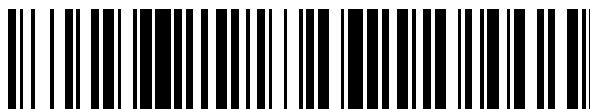


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 739 426**

51 Int. Cl.:

G01N 33/50 (2006.01)

G01N 33/574 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.11.2014** **PCT/EP2014/075427**

87 Fecha y número de publicación internacional: **28.05.2015** **WO15075242**

96 Fecha de presentación y número de la solicitud europea: **24.11.2014** **E 14820760 (8)**

97 Fecha y número de publicación de la concesión europea: **27.03.2019** **EP 3071966**

54 Título: **Olfactomedina-4, neudesina y desmoplaquina como biomarcadores del cáncer de mama**

30 Prioridad:

22.11.2013 EP 13306604

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
31.01.2020

73 Titular/es:

**INSTITUT DE CANCÉROLOGIE DE L'OUEST
(33.3%)
2 rue Moll
49000 Angers, FR;
UNIVERSITÉ D'ANGERS (33.3%) y
INSTITUT NATIONAL DE LA SANTÉ ET DE LA
RECHERCHE MÉDICALE (33.3%)**

72 Inventor/es:

**GUETTE, CATHERINE;
RARO, PEDRO;
COQUERET, OLIVIER;
BARRE, BENJAMIN y
CAMPONE, MARIO**

74 Agente/Representante:

CURELL SUÑOL, S.L.P.

ES 2 739 426 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

Olfactomedina-4, neudesina y desmoplaquina como biomarcadores del cáncer de mama.

5 Introducción

La presente invención se encuentra en el campo técnico de la gestión del cáncer de mama, y más particularmente se refiere al diagnóstico de cáncer de mama. La invención se basa más particularmente en el hallazgo de que biomarcadores específicos se expresan de manera aberrante en la sangre de pacientes con

Con más de 1.3 millones de casos de cánceres de mama invasivos diagnosticados anualmente, y más de 450,000 muertes notificadas al año, el cáncer de mama es el tumor maligno más común diagnosticado en mujeres y una de las principales causas de muerte relacionada con el cáncer en mujeres.

La detección temprana del cáncer de mama es la piedra angular para reducir las tasas de mortalidad en este cáncer que afecta a una de cada nueve mujeres. Actualmente, las campañas de cribado del cáncer de mama se realizan a través de una mamografía y, aunque no existe duda de su eficacia, este enfoque presenta limitaciones en cuanto a sensibilidad en mujeres con tejido mamario muy denso y en mujeres jóvenes consideradas "en riesgo" (historia familiar o predisposición genética) para quienes no se recomienda la utilización regular de radiación ionizante. Además, según un trabajo reciente publicado en Lancet (Independent UK Panel on Breast Cancer Screening, 2012), el cribado de mamografía conduce a un sobrediagnóstico en el 19% de las mujeres. En otras palabras, se dice que uno de cada cinco diagnósticos es un sobrediagnóstico. Están disponibles otras técnicas de obtención de imágenes, tales como obtención de imágenes por ecografía y resonancia magnética nuclear, pero generalmente no se utilizan para la detección, usándose en su lugar como un cribado adicional después de la mamografía.

Además, a pesar de la mejora en las terapias contra el cáncer de mama, se produce recidiva local, en la mama contralateral o a distancia (asimismo conocida como metástasis) en del 10 al 20% de los pacientes en los tres a diez años posteriores al tratamiento adyuvante inicial. Sin embargo, dicha recidiva a menudo se omite o se identifica como falso positivo por la mamografía, y se realizan biopsias innecesarias en pacientes sospechosos de recaída.

Por tanto, resulta crítico identificar biomarcadores fiables que permitan, sin recurrir de manera rutinaria a técnicas de obtención de imágenes o biopsias invasivas, no solo la detección temprana de un tumor de mama, sino asimismo la monitorización de la progresión del cáncer.

Junto con las técnicas de obtención de imágenes, se ha realizado una gran cantidad de trabajo examinando la expresión de genes o proteínas en el tejido tumoral mamario, pero el número de biomarcadores que podrían utilizarse potencialmente para detectar de manera fiable el cáncer de mama fue muy limitado, principalmente porque carecían de sensibilidad en el contexto clínico. En este sentido, los biomarcadores séricos antígeno prostático específico (PSA), CA 15-3 y antígeno carcinoembrionario (CEA), que han demostrado cierto valor en el diagnóstico y tratamiento de otros cánceres, no demostraron ser útiles en la detección y monitorización del cáncer de mama ya que carecían de la sensibilidad y especificidad deseadas. Se han descrito notablemente biomarcadores adicionales para la detección del cáncer de mama las solicitudes PCT n^{os} WO 2013/033609 y WO 2011/113047, y en las solicitudes US n^{os} 2007/037228 y 2011/085982. Sin embargo, ni OLFM4 ni neudesina se dan a conocer como biomarcadores de cáncer de mama en ninguno de estos documentos.

Por tanto, existe la necesidad urgente de identificar biomarcadores de cáncer de mama que sean fácilmente detectables, lo suficientemente sensibles para detectar la presencia de un tumor en pacientes con cáncer de mama, y lo suficientemente específicos para no detectar un tumor de este tipo en los que no presentan cáncer.

Las necesidades mencionadas anteriormente se abordan mediante la presente invención, con los resultados de una investigación realizada mediante mapeo comparativo de proteoma de tumores de mama y mediante validación de biomarcadores proteómicos secretados desregulados en una gran cohorte de pacientes con cáncer de mama notificados en la presente memoria. A diferencia de los biomarcadores genómicos, los biomarcadores proteómicos son de hecho particularmente ventajosos ya que reflejan mejor el microentorno tumoral y pueden experimentar modificaciones postraduccionales específicas del cáncer.

En particular, en el contexto de la presente invención se ha demostrado que la desregulación en el nivel de expresión de proteínas de olfactomedina-4 (OLFM4), neudesina (NENF) y/o desmoplaquina (DSP) se correlaciona con el cáncer de mama, y que tales biomarcadores son detectables en muestras de sangre de pacientes. Se ha descubierto notablemente que el nivel de expresión de olfactomedina-4 y neudesina es mayor en pacientes con cáncer de mama a lo largo de la progresión de la enfermedad, mientras que el nivel de expresión de desmoplaquina es más bajo en un estadio temprano y más alto en el caso de recidiva, en comparación con sujetos sanos.

La olfactomedina-4 es una proteína N-glicosilada secretada que pertenece a la familia de proteínas que contienen dominio de olfactomedina, que se caracteriza por un dominio N-terminal de dominio hélice-hélice y un dominio de la olfactomedina C-terminal bien conservado. La proteína OLFM4 se ha descrito en la bibliografía como mediadora de la adhesión celular a través de la unión a cadherinas y lectinas. (Liu *et al.*, 2006), y como implicada en la regulación de la apoptosis celular y en la proliferación de células cancerosas. (Zhang *et al.*, 2004; Kobayashi *et al.*, 2007). OLFM4 se ha descrito como biomarcador del cáncer de colon (documento WO 2012/117267). Además, se ha encontrado que la expresión génica de OLFM4 aumenta en diversos cánceres, incluyendo cánceres de mama y pulmón (Koshida *et al.*, 2007).

La neudesina, por otro lado, es una proteína extracelular de unión a grupo hemo que se ha descrito como que presenta actividad neurotrófica en las neuronas por medio de la proteína cinasa activada por mitógeno (MAPK) y fosfatidilinositol 3-cinasa (PI3K) (Kimura *et al.*, 2013). Se ha notificado que la neudesina está implicada en la tumorigénesis de la mama (Han *et al.*, 2012).

La desmoplaquina es un miembro fundador de la familia de plaquina, y se conoce como la principal proteína de placas de los desmosomas. (Leung *et al.*, 2002). Por tanto, está especializada en uniones de adhesión encontradas en diversos tejidos y desempeña un papel fundamental en el mantenimiento de la integridad del tejido epitelial. Recientemente, estudios sugirieron que los desmosomas participan en la regulación de la motilidad, el crecimiento, la diferenciación y la apoptosis celulares. (Allen *et al.*, 1996; Párido *et al.*, 2007; y Rickelt *et al.*, 2009). Hasta la fecha se ha informado de dos isoformas de desmoplaquina, desmoplaquina I (322 kDa) y desmoplaquina II (259 kDa), ambas codificadas por el gen de desmoplaquina en el cromosoma humano 6p24.3. Las proteínas desmoplaquina interaccionan con placoglobina (γ -catenina), placofilinas y filamentos intermedios, proporcionando el vínculo íntimo entre las cadherinas desmosómicas y el citoesqueleto (Junkman *et al.*, 2005; y Kowalczyk *et al.*, 1997).

Los biomarcadores anteriores pueden utilizarse en la presente memoria para detectar cáncer de mama, de una simple muestra de sangre, para monitorizar la progresión de la enfermedad, pero asimismo para desarrollar y adaptar un tratamiento para el cáncer de mama. Asimismo pueden utilizarse como dianas terapéuticas para diseñar nuevos fármacos.

Por tanto, basándose en los hallazgos dados a conocer en la presente memoria, por primera vez se proporciona un método de diagnóstico fiable y fácil de realizar para el cáncer de mama, que se basa en la determinación del nivel de expresión del/de los biomarcador(es) mencionados anteriormente. Asimismo se proporciona un método de cribado para identificar fármacos. Asimismo se proporcionan en la presente memoria kits y micromatrices de proteínas para llevar a cabo los métodos.

Sumario de la invención

La presente invención se define en las reivindicaciones 1 a 13. Cualquier otro aspecto o forma de realización expuestos en la presente memoria son únicamente informativos.

Descripción detallada de la invención

A menos que se indique lo contrario, los términos científicos y técnicos utilizados en relación con la presente invención presentarán los significados que entienden comúnmente los expertos habituales en la materia. Además, a menos que el contexto lo requiera, las nomenclaturas utilizadas en la presente memoria y las técnicas de biología molecular y cultivo celular son las bien conocidas y comúnmente utilizadas en la materia.

No obstante, con respecto a la utilización de diferentes términos a lo largo de la presente memoria descriptiva, las siguientes definiciones se aplican más particularmente.

Según los diferentes aspectos y formas de realización proporcionados en la presente memoria, el término “que comprende” o “que contiene” significa la inclusión del referente y no excluye la presencia de ningún otro elemento. En contraste con el término “que comprende”, el término “que consiste en” significa la inclusión exclusiva del referente y, por tanto, excluye la presencia de cualquier otro elemento.

El término “sujeto” o “paciente” se utiliza en la presente memoria para describir cualquier miembro del reino animal, preferentemente un ser humano, más preferentemente una mujer.

El término “diagnosticar” o “diagnóstico”, tal como se utiliza en el presente contexto, significa evaluar si un sujeto padece o no una enfermedad. Tal como apreciarán los expertos en la materia, una evaluación de este tipo, aunque se prefiere que lo sea, puede no ser correcta para el 100% de los sujetos investigados. Este término requiere, sin embargo, que una parte estadísticamente significativa de los sujetos pueda evaluarse correctamente y, por tanto, diagnosticarse. El experto en la materia puede determinar fácilmente si una parte es estadísticamente significativa utilizando diversas herramientas de evaluación estadística bien conocidas, tal

como la determinación de los intervalos de confianza, la determinación del valor de p, la prueba de la t de Student, la prueba de Mann-Whitney, etc. Los detalles de tales métodos estadísticos pueden encontrarse en Dowdy y Wearden (1983). Los métodos estadísticos pueden permitir notablemente la determinación de la sensibilidad y la especificidad de una prueba de diagnóstico. La sensibilidad de una prueba de diagnóstico puede definirse como la proporción de sujetos que padecen una enfermedad que presentará un resultado positivo, mientras que la especificidad de una prueba de diagnóstico puede definirse como la proporción de sujetos sin la enfermedad que presentarán un resultado negativo.

Mediante “cáncer de mama” se hace referencia en la presente memoria a un cáncer que se forma en los tejidos de la mama, tal como lo define el National Cancer Institute. Los tipos de cáncer de mama incluyen, sin limitación, carcinoma ductal, que comienza en el revestimiento de los conductos lácteos (tubos delgados que transportan la leche desde los lóbulos de la mama hasta el pezón); carcinoma lobular, que comienza en los lóbulos (glándulas lácteas) de la mama; y cáncer de mama invasivo (cáncer de mama que se ha diseminado desde donde comenzó en los conductos o lóbulos mamarios hasta el tejido normal circundante).

Mediante “cáncer de mama temprano”, se hace referencia en la presente memoria a un cáncer de mama que no se ha diseminado más allá de la mama o de los ganglios linfáticos axilares. Según la clasificación internacional de cáncer de mama TNM (tumor, ganglios, metástasis (*Tumor, Node, Metastasis*)), esto incluye carcinoma ductal o lobular *in situ* (pTis N0 M0) y cáncer de mama en estadio I (T1 N0 M0). Más particularmente, un cáncer de mama pTis N0 M0 se refiere a un cáncer de mama, en el que las células cancerosas solo pueden encontrarse dentro de los conductos o lóbulos mamarios (T0), sin que el tumor atraviese la membrana basal. Un cáncer de mama en estadio I (T1 N0 M0) se refiere a un cáncer de mama, en el que las células cancerosas se han infiltrado en el tejido que rodea los conductos y lóbulos mamarios, formando un tumor cuyo diámetro es inferior o igual a 2 cm. La abreviatura N0 significa que el cáncer no se ha diseminado a los ganglios linfáticos, mientras que la abreviatura M0 significa que no hay metástasis a distancia. Un cáncer de mama temprano generalmente se caracteriza por una tasa de supervivencia del 100%, en el plazo de cinco años desde el diagnóstico inicial.

Mediante “recidiva del cáncer”, “cáncer recidivante”, “recaída del cáncer” o “cáncer con recaída”, se hace referencia en la presente memoria, en el contexto de los posibles desenlaces clínicos del cáncer y tal como define el National Cancer Institute, a que el cáncer ha experimentado recaída (ha regresado), habitualmente después de un período de tiempo durante el cual no pudo detectarse el cáncer. Un cáncer recidivante puede referirse a un cáncer que regresa al mismo lugar que el tumor original (primario) o a otro lugar en el cuerpo (asimismo conocido como metástasis).

Una “muestra de líquido biológico” tal como se describe en la presente memoria puede ser cualquier muestra de líquido que puede aislarse de un sujeto, incluyendo, sin limitación, sangre o un componente fraccional de la misma (suero, plasma, extracto celular), linfa, líquido intersticial de tumor, saliva, moco, esputos, sudor u orina. Además, debe indicarse que, en el caso de una recidiva del cáncer local o distante, pueden aislarse células tumorales circulantes (CTC) de un líquido biológico tal como se definió anteriormente, preferentemente de sangre, mediante técnicas bien conocidas en la materia. Un ejemplo de una técnica que permite el aislamiento de células tumorales circulantes (CTC) es fraccionamiento de flujo de campo (DFF), tal como establecen Hou *et al.* (2013). En el contexto de la presente invención, la muestra de líquido biológico es preferentemente una muestra de sangre, tal como una muestra de suero o plasma, e incluso más preferentemente una muestra de suero.

El término “biomarcador” tal como se describe en la presente memoria se refiere a un polipéptido o proteína, fragmento del mismo o epítipo que está presente de manera diferencial en un sujeto en comparación con sujetos sanos, incluidos biomarcadores expresados y modificados (por ejemplo, glicosilados) de manera diferencial. Debe indicarse que el término “biomarcador” incluye biomarcadores solubles, es decir, biomarcadores que se escinden, secretan, liberan o eliminan de manera diferencial de una célula tumoral en un sujeto, y por tanto son detectables en un líquido biológico tal como se definió anteriormente.

En la siguiente tabla 1 se enumeran biomarcadores particularmente preferidos asociados con el cáncer de mama.

Tabla 1. Biomarcadores de cáncer de mama

Símbolo	Nombre completo	Número de registro de UniprotKB/Swiss-Prot (SEC ID número)
OLM4 o OLFM4	Olfactomedina-4 (OLFM4), nombre(s) alternativo(s): proteína antiapoptótica GW112, proteína de clon 1 estimulada por G-CSF (hGC-1) hOLF4	Q6UX06 (SEC ID NO:1)

Símbolo	Nombre completo	Número de registro de UniprotKB/Swiss-Prot (SEC ID número)
NENF	Neudesina, nombre(s) alternativos(s): proteína 2 relacionada con inmortalización celular, proteína secretada de factor neurotrófico derivado de neuronas de función desconocida (proteína SPUF)	Q9UMX5 (SEC ID NO:2)
DSP o DP	Desmoplaquina: Isoformas 1 y 2, nombre(s) alternativo(s): antígeno de pénfigo paraneoplásico de 250/210 kDA	P15924-1 (isoforma 1: SEC ID NO: 3) P15924-2 (isoforma 2: SEC ID NO: 4)

El término “nivel de expresión”, tal como se aplica a un biomarcador, se refiere en la presente memoria a la cantidad o el nivel de un biomarcador de interés expresado en una célula, tejido, líquido biológico u órgano(s). El término “nivel” tal como se utiliza en la presente memoria se refiere a una cantidad (por ejemplo cantidad relativa o concentración) de un biomarcador que puede detectarse o medirse en una muestra. Por ejemplo, el nivel puede ser una concentración tal como µg/l o una cantidad relativa en comparación con un nivel de expresión de referencia. El acto de “determinar el nivel de expresión” realmente de un biomarcador en una muestra biológica se refiere a al acto de detectar de manera activa si un biomarcador se expresa en dicha muestra o no, y notablemente permite detectar si la expresión del biomarcador está regulada por incremento, regulada por disminución o sustancialmente sin cambios en comparación con un nivel de expresión de referencia. Un “nivel de expresión desregulado” de un biomarcador dado, tal como se describe en la presente memoria, es un nivel de expresión regulado por disminución o regulado por incremento en comparación con un nivel de expresión de referencia.

Mediante “nivel de expresión de referencia” o “nivel de expresión de control” de un biomarcador, se hace referencia a un nivel de expresión predeterminado de dicho biomarcador, que puede utilizarse como referencia en cualquier método descrito en la presente memoria. Por ejemplo, un nivel de expresión de referencia puede ser el nivel de expresión de un biomarcador en una muestra biológica de un sujeto sano, o el promedio o la mediana, preferentemente la mediana, del nivel de expresión en una muestra biológica de una población de sujetos sanos.

Se proporcionan definiciones adicionales a lo largo de toda la memoria descriptiva.

La presente invención puede ponerse más claramente de manifiesto haciendo referencia a la siguiente descripción detallada, incluyendo las formas de realización preferidas de la invención, y los ejemplos incluidos en la presente memoria.

En el contexto de la presente invención se ha demostrado que el nivel de expresión de olfactomedina-4 (OLFM4), neudesina (NENF) y/o desmoplaquina (DSP) circulantes en la sangre está desregulado en sujetos que padecen cáncer de mama. Se propone por tanto un método para diagnosticar de manera rápida y fácil cáncer de mama en un sujeto basándose en el descubrimiento anterior, determinando el nivel de expresión de dicho(s) biomarcador(s), a partir de una mera muestra de líquido biológico tal como sangre. Tal método de diagnóstico permite de ese modo evitar la utilización de métodos de diagnóstico convencionales, engorrosos o incluso invasivos tales como biopsia, obtención de imágenes por resonancia magnética (MRI), tomografía computarizada (CT) o exploración por CT potenciada por contraste intratecal.

En un primer aspecto se da a conocer un método *in vitro* para diagnosticar un cáncer de mama en un sujeto, que comprende las etapas de:

a) determinar a partir de una muestra de líquido biológico de un sujeto el nivel de expresión de proteínas de por lo menos un biomarcador seleccionado de entre el grupo que consiste en olfactomedina-4, neudesina, desmoplaquina, y cualquier combinación de las mismas; y

b) comparar dicho nivel de expresión con un nivel de expresión de referencia de dicho biomarcador.

El método anterior puede comprender opcionalmente además la etapa c) de determinar si dicho sujeto padece cáncer de mama, basándose en la comparación de la etapa b).

Cada uno de los biomarcadores anteriores es suficiente para realizar un diagnóstico tal como se describe en la presente memoria. No obstante, el experto en la materia comprenderá fácilmente que los biomarcadores anteriores pueden combinarse como un panel de biomarcadores, contribuyendo cada uno de los cuales al diagnóstico final.

Como ejemplo, un nivel de expresión de proteínas de dicho por lo menos un biomarcador desregulado en comparación con un nivel de expresión de referencia de dicho biomarcador obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, es indicativo de que dicho sujeto padece cáncer de mama.

Preferentemente, dicho nivel de expresión de proteínas de:

- olfactomedina-4 es superior a dicho nivel de expresión de referencia; y/o
- neudesina es superior a dicho nivel de expresión de referencia; y/o
- desmoplaquina es inferior o superior a dicho nivel de expresión de referencia;

en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama.

Es decir, un nivel de expresión de proteínas de olfactomedina-4 en etapa a) superior al nivel de expresión de referencia de olfactomedina-4 obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, es indicativo de que el sujeto sometido a prueba padece cáncer de mama.

De manera similar, un nivel de expresión de proteínas de neudesina en etapa a) superior al nivel de expresión de referencia de neudesina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, es indicativo de que el sujeto sometido a prueba padece cáncer de mama.

Esto significa asimismo que un nivel de expresión de proteínas de desmoplaquina en etapa a) inferior al nivel de expresión de referencia de desmoplaquina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, es indicativo de que el sujeto sometido a prueba padece cáncer de mama.

Esto asimismo significa que un nivel de expresión de proteínas de desmoplaquina en la etapa a) superior al nivel de expresión de referencia de desmoplaquina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, es indicativo de que el sujeto sometido a prueba padece cáncer de mama.

Se entenderá además que el presente método, así como otros métodos descritos en la presente memoria, comprenden la utilización de cualquier combinación de los biomarcadores anteriores.

El nivel de expresión de proteínas de una cualquiera de olfactomedina-4, neudesina y desmoplaquina, o de cualquier combinación de las mismas, puede indicar además el estadio del cáncer de mama.

Por lo tanto, como ejemplo, dicho nivel de expresión de olfactomedina-4 superior a dicho nivel de expresión de referencia es indicativo de que el cáncer de mama es un cáncer de mama temprano.

En otro ejemplo, dicho nivel de expresión de neudesina superior a dicho nivel de expresión de referencia es indicativo de que el cáncer de mama es un cáncer de mama temprano.

Aún, en otro ejemplo, dicho nivel de expresión de desmoplaquina inferior a dicho nivel de expresión de referencia es indicativo de que el cáncer de mama es un cáncer de mama temprano.

Tal prueba de diagnóstico de cáncer de mama temprano es particularmente útil para pacientes en riesgo (por ejemplo que presentan una historia familiar de cáncer de mama), y para pacientes para los que no pueden detectarse de manera precisa tumores de mama pequeños (por ejemplo tamaño tumoral por debajo de 1 cm) mediante métodos de diagnóstico convencionales tales como ultrasonidos.

Todavía, en otro ejemplo, dicho nivel de expresión de desmoplaquina superior a dicho nivel de expresión de referencia es indicativo de que el cáncer de mama es un cáncer de mama recidivante. Tal prueba de diagnóstico de un cáncer de mama recidivante es particularmente útil para la monitorización de un paciente que padece previamente cáncer de mama. Notablemente, si se realiza de manera temprana, tal prueba de diagnóstico puede ayudar a mejorar el pronóstico y la supervivencia del paciente.

Mediante superior a un nivel de expresión de referencia, se hace referencia a que la razón entre el nivel de expresión de dicho biomarcador y el nivel de expresión de referencia está por encima de 1.

Mediante inferior a un nivel de expresión de referencia, se hace referencia a que la razón entre el nivel de expresión de dicho biomarcador y el nivel de expresión de referencia está por debajo de 1.

Alternativamente, dicho nivel de expresión puede indicarse como la concentración de biomarcador en el líquido biológico sometido a prueba.

Por lo tanto, dicho nivel de expresión de proteínas de olfactomedina-4 es preferentemente superior a 30, 31, 32, 34, 35, 36, 37, 38, 39, 40, 50, 60, 70, 80, 90, 100, 150, 200 ng/ml, más preferentemente superior a 30, 31, 32, 34, 35, 36, 37, 38, 39, 40 ng/ml, todavía más preferentemente superior a 31 ng/ml, en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama, preferentemente cáncer de mama temprano. De hecho, tal como se ilustra en los resultados experimentales, en el contexto de la presente invención se ha demostrado en una primera cohorte de 335 pacientes con cáncer de mama y 65 sujetos sanos que la sensibilidad de una prueba de diagnóstico de cáncer de mama basada en olfactomedina-4 es del 67%, mientras que la especificidad de tal

prueba es del 88%, para una concentración sérica de olfactomedina-4 por encima de 40 ng/ml. Se demuestra además en una segunda cohorte de 766 pacientes con cáncer de mama y 195 sujetos sanos que la sensibilidad de una prueba de diagnóstico de cáncer de mama basada en olfactomedina-4 oscila entre el 64 y el 78%, con una especificidad de aproximadamente el 80-90% para una concentración sérica de olfactomedina-4 por encima de 31 ng/ml. Se observaron valores similares en pacientes con cáncer de mama temprano.

Todavía, dicho nivel de expresión de proteínas de neudesina es preferentemente superior a 15, 16, 17, 18, 19, 20, 25, 30, 35, 40 ng/ml, más preferentemente superior a 15, 16, 17, 18, 19, 20 ng/ml, todavía más preferentemente superior a 16 ng/ml, en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama, preferentemente cáncer de mama temprano. De hecho, tal como se ilustra en los resultados experimentales, en el contexto de la presente invención se ha demostrado en una primera cohorte de 335 pacientes con cáncer de mama y 65 sujetos sanos, que la sensibilidad de una prueba de diagnóstico de cáncer de mama basada en neudesina es del 47%, mientras que la especificidad de tal prueba alcanzó el 91%, para una concentración sérica de neudesina por encima de 20 ng/ml. Se demuestra además en una segunda cohorte de 766 pacientes con cáncer de mama y 195 sujetos sanos que la sensibilidad de una prueba de diagnóstico de cáncer de mama basada en neudesina oscila entre el 52 y el 60%, con una especificidad que alcanza aproximadamente el 75-80% para una concentración sérica de neudesina por encima de 16 ng/ml. Se observaron valores similares en pacientes con cáncer de mama temprano.

Todavía, dicho nivel de expresión de proteínas de desmoplaquina es preferentemente inferior a 600, 500, 400, 300 pg/ml, más preferentemente inferior a 600 pg/ml, en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama temprano.

Aún, todavía, dicho nivel de expresión de proteínas de desmoplaquina es preferentemente superior a 1800, 1900, 2000 pg/ml, más preferentemente superior a 1800 pg/ml, en la muestra de líquido biológico de dicho sujeto que padece un cáncer de mama recidivante.

Tal como se menciona anteriormente, los biomarcadores anteriores pueden combinarse como un panel de biomarcadores, contribuyendo cada uno de los cuales al diagnóstico final. De hecho, en el contexto de la presente invención se ha demostrado que tal combinación aumenta la sensibilidad y/o la especificidad de la prueba de diagnóstico.

Por lo tanto, en un ejemplo, el nivel de expresión de proteínas de por lo menos dos, preferentemente tres, de dichos biomarcadores se determina en la etapa a).

Preferentemente, el nivel de expresión de proteínas de olfactomedina-4 y neudesina se determina en la etapa a) del método anterior. La combinación de estos biomarcadores es particularmente útil para diagnosticar cáncer de mama, tal como para diagnosticar un cáncer de mama temprano. Es decir, en una forma de realización preferida, dicho nivel de expresión de proteínas de:

- olfactomedina-4 es superior a dicho nivel de expresión de referencia, y
- neudesina es superior a dicho nivel de expresión de referencia, en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama, tal como un cáncer de mama temprano.

En otras palabras, un nivel de expresión de proteínas de:

- olfactomedina-4 superior al nivel de expresión de referencia de olfactomedina-4 obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, y
- neudesina superior al nivel de expresión de referencia de neudesina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano,

es indicativo de que el sujeto sometido a prueba padece cáncer de mama, tal como un cáncer de mama temprano.

De hecho, tal como se ilustra en los resultados experimentales, en el contexto de la presente invención se ha demostrado en una primera cohorte de 335 pacientes con cáncer de mama y 65 sujetos sanos que la sensibilidad de una prueba de diagnóstico de cáncer de mama basada en la combinación de olfactomedina-4 y neudesina es de 74%, mientras que la especificidad de tal prueba es de 78%, para una concentración sérica de esos biomarcadores por encima de 44 ng/ml. Se demuestra además en una segunda cohorte de 766 pacientes con cáncer de mama y 195 sujetos sanos que la sensibilidad de una prueba de diagnóstico de cáncer de mama basada en la combinación de olfactomedina-4 y neudesina oscila entre 75 y 85%, con una especificidad que alcanza e 87% para una concentración sérica de esos biomarcadores por encima de 38 ng/ml. Se observaron valores similares en pacientes con cáncer de mama temprano. La combinación de olfactomedina-4 y neudesina aumenta por tanto los valores de sensibilidad y/o de la prueba de diagnóstico de la invención, en comparación

con una prueba basada sólo en uno de estos biomarcadores.

Alternativamente, el nivel de expresión de proteínas de olfactomedina-4 y desmoplaquina se determina preferentemente en la etapa a) del método anterior. La combinación de estos biomarcadores es particularmente útil para diagnosticar cáncer de mama, y más particularmente un cáncer de mama temprano. Es decir, en un ejemplo, dicho nivel de expresión de proteínas de:

- olfactomedina-4 es superior a dicho nivel de expresión de referencia, y
- desmoplaquina es inferior a dicho nivel de expresión de referencia,

en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama, y preferentemente un cáncer de mama temprano. En otras palabras, un nivel de expresión de proteínas de:

- olfactomedina-4 superior al nivel de expresión de referencia de olfactomedina-4 obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, y
- desmoplaquina inferior al nivel de expresión de referencia de desmoplaquina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano,

es indicativo de que el sujeto sometido a prueba padece cáncer de mama, y preferentemente un cáncer de mama temprano.

De hecho, tal como se ilustra en los resultados experimentales, en el contexto de la presente invención se ha demostrado que la sensibilidad de una prueba de diagnóstico de cáncer de mama temprano basada en la combinación de olfactomedina-4 y desmoplaquina es de 87%, mientras que la especificidad de tal prueba es de 84%.

Alternativamente, el nivel de expresión de proteínas de neudesina y desmoplaquina se determina preferentemente en la etapa a) del método anterior. La combinación de estos biomarcadores es particularmente útil para diagnosticar cáncer de mama, y más particularmente un cáncer de mama temprano. Es decir, en un ejemplo, dicho nivel de expresión de proteínas de:

- neudesina es superior a dicho nivel de expresión de referencia, y
- desmoplaquina es inferior a dicho nivel de expresión de referencia,

en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama, y preferentemente un cáncer de mama temprano.

Es decir, un nivel de expresión de proteínas de:

- neudesina superior al nivel de expresión de referencia de neudesina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, y
- desmoplaquina inferior al nivel de expresión de referencia de desmoplaquina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano,

es indicativo de que el sujeto sometido a prueba padece cáncer de mama, y preferentemente un cáncer de mama temprano.

Todavía preferentemente, el nivel de expresión de proteínas de olfactomedina-4, neudesina y desmoplaquina se determina en la etapa a) del método anterior. La combinación de estos tres biomarcadores es particularmente útil para diagnosticar cáncer de mama, más particularmente un cáncer de mama temprano. Es decir, en un ejemplo particular, dicho nivel de expresión de proteínas de:

- olfactomedina-4 es superior a dicho nivel de expresión de referencia,
- neudesina es superior a dicho nivel de expresión de referencia, y
- desmoplaquina es inferior a dicho nivel de expresión de referencia, en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama, y preferentemente un cáncer de mama temprano.

En otras palabras, un nivel de expresión de proteínas de:

- olfactomedina-4 superior al nivel de expresión de referencia de olfactomedina-4 obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano,
- neudesina superior al nivel de expresión de referencia de neudesina obtenido a partir de una muestra de

líquido biológico de por lo menos un sujeto sano, y

- desmoplaquina inferior al nivel de expresión de referencia de desmoplaquina obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, es indicativo de que el sujeto sometido a prueba padece un cáncer de mama, y preferentemente un cáncer de mama temprano.

Debe apreciarse además que la información obtenida utilizando los métodos descritos en la presente memoria pueden utilizarse en combinación con otra información, tal como, pero sin limitarse a, los niveles de expresión de biomarcadores adicionales que pueden ser biomarcadores convencionales, parámetros químicos clínicos, parámetros histopatológicos, o la edad, el género y/o el peso del sujeto.

Por lo tanto, en una forma de realización adicional preferida, el método de diagnóstico *in vitro* de la invención comprende además la etapa de determinar el nivel de expresión de proteínas de por lo menos un biomarcador convencional asociado con cáncer de mama, tal como receptor de estrógenos (ER), receptor de progesterona (PR) o receptor 2 de factor de crecimiento epidérmico humano (HER2).

Tal como se menciona anteriormente, en el presente contexto, el nivel de expresión se mide al nivel de proteína. Se conocen bien en la materia métodos para medir los niveles de expresión de proteínas y los revisan de manera notable Reeves *et al.* (2000) y Schena (2005). Esos métodos implican generalmente poner en contacto una muestra biológica de interés con uno o más reactivos detectables que es o son adecuados para medir el nivel de expresión de proteínas, tal como un anticuerpo, y posteriormente determinar el nivel de expresión de proteínas basándose en el nivel de reactivo detectado, preferentemente tras la normalización. Los ejemplos de métodos que implican generalmente la utilización de un anticuerpo incluyen, sin limitación, transferencia de tipo Western, inmunotransferencia, ensayo de inmuoadsorción ligado a enzimas (ELISA), Immunospot ligado a enzimas (ELISPOT), radioinmunoensayo (RIA), inmunohistoquímica e inmunoprecipitación. Pueden utilizarse otros métodos adecuados para medir un nivel de expresión de proteínas, que no implican necesariamente la utilización de un anticuerpo, incluyendo, sin limitación, clasificación celular activada por fluorescencia (FACS), microscopía tal como microscopía de fuerza atómica, citometría de flujo, microcitometría, ensayo de unión a proteínas, ensayo de unión a ligandos, micromatriz, electroforesis en gel de poliácridamida tal como SDS-PAGE, resonancia de plasmón superficial (SPR), transferencia de energía por resonancia de Förster (FRET), transferencia de energía por resonancia de bioluminiscencia (BRET), quimioluminiscencia, polarización de fluorescencia, fosforescencia, espectrometría de masas tal como cromatografía de líquidos con espectrometría de masas (CL-EM) o cromatografía de líquidos/espectrometría de masas/espectrometría de masas (CL-EM-EM), ionización/desorción por láser asistida por matriz-tiempo de vuelo (MALDI-TOF), desorción/ionización por láser potenciada por superficie-tiempo de vuelo (SELDI-TOF) y obtención de imágenes por resonancia magnética (MRI).

Según el aspecto diferente, la etapa de determinar el nivel de expresión de un biomarcador de interés comprende además preferentemente una subetapa de normalizar el nivel de expresión de dicho biomarcador. El método para normalizar el nivel de expresión puede seleccionarse basándose en el método utilizado para medir el nivel de expresión. Por ejemplo, si se realiza una inmunotransferencia de tipo Western, el nivel de expresión de un biomarcador de interés en una muestra biológica puede normalizarse evaluando en paralelo en dicha muestra el nivel de expresión de una proteína que se expresa habitualmente de manera constitutiva en cualquier célula de un organismo vivo, preferentemente al mismo nivel de expresión ya sea la célula sana o no (por ejemplo cancerosa o no). Un ejemplo de proteína expresada de manera constitutiva es una proteína de mantenimiento, que puede seleccionarse, sin limitación, de actina, beta-tubulina y gliceraldehído-3-fosfato deshidrogenasa (GAPDH), por nombrar unas pocas. Alternativamente, si se realiza un ELISA, que implica por ejemplo un método de detección colorimétrica, el nivel de expresión de proteínas puede normalizarse mediante el número de células total. Aún, todavía alternativamente, si se realiza una micromatriz, el nivel de expresión de proteínas puede normalizarse, por ejemplo, mediante regresión de Loess. Para una revisión detallada de métodos de normalización del nivel de expresión de proteínas en una micromatriz de anticuerpos, un experto en la materia puede hacer referencia a Hamelinck *et al.* (2005).

Todos estos métodos para medir y normalizar nivel de expresión de proteínas los conoce bien el experto, y por tanto no es necesario que se detallen adicionalmente en la presente memoria. Si el experto desea utilizar cualquiera de los métodos anteriores que implican la utilización de un anticuerpo para medir un nivel de expresión de proteínas biomarcadoras, puede utilizarse cualquier anticuerpo comercial apropiado específico para dicho biomarcador. Alternativamente, basándose en el conocimiento de la secuencia de aminoácidos de un biomarcador de interés, al experto le resulta fácil diseñar reactivo(s) adecuado(s) para medir el nivel de expresión en cualquier muestra biológica. Por ejemplo, puede prepararse un anticuerpo dirigido contra un biomarcador específico mediante cualquier método convencional, por ejemplo inmunizando un animal, tal como un ratón, con una forma inmunógena de dicho biomarcador que provoca una respuesta de anticuerpos en dicho animal. En la bibliografía se describen bien métodos para producir anticuerpos monoclonales y policlonales (ver particularmente Kohler y Milstein, 1975; Kozbor *et al.*, 1983; Roder *et al.*, 1986; y Huse *et al.*, 1986), y por tanto no es necesario que se detallen adicionalmente en la presente memoria.

Tal como se menciona anteriormente, la comparación de un nivel de expresión determinado o sometido a prueba con un nivel de expresión de referencia puede hacerse simplemente calculando la razón entre el nivel de expresión de un biomarcador de interés en la muestra biológica sometida a prueba y en por lo menos una muestra de referencia, preferentemente tras la normalización tal como se describió anteriormente. Por lo tanto, una razón por encima de 1 es indicativa de que el biomarcador se sobreexpresa, mientras que una razón por debajo de 1 es indicativa de que el biomarcador se subexpresa (es decir, regulado por disminución).

Los biomarcadores dados a conocer en la presente memoria podrían utilizarse asimismo para determinar si un paciente responderá o no a una terapia contra el cáncer. La asociación de la respuesta de un paciente al tratamiento con tal(es) biomarcador(es) puede de hecho dilucidar nuevas oportunidades para el tratamiento en pacientes que no responden o indicar un tratamiento sobre otras opciones de tratamiento.

Por tanto, se describe además en la presente memoria un método *in vitro* para determinar un fenotipo que responde o no responde al fármaco en un sujeto que padece cáncer de mama, que comprende las etapas de:

- a) determinar a partir de una muestra de líquido biológico de dicho sujeto el nivel de expresión de proteínas de por lo menos un biomarcador seleccionado de entre el grupo que consiste en olfactomedina-4, neudesina, desmoplaquina y cualquier combinación de las mismas;
- b) comparar el nivel de expresión de proteínas en etapa a) con un nivel de expresión de referencia de dicho biomarcador; y
- c) determinar a partir de dicha comparación el fenotipo que responde o no responde al fármaco.

Tal como se describe en la presente memoria, un “fenotipo que responde al fármaco” se refiere a un estado de respuesta de un sujeto a la administración de un fármaco. Un “estado de respuesta” significa que dicho sujeto responde al tratamiento, es decir que dicho tratamiento es eficaz en dicho sujeto. Un fenotipo que responde se caracteriza por tanto por una mejora en los signos clínicos, es decir en el presente contexto, un fenotipo que responde se caracteriza por ejemplo por una regresión o desaparición de las células de cáncer de mama células y metástasis de las mismas, si existen. En cambio, un “fenotipo que no responde al fármaco” se refiere a la ausencia en dicho sujeto de un estado de respuesta, lo que significa que dicho sujeto es resistente al tratamiento.

El nivel de expresión de proteínas de los biomarcadores mencionados anteriormente en un sujeto que padece cáncer de mama es tal como se describió anteriormente.

Los biomarcadores dados a conocer en la presente memoria podrían utilizarse asimismo para diseñar o adaptar un tratamiento del cáncer de mama. En particular, tal tratamiento puede diseñarse o adaptarse una vez que al sujeto se le ha diagnosticado que padece cáncer de mama, según el método proporcionado en la presente memoria.

Por lo tanto, se describe además en la presente memoria un método para diseñar o adaptar un régimen de tratamiento para un sujeto que padece cáncer de mama, que comprende las etapas de:

- a) determinar a partir de una muestra biológica de dicho sujeto un fenotipo que responde o que no responde al fármaco, según los métodos *in vitro* descritos anteriormente; y
- b) diseñar o adaptar un régimen de tratamiento para dicho sujeto basándose en dicho fenotipo que responde o no responde.

El presente método es particularmente útil para ofrecer una terapia adaptada a cada paciente afectado por cáncer de mama.

El término “régimen de tratamiento” se refiere en la presente memoria a un plan de tratamiento que especifica el de tratamiento (es decir, tipo de fármaco o combinación de fármacos, y modo de administración de dicho(s) fármaco(s)), dosificación, programa y/o duración de un tratamiento proporcionado a un sujeto que lo necesita. Una dosificación, programa y/o duración de tratamiento puede variar dependiendo de la progresión de la enfermedad y el tipo seleccionado de tratamiento. En este sentido, además de los fármacos que pueden identificarse según el método de cribado descrito en la presente memoria, los agentes terapéuticos que pueden utilizarse en el régimen de tratamiento descrito en la presente memoria incluyen, sin limitación, agentes quimioterápicos; agentes terapéuticos hormonales tales como tamoxifeno o inhibidores de aromatasa (por ejemplo raloxifeno, toremifeno, fulvestrant, anastrozol, exemestano, letrozol); inhibidores de receptor 2 de factor de crecimiento epidérmico humano (HER2) tales como trastuzumab (Herceptin), pertuzumab o lapatinib; inhibidores de receptor de factor de crecimiento endotelial vascular (VEGFR) tales como bevacizumab; inhibidores de receptor de factor de crecimiento epidérmico (EGFR) tales como cetuximab y panitumumab; y cualquier combinación de los mismos.

Los agentes quimioterápicos convencionales para tratar cáncer de mama incluyen, sin limitación, agentes a base de platino tales como oxaliplatino, cisplatino, carboplatino, espiroplatino, iroplatino y satraplatino; agentes alquilantes tales como ciclofosfamida, ifosfamida, clorambucilo, busulfán, melfalán, mecloretamina, uramustina, tiotepa y nitrosoureas; antimetabolitos tales como 5-fluorouracilo, azatioprina, 6-mercaptopurina, metotrexato, leucovorina, capecitabina, citarabina, floxuridina, fludarabina, gemcitabina, pemetrexed o raltitrexed; alcaloides vegetales tales como vincristina, vinblastina, vinorelbina, vindesina, podofilotoxina, o taxanos tales como paclitaxel y docetaxel; inhibidores de topoisomerasa tales como irinotecán, topotecán, amsacrina, etopósido (VP16), fosfato de etopósido o tenipósido; antibióticos antitumorales tales como antraciclinas (por ejemplo doxorubicina, daunorubicina, epirubicina, mitoxantrona), actinomicina, bleomicina, mitomicina o plicamicina; y cualquier combinación de los mismos.

En el método anterior, el régimen de tratamiento que se diseña y se adapta y opcionalmente se administra al sujeto depende del fenotipo que responde o que no responde.

En particular, puede seleccionarse un régimen de tratamiento por primera vez, continuarse, ajustarse o detenerse basándose en dicho fenotipo. Por ejemplo, puede ajustarse un régimen de tratamiento aumentando la dosis que va a administrarse, o interrumpirse y cambiarse a un régimen de tratamiento alternativo, si el sujeto no responde. Todavía, alternativamente, puede seleccionarse un régimen de tratamiento por primera vez o continuarse si un sujeto responde. Un experto en la materia no obstante diseñaría o ajustaría fácilmente el tipo de tratamiento con la dosificación, el programa y la duración del tratamiento, dependiendo del fenotipo del sujeto.

Además, basándose en dicho fenotipo, el régimen de tratamiento seleccionado puede ser uno agresivo que se espera que dé como resultado el mejor desenlace clínico (por ejemplo, regresión y/o desaparición del cáncer de mama) y que puede estar asociado con alguna molestia para el sujeto o efectos secundarios adversos (por ejemplo, daño para las células o el tejido sano), o uno más moderado que puede sólo ralentizar la progresión de la enfermedad. Un ejemplo de régimen de tratamiento agresivo incluye un régimen de tratamiento tal como se describió anteriormente combinado con intervención quirúrgica para eliminar células tumorales, tejido u órganos y/o una exposición a radioterapia. Un régimen de tratamiento agresivo puede incluir asimismo una mayor dosificación del/de los agente(s) terapéutico(s), una administración más frecuente de dicho(s) agente(s) y/o una mayor duración del tratamiento.

Por tanto, una vez que se ha determinado un régimen de tratamiento según las enseñanzas descritas en la presente memoria, el sujeto puede recibir el tratamiento apropiado. Por tanto, asimismo se describe un método en la presente memoria para tratar cáncer de mama en un sujeto que lo necesita, que comprende las etapas de:

- a) determinar a partir de una muestra biológica de dicho sujeto un fenotipo que responde o que no responde al fármaco, según el método descrito anteriormente; y
- b) administrar a dicho sujeto dicho fármaco si el fenotipo es un fenotipo que responde.

El término “administrar” tal como se utiliza en la presente memoria significa que el/los fármaco(s) de interés se administra(n) se dispensa(n) a un sujeto por vía oral, o por vía parenteral tal como mediante inyección subcutánea, intravenosa, intramuscular, intratecal o intraperitoneal.

En otro aspecto, los biomarcadores dados a conocer en la presente memoria pueden utilizarse para fines de cribado de fármacos. En particular, pueden proporcionarse nuevos ensayos de fármacos que identifican productos terapéuticos que interfieren eficazmente con la proliferación de células de cáncer de mama que expresan de manera aberrante esos biomarcadores. El tratamiento de cáncer de mama actual se basa principalmente en quimioterapia y/o fármacos antiangiogénicos, que pueden combinarse, si es necesario, con cirugía.

Por lo tanto, en el presente aspecto, se proporciona en la presente memoria un método de cribado para identificar un fármaco o combinación de fármacos adecuados para tratar cáncer de mama, que comprende las etapas de:

- a) poner en contacto células de cáncer de mama aisladas o una estirpe celular que presenta fenotipo de cáncer de mama con un fármaco candidato o combinación de fármacos candidatos;
- b) determinar, a partir de dichas células o estirpe celular en contacto con dicho fármaco o combinación de fármacos, el nivel de expresión de proteínas de por lo menos un biomarcador seleccionado de entre el grupo que consiste en olfactomedina-4, neudesina, desmoplaquina, y cualquier combinación de las mismas; y
- c) comparar el nivel de expresión de proteínas de dicho biomarcador en la etapa b) con su nivel de expresión en ausencia de dicho fármaco o combinación de fármacos.

Mediante “fármaco” o “agente” se hace referencia en la presente memoria a un compuesto tal como una molécula química o biológica que puede administrarse o someterse a prueba. Un compuesto químico puede ser cualquier composición tal como inorgánica u orgánica. Una molécula biológica puede ser una molécula de cualquier origen biológico que puede encontrarse o producirse mediante, por lo menos en parte, una célula, tal como, sin limitación, péptidos o proteínas tales como anticuerpos o aficuerpos, lípidos, ácidos nucleicos tales como iARN o aptámeros, hidratos de carbono y cualquier combinación de los mismos.

Mediante “fármaco adecuado para tratar cáncer de mama” se hace referencia en la presente memoria a un fármaco que puede ralentizar o detener el crecimiento de células de cáncer de mama y metástasis de las mismas, si existen, o bien destruyendo dichas células o bien ralentizando o deteniendo su división incontrolada.

Además, se entenderá que mediante “células de cáncer de mama o estirpe celular” tal como se describe en la presente memoria se hace referencia preferentemente a células de cáncer de mama o una estirpe celular en las que el nivel de expresión de proteínas de olfactomedina-4, neudesina y/o desmoplaquina está desregulado en comparación con un nivel de expresión de referencia de dicho(s) biomarcador(es) en las células de mama de por lo menos un sujeto sano. Preferentemente, las células o la estirpe celular utilizadas en el presente método de cribado son células de cáncer de mama aisladas de un sujeto al que se le diagnostica que padece cáncer de mama según el método proporcionado en la presente memoria.

El método de cribado descrito anteriormente es preferentemente un método de cribado *in vitro*. Por ejemplo, las células o la estirpe celular utilizadas en el presente método pueden cultivarse en un sistema de cultivo tridimensional (3D), para imitar el microentorno del tumor de mama. Para hacer eso, dichas células pueden incrustarse en una matriz extracelular (ECM) tal como describen Weigelt *et al.* (2008), Kenny *et al.* (2007) y Li *et al.* (2010).

Con el fin de evaluar la eficacia del agente anticancerígeno candidato, dichas células o estirpe celular, como alternativa o como prueba de validación, pueden injertarse en un animal no humano, tal como un ratón. Si tal xenoinjerto se lleva a cabo, el método de cribado descrito anteriormente comprende además preferentemente la etapa de matar a dicho animal no humano.

En un ejemplo del método anterior, un nivel de expresión de proteínas de olfactomedina-4 en la etapa b) inferior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar cáncer de mama.

En un ejemplo, un nivel de expresión de proteínas de neudesina en la etapa b) inferior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar cáncer de mama.

En un ejemplo adicional, un nivel de expresión de proteínas de desmoplaquina en la etapa b) superior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar cáncer de mama temprano.

Aún, en otro ejemplo, un nivel de expresión de proteínas de desmoplaquina en la etapa b) inferior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar un cáncer de mama recidivante.

Un experto en la materia entendería fácilmente a partir de los datos proporcionados en la presente memoria que los biomarcadores mencionados anteriormente pueden combinarse para ayudar en la identificación de un fármaco o combinación de fármacos. Está dentro de la experiencia del experto en la materia seleccionar el biomarcador apropiado que va a combinarse.

Preferentemente, un nivel de expresión de proteínas de olfactomedina-4 y neudesina en la etapa b) inferior al nivel de expresión de proteínas de dichos biomarcadores en ausencia de dicho fármaco o combinación de fármacos es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar cáncer de mama tal como un cáncer de mama temprano.

Preferentemente, un nivel de expresión de proteínas de:

- desmoplaquina en la etapa b) superior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos, y
- olfactomedina-4 en la etapa b) inferior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos,

es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar cáncer de mama, y preferentemente un cáncer de mama temprano.

5 Preferentemente, un nivel de expresión de proteínas de:

- desmoplaquina en la etapa b) superior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos, y

10 • neudesina en la etapa b) inferior al nivel de expresión de proteínas de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos,

es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar cáncer de mama, y preferentemente un cáncer de mama temprano.

15 Todavía preferentemente, un nivel de expresión de proteínas de:

- olfactomedina-4 y neudesina en la etapa b) inferior al nivel de expresión de proteínas de dichos biomarcadores en ausencia de dicho fármaco o combinación de fármacos, y

20 • desmoplaquina en la etapa b) superior al nivel de expresión de dicho biomarcador en ausencia de dicho fármaco o combinación de fármacos,

es indicativo de que dicho fármaco o combinación de fármacos es adecuado para tratar cáncer de mama, y preferentemente un cáncer de mama temprano.

En otro aspecto, se proporcionan kits que pueden emplearse en los métodos descritos en la presente memoria. En este sentido, se proporciona en la presente memoria un kit para su utilización en cualquier método descrito anteriormente, comprendiendo o consistiendo dicho kit en:

30 a) por lo menos un reactivo apto para determinar específicamente el nivel de expresión de proteínas de por lo menos un biomarcador seleccionado de entre el grupo que consiste en olfactomedina-4, neudesina, desmoplaquina, y cualquier combinación de las mismas; y

35 b) instrucciones para realizar dicho método.

Tal como se utiliza en la presente memoria, el término “instrucciones” se refiere a una publicación, una grabación, un diagrama o cualesquier otros medios que puedan utilizarse para comunicar cómo realizar un método descrito en la presente memoria. Dichas instrucciones, por ejemplo, pueden fijarse a un recipiente que contiene dicho kit. Preferentemente, las instrucciones para utilizar dicho kit incluyen un nivel de expresión de referencia de dicho(s) biomarcador(s).

El término “reactivo apto para determinar específicamente el nivel de expresión de proteínas [de un biomarcador dado]” designa un reactivo o conjunto de reactivos que reconocen específicamente dicho biomarcador y permiten la cuantificación de su nivel de expresión de proteínas. Estos reactivos pueden ser por ejemplo anticuerpos, aptámeros o aficuerpos que reconocen específicamente un biomarcador. En el presente contexto, se dice que un reactivo de este tipo es “específico” para su diana (es decir biomarcador) o “reconoce específicamente” su diana si 1) presenta un nivel umbral de actividad de unión, y/o 2) no reacciona de manera cruzada significativamente con moléculas diana que se sabe que están relacionadas con el biomarcador de interés. La afinidad de unión de tal reactivo puede determinarla fácilmente un experto en la materia, por ejemplo, mediante análisis de Scatchard. Asimismo, un experto en la materia puede determinar fácilmente la reactividad cruzada de un reactivo, y por tanto no es necesario que se detalle adicionalmente en la presente memoria.

En un ejemplo particular, el kit puede comprender además:

55 c) por lo menos un reactivo apto para determinar específicamente el nivel de expresión de proteínas de por lo menos un biomarcador de cáncer de mama convencional, tal como receptor de estrógenos (ER), receptor de progesterona (PR) o receptor 2 de factor de crecimiento epidérmico humano (HER2).

60 Con el fin de normalizar nivel de expresión de proteínas, el kit puede comprender asimismo opcionalmente por lo menos un reactivo apto para determinar específicamente el nivel de expresión de proteínas de una proteína de mantenimiento, tal como actina, beta-tubulina o gliceraldehído-3-fosfato deshidrogenasa (GAPDH).

En aún otro aspecto, los métodos descritos en la presente memoria pueden ponerse en práctica utilizando una micromatriz, para determinar de manera notable el nivel de expresión de biomarcadores de interés.

El término “micromatriz” se refiere en la presente memoria a una colección espacialmente definida y separada de moléculas biológicas individuales que están inmovilizadas sobre una superficie sólida, y a la que se unen uno o varios biomarcadores de interés específicamente. Esas moléculas biológicas permiten la determinación del nivel de expresión de dicho(s) biomarcador(es), y pueden ser anticuerpos, aficuerpos o aptámeros si la micromatriz es una micromatriz de proteínas, que es un tipo preferido de micromatriz. El experto en la materia conoce bien las tecnologías de micromatrices de proteínas, y se describen de manera notable en Mitchell (2002), Haab (2005) y Eckel-Passow *et al.* (2005), y en las patentes US n^{os} 6.087.102, 6.139.831 y 6.087.103. Para la determinación del nivel de expresión de proteínas de uno o varios biomarcadores utilizando dicha matriz, normalmente pueden utilizarse dos tecnologías: 1) marcaje directo y 2) marcaje indirecto, tal como describen por ejemplo Kingsmore *et al.* (2006). En el método de “marcaje directo”, la proteína de interés (es decir, el biomarcador descrito en la presente memoria, o la diana) obtenida a partir de una muestra, tal como una muestra biológica, se marca con un marcador específico (por ejemplo un marcador fluorescente o de radioisótopo), y posteriormente se hibrida con la micromatriz mediante unión específica a un reactivo que reconoce a dicho biomarcador, estando dicho reactivo conjugado con la superficie de la micromatriz de proteína. Si va a evaluarse el nivel de expresión de varios biomarcadores, cada biomarcador se marca con un marcador distinto. En el método de “marcaje indirecto”, la muestra que contiene el biomarcador de interés se hibrida con la micromatriz mediante unión específica a un reactivo no marcado que reconoce a dicho biomarcador, estando dicho reactivo conjugado con la superficie de la micromatriz de proteínas, y entonces se añade un reactivo marcado secundario, que reconoce específicamente asimismo a dicho biomarcador. La especificidad y sensibilidad de tal método de marcaje indirecto puede mejorarse además utilizando un tercer reactivo marcado, que reconoce al reactivo secundario (ensayo de tipo sándwich). De manera similar, si va a evaluarse el nivel de expresión de varios biomarcadores en el método de marcaje indirecto, cada reactivo secundario o terciario se marca con un marcador distinto. Asimismo pueden utilizarse sistemas sin marcador para determinar el nivel de expresión de un biomarcador en una micromatriz de proteína; en tal sistema, la detección del biomarcador, y por tanto de su nivel de expresión, puede hacerse mediante resonancia de plasmón superficial (SPR), biodetección por microménsula, SELDI-TOF-EM o microscopía de fuerza atómica (Chandra *et al.*, 2011).

Por tanto, se proporciona además una micromatriz de proteína para su utilización en cualquier método descrito anteriormente, que comprende o consiste en:

- a) por lo menos un reactivo apto para determinar específicamente el nivel de expresión de proteínas de por lo menos un biomarcador seleccionado de entre el grupo que consiste en olfactomedina-4, neudesina, desmoplaquina, y cualquier combinación de las mismas.

En un ejemplo particular, dicha micromatriz de proteínas puede comprender además:

- b) por lo menos un reactivo apto para determinar específicamente el nivel de expresión de proteínas de por lo menos un biomarcador de cáncer de mama convencional, tal como receptor de estrógenos (ER), receptor de progesterona (PR) o receptor 2 de factor de crecimiento epidérmico humano (HER2).

Con el fin de normalizar el nivel de expresión de proteínas, la micromatriz puede comprender asimismo opcionalmente por lo menos un reactivo apto para determinar específicamente el nivel de expresión de una proteína de mantenimiento, tal como actina, beta-tubulina, o gliceraldehído-3-fosfato-deshidrogenasa (GAPDH).

La presente invención se pondrá más claramente de manifiesto a partir de la siguiente descripción detallada de experimentos, incluyendo ejemplos. No obstante, el experto en la materia apreciará que la descripción detallada no es limitativa y que pueden introducirse diversas modificaciones, sustituciones, omisiones y cambios sin apartarse del alcance de la invención.

Breve descripción de los dibujos

Figura 1: Sistema de puntuación para la selección de biomarcadores candidatos para el cáncer de mama

Diagrama de flujo que muestra el algoritmo de filtrado utilizado para identificar 5 posibles biomarcadores candidatos, a saber, versicano (VCAN), tenascina (TNC), olfactomedina-4 (OLFM4), neudesina (NENF) y desmoplaquina (DSP).

Figura 2: Verificación preliminar de la expresión de versicano (A), tenascina (B), olfactomedina-4 (C), neudesina (D) y desmoplaquina (E) en pacientes con cáncer de mama frente a controles sanos

Las concentraciones séricas de estos cinco biomarcadores candidatos en pacientes diagnosticados con cáncer de mama y sueros de controles sanos se midieron utilizando ELISA. Las medianas de concentración correspondientes se representan por una línea horizontal. Los candidatos más prometedores son OLFM4, NENF y DSP.

Figura 3: Niveles de expresión de OLFM4 y NENF en muestras de suero de pacientes con cáncer de mama

(primer estudio)

A: Nivel de expresión de olfactomedina-4 circulante en el suero de 335 pacientes con cáncer de mama y en el de 65 controles sanos; se utilizó la prueba de Mann-Whitney para muestras independientes para calcular la significación de los resultados. La olfactomedina-4 se sobreexpresaba significativamente ($p < 0.0001$) en los pacientes con cáncer de mama y en cada uno de los subgrupos por grado.

B: Nivel de expresión de neudesina circulante en el suero de 335 pacientes con cáncer de mama y en el de 65 controles sanos; se utilizó la prueba de Mann-Whitney para muestras independientes para calcular la importancia de los resultados. La neudesina se sobreexpresaba significativamente ($p < 0.0001$) en los pacientes con cáncer de mama y en cada uno de los subgrupos por grado ($p = 0.008$ para el grado 1 y $p < 0.0001$ para el grado 2 y 3).

C: Niveles de expresión de olfactomedina-4 circulante y neudesina en el suero de 335 pacientes con cáncer de mama y en el de 65 controles sanos; la combinación de estas dos proteínas se sobreexpresa significativamente en el grupo de pacientes con cáncer de mama ($p < 0.0001$).

D: Curva ROC del análisis de olfactomedina-4 para distinguir cánceres de mama ($n = 335$) de controles sanos ($n = 65$); el área bajo la curva (AUC) para pacientes de todos los estadios combinados en comparación con la población sin cáncer fue de 0.78 (IC 95: 0.74 - 0.85); la sensibilidad de la prueba fue del 67%, mientras que la especificidad fue del 88% para una concentración > 40 ng/ml.

E: Curva ROC del análisis de neudesina para distinguir cánceres de mama ($n = 241$) de controles sanos ($n = 65$); el área bajo la curva (AUC) para pacientes de todos los estadios combinados en comparación con la población sin cáncer fue de 0.73 (IC 95: 0.70 - 0.80); la sensibilidad de la prueba fue del 47%, mientras que la especificidad fue del 91% para una concentración > 20 ng/ml.

F: Curva ROC del análisis de olfactomedina-4 + neudesina para distinguir cánceres de mama ($n = 241$) de controles sanos ($n = 65$); el área bajo la curva (AUC) para pacientes de todos los estadios combinados en comparación con la población sin cáncer fue de 0.81 (IC 95: 0.74 - 0.86); la sensibilidad de la prueba fue de 74%, mientras que la especificidad fue de 78% para una concentración > 44 ng/ml.

Figura 4: Niveles de expresión de DSP en muestras de suero de pacientes con cáncer de mama (primer estudio)

Nivel de expresión de DSP circulante en el suero de 241 pacientes con cáncer de mama y en el de 65 controles sanos; se utilizó la prueba de Mann-Whitney para muestras independientes para calcular la significación de los resultados. La desmoplaquina se sobreexpresaba significativamente ($p = 0.0037$) en el grupo de cáncer de mama pT1a-pTib y se sobreexpresaba significativamente ($p = 0.0069$) en el grupo de recidiva.

Figura 5: Desregulación de OLFM4, NENF y DSP en pacientes con cáncer de mama (primer estudio)

Proporción de pacientes con cáncer de mama que dieron positivo para la elevación de OLFM4 y NENF y la disminución de DSP.

Figura 6: OLFM4 y baja-DSP son biomarcadores tempranos de cáncer de mama (primer estudio)

A: Proporción de pacientes con cáncer de mama temprano que dieron positivo para alta expresión de OLFM4 y NENF y baja expresión de DSP.

B: Nivel de expresión de OLFM4 circulante en el suero de 335 pacientes con cáncer de mama y en 81 de cáncer de mama temprano y en el de 65 controles sanos; se utilizó la prueba de Mann-Whitney para muestras independientes para calcular la significación de los resultados. OLFM4 se sobreexpresaba significativamente ($p < 0.0001$) en pacientes con cáncer de mama temprano (tamaño tumoral < 1 cm).

C: Curva ROC del análisis OLFM4 para distinguir cánceres de mama tempranos ($n = 81$) de controles sanos ($n = 65$); el AUC para los pacientes en comparación con la población sin cáncer fue de 0.83 (IC 95: 0.75 - 0.89); la sensibilidad de la prueba fue del 67%, mientras que la especificidad fue del 87% para una concentración > 40 ng/ml.

D: Curva ROC obtenida con el factor de predicción que combina OLFM4 y DSP. El AUC para los pacientes fue de 0.92; La sensibilidad de la prueba fue de 87% y la especificidad fue de 84%.

Figura 7: Niveles séricos de OLFM4 y NENF en muestras de cáncer de mama (segundo estudio)

- A: Nivel de olfactomedina-4 circulante en los sueros de pacientes con cáncer de mama BC-1 (cohorte de prueba, n = 277), BC-2 (cohorte de validación-1, n = 171) y BC-3 (cohorte de validación-2, n = 318) y en los de 195 controles sanos; se utilizó la prueba de Mann-Whitney para muestras independientes para calcular la significación de los resultados. La olfactomedina-4 se expresaba significativamente ($p < 0.0001$) en la cohorte de prueba y en ambas cohortes de validación.
- B: Curvas ROC del análisis de olfactomedina-4 para distinguir BC-1, BC-2 y BC-3 de controles sanos; el área bajo la curva (AUC) para los pacientes en comparación con la población sin cáncer fue de 0.89 (IC 95: 0.87 - 0.92) para BC-1; 0.85 (IC 95: 0.81 - 0.89) para BC-2; 0.89 (IC 95: 0.86 - 0.91) para BC-3 y 0.88 (CI 95: 0.85 - 0.90) para las 3 cohortes.
- C: Nivel de neudesina circulante en el suero de pacientes con cáncer de mama BC-1 (cohorte de prueba, n = 277), BC-2 (cohorte de validación-1, n = 171) y BC-3 (cohorte de validación-2, n = 318) y en el de 195 controles sanos; se utilizó la prueba de Mann-Whitney para muestras independientes para calcular la significación de los resultados. La neudesina se sobreexpresaba significativamente ($p < 0.0001$) en la cohorte de prueba y en ambas cohortes de validación.
- D: Curvas ROC de la neudesina para BC-1, BC-2 y BC-3; el AUC para los pacientes en comparación con la población sin cáncer fue de 0.74 (IC 95: 0.68 - 0.79) para BC-1; 0.72 (IC 95: 0.66 - 0.77) para BC-2; 0.74 (IC 95: 0.70 - 0.79) para BC-3 y 0.73 (CI 95: 0.69 - 0.77) para las 3 cohortes.
- E: Curvas ROC de olfactomedina-4 + neudesina para BC-1, BC-2 y BC-3; el AUC para los pacientes en comparación con la población sin cáncer fue de 0.92 (IC 95: 0.89 - 0.95) para BC-1; 0.88 (IC 95: 0.84 - 0.92) para BC-2; 0.91 (IC 95: 0.88 - 0.94) para BC-3 y 0.91 (CI 95: 0.88 - 0.93) para las 3 cohortes.
- F: Curvas ROC de olfactomedina-4 sola, neudesina sola y la asociación de ambas proteínas; el AUC fue de 0.88 (IC 95: 0.85 - 0.90) para la olfactomedina-4 sola, 0.73 (CI 95: 0.69 - 0.77) para la neudesina sola y 0.91 (IC 95: 0.88 - 0.93) para ambas proteínas.

Figura 8: Pacientes positivos para OLFM4, NENF y OLFM4 + NENF (segundo estudio)

Proporción de pacientes que dieron positivo para niveles aumentados de OLFM4 (A), NENF (B) y OLFM4 + NENF (C).

Figura 9: Niveles séricos de OLFM4 y NENF en sueros de pacientes con un tumor pequeño (<1 cm) (segundo estudio)

- A: Nivel de olfactomedina-4 circulante en los sueros de pacientes T1a-T1b-1 (cohorte de prueba, n = 105), T1a-T1b-2 (cohorte de validación-1, n = 123) y T1a-T1b-3 (cohorte de validación -2, n = 108) y en los de 195 controles sanos; se utilizó la prueba de Mann-Whitney para muestras independientes para calcular la significación de los resultados. La olfactomedina-4 se sobreexpresaba significativamente ($p < 0.0001$) en la cohorte de prueba y en ambas cohortes de validación.
- B: Curvas ROC de OLFM4 en las cohortes BC y T1a-T1b; el AUC fue de 0.88 (IC 95: 0.85 - 0.90) en la cohorte BC y 0.89 (IC 95: 0.86 - 0.92) en la T1a-T1b.
- C: Nivel de la neudesina circulante en las mismas cohortes. La neudesina se sobreexpresaba significativamente ($p < 0.0001$) en la cohorte de prueba y en ambas cohortes de validación.
- D: Curvas ROC de NENF en las cohortes BC y T1a-T1b; el AUC fue de 0.73 (IC 95: 0.69 - 0.77) en la cohorte BC y 0.72 (IC 95: 0.68 - 0.77) en la T1a-T1b.

Figura 10: OLFM4 y NENF son biomarcadores tempranos de cáncer de mama (segundo estudio)

- A: Niveles de OLFM4 en sueros de las cohortes T1a-T1b-3 (n = 336), en las cohortes BC-3 (n = 766) y en controles sanos (n = 195).
- B: Niveles de NENF en sueros de las cohortes T1a-T1b-3 (n = 336), en las cohortes BC-3 (n = 766) y en controles sanos (n = 195).
- C: Niveles de OLFM4 + NENF en sueros de las cohortes T1a-T1b-3 (n = 336), en las cohortes BC-3 (n = 766) y en controles sanos (n = 195).
- D: Curvas ROC de OLFM4, NENF y OLFM4 + NENF en las cohortes BC-3 y en las cohortes T1a-T1b-3.

Ejemplos

1. Materiales y métodos

1.1. Selección de pacientes

- Se realizó un estudio preliminar en 20 controles sanos y (20-50) muestras de suero de cáncer de mama de pacientes con cáncer de mama femeninos. Este estudio se dirigía a evaluar el valor de diagnóstico de posibles biomarcadores de cáncer de mama identificados por medio de mapeo proteómico de una estirpe celular de cáncer de mama transformada y de varios tumores de mama.
- Posteriormente se realizó un primer estudio en 65 controles sanos y 335 muestras de suero de cáncer de mama de pacientes con cáncer de mama femeninos. Este estudio se dirigía a evaluar con mayor detalle los valores de diagnóstico de tres posibles biomarcadores de cáncer de mama específicos.

Tabla 2 - Características clínicas/patológicas de pacientes con carcinoma de mama (primer estudio)

Características de los pacientes	BC-1 (n = 335)
Mediana de edad (años) [min-max]	60 [31-90]
< 50 (%)	72 (21.5)
>= 50 (%)	263 (78.5)
HR	
negativo (%)	28 (8.3)
positivo (%)	307 (91.7)
Sobreexpresión de Her2 (%)	40 (11.9)
Positivo para el estado de ganglios linfáticos (%)	41 (12.2)

- En un segundo estudio dirigido a validar adicionalmente dos biomarcadores del primer estudio, se recogieron sueros de pacientes con cáncer de mama femeninos en el Institut de Cancerologie de l'Ouest (ICO) Paul Papin en Angers para la primera cohorte, en ICO René Gauducheau en Nantes para la segunda cohorte y en ICO Paul Papin en Angers para la tercera cohorte. La primera población establecida consistía en 277 sujetos (BC-1) entre los cuales 105 mujeres presentaban un tumor pequeño <1 cm (T1ab-1); la segunda población establecida consistía en 171 individuos con 123 (BC-2) mujeres con un tumor pequeño <1 cm (T1ab-2); la tercera cohorte consistía en 318 mujeres (BC-3) entre las cuales 108 presentaban un tumor pequeño <1 cm (T1ab-3) (ver sus características en la tabla 3 a continuación). Asimismo se evaluaron un total de 195 controles sanos del Etablissement Français du Sang.

Tabla 3 - Características clínicas/patológicas de pacientes con carcinoma de mama (segundo estudio)

Características de los pacientes	BC-1 (n=277)	BC-2 (n=171)	BC-3 (n=318)
Mediana de edad (años) [min-max]	60 [31-87]	61 [33-91]	58 [19-90]
< 50 (%)	60 (21.6)	31 (18.1)	81 (25.5)
>= 50 (%)	217 (78.4)	140 (81.9)	237 (74.5)
HR			
negativo (%)	23 (8.3)	7 (4.1)	29 (9.1)
positivo (%)	254 (91.7)	164 (95.9)	289 (90.9)
Sobreexpresión de Her2 (%)	33 (11.9)	19 (11.1)	41 (12.9)
Positivo para el estado de ganglios linfáticos (%)	33 (11.9)	19 (11.1)	41 (12.9)
T1a-T1b (<1 cm)	105	123	108
	T1ab-1 (n=105)	T1ab-2 (n=123)	T1ab-3 (n=108)
Mediana de edad (años) [min-max]	61 [34-84]	63 [36-90]	60 [37-84]
< 50 (%)	19 (18.1)	17 (13.8)	18 (16.6)
>= 50 (%)	86 (81.9)	106 (86.2)	90 (83.4)
HR			
negativo (%)	7 (6.7)	5 (4.0)	9 (8.3)
positivo (%)	98 (93.3)	118 (96.0)	99 (91.7)
Sobreexpresión de Her2 (%)	12 (11.4)	7 (5.7)	12 (11.1)
Positivo para el estado de ganglios linfáticos (%)	12 (11.4)	6 (4.9)	15 (13.9)

En cada estudio, se recogieron todos los sueros tras obtener el consentimiento informado por escrito. El protocolo del estudio fue aprobado por el Comité de revisión institucional. Todas las muestras se recogieron, se procesaron y se almacenaron de una manera similar. En resumen, se centrifugó la muestra de sangre a 3500 rpm durante 10 minutos, y se almacenó el suero a -80°C. Todos los sueros se obtuvieron antes de la cirugía o el tratamiento neoadyuvante.

1.2. Cultivo celular

Las estirpes celulares epiteliales mamarias humanas MCF10A LXSN (MCF10A) (estirpe celular epitelial mamaria no tumorigénica, que expresa un vector vacío de control) y MCF10A KRASV12 (MCF10A-RAS, que es una estirpe celular epitelial mamaria tumorigénica) se obtuvieron mediante infección retroviral tal como describió anteriormente (Konishi *et al.*, 2007). Las proporcionó amablemente el Dr. Ben Ho'Park. Las estirpes celulares MCF10A LXSN y MCF10A KRASV12 se cultivaron en DMEM/F12 (1:1) (Life Technologies) complementado con Hepes 50 mM, suero de caballo donante al 5% (DHS; Eurobio), L-glutamina al 1%, EGF 20 ng/ml (PeproTech), insulina 10 µg/ml, hidrocortisona 0.5 µg/ml y toxina del cólera 0.1 µg/ml. Todos los suplementos se adquirieron de Sigma-Aldrich a menos que se indique lo contrario. Las células se recogieron utilizando tripsina EDTA, la reacción enzimática se detuvo con 2 volúmenes de medio DMEM/F12 complementado, las células se lavaron dos veces con PBS y los sedimentos celulares se congelaron.

1.3. Extracción de proteínas de las estirpes celulares epiteliales mamarias MCF10A y MCF10A-RAS

Se lisaron aproximadamente 5×10^6 células en 0.6 ml de SDS al 4% y DTT 0.1 M en Tris-HCl 0.1 M, pH 7.6 a temperatura ambiente durante 30 minutos y se sonicaron brevemente para reducir la viscosidad del lisado. El detergente se eliminó de los lisados y las proteínas se digirieron con tripsina según el protocolo de FASP. (Wiśniewski *et al.*, 2009) utilizando unidades de ultrafiltración por centrifugación con un corte de peso molecular nominal de 30000 Daltons. Brevemente, el lisado de proteínas se aplicó a unidades de filtro Microcon YM-30 (n.º de cat. MRCF0R030, Millipore), se centrifugó y se lavó tres veces con 200 µl de urea 8 M en Tris/HCl 0.1 M, pH 8.5. Luego, se añadieron a los filtros 6 µl de MMTS 200 mM en urea 8 M y se incubaron las muestras durante 20 min. Se lavaron los filtros tres veces con 200 µl de urea 8 M en Tris/HCl 0.1 M, pH 8.5, seguidos de seis lavados con 100 µl de TEAB 0.5M. Finalmente, se añadió tripsina (AB sciex) en 100 µl de TEAB 0.5 M a cada filtro. La razón de proteína con respecto a enzima era de 100:1. Las muestras se incubaron durante la noche a 37°C y los péptidos liberados se recogieron por centrifugación. Las muestras se secaron entonces completamente utilizando un instrumento Speed-Vac y se resuspendieron en 100 µl de ácido trifluoroacético (TFA) a 0.5% en acetonitrilo a 5%, y se desalaron por medio de columnas de centrifugación PepClean C-18 (Pierce Biotechnology, Rockford, IL). El contenido de péptidos se determinó utilizando un kit de ensayo de proteínas Micro BCA (Pierce - Thermo Scientific, Rockford, IL).

1.4. Extracción de proteínas de tejidos congelados (tumores mamarios y tejidos mamarios sanos)

Se cortaron secciones congeladas (12 µm de grosor) de tumores de mama o área de mama normal en un criostato (Bright Instrument Co Ltd, St Margarets Way, Reino Unido). Se tiñeron secciones específicas con azul de toluidina para referencia visual. Para considerar la heterogeneidad tumoral, se lisaron diez secciones congeladas por tumor de tumores de mama luminales A, con Her-2 sobreexpresado y triples negativos en un tampón que consistía en Tris-HCl 0.1 M, pH 8.0, DTT 0.1 M y SDS al 4% a 95°C durante 90 min. Se eliminó el detergente de los lisados y se digirieron las proteínas con tripsina según el protocolo de FASP. (Wisniewski *et al.*, 2009) utilizando unidades de ultrafiltración por centrifugación con un corte de peso molecular nominal de 30000 Daltons. A las unidades de filtro Microcon YM-30 (n.º de cat. MRCF0R030, Millipore) que contenían concentrados de proteínas, se les añadieron 200 µl de urea 8 M en Tris/HCl 0.1 M, pH 8.5 (UA) y se centrifugaron las muestras a 14000 g a 20°C durante 8 min. Esta etapa se realizó tres veces. Luego, se añadieron a los filtros 6 µl de MMTS 200 mM en urea 8 M y se incubaron las muestras durante 20 min. Se lavaron los filtros tres veces con 200 µl de UA 8 M seguido de seis lavados con 100 µl de TEAB 0.5 M. Finalmente, se añadió tripsina (AB sciex) en 100 µl de TEAB 0.5 M a cada filtro. La razón de proteína con respecto a enzima era de 100:1. Las muestras se incubaron durante la noche a 37°C y los péptidos liberados se recogieron por centrifugación. Las muestras se secaron entonces completamente utilizando un instrumento Speed-Vac y se resuspendieron en 100 µl de ácido trifluoroacético (TFA) al 0.5% en acetonitrilo al 5%, y se desalaron por medio de columnas de centrifugación PepClean C-18 (Pierce Biotechnology, Rockford, IL). El contenido de péptidos se determinó utilizando el kit de ensayo de proteínas Micro BCA (Pierce - Thermo Scientific, Rockford, IL).

1.5. Marcaje de péptidos con reactivos iTRAQ.

Para el marcaje con iTRAQ, se marcaron 100 µg de cada disolución de péptidos a temperatura ambiente durante 2 h con un vial de reactivo iTRAQ previamente reconstituido con 70 µl de etanol para el reactivo 4TRA iTRAQ. Los péptidos marcados se mezclaron posteriormente en una razón de 1:1:1:1 y se secaron completamente utilizando un instrumento Speed-Vac.

1.6. Fraccionamiento de péptido OFFGEL

Para la separación de péptidos basada en pl, se utilizó el fraccionador OFFGEL 3100 (Agilent Technologies, Böblingen, Alemania) con una configuración de 12 o 24 pocillos utilizando el siguiente protocolo. En primer lugar, las muestras se desalaron en un cartucho Sep-Pak C18 (Waters). Para la configuración de 24 pocillos, las muestras de péptidos se diluyeron en la disolución de muestra de péptidos OFFGEL hasta un volumen final de 3.6 ml. Luego, la tira de gel de IPG de 24 cm de largo (GE Healthcare, Múnich, Alemania) con un intervalo de pH

lineal de 3-10 se rehidrató con la disolución de rehidratación de tiras de péptidos IPG, según el protocolo del fabricante, durante 15 min. Se cargaron 150 µl de muestra en cada pocillo. Se realizó el electroenfoque de los péptidos se realizó a 20°C y 50 µA hasta alcanzar el nivel de 50 kVh. Después del enfoque, se retiraron las 24 fracciones de péptidos y se lavaron los pocillos con 200 µl de una disolución de agua/metanol/ácido fórmico (49/50/1). Después de 15 minutos, se combinó cada disolución de lavado con su correspondiente fracción de péptido. Todas las fracciones se evaporaron por centrifugación a vacío y se mantuvieron a -20°C. Para el enfoque de 2D-OFFGEL, los péptidos se fraccionaron en primer lugar en 12 fracciones en el intervalo de pH 3-10. Luego, las fracciones F1-F2, la fracción F3 a F8 y las fracciones F9 a F12 se agruparon y volvieron a fraccionarse en 24 fracciones en el intervalo de pH 3.5-4.5, 4-7 y 6-9, respectivamente. Se obtuvieron 72 fracciones que se analizaron posteriormente mediante nanoCL-EM/EM.

1.7. Separación por CL capilar

Justo antes del análisis por nano-CL, cada fracción se resuspendió en 20 µl de H₂O con TFA al 0.1% (v/v). Las muestras se separaron en un sistema de nano-CL Ultimate 3000 (Dionex, Sunnyvale, EE. UU.) Utilizando una columna C18 (PepMap100, 3 µm, 100A, 75 µm di x 15 cm, Dionex) a 300 nl/min de velocidad de flujo. El tampón A era ACN al 2% en agua con TFA al 0.05% y el tampón B fue ACN al 80% en agua con TFA al 0.04%. Los péptidos se desalaron durante 3 minutos utilizando solo tampón A en la precolumna, seguido de una separación durante 105 minutos utilizando el siguiente gradiente: del 0 al 20% de B en 10 minutos, del 20% al 45% de B en 85 minutos y del 45% al 100% B en 10 min. Los cromatogramas se registraron a la longitud de onda de 214 nm. Se recogieron fracciones de péptidos utilizando un colector de microfracciones Probot (Dionex). Se utilizó CHCA (LaserBioLabs, Sophia-Antipolis, Francia) como matriz de MALDI. La matriz (concentración de 2 mg/ml en ACN al 70% en agua con TFA al 0.1%) se añadió continuamente al efluente de la columna a través de una pieza de mezclado micro "T" a una velocidad de flujo de 1.2 µl/min. Después de 12 minutos de ejecución, se envió una señal de inicio al Probot para iniciar el fraccionamiento. Se recogieron las fracciones durante 10 segundos y se dispusieron en manchas sobre una placa de muestra de MALDI (1.664 manchas por placa, ABSciex, Foster City, CA.).

1.8. MALDI-EM/EM

Los análisis de EM y EM/EM de muestras de péptidos dispuestos en manchas fuera de línea se realizaron utilizando el analizador de MALDI-TOF/TOF 5800 (ABSciex) y el software 4000 Series Explorer, versión 4.0. El instrumento se hizo funcionar en modo de ion positivo y se calibró externamente utilizando un kit estándar de calibración de masa (ABSciex). La potencia del láser se estableció entre 2800 y 3400 para EM y entre 3600 y 4200 para adquisición de EM/EM. Después de examinar todas las posiciones de muestra de CL-MALDI en modo de reflector EM-positivo utilizando 2000 disparos de láser, se realizó la fragmentación de precursores seleccionados automáticamente a una energía de colisión de 1 kV utilizando aire como gas de colisión (presión de ~ 2 x 10⁻⁶ Torr) con una acumulación de 3000 disparos para cada espectro. Los espectros de EM se adquirieron entre m/z 1000 y 4000. El ion original de Glu1-fibrinopéptido se utilizó a m/z 1570.677 diluido en la matriz (30 femtomoles por punto) para la calibración interna. Se seleccionaron hasta 12 de las señales de iones más intensas por posición de mancha que presentaban un S/N > 20 como precursores para la adquisición de EM/EM. La identificación de péptidos y proteínas se realizó mediante el software ProteinPilot™ V 4.0 (AB Sciex) utilizando el algoritmo Paragon como motor de búsqueda. (Shilov *et al.*, 2007). En cada espectro de EM/EM se buscaron especies de *Homo sapiens* en la base de datos Uniprot/swissprot (UniProtKB/Sprot 20110208 versión 01, con 525997 entradas de secuencia). Las búsquedas se realizaron utilizando la modificación fija del parámetro de cisteína marcada con metilmetanotiosulfato habilitado. Otros parámetros, tales como la especificidad de escisión triptica, la precisión de la masa del ion precursor y la precisión de la masa del ion del fragmento, fueron las funciones integradas en MALDI 5800 del software ProteinPilot. El umbral de proteínas detectado (protscore no utilizado (confianza)) en el software se estableció en 1.3 para lograr un 95% de confianza, y las proteínas identificadas se agruparon mediante el algoritmo ProGroup (ABSciex) para minimizar la redundancia. Se ejecutó la opción de corrección de sesgo.

Para estimar la tasa de descubrimiento falso (FDR), se utilizó una estrategia de búsqueda de base de datos de señuelo. La FDR se define como el porcentaje de proteínas señuelo identificadas frente a la identificación total de proteínas. La FDR se calculó buscando el espectro en la base de datos de señuelos de *Homo sapiens* de Uniprot.

1.9. Pruebas de ELISA

Se utilizaron kits de ELISA disponibles comercialmente de USCN Life Science Inc. o R&D para analizar las concentraciones de OLFM4, NENF y DSP. Los kits consistían en placas de microtitulación de 96 pocillos recubiertas con un anticuerpo específico para cada tipo de molécula, anticuerpos de detección para identificar el anticuerpo-proteína en la placa mediante marcaje con estreptavidina-biotina y sustrato TMB que generó un producto coloreado. La muestra se añadió y el ensayo se realizó según las instrucciones del fabricante. La absorbancia del producto coloreado desarrollado al final del ensayo se cuantificó a una longitud de onda de 450 nm en un lector de ELISA (Tecan Magellan Sunrise).

1.10. Cuantificación estadística de la expresión de proteínas relativa

Para la cuantificación de la expresión de proteínas relativa, se utilizaron el paquete de software personalizado iQuantitator (Schwacke *et al.*, 2009; Grant *et al.* 2009) así como los softwares TANAGRA (V1.4) y GraphPad Prism 5 para inferir la magnitud del cambio en la expresión de proteínas. Esos softwares infieren cambios dependientes del tratamiento en la expresión utilizando métodos estadísticos bayesianos, más específicamente, la prueba de Mann-Whitney para muestras independientes y curvas de características operativas del receptor (ROC). Básicamente, se utilizó este enfoque para generar medias, medianas, el área bajo la curva ROC (AUC) e intervalos de confianza del 95% (superior e inferior) para someter a prueba la hipótesis de que el AUC era superior a 0.5 para cada cambio dependiente del tratamiento en la expresión de proteínas mediante la utilización de datos a nivel de péptido para cada péptido componente.

Para proteínas cuyas razones de iTRAQ estaban reguladas por disminución, la extensión de la regulación por disminución se consideró adicionalmente si el límite superior del intervalo creíble presentaba un valor inferior a 1. A la inversa, para las proteínas cuyas razones de iTRAQ aumentaron, la extensión de la regulación por incremento se consideró adicionalmente si el límite inferior del intervalo de confianza presentaba un valor mayor de 1. La anchura de estos intervalos de confianza dependía de los datos disponibles para una proteína dada. El software iQuantitator tenía en consideración todos los péptidos observados y el número de espectros utilizados para cuantificar el cambio en la expresión de una proteína dada. En estas condiciones, fue posible detectar cambios pequeños pero significativos en la regulación por incremento y por disminución cuando estaban disponibles muchos péptidos. Los criterios de selección de péptidos para la cuantificación relativa se realizaron tal como sigue. Solo se consideraron péptidos únicos para una proteína dada para la cuantificación relativa, excluyendo los comunes a otras isoformas o proteínas de la misma familia. Las proteínas se identificaron basándose en presentar por lo menos dos péptidos con una puntuación de iones por encima del 95% de confianza.

2. Resultados

La estrategia utilizada para identificar nuevos biomarcadores candidatos para tumores de mama se resume en la figura 1. En primer lugar, en el contexto de la presente invención se traza un mapa proteómico de una estirpe celular de mama transformada utilizando varios tumores de mama con diferentes estados con el fin de establecer la lista más exhaustiva posible de las proteínas que caracterizan a los tumores de mama. Esta lista se comparó con la lista de proteínas secretadas que ya se han identificado en la sangre humana (Schenk *et al.*, 2008; Cao *et al.*, 2013; sitio web del HUPO Plasma Proteome Project <http://www.ccmb.med.umich.edu/PPP>), luego se examinó en la "Base de datos de cáncer de mama" (<http://www.itb.cnr.it/breastcancer/>) para que solo se seleccionaran las nuevas proteínas secretadas en la patología de cáncer de mama.

2.1. Análisis proteómico de las estirpes celulares MCF10A y MCF10A-RAS (Imbalzano *et al.*, 2013)

Utilizando el enfoque OFFGEL de dos etapas, en el contexto de la presente invención se mapeó el proteoma de la estirpe celular de mama no transformada (MCF10A), así como la de esta misma estirpe celular transformada con el oncogén KRAS con el fin de imitar la activación oncogénica y la supervivencia anómala. Aunque no son comunes mutaciones en el gen RAS en los cánceres de mama, la ruta de RAS se activa en esta enfermedad por la sobreexpresión de la señalización de receptores de factores de crecimiento tales como el receptor ErbB2, que se activa en el 30% de los cánceres de mama. Los tumores de mama inducidos por RAS se caracterizan por la activación de la señalización de proteína cinasas activadas por mitógeno, que es conocido que está asociada con neoplasia temprana y mal pronóstico. Utilizando el enfoque proteómico descrito en la presente memoria, se identificaron 2152 proteínas con por lo menos dos péptidos. De estas proteínas, 262 se encontraron en las bases de datos de proteínas secretadas (datos no representados).

2.2. Análisis proteómico de tumores de mama luminales A

A través de un mapeo proteómico similar de tumores de mama luminales A, en el contexto de la presente invención se pudieron identificar 1093 proteínas con por lo menos dos péptidos. De estas proteínas, 246 eran proteínas secretadas (datos no representados).

2.3. Análisis proteómico de tumores de mama que expresan HER2

A través de un mapeo proteómico similar de tumores de mama con sobreexpresión del receptor HER2, en el contexto de la presente invención se pudieron identificar 624 proteínas con por lo menos dos péptidos. De estas proteínas, 225 eran proteínas secretadas (datos no representados).

2.4. Análisis proteómico de tumores de mama triple negativos

A través de un mapeo proteómico similar de tumores de mama triple negativos, en el contexto de la presente

invención se identificaron 2407 proteínas con por lo menos dos péptidos. De estas proteínas, 266 eran proteínas secretadas (datos no representados).

2.5. Análisis proteómico comparativo

Una comparación de todas las proteínas secretadas identificadas en la estirpe celular MCF10A-RAS, así como en tumores de mama luminales A, positivos para HER2 y triple negativos frente a la "base de datos de cáncer de mama" (<http://www.itb.cnr.it/breastcancer/>) permitió seleccionar 125 proteínas nuevas en la patología de cáncer de mama en el primer estudio, de las cuales 121 se validaron en el segundo estudio (datos no representados).

2.6. Nuevas proteínas secretadas desreguladas en la estirpe celular transformada MCF10A transformada mediante el oncogén KRAS (MCF10-RAS)

Comparando la lista de las 121 proteínas secretadas con la lista de proteínas significativamente desreguladas en la estirpe celular MCF10A transformada mediante el oncogén KRAS, se identificaron 12 proteínas (tabla 4). Siete se subexpresaban (DSP, JUP, ACTN1, CTNNA1, METTL13, HSPD1 y GSTP1) y cinco se sobreexpresaban (COPA, TLN1, PYGB, HPCAL1 e IGF2R).

Tabla 4. Proteínas secretadas desreguladas en la estirpe celular MCF10A transformada mediante el oncogén KRAS

Símbolo	N.º de registro	Nombre completo del biomarcador	Razón de iTRAQ
DSP	sp P15924 DESP_HUMAN	Desmoplaquina	0.0413
JUP	sp P14923 PLAK_HUMAN	Placoglobina de uniones	0.0643
ACTN1	sp P12814 ACTN1_HUMAN	Alfa-actinina-1	0.2559
CTNNA1	sp P35221 CTNA1_HUMAN	Catenina alfa-1	0.2606
METTL13	sp Q8N6R0 MTL13_HUMAN	Proteína 13 similar a metiltransferasa	0.6668
HSPD1	sp P10809 CH60_HUMAN	Proteína de choque térmico de 60 kDa, mitocondrial	0.6792
GSTP1	sp P09211 GSTP1_HUMAN	Glutación-S-transferasa P	0.879
COPA	sp P53621 COPA_HUMAN	Subunidad alfa de coatómero	3.4995
TLN1	sp Q9Y490 TLN1_HUMAN	Talina-1	5.7016
PYGB	sp P11216 PYGB_HUMAN	Glucógeno fosforilasa, forma cerebral	8.4723
HPCAL1	sp P37235 HPCL1_HUMAN	Proteína 1 similar a hipocalcina	11.3763
IGF2R	sp P11717 MPRI_HUMAN	Receptor de manosa-6-fosfato independiente de catión	17.5388

2.7. Proteínas secretadas nuevas desreguladas en tumores triples negativos

Comparando la lista de las 121 proteínas secretadas frente a la lista de proteínas significativamente desreguladas en los tumores triples negativos, se identificaron 15 proteínas. Dos estaban reguladas por disminución (APOH y CFH) y 13 se sobreexpresaban (CMPK1, ALDOA, COPA, DDT, CFL1, GSTO1, ARF1, COTL1, FTL, DSTN, DSP, ACTN1, TNC) (tabla 5).

Tabla 5. Proteínas secretadas desreguladas en tumores triples negativos

Símbolo	Nombre completo del biomarcador	Razón de iTRAQ
APOH	Beta-2-glicoproteína 1	0.602
CFH	Factor del complemento	0.693
CMPK1	UMP-CMP cinasa	1.332
ALDOA	Fructosa-bisfosfato aldolasa A	1.37
COPA	Subunidad alfa de coatómero	1.496
DDT	D-dopacromo descarboxilasa	1.504
CFL1	Cofilina-1	1.569
GSTO1	Glutación-S-transferasa omega-1	1.586
ARF1	Factor 1 de ADP-ribosilación	1.624
COTL1	Proteína similar a coactosina	1.703
FTL	Cadena ligera de ferritina	1.743
DSTN	Destrina	1.978
DSP	Desmoplaquina	2.119
ACTN1	Alfa-actinina-1	2.196
TNC	Tenascina	1.616

2.8. Proteínas secretadas nuevas desreguladas en un tumor HER2+ en comparación con el tejido sano

Comparando el proteoma de proteínas desreguladas frente a la lista de las 125 proteínas, se caracterizaron cinco proteínas; una proteína subexpresada (CFH) y cuatro proteínas sobreexpresadas (ANXA2, FTL, TAGLN2, TNC) (tabla 6).

Tabla 6. Proteínas secretadas desreguladas en tumor HER2+

Símbolo	Nombre completo del biomarcador	Razón de iTRAQ
ANXA2	Anexina A2	3.1623
CFH	Factor del complemento H	0.0474
FTL	Cadena ligera de ferritina	20.7014
TAGLN2	Transgelina-2	14.5881
TNC	Tenascina	9.6383

2.9. Análisis glicoproteómico

El propósito de este análisis fue completar el subproteoma de proteínas secretadas. Utilizando tres tumores de mama, en el contexto de la presente invención se establece un glicoproteoma. Utilizando este glicoproteoma, se caracterizaron 5 proteínas secretadas (HPX, OLFM4, OLFML3, TNC, VCAN) que nunca se han estudiado como biomarcadores de cáncer de mama.

2.10. Selección de biomarcadores candidatos.

En el contexto de la presente invención se utiliza un sistema de puntuación sistemático para segregar 5 candidatos: DSP, NENF, OLFM4, TNC, VCAN.

2.11. Validación de candidatos en pacientes con cáncer de mama

Se realizó una verificación preliminar en 20 controles sanos y (20-50) muestras de suero de cáncer de mama. Las medianas de concentración para las muestras de cáncer con TNC y VCAN no fueron significativamente diferentes ($p > 0.05$) de las de los controles sanos. Las medianas de concentración para OLFM4 y NENF fueron 2.2 y 3.1 veces más altas que los sueros de controles sanos, respectivamente, con un valor de $p < 0.005$. La mediana de concentración para las muestras de cáncer de mama con DSP no fue significativamente diferente de las de los controles sanos, pero se identificó una diferencia significativa (valor de $p < 0.032$) entre las muestras de controles sanos y el grupo de tumores pequeños (tamaño < 2 cm, pT1) (figura 2). En este caso, la concentración de DSP fue menor que en el grupo de control. A la inversa, entre las 50 muestras de cáncer de mama, se encontraron 3 muestras de DSP sobreexpresadas que coincidieron con 3 tumores de mama recidivantes.

2.12. Elevación de OLFM4 y NENF en sueros de cáncer de mama

- Para evaluar adicionalmente el potencial de OLFM4 y NENF como biomarcadores de cáncer de mama en suero, sus concentraciones séricas se determinaron en un primer estudio basado en 65 sujetos sanos y en 335 con cáncer de mama. Se encontró que OLFM4 y NENF estaban significativamente elevados ($p < 0.0001$) en sueros de cáncer de mama (independientemente del grado del tumor) en comparación con los sueros sanos (figura 3A y B). La concentración sérica de OLFM4 y NENF se combinó entonces para cada paciente ($n = 335$) y este valor se comparó con el obtenido en sueros de controles sanos: se encontró que la concentración de OLFM4 + NENF estaba significativamente elevada ($p < 0.0001$) en los sueros de cáncer de mama (figura 3C).
- En un segundo estudio de validación, 766 participantes se dividieron en 3 cohortes independientes: BC-1 reclutados en el Angers ICO Cancer Canter, BC-2 reclutados en el Nantes ICO Cancer Canter y BC-3 reclutados en el Angers ICO Cancer Center. Las concentraciones de ambos marcadores asimismo se determinaron en 195 sujetos sanos. Para la cohorte de control, la mediana de concentración de OLFM4 fue de 9.96 ng/ml (IQR 1.00-21.94) y para NENF, la mediana de concentración fue de 6.77 ng/ml (IQR 1.46-13.14). Se encontró que las medianas de concentración para OLFM4 y NENF estaban significativamente elevadas ($p < 0.0001$) en los sueros de cáncer de mama en comparación con las muestras sanas (figuras 7A y 7C); Los valores no difirieron significativamente entre las tres cohortes. Cuando se combinaron las 3 cohortes de sueros de cáncer de mama independientes, se determinó una mediana de concentración de OLFM4 de 47.00 ng/ml (IQR 25.00-75.00) y una mediana de concentración de NENF de 16.82 ng/ml (IQR 8.05-31.69).

Las curvas ROC mostraron que el corte de diagnóstico óptimo para OLFM4 fue de 29.8, 30.0 y 29.4 ng/ml para BC-1, BC-2 y BC-3 respectivamente. El valor de corte óptimo para NENF fue de 13.8, 15.6 y 13.8 ng/ml para BC-1, BC-2 y BC-3, respectivamente. Cuando se combinaron las 3 cohortes (figura 7B y 7D), los valores de diagnóstico óptimos fueron 30.6 y 15.6 ng/ml para OLFM4 y NENF, respectivamente. Estos valores fueron muy

similares a los obtenidos para cada cohorte independiente, los valores de corte en este estudio se eligieron para fueran de 31 ng/ml para OLFM4 y 16 ng/ml para NENF (tabla 7). Con estos valores de corte, la sensibilidad oscilaba entre el 64 y el 78% para OLFM4 y entre el 52% y el 60% para NENF.

5 Con el fin de desarrollar una prueba específica, se calculó el valor de corte a una especificidad del 90 y 95%:

- para OLFM4, con una especificidad del 90%, el valor de corte alcanzó 33.9 ng/ml en las tres cohortes independientes y en la cohorte total con una sensibilidad que oscila entre el 57 y el 71%; con una especificidad del 95%, el valor de corte se elevó hasta 41 o 42 ng/ml con una sensibilidad del 49 al 63% (tabla 7);
- para NENF, con una especificidad del 90%, el valor de corte alcanzó 13.8 o 15.6 ng/ml para las diferentes cohortes con una sensibilidad que oscila entre el 53 y el 60% y con una especificidad del 95%, el valor de corte se elevó hasta 39 ng/ml con una sensibilidad del 14 al 22% (tabla 7).

Para someter a prueba si estos marcadores eran complementarios, los marcadores combinados OLFM4 y NENF se estimaron mediante regresión logística binaria y los valores de esta función se utilizaron como marcador y se sometieron a análisis ROC. Los valores correspondientes a la adición de las concentraciones de Elisa de OLFM4 y NENF en el mismo suero asimismo se sometieron a prueba y las curvas ROC eran tal como se mostró equivalentes (datos no representados). Entonces, al combinar las concentraciones de Elisa de OLFM4 y NENF, las curvas ROC mostraron que el corte de diagnóstico óptimo fue de 38.3 o 38.4 ng/ml en las tres cohortes independientes o en la cohorte combinada con una especificidad del 87% y una sensibilidad que oscila entre el 75 y el 85% (figura 7E).

Cuando se compararon las curvas ROC, el AUC para OLFM4 + NENF y OLFM4 solo parecían estar muy cerca (figura 7F). La proporción de pacientes que eran positivos para OLFM4, NENF y OLFM4 + NENF se compararon en las diferentes condiciones de especificidad (figura 8). Para OLFM4, la proporción de pacientes positivos pareció ser mayor en la cohorte de prueba (BC-1). A través de las 3 cohortes, una proporción de pacientes positivos del 70% para el corte óptimo, y del 66% y el 56% de pacientes positivos a una especificidad del 90% y el 95%, respectivamente (figura 8A). Para NENF, la proporción de pacientes positivos fue bastante similar en las 3 cohortes y se alcanzó un promedio del 53% de pacientes positivos en el corte óptimo, y se alcanzaron del 32% y el 19% para una especificidad del 90% y el 95%, respectivamente (figura 8B). Cuando se combinaron los dos marcadores, parecía que, en la cohorte de prueba y en la cohorte de validación 2 (BC-2), más del 80% de los pacientes fueron positivos (el 75% para la cohorte BC-2), en el valor de corte óptimo. Cuando se evaluó la proporción a una especificidad del 90%, se alcanzó todavía entre el 70 y el 81% de pacientes positivos. Con una especificidad del 95%, entre 62 y 73 pacientes fueron positivos para OLFM4 + NENF (figura 8C). La proporción de pacientes que fueron positivos para OLFM4 en el corte óptimo (31 ng/ml) fue del 70% y esta proporción alcanzó el 81% cuando se combinaron los dos marcadores (en el corte óptimo de 39 ng/ml, figura 8A y 8C). El número de pacientes positivos para OLFM4 + NENF fue superior al 10% por lo menos para cada valor de especificidad, comparando la proporción de pacientes positivos para OLFM4 solo.

Tabla 7. Valores de área bajo la curva ROC, sensibilidad y especificidad para pruebas de diagnóstico basadas en OFLM4, NENF o OLFM4+NENF (segundo estudio)

	AUC (IC del 95%)	Sensibilidad (%)	Especificidad (%)	Corte (ng/ml)	Especificidad del 90%		Especificidad del 95%	
					Sensibilidad (%) (IC del 95%)	Corte (ng/ml)	Sensibilidad (%) (IC del 95%)	Corte (ng/ml)
<i>BC-1</i>								
OLFM4	0.8945	76.5	87.7	29.8	71.5	33.9	63.54	40.9
NENF	0.7367	57.24	76.92	13.8	29.66	25.9	14.48	38.4
OLFM4+NENF	0.9188	85.6	87.3	38.3	81.48	43.6	73.61	53.3
<i>BC-Nantes</i>								
OLFM4	0.8543	64.3	87.7	30	57.89	33.9	49.71	41
NENF	0.7169	53.8	81.5	15.6	30.41	25.9	17.54	39.0
OLFM4+NENF	0.8811	75.8	87.34	38.4	71.34	43.5	63.06	53.7
<i>BC-2</i>								
OLFM4	0.8869	78.62	82.5	29.4	64.78	33.9	53.46	40.9
NENF	0.7442	60.33	76.92	13.8	34	25.9	22.33	39
OLFM4+NENF	0.9146	82.32	87.34	38.3	77.17	43.5	62.06	53.7
<i>T1a-T1b-1</i>								
OLFM4	0.9189	83.81	88.21	30.8	79	33.6	66.67	42
NENF	0.7111	51.06	82.56	16.3	29.79	25.9	14.89	39.5
OLFM4+NENF	0.9177	87.62	86.08	37.8	80	43.5	68.57	53.6
<i>T1a-T1b-Nantes</i>								
OLFM4	0.8828	77.24	88.21	30.6	73.17	33.9	67.48	41.9
NENF	0.7363	59.35	81.54	15.6	37.4	26	19.51	39
OLFM4+NENF	0.9029	78.05	92.41	45.6	78.05	43.5	71.54	53.7
<i>T1a-T1b-2</i>								
OLFM4	0.8778	79.61	83.08	27.5	68.93	33.9	51.46	41.8
NENF	0.7141	43.3	87.69	19.8	30.93	26.1	20.62	39.7
OLFM4+NENF	0.8986	83.33	84.18	35.4	74.07	43.5	60.19	53.7
<i>BC-3 cohortes</i>								
OLFM4	0.8769	70.17	88.21	30.8	65.67	34	55.67	41.4
NENF	0.7348	52.92	81.54	15.6	33.71	25.9	19.1	39
OLFM4+NENF	0.9083	81.87	87.34	38.3	77.19	43.5	65.94	53.6
<i>T1a-T1b-3 cohortes</i>								
OLFM4	0.8927	77.64	88.21	30.6	73.72	33.6	63.75	40.9
NENF	0.7238	53.18	81.54	15.6	31.98	26	19.16	38.7
OLFM4+NENF	0.9061	82.14	87.34	38.4	77.38	43.5	66.96	53.6

2.13. La DSP disminuye en los sueros de pacientes con tumores tempranos y aumenta en los sueros de pacientes con tumores recidivantes

Para evaluar el potencial del DSP para discriminar tumores pequeños (baja DSP) y tumores recidivantes (alta DSP) de los controles, se determinaron las concentraciones séricas en 65 sujetos sanos y en 384 con cáncer de mama (primer estudio). Se encontró que la DSP estaba significativamente elevada ($p = 0.0069$) en sueros de tumores recidivantes, según el enfoque proteómico del tumor. A la inversa, se encontró una disminución significativa de la DSP ($p = 0.0037$) en tumores pequeños con un tamaño <1 cm (tumores pT1a y pT1b (figura 4)). Esta concentración sérica de DSP baja fue consecuente con los resultados proteómicos de la estirpe celular mamaria.

2.13. OLFM4, NENF y DSP son biomarcadores de cáncer de mama

La proporción de pacientes que dieron positivo para OLFM4, NENF y DSP se mostró en la figura 5 (primer estudio). OLFM4 mostró una elevación significativa en 208 sueros de cáncer de mama (335 sueros en la cohorte total). Entre 127 pacientes que fueron negativos para OLFM4, 58 fueron positivos para NENF. Entre los últimos 69 pacientes que fueron negativos para OLFM4 y NENF, 39 fueron positivos para baja DSP. Por lo menos, 30 pacientes fueron negativos para los tres biomarcadores (figura 5). Debe indicarse que entre esta cohorte de 335 pacientes, 32 pacientes presentan un nivel de DSP alto (> 1800 pg/ml). Entre estos 32 pacientes, 12 estaban en un estado recidivante (38%).

2.14. OLFM4, baja DSP y NENF son biomarcadores para el cáncer de mama en la fase temprana

- Para evaluar adicionalmente el potencial de OLFM4, NENF y baja DSP como biomarcadores de cáncer de mama en suero en la fase temprana, se determinaron sus concentraciones en suero en un primer estudio basado en 65 sujetos sanos y en 81 pacientes con un tumor pequeño (<1 cm), que representa cáncer de mama temprano. La proporción de sujetos que dieron positivo para OLFM4 fue del 65% y esta proporción no aumentó cuando se añadieron pacientes que dieron positivo para NENF, lo que sugiere que NENF podría no ser un biomarcador del tumor de mama temprano. No obstante, al añadir a los sujetos positivos para baja DSP, el 91% de los pacientes se detectaron como positivos (figura 6A). Se construyó un factor de predicción que combina información de OLFM4 y baja DSP utilizando un modelo de regresión logística. Este factor de predicción mostró un AUC = 0.92 para pacientes con cáncer de mama en comparación con una población sin cáncer. La sensibilidad de la prueba fue del 87% y la especificidad fue del 84% (figura 6D).
- En el segundo estudio de validación que se llevó a cabo en una cohorte mucho mayor de pacientes, se evaluó adicionalmente el potencial de OLFM4 y NENF como biomarcadores séricos en la fase temprana del cáncer de mama. Para este fin, se determinaron sus concentraciones séricas en 336 pacientes con un tumor pequeño (<1 cm) dividido en 3 cohortes independientes (105 pacientes para T1ab-1, 123 pacientes para T1ab-2 y 108 pacientes para T1ab-3).

En cuanto a las cohortes BC, se encontró que la mediana de concentración de OLFM4 y NENF estaba significativamente elevada ($p < 0.0001$) en los sueros de tumores de mama T1a-T1b en comparación con las muestras sanas (figuras 9A y 9C). Los valores no difirieron significativamente entre las tres cohortes, y fueron similares a las cohortes BC. Cuando se combinaron las 3 cohortes de sueros de cáncer de mama independientes, se determinó que la mediana de concentración de OLFM4 era de 51.89 ng/ml (IQR 32.69-83.52) y la mediana de concentración de NENF de 17.40 ng/ml (IQR 7.56-30.99).

Las curvas ROC mostraron que el AUC para OLFM4 en la cohorte T1a-T1b era ligeramente más alta que en la cohorte BC (0.89 y 0.88 respectivamente) mientras que los valores de AUC para NENF eran idénticos (tabla 7 y figura 9D). Cuando se combinaron ambos marcadores, las medianas de concentraciones fueron idénticas en las cohortes T1a-T1b y en las cohortes BC (figura 10). De la misma manera, las curvas ROC eran superponibles. Cuando se comparó la proporción de pacientes que fueron positivos para OLFM4, NENF y OLFM4 + NENF en las diferentes condiciones de especificidad, los resultados parecían estar muy cerca de los obtenidos para las cohortes BC. Además, las curvas ROC para OLFM4 en las cohortes BC y T1a-T1b estaban superpuestas.

Los análisis de pacientes positivos para OLFM4, NENF y OLFM4 más NENF mostraron resultados comparables a los obtenidos en cohortes BC. La proporción de pacientes que fueron positivos para OLFM4 o NENF en el valor de corte óptimo de la curva ROC fue del 86%. A una especificidad del 90%, esta proporción fue del 78% y a una especificidad del 95%, del 67%. Al combinar OLFM4 más NENF, la proporción de pacientes que fueron positivos para OLFM4 + NENF, u OLFM4 solo o NENF solo asimismo fue del 86% al valor de corte óptimo, este valor fue del 79% a una especificidad del 90% y del 69% al 95% de especificidad.

3. Conclusión

El objetivo de los presentes estudios fue utilizar un enfoque proteómico innovador para identificar nuevos biomarcadores de cáncer de mama. Partiendo del hecho de que ningún candidato serio se había destacado con los enfoques habituales, el presente trabajo se centró en proteínas secretadas que nunca se habían descrito como posibles biomarcadores de cáncer de mama. La primera etapa fue crear una base de datos de proteínas a partir de una estirpe celular de mama transformada, y de tumores de mama luminales A, con Her2 sobreexpresado y triple negativos con un enfoque proteómico global. Las proteínas identificadas se compararon con la base de datos del proyecto HUPO Plasma y el trabajo de Mann para asegurarse de que estas proteínas pudieran detectarse en la sangre de los pacientes. Entonces, las proteínas identificadas se compararon con la base de datos de cáncer de mama y se identificaron más de cien proteínas secretadas que nunca se sometieron a prueba en patología mamaria.

Entre estas proteínas, en el contexto de la presente invención se han identificado más específicamente OLFM4, NENF y DSP como nuevos biomarcadores de cáncer de mama, fácilmente detectables en sueros de pacientes que padecen tal cáncer.

De hecho, basándose en el primer estudio realizado por los inventores, la elevación de OLFM4 y NENF, así como la desregulación de DSP, permiten la detección de cáncer de mama. En el caso de la prevención del cáncer de mama (detección de tumores pequeños), parece que debe buscarse preferentemente una elevación de OLFM4 y una disminución de DSP (<600 pg/ml). En el caso de la monitorización del cáncer de mama, debe buscarse una elevación de DSP (> 1800 pg/ml) para identificar una recidiva de cáncer de mama.

El segundo estudio lo llevaron a cabo los inventores en una cohorte más grande de pacientes con tres cohortes de sueros independientes, con el fin de evaluar adicionalmente los biomarcadores de cáncer de mama NENF y OLFM4. Cuando se evaluó la precisión de diagnóstico de la combinación de ambos biomarcadores para distinguir a los pacientes con cáncer de mama de los controles sanos utilizando un criterio de ROC generalizado, se observó una precisión de diagnóstico global muy significativa. Al reunir los 3 conjuntos de sueros, se obtuvo un AUC del 91% con una sensibilidad del 82% y una especificidad del 87%. De manera muy interesante, estos valores permanecieron elevados en pacientes con cáncer de mama en estadio temprano (tamaño tumoral inferior a 1 cm), y la sensibilidad y especificidad difirieron ligeramente entre las diferentes cohortes de sueros. Las capacidades de diagnóstico de OLFM4 y NENF séricos fueron similares en todas las cohortes de pacientes con cáncer de mama y cáncer de mama en estadio temprano. Además, la asociación de los dos marcadores fue realmente beneficiosa para la predicción del diagnóstico. En el conjunto de cáncer de mama, 421 pacientes (70%) fueron positivos para OLFM4 sola y 342 pacientes (53%) fueron positivos para NENF sola. Cuando se combinaron ambos marcadores, se detectaron 556 pacientes (85%). De la misma manera, para la cohorte de cáncer de mama en estadio temprano, el 78% de los pacientes fueron positivos para OLFM4 y el 53% para NENF sola. Este número se elevó hasta el 82% cuando se combinaron las concentraciones de OLFM4 y NENF.

Por tanto, los presentes datos indican que los biomarcadores séricos OLFM4, NENF y DSP, y más preferentemente OLFM4 y NENF, pueden utilizarse para detectar el cáncer de mama, especialmente en el diagnóstico en estadio temprano. Puesto que esas proteínas se secretan, su expresión puede medirse simplemente de manera fiable, sin la necesidad de técnicas invasivas con el fin de detectar el cáncer de mama de manera temprana, o bien combinada con una mamografía para aumentar la tasa de cáncer oculto detectado, o bien espaciando el número de mamografías en la monitorización del paciente.

El posible beneficio de una metodología de detección diseñada para identificar el cáncer de mama en estadio temprano es claro. Se ha demostrado que la mamografía es la herramienta de cribado más eficaz para detectar el cáncer de mama de manera temprana y para salvar vidas. Sin embargo, la mamografía presenta limitaciones intrínsecas que pueden ser difíciles de superar y su sensibilidad oscila entre el 63 y el 87%, dependiendo de la edad, la densidad mamaria y las características del tumor. Por tanto, se necesitan pruebas complementarias para detectar mujeres con cáncer de mama y para aumentar la sensibilidad de diagnóstico de los enfoques de cribado. Los biomarcadores séricos pueden ser útiles para aumentar el valor predictivo positivo de las lesiones mamográficas, disminuyendo de ese modo el número de mujeres que se someten a biopsias innecesarias. Además, asimismo pueden utilizarse biomarcadores para seleccionar casos de técnicas de diagnóstico más sensibles, tales como obtención de imágenes por resonancia magnética. Otra aplicación significativa será la monitorización de mujeres jóvenes "en riesgo". Las mujeres con un alto riesgo de desarrollar cáncer de mama son esencialmente las portadoras de mutaciones en los genes BRCA1 y BRCA2 o con una alta probabilidad de una predisposición hereditaria al cáncer de mama. En consecuencia, el cribado se refuerza en mujeres que portan estos genes mutados: deben someterse a exámenes clínicos dos veces al año y a pruebas de obtención de imágenes tan pronto como alcancen la edad de 30 años.

Desde 2004, varios ensayos prospectivos han comparado las técnicas de obtención de imágenes mamarias en mujeres que expresan estas mutaciones o que presentan un alto riesgo de cáncer de mama. Todos los ensayos encontraron que la IRM era superior a las otras técnicas para la detección temprana del cáncer de mama. Sardanelli *et al.* (2007) analizaron los resultados de 5 estudios prospectivos que evaluaron el rendimiento de la

mamografía, la ecografía y la IRM de mama: la sensibilidad de la IRM de mama es alta (el 80% frente al 40% para mamografía), pero su valor predictivo de IRM es bajo, es decir, solo del 53%, lo que indica que el número de biopsias para resultados falsos positivos aumentó con la prueba. La mastectomía bilateral profiláctica puede reducir el riesgo de aparición de cáncer de mama en del 85 al 100%, pero no hay datos que demuestren los beneficios de un procedimiento de este tipo en cuanto a supervivencia global en comparación con la monitorización y el cribado temprano. Los estudios tampoco encontraron beneficios de supervivencia cuando se realizó mastectomía profiláctica en un estadio temprano, tan pronto como a los 25 años.

Una monitorización simple que implica dos exámenes clínicos anuales, IRM de mama y una mamografía es una alternativa asociada con casi ninguna complicación. El procedimiento no reduce el riesgo de cáncer, sino que se dirige a detectar y tratar lesiones precancerosas y cancerosas lo antes posible.

Los datos presentados en la presente memoria demuestran claramente que la utilización de los biomarcadores séricos OLFM4 y/o NENF, potencialmente junto con DSP, aumentaría este valor predictivo positivo mientras se mantiene una buena sensibilidad.

Bibliografía

Independent UK Panel on Breast Cancer Screening (2012). *The Lancet*; 380(9855): 1778 - 1786.

Liu W., Chen L., Zhu J., y Rodgers G.P. (2006). *Exp. Cell Res.*; 312: 1785-1797.

Zhang X., Huang Q., Yang Z., Li Y., and Li C.Y. (2004). *Cancer Res.*; 64:2474-2481.

Kobayashi D., Koshida S., Moriai R., Tsuji N., y Watanabe N. (2007). *Cancer Sci.*; 98(3): 334-340.

Kimura I., Nakayama Y., Zhao Y., Konishi M., y Itoh N. (2013). *Front Neurosci.*; 25;7:111.

Leung C.L., Green K.J., y Liem R.K. (2002). *Trends Cell Biol.*; 12: 37-45.

Allen E., Yu Q.C., y Fuchs E. (1996). *J. Cell Biol.*; 133: 1367-1382.

Wan H., South A.P., y Hart I.R. (2007). *Exp. Cell Res.*; 313: 2336-2344.

Rickelt S., Winter-Simanowski S., Noffz E., Kuhn C., y Franke W.W. (2009). *Int.J. Cancer*, 125: 2036-2048.

Jonkman M.F., Pasmooij A.M., Pasmans S.G., Van Den Berg M.P., Ter Horst H.J., Timmer A., y Pas H.H. (2005). *Am. J. Hum. Genet.*; 77: 653-660.

Kowalczyk A.P., Bornslaeger E.A., Borgwardt J.E., Palka H.L., Dhaliwal A.S., Corcoran C.M., Denning M.F., y Green K.J. (1997). *J. Cell Biol.*; 139: 773-784.

Dowdy S.M., y Wearden S. (1983). *Statistics for Research*, John Wiley & Sons, Nueva York.

Hou H.W., Warkiani M.E., Khoo B.L., Li Z.R., Soo R.A., Tan D.S., Lim W.T., Han J., Bhagat A.A., Lim C.T. (2013). *Sci. Rep.*; 3:1259.

Reeves J.R. y Bartlett J.M.S. (2000). *Methods in Molecular Medicine*; vol. 39, capítulo 51, p. 471-483.

Schena M. (2005). *Protein microarrays*; Jones and Bartlett Learning.

Hamelinck D., Zhou H., Li L., Verweij C., Dillon D., Feng Z., Costa J., y Haab B.B. (2005). *Mol. Cell Proteomics*; 4(6):773-84.

Köhler G. y Milstein C. (1975). *Nature*; 256 (5517): 495-7.

Kozbor D., Roder J.C. (1983). *Immunology Today*; vol. 4: p. 72-79.

Roder J.C., Cole S.P., y Kozbor D. (1986). *Methods Enzymol.*; 121:140-167.

Huse W.D., Sastry L., Iverson S.A., Kang A.S., Altling-Mees M., Burton D.R., Benkovic S.J., y Lerner R.A. (1989). *Science*; 246:1275-1281.

Weigelt B. y Bissell M.J. (2008). *Semin Cancer Biol.*; 18(5): 311-321.

Kenny P.A., Lee G.Y., Myers C.A., Neve R.M., Semeiks J.R., Spellman P.T., Lorenz K., Lee E.H., Barcellos-

- Hoff M.H., Petersen O.W., Gray J.W., y Bissell M.J. (2007). *Mol Oncol.*; 1(1):84-96.
- Li Q., Chow A.B., y Mattingly R.R. (2010). *J Pharmacol Exp Ther.*; 332(3): 821-828.
- 5 Mitchell P. (2002). *Nature Biotech*; 20: 225-229.
- Haab B.B. (2005). *Mol Cell Proteomics* ;4(4):377-83.
- Eckel-Passow J.E., Hoering A., Therneau T.M., y Ghobrial I. (2005). *Cancer Res*; 15;65(8):2985-9.
- 10 Kingsmore S.F. (2006). *Nat Rev Drug Discov.*; 5(4):310-20.
- Chandra H., Reddy P.J., y Srivastava S. (2011). *Expert Rev Proteomics*; 8(1):61-79.
- 15 Schenk S., Schoenhals G.J., de Souza G., y Mann (2008). *BMC Med Genomics*; 15;1:41.
- Cao Z., Yende S., Kellum J.A., y Robinson R.A.S. (2013). *Int J Proteomics*: 2013:654356.
- 20 Imbalzano K.M., Tatarkova I., Imbalzano A.N., y Nickerson J.A. (2009). *Cancer Cell Int.* 9:7. doi: 10.1186/1475-2867-9-7.
- Wisniewski JR., Zougman A., Nagaraj N., y Mann M. (2009). *Nat Methods*; 6, 359-362.
- 25 Shilov I. V., Seymour S. L., Patel A. A., Loboda A., Tang W. H., Keating S. P., Hunter C. L., Nuwaysir L. M., y Schaeffer D. A. (2007). *Mol Cell Proteomics*; 6, 1638-1655.
- Schwacke, J. H., Hill, E. G., Krug, E. L., Comte-Walters, S., y Schey, K. L. (2009). *BMC Bioinformatics*; 10, 342.
- 30 Grant J.E., Bradshaw A.D., Schwacke, J.H, Baicu, C.F., Zile, M.R., y Schey, K.L. (2009). *J Proteome Res.*; 9, 4252-63.
- Konishi H., Karakas B., y Abde M. Abukhdeir A.M. (2007). *Cancer Res.*; 67:8460-8467.
- 35 Sardanelli F., y Podo F. (2007). *Eur Radiol.*; 17:873-87.
- Han K., *et al.* (2012). *BMC Cancer*, 12(274): 1-12.
- Koshida S., *et. al.* (2007). *Cancer Sci*, 98(3): 315 - 320.

Listado de secuencias

<110> INSTITUT DE CANCEROLOGIE DE L'OUEST UNIVERSITE D'ANGERS INSTITUT NATIONAL DE LA
SANTÉ ET DE LA RECHERCHE MÉDICALE (INSERM)

45 <120> Olfactomedina-4, neudesina y desmoplaquina como nuevos biomarcadores de cáncer de mama.

<130> 366580 D33046

50 <150> EP 13306604.3
<151> 2013-11-22

<160> 4

55 <170> PatentIn version 3.5

<210> 1
<211> 510
<212> PRT

60 <213> homo sapiens

<220>
<221> MISC_FEATURE
<223> secuencia de aminoácidos de olfactomedina 4

65

ES 2 739 426 T3

<400> 1

Met	Arg	Pro	Gly	Leu	Ser	Phe	Leu	Leu	Ala	Leu	Leu	Phe	Phe	Leu	Gly
1				5					10					15	
Gln	Ala	Ala	Gly	Asp	Leu	Gly	Asp	Val	Gly	Pro	Pro	Ile	Pro	Ser	Pro
			20					25					30		
Gly	Phe	Ser	Ser	Phe	Pro	Gly	Val	Asp	Ser	Ser	Ser	Ser	Phe	Ser	Ser
		35					40					45			
Ser	Ser	Arg	Ser	Gly	Ser	Ser	Ser	Ser	Arg	Ser	Leu	Gly	Ser	Gly	Gly
	50					55					60				
Ser	Val	Ser	Gln	Leu	Phe	Ser	Asn	Phe	Thr	Gly	Ser	Val	Asp	Asp	Arg
65					70					75					80
Gly	Thr	Cys	Gln	Cys	Ser	Val	Ser	Leu	Pro	Asp	Thr	Thr	Phe	Pro	Val
			85						90					95	
Asp	Arg	Val	Glu	Arg	Leu	Glu	Phe	Thr	Ala	His	Val	Leu	Ser	Gln	Lys
		100						105					110		
Phe	Glu	Lys	Glu	Leu	Ser	Lys	Val	Arg	Glu	Tyr	Val	Gln	Leu	Ile	Ser
		115					120					125			

ES 2 739 426 T3

Val	Tyr	Glu	Lys	Lys	Leu	Leu	Asn	Leu	Thr	Val	Arg	Ile	Asp	Ile	Met	130	135	140
Glu	Lys	Asp	Thr	Ile	Ser	Tyr	Thr	Glu	Leu	Asp	Phe	Glu	Leu	Ile	Lys	145	150	155
Val	Glu	Val	Lys	Glu	Met	Glu	Lys	Leu	Val	Ile	Gln	Leu	Lys	Glu	Ser	165	170	175
Phe	Gly	Gly	Ser	Ser	Glu	Ile	Val	Asp	Gln	Leu	Glu	Val	Glu	Ile	Arg	180	185	190
Asn	Met	Thr	Leu	Leu	Val	Glu	Lys	Leu	Glu	Thr	Leu	Asp	Lys	Asn	Asn	195	200	205
Val	Leu	Ala	Ile	Arg	Arg	Glu	Ile	Val	Ala	Leu	Lys	Thr	Lys	Leu	Lys	210	215	220
Glu	Cys	Glu	Ala	Ser	Lys	Asp	Gln	Asn	Thr	Pro	Val	Val	His	Pro	Pro	225	230	235
Pro	Thr	Pro	Gly	Ser	Cys	Gly	His	Gly	Gly	Val	Val	Asn	Ile	Ser	Lys	245	250	255
Pro	Ser	Val	Val	Gln	Leu	Asn	Trp	Arg	Gly	Phe	Ser	Tyr	Leu	Tyr	Gly	260	265	270
Ala	Trp	Gly	Arg	Asp	Tyr	Ser	Pro	Gln	His	Pro	Asn	Lys	Gly	Leu	Tyr	275	280	285
Trp	Val	Ala	Pro	Leu	Asn	Thr	Asp	Gly	Arg	Leu	Leu	Glu	Tyr	Tyr	Arg	290	295	300
Leu	Tyr	Asn	Thr	Leu	Asp	Asp	Leu	Leu	Leu	Tyr	Ile	Asn	Ala	Arg	Glu	305	310	315
Leu	Arg	Ile	Thr	Tyr	Gly	Gln	Gly	Ser	Gly	Thr	Ala	Val	Tyr	Asn	Asn	325	330	335
Asn	Met	Tyr	Val	Asn	Met	Tyr	Asn	Thr	Gly	Asn	Ile	Ala	Arg	Val	Asn	340	345	350
Leu	Thr	Thr	Asn	Thr	Ile	Ala	Val	Thr	Gln	Thr	Leu	Pro	Asn	Ala	Ala	355	360	365
Tyr	Asn	Asn	Arg	Phe	Ser	Tyr	Ala	Asn	Val	Ala	Trp	Gln	Asp	Ile	Asp	370	375	380

ES 2 739 426 T3

Phe Ala Val Asp Glu Asn Gly Leu Trp Val Ile Tyr Ser Thr Glu Ala
385 390 395 400

Ser Thr Gly Asn Met Val Ile Ser Lys Leu Asn Asp Thr Thr Leu Gln
405 410 415

Val Leu Asn Thr Trp Tyr Thr Lys Gln Tyr Lys Pro Ser Ala Ser Asn
420 425 430

Ala Phe Met Val Cys Gly Val Leu Tyr Ala Thr Arg Thr Met Asn Thr
435 440 445

Arg Thr Glu Glu Ile Phe Tyr Tyr Tyr Asp Thr Asn Thr Gly Lys Glu
450 455 460

Gly Lys Leu Asp Ile Val Met His Lys Met Gln Glu Lys Val Gln Ser
465 470 475 480

Ile Asn Tyr Asn Pro Phe Asp Gln Lys Leu Tyr Val Tyr Asn Asp Gly
485 490 495

Tyr Leu Leu Asn Tyr Asp Leu Ser Val Leu Gln Lys Pro Gln
500 505 510

<210> 2

<211> 172

<212> PRT

<213> homo sapiens

<220>

<221> MISC_FEATURE

<223> secuencia de aminoácidos de neudesina

<400> 2

Met Val Gly Pro Ala Pro Arg Arg Arg Leu Arg Pro Leu Ala Ala Leu
1 5 10 15

Ala Leu Val Leu Ala Leu Ala Pro Gly Leu Pro Thr Ala Arg Ala Gly
20 25 30

Gln Thr Pro Arg Pro Ala Glu Arg Gly Pro Pro Val Arg Leu Phe Thr
35 40 45

Glu Glu Glu Leu Ala Arg Tyr Gly Gly Glu Glu Glu Asp Gln Pro Ile
50 55 60

Tyr Leu Ala Val Lys Gly Val Val Phe Asp Val Thr Ser Gly Lys Glu
65 70 75 80

ES 2 739 426 T3

Phe Tyr Gly Arg Gly Ala Pro Tyr Asn Ala Leu Thr Gly Lys Asp Ser
85 90 95

Thr Arg Gly Val Ala Lys Met Ser Leu Asp Pro Ala Asp Leu Thr His
100 105 110

Asp Thr Thr Gly Leu Thr Ala Lys Glu Leu Glu Ala Leu Asp Glu Val
115 120 125

Phe Thr Lys Val Tyr Lys Ala Lys Tyr Pro Ile Val Gly Tyr Thr Ala
130 135 140

Arg Arg Ile Leu Asn Glu Asp Gly Ser Pro Asn Leu Asp Phe Lys Pro
145 150 155 160

Glu Asp Gln Pro His Phe Asp Ile Lys Asp Glu Phe
165 170

<210> 3

<211> 2871

<212> PRT

<213> homo sapiens

<220>

<221> MISC_FEATURE

<223> sequencia de aminoácidos de isoforma 1 de desmoplaquina

<400> 3

Met Ser Cys Asn Gly Gly Ser His Pro Arg Ile Asn Thr Leu Gly Arg
1 5 10 15

Met Ile Arg Ala Glu Ser Gly Pro Asp Leu Arg Tyr Glu Val Thr Ser
20 25 30

Gly Gly Gly Gly Thr Ser Arg Met Tyr Tyr Ser Arg Arg Gly Val Ile
35 40 45

Thr Asp Gln Asn Ser Asp Gly Tyr Cys Gln Thr Gly Thr Met Ser Arg
50 55 60

His Gln Asn Gln Asn Thr Ile Gln Glu Leu Leu Gln Asn Cys Ser Asp
65 70 75 80

Cys Leu Met Arg Ala Glu Leu Ile Val Gln Pro Glu Leu Lys Tyr Gly
85 90 95

Asp Gly Ile Gln Leu Thr Arg Ser Arg Glu Leu Asp Glu Cys Phe Ala
100 105 110

ES 2 739 426 T3

Gln Ala Asn Asp Gln Met Glu Ile Leu Asp Ser Leu Ile Arg Glu Met
 115 120 125
 Arg Gln Met Gly Gln Pro Cys Asp Ala Tyr Gln Lys Arg Leu Leu Gln
 130 135 140
 Leu Gln Glu Gln Met Arg Ala Leu Tyr Lys Ala Ile Ser Val Pro Arg
 145 150 155 160
 Val Arg Arg Ala Ser Ser Lys Gly Gly Gly Tyr Thr Cys Gln Ser
 165 170 175
 Gly Ser Gly Trp Asp Glu Phe Thr Lys His Val Thr Ser Glu Cys Leu
 180 185 190
 Gly Trp Met Arg Gln Gln Arg Ala Glu Met Asp Met Val Ala Trp Gly
 195 200 205
 Val Asp Leu Ala Ser Val Glu Gln His Ile Asn Ser His Arg Gly Ile
 210 215 220
 His Asn Ser Ile Gly Asp Tyr Arg Trp Gln Leu Asp Lys Ile Lys Ala
 225 230 235 240
 Asp Leu Arg Glu Lys Ser Ala Ile Tyr Gln Leu Glu Glu Glu Tyr Glu
 245 250 255
 Asn Leu Leu Lys Ala Ser Phe Glu Arg Met Asp His Leu Arg Gln Leu
 260 265 270
 Gln Asn Ile Ile Gln Ala Thr Ser Arg Glu Ile Met Trp Ile Asn Asp
 275 280 285
 Cys Glu Glu Glu Glu Leu Leu Tyr Asp Trp Ser Asp Lys Asn Thr Asn
 290 295 300
 Ile Ala Gln Lys Gln Glu Ala Phe Ser Ile Arg Met Ser Gln Leu Glu
 305 310 315 320
 Val Lys Glu Lys Glu Leu Asn Lys Leu Lys Gln Glu Ser Asp Gln Leu
 325 330 335
 Val Leu Asn Gln His Pro Ala Ser Asp Lys Ile Glu Ala Tyr Met Asp
 340 345 350

ES 2 739 426 T3

Thr Leu Gln Thr Gln Trp Ser Trp Ile Leu Gln Ile Thr Lys Cys Ile
 355 360 365
 Asp Val His Leu Lys Glu Asn Ala Ala Tyr Phe Gln Phe Phe Glu Glu
 370 375 380
 Ala Gln Ser Thr Glu Ala Tyr Leu Lys Gly Leu Gln Asp Ser Ile Arg
 385 390 395 400
 Lys Lys Tyr Pro Cys Asp Lys Asn Met Pro Leu Gln His Leu Leu Glu
 405 410 415
 Gln Ile Lys Glu Leu Glu Lys Glu Arg Glu Lys Ile Leu Glu Tyr Lys
 420 425 430
 Arg Gln Val Gln Asn Leu Val Asn Lys Ser Lys Lys Ile Val Gln Leu
 435 440 445
 Lys Pro Arg Asn Pro Asp Tyr Arg Ser Asn Lys Pro Ile Ile Leu Arg
 450 455 460
 Ala Leu Cys Asp Tyr Lys Gln Asp Gln Lys Ile Val His Lys Gly Asp
 465 470 475 480
 Glu Cys Ile Leu Lys Asp Asn Asn Glu Arg Ser Lys Trp Tyr Val Thr
 485 490 495
 Gly Pro Gly Gly Val Asp Met Leu Val Pro Ser Val Gly Leu Ile Ile
 500 505 510
 Pro Pro Pro Asn Pro Leu Ala Val Asp Leu Ser Cys Lys Ile Glu Gln
 515 520 525
 Tyr Tyr Glu Ala Ile Leu Ala Leu Trp Asn Gln Leu Tyr Ile Asn Met
 530 535 540
 Lys Ser Leu Val Ser Trp His Tyr Cys Met Ile Asp Ile Glu Lys Ile
 545 550 555 560
 Arg Ala Met Thr Ile Ala Lys Leu Lys Thr Met Arg Gln Glu Asp Tyr
 565 570 575
 Met Lys Thr Ile Ala Asp Leu Glu Leu His Tyr Gln Glu Phe Ile Arg
 580 585 590
 Asn Ser Gln Gly Ser Glu Met Phe Gly Asp Asp Asp Lys Arg Lys Ile
 595 600 605

ES 2 739 426 T3

Gln Ser Gln Phe Thr Asp Ala Gln Lys His Tyr Gln Thr Leu Val Ile
 610 615 620
 Gln Leu Pro Gly Tyr Pro Gln His Gln Thr Val Thr Thr Thr Glu Ile
 625 630 635 640
 Thr His His Gly Thr Cys Gln Asp Val Asn His Asn Lys Val Ile Glu
 645 650 655
 Thr Asn Arg Glu Asn Asp Lys Gln Glu Thr Trp Met Leu Met Glu Leu
 660 665 670
 Gln Lys Ile Arg Arg Gln Ile Glu His Cys Glu Gly Arg Met Thr Leu
 675 680 685
 Lys Asn Leu Pro Leu Ala Asp Gln Gly Ser Ser His His Ile Thr Val
 690 695 700
 Lys Ile Asn Glu Leu Lys Ser Val Gln Asn Asp Ser Gln Ala Ile Ala
 705 710 715 720
 Glu Val Leu Asn Gln Leu Lys Asp Met Leu Ala Asn Phe Arg Gly Ser
 725 730 735
 Glu Lys Tyr Cys Tyr Leu Gln Asn Glu Val Phe Gly Leu Phe Gln Lys
 740 745 750
 Leu Glu Asn Ile Asn Gly Val Thr Asp Gly Tyr Leu Asn Ser Leu Cys
 755 760 765
 Thr Val Arg Ala Leu Leu Gln Ala Ile Leu Gln Thr Glu Asp Met Leu
 770 775 780
 Lys Val Tyr Glu Ala Arg Leu Thr Glu Glu Glu Thr Val Cys Leu Asp
 785 790 795 800
 Leu Asp Lys Val Glu Ala Tyr Arg Cys Gly Leu Lys Lys Ile Lys Asn
 805 810 815
 Asp Leu Asn Leu Lys Lys Ser Leu Leu Ala Thr Met Lys Thr Glu Leu
 820 825 830
 Gln Lys Ala Gln Gln Ile His Ser Gln Thr Ser Gln Gln Tyr Pro Leu
 835 840 845
 Tyr Asp Leu Asp Leu Gly Lys Phe Gly Glu Lys Val Thr Gln Leu Thr
 850 855 860

ES 2 739 426 T3

Asp Arg Trp Gln Arg Ile Asp Lys Gln Ile Asp Phe Arg Leu Trp Asp
 865 870 875 880
 Leu Glu Lys Gln Ile Lys Gln Leu Arg Asn Tyr Arg Asp Asn Tyr Gln
 885 890 895
 Ala Phe Cys Lys Trp Leu Tyr Asp Ala Lys Arg Arg Gln Asp Ser Leu
 900 905 910
 Glu Ser Met Lys Phe Gly Asp Ser Asn Thr Val Met Arg Phe Leu Asn
 915 920 925
 Glu Gln Lys Asn Leu His Ser Glu Ile Ser Gly Lys Arg Asp Lys Ser
 930 935 940
 Glu Glu Val Gln Lys Ile Ala Glu Leu Cys Ala Asn Ser Ile Lys Asp
 945 950 955 960
 Tyr Glu Leu Gln Leu Ala Ser Tyr Thr Ser Gly Leu Glu Thr Leu Leu
 965 970 975
 Asn Ile Pro Ile Lys Arg Thr Met Ile Gln Ser Pro Ser Gly Val Ile
 980 985 990
 Leu Gln Glu Ala Ala Asp Val His Ala Arg Tyr Ile Glu Leu Leu Thr
 995 1000 1005
 Arg Ser Gly Asp Tyr Tyr Arg Phe Leu Ser Glu Met Leu Lys Ser
 1010 1015 1020
 Leu Glu Asp Leu Lys Leu Lys Asn Thr Lys Ile Glu Val Leu Glu
 1025 1030 1035
 Glu Glu Leu Arg Leu Ala Arg Asp Ala Asn Ser Glu Asn Cys Asn
 1040 1045 1050
 Lys Asn Lys Phe Leu Asp Gln Asn Leu Gln Lys Tyr Gln Ala Glu
 1055 1060 1065
 Cys Ser Gln Phe Lys Ala Lys Leu Ala Ser Leu Glu Glu Leu Lys
 1070 1075 1080
 Arg Gln Ala Glu Leu Asp Gly Lys Ser Ala Lys Gln Asn Leu Asp
 1085 1090 1095
 Lys Cys Tyr Gly Gln Ile Lys Glu Leu Asn Glu Lys Ile Thr Arg
 1100 1105 1110

ES 2 739 426 T3

Leu Thr	Tyr Glu Ile Glu Asp	Glu Lys Arg Arg Arg	Lys Ser Val
1115	1120	1125	
Glu Asp	Arg Phe Asp Gln Gln	Lys Asn Asp Tyr Asp	Gln Leu Gln
1130	1135	1140	
Lys Ala	Arg Gln Cys Glu Lys	Glu Asn Leu Gly Trp	Gln Lys Leu
1145	1150	1155	
Glu Ser	Glu Lys Ala Ile Lys	Glu Lys Glu Tyr Glu	Ile Glu Arg
1160	1165	1170	
Leu Arg	Val Leu Leu Gln Glu	Glu Gly Thr Arg Lys	Arg Glu Tyr
1175	1180	1185	
Glu Asn	Glu Leu Ala Lys Val	Arg Asn His Tyr Asn	Glu Glu Met
1190	1195	1200	
Ser Asn	Leu Arg Asn Lys Tyr	Glu Thr Glu Ile Asn	Ile Thr Lys
1205	1210	1215	
Thr Thr	Ile Lys Glu Ile Ser	Met Gln Lys Glu Asp	Asp Ser Lys
1220	1225	1230	
Asn Leu	Arg Asn Gln Leu Asp	Arg Leu Ser Arg Glu	Asn Arg Asp
1235	1240	1245	
Leu Lys	Asp Glu Ile Val Arg	Leu Asn Asp Ser Ile	Leu Gln Ala
1250	1255	1260	
Thr Glu	Gln Arg Arg Arg Ala	Glu Glu Asn Ala Leu	Gln Gln Lys
1265	1270	1275	
Ala Cys	Gly Ser Glu Ile Met	Gln Lys Lys Gln His	Leu Glu Ile
1280	1285	1290	
Glu Leu	Lys Gln Val Met Gln	Gln Arg Ser Glu Asp	Asn Ala Arg
1295	1300	1305	
His Lys	Gln Ser Leu Glu Glu	Ala Ala Lys Thr Ile	Gln Asp Lys
1310	1315	1320	
Asn Lys	Glu Ile Glu Arg Leu	Lys Ala Glu Phe Gln	Glu Glu Ala
1325	1330	1335	

ES 2 739 426 T3

Lys Arg	Arg Trp Glu Tyr Glu	Asn Glu Leu Ser Lys	Val Arg Asn
1340	1345	1350	
Asn Tyr	Asp Glu Glu Ile Ile	Ser Leu Lys Asn Gln	Phe Glu Thr
1355	1360	1365	
Glu Ile	Asn Ile Thr Lys Thr	Thr Ile His Gln Leu	Thr Met Gln
1370	1375	1380	
Lys Glu	Glu Asp Thr Ser Gly	Tyr Arg Ala Gln Ile	Asp Asn Leu
1385	1390	1395	
Thr Arg	Glu Asn Arg Ser Leu	Ser Glu Glu Ile Lys	Arg Leu Lys
1400	1405	1410	
Asn Thr	Leu Thr Gln Thr Thr	Glu Asn Leu Arg Arg	Val Glu Glu
1415	1420	1425	
Asp Ile	Gln Gln Gln Lys Ala	Thr Gly Ser Glu Val	Ser Gln Arg
1430	1435	1440	
Lys Gln	Gln Leu Glu Val Glu	Leu Arg Gln Val Thr	Gln Met Arg
1445	1450	1455	
Thr Glu	Glu Ser Val Arg Tyr	Lys Gln Ser Leu Asp	Asp Ala Ala
1460	1465	1470	
Lys Thr	Ile Gln Asp Lys Asn	Lys Glu Ile Glu Arg	Leu Lys Gln
1475	1480	1485	
Leu Ile	Asp Lys Glu Thr Asn	Asp Arg Lys Cys Leu	Glu Asp Glu
1490	1495	1500	
Asn Ala	Arg Leu Gln Arg Val	Gln Tyr Asp Leu Gln	Lys Ala Asn
1505	1510	1515	
Ser Ser	Ala Thr Glu Thr Ile	Asn Lys Leu Lys Val	Gln Glu Gln
1520	1525	1530	
Glu Leu	Thr Arg Leu Arg Ile	Asp Tyr Glu Arg Val	Ser Gln Glu
1535	1540	1545	
Arg Thr	Val Lys Asp Gln Asp	Ile Thr Arg Phe Gln	Asn Ser Leu
1550	1555	1560	
Lys Glu	Leu Gln Leu Gln Lys	Gln Lys Val Glu Glu	Glu Leu Asn
1565	1570	1575	

ES 2 739 426 T3

Arg Leu 1580	Lys Arg Thr Ala Ser 1585	Glu Asp Ser Cys Lys 1590	Arg Lys Lys
Leu Glu 1595	Glu Glu Leu Glu Gly 1600	Met Arg Arg Ser Leu 1605	Lys Glu Gln
Ala Ile 1610	Lys Ile Thr Asn Leu 1615	Thr Gln Gln Leu Glu 1620	Gln Ala Ser
Ile Val 1625	Lys Lys Arg Ser Glu 1630	Asp Asp Leu Arg Gln 1635	Gln Arg Asp
Val Leu 1640	Asp Gly His Leu Arg 1645	Glu Lys Gln Arg Thr 1650	Gln Glu Glu
Leu Arg 1655	Arg Leu Ser Ser Glu 1660	Val Glu Ala Leu Arg 1665	Arg Gln Leu
Leu Gln 1670	Glu Gln Glu Ser Val 1675	Lys Gln Ala His Leu 1680	Arg Asn Glu
His Phe 1685	Gln Lys Ala Ile Glu 1690	Asp Lys Ser Arg Ser 1695	Leu Asn Glu
Ser Lys 1700	Ile Glu Ile Glu Arg 1705	Leu Gln Ser Leu Thr 1710	Glu Asn Leu
Thr Lys 1715	Glu His Leu Met Leu 1720	Glu Glu Glu Leu Arg 1725	Asn Leu Arg
Leu Glu 1730	Tyr Asp Asp Leu Arg 1735	Arg Gly Arg Ser Glu 1740	Ala Asp Ser
Asp Lys 1745	Asn Ala Thr Ile Leu 1750	Glu Leu Arg Ser Gln 1755	Leu Gln Ile
Ser Asn 1760	Asn Arg Thr Leu Glu 1765	Leu Gln Gly Leu Ile 1770	Asn Asp Leu
Gln Arg 1775	Glu Arg Glu Asn Leu 1780	Arg Gln Glu Ile Glu 1785	Lys Phe Gln
Lys Gln 1790	Ala Leu Glu Ala Ser 1795	Asn Arg Ile Gln Glu 1800	Ser Lys Asn
Gln Cys 1805	Thr Gln Val Val Gln 1810	Glu Arg Glu Ser Leu 1815	Leu Val Lys

ES 2 739 426 T3

Ile Lys	Val Leu Glu Gln Asp	Lys Ala Arg Leu Gln	Arg Leu Glu
1820	1825	1830	
Asp Glu	Leu Asn Arg Ala Lys	Ser Thr Leu Glu Ala	Glu Thr Arg
1835	1840	1845	
Val Lys	Gln Arg Leu Glu Cys	Glu Lys Gln Gln Ile	Gln Asn Asp
1850	1855	1860	
Leu Asn	Gln Trp Lys Thr Gln	Tyr Ser Arg Lys Glu	Glu Ala Ile
1865	1870	1875	
Arg Lys	Ile Glu Ser Glu Arg	Glu Lys Ser Glu Arg	Glu Lys Asn
1880	1885	1890	
Ser Leu	Arg Ser Glu Ile Glu	Arg Leu Gln Ala Glu	Ile Lys Arg
1895	1900	1905	
Ile Glu	Glu Arg Cys Arg Arg	Lys Leu Glu Asp Ser	Thr Arg Glu
1910	1915	1920	
Thr Gln	Ser Gln Leu Glu Thr	Glu Arg Ser Arg Tyr	Gln Arg Glu
1925	1930	1935	
Ile Asp	Lys Leu Arg Gln Arg	Pro Tyr Gly Ser His	Arg Glu Thr
1940	1945	1950	
Gln Thr	Glu Cys Glu Trp Thr	Val Asp Thr Ser Lys	Leu Val Phe
1955	1960	1965	
Asp Gly	Leu Arg Lys Lys Val	Thr Ala Met Gln Leu	Tyr Glu Cys
1970	1975	1980	
Gln Leu	Ile Asp Lys Thr Thr	Leu Asp Lys Leu Leu	Lys Gly Lys
1985	1990	1995	
Lys Ser	Val Glu Glu Val Ala	Ser Glu Ile Gln Pro	Phe Leu Arg
2000	2005	2010	
Gly Ala	Gly Ser Ile Ala Gly	Ala Ser Ala Ser Pro	Lys Glu Lys
2015	2020	2025	
Tyr Ser	Leu Val Glu Ala Lys	Arg Lys Lys Leu Ile	Ser Pro Glu
2030	2035	2040	
Ser Thr	Val Met Leu Leu Glu	Ala Gln Ala Ala Thr	Gly Gly Ile
2045	2050	2055	

ES 2 739 426 T3

Ile	Asp	Pro	His	Arg	Asn	Glu	Lys	Leu	Thr	Val	Asp	Ser	Ala	Ile
2060						2065					2070			
Ala	Arg	Asp	Leu	Ile	Asp	Phe	Asp	Asp	Arg	Gln	Gln	Ile	Tyr	Ala
2075						2080					2085			
Ala	Glu	Lys	Ala	Ile	Thr	Gly	Phe	Asp	Asp	Pro	Phe	Ser	Gly	Lys
2090						2095					2100			
Thr	Val	Ser	Val	Ser	Glu	Ala	Ile	Lys	Lys	Asn	Leu	Ile	Asp	Arg
2105						2110					2115			
Glu	Thr	Gly	Met	Arg	Leu	Leu	Glu	Ala	Gln	Ile	Ala	Ser	Gly	Gly
2120						2125					2130			
Val	Val	Asp	Pro	Val	Asn	Ser	Val	Phe	Leu	Pro	Lys	Asp	Val	Ala
2135						2140					2145			
Leu	Ala	Arg	Gly	Leu	Ile	Asp	Arg	Asp	Leu	Tyr	Arg	Ser	Leu	Asn
2150						2155					2160			
Asp	Pro	Arg	Asp	Ser	Gln	Lys	Asn	Phe	Val	Asp	Pro	Val	Thr	Lys
2165						2170					2175			
Lys	Lys	Val	Ser	Tyr	Val	Gln	Leu	Lys	Glu	Arg	Cys	Arg	Ile	Glu
2180						2185					2190			
Pro	His	Thr	Gly	Leu	Leu	Leu	Leu	Ser	Val	Gln	Lys	Arg	Ser	Met
2195						2200					2205			
Ser	Phe	Gln	Gly	Ile	Arg	Gln	Pro	Val	Thr	Val	Thr	Glu	Leu	Val
2210						2215					2220			
Asp	Ser	Gly	Ile	Leu	Arg	Pro	Ser	Thr	Val	Asn	Glu	Leu	Glu	Ser
2225						2230					2235			
Gly	Gln	Ile	Ser	Tyr	Asp	Glu	Val	Gly	Glu	Arg	Ile	Lys	Asp	Phe
2240						2245					2250			
Leu	Gln	Gly	Ser	Ser	Cys	Ile	Ala	Gly	Ile	Tyr	Asn	Glu	Thr	Thr
2255						2260					2265			
Lys	Gln	Lys	Leu	Gly	Ile	Tyr	Glu	Ala	Met	Lys	Ile	Gly	Leu	Val
2270						2275					2280			

ES 2 739 426 T3

Arg Pro 2285	Gly Thr Ala Leu 2290	Glu Leu Leu Glu Ala Gln 2295	Ala Ala Thr
Gly Phe 2300	Ile Val Asp Pro Val 2305	Ser Asn Leu Arg Leu 2310	Pro Val Glu
Glu Ala 2315	Tyr Lys Arg Gly Leu 2320	Val Gly Ile Glu Phe 2325	Lys Glu Lys
Leu Leu 2330	Ser Ala Glu Arg Ala 2335	Val Thr Gly Tyr Asn 2340	Asp Pro Glu
Thr Gly 2345	Asn Ile Ile Ser Leu 2350	Phe Gln Ala Met Asn 2355	Lys Glu Leu
Ile Glu 2360	Lys Gly His Gly Ile 2365	Arg Leu Leu Glu Ala 2370	Gln Ile Ala
Thr Gly 2375	Gly Ile Ile Asp Pro 2380	Lys Glu Ser His Arg 2385	Leu Pro Val
Asp Ile 2390	Ala Tyr Lys Arg Gly 2395	Tyr Phe Asn Glu Glu 2400	Leu Ser Glu
Ile Leu 2405	Ser Asp Pro Ser Asp 2410	Asp Thr Lys Gly Phe 2415	Phe Asp Pro
Asn Thr 2420	Glu Glu Asn Leu Thr 2425	Tyr Leu Gln Leu Lys 2430	Glu Arg Cys
Ile Lys 2435	Asp Glu Glu Thr Gly 2440	Leu Cys Leu Leu Pro 2445	Leu Lys Glu
Lys Lys 2450	Lys Gln Val Gln Thr 2455	Ser Gln Lys Asn Thr 2460	Leu Arg Lys
Arg Arg 2465	Val Val Ile Val Asp 2470	Pro Glu Thr Asn Lys 2475	Glu Met Ser
Val Gln 2480	Glu Ala Tyr Lys Lys 2485	Gly Leu Ile Asp Tyr 2490	Glu Thr Phe
Lys Glu 2495	Leu Cys Glu Gln Glu 2500	Cys Glu Trp Glu Glu 2505	Ile Thr Ile
Thr Gly 2510	Ser Asp Gly Ser Thr 2515	Arg Val Val Leu Val 2520	Asp Arg Lys

ES 2 739 426 T3

Thr Gly 2525	Ser Gln Tyr Asp 2530	Gln Asp Ala Ile 2535	Lys Gly Leu
Val Asp 2540	Arg Lys Phe Phe 2545	Gln Tyr Arg Ser 2550	Ser Leu Ser
Leu Thr 2555	Gln Phe Ala Asp 2560	Ile Ser Leu Lys 2565	Gly Val Gly
Thr Ser 2570	Ser Ser Met Gly 2575	Gly Val Ser Asp 2580	Val Phe Ser
Ser Ser 2585	Arg His Glu Ser 2590	Ser Lys Ile Ser 2595	Ile Ser Ser
Val Arg 2600	Asn Leu Thr Ile 2605	Ser Ser Ser Phe 2610	Asp Thr Leu
Glu Glu 2615	Ser Ser Pro Ile 2620	Ala Ile Phe Asp 2625	Glu Asn Leu
Glu Lys 2630	Ile Ser Ile Thr 2635	Gly Ile Glu Arg 2640	Ile Val Asp
Ser Ile 2645	Thr Gly Gln Arg 2650	Leu Glu Ala Gln 2655	Cys Thr Gly
Gly Ile 2660	Ile His Pro Thr 2665	Gly Gln Lys Leu 2670	Leu Gln Asp
Ala Val 2675	Ser Gln Gly Val 2680	Asp Gln Asp Met 2685	Thr Arg Leu
Lys Pro 2690	Ala Gln Lys Ala 2695	Ile Gly Phe Glu 2700	Val Lys Gly
Lys Lys 2705	Lys Met Ser Ala 2710	Glu Ala Val Lys 2715	Lys Trp Leu
Pro Tyr 2720	Glu Ala Gly Gln 2725	Phe Leu Glu Phe 2730	Tyr Leu Thr
Gly Gly 2735	Leu Val Asp Pro 2740	Val His Gly Arg 2745	Ser Thr Glu
Glu Ala 2750	Ile Arg Lys Gly 2755	Ile Asp Gly Arg 2760	Ala Gln Arg

ES 2 739 426 T3

Leu Gln Asp Thr Ser Ser Tyr Ala Lys Ile Leu Thr Cys Pro Lys
2765 2770 2775

Thr Lys Leu Lys Ile Ser Tyr Lys Asp Ala Ile Asn Arg Ser Met
2780 2785 2790

Val Glu Asp Ile Thr Gly Leu Arg Leu Leu Glu Ala Ala Ser Val
2795 2800 2805

Ser Ser Lys Gly Leu Pro Ser Pro Tyr Asn Met Ser Ser Ala Pro
2810 2815 2820

Gly Ser Arg Ser Gly Ser Arg Ser Gly Ser Arg Ser Gly Ser Arg
2825 2830 2835

Ser Gly Ser Arg Ser Gly Ser Arg Arg Gly Ser Phe Asp Ala Thr
2840 2845 2850

Gly Asn Ser Ser Tyr Ser Tyr Ser Tyr Ser Phe Ser Ser Ser Ser
2855 2860 2865

Ile Gly His
2870

<210> 4
<211> 2272
<212> PRT
<213> homo sapiens

<220>
<221> MISC_FEATURE
<223> secuencia de aminoácidos de isoforma de desmoplaquina 2

<400> 4
Met Ser Cys Asn Gly Gly Ser His Pro Arg Ile Asn Thr Leu Gly Arg
1 5 10 15

Met Ile Arg Ala Glu Ser Gly Pro Asp Leu Arg Tyr Glu Val Thr Ser
20 25 30

Gly Gly Gly Gly Thr Ser Arg Met Tyr Tyr Ser Arg Arg Gly Val Ile
35 40 45

Thr Asp Gln Asn Ser Asp Gly Tyr Cys Gln Thr Gly Thr Met Ser Arg
50 55 60

His Gln Asn Gln Asn Thr Ile Gln Glu Leu Leu Gln Asn Cys Ser Asp
65 70 75 80

ES 2 739 426 T3

Cys	Leu	Met	Arg	Ala	Glu	Leu	Ile	Val	Gln	Pro	Glu	Leu	Lys	Tyr	Gly	85	90	95
Asp	Gly	Ile	Gln	Leu	Thr	Arg	Ser	Arg	Glu	Leu	Asp	Glu	Cys	Phe	Ala	100	105	110
Gln	Ala	Asn	Asp	Gln	Met	Glu	Ile	Leu	Asp	Ser	Leu	Ile	Arg	Glu	Met	115	120	125
Arg	Gln	Met	Gly	Gln	Pro	Cys	Asp	Ala	Tyr	Gln	Lys	Arg	Leu	Leu	Gln	130	135	140
Leu	Gln	Glu	Gln	Met	Arg	Ala	Leu	Tyr	Lys	Ala	Ile	Ser	Val	Pro	Arg	145	150	155
Val	Arg	Arg	Ala	Ser	Ser	Lys	Gly	Gly	Gly	Gly	Tyr	Thr	Cys	Gln	Ser	165	170	175
Gly	Ser	Gly	Trp	Asp	Glu	Phe	Thr	Lys	His	Val	Thr	Ser	Glu	Cys	Leu	180	185	190
Gly	Trp	Met	Arg	Gln	Gln	Arg	Ala	Glu	Met	Asp	Met	Val	Ala	Trp	Gly	195	200	205
Val	Asp	Leu	Ala	Ser	Val	Glu	Gln	His	Ile	Asn	Ser	His	Arg	Gly	Ile	210	215	220
His	Asn	Ser	Ile	Gly	Asp	Tyr	Arg	Trp	Gln	Leu	Asp	Lys	Ile	Lys	Ala	225	230	235
Asp	Leu	Arg	Glu	Lys	Ser	Ala	Ile	Tyr	Gln	Leu	Glu	Glu	Glu	Tyr	Glu	245	250	255
Asn	Leu	Leu	Lys	Ala	Ser	Phe	Glu	Arg	Met	Asp	His	Leu	Arg	Gln	Leu	260	265	270
Gln	Asn	Ile	Ile	Gln	Ala	Thr	Ser	Arg	Glu	Ile	Met	Trp	Ile	Asn	Asp	275	280	285
Cys	Glu	Glu	Glu	Glu	Leu	Leu	Tyr	Asp	Trp	Ser	Asp	Lys	Asn	Thr	Asn	290	295	300
Ile	Ala	Gln	Lys	Gln	Glu	Ala	Phe	Ser	Ile	Arg	Met	Ser	Gln	Leu	Glu	305	310	315

ES 2 739 426 T3

Val Lys Glu Lys Glu Leu Asn Lys Leu Lys Gln Glu Ser Asp Gln Leu
 325 330 335
 Val Leu Asn Gln His Pro Ala Ser Asp Lys Ile Glu Ala Tyr Met Asp
 340 345 350
 Thr Leu Gln Thr Gln Trp Ser Trp Ile Leu Gln Ile Thr Lys Cys Ile
 355 360 365
 Asp Val His Leu Lys Glu Asn Ala Ala Tyr Phe Gln Phe Phe Glu Glu
 370 375 380
 Ala Gln Ser Thr Glu Ala Tyr Leu Lys Gly Leu Gln Asp Ser Ile Arg
 385 390 395 400
 Lys Lys Tyr Pro Cys Asp Lys Asn Met Pro Leu Gln His Leu Leu Glu
 405 410 415
 Gln Ile Lys Glu Leu Glu Lys Glu Arg Glu Lys Ile Leu Glu Tyr Lys
 420 425 430
 Arg Gln Val Gln Asn Leu Val Asn Lys Ser Lys Lys Ile Val Gln Leu
 435 440 445
 Lys Pro Arg Asn Pro Asp Tyr Arg Ser Asn Lys Pro Ile Ile Leu Arg
 450 455 460
 Ala Leu Cys Asp Tyr Lys Gln Asp Gln Lys Ile Val His Lys Gly Asp
 465 470 475 480
 Glu Cys Ile Leu Lys Asp Asn Asn Glu Arg Ser Lys Trp Tyr Val Thr
 485 490 495
 Gly Pro Gly Gly Val Asp Met Leu Val Pro Ser Val Gly Leu Ile Ile
 500 505 510
 Pro Pro Pro Asn Pro Leu Ala Val Asp Leu Ser Cys Lys Ile Glu Gln
 515 520 525
 Tyr Tyr Glu Ala Ile Leu Ala Leu Trp Asn Gln Leu Tyr Ile Asn Met
 530 535 540
 Lys Ser Leu Val Ser Trp His Tyr Cys Met Ile Asp Ile Glu Lys Ile
 545 550 555 560
 Arg Ala Met Thr Ile Ala Lys Leu Lys Thr Met Arg Gln Glu Asp Tyr
 565 570 575

ES 2 739 426 T3

Met Lys Thr Ile Ala Asp Leu Glu Leu His Tyr Gln Glu Phe Ile Arg
580 585 590

Asn Ser Gln Gly Ser Glu Met Phe Gly Asp Asp Asp Lys Arg Lys Ile
595 600 605

Gln Ser Gln Phe Thr Asp Ala Gln Lys His Tyr Gln Thr Leu Val Ile
610 615 620

Gln Leu Pro Gly Tyr Pro Gln His Gln Thr Val Thr Thr Thr Glu Ile
625 630 635 640

Thr His His Gly Thr Cys Gln Asp Val Asn His Asn Lys Val Ile Glu
645 650 655

Thr Asn Arg Glu Asn Asp Lys Gln Glu Thr Trp Met Leu Met Glu Leu
660 665 670

Gln Lys Ile Arg Arg Gln Ile Glu His Cys Glu Gly Arg Met Thr Leu
675 680 685

Lys Asn Leu Pro Leu Ala Asp Gln Gly Ser Ser His His Ile Thr Val
690 695 700

Lys Ile Asn Glu Leu Lys Ser Val Gln Asn Asp Ser Gln Ala Ile Ala
705 710 715 720

Glu Val Leu Asn Gln Leu Lys Asp Met Leu Ala Asn Phe Arg Gly Ser
725 730 735

Glu Lys Tyr Cys Tyr Leu Gln Asn Glu Val Phe Gly Leu Phe Gln Lys
740 745 750

Leu Glu Asn Ile Asn Gly Val Thr Asp Gly Tyr Leu Asn Ser Leu Cys
755 760 765

Thr Val Arg Ala Leu Leu Gln Ala Ile Leu Gln Thr Glu Asp Met Leu
770 775 780

Lys Val Tyr Glu Ala Arg Leu Thr Glu Glu Glu Thr Val Cys Leu Asp
785 790 795 800

Leu Asp Lys Val Glu Ala Tyr Arg Cys Gly Leu Lys Lys Ile Lys Asn
805 810 815

Asp Leu Asn Leu Lys Lys Ser Leu Leu Ala Thr Met Lys Thr Glu Leu
820 825 830

ES 2 739 426 T3

Gln Lys Ala Gln Gln Ile His Ser Gln Thr Ser Gln Gln Tyr Pro Leu
 835 840 845
 Tyr Asp Leu Asp Leu Gly Lys Phe Gly Glu Lys Val Thr Gln Leu Thr
 850 855 860
 Asp Arg Trp Gln Arg Ile Asp Lys Gln Ile Asp Phe Arg Leu Trp Asp
 865 870 875 880
 Leu Glu Lys Gln Ile Lys Gln Leu Arg Asn Tyr Arg Asp Asn Tyr Gln
 885 890 895
 Ala Phe Cys Lys Trp Leu Tyr Asp Ala Lys Arg Arg Gln Asp Ser Leu
 900 905 910
 Glu Ser Met Lys Phe Gly Asp Ser Asn Thr Val Met Arg Phe Leu Asn
 915 920 925
 Glu Gln Lys Asn Leu His Ser Glu Ile Ser Gly Lys Arg Asp Lys Ser
 930 935 940
 Glu Glu Val Gln Lys Ile Ala Glu Leu Cys Ala Asn Ser Ile Lys Asp
 945 950 955 960
 Tyr Glu Leu Gln Leu Ala Ser Tyr Thr Ser Gly Leu Glu Thr Leu Leu
 965 970 975
 Asn Ile Pro Ile Lys Arg Thr Met Ile Gln Ser Pro Ser Gly Val Ile
 980 985 990
 Leu Gln Glu Ala Ala Asp Val His Ala Arg Tyr Ile Glu Leu Leu Thr
 995 1000 1005
 Arg Ser Gly Asp Tyr Tyr Arg Phe Leu Ser Glu Met Leu Lys Ser
 1010 1015 1020
 Leu Glu Asp Leu Lys Leu Lys Asn Thr Lys Ile Glu Val Leu Glu
 1025 1030 1035
 Glu Glu Leu Arg Leu Ala Arg Asp Ala Asn Ser Glu Asn Cys Asn
 1040 1045 1050
 Lys Asn Lys Phe Leu Asp Gln Asn Leu Gln Lys Tyr Gln Ala Glu
 1055 1060 1065
 Cys Ser Gln Phe Lys Ala Lys Leu Ala Ser Leu Glu Glu Leu Lys
 1070 1075 1080

ES 2 739 426 T3

Arg	Gln	Ala	Glu	Leu	Asp	Gly	Lys	Ser	Ala	Lys	Gln	Asn	Leu	Asp
1085						1090					1095			
Lys	Cys	Tyr	Gly	Gln	Ile	Lys	Glu	Leu	Asn	Glu	Lys	Ile	Thr	Arg
1100						1105					1110			
Leu	Thr	Tyr	Glu	Ile	Glu	Asp	Glu	Lys	Arg	Arg	Arg	Lys	Ser	Val
1115						1120					1125			
Glu	Asp	Arg	Phe	Asp	Gln	Gln	Lys	Asn	Asp	Tyr	Asp	Gln	Leu	Gln
1130						1135					1140			
Lys	Ala	Arg	Gln	Cys	Glu	Lys	Glu	Asn	Leu	Gly	Trp	Gln	Lys	Leu
1145						1150					1155			
Glu	Ser	Glu	Lys	Ala	Ile	Lys	Glu	Lys	Glu	Tyr	Glu	Ile	Glu	Arg
1160						1165					1170			
Leu	Arg	Val	Leu	Leu	Gln	Glu	Glu	Gly	Thr	Arg	Lys	Arg	Glu	Tyr
1175						1180					1185			
Glu	Asn	Glu	Leu	Ala	Lys	Ala	Ser	Asn	Arg	Ile	Gln	Glu	Ser	Lys
1190						1195					1200			
Asn	Gln	Cys	Thr	Gln	Val	Val	Gln	Glu	Arg	Glu	Ser	Leu	Leu	Val
1205						1210					1215			
Lys	Ile	Lys	Val	Leu	Glu	Gln	Asp	Lys	Ala	Arg	Leu	Gln	Arg	Leu
1220						1225					1230			
Glu	Asp	Glu	Leu	Asn	Arg	Ala	Lys	Ser	Thr	Leu	Glu	Ala	Glu	Thr
1235						1240					1245			
Arg	Val	Lys	Gln	Arg	Leu	Glu	Cys	Glu	Lys	Gln	Gln	Ile	Gln	Asn
1250						1255					1260			
Asp	Leu	Asn	Gln	Trp	Lys	Thr	Gln	Tyr	Ser	Arg	Lys	Glu	Glu	Ala
1265						1270					1275			
Ile	Arg	Lys	Ile	Glu	Ser	Glu	Arg	Glu	Lys	Ser	Glu	Arg	Glu	Lys
1280						1285					1290			
Asn	Ser	Leu	Arg	Ser	Glu	Ile	Glu	Arg	Leu	Gln	Ala	Glu	Ile	Lys
1295						1300					1305			

ES 2 739 426 T3

Arg Ile 1310	Glu Glu Arg Cys Arg 1315	Arg Lys Leu Glu Asp 1320	Ser Thr Arg
Glu Thr 1325	Gln Ser Gln Leu Glu 1330	Thr Glu Arg Ser Arg 1335	Tyr Gln Arg
Glu Ile 1340	Asp Lys Leu Arg Gln 1345	Arg Pro Tyr Gly Ser 1350	His Arg Glu
Thr Gln 1355	Thr Glu Cys Glu Trp 1360	Thr Val Asp Thr Ser 1365	Lys Leu Val
Phe Asp 1370	Gly Leu Arg Lys Lys 1375	Val Thr Ala Met Gln 1380	Leu Tyr Glu
Cys Gln 1385	Leu Ile Asp Lys Thr 1390	Thr Leu Asp Lys Leu 1395	Leu Lys Gly
Lys Lys 1400	Ser Val Glu Glu Val 1405	Ala Ser Glu Ile Gln 1410	Pro Phe Leu
Arg Gly 1415	Ala Gly Ser Ile Ala 1420	Gly Ala Ser Ala Ser 1425	Pro Lys Glu
Lys Tyr 1430	Ser Leu Val Glu Ala 1435	Lys Arg Lys Lys Leu 1440	Ile Ser Pro
Glu Ser 1445	Thr Val Met Leu Leu 1450	Glu Ala Gln Ala Ala 1455	Thr Gly Gly
Ile Ile 1460	Asp Pro His Arg Asn 1465	Glu Lys Leu Thr Val 1470	Asp Ser Ala
Ile Ala 1475	Arg Asp Leu Ile Asp 1480	Phe Asp Asp Arg Gln 1485	Gln Ile Tyr
Ala Ala 1490	Glu Lys Ala Ile Thr 1495	Gly Phe Asp Asp Pro 1500	Phe Ser Gly
Lys Thr 1505	Val Ser Val Ser Glu 1510	Ala Ile Lys Lys Asn 1515	Leu Ile Asp
Arg Glu 1520	Thr Gly Met Arg Leu 1525	Leu Glu Ala Gln Ile 1530	Ala Ser Gly
Gly Val 1535	Val Asp Pro Val Asn 1540	Ser Val Phe Leu Pro 1545	Lys Asp Val

ES 2 739 426 T3

Ala Leu 1550	Ala Arg Gly Leu Ile 1555	Asp Arg Asp Leu Tyr 1560	Arg Ser Leu
Asn Asp 1565	Pro Arg Asp Ser Gln 1570	Lys Asn Phe Val Asp 1575	Pro Val Thr
Lys Lys 1580	Lys Val Ser Tyr Val 1585	Gln Leu Lys Glu Arg 1590	Cys Arg Ile
Glu Pro 1595	His Thr Gly Leu Leu 1600	Leu Leu Ser Val Gln 1605	Lys Arg Ser
Met Ser 1610	Phe Gln Gly Ile Arg 1615	Gln Pro Val Thr Val 1620	Thr Glu Leu
Val Asp 1625	Ser Gly Ile Leu Arg 1630	Pro Ser Thr Val Asn 1635	Glu Leu Glu
Ser Gly 1640	Gln Ile Ser Tyr Asp 1645	Glu Val Gly Glu Arg 1650	Ile Lys Asp
Phe Leu 1655	Gln Gly Ser Ser Cys 1660	Ile Ala Gly Ile Tyr 1665	Asn Glu Thr
Thr Lys 1670	Gln Lys Leu Gly Ile 1675	Tyr Glu Ala Met Lys 1680	Ile Gly Leu
Val Arg 1685	Pro Gly Thr Ala Leu 1690	Glu Leu Leu Glu Ala 1695	Gln Ala Ala
Thr Gly 1700	Phe Ile Val Asp Pro 1705	Val Ser Asn Leu Arg 1710	Leu Pro Val
Glu Glu 1715	Ala Tyr Lys Arg Gly 1720	Leu Val Gly Ile Glu 1725	Phe Lys Glu
Lys Leu 1730	Leu Ser Ala Glu Arg 1735	Ala Val Thr Gly Tyr 1740	Asn Asp Pro
Glu Thr 1745	Gly Asn Ile Ile Ser 1750	Leu Phe Gln Ala Met 1755	Asn Lys Glu
Leu Ile 1760	Glu Lys Gly His Gly 1765	Ile Arg Leu Leu Glu 1770	Ala Gln Ile
Ala Thr 1775	Gly Gly Ile Ile Asp 1780	Pro Lys Glu Ser His 1785	Arg Leu Pro

ES 2 739 426 T3

Val Asp 1790	Ile Ala Tyr Lys Arg 1795	Gly Tyr Phe Asn Glu 1800	Glu Leu Ser
Glu Ile 1805	Leu Ser Asp Pro Ser 1810	Asp Asp Thr Lys Gly 1815	Phe Phe Asp
Pro Asn 1820	Thr Glu Glu Asn Leu 1825	Thr Tyr Leu Gln Leu 1830	Lys Glu Arg
Cys Ile 1835	Lys Asp Glu Glu Thr 1840	Gly Leu Cys Leu Leu 1845	Pro Leu Lys
Glu Lys 1850	Lys Lys Gln Val Gln 1855	Thr Ser Gln Lys Asn 1860	Thr Leu Arg
Lys Arg 1865	Arg Val Val Ile Val 1870	Asp Pro Glu Thr Asn 1875	Lys Glu Met
Ser Val 1880	Gln Glu Ala Tyr Lys 1885	Lys Gly Leu Ile Asp 1890	Tyr Glu Thr
Phe Lys 1895	Glu Leu Cys Glu Gln 1900	Glu Cys Glu Trp Glu 1905	Glu Ile Thr
Ile Thr 1910	Gly Ser Asp Gly Ser 1915	Thr Arg Val Val Leu 1920	Val Asp Arg
Lys Thr 1925	Gly Ser Gln Tyr Asp 1930	Ile Gln Asp Ala Ile 1935	Asp Lys Gly
Leu Val 1940	Asp Arg Lys Phe Phe 1945	Asp Gln Tyr Arg Ser 1950	Gly Ser Leu
Ser Leu 1955	Thr Gln Phe Ala Asp 1960	Met Ile Ser Leu Lys 1965	Asn Gly Val
Gly Thr 1970	Ser Ser Ser Met Gly 1975	Ser Gly Val Ser Asp 1980	Asp Val Phe
Ser Ser 1985	Ser Arg His Glu Ser 1990	Val Ser Lys Ile Ser 1995	Thr Ile Ser
Ser Val 2000	Arg Asn Leu Thr Ile 2005	Arg Ser Ser Ser Phe 2010	Ser Asp Thr
Leu Glu 2015	Glu Ser Ser Pro Ile 2020	Ala Ala Ile Phe Asp 2025	Thr Glu Asn

ES 2 739 426 T3

Leu	Glu	Lys	Ile	Ser	Ile	Thr	Glu	Gly	Ile	Glu	Arg	Gly	Ile	Val
2030						2035					2040			
Asp	Ser	Ile	Thr	Gly	Gln	Arg	Leu	Leu	Glu	Ala	Gln	Ala	Cys	Thr
2045						2050					2055			
Gly	Gly	Ile	Ile	His	Pro	Thr	Thr	Gly	Gln	Lys	Leu	Ser	Leu	Gln
2060						2065					2070			
Asp	Ala	Val	Ser	Gln	Gly	Val	Ile	Asp	Gln	Asp	Met	Ala	Thr	Arg
2075						2080					2085			
Leu	Lys	Pro	Ala	Gln	Lys	Ala	Phe	Ile	Gly	Phe	Glu	Gly	Val	Lys
2090						2095					2100			
Gly	Lys	Lys	Lys	Met	Ser	Ala	Ala	Glu	Ala	Val	Lys	Glu	Lys	Trp
2105						2110					2115			
Leu	Pro	Tyr	Glu	Ala	Gly	Gln	Arg	Phe	Leu	Glu	Phe	Gln	Tyr	Leu
2120						2125					2130			
Thr	Gly	Gly	Leu	Val	Asp	Pro	Glu	Val	His	Gly	Arg	Ile	Ser	Thr
2135						2140					2145			
Glu	Glu	Ala	Ile	Arg	Lys	Gly	Phe	Ile	Asp	Gly	Arg	Ala	Ala	Gln
2150						2155					2160			
Arg	Leu	Gln	Asp	Thr	Ser	Ser	Tyr	Ala	Lys	Ile	Leu	Thr	Cys	Pro
2165						2170					2175			
Lys	Thr	Lys	Leu	Lys	Ile	Ser	Tyr	Lys	Asp	Ala	Ile	Asn	Arg	Ser
2180						2185					2190			
Met	Val	Glu	Asp	Ile	Thr	Gly	Leu	Arg	Leu	Leu	Glu	Ala	Ala	Ser
2195						2200					2205			
Val	Ser	Ser	Lys	Gly	Leu	Pro	Ser	Pro	Tyr	Asn	Met	Ser	Ser	Ala
2210						2215					2220			
Pro	Gly	Ser	Arg	Ser	Gly	Ser	Arg	Ser	Gly	Ser	Arg	Ser	Gly	Ser
2225						2230					2235			
Arg	Ser	Gly	Ser	Arg	Ser	Gly	Ser	Arg	Arg	Gly	Ser	Phe	Asp	Ala
2240						2245					2250			
Thr	Gly	Asn	Ser	Ser	Tyr	Ser	Tyr	Ser	Tyr	Ser	Phe	Ser	Ser	Ser
2255						2260					2265			
Ser	Ile	Gly	His											
2270														

REIVINDICACIONES

1. Método *in vitro* para diagnosticar un cáncer de mama en un sujeto, que comprende las etapas de:

a) determinar a partir de una muestra de líquido biológico de un sujeto el nivel de expresión de proteínas de por lo menos dos biomarcadores seleccionados de entre el grupo que consiste en olfactomedina-4, neudesina y desmoplaquina, en el que dichos por lo menos dos biomarcadores son olfactomedina-4 y neudesina; y

b) comparar dicho nivel de expresión con un nivel de expresión de referencia de dicho biomarcador.

2. Método según la reivindicación 1, en el que un nivel de expresión de proteínas de dichos por lo menos dos biomarcadores desregulados mediante comparación con dicho nivel de expresión de referencia obtenido a partir de una muestra de líquido biológico de por lo menos un sujeto sano, es indicativo de que dicho sujeto padece cáncer de mama.

3. Método según la reivindicación 2, en el que dicho nivel de expresión de:

- olfactomedina-4 es superior a dicho nivel de expresión de referencia, y
- neudesina es superior a dicho nivel de expresión de referencia, y opcionalmente
- desmoplaquina es inferior o superior a dicho nivel de expresión de referencia,

en la muestra de líquido biológico de dicho sujeto que padece cáncer de mama, tal como un cáncer de mama temprano.

4. Método según la reivindicación 3, en el que dicho nivel de expresión de desmoplaquina inferior a dicho nivel de expresión de referencia es indicativo de que dicho cáncer de mama es un cáncer de mama temprano.

5. Método según la reivindicación 3, en el que dicho nivel de expresión de desmoplaquina superior a dicho nivel de expresión de referencia es indicativo de que dicho cáncer de mama es un cáncer de mama recidivante.

6. Método según cualquiera de las reivindicaciones 1 a 5, en el que el nivel de expresión de proteínas de por lo menos tres de dichos biomarcadores se determina en la etapa a).

7. Método según cualquiera de las reivindicaciones 1 a 6, que comprende además la etapa de determinar el nivel de expresión de proteínas de por lo menos un biomarcador estándar asociado con cáncer de mama, seleccionándose preferentemente dicho biomarcador estándar de entre el grupo que consiste en receptor de estrógenos (ER), receptor de progesterona (PR) y receptor 2 de factor de crecimiento epidérmico humano (HER2).

8. Método de cribado para identificar un fármaco o una combinación de fármacos adecuados para tratar cáncer de mama, que comprende las etapas de:

a) poner en contacto unas células de cáncer de mama aisladas o una estirpe celular que presenta un fenotipo de cáncer de mama con un fármaco candidato o combinación de fármacos candidatos;

b) determinar, a partir de dichas células o estirpe celular en contacto con dicho/a fármaco o combinación de fármacos, el nivel de expresión de proteínas de por lo menos dichos dos biomarcadores como se define en la reivindicación 1; y

c) comparar el nivel de expresión de proteínas de dicho biomarcador en la etapa b) con su nivel de expresión en ausencia de dicho/a fármaco o combinación de fármacos.

9. Método según cualquiera de las reivindicaciones 1 a 7, en el que dicho líquido biológico se selecciona de entre el grupo que consiste en sangre, suero, plasma, linfa, líquido intersticial de tumor, saliva, moco, esputo, sudor y orina.

10. Método según la reivindicación 9, en el que dicho líquido biológico es el suero.

11. Método según cualquiera de las reivindicaciones 1 a 10, en el que el nivel de expresión de proteínas se determina mediante un método seleccionado de entre el grupo que consiste en transferencia de tipo Western, inmunotransferencia, ensayo de inmunoadsorción ligado a enzimas (ELISA), Immunospot ligado a enzimas (ELISPOT), radioinmunoensayo (RIA), inmunohistoquímica, inmunoprecipitación, clasificación celular activada por fluorescencia (FACS), microscopía, citometría de flujo, microcitometría, ensayo de unión a proteínas, ensayo

de unión a ligandos, micromatriz, electroforesis en gel de poliacrilamida tal como SDS-PAGE, resonancia de plasmón superficial (SPR), transferencia de energía por resonancia de Forster (FRET), transferencia de energía por resonancia de bioluminiscencia (BRET), quimioluminiscencia, polarización de fluorescencia, fosforescencia, espectrometría de masas, obtención de imágenes por resonancia magnética (MRI), y cualquier combinación de los mismos.

5

12. Utilización de un kit en un método como se define en cualquiera de las reivindicaciones 1 a 11, comprendiendo dicho kit:

10

a) por lo menos un reactivo apto para determinar específicamente el nivel de expresión de proteínas de por lo menos dichos dos biomarcadores como se define en la reivindicación 1; y

b) unas instrucciones para realizar dicho método.

15

13. Utilización de una micromatriz de proteínas en un método como se define en cualquiera de las reivindicaciones 1 a 11, comprendiendo dicha micromatriz de proteínas:

a) por lo menos un reactivo apto para determinar específicamente el nivel de expresión de proteínas de por lo menos dichos dos biomarcadores como se define en la reivindicación 1.

20

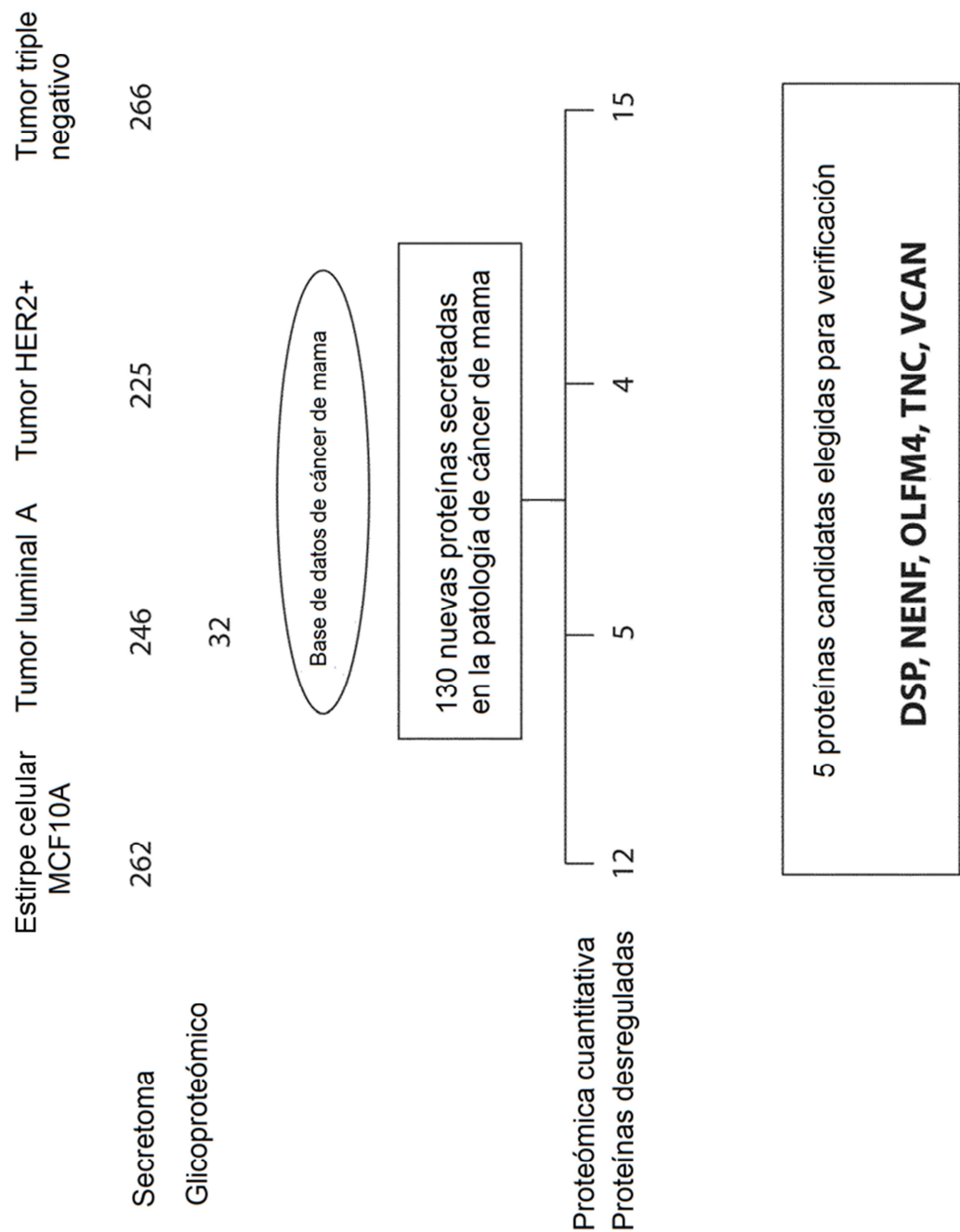


Figura 1

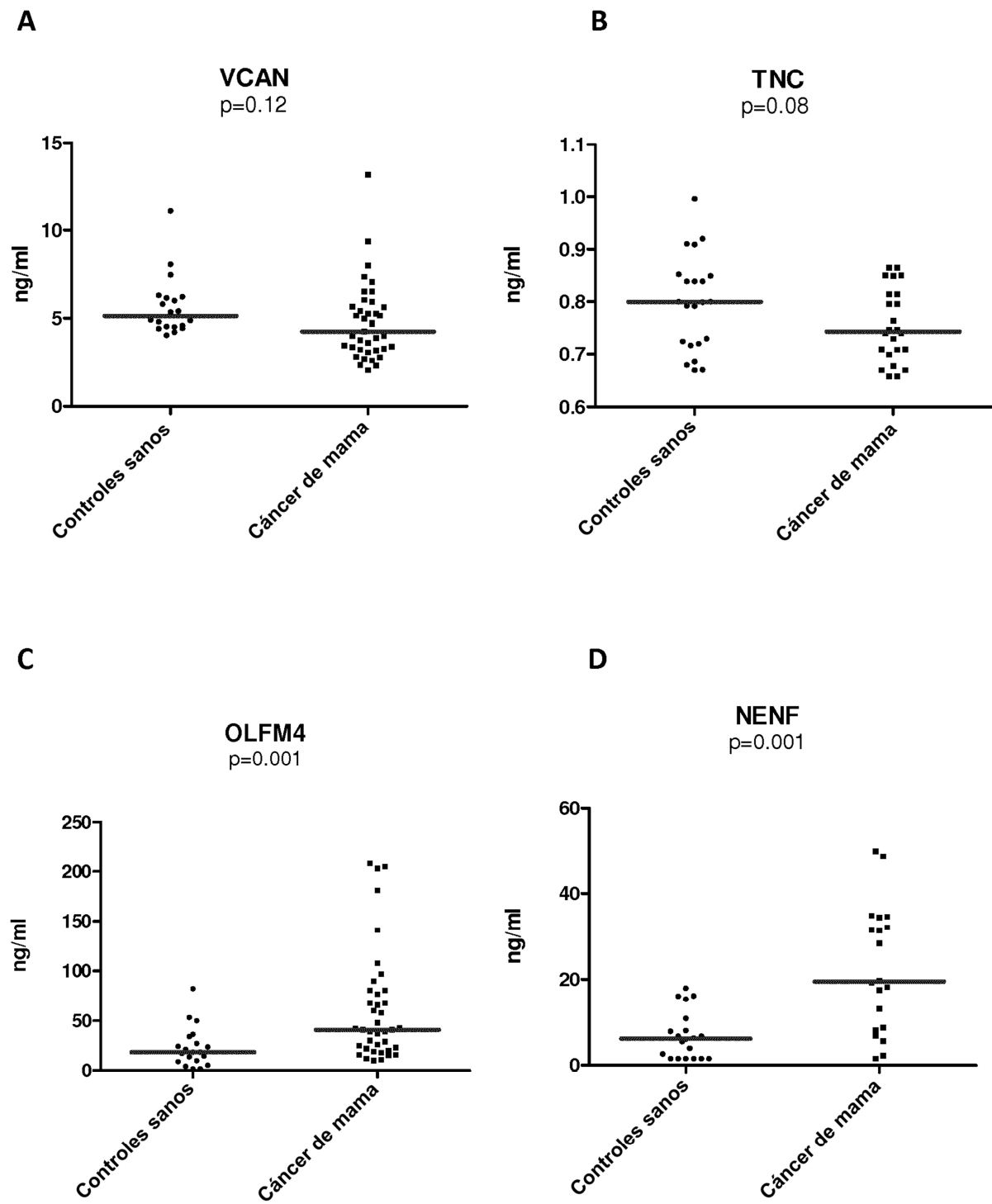


Figura 2

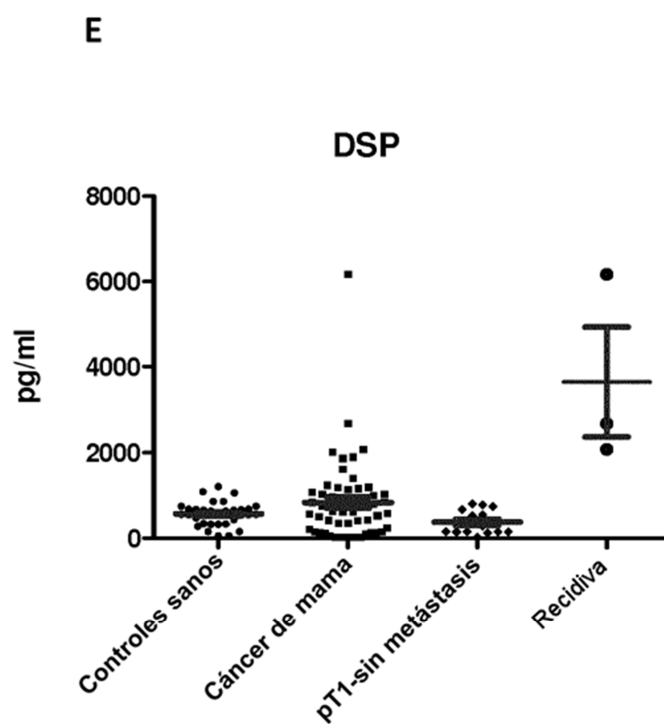


Figura 2 (continuación)

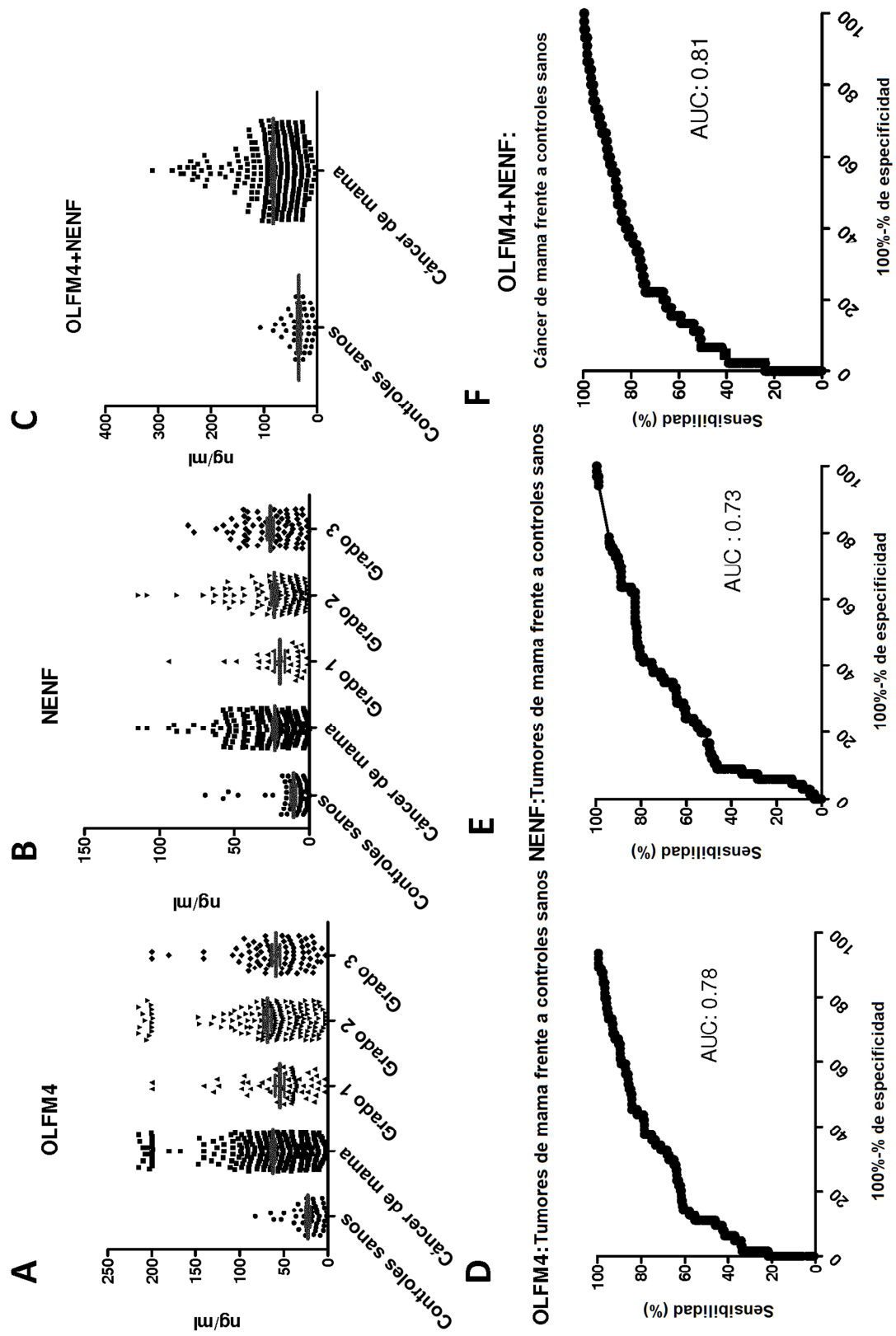


Figura 3

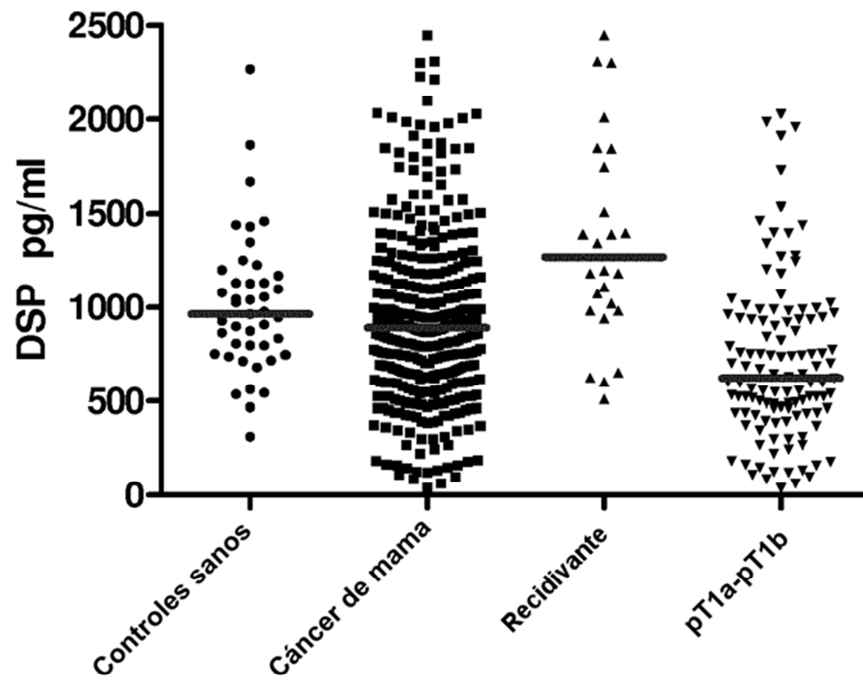


Figura 4

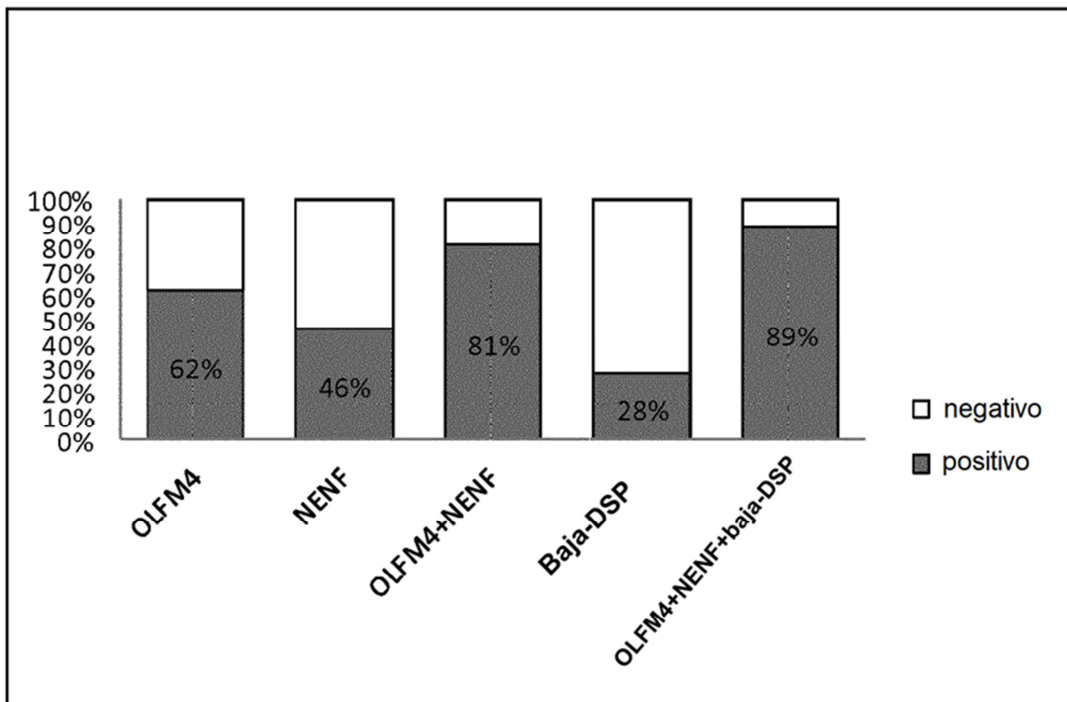
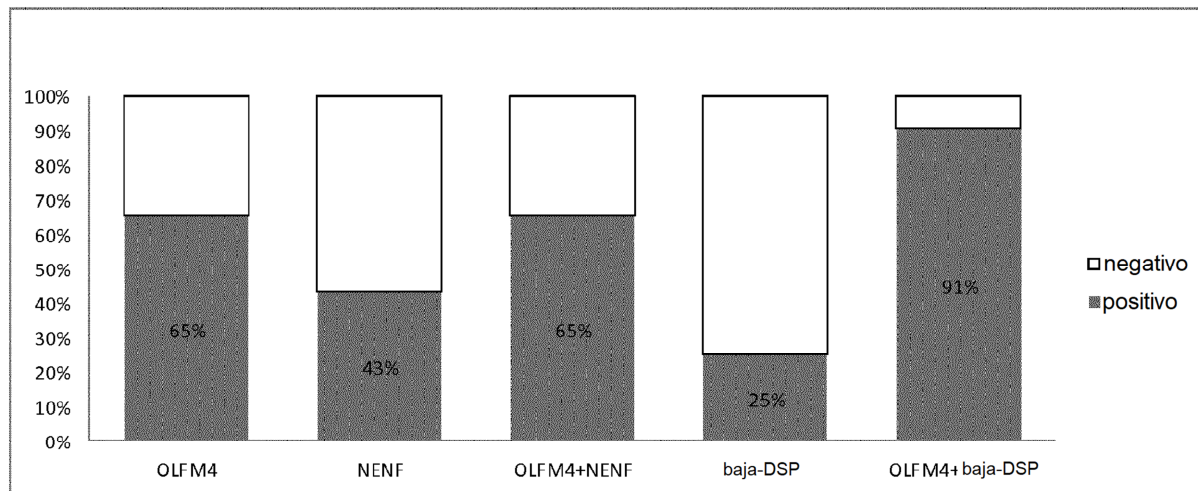


Figura 5

A



B

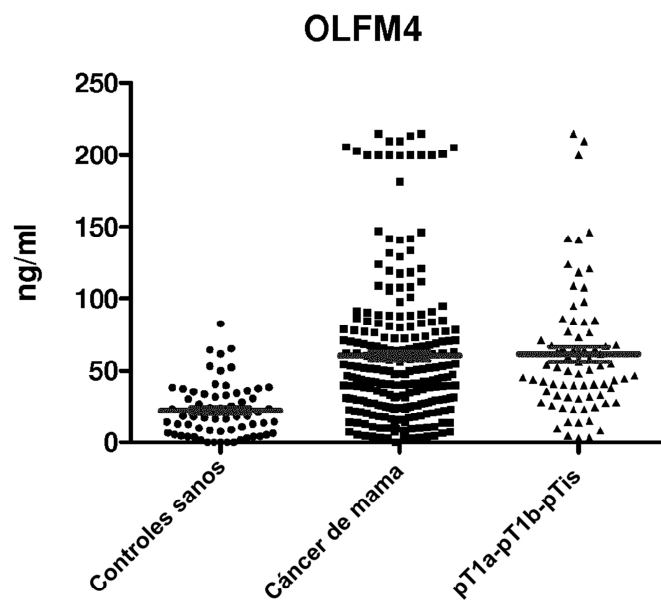
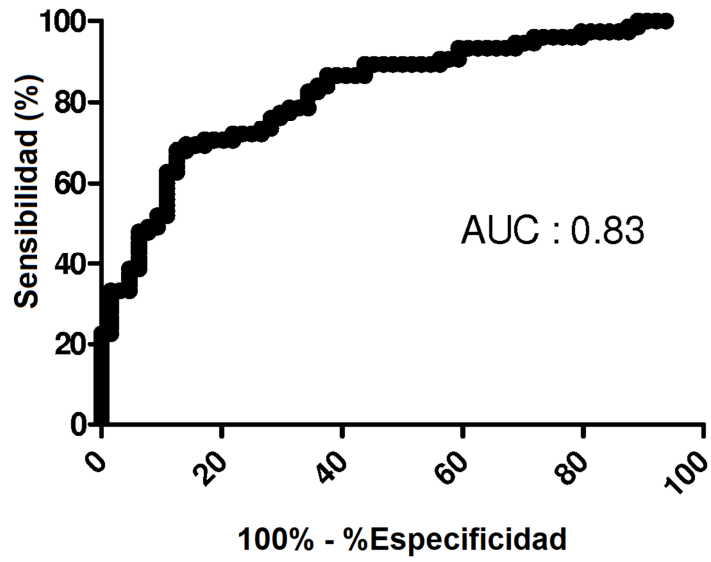


Figura 6

C

OLFM4:pT1a-pT1b-Tis frente a controles sanos



D

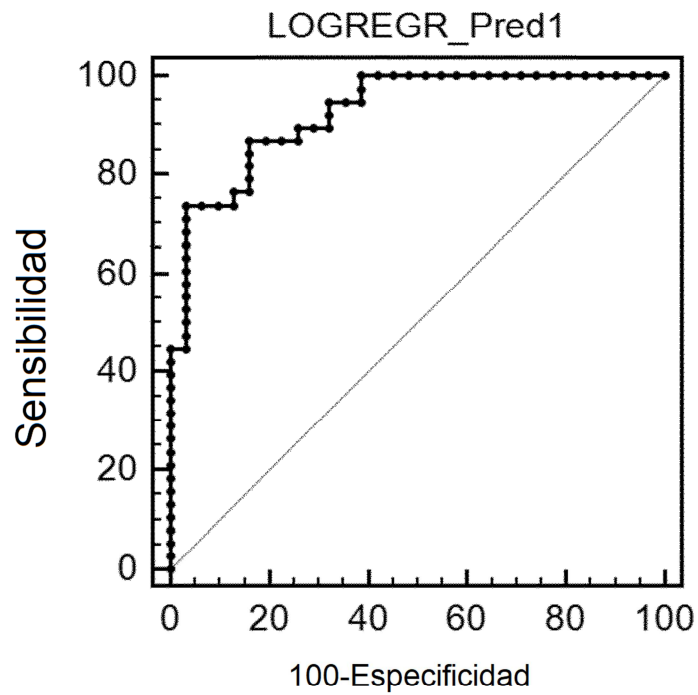
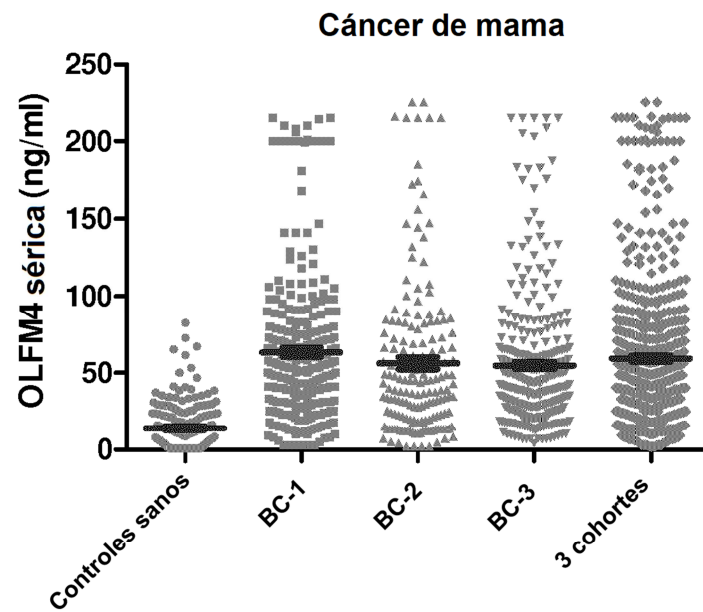


Figura 6 (continuación)

A



B

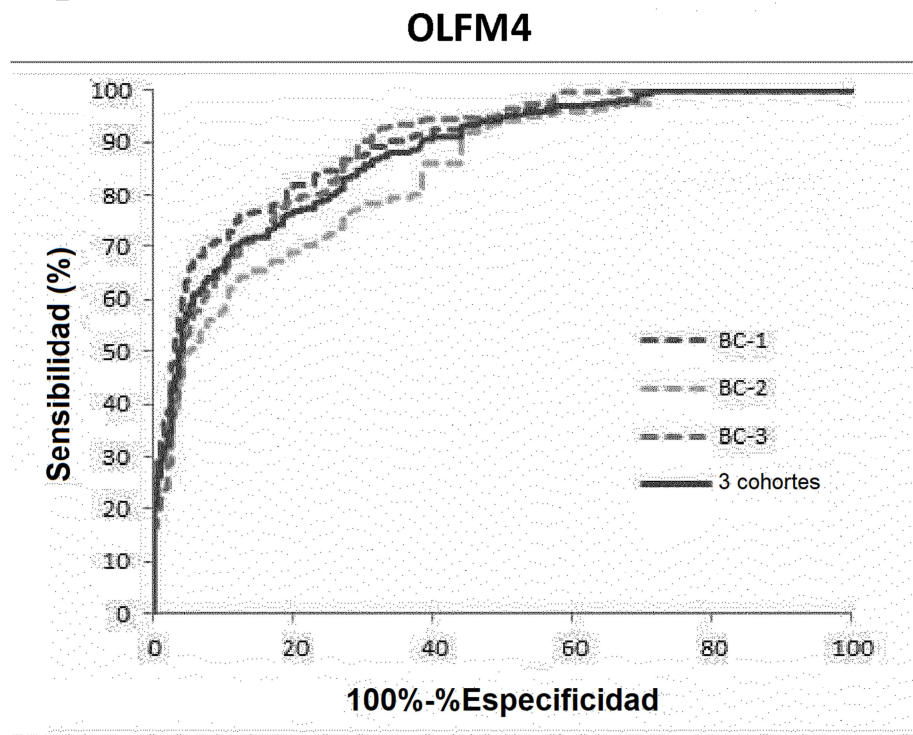


Figura 7

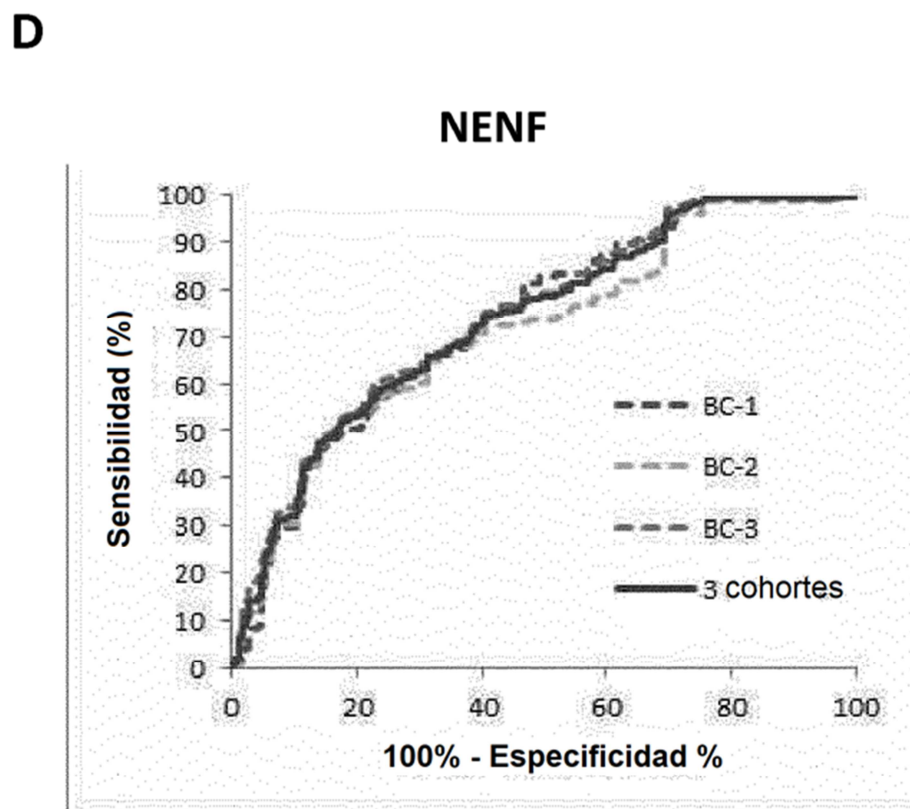
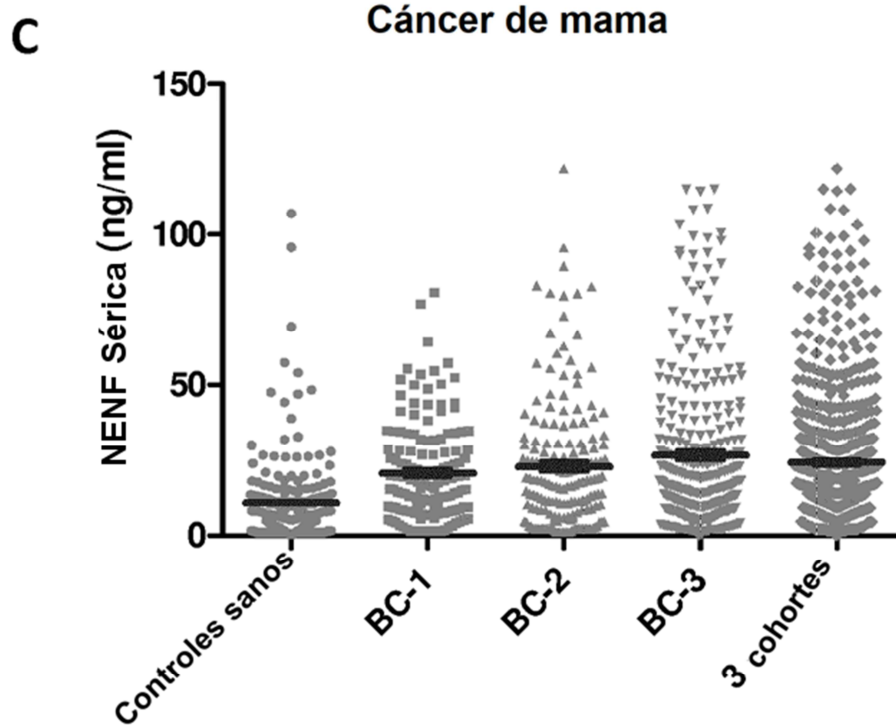


Figura 7 (continuación)

OLFM4 + NENF

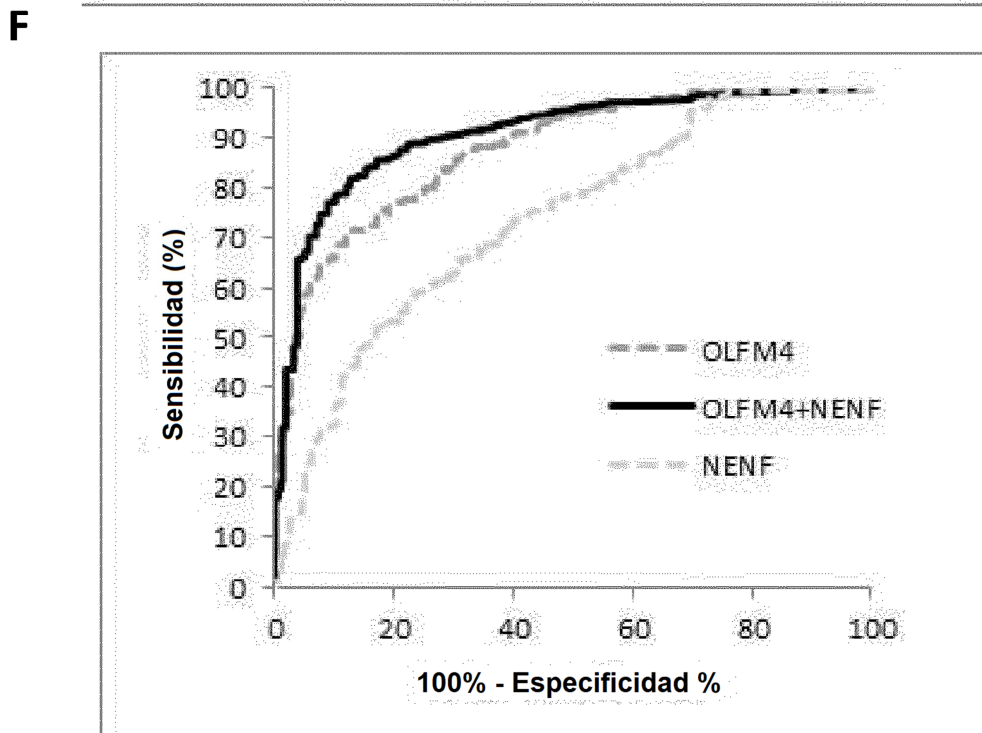
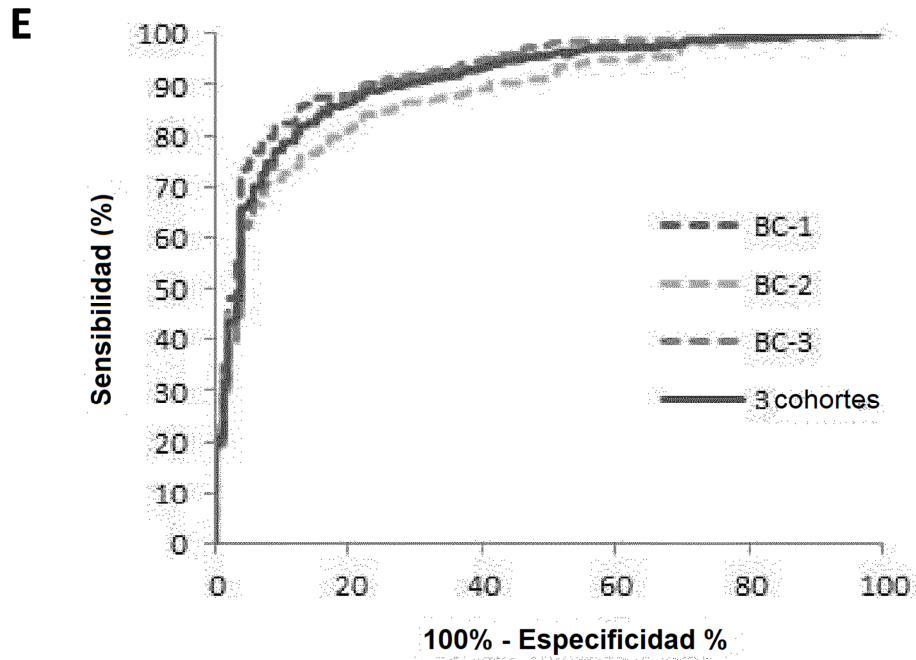


Figura 7 (continuación)

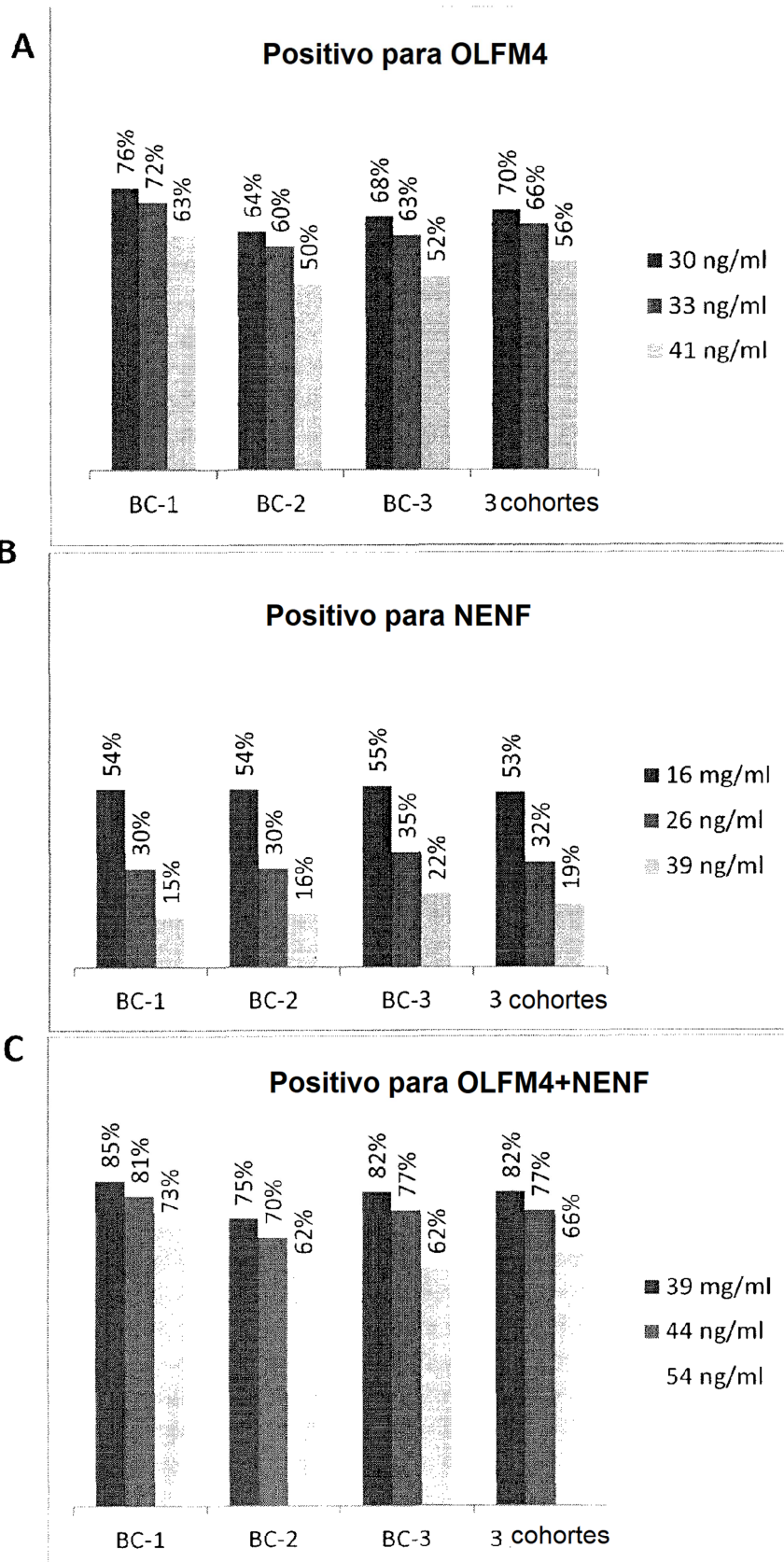
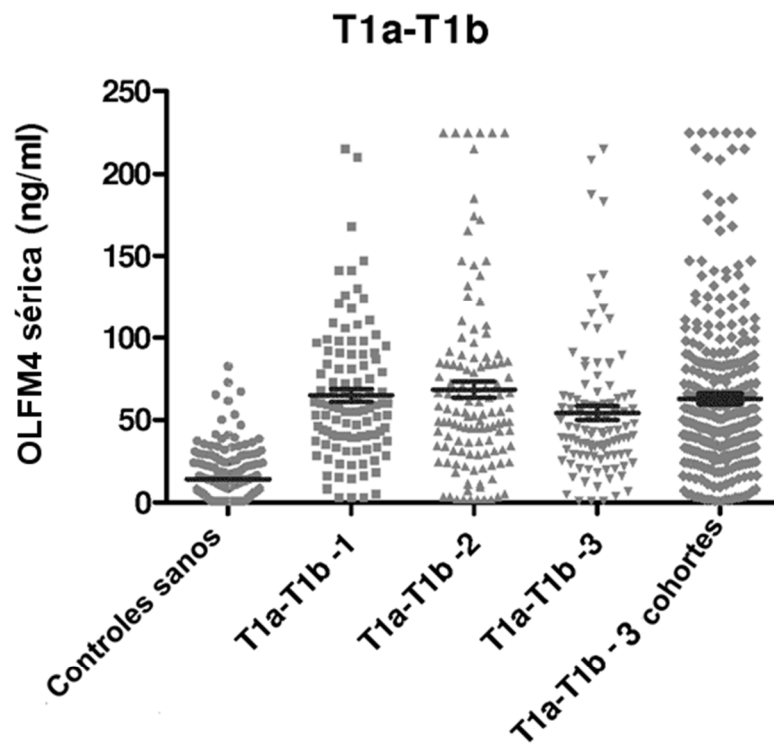


Figura 8

A



B

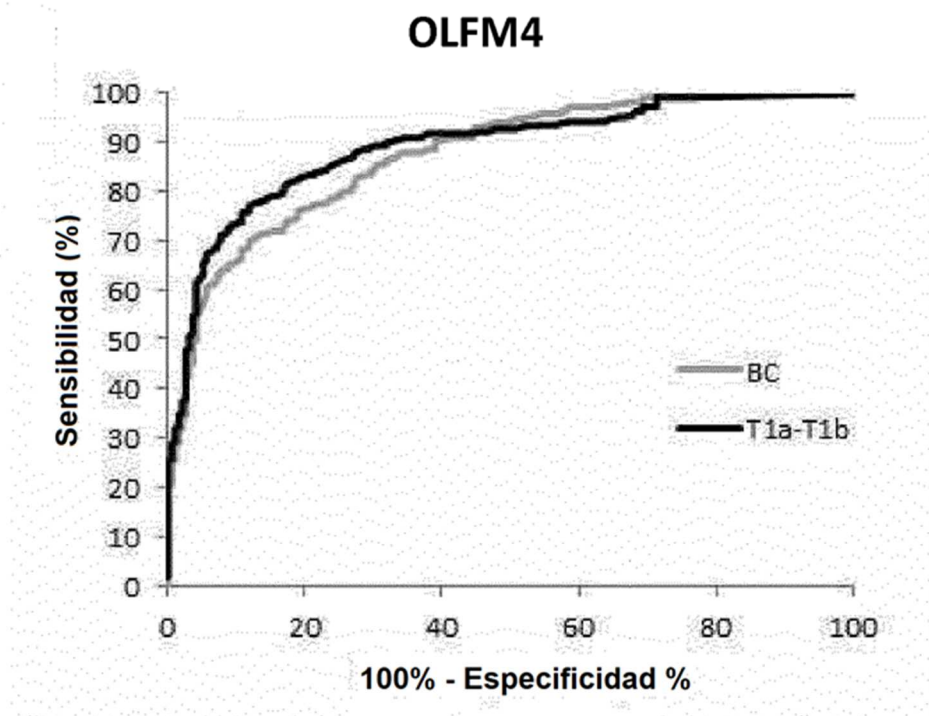
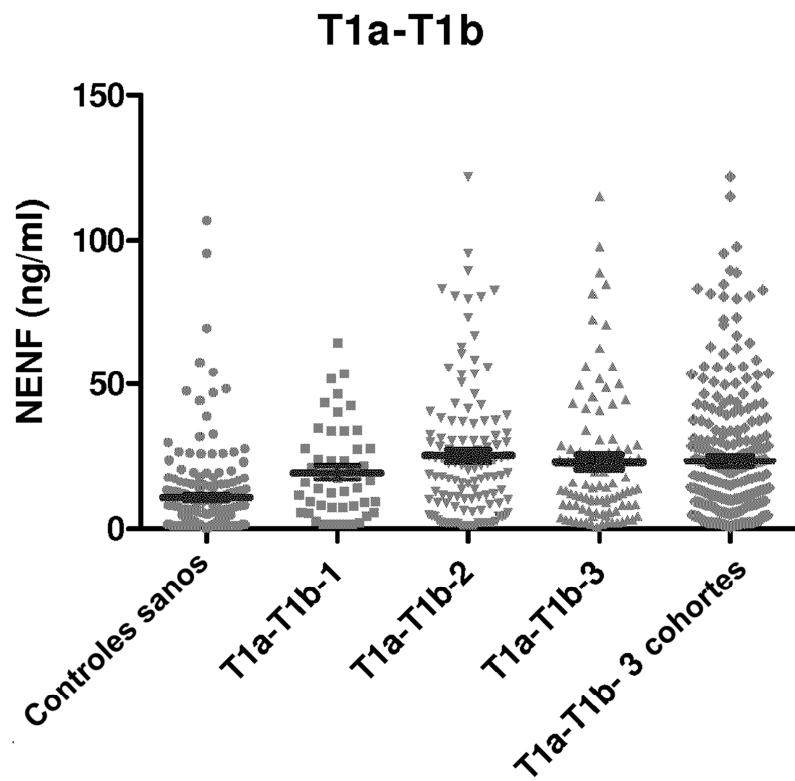


Figura 9

C



D

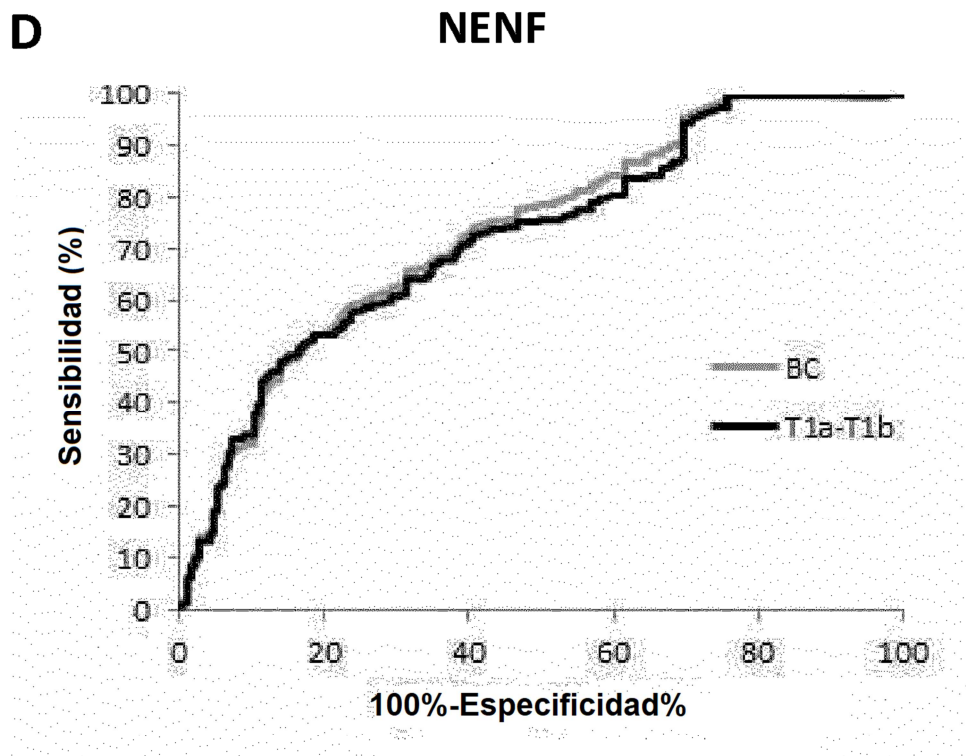


Figura 9 (continuación)

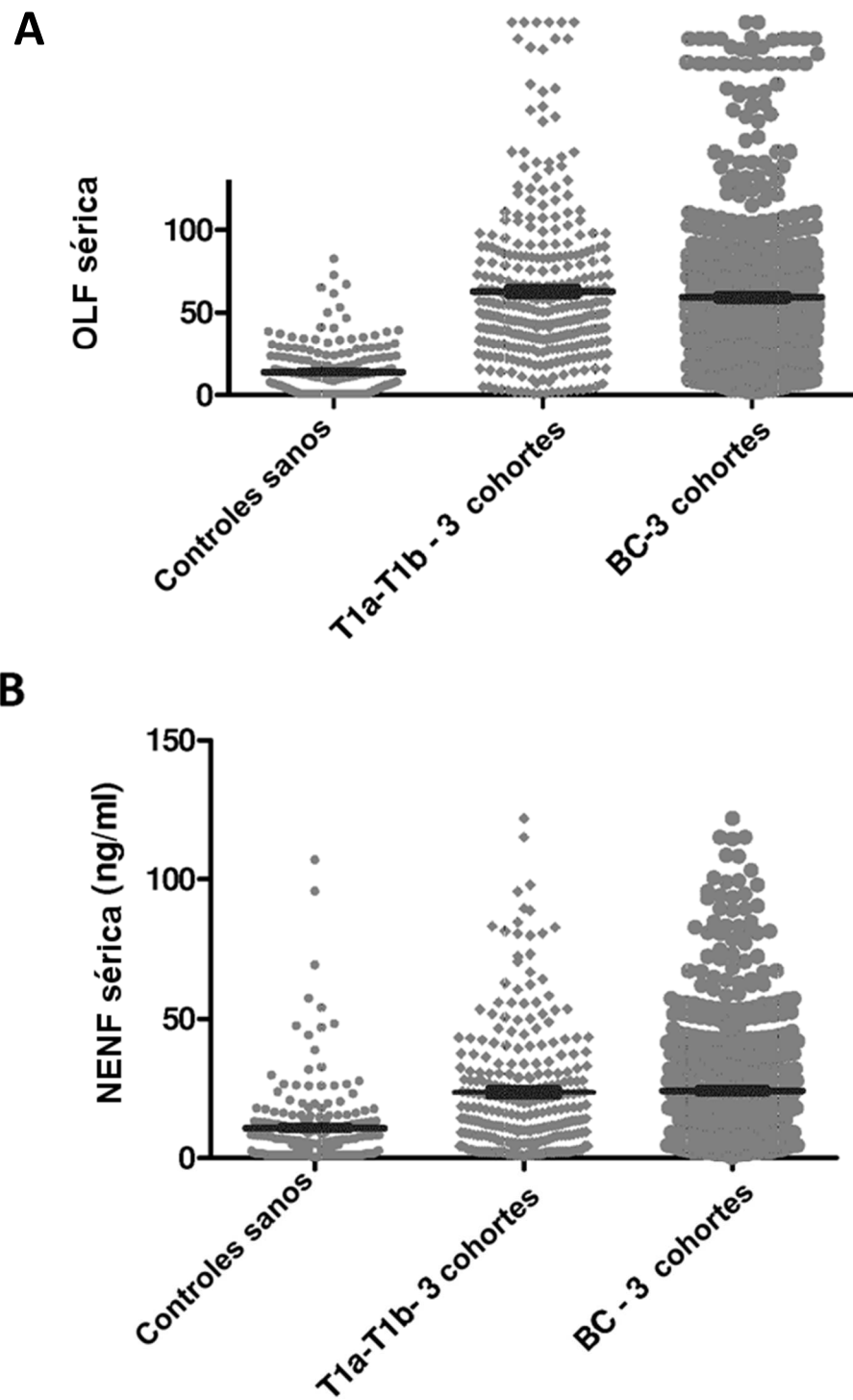


Figura 10

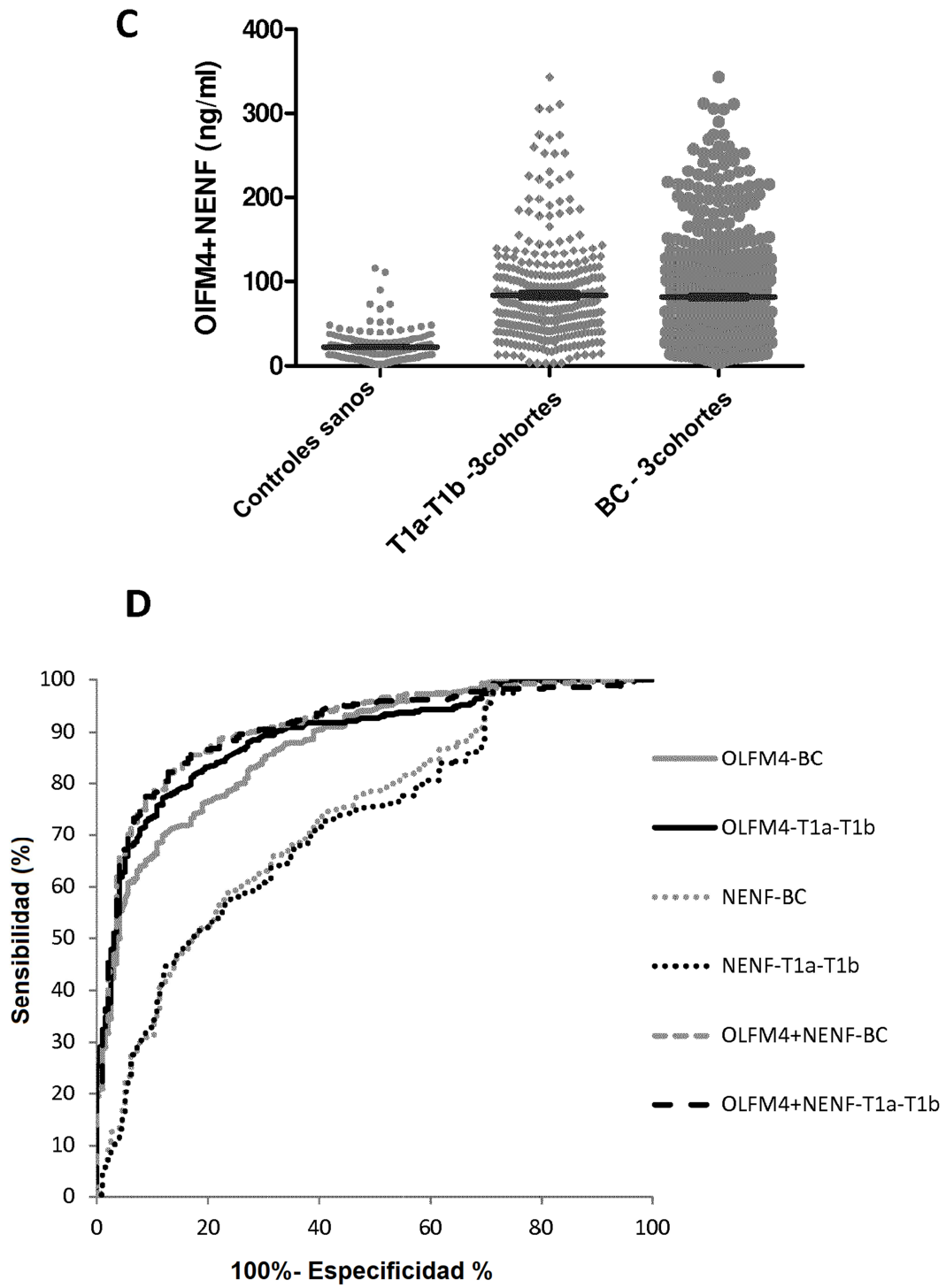


Figura 10 (continuación)