cromatografía de afinidad (por ejemplo, con anticuerpos) y similares. El procedimiento de fraccionamiento depende del tipo de procedimiento de detección

usado. Se puede usar cualquier procedimiento que enriquece la proteína de interés. Las preparaciones de muestra, tales como los protocolos de pre-fraccionamiento, son opcionales y pueden no ser necesarios para aumentar la capacidad de detección de los marcadores de acuerdo con los procedimientos de detección utilizados. Por ejemplo, la preparación de la muestra puede no ser necesaria si los anticuerpos que unen específicamente los marcadores se usan para detectar la presencia de los marcadores en una muestra.

Normalmente, la preparación de la muestra involucra el fraccionamiento de la muestra y la recolección de fracciones determinadas para obtener los biomarcadores. Los procedimientos de pre-fraccionamiento incluyen, por ejemplo, cromatografía por exclusión de tamaño, cromatografía de intercambio iónico, cromatografía con heparina, cromatografía de afinidad, extracción secuencial, electroforesis en gel y cromatografía líquida. Los analitos también se pueden modificar antes de la detección. Estos procedimientos son útiles para simplificar la muestra para el análisis posterior. Por ejemplo, puede ser útil para eliminar proteínas presentes en gran abundancia, tales como la albúmina, de la sangre antes del análisis. Los ejemplos de procedimientos de fraccionamiento se describen en PCT/US03/00531.

Preferiblemente, la muestra se pre-fracciona por cromatografía de intercambio aniónico. La cromatografía de intercambio aniónico permite el pre-fraccionamiento de las proteínas en una muestra aproximadamente de acuerdo con sus características de carga. Por ejemplo, se puede usar una resina de intercambio aniónico Q (por ejemplo, Q Hyped F, Biosepra), y una muestra se puede eluir en forma secuencial con eluyentes que tienen pH diferentes. La cromatografía de intercambio aniónico permite la separación de biomoléculas en una muestra que tienen más carga negativa que otros tipos de biomoléculas. Las proteínas que se eluyen con un eluyente que tiene un pH alto probablemente tienen carga negativa débil, y una fracción que eluye con un eluyente que tiene un pH bajo probablemente tiene carga negativa fuerte. En consecuencia, además de reducir la complejidad de una muestra, la cromatografía de intercambio aniónico separa proteínas de acuerdo con sus características de unión.

En realizaciones preferidas, las muestras de suero se fraccionan por cromatografía de intercambio aniónico. La supresión de la señal de las proteínas de menor abundancia por las proteínas de mayor abundancia representa un desafío significativo para la espectrometría de masas SELDI. El fraccionamiento de una muestra reduce la complejidad de los constituyentes de cada fracción. Este procedimiento también se puede usar para intentar aislar proteínas presentes en gran abundancia en una fracción, y de este modo reduce su efecto de supresión de señal sobre las proteínas de menor abundancia. El fraccionamiento de intercambio aniónico separa las proteínas por su punto isoelectrico (pI). Las proteínas están compuestas de aminoácidos, que son ambivalentes, - sus cargas cambian sobre la base del pH del ambiente al que se exponen. El pI de la proteína es el pH al que la proteína no tiene carga neta. Se supone que una proteína tiene una carga neutra cuando el pH del ambiente es equivalente al pI de la proteína. Cuando el pH aumenta por encima del pI de la proteína, la proteína asume una carga neta negativa. De modo similar, cuando el pH del ambiente cae por debajo del pI de la proteína, la proteína tiene una carga neta positiva. Las muestras de suero se fraccionaron de acuerdo con el protocolo que se expone en los siguientes Ejemplos para obtener los marcadores descritos en la presente.

Después de la captura por intercambio aniónico, las proteínas se eluyeron en una serie de etapas de lavados a pH 9, pH 7, pH 5, pH 4 y pH 3. Se descubrió una panel de tres potenciales biomarcadores por análisis UMSA de datos de perfiles de tres fracciones (pH 9/flujo continuo, pH 4, y solvente orgánico). Dos de los picos eran de la fracción pH 4 a m/z de 12828 y 28043, ambos regulados por disminución en el grupo de cáncer, y el tercero era de la fracción de pH 9/flujo continuo con m/z de 3272, regulado por aumento en el grupo de cáncer. El total unido a la matriz de cromatografía de afinidad con metal inmovilizado se cargó con iones cobre (IMAC3-Cu) (espectros de la Figura 1).

Las biomoléculas de una muestra también se pueden separar por electroforesis de alta resolución, por ejemplo, electroforesis en gel de una o dos dimensiones. Una fracción que contiene un marcador se puede aislar y analizar posteriormente por espectrometría de iones en fase gaseosa. Preferiblemente, se usa electroforesis en gel de dos dimensiones para generar una matriz de dos dimensiones de las manchas de las biomoléculas que incluyen uno o más marcadores. Ver, por ejemplo, Jungblut y Thiede, Mass Spectr. Rev. 16:145-162 (1997).

La electroforesis en gel de dos dimensiones se puede realizar usando procedimientos conocidos en la técnica. Ver, por ejemplo, Deutscher ed., Methods In Enzymology vol. 182. Normalmente, las biomoléculas de una muestra se separan, por ejemplo, por isoelectroenfoque, durante el cual se separan las biomoléculas de una muestra en un gradiente de pH hasta que alcanzan un punto en el que su carga neta es cero (es decir, punto isoelectrico). Esta primera etapa de separación produce una matriz unidimensional de biomoléculas. Las biomoléculas de la matriz unidimensional se separan posteriormente usando una técnica generalmente diferente de la usada en la primera etapa de separación. Por ejemplo, en la segunda dimensión, las biomoléculas separadas por isoelectroenfoque se separan adicionalmente usando un gel de poliacrilamida, tal como la electroforesis en gel de poliacrilamida en presencia de dodecilsulfato de sodio (SD PAGE). El gel SDS-PAGE permite la separación adicional sobre la base de la masa molecular de las biomoléculas. Normalmente, la electroforesis en gel de dos dimensiones puede separar biomoléculas químicamente diferentes en un intervalo de masa molecular de 1.000-200.000 Da dentro de las mezclas del complejo. El intervalo de pI de estos geles es aproximadamente 3-10 (geles de intervalo amplio).

Las biomoléculas de la matriz de dos dimensiones se pueden detectar usando cualquiera de los procedimientos adecuados conocidos en la técnica. Por ejemplo, las biomoléculas se pueden marcar o teñir en un gel (por ejemplo, azul de Coomassie o tinción con plata). Si la electroforesis en gel genera manchas que corresponden al peso molecular de uno o más marcadores de la invención, la mancha se puede analizar posteriormente por espectrometría de iones en

fase gaseosa. Por ejemplo, las manchas se pueden cortar del gel y analizar por espectrometría de iones en fase gaseosa. Alternativamente, el gel que contiene las biomoléculas se puede transferir a una membrana inerte por la aplicación de un campo eléctrico. Posteriormente una mancha de la membrana que corresponde aproximadamente al peso molecular de un marcador se puede analizar por espectrometría de iones en fase gaseosa. En la espectrometría de iones en fase gaseosa, las manchas se pueden analizar usando cualquier técnica adecuada, tal como MALDI o SELDI (por ejemplo, usando una matriz ProteinChip®) como se describe en la presente.

Antes del análisis de espectrometría de iones en fase gaseosa, puede ser conveniente escindir las biomoléculas de la mancha en fragmentos más pequeños usando reactivos de escisión, tales como proteasas (por ejemplo, tripsina). La digestión de las biomoléculas en fragmentos pequeños proporciona una huella digital de la masa de las biomoléculas en la mancha, que se puede usar para determinar la identidad de los marcadores, según corresponda.

La cromatografía líquida de alta resolución (HPLC) se puede usar para separar una mezcla de biomoléculas en una muestra sobre la base de sus diferentes propiedades físicas, tales como polaridad, carga y tamaño. Los instrumentos de HPLC normalmente consisten en un reservorio de fase móvil, una bomba, un inyector, una columna de separación y un detector. Las biomoléculas de una muestra se separan por la inyección de una alícuota de la muestra en la columna. Las diferentes biomoléculas de la mezcla pasan a través de la columna a velocidades diferentes debido a las diferencias en su comportamiento de partición entre la fase líquida móvil y la fase estacionaria. Se puede recolectar una fracción que corresponde al peso molecular y/o propiedades físicas de uno o más marcadores. La fracción se puede analizar posteriormente por espectrometría de iones en fase gaseosa para detectar los marcadores. Por ejemplo, las manchas se pueden analizar usando MALDI o SELDI (por ejemplo, usando matriz ProteinChip®) como se describe en la presente.

Opcionalmente, un marcador se puede modificar antes del análisis para mejorar su resolución o para determinar su identidad. Por ejemplo, los marcadores se pueden someter a la digestión proteolítica antes del análisis. Se puede usar cualquier proteasa. Las proteasas, tales como tripsina, que probablemente escinden los marcadores en un número discreto de fragmentos son particularmente útiles. Los fragmentos que provienen de la digestión funcionan como una huella digital para los marcadores, de este modo permiten su detección en forma indirecta. Esto es particularmente útil cuando hay marcadores con masas moleculares similares que podrían confundir al marcador en cuestión. Asimismo, la fragmentación proteolítica es útil para los marcadores de alto peso molecular ya que los marcadores más pequeños se resuelven más fácilmente por espectrometría de masas. En otro ejemplo, las biomoléculas se pueden modificar para aumentar la resolución de la detección. Por ejemplo, se puede usar neuraminidasa para eliminar los residuos de ácido siálico terminales de las glicoproteínas para aumentar la unión a un adsorbente aniónico (por ejemplo, matrices de intercambio catiónico ProteinChip®) y para aumentar la resolución de la detección. En otro ejemplo, los marcadores se pueden modificar por la fijación de una marca de peso molecular particular que se une específicamente a los marcadores moleculares, además de distinguirlos. Opcionalmente, después de detectar tales marcadores modificados, se puede determinar adicionalmente la identidad de los marcadores por coincidencia de las características físicas y químicas de los marcadores modificados en una base de datos de proteínas (por ejemplo, SwissProt).

III. Captura de marcadores

Los biomarcadores preferiblemente se capturan con reactivos de captura inmovilizados en un soporte sólido, tal como cualquier biochip descrito en la presente, en una placa de microtitulación multipocillo o mediante una resina. En particular, los biomarcadores de esta invención preferiblemente se capturan en un biochip de proteína para SELDI. La captura puede ser en una superficie cromatográfica o una superficie bioespecífica. Cualquier biochip de proteína para SELDI que comprende superficies reactivas se puede usar para capturar y detectar los biomarcadores usados en esta invención. Sin embargo, los biomarcadores usados en esta invención se unen bien a los quelatos metálicos inmovilizados. Los biochips IMAC-3 e IMAC 30, cuyas funcionalidades de ácido nitriloacético que adsorben los iones de metales de transición tales como Cu⁺⁺ y Ni⁺⁺ por quelación, son los biochips de SELDI preferidos para capturar los biomarcadores. Se puede usar cualquiera de los biochips de proteína de SELDI que comprende superficies reactivas para capturar y detectar los biomarcadores. Estos biochips se pueden derivar con los anticuerpos que capturan específicamente los biomarcadores, o se pueden derivar con reactivos de captura, tales como proteína A o proteína G que se une a las inmunoglobulinas. Posteriormente los biomarcadores se pueden capturar en solución usando anticuerpos específicos y los marcadores capturados aislados sobre el chip mediante el reactivo de captura.

En general, una muestra que contiene los biomarcadores, tal como suero, se coloca en la superficie activa de un biochip durante un tiempo suficiente para permitir la unión. Posteriormente, las moléculas no unidas se lavan de la superficie usando un eluyente adecuado, tal como solución salina tamponada con fosfato. En general, cuanto más riguroso es el eluyente, más estrechamente unidas deben estar las proteínas para ser retenidas después del lavado. Los biomarcadores proteicos retenidos se pueden detectar por medios apropiados.

IV. Detección y medición de marcadores

Una vez capturado un sustrato, por ejemplo, biochip o anticuerpo, se puede usar cualquier procedimiento adecuado para medir un marcador o marcadores en una muestra. Por ejemplo, los marcadores se pueden detectar y/o medir por una variedad de procedimientos de detección que incluyen por ejemplo, procedimientos de espectrometría

de iones en fase gaseosa, procedimientos ópticos, procedimientos electroquímicos, microscopia de fuerza atómica y procedimientos de radiofrecuencia. Usando estos procedimientos, se pueden detectar uno o más marcadores.

5 A) SELDI

Un procedimiento de detección y/o medición preferido de los biomarcadores usa espectrometría de masas y, en particular, “Desorción/ionización láser potenciada por superficie” o “SELDI”. SELDI se refiere a un procedimiento de espectrometría de iones en fase gaseosa con desorción/ionización (por ejemplo, espectrometría de masas) en el que el analito es capturado sobre la superficie de una sonda de SELDI que capta la interfaz de la sonda. En “EM por SELDI” el espectrómetro de iones en fase gaseosa es un espectrómetro de masas. La tecnología SELDI se describe con más detalle antes. La Apo A1, la transtiretina $\Delta N10$ y el fragmento IAIH4 se detectan como picos a m/z de 28043, m/z de aproximadamente 12870.9, y m/z de 3272, respectivamente.

15 B) Inmunoensayo

En otra realización, se puede usar un inmunoensayo para detectar y analizar marcadores en una muestra. Este procedimiento comprende: (a) proporcionar un anticuerpo que se une específicamente a un marcador; (b) poner en contacto una muestra con el anticuerpo; y (c) detectar la presencia de un complejo del anticuerpo unido al marcador de la muestra.

Un inmunoensayo es un ensayo que utiliza un anticuerpo para unir específicamente un antígeno (por ejemplo, un marcador). El inmunoensayo se caracteriza por el uso de propiedades de unión específica de un anticuerpo particular para aislar, localizar selectivamente y/o cuantificar el antígeno. La frase “se une específicamente (o selectivamente)” a un anticuerpo o “es específicamente (o selectivamente) inmunorreactivo con” cuando se refiere a una proteína o péptido, se refiere a una reacción de unión que es determinante de la presencia de la proteína en una población heterogénea de proteínas y otras sustancias biológicas. En consecuencia, en las condiciones diseñadas del inmunoensayo, los anticuerpos especificados se unen a una proteína particular al menos el doble del valor umbral y no se unen sustancialmente en una cantidad significativa a otras proteínas presentes en la muestra. La unión específica a un anticuerpo en tales condiciones puede requerir un anticuerpo que se selecciona por su especificidad por una proteína particular. Por ejemplo, se pueden seleccionar anticuerpos policlonales originados para un marcador de especies específicas tales como rata, ratón o ser humano para obtener solo estos anticuerpos policlonales que son específicamente inmunorreactivos con este marcador y no con otras proteínas, excepto las variantes polimórficas y los alelos del marcador. Esta selección se puede llevar a cabo por la sustracción de los anticuerpos que reaccionan en forma cruzada con las moléculas del marcador de otras especies.

Usando los marcadores purificados o sus secuencias de ácidos nucleicos, se pueden preparar anticuerpos que se unen específicamente a un marcador usando cualquiera de los procedimientos adecuados conocidos en la técnica. Ver, por ejemplo, Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies: A Laboratory Manual* (1988); Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); y Kohler & Milstein, *Nature* 256:495-497 (1975). Tales técnicas incluyen, pero sin limitación, la preparación de anticuerpos por selección de anticuerpos de bibliotecas de anticuerpos recombinantes en fagos o vectores similares, así como una preparación de anticuerpos policlonales y monoclonales por inmunización de conejos o ratones (ver, por ejemplo, Huse *et al.*, *Science* 246: 1275-1281 (1989); Ward *et al.*, *Nature* 341:544-546 (1989)). Normalmente, una reacción específica o selectiva será al menos el doble del valor de la señal o ruido umbral y con mayor frecuencia, más de 10 a 100 veces el valor umbral.

En general, una muestra obtenida de un sujeto se puede poner en contacto con el anticuerpo que se une específicamente al marcador. Opcionalmente, el anticuerpo se puede fijar a un soporte sólido para facilitar el lavado y posterior el aislamiento del complejo, antes de poner en contacto el anticuerpo con una muestra. Los ejemplos de soportes sólidos incluyen vidrio o plástico en la forma de, por ejemplo, una placa de microtitulación, una vara, una perla, o una a microperla. Los anticuerpos también se pueden unir a un sustrato sonda o a una matriz ProteinChip® descrita anteriormente. La muestra es preferiblemente una muestra biológica líquida tomada de un sujeto. Los ejemplos de muestras biológicas líquidas incluyen sangre, suero, plasma, aspirado del pezón, orina, lágrimas, saliva etc. En la invención, la muestra es una muestra de suero. La muestra se puede diluir con un eluyente adecuado antes de poner en contacto la muestra con el anticuerpo.

Después de incubar la muestra con los anticuerpos, la mezcla se lava y se puede detectar el complejo de anticuerpo-marcador formado. Esto se puede obtener por la incubación de la mezcla lavada con un reactivo de detección. Este reactivo de detección puede ser, por ejemplo, un segundo anticuerpo que se marca con una marca detectable. Los ejemplos de marcas detectables incluyen perlas magnéticas (por ejemplo, DYNABEADS™), colorantes fluorescentes, radiomarcas, enzimas (por ejemplo, peroxidasa de rábano picante, fosfatasa alcalina y otras comúnmente usadas en un ELISA), y marcas colorimétricas tales como oro coloidal o vidrio coloreado o perlas de plástico. Alternativamente, se puede detectar el marcador en la muestra usando un ensayo indirecto, en el que, por ejemplo, se usa un segundo anticuerpo marcado para detectar el marcador-anticuerpo específico unido, y/o en un ensayo de competición o inhibición en el que, por ejemplo, un anticuerpo monoclonal que se une a un epítopo distinto del marcador se incubaba simultáneamente con la mezcla.

Los procedimientos para medir la cantidad, o presencia del complejo anticuerpo-marcador incluyen, por ejemplo, detección de fluorescencia, luminiscencia, quimioluminiscencia, absorbancia, reflectancia, transmitancia, birrefringencia o índice de refracción (por ejemplo, resonancia del plasmón superficial, elipsometría, un procedimiento de espejo resonante, un procedimiento de acoplador direccional de guía de onda o interferometría). Los procedimientos ópticos incluyen microscopia (confocal y no confocal), procedimientos de imágenes y procedimientos sin imagen. Los procedimientos electroquímicos incluyen procedimientos voltimétricos y amperométricos. Los procedimientos de radiofrecuencia incluyen espectroscopia de resonancia multipolar. Los procedimientos para realizar estos ensayos son fácilmente conocidos en la técnica. Los ensayos útiles incluyen, por ejemplo, un ensayo inmunoenzimático (EIA) tal como el ensayo inmunoabsorbente ligado a enzima (ELISA), un radioinmunoensayo (RIA), un ensayo de transferencia Western, o un ensayo de transferencia de ranura. Estos procedimientos se describen en, por ejemplo, *Methods in Cell Biology: Antibodies in Cell Biology*, volumen 37 (Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991); y Harlow & Lane, *supra*.

En todos los ensayos, se pueden requerir etapas de incubación y/o lavado después de cada combinación de reactivos. Las etapas de incubación pueden variar de aproximadamente 5 segundos a varias horas, preferiblemente de aproximadamente 5 minutos a aproximadamente 24 horas. Sin embargo, el tiempo de incubación dependerá del formato del ensayo, marcador, volumen de la solución, concentraciones y similares. Usualmente los ensayos se llevarán a cabo a temperatura ambiente, aunque se pueden realizar en un intervalo de temperaturas, tal como 10°C a 40°C.

Los inmunoensayos se pueden usar para determinar la presencia o ausencia de un marcador en una muestra así como la cantidad de un marcador en una muestra. La cantidad de un complejo de anticuerpo-marcador se puede determinar por la comparación con un estándar. Un estándar puede ser, por ejemplo, un compuesto conocido u otra proteína que se sabe está presente en una muestra. Como se indicó antes, la cantidad de ensayo del marcador no se necesita medir en unidades absolutas, siempre que la unidad de medición se pueda comparar con un control.

Los procedimientos para detectar estos marcadores en una muestra tienen muchas aplicaciones. Por ejemplo, uno o más marcadores se pueden medir para colaborar en el diagnóstico o pronóstico del cáncer humano. En otro ejemplo, los procedimientos para la detección de los marcadores se pueden usar para controlar las respuestas al tratamiento del cáncer en un sujeto. En otro ejemplo, los procedimientos para detectar marcadores se pueden usar para analizar e identificar compuestos que modulan la expresión de estos marcadores *in vivo* o *in vitro*. En un ejemplo preferido, los biomarcadores se usan para diferenciar entre las diferentes etapas del avance del tumor, de este modo ayuda a determinar el tratamiento apropiado y el grado de metástasis del tumor.

V. *Análisis de datos*

Cuando se mide la muestra y se generan los datos, por ejemplo, por espectrometría de masas, los datos se analizan posteriormente con un programa de computación. En general, el programa de computación puede comprender un código que convierte la señal del espectrómetro de masas en formas legibles por el ordenador. El programa de computación también puede incluir un código que aplica un algoritmo al análisis de la señal para determinar si la señal representa un “pico” de la señal correspondiente a un marcador de esta invención, u otros marcadores útiles. El programa de computación también puede incluir un código que ejecuta un algoritmo que compara la señal de una muestra de ensayo con una señal característica y típica “normal” y de cáncer humano y determine la cercanía del ajuste entre las dos señales. El programa de computación también puede incluir un código que indica a cuál es la muestra de ensayo más cercana, y de este modo proporciona un diagnóstico probable.

En los procedimientos de la presente invención, se miden múltiples biomarcadores. El uso de múltiples biomarcadores incrementa el valor predictivo de la muestra y proporciona mayor utilidad en diagnóstico, toxicología, estratificación y control de la paciente. El proceso llamado “reconocimiento del patrón” detecta los patrones formados por múltiples biomarcadores que aumenta mucho la sensibilidad y especificidad de la proteómica clínica para la medicina predictiva. Las variaciones sutiles en los datos de las muestras clínicas que se pueden obtener, por ejemplo, usando SELDI indican que ciertos patrones de expresión de proteína pueden predecir fenotipos tales como la presencia o ausencia de una determinada enfermedad, una etapa particular de la progresión del cáncer, o una respuesta positiva o adversa a los tratamientos farmacológicos.

La generación de datos en la espectrometría de masas comienza con la detección de los iones por un detector de iones como se describió anteriormente. Los iones que golpean en el detector generan un potencial eléctrico que se digitaliza con un dispositivo de registro de matriz de tiempo de alta velocidad que captura digitalmente la señal analógica. El sistema ProteinChip® de Ciphergen emplea un convertidor de señal analógica a digital (ADC) para obtener este proceso. El ADC integra el resultado del detector en intervalos de tiempo separados regularmente por intervalos dependientes del tiempo. Los intervalos de tiempo normalmente son de uno a cuatro nanosegundos. Por otra parte, el espectro del tiempo de vuelo finalmente analizado, normalmente no representa la señal de un solo pulso de energía de ionización contra una muestra, sino más bien la suma de señales de numerosos pulsos. Esto reduce el ruido e incrementa el intervalo dinámico. Estos datos de tiempo de vuelo posteriormente se someten al procesamiento de datos. En el programa de computación de ProteinChip® de Ciphergen, el procesamiento de datos normalmente incluye la transformación de TOF a M/Z, sustracción del valor basal, filtración del ruido de alta frecuencia.

La transformación de TOF a M/Z involucra la aplicación de un algoritmo que transforma tiempos de vuelo en relación de masa a carga (M/Z). En esta etapa, las señales se convierten del dominio del tiempo al dominio de la masa. Es decir, cada tiempo de vuelo se convierte en la relación masa a carga, o M/Z. La calibración se puede realizar en forma interna o externa. En la calibración interna, la muestra analizada contiene uno o más analitos de M/Z conocida.

Los picos de la señal de tiempos de vuelo que representan los analitos con masa se asignan a la M/Z conocida. Sobre la base de estas relaciones M/Z asignadas, se calculan los parámetros para una función matemática que convierte los tiempos de vuelo a M/Z. En la calibración externa, una función que convierte tiempos de vuelo a M/Z, tal como la creada por la calibración interna previa, se aplica a un espectro de tiempo de vuelo sin el uso de calibradores internos.

La sustracción del valor basal mejora la cuantificación de los datos al eliminar las compensaciones artificiales y reproducibles del instrumento que perturban el espectro. Esto involucra calcular un espectro basal usando un algoritmo que incorpora parámetros tales como ancho de pico, y posteriormente sustraer el valor basal del espectro de masas.

Las señales de ruido de alta frecuencia se eliminan por la aplicación de una función de atenuación. Una función de atenuación típica aplica una función promedio de movimiento a cada intervalo dependiente del tiempo. En una versión mejorada, el filtro de movimiento promedio es un filtro digital de ancho variable en el que el ancho de banda del filtro varía en función de, por ejemplo, ancho de banda del pico, que generalmente es más ancho con el aumento del tiempo de vuelo. Ver, por ejemplo, WO 00/70648, 23 de noviembre de 2000 (Gavin *et al.*, "Variable Width Digital Filter for Time-of-flight Mass Spectrometry").

El análisis generalmente involucra la identificación de picos en el espectro que representa la señal de un analito. La selección del pico, obviamente, se puede realizar en forma ocular. Sin embargo, el programa de computación está disponible como parte del programa de computación de ProteinChip® de Ciphergen que puede automatizar la detección de los picos. En general, este programa de computación funciona por la identificación de señales que tienen una relación de señal a ruido superior a un umbral seleccionado y la marcación de la masa del pico en el centroide de la señal del pico. En una aplicación útil se comparan muchos espectros para identificar picos idénticos presentes en algún porcentaje seleccionado del espectro de masa. Una versión de este programa de computación agrupa todos los picos que aparecen en los diversos espectros dentro de un intervalo de masa definido, y asigna una masa (M/Z) a todos los picos que están cerca del punto medio del grupo de masa (M/Z).

Los datos del pico de uno o más espectros se pueden someter a un análisis posterior, por ejemplo, creando una planilla de cálculo en la que cada fila representa un espectro de masa particular, cada columna representa un pico del espectro definido por la masa y cada celda incluye la intensidad del pico en este espectro particular. Se pueden aplicar varios procedimientos estadísticos o de reconocimiento de patrón a los datos.

En un ejemplo, el programa de computación Biomarker Patterns™ de Ciphergen se usa para detectar un patrón en el que se generan los espectros. Los datos se clasifican usando un proceso de reconocimiento de patrón que usa un modelo de clasificación. En general, los espectros representarán las muestras de al menos dos grupos diferentes para los que se busca un algoritmo de clasificación. Por ejemplo, los grupos pueden ser patológicos *versus* no patológicos (por ejemplo, cáncer *versus* no cáncer), que responden al fármaco *versus* que no responden al fármaco, respuesta tóxica *versus* respuesta no tóxica, que progresa a estado de enfermedad *versus* que no progresa a estado de enfermedad, condición fenotípica presente *versus* condición fenotípica ausente.

Los espectros que se generan se pueden clasificar usando un proceso de reconocimiento de patrón que usa un modelo de clasificación. En algunas realizaciones, los datos derivados de los espectros (por ejemplo, espectros de masa o espectros de tiempo de vuelo) que se generan usando muestras tales como "muestras conocidas" se pueden usar posteriormente para "entrenar" un modelo de clasificación. Una "muestra conocida" es una muestra que está preclasificada (por ejemplo, cáncer o no cáncer). Los datos derivados de los espectros (por ejemplo, espectros de masa o espectros de tiempo de vuelo) que se generan usando las muestras tales como "muestras conocidas" posteriormente se pueden usar para "entrenar" un modelo de clasificación. Una "muestra conocida" es una muestra que está preclasificada. Los datos que se extraen de los espectros y se usan para formar el modelo de clasificación se pueden denominar "conjuntos de datos de entrenamiento" (del inglés, training data set). Una vez diseñado, el modelo de clasificación puede reconocer patrones de datos derivados de los espectros generados usando muestras desconocidas. El modelo de clasificación posteriormente se puede usar para clasificar las muestras desconocidas en clases. Esto puede ser útil, por ejemplo, para predecir si una muestra biológica en particular está asociada o no a determinada condición biológica (por ejemplo, enfermo *versus* no enfermo).

El conjunto de datos del entrenamiento que se usa para formar el modelo de clasificación puede comprender datos crudos o datos preprocesados. En algunas realizaciones, los datos crudos se pueden obtener directamente de los espectros de tiempo de vuelo o espectros de masa, y después opcionalmente se puede "preprocesar" de cualquiera manera adecuada. Por ejemplo, se pueden seleccionar señales superiores a una relación señal a ruido predeterminada de modo de seleccionar un subconjunto de picos del espectro, más que de seleccionar todos los picos del espectro. En otro ejemplo, se puede usar un número predeterminado de "grupos" de picos en un valor común (por ejemplo, un valor de tiempo de vuelo o valor de relación masa a carga en particular) para seleccionar picos. De modo ilustrativo, si un pico de una relación de masa a carga determinada está en menos de 50% de los espectros de masa de un grupo de espectros de masa, por lo tanto se puede omitir el pico de esta relación masa a carga del conjunto de datos del entrenamiento. Se

pueden usar tales etapas de pre-procesamiento para reducir la cantidad de datos que se usan para diseñar el modelo de clasificación. Los modelos de clasificación se pueden formar usando cualquier procedimiento adecuado de clasificación (o “conocimiento”) estadístico que intenta segregar cuerpos de datos en clases basadas en parámetros objetivos presentes en los datos. Los procedimientos de clasificación pueden ser supervisados o no supervisados. Los ejemplos de procesos de clasificación supervisados y no supervisados se describen en Jain, “Statistical Pattern Recognition: A Review”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, enero de 2000.

En la clasificación supervisada, los ejemplos que contienen datos de entrenamiento de categorías conocidas se presentan a un mecanismo de aprendizaje, que enseña más conjuntos de relaciones que definen cada una de las clases conocidas. Los nuevos datos se pueden aplicar posteriormente al mecanismo de aprendizaje, que luego clasifica los nuevos datos usando las relaciones aprendidas. Los ejemplos de procesos de clasificación supervisada incluyen el proceso de regresión lineal (por ejemplo, regresión lineal múltiple (MLR), regresión de cuadrados mínimos parciales (PLS) y regresión de componentes principales (PCR)), árboles de decisión binarios (por ejemplo, procesos de partición recursiva tales como CART - árboles de clasificación y regresión, redes neurales artificiales tales como redes de retropropagación, análisis discriminante (por ejemplo, clasificador de Bayesian o análisis de Fischer), clasificadores logísticos, y clasificadores del vector soporte (máquinas del vector soporte).

Un procedimiento de clasificación supervisada preferido es un proceso de partición recursiva. Los procesos de partición recursivos usan árboles de partición recursiva para clasificar los espectros derivados de las muestras desconocidas. Otros detalles acerca de los procesos de partición recursivos se proporcionan en U.S. 2002 0138208 A1 (Paulse *et al.*, “Method for analyzing mass spectra”, 26 de septiembre de 2002).

En otras realizaciones, los modelos de clasificación que se crean se pueden formar usando procedimientos de aprendizaje supervisados. La clasificación no supervisada intenta informar clasificaciones basadas en las semejanzas del conjunto de datos de entrenamiento, sin preclasificar los espectros de los cuales se derivó el conjunto de datos de entrenamiento. Los procedimientos de aprendizaje no supervisado incluyen los análisis del grupo. Un análisis del grupo intenta dividir los datos en “agrupamientos” o grupos que idealmente tendrían miembros que son similares entre sí, y muy distintos de los otros agrupamientos. La semejanza se mide posteriormente usando alguna distancia métrica, que mide la distancia entre los ítems de datos, y los ítems de datos junto con los grupos que están más próximos entre sí. Las técnicas de agrupamiento incluyen el algoritmo K media de MacQueen y el algoritmo del mapa autoorganizado de Kohonen.

Los algoritmos de aprendizaje indicados para usar en la clasificación de información biológica se describen en, por ejemplo, el documento WO 01/31580 (Barnhill *et al.*, “Methods and devices for identifying patterns in biological systems and methods of use thereof”, 3 de mayo de 2001); U.S. 2002/0193950 A1 (Gavin *et al.*, “Method or analyzing mass spectra”, 19 de diciembre de 2002); U.2003/0004402 A1 (Hitt *et al.*, “Process for discriminating between biological states based on hidden patterns from biological data”, January 2, 2003); y U.S. 2003/0055615 A1 (Zhang y Zhang, “Systems and methods for processing biological expression data” 20 de marzo de 2003).

Más específicamente, para obtener los biomarcadores Apo A1, transtiretina $\Delta N10$ y fragmento IAIH4, se usaron los datos de intensidad de pico de las muestras de pacientes con cáncer y controles sanos como “conjunto de descubrimiento”. Estos datos se combinaron y dividieron aleatoriamente en un conjunto de entrenamiento y un conjunto de ensayo para construir y analizar modelos predictivos multivariados que usan una versión no lineal de los clasificadores del análisis de separabilidad máxima unificada (“USMA”). Los detalles de los clasificadores de USMA se describen en U.S. 2003/0055615 A1.

En general, los datos generados de la Sección IV anterior se ingresaron en un algoritmo de diagnóstico (es decir, algoritmo de clasificación descrito anteriormente). El algoritmo de clasificación posteriormente se genera sobre la base del algoritmo de aprendizaje. El proceso involucra el desarrollo de un algoritmo que puede generar el algoritmo de clasificación. Los procedimientos de la presente invención generan un algoritmo de clasificación más preciso al tener acceso a numerosas muestras de cáncer ovárico y muestras normales de una cantidad suficiente basada en los cálculos estadísticos de la muestra. Las muestras se usan como un conjunto de datos de entrenamiento sobre el algoritmo de aprendizaje.

La generación de la clasificación, es decir, el algoritmo de diagnóstico es dependiente del protocolo del ensayo usado para analizar las muestras y generan los datos obtenidos en la Sección IV anterior. Es imperativo que el protocolo para la detección y/o medición de los marcadores (por ejemplo, en la etapa IV) sea el mismo que se usa para obtener los datos usados para desarrollar el algoritmo de clasificación. Las condiciones de ensayo, que se deben mantener en todos los sistemas de entrenamiento y clasificación incluyen tipo de chip y parámetros del espectrómetro de masas, así como protocolos generales para la preparación y análisis de la muestra. Si se cambia el protocolo para la detección y/o medición de los marcadores (etapa IV), el algoritmo de aprendizaje y el algoritmo de clasificación también deben cambiar. De modo similar, si el algoritmo de aprendizaje y el algoritmo de clasificación cambian, entonces también debe cambiar el protocolo de detección y/o medición de los marcadores (etapa IV) para ser compatible con el que se usa para generar el algoritmo de clasificación. El desarrollo de un nuevo modelo de clasificación debe requerir el acceso a un número suficiente de muestras de cáncer ovárico y normales, desarrollar un nuevo conjunto de entrenamiento de datos basados en un nuevo protocolo de detección, generar un nuevo algoritmo de clasificación usando los datos y finalmente, verificar el algoritmo de clasificación con un estudio multicéntrico.

Los modelos de clasificación se pueden formar y usar en cualquier ordenador digital adecuado. Los ordenadores digitales adecuados incluyen micro-, mini- o grandes ordenadores que usan cualquier sistema operativo estándar o especializado tal como un sistema operativo basado en Unix, Windows™ o Linux™. El ordenador digital que se usa puede estar separado físicamente del espectrómetro de masas que se usa para crear los espectros de interés, o se puede acoplar al espectrómetro de masas. Si está separado del espectrómetro de masas, los datos se deben ingresar en el ordenador por algún otro medio, ya sea manual o automatizado.

El conjunto de datos de entrenamiento y los modelos de clasificación de acuerdo con realizaciones de la invención se pueden realizar por el código de ordenador que se ejecuta o usa con un ordenador digital. El código del ordenador se puede almacenar en cualquier medio legible por el ordenador adecuado que incluye discos ópticos o magnéticos, tarjetas, cintas, etc., y se puede escribir en cualquier lenguaje de programación de ordenador adecuado que incluye C, C++, básico visual, etc.

VI. Ejemplos de realizaciones preferidas

En una realización preferida, una muestra de suero se recolecta de una paciente y posteriormente se fracciona usando una resina de intercambio iónico descrita anteriormente. Los biomarcadores en la muestra se capturan usando una matriz ProteinChip de cobre IMAC. Los marcadores posteriormente se detectan usando SELDI. En tal prueba se pueden detectar Apo A1, transtiretina Δ N10 y fragmento IAIH4. Los resultados posteriormente se ingresan en un sistema de computación, que contiene un algoritmo que se diseña usando los mismos parámetros que se usaron en el algoritmo de aprendizaje y el algoritmo de clasificación para determinar originalmente los biomarcadores. El algoritmo produce un diagnóstico basado en los datos recibidos referentes a cada biomarcador.

En realizaciones especialmente preferidas, también se detecta la cantidad de biomarcador CA125II, ya sea por medio de un procedimiento conocido, por ejemplo, inmunoensayo, o por medio de una matriz de chip de proteína de SELDI. En estas realizaciones, los resultados para el marcador CA125II también se ingresan en el algoritmo del ordenador y se usan para preparar un diagnóstico. Una prueba diagnóstica que se basa en la detección de los cuatro biomarcadores, Apo A1, transtiretina Δ N10, fragmento IAIH4 y CA125II tiene una especificidad de al menos aproximadamente 80%.

El diagnóstico se determina por el examen de los datos producidos de las pruebas de SELDI con el algoritmo de clasificación que se desarrolla usando los biomarcadores. El algoritmo de clasificación depende de las particularidades del protocolo de ensayo usado para detectar los biomarcadores. Estas particularidades incluyen, por ejemplo, la preparación de la muestra, tipo de chip y parámetros del espectrómetro de masas. Si los parámetros de ensayo cambian, el algoritmo debe cambiar. De modo similar, si el algoritmo cambia, el protocolo de ensayo debe cambiar.

En otra realización, la muestra se recolecta de la paciente. Los biomarcadores se capturan usando una matriz ProteinChip de anticuerpo que se describió anteriormente. Los marcadores se detectan usando un sistema de ensayo SELDI bioespecífico. En tal prueba se puede detectar Apo A1, transtiretina Δ N10 y fragmento IAIH4. Los resultados posteriormente se ingresan en un sistema de computación que contiene un algoritmo, que se diseña usando los mismos parámetros que se usaron en el algoritmo de aprendizaje y el algoritmo de clasificación para determinar originalmente los biomarcadores. El algoritmo produce un diagnóstico basado en los datos recibidos referentes a cada biomarcador.

En aún otra realización preferida, los marcadores se capturan y analizan usando los formatos no SELDI. En un ejemplo, la muestra se recolecta de la paciente. Los biomarcadores se capturan sobre un sustrato usando otros medios conocidos, por ejemplo, anticuerpos para los marcadores. Los marcadores se detectan usando procedimientos conocidos en la técnica, por ejemplo, procedimientos ópticos y del índice de refracción. Los ejemplos de procedimientos ópticos incluyen la detección de fluorescencia, por ejemplo, ELISA. Los ejemplos de índice de refracción incluyen la resonancia de plasmones superficiales. Los resultados para los marcadores posteriormente se someten a un algoritmo, que puede requerir o no inteligencia artificial. El algoritmo produce un diagnóstico basado en los datos recibidos referentes a cada biomarcador.

En cualquiera de los procedimientos anteriores, los datos de la muestra se pueden incorporar directamente de los medios de detección en un ordenador que contiene el algoritmo de diagnóstico. Alternativamente, los datos obtenidos se pueden incorporar en forma manual, o por un medio automatizado, en un ordenador separado que contiene el algoritmo de diagnóstico.

VII. Diagnóstico del sujeto y determinación del estado del cáncer ovárico

Cualquier biomarcador, en forma individual, es útil para ayudar a la determinación del estado del cáncer ovárico. Primero, el biomarcador seleccionado se mide en una muestra del sujeto usando los procedimientos descritos en la presente, por ejemplo, captura en biochips de SELDI seguido por la detección mediante espectrometría de masas. Posteriormente, la medición se compara con una cantidad diagnóstica o control que distingue un estado del cáncer ovárico de un estado no canceroso. En la presente la cantidad diagnóstica reflejará la información de que un biomarcador particular está regulado por aumento o regulado por disminución en un estado de cáncer en comparación con un

estado no canceroso. Tal como es bien entendido en la técnica, la cantidad diagnóstica particular usada se puede ajustar para aumentar la sensibilidad o la especificidad del ensayo diagnóstico de acuerdo con la preferencia del diagnóstico. La cantidad de ensayo en comparación con la cantidad diagnóstica en consecuencia indica el estado del cáncer ovárico.

Aunque los biomarcadores individuales son marcadores diagnósticos útiles, se ha hallado que una combinación de biomarcadores proporciona mayor valor predictivo que los marcadores individuales solos. Específicamente, la detección de una pluralidad de marcadores en una muestra incrementa el porcentaje de diagnósticos verdaderos positivos y negativos verdaderos y debería disminuir el porcentaje de diagnósticos falsos positivos o falsos negativos. En consecuencia, los procedimientos preferidos de la presente invención comprenden la medición de más de un biomarcador. Por ejemplo, los procedimientos de la presente invención tienen una AUC del análisis de ROC mayor de 0,50, los procedimientos más preferidos tienen una AUC mayor de 0,60, los procedimientos más preferidos tienen una AUC mayor de 0,70. Los procedimientos especialmente preferidos tienen una AUC mayor de 0,70 y los procedimientos de máxima preferencia tienen una AUC mayor de 0,80.

Por otra parte, usando un procedimiento que mide la combinación de los tres biomarcadores usados en la presente invención con el Marcador 4, por ejemplo, CA 125, mejora significativamente después de la realización del diagnóstico de CA 125, lo que proporciona una prueba que tiene una AUC mayor de 0,50, las pruebas más preferidas tienen una AUC mayor de 0,60, las pruebas más preferidas tienen una AUC mayor de 0,70.

A fin de usar los biomarcadores en combinación, es útil un algoritmo de regresión logística. El algoritmo UMSA es particularmente útil para generar un algoritmo de diagnóstico de los datos de ensayo. Este algoritmo se describe en Z. Zhang *et al.*, Applying classification separability analysis to microarray data. En: Lin SM, Johnson KF, eds. Methods of Microarray data analysis: papers from CAMDA '00. Boston: Kluwer Academic Publishers, 2001:125-136; y Z. Zhang *et al.*, Fishing Expedition -a Supervised Approach to Extract Patterns from a Compendium of Expression Profiles. En Lin SM, Johnson, KF, eds. Microarray Data Analysis II: Papers from CAMDA'01. Boston: Kluwer Academic Publishers, 2002.

El algoritmo de aprendizaje generará un algoritmo de clasificación multivariado (diagnóstico) ajustado a la especificidad y sensibilidad particulares deseadas por el operador. El algoritmo de clasificación posteriormente se puede usar para determinar el estado del cáncer ovárico. El procedimiento también involucra medir los biomarcadores seleccionados en una muestra (por ejemplo, Apo A1, transtiretina y fragmento IAIH4). Estas mediciones se someten al algoritmo de clasificación. El algoritmo de clasificación genera una puntuación indicadora que indica el estado del cáncer ovárico.

Asimismo, la mera presencia o ausencia de un marcador, sin cuantificar la cantidad de marcador, es útil y se puede correlacionar con un probable diagnóstico de cáncer ovárico. Por ejemplo, el fragmento IAIH4 se puede detectar con más frecuencia en pacientes humanos con cáncer ovárico que en sujetos normales. De la misma manera, por ejemplo, los biomarcadores Apo A1 y transtiretina ΔN10, se pueden detectar con menos frecuencia en los pacientes humanos con cáncer ovárico que en los sujetos normales. En consecuencia, una presencia o ausencia detectada respectivamente, de estos marcadores en un sujeto analizado indica que el sujeto tiene una mayor probabilidad de tener cáncer ovárico.

La medición de los marcadores involucra cuantificar los marcadores para correlacionar la detección de los marcadores con un probable diagnóstico de cáncer ovárico. En consecuencia, si la cantidad de marcadores detectados en un sujeto analizado es diferente en comparación con una cantidad control (es decir, mayor o menor que el control, de acuerdo con el marcador), posteriormente el sujeto analizado tiene una mayor probabilidad de tener cáncer ovárico.

La correlación puede tomar en cuenta la cantidad de marcador o marcadores en la muestra en comparación con una cantidad control del marcador o marcadores (regulación por aumento o disminución del marcador o marcadores) (por ejemplo, en los sujetos normales en los que el cáncer humano es indetectable). Un control puede ser, por ejemplo, la cantidad promedio o media del marcador presente en muestras comparables de sujetos normales en los que el cáncer humano es indetectable. La cantidad control se mide en condiciones experimentales iguales o sustancialmente similares que en la medición de la cantidad de ensayo. La correlación puede tomar en cuenta la presencia o ausencia de los marcadores en una muestra de ensayo y la frecuencia de detección de los mismos marcadores en un control.

La correlación puede tomar en cuenta a ambos factores para facilitar la determinación del estado del cáncer ovárico.

Los procedimientos de calificación del estado del cáncer ovárico también pueden servir para manejar el tratamiento del sujeto sobre la base del estado. Como se mencionó anteriormente, tal tratamiento describe las acciones del médico o clínico posteriores a la determinación del estado del cáncer ovárico. Por ejemplo, si el resultado de los procedimientos de la presente invención no es concluyente o existe una razón de que sea necesaria la confirmación del estado, el médico puede ordenar más pruebas. Alternativamente, si el estado indica que la cirugía es apropiada, el médico puede programar la cirugía para la paciente. En otros casos, la paciente puede recibir quimioterapia o tratamientos de radiación, sea en reemplazo de la cirugía o además de ella. Asimismo, si el resultado es negativo, por ejemplo, el estado indica cáncer ovárico de etapa final o si el estado es de otro modo agudo, no se justificaría ninguna acción adicional.

Por otra parte, si los resultados muestran que el tratamiento ha sido exitoso, puede no ser necesario un tratamiento adicional.

La invención también permite que tales procedimientos en que se miden los biomarcadores (o combinación de biomarcadores específicos) nuevamente después del tratamiento del sujeto. En estos casos, los procedimientos se usan para controlar el estado del cáncer, por ejemplo, respuesta al tratamiento del cáncer, remisión de la enfermedad o progresión de la enfermedad. Debido a la facilidad de uso de los procedimientos y la falta de invasividad de los procedimientos, los procedimientos se pueden repetir después de cada tratamiento que recibe la paciente. Esto permite al médico seguir la efectividad del tratamiento. Si los resultados muestran que el tratamiento no es efectivo, por consiguiente se puede alterar el curso del tratamiento. Esto permite al médico ser flexible en las opciones de tratamiento.

En otro ejemplo, los procedimientos para detectar marcadores se pueden usar para analizar e identificar los compuestos que modulan la expresión de estos marcadores *in vivo* o *in vitro*.

Los procedimientos descritos tienen también otras aplicaciones. Por ejemplo, los marcadores se pueden usar para analizar los compuestos que modulan la expresión de los marcadores *in vitro* o *in vivo*, estos compuestos a su vez pueden ser útiles para tratar o prevenir el cáncer ovárico en los pacientes. En otro ejemplo, los marcadores se pueden usar para controlar la respuesta a los tratamientos para el cáncer ovárico. En aún otro ejemplo, los marcadores se pueden usar en estudios hereditarios para determinar si el sujeto está en riesgo de desarrollar el cáncer ovárico. Por ejemplo, ciertos marcadores se pueden ligar genéticamente. Esto se puede determinar, por ejemplo, analizando las muestras de una población de pacientes con cáncer ovárico cuyas familias tienen antecedentes de cáncer ovárico. Los resultados posteriormente se pueden comparar con los datos obtenidos, por ejemplo, de pacientes con cáncer ovárico cuyas familias no tienen antecedentes de cáncer ovárico. Los marcadores que se ligan genéticamente se pueden usar para determinar si un sujeto cuya familia tiene antecedentes de cáncer ovárico está predispuesto a tener cáncer ovárico.

VIII. Kits

En aún otro aspecto, la presente invención proporciona un kit para calificar el estado del cáncer ovárico, en el que el kit se puede usar para medir los marcadores de la presente invención. Por ejemplo, el kit se puede usar para medir los marcadores descritos en la presente, dichos marcadores están presentes en forma diferencial en muestras de pacientes con cáncer ovárico y sujetos normales.

El kit de la invención tiene muchas aplicaciones. Por ejemplo, el kit se puede usar para diferenciar si un sujeto tiene cáncer ovárico o tiene un diagnóstico negativo, de este modo permite al médico o clínico diagnosticar la presencia o ausencia del cáncer. El kit también se puede usar para controlar la respuesta de la paciente en un curso del tratamiento que permite al médico modificar el tratamiento basado en los resultados de la prueba. En otro ejemplo, el kit se puede usar para identificar compuestos que modulan la expresión de uno o más de los marcadores en modelos animales *in vitro* o *in vivo* para el cáncer ovárico.

La presente invención en consecuencia proporciona un kit reivindicado, es decir, que comprende (a) un reactivo de captura o reactivos de captura que se unen a Apo A1, transtiretina ΔN10, y fragmento IAIH4; y (b) un recipiente que comprende Apo A1, transtiretina ΔN10 y fragmento IAIH4. El kit además puede comprender un reactivo de captura que une CA125.

Si bien el reactivo de captura puede ser cualquier tipo de reactivo, preferiblemente el reactivo es una sonda de SELDI.

También se describen kits, en los que el reactivo de captura comprende un IMAC. Cada uno de los tres marcadores identificados aquí se une a la matriz IMAC ProteinChip®. En consecuencia, una realización preferida incluye una prueba de alto rendimiento para la detección precoz del cáncer ovárico, que analiza una muestra de la paciente en la matriz IMAC ProteinChip® para estos tres analitos, así como el ELISA de CA-125 tradicional (o el ELISA CA-125 se puede transferir a la plataforma de la matriz ProteinChip®).

Los kits de la presente invención también pueden comprender una solución de lavado, o eluyente, que permite la retención selectiva del biomarcador unido al reactivo de captura en comparación con otros biomarcadores después del lavado. Alternativamente, el kit puede contener instrucciones para preparar una solución de lavado, en el que la combinación del adsorbente y la solución de lavado permiten la detección de los marcadores usando espectrometría de iones en fase gaseosa.

Preferiblemente, el kit comprende instrucciones escritas para su uso para la detección del cáncer y las instrucciones indican poner en contacto una muestra de ensayo con el reactivo de captura y detectar uno o más biomarcadores retenidos por el reactivo de captura. Por ejemplo, el kit puede tener instrucciones estándares que informan al consumidor cómo lavar el reactivo de captura (por ejemplo, sonda) después de que una muestra de suero sanguíneo se pone en contacto con el reactivo de captura. En otro ejemplo, el kit puede tener instrucciones para el pre-fraccionamiento de una muestra para reducir la complejidad de las proteínas de la muestra. En otro ejemplo, el kit puede tener instrucciones para automatizar el fraccionamiento u otros procesos.

Dichos kits se pueden preparar a partir de los materiales descritos anteriormente y la discusión previa de estos materiales (por ejemplo, sustratos de sonda, reactivos de captura, adsorbentes, soluciones de lavado, etc.) es completamente aplicable a esta sección y no se repetirá.

El kit puede comprender un primer sustrato que comprende un adsorbente en este (por ejemplo, una partícula funcionalizada con un adsorbente) y un segundo sustrato en el que el primer sustrato se puede ubicar para formar una sonda, que se puede insertar en forma removible en un espectrómetro de iones en fase gaseosa. El kit también puede comprender un solo sustrato, que está en forma de una sonda que se puede insertar en forma removible con adsorbentes sobre el sustrato. El kit también puede comprender una columna de pre-fraccionamiento (por ejemplo, columna de agarosa con azul de Cibacron, columna de agarosa anti-HSA, columna de exclusión por tamaño K-30, columna de intercambio aniónico Q, columna de ADN monocatenario, columna de lectina, etc.).

Opcionalmente, el kit también puede comprender una información estándar o control de modo que la muestra de ensayo se puede comparar con la información de control estándar para determinar si la cantidad de ensayo de un marcador detectado en una muestra es una cantidad diagnóstica compatible con un diagnóstico de cáncer ovárico.

También se describe un sistema que comprende una pluralidad de reactivos de captura, cada uno de los cuales tiene unido a este un biomarcador diferente seleccionado de Apo A1, transtiretina Δ N10, fragmento IAIH4, y al menos un marcador que se ajusta a la categoría del Marcador 4. Un ejemplo de tal sistema incluye, pero sin limitación, un conjunto de matrices ProteinChip[®], que comprende adsorbentes que se unen a uno o más de los biomarcadores seleccionados de Apo A1, transtiretina Δ N10, y fragmento IAIH4. En este tipo de sistema, hay una matriz ProteinChip[®] para cada uno de los biomarcadores. O, alternativamente, existe una matriz ProteinChip[®] para una pluralidad de marcadores del grupo de Apo A1, transtiretina Δ N10 y fragmento IAIH4 y una segunda matriz ProteinChip[®] para CA 125. Los ejemplos de otros sistemas incluyen aquellos en que los reactivos de captura son tubos de ensayo que contienen un anticuerpo para cada uno de los biomarcadores, sea en forma separada o en grupos. Los expertos en la técnica fácilmente podrían fabricar otros de estos artículos de acuerdo con las enseñanzas descritas en la presente.

Los siguientes ejemplos se ofrecen a modo de ilustración, no a modo de limitación. Si bien se han provisto ejemplos específicos, la anterior descripción es ilustrativa y no restrictiva. Alguna o más de las características de las realizaciones previamente descritas se pueden combinar de alguna manera con una o más características de alguna de las otras realizaciones de la presente invención.

Ejemplos

Materiales y Métodos

Muestras

Los datos del perfil proteómico se obtuvieron retrospectivamente de un total de 503 especímenes de suero recolectados en Groningen University Hospital (Groningen, Países Bajos), Duke University Medical Center (Durham, NC), Royal Hospital for Women (Sydney, Australia), y MD Anderson Cancer Center (Houston, TX). El grupo de cáncer ovárico estaba constituido por 65 pacientes con cáncer ovárico epitelial invasivo de etapas I/II y 88 pacientes con cáncer ovárico epitelial invasivo de etapas III/IV, 28 pacientes con tumores de bajo potencial maligno, y 14 pacientes con enfermedad recurrente. Los casos de cáncer fueron clasificados óptimamente por patólogos basados en los criterios FIGO. Entre los 65 pacientes con cáncer ovárico epitelial invasivo de etapas I/II, 20 eran serosos, 17 eran mucinosos, 15 eran endometrioides, 8 eran de células claras, 1 era carcinosarcoma y 4 eran carcinoma epitelial mixto. Las muestras también incluyeron 166 pacientes diagnosticados con masas pélvicas benignas y 142 controles sanos. Las características y los estadísticos básicos descriptivos de la población de estudio, que incluyen edad y niveles de CA125, se listan en la Tabla 1.

Todas las muestras de pacientes se recogieron antes de la cirugía o tratamiento y los especímenes de voluntarios sanos se recolectaron con aprobación institucional. La sangre se dejó coagular y se separó rápidamente el suero. Todas las muestras se conservaron a -70°C y se descongelaron inmediatamente antes del ensayo. Los niveles de CA 125 de todos los pacientes estuvieron disponibles a partir de un estudio previo usando un kit de radioinmunoensayo CA125II (Centocor).

Además de los 503 especímenes para el perfil proteómico, 142 especímenes de suero archivados recolectados para el ensayo de laboratorio clínico de rutina en Johns Hopkins Medical Institutions se analizaron para determinar los niveles de los biomarcadores identificados para lo cual estaba disponible la prueba de inmunoensayo. De estas muestras, 41 eran de pacientes con cáncer ovárico etapa final y 41 eran de mujeres sanas. Las restantes 60 muestras consistían en 20 de cada grupo de pacientes con cáncer de mamas, cáncer de colon y cáncer de próstata y se usaron para analizar la especificidad del sitio del tumor de los biomarcadores identificados (Tabla 3). Todas las muestras se procesaron dentro de dos a cuatro horas después de la recolección y se conservaron a 2-8°C durante un máximo de 48 horas antes de la congelación a -70°C. También se realizó el ensayo de CA125II usando un ensayo inmunoenzimométrico de dos sitios en el analizador Tosoh AIA-600 II (Tosoh Medics).

Ejemplo 1

Perfil de expresión de proteínas

- 5 *Fraccionamiento del suero:* Las muestras de suero se descongelaron en hielo y posteriormente se centrifugaron a 20000 g durante 10 minutos para eliminar el precipitado. Se mezclaron 20 μ l de suero con 30 μ l de un tampón desnaturante (U9: 9 M de urea, 2% de CHAPS, 50 mM de Tris pH 9,0) y se agitaron en vórtex durante veinte minutos a 4 grados. Para cada muestra, 180 μ l de resina de intercambio iónico Hyper Q DF se equilibraron en 200 μ l de tampón U1 (U9 que se diluyó 1:9 en 50 mM de Tris pH 9,0) tres veces. El suero desnaturado se aplicó a la resina y se dejó unir durante treinta minutos. El material no unido se recolectó y posteriormente 100 μ l de Tris 9,0 50 mM que contiene 0,1% de OGP se añadieron a la resina. Este lavado se recolectó y combinó con el material no unido (flujo continuo; fracción 1). Las fracciones posteriormente se recolectaron en un gradiente de pH escalonado usando dos veces 100 μ l de alícuotas de tampón de lavado a pH 7, 5, 4, 3, y solvente orgánico). Esto llevó a la recolección de un total de seis fracciones. El fraccionamiento se realizó en un dispensador de líquidos automatizado Biomek 2000 (Beckman) y un agitador Micromix (DPC). Una muestra de suero humano combinado control (Intergen) se procesó en forma idéntica para controlar el rendimiento del ensayo.

A. *Materiales para el perfil de expresión de proteínas*

- 20 Estación de trabajo Beckman Biomek 2000 Automatizada.
- Resina de intercambio iónico cerámica Q Hyper DF (Biosepra, France), microplaca de fondo en V de 96 pocillos.
- 25 Placa de filtro de membrana loprodina 96 pocillos (Silent Screen, Nalge Nunc).
- Tampón de equilibrado - 50 mM de Tris-HCl pH 9,0.
- 30 U9 - 9 M de Urea, 2,0% de CHAPS, 50 mM de Tris-HCl pH 9,0 U1 - 1 M de Urea, 0,22% de CHAPS, 50 mM de Tampón de Tris-HCl pH 9,0 - 100 mM de Tampón de Tris-HCl, 0,1% de OGP pH 9,0 pH 7,0 - 100 mM de Tampón de HEPES, 0,1% de OGP pH 7,0 pH 5,0 - 100 mM de Tampón de Na Acetato, 0,1% de OGP pH 5,0 pH 4,0 - 100 mM de Tampón de acetato de sodio, 0,1% de OGP pH 4,0 pH 3,0 - 50 mM de citrato de Na, 0,1% OGP pH 3,0.
- 35 Tampón Org - 33,3% de isopropanol/16,67% de acetonitrilo/0,5% de ácido trifluoroacético (TFA).

B. *Procedimiento*40 *Desnaturalización del suero*

Se pipetea 20 μ l de suero en una placa de parte inferior en V de 96 pocillos. Se añaden 30 μ l de U9 a cada pocillo que contiene suero. Se cubre la placa de 96 pocillos con película selladora de placa. Se agita en vortex a 4°C durante al menos 20 minutos mientras que se equilibra la resina.

45

Equilibrio de resina

- 50 Se lava la resina 5 veces con tres volúmenes de lecho de 50 mM de Tris-HCl pH 9,0. Esto se puede realizar en un tubo de centrífuga de 50 ml. Se crea una suspensión 50/50 de resina por el añadido de un volumen equivalente de 50 mM de Tris-HCl pH 9,0 a la resina. Se añaden 180 μ l de la suspensión 50/50 a cada pocillo de una placa de filtro de 96 pocillos. Se agita con vórtex en un tubo que contiene la suspensión en forma regular (cada dos o tres alícuotas) para asegurar una relación constante de resina a tampón. Posteriormente se filtra el tampón y se añaden 200 μ l de U1 y se filtra una vez más. Esto posteriormente se realiza dos veces más de la misma manera.

55

Aplicación e incubación de la muestra

- 60 La próxima etapa es unir el suero a la resina. La primera etapa de este proceso es pipetear 50 μ l de cada muestra en un pocillo correspondiente en la placa de filtro. Después se añaden 50 μ l de U1 a cada pocillo de la placa de muestra y se mezcla 5 veces. Posteriormente se pipetea 50 μ l de cada pocillo de la placa de muestra al pocillo correspondiente de la placa de filtro. Se agita en vórtex durante 30 minutos a 4°C.

- 65 La próxima etapa es recolectar las fracciones. Se coloca una placa de parte inferior en V de 96 pocillos bajo la placa de filtro. Se recolecta el flujo continuo de la placa de filtro. Posteriormente se añaden 100 μ l de tampón de lavado 1 a cada pocillo de la placa de filtro. Después se agita en vórtex durante 10 minutos a temperatura ambiente. La fracción 1 contiene el flujo continuo y eluyente a pH 9. Después se añaden 100 μ l de tampón de lavado 2 a cada pocillo de la placa de filtro. Se agita en vórtex durante 10 minutos a temperatura ambiente. Se coloca una placa de parte inferior en

- V limpia bajo la placa de filtro y se recolecta la fracción 2 en la placa. Se añaden 100 μ l de tampón de lavado 2 a cada pocillo en la placa de filtro. Se agita en vórtex durante 10 minutos a temperatura ambiente. Se recolecta el resto de la fracción 2 en la placa de parte inferior en V de 96 pocillos. La fracción 2 contiene eluyente a pH 7. Se añaden 100 μ l de tampón de lavado 3 a cada pocillo de la placa de filtro. Se agita en vórtex durante 10 minutos a temperatura ambiente.
- 5 Se coloca una placa de parte inferior en V limpia bajo la placa de filtro y se recolecta la fracción 3. Se añaden 100 μ l de tampón de lavado 3 a cada pocillo de la placa de filtro y se agita en vórtex durante 10 minutos a temperatura ambiente. Se recolecta el resto de la fracción 3 en la placa de parte inferior en V. La fracción 3 contiene eluyente de pH 5. Se añaden 100 μ l de tampón de lavado 4 a la placa de filtro y se agita en vórtex durante 10 minutos a temperatura ambiente. Se coloca una placa de parte inferior en V limpia debajo de la placa de filtro y se recolecta la fracción 4.
- 10 Posteriormente se añaden 100 μ l de tampón de lavado 4 a la placa de filtro y se agita en vórtex durante 10 minutos a temperatura ambiente. Se recolecta la fracción restante 4 en la placa de parte inferior en V. Fracción 4 contiene eluyente de pH 4. Posteriormente se añaden 100 μ l de tampón de lavado 5 a cada pocillo de la placa de filtro y se agita en vórtex durante 10 minutos a temperatura ambiente. Se coloca una placa de parte inferior en V limpia debajo de la placa de filtro y se recolecta la fracción 5. Posteriormente se añaden 100 μ l de tampón de lavado 5 a la placa de filtro y se agita en vórtex durante 10 minutos a temperatura ambiente. Se recolecta la fracción restante 5 en la placa de parte inferior en V. Fracción 5 contiene eluyente de pH 3. Se añaden 100 μ l de tampón de lavado 6 a la placa de filtro y se agita en vórtex durante 10 minutos a temperatura ambiente. Después se coloca una placa de parte inferior en V limpia debajo de la placa de filtro y se recolecta la fracción 6. Se añaden 100 μ l de tampón de lavado 6 a la placa de filtro y una vez más se agita en vórtex durante 10 minutos a temperatura ambiente. Se recolecta la fracción restante.
- 20 La fracción 6 contiene eluyente de solvente orgánico.

Se congelan las fracciones hasta que estén listas para el protocolo de unión del chip.

- Unión a la matriz:* 10% 1 de cada fracción se mezcló con 90 μ l de tampón de unión y se unió por triplicado a matrices IMAC, SAX, H50 y WCX ProteinChip (Ciphergen Biosystems). Para IMAC, el tampón de unión era 100 mM de fosfato de sodio pH 7,0 que contiene 500 mM de NaCl; para SAX, el tampón de unión era 100 mM de fosfato de sodio, pH 7; H50, el tampón de unión era 50% de acetonitrilo en H₂O; y para WCX, el tampón era 100 mM de acetato de Na pH 4,0. Se dejó que la unión ocurriera durante treinta minutos a temperatura ambiente. Posteriormente los chips se lavaron tres veces con tampón de unión y posteriormente tres veces con agua. La matriz usada fue ácido sinapínico.
- 30

- Obtención y análisis de datos:* En ambos análisis SELDI, todas las matrices se leyeron usando un lector de matriz Ciphergen PBS II ProteinChip®, un espectrómetro de masas de enfoque de tiempo diferido, lineal, desorción/ionización láser-tiempo de vuelo. Todos los espectros se obtuvieron en el modo de ion positivo. Los tiempos de retraso del enfoque de tiempo diferido se ajustaron a 400 ns para los péptidos y 1900 ns para las proteínas. Los iones se extrajeron usando un pulso de extracción de 3 kV, y se aceleraron a una velocidad final usando 20 kV de potencial de aceleración. El sistema empleó un láser de nitrógeno pulsado a velocidades de repetición que varían de 2 a 5 pulsos por segundos. La fluencia láser típica varió de 30-150 μ J/mm². Se usó un protocolo analítico automatizado para controlar el proceso de obtención de datos en la mayor parte del análisis de la muestra. Cada espectro tuvo un promedio de al menos 100 disparos de láser y se calibró en forma externa contra una mezcla de péptidos o proteínas conocidos. Los instrumentos se controlaron en forma semanal respecto del rendimiento usando estándares de insulina e inmunoglobulina. Cada chip se leyó con dos energías láser, baja y alta. Los espectros se calibraron en forma externa, se sustrajo el valor basal con un ajuste de 8 veces el ancho de ajuste, y posteriormente se normalizaron con la corriente iónica total (que excluye la región de la matriz).
- 45

Ejemplo 2

Análisis estadístico

- Descubrimiento del biomarcador:* Se seleccionaron los picos de masa calificados (S/N > 5, ventana de masa del grupo a 0,3%) dentro del intervalo de masa de M/Z 2 kD – 50 kD de los espectros de SELDI. A fin de obtener un nivel de varianza de datos más constante en todo el intervalo del espectro de interés, se aplicó la transformación logarítmica a la intensidad del pico antes del análisis posterior. Los datos de la intensidad de pico de los pacientes con cáncer ovárico epitelial de etapa temprana y los controles sanos de Duke University Medical Center (Ca n=36, HC n=47) y Groningen University Hospital (Ca n=20, HC n=30) se analizaron usando el algoritmo del análisis de separabilidad máxima unificada (UMSA) que se usó primero para el análisis de datos de la micromatriz y posteriormente para los datos de expresión de proteínas (ProPeak, 3Z Informatics). ((Li J, *et al.*, Clin Chem 2002; 48:1296-304; Rai AJ, *et al.*, Zhang Z, *et al.* Arch Pathol Lab Med 2002; 126:1518-26; Zhang Z, *et al.*, Applying classification separability analysis to microarray data. En: Lin SM, Johnson KF, eds. Methods of Microarray data analysis: papers from CAMDA'00. Boston: Kluwer Academic Publishers, 2001:125-136; Zhang Z, *et al.*, Fishing Expedition - a Supervised Approach to Extract Patterns from a Compendium Of Expression Profiles. En: Lin SM, Johnson KF, eds. Microarray Data Analysis II: Papers from CAMDA'01. Boston: Kluwer Academic Publishers, 2002).
- 60

- Para reducir la posibilidad de elegir picos como resultado de sesgos o artefactos en los datos, se analizaron los datos de los dos sitios en forma independiente. Se usó un procedimiento de re-muestreo repetitivo para seleccionar picos que contribuyeron en forma significativa y constante a la separación de los pacientes con cáncer ovárico de etapa temprana y los controles sanos. En cada corrida repetitiva, un porcentaje fijo de las muestras de cáncer y control se

seleccionó aleatoriamente con reemplazo para el análisis. Los picos individuales se clasificaron de acuerdo con sus contribuciones a una versión lineal del clasificador UMSA. La media y el desvío estándar de cada rango de pico se estimaron respecto de múltiples corridas (20-40). Los picos con rangos medios altos y desvíos estándares pequeños se seleccionaron para formar una lista breve de picos candidatos. Los resultados de los dos sitios posteriormente se compararon en forma cruzada para determinar un conjunto final de picos con patrones de expresión constantes como potenciales biomarcadores.

Modelos predictivos multivariados: Para construir modelos predictivos multivariados, los datos de los dos sitios se combinaron y posteriormente se dividieron aleatoriamente en un conjunto de entrenamiento y un conjunto de ensayo. El rendimiento del panel de biomarcadores potenciales y los modelos predictivos derivados se evaluaron primero en el conjunto de ensayo y finalmente se validaron sobre los datos independientes provenientes de los dos sitios restantes que no estaban involucrados en el descubrimiento del biomarcador y el proceso de construcción del modelo. Los procedimientos estadísticos para la evaluación incluyeron estimación de la sensibilidad y especificidad y el análisis de la curva de características de receptor-operador (ROC).

Ejemplo 3

Purificación de biomarcadores

Para todos los marcadores, el suero se fraccionó inicialmente usando el protocolo de intercambio aniónico utilizado para el perfil de expresión de proteínas. En cada etapa de purificación, las fracciones se controlaron en matrices NP20 o IMAC-cobre ProteinChip.

Purificación del marcador de 28 kD: 1 ml de la fracción de pH 4 de la separación de intercambio aniónico se añadió a 500 µl de RPC PoliBio 10-15 (Biosepra) y se incubó a 4°C durante 1 hora. Se recolectaron las fracciones que contienen cantidades crecientes de acetonitrilo con 0,1% de ácido trifluoroacético. La fracción de 75% de acetonitrilo/0,1% de ácido trifluoroacético se secó por speed-vac y se rehidrató en 100 µl de tampón de carga de muestra SDS-tricina sin DTT. Se cargaron 40 µl de muestra en 16% de gel de tricina y se corrió a 100 mV durante 4 horas. El gel se decoloró con el kit de azul coloidal (Pierce) y se escindió el de 28 kD.

Purificación del marcador de 12,8 kDa: 10 ml de la fracción de pH 4 de la separación de intercambio aniónico se ajustaron a pH 7,5 con 1 M de Tris HCl, pH 11 y se cargaron en 10 ml de perlas de MEP (Biospra) que habían sido prelavadas con 20 ml de PBS, pH 7,2 tres veces. La fracción del flujo continuo que contiene el pico se obtuvo después de agitar a 4°C durante 30 minutos. Debido a que esta fracción contiene una gran cantidad de albúmina, se realizó una inmunodepleción de albúmina. Las perlas de Proteína A se prelavaron tres veces con 1,5 ml de PBS que contiene 0,1% de triton-100 seguido por tres lavados con 1,5 ml de PBS. Se añadió 4 ml de anticuerpo anti-HSA (ICN) a 1,5 ml de las perlas de proteína A y se dejó acoplar durante la noche. Las perlas acopladas se lavaron con 1 ml PBS con 0,1% triton-100 tres veces y posteriormente tres veces con 1 ml de PBS. El flujo continuo de la columna MEP se añadió a las perlas y se incubó durante una hora a 4°C. El flujo continuo se obtuvo por centrifugado a 3000 rcf durante 1 minuto. La fracción del flujo continuo de la columna de proteína-A-anticuerpo antiHSA se añadió a una columna que contiene 1,5 ml de resina RPC PoliBio10-15 (Biosepra) que había sido prelavada cuatro veces con 1,5 ml de 0,1% de TFA. El flujo continuo se extrajo por centrifugado a 3000 rcf después de la incubación a 4°C durante 40 minutos con agitación suave y la perla se lavó con 0,8 ml de 0,1% de TFA. Se recolectaron las fracciones que contienen cantidades crecientes de acetonitrilo con 0,1% de ácido trifluoroacético. La fracción de 75% de acetonitrilo/0,1% de ácido trifluoroacético se secó por speed-vac y se rehidrató en 100 µl de tampón de carga de muestra SDS-tricina sin DTT. Se cargó 40 µl de muestra en 16% de gel de tricina y se corrió a 100 mV durante 4 horas. El gel se decoloró con el kit de azul coloidal (Pierce) y se escindió el de 12,8 kDa.

Purificación del biomarcador de 3272 dalton: 1 ml del flujo continuo del fraccionamiento de intercambio aniónico se cargó en 125 µl (250 µl de suspensión al 50%) de celulosa IMAC (Biosepra) acoplada con sulfato de cobre y se incubó a 4°C durante 1 hora. Las perlas posteriormente se lavaron con un gradiente creciente escalonado de imidazol (250 µl de 20 mM, 50 mM, 100 mM 150 mM y 200 mM de imidazol en 100 mM de NaPO₄, pH 7 con 500 mM de NaCl). Se cargó 200 µl de las fracciones que contienen el biomarcador (50-150 mM de imidazol) en una columna C18 (tecnologías ANSYS, Metachem polaris C18-A5U) y se lavó con TFA al 0,1% durante 5 minutos a 1 ml/min seguido por un gradiente de diez minutos desde 0% de ACN a 9% de ACN con 0,1% de TFA a 1 ml/min. La columna posteriormente se eluyó con un gradiente lineal desde 9% de ACN con 0,1% de TFA a 45% de ACN con 0,1% de TFA en 30 minutos a 1 ml/min. Las fracciones se recolectaron en alícuotas de 1 ml y el marcador eluyó en la fracción 38 (en la que la concentración de ACN es 34,2%).

Ejemplo 4

Identificación de biomarcadores

Las proteínas purificadas se digirieron con tripsina, y los fragmentos de tripsina se analizaron en el lector ProteinChip. Cada espectro tenía un promedio de al menos 250 disparos láser y se calibró en forma externa contra una mezcla de péptidos conocidos o se calibró en forma interna usando autólisis trípica y picos de la matriz. Las masas de

los picos se sometieron a la búsqueda del sitio de mapeo de péptidos ProFound (http://129.85.19.192/profound_bin/WebProFound.exe). Las secuencias de proteínas se recuperaron usando la base de datos NCBI. La confirmación de estas coincidencias con la base de datos se realizó usando un PE Sciex QStar (Concord, Canadá) equipado con una interfaz de la matriz ProteinChip (Ciphergen). En los experimentos de EM/EM, los espectros se obtuvieron en un espectrómetro de masas cuadripolar en tándem-tiempo de vuelo ScieQStar (Concord, Ontario, Canadá) equipado con una interfaz de la matriz Ciphergen PCI 1000 ProteinChip®. Los iones se crearon usando un láser de nitrógeno pulsado (Laser Science VSL 337 ND S, Franklin, MA, USA) operado a 30 pulsos por segundo que administra una fluencia de pulso promedio de 130 $\mu\text{J}/\text{mm}^2$. Se usó gas nitrógeno, a 10 mtorr de presión para el enfriamiento colisional de los iones formados así como para todos los experimentos de disociación inducida por colisión de baja energía (CID). La energía de colisión aplicada generalmente siguió la regla de 50 eV/kDa. Para los modos de EM y EM/EM, el sistema se calibró en forma externa usando una mezcla de péptidos conocidos. La identificación de la proteína se llevó a cabo usando el programa UCSF ProteinProspector EM-Tag (<http://prospector.ucsf.edu>). Las búsquedas en bases de datos con EM-Tag se realizaron usando los siguientes valores: *Homo sapiens*, digesto de tripsina (dos escisiones faltantes permitidas), cisteínas modificadas por carbamidometilación, tolerancia de masa iónica original y fragmento 50 ppm, y bases de datos NCBI o Swiss-Prot.

La confirmación de estas identidades se realizó por EIA o usando un inmunoensayo de la matriz ProteinChip.

Si bien estas proteínas se han caracterizado generalmente como reactantes de fase aguda, cabe mencionar que en estudios preliminares usando inmunoensayos, el nivel de apolipoproteína A1 no se halló alterado en pacientes con cáncer de mamas o colon y el nivel de prealbúmina tampoco ha sido alterado en pacientes con cáncer de mamas o próstata.

La transtiretina es una proteína de fase aguda negativa y se ha informado previamente que sus niveles están disminuidos en el cáncer ovárico epitelial. (Mahlck CG, *et al.*, Gynecol Obstet Invest, 1994; 37:135-40). La transtiretina es el principal transportador para la tiroxina y triiodotironina séricas, y facilita el transporte del retinol por medio de su interacción con la proteína de unión al retinol. Los ratones transgénicos que carecen de la expresión de transtiretina tienen niveles extremadamente bajos de retinol y proteína de unión al retinol, (van Bennekum AM, *et al.*, J Biol Chem 2001; 276:1107-13) y se ha demostrado que los niveles disminuidos de proteína de unión al retinol así como de proteína de unión al retinol celular se asocian con un aumento de la tasa de transformación maligna del epitelio ovárico (van Bennekum AM, *et al.*, J Biol Chem 2001; 276: 1107-13; Roberts D, *et al.*, DNA Cell Biol 2002; 21:11-9). Además, se ha informado que los niveles de las proteínas de unión al retinol cambian en el cáncer ovárico, por el análisis de la matriz de oligonucleótidos. (Giordano TJ, *et al.*, Am J Pathol 2001; 159: 1231-8).

Se ha demostrado que la porción carboxilo de ITIH4, de la cual deriva el biomarcador de m/z 3272 es un sustrato para la calicreína del plasma. (Pu XP, *et al.*, Biochim Biophys Acta 1994; 1208:338-43; Nishimura H, *et al.*, FEBS Lett 1995; 357: 207-11). Las proteasas de calicreína consisten en calicreínas plasmáticas y calicreínas tisulares, que tienen superposición de especificidad de sustrato. (Diamandis EP, *et al.*, Clin Chem 2002; 48: 1198-205). Las calicreínas tisulares son productos de una gran familia multigénica que incluye el antígeno específico de próstata (PSA; hK3), un marcador tumoral para el cáncer de próstata. Se ha hallado que varias calicreínas tisulares están desreguladas en el cáncer ovárico, incluyendo hK4, hK5, hK7, hK8, y hK9. (Yousef GM, *et al.*, Minerva Endocrinol 2002; 27:157-66).

La transtiretina $\Delta\text{N}10$ y el fragmento IAIH4 son productos del truncamiento de las proteínas maduras. Estos marcadores pueden ser el producto de la escisión con una o más proteasas, que incluyen calicreína plasmática, calicreínas tisulares, metaloproteasas de matriz, o prostatina, recientemente se informó una serina proteasa tipo tripsina que está aumentada en los casos de cáncer ovárico. (Mok SC, *et al.*, J Natl Cáncer Inst 2001; 93: 1458-64). Las proteasas que generan estos marcadores también se pueden usar como marcadores que se pueden combinar con los Marcadores 1-4 para conferir aun mayor sensibilidad y especificidad a un modelo predictivo.

Ejemplo 5

Poder discriminatorio de los biomarcadores individuales

En el conjunto de descubrimiento, las diferencias de los niveles de expresión de los tres biomarcadores entre los pacientes con cáncer ovárico de etapa temprana y los controles sanos fueron estadísticamente significativas ($P < 0,000001$ para los marcadores a m/z 12828 y 28043 y $P < 0,003$ para el marcador a m/z 3272) (Tabla 1).

Tabla 1. Características de las muestras de estudio y la distribución de edad, CA125, y los tres biomarcadores identificados.

Grupos	Número y origen* de muestra				Edad**Edad*		CA125 (U/ml)		Nivel de biomarcador (intensidad de pico normalizada) (media ± DE, mediana) & valor de P†					
									m/z 12828		m/z 28043		m/z 3272	
	A	B	C	D	Tot	<50 a	≥50 a		P		P		P	P
Conjunto de descubrimiento de biomarcadores														
Control sano	30	49			79	82	16	13,0±10,2,10,2		0,40±0,09, 0,42		0,58±0,25, 0,52		0,57±1,06, 0,32
Cáncer ovárico epitelial	20	37			57	24	33	104,1±145,4, 49,7	<0,000001	0,27±0,09,0,28	<0,000001	0,36±0,26,029	0,000002	2,70±2,61,1,76
Cáncer invasivo etapas I/II	20	22			42	17	25	98,2±143,7, 47,4	0,000001	0,27±0,10, 0,28	<0,000001	0,34±0,27, 0,26	0,000004	2,11 ±2,01. 1,54
Cáncer invasivo etapa IIIA	2				2	1	1	116,9±115,5,118,9	n.c.*	0,24±0,04,0, 0,24	n.c.*	0,29±0,16, 0,29	n.c.*	5,78 ±7,87,5,78
Tumor bajo potencial maligno I/II	13				13	6	7	120,8±162,9, 82,6	<0,000001	0,25±0,08. 0,28	<0,000001	0,42±0,22, 0,31	0,047136	4,11±2,82, 3,95
Conjunto de validación independiente														
Control sano			63		63	41	22	13,0±12,2, 10,5		0,39:1:0,09, 0,38		0,BO:0,4t, 0,74		0,64*0,81, 0,28
Cáncer ovárico epitelial total		138			138	35	103	330,7±337,2,195, 1	<0,000001	0,20±0,09, 0,20	<0,000001	0,42±0,30, 0,32	<0,000001	0,98±1,30. 0,41
Cáncer invasivo etapa VII		23			23	9	14	180,5±224,7,74,5	<0,000001	0,22±-0,09, 0,21	<0,000001	0,42±0,26,0,32	<0,000001	1,10±1,59, 0,37
Cáncer invasivo etapas III/IV		86			86	17	69	433,2±1:357,3,34 7,4	<0,000001	0,20_0,08, 0,19	<0,000001	0,39±0,26, 0,31	<0,000001	0,76±0,99, 0,37

Conjunto de validación independiente											
Recurrente	14	14	9	208,2±249,1, 95,4	<0,000001	0,20±0,10, 0,20	<0,000001	0,43±0,29, 0,36	0,002050	1,52±1,71, 0,74	0,004875
Tumor bajo potencial maligno etapas VII	12	12	3	88,3±152,9, 12,2	0,000173	0,221-0,11, 0,23	<0,000001	0,66±0,50, 0,37	0,319724	1,72±1,84, 0,91	0,001435
Tumor bajo potencial Maligno Etapas III/IV	3	3	1	84,2±124,2, 18,7	0,000009	0,19±0,04, 0,21	0,000155	0,55±0,24, 0,83	0,298460	0,64±0,42, 0,60	0,994928
Masas pélvicas benignas	50	90	26	166	73	91	36,1±5,45, 15,7	<0,000001†	0,301-0,12, 0,31	2,94-2,68, 2,16	n.c,‡
Grupos etarios-Controles sanos totales											
Edad < 50 a	25	37	41	103	103	14,0±12, 0,11,0	0,0087	0,41±0,08, 0,42	0,68 ± 0,3410,59	0,52±0,6 4, 0,32	0,101
Edad ≤50 a	5	11	22	38	38	10,4 ±7,7,9,5	0,0087	0,37±0,08, 0,36	0,67±0,36, 0,60	0,82±1,5 0,0,029	0,101
Grupos etarios – Cánceres ováricos epiteliales invasivos de etapas I/II											
Edad < 50 a	10	7	9	26	26	1101,9±16 4,7, 43,8	0,356	0,27±10,11, 0,27	0,41±0,31, 0,37	1,27±1,66, 0,33	0,100
Edad ≤50 a	10	15	14	39	39	144±188,7, 71,5	0,356	0,24± 0,09, 0,24	0,34±0,24,0,28	2,08±2,04, 1,52	0,100

*Origen de las muestras: (A) Groningen University Hospital, Groningen, Países bajos, (B) Duke University Medical Center. Durham, NC, (C) Royal Hospital for Women, Sydney, Australia, y (D) MD Anderson Cancer Center, Houston, TX. "Edad: Tres muestras con datos de edad faltantes, 1 control sano del conjunto de descubrimiento del biomarcador y 2 pacientes con masas pélvicas benignas en el conjunto de validación independiente. § Valores P estimados entre pacientes con cáncer ovárico y controles sanos o entre dos grupos etarios a menos que se especifique lo contrario. # Valor de P no computado debido al tamaño pequeño de la muestra. † Valores de P estimados de la prueba T de dos grupos entre los pacientes con masa pélvica benigna y los pacientes con cáncer ovárico epitelial en el conjunto de validación independiente. ‡ Valor de P no computado ya que la intensidad media del pico de m/z 3272 fue superior entre los pacientes con masa pélvica benigna que en los pacientes con cáncer ovárico epitelial en el conjunto de validación independiente.

La Figura 2 (paneles A-D) compara el poder discriminatorio de los tres biomarcadores individuales con el de CA125 usando los datos del análisis de la curva de características del receptor-operador (ROC) de los pacientes con cáncer ovárico de etapa temprana y controles sanos. En los paneles A-D, 1: CA125, 2: m/z 12,8 kD, 3: m/z 28 kD, 4: m/z 3272D. CA125 y m/z 12828 se desempeñaron en forma comparable en ambos conjuntos de descubrimiento y de validación independiente, mientras que los otros dos marcadores presentaron un área debajo de la curva (AUC) menor que CA125 en uno o ambos conjuntos de datos. Sin embargo, las correlaciones estimadas y los tres biomarcadores y CA125 fueron bajas (datos no mostrados) lo que indica la posibilidad de que estos fueran complementarios entre sí y una estrategia multivariada podría superar el ensayo individual de CA125.

Debido a que el 27% de las muestras de los controles sanos eran de mujeres de 50 o más en comparación con el 61% de ellas en el grupo de cáncer ovárico de etapa temprana ($P < 0,000001$), nosotros consideramos que estos marcadores podrían reflejar los cambios relacionados con la edad. Sin embargo, los biomarcadores identificados no fueron significativamente diferentes entre los grupos etarios o fueron diferentes en un nivel comparable al del CA125 (Tabla 1). En estudios previos basados en la población han mostrado que los niveles de apolipoproteína A1 realmente apenas aumentan con la edad. (Jungner I, I, *et al.*, Clin Chem 1998; 44:1641-9; Bachorik PS, *et al.*, Clin Chem 1997; 43:2364-78).

Ejemplo 6

Modelos predictivos multivariados

Se construyeron modelos predictivos multivariados usando clasificadores UMSA no lineales. El primero utilizó solo los tres biomarcadores como su entrada y el segundo usó los tres biomarcadores junto con el nivel de CA125. Los paneles E-H de la Figura 2 comparan el desempeño diagnóstico global de los dos modelos con el de CA125 usando el análisis de ROC. En los paneles E-H, O: CA125, K: modelo multivariado que usa los tres biomarcadores, K: modelo multivariado que usa los tres biomarcadores y CA125. En los datos de entrenamiento, el valor límite de 0,5 maximizó aproximadamente la suma de la sensibilidad y la especificidad. Usando este valor límite, los modelos se aplicaron a los datos de ensayo y los datos de validación independiente (Tabla 2). Para la discriminación entre los controles sanos y de cáncer ovárico invasivo de las etapas I/II en el conjunto de validación independiente, el modelo multivariado que usa los tres biomarcadores y CA125, con una sensibilidad de 82,6% (IC de 95%, 61,2-95,1%), presentó una especificidad de 93,7% (84,5-98,2%). En comparación, CA125 en el valor límite de 11 U/ml presentó la misma sensibilidad (82,6%), aún su especificidad fue solo 52,4% (39,4-65,1%). La Tabla 2 también incluye los resultados de las pacientes en condiciones benignas, cáncer invasivo en etapa final, o tumores con bajo potencial maligno en el conjunto de validación independiente.

(Tabla pasa a página siguiente)

Tabla 2. Distribuciones de los dos modelos predictivos multivariados y comparación del desempeño diagnóstico.

Conjunto de datos	N	CA125 a 35 U/ml	Modelo con 3 marcadores ^s		Modelo con 3 marcadores + CA125 ^s		CA125 a 11 U/ml*
		% de correctos (IC 95%) [†]	media ± DE, mediana	% de correctos (IC 95%)*	media ± DE, mediana	% de correctos (IC 95%) [†]	
Conjunto de entrenamiento (Parte de conjunto de descubrimiento de biomarcadores)							
Control sano	33	97,0% (84,2-99,9)	0,46 ±0,04, 0,44	87,9% (71,8-96,6)	0,21±0,15, 0,14 4	97,0% (84,2-99,9)	51,5% (33,5-691)
Cáncer ovárico epitelial	28	67,0% (47,7-84,1)	0,58±0,08, 0,60	89,3% (71,8-97,7)	0,85±0,12, 0,89	100,0%(87,7-100,0)	(71,8-97,7)
Cáncer invasivo etapas I/II	22	63,6% (40,7-82,8)	0,68±0,06, 0,60	86,4% (65,1-97,1)	0,84±0,12, 0,87	100,0% (84,6-100,0)	86,4% (65,1-97,1)
Cáncer invasivo etapa IIIA	1	n.c*	0,60±0,00,0,60	n.c*	0,84±0,00, 0,84	n.c*	n.c*
Conjunto de ensayo (Parte del Conjunto de descubrimiento de biomarcadores)							
Control sano	46	97,8% (88,5-99,9)	0,46±0,04,0,44	91,3% (79,2-97,6)	0,19±0,11, 0,16	97,8% (88,5-99,9)	52,2% (37,0-66,9)
Tumor bajo potencial maligno etapas I/II	5	80,0% (28,4-99,5)	0,58±0,04,0,57	100% (47,8-100,0)	0,93±0,05, 0,93	100% (47,8-100,0)	100% (47,8-100,0)
Cáncer ovárico epitelial	29	65,5% (45,7-82,1)	0,57±0,07, 0,60	82,8% (64,2-94,2)	0,79±0,27, 0,91	86,2% (68,3-96,1)	100,0%(86,1-100,0)
Cáncer invasivo etapas I/II	20	70,0% (45,7-88,1)	0,56±0,07,0,59	75,0% (50,9-91,3)	0,7410,31, 0,89	80,0% (56,3-94,3)	100,0%(83,2-100,0)
Cáncer invasivo Etapa IIIA	1	n.c.*	0,61±0,00, 0,61	n.c.*	0,97±0,00, 0,97	n.c.*	n.c.*
Tumor bajo potencial maligno etapas I/II	8	50,0% (15,7-84,3)	0,58±0,05, 0,58	100,0%(63,1-100,0)	0,87±0,13, 0,91	100,0%(63,1-100,0)	100,0%(63,1-100,0)
Conjunto de validación independiente							
Control sano	63	96,8% ^h (89,0-99,6)	0,46±0,03,0,44	87,3% (76,5-94,4)	0,268±0,15, 0,23	93,7% (84,5-98,2)	52,4% (39,4-65,1)
Cánceres ováricos epiteliales totales	138	76,8% (68,9-83,6)	0,598±0,07, 0,62	83,3% (76,1-89,1)	0,818±0,27,0,98	81,9% (74,4-78,9)	88,4% 81,9-93,2)

(continuación) **Conjunto de validación independiente**

Cáncer invasivo etapas I/II	23	65,2% (42,7-83,6)	0,58±0,07, 0,62	87,0% (66,4-97,2)	0,73±0,28, 0,81	82,6% (61,2-95,1)	82,6% (61,2-95,1)
Cáncer invasivo etapas III/IV	86	87,2% (78,3-93,4)	0,60±0,07, 0,62	87,2% (78,3-93,4)	0,88±0,23, 0,99	87,2% (78,3-93,4)	94,2% (87,0-98,1)
Recurrente	14	78,6% (49,2-95,3)	0,58±0,08, 0,61	71,4% (41,9-91,6)	0,81±0,25, 0,91	85,7% (57,2-98,2)	82,9% (66,1-99,8)
Tumor bajo potencial maligno etapas I/II	12	33,3% (9,9-65,1)	0,58±0,07, 0,60	75,0% (42,8-94,5)	0,62±0,31, 0,63	50,0% (21,1-78,9)	58,3% (27,7-84,8)
Tumor bajo potencial maligno etapas III/IV	3	33,3% (0,8-90,6)	0,52±0,10, 0,47	33,3% (0,8-90,6)	0,43±0,48, 0,17-	33,3% (0,8-90,6)	66,7% (9,4-99,2)
Masas pélvicas benignas	166	75,9% (68,7-82,2)	0,54±0,07, 0,531	28,9% (22,2-36,5)	0,62±0,27, 0,66t	33,1% (26,0-40,9)	31,9% (24,9-39,6)

§ Los valores de P estimados entre los pacientes con cáncer ovárico y controles sanos se calcularon en <0,000001 en todas las comparaciones en ambos modelos, excepto en la comparación entre el tumor de bajo potencial maligno de etapa III/IV con los controles sanos para los que los valores de p fueron 0,316324 y 0,085588 para el modelo con tres marcadores y para el modelo con tres marcadores más CA125, respectivamente. # No computado debido al pequeño tamaño de la muestra. † Los valores de P estimados de la prueba T de dos grupos entre los pacientes con masa pélvica benigna y los pacientes con cáncer ovárico epitelial en el conjunto de validación independiente se calcularon que son 0,011870 para el modelo con tres marcadores y <0,000001 para el modelo con tres marcadores y CA125. ‡ Especificidad para los controles sanos y pacientes con masa pélvica benigna y sensibilidad para los pacientes con cáncer ovárico.

*Valor límite de CA125 a 11 U/ml seleccionado para coincidir con la sensibilidad (82,6%) en el conjunto de validación independiente del modelos con tres marcadores más CA125 para los pacientes con cáncer ovárico invasivo de las etapas I/II.

La Figura 3 grafica los patrones de distribución de CA125, los tres biomarcadores, y el resultados de los dos modelos para las muestras en todos los grupos de diagnóstico. El eje y es de intensidad relativa en escala lineal para los tres biomarcadores, los niveles séricos en escala log para CA125, y el valor continuo entre 0 (menor riesgo de cáncer) y 1 (mayor riesgo de cáncer) para los dos modelos. Los grupos de muestra incluyeron: A) controles sanos, B) benigno, C) cáncer invasivo de etapas I/II, D) cáncer invasivo de etapas III/IV, E) recurrente, F) tumor con bajo potencial maligno de etapas I/II. Dos casos de casos de cáncer invasivo IIIc del conjunto de descubrimiento del biomarcador y tres tumores con bajo potencial maligno de etapas III/IV del conjunto de validación independiente no se graficaron.

Cabe mencionar que con la excepción de m/z 3272, los otros dos biomarcadores así como los dos modelos predictivos fueron moderadamente capaces de detectar cáncer invasivo de etapas I/II de los casos benignos ($P = 0,004$ y $0,001$ para m/z 12828 y 28043, respectivamente, y $P = 0,003$ y $0,0001$ para los modelos sin CA125 y con CA125, respectivamente).

Ejemplo 7

Validación independiente mediante inmunoensayos

Los 142 especímenes archivados se analizaron para determinar apolipoproteína A1 usando un inmunoensayo turbidimétrico en un formato de placa de microtitulación (Wako Chemical USA), y transtiretina $\Delta N10$ usando un inmunoensayo turbidimétrico potenciado por partícula en el Dimension RxL Instrument (Dade-Behring) (Tabla 3).

(Tabla pasa a página siguiente)

Tabla 3. Distribuciones de niveles séricos de CA125, apolipoproteína A1, y transtiretina en 142 muestras archivadas para la validación independiente mediante inmunoensayos.

Grupo de diagnóstico	N	Edad (años) (media± DE)	CA125 (U/mL)				Apolipoproteína A1 (mg/dl)				Prealbúmina (mg/dl)			
			Media	DE	Median	P	Media	DE	Mediana	P	Media	DE	Median	P
Mujeres sanas	41	33±10	17,9	23,0	10,2		152,8	28,1	154,3		26,7	5,8	25,1	
Cáncer ovárico de etapa IIIC/IV	41	59±11	2387,9	4723,3	426,6	0,001895	121,5	41,9	123,9	0,000151	19,8	7,1	18,7	0,000006
Cáncer de mamas	20	53±14	13,4	13,0	7,9	0,416705	154,4	31,8	145,9	0,844163	26,8	4,8	26,4	0,928519
Cáncer de colon	20	70 ±15	17,0	14,2	10,9	0,867806	144,2	39,7	149,5	0,330148	22,4	5,1	22,8	0,006889
Cáncer de próstata	20	57 ±7	13,1	21,2	7,6	0,430720	138,2	21,2	136,1	0,043676	27,8	7,8	29,8	0,546918

Los niveles séricos de CA125 estaba regulados por aumento, mientras que los niveles de apolipoproteína A1 y transtiretina $\Delta N10$ estaba regulados por disminución entre los 41 pacientes con cáncer ovárico en etapa final en comparación con los 41 controles sanos ($P = 0,001895$, $0,000151$, y $0,000006$, respectivamente). El nivel sérico medio de apolipoproteína A1 entre los controles sanos no fue significativamente diferente del de los pacientes con cáncer de mamas o colorrectal ($P = 0,844163$, $0,330148$, respectivamente) y solo con diferencia marginal del de los pacientes con cáncer de próstata ($P = 0,043676$). El nivel sérico medio de transtiretina estaba regulado por disminución entre los pacientes con cáncer colorrectal ($P = 0,006889$) aunque en un grado menor que en los pacientes con cáncer ovárico. No hubo diferencias significativas en los niveles séricos medios de transtiretina $\Delta N10$ entre los controles sanos y los pacientes con cáncer de mamas o próstata ($P = 0,928519$, $0,546918$, respectivamente).

Ejemplo 8

Algoritmo de clasificación

Con referencia al algoritmo de clasificación ilustrado en forma de diagrama en la Figura 4, los Módulos 1-3 se entrenaron con el algoritmo de aprendizaje UMSA. El módulo clasificador final, sin embargo, tiene la misma forma matemática que un módulo clasificador 1 con máquina del vector soporte regular.

Módulo clasificador UMSA 1

$$CA125 \text{ nm} = \log(CA125 + 0,01)$$

$$m/z \ 12,9 \text{ Knm} = (m/z \ 12828 - 61,103)/239,031 \quad m/z \ 28 \text{ Knm} = (m/z \ 28043 - 61,3043)/238,9799 \quad m/z \ 3272 \text{ nm} = \log(m/z \ 3272 + 0,01)$$

$\log_0$: logaritmo natural

Función Kernel: polinómica $\langle X(:,i), X(:,j) \rangle^3, 0$.

Vectores y coeficientes de soporte

CA125nm	m/z12,9Knm	m/z28Knm	m/z3272nm	y	alfa
2,83966	-0,25465	-0,25597	-1,34323	1	0,00409
3,05918	-0,25416	-0,25435	-1,68740	-1	0,03389
2,39880	-0,25428	-0,25529	0,42657	1	0,05926
3,61658	-0,25450	-0,25578	-1,51413	1	0,22118
3,23120	-0,25417	-0,25493	-1,20065	-1	0,29988

ES 2 346 202 T3

Módulo clasificador UMSA 2

CA125 nm = $\log(\text{CA125} + 0,01)$

5 m/z 12,9 Knm = $(m/z$ 12828 - 0,345)/0,1114 m/z 28 Knm = $(m/z$ 28043 - 0,4834)/0,2792 m/z 3272 nm = $\log(m/z$ 3272 + 0,01)

Log₀: logaritmo natural

10 Función Kernel: $\exp(-|X(:,i)-X(i,j)|^2/(2*(1,0)^2))$.

Vectores y coeficientes de soporte

15

20

25

30

35

40

45

50

55

60

65

Valores del marcador					
CA125nm	m/z 12,9nm	m/z 28nm	m/z 3272nm	y	alfa
2,3618	0,1795	-0,3990	0,8717	1	0,0151
2,5734	0,4758	1,2307	-1,0328	-1	0,0582
1,6114	-1,1939	0,1311	-1,3318	-1	0,2112
2,1175	0,1975	-0,6103	-1,2588	-1	0,2705
3,5178	-0,6014	-1,3195	1,4120	1	0,2797
3,8525	-1,2926	-0,9327	-0,2890	1	0,3556
2,9658	-0,6014	0,2958	1,7483	1	0,3573
2,0028	-0,1167	1,4885	-1,7603	-1	0,4172
4,1061	-0,4668	1,0766	1,1743	1	0,4281
4,1776	-1,1400	-0,6927	1,3463	1	0,4927
1,5497	0,3142	1,6390	-0,6714	-1	0,6698
1,4839	-0,8707	-0,1519	-0,9113	-1	0,6860
3,9062	0,0987	1,9542	1,8925	1	0,7588
2,3988	-0,1975	-0,6748	0,4266	1	1,1697
3,6166	-0,6732	-1,0974	-1,5141	1	1,8394
3,2312	0,0269	-0,3703	-1,2006	-1	1,8394

ES 2 346 202 T3

Módulo clasificador UMSA 3

Función Kernel: polinómica $\langle X(:,i), X(:,j) \rangle^{2.0}$

5 $X1 = \exp(\text{resultado del módulo } 1) / (1 + \exp(\text{resultado del módulo } 1))$

$X2 = \text{resultado del módulo } 2.$

10 *Vectores y coeficientes de soporte*

	X1	X2	y	alfa
15	0,72862	0,41333	1	0,830900
	0,99835	0,25941	1	1,641283
20	0,39802	0,57799	1	1,839397
	0,96185	0,23167	1	1,839397
	0,58865	0,18582	-1	1,839397
25	0,96194	0,21066	1	1,839397
	0,55377	0,10709	-1	1,839397
	0,78444	0,05422	-1	1,839397
30	0,95604	-0,01117	-1	1,839397
	0,48706	0,28531	-1	2,343210

35 *Pos-procesamiento*

$$40 \quad Y = \begin{cases} \text{Resultado del módulo, si } (CA125 \leq 75); \\ \text{Resultado del módulo.} + \log((CA125 + 0.0001)/75) * 8 / \log(10/3), \end{cases}$$

45 En otro caso,

$$50 \quad \text{Resultado del módulo} = \exp(Y/2) / (1 + \exp(Y/2))$$

55

60

65

REIVINDICACIONES

1. Un procedimiento para calificar el estado del cáncer ovárico en un sujeto, que comprende:

- (a) medir la cantidad de Apo A1 (apolipoproteína A1), transtiretina Δ N10 y fragmento IAIH4 en una muestra de suero, y
- (b) correlacionar la medición con el estado del cáncer ovárico, en el que el fragmento IAIH4 consiste en la secuencia de aminoácidos de la SEQ ID NO: 1, y la transtiretina Δ N10 es una forma truncada de prealbúmina que carece de los diez aminoácidos N-terminales, y en el que
 - (i) la ausencia de Apo A1, o una disminución estadísticamente significativa de la cantidad de Apo A1, en comparación con un control normal, se correlaciona con un estado del cáncer ovárico;
 - (ii) la ausencia de transtiretina Δ N10, o una disminución estadísticamente significativa de la cantidad de transtiretina Δ N10, en comparación con un control normal, se correlaciona con un estado del cáncer ovárico;
 - (iii) la presencia del fragmento IAIH4, o un aumento de la cantidad de fragmento IAIH4, en comparación con un control normal, se correlaciona con un estado del cáncer ovárico.

2. El procedimiento de acuerdo con la reivindicación 1, en el que el estado del cáncer ovárico se selecciona del grupo que consiste en el sujeto con riesgo de cáncer, la presencia o ausencia de enfermedad, la etapa de la enfermedad y la efectividad del tratamiento de la enfermedad.

3. El procedimiento de acuerdo con la reivindicación 1, que además comprende medir al menos un biomarcador conocido en una muestra del sujeto y correlacionar la medición del biomarcador conocido y la medición de Apo A1, transtiretina Δ N10 y fragmento IAIH4 con el estado del cáncer ovárico, en el que el biomarcador conocido se selecciona de CA125, CA125 II, CA15-3, CA19-9, CA72-4, CA 195, inhibidor de tripsina asociado al tumor (TATI), CEA, fosfatasa alcalina placentaria (PLAP), Sialil TN, galactosiltransferasa, factor estimulante de colonias de macrófagos (M-CSF, CSF-1), ácido lisofosfatídico (LPA), componente de 110 kD del dominio extracelular del receptor del factor de crecimiento epidérmico (p110EGFR), calcireínas del tejido, por ejemplo calcireína 6 y calcireína 10 (NES-1), prostasina, HE4, creatina quinasa B (CKB), LASA, HER-2/neu, péptido de gonadotropina urinaria, Dianon NB 70/K, antígeno peptídico del tejido (TPA), osteopontina y haptoglobina, y variantes de proteína (por ejemplo, formas de escisión, isoformas) de los marcadores.

4. El procedimiento de acuerdo con la reivindicación 1, en el que la medición comprende: (a) proveer una muestra de sangre o un derivado de sangre de un sujeto; (b) fraccionar proteínas de la muestra en una resina de intercambio iónico y recolectar fracciones que contienen ApoA1, transtiretina Δ N10 y fragmento IAIH4; y (c) capturar Apo A1, transtiretina Δ N10 y fragmento IAIH4 de las fracciones sobre una superficie de un sustrato que comprende reactivos de captura que se unen a los biomarcadores proteicos.

5. El procedimiento de acuerdo con la reivindicación 4, en el que el sustrato es una sonda de SELDI que comprende una superficie de cobre IMAC y en el que los biomarcadores proteicos se detectan por SELDI.

6. El procedimiento de acuerdo con la reivindicación 4, en el que el sustrato es una sonda de SELDI que comprende reactivos de afinidad bioespecíficos que se unen a Apo A1, transtiretina Δ N10 y fragmento IAIH4 y en el que los biomarcadores proteicos se detectan por SELDI.

7. El procedimiento de acuerdo con la reivindicación 4, en el que el sustrato es una placa de microtitulación que comprende reactivos de afinidad bioespecíficos que se unen a Apo A1, transtiretina Δ N10 y fragmento IAIH4 y los biomarcadores proteicos se detectan por inmunoensayo.

8. El procedimiento de acuerdo con la reivindicación 1, en el que al menos un biomarcador se mide utilizando una matriz de biochip.

9. El procedimiento de acuerdo con la reivindicación 8, en el que la matriz de biochip es una matriz de chip de proteína.

10. El procedimiento de acuerdo con la reivindicación 8, en el que la matriz de biochip es una matriz de ácido nucleico.

11. El procedimiento de acuerdo con la reivindicación 8, en el que al menos un biomarcador se inmoviliza en la matriz de biochip.

ES 2 346 202 T3

12. El procedimiento de acuerdo con la reivindicación 1, en el que los biomarcadores proteicos se miden por SELDI.

5 13. El procedimiento de acuerdo con la reivindicación 1, en el que los biomarcadores proteicos se miden por inmunoensayo.

14. El procedimiento de acuerdo con la reivindicación 1, en el que la correlación se realiza con un algoritmo de clasificación de programas de computación.

10 15. Un kit que comprende: (a) un reactivo de captura o reactivos de captura que se une(n) a Apo A1, transtiretina Δ N10 y fragmento IAIH4, y (b) un recipiente que comprende Apo A1, transtiretina Δ N10 y fragmento IAIH4, en el que el fragmento IAIH4 consiste en la secuencia de aminoácidos de la SEQ ID NO:1, y la transtiretina Δ N10 es una forma truncada de prealbúmina que carece de los diez aminoácidos N-terminales.

15 16. El kit de acuerdo con la reivindicación 15, en el que el reactivo de captura es una sonda de SELDI.

17. El kit de acuerdo con la reivindicación 15, que además comprende un reactivo de captura que se une a CA125.

20

25

30

35

40

45

50

55

60

65

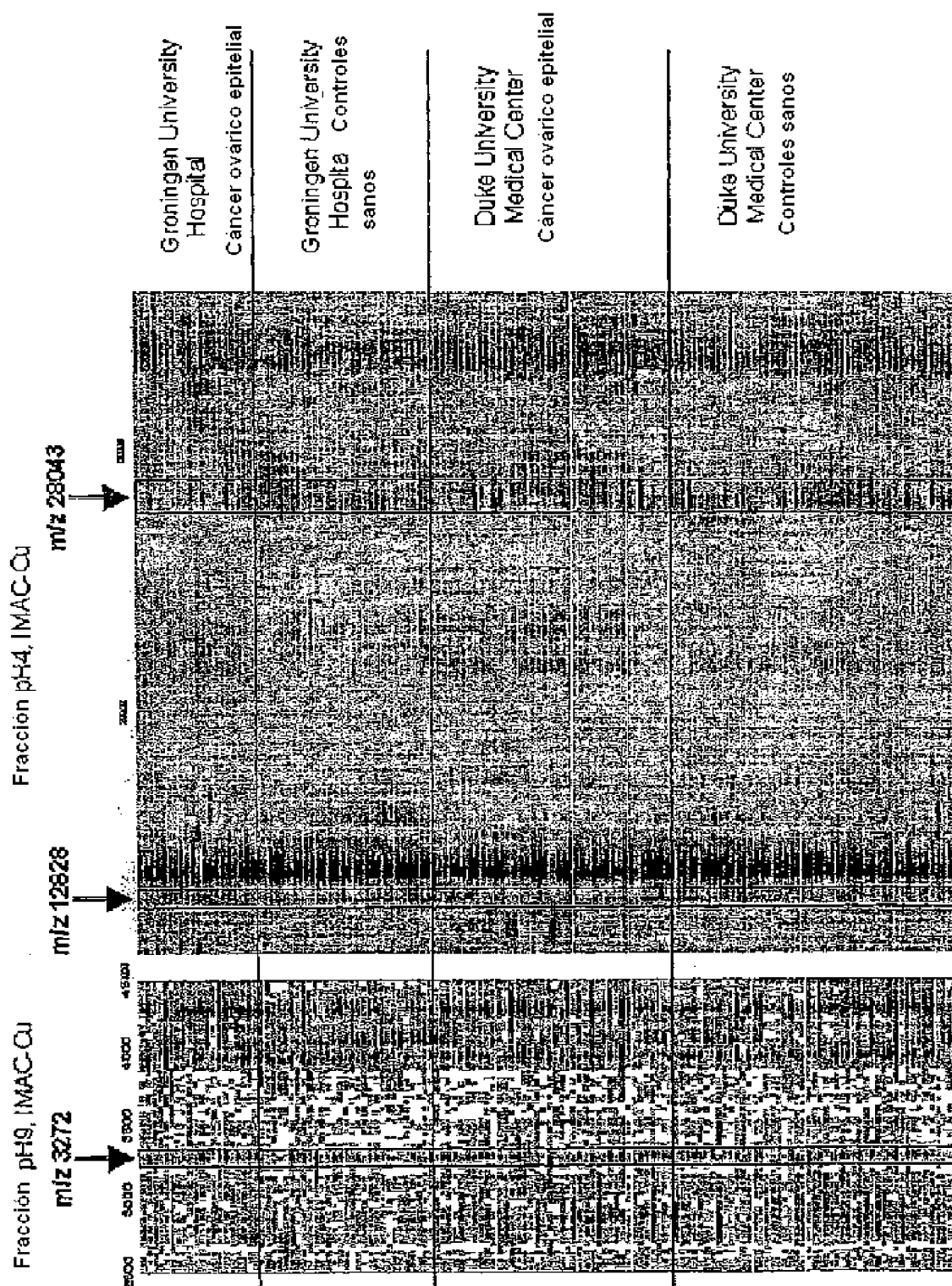
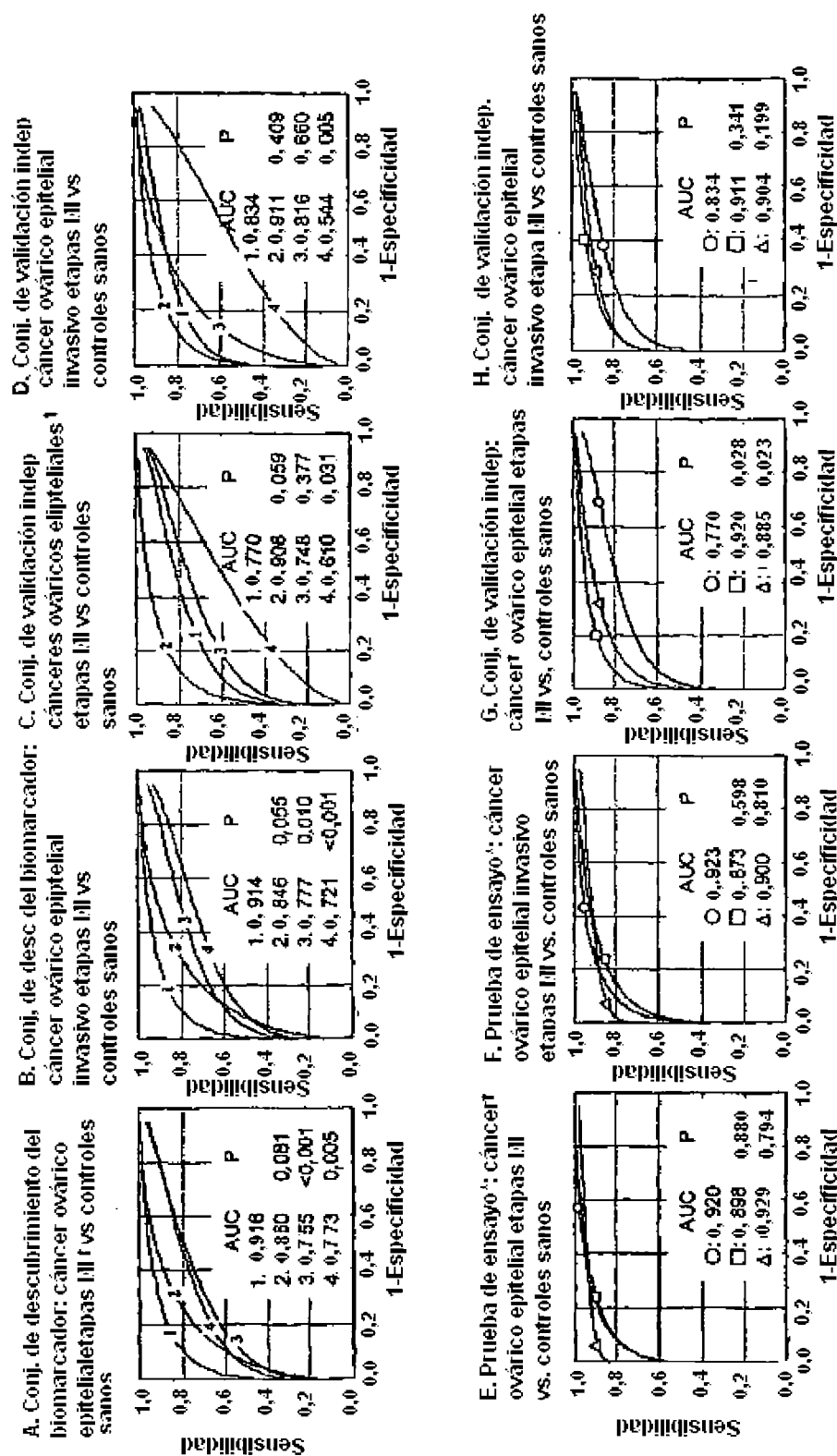


Figura 1



[†]Parte del conjunto de descubrimiento de biomarcadores. † En los paneles A, C, E y G, el cáncer ovárico epitelial incluye los tumores invasivos de bajo potencial maligno.

Figura 2

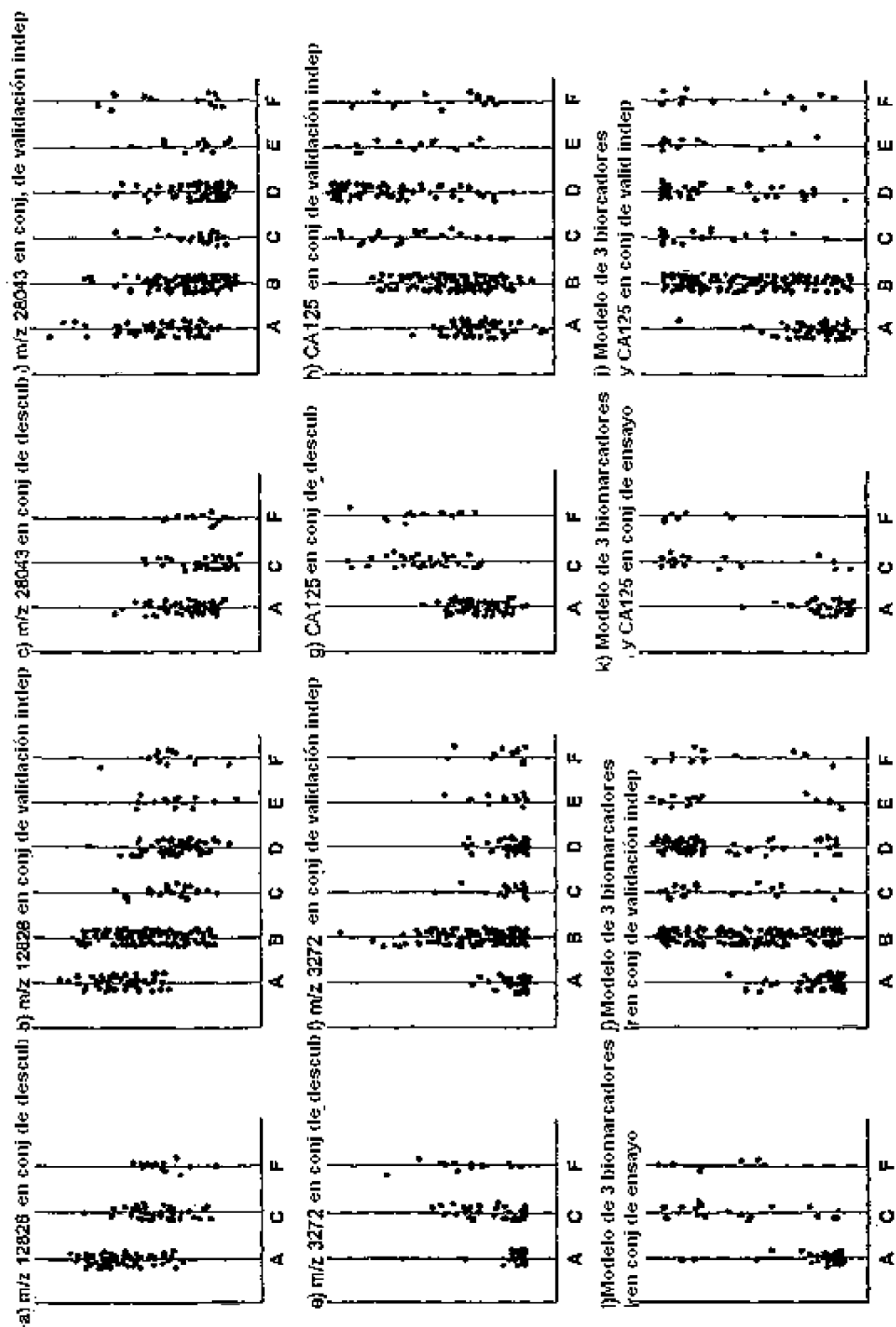


Figura 3

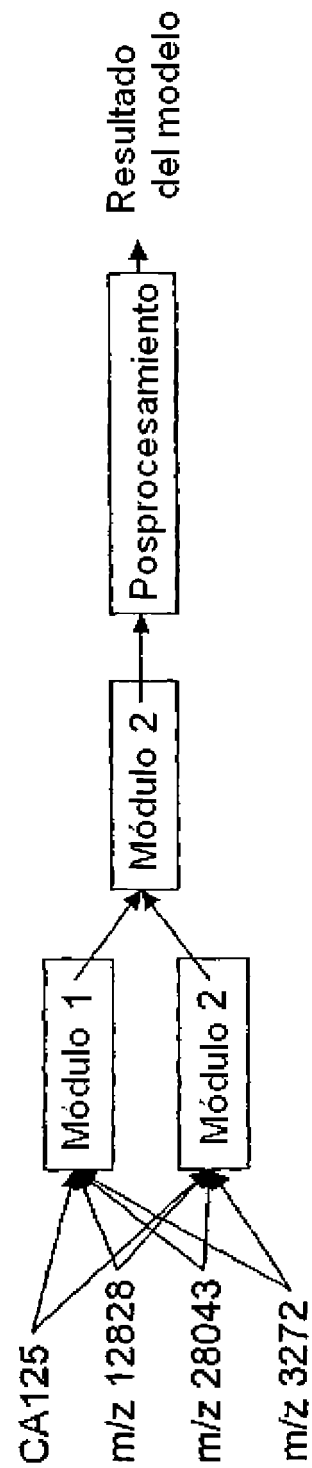


Figura 4