

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 375 998**

51 Int. Cl.:
C12Q 1/68 (2006.01)
G01N 33/574 (2006.01)
G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **06829113 .7**
96 Fecha de presentación: **23.11.2006**
97 Número de publicación de la solicitud: **1954830**
97 Fecha de publicación de la solicitud: **13.08.2008**

54 Título: **PROCEDIMIENTOS Y KITS PARA EL PRONÓSTICO DE CÁNCER DE MAMA.**

30 Prioridad:
03.12.2005 EP 05026423

45 Fecha de publicación de la mención BOPI:
08.03.2012

45 Fecha de la publicación del folleto de la patente:
08.03.2012

73 Titular/es:
Sividon Diagnostics GmbH
Nattermann Allee 1
50829 Köln, DE

72 Inventor/es:
GEHRMANN, Mathias y
VON TOERNE, Christian

74 Agente/Representante:
Carpintero López, Mario

ES 2 375 998 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

Procedimientos y kits para el pronóstico de cáncer de mama

Campo técnico de la invención

5 La presente invención se refiere al campo del pronóstico de una enfermedad proliferativa en una paciente. Más específicamente, la presente invención se refiere a procedimientos para la identificación del resultado probable de cáncer de mama en una paciente con cáncer de mama. La invención también se refiere a procedimientos para la identificación del resultado probable de cáncer de mama en una paciente en base al estado de ERBB2 de la paciente y los niveles de expresión de genes inmunes en el tejido tumoral de dicha paciente.

Antecedentes de la invención

10 El cáncer de mama es una de las principales causas de muerte por cáncer en mujeres en los países occidentales. Más específicamente, el cáncer de mama se cobra la vida de aproximadamente 40.000 mujeres y se diagnostica a unas 200.000 mujeres cada año, solo en los Estados Unidos. En las últimas décadas, la terapia sistémica complementaria ha provocado una notoria mejora en la supervivencia en cáncer de mama precoz [2, 3]. Esta experiencia clínica ha llevado a recomendaciones de consenso que ofrecen terapia sistémica complementaria a la gran mayoría de las pacientes con cáncer de mama [1]. Para tratar el cáncer de mama se dispone de una diversas opciones de tratamiento que pueden aplicarse además de la extirpación quirúrgica del tumor que se realiza comúnmente y la posterior radiación del lecho del tumor. El tratamiento endocrino, la quimioterapia y el tratamiento con terapias dirigidas son tres estrategias principales y conceptualmente diferentes. Un requisito para el tratamiento con agentes endocrinos es la expresión de receptores de hormonas en el tejido tumoral, ya sea estrógeno, progesterona o ambas. Se encuentran disponibles muchos agentes endocrinos con distintos modos de acción y diferencias en el resultado de la enfermedad cuando se evalúan en grandes cohortes de pacientes. El tamoxifén es uno de los fármacos endocrinos más antiguos, que redujo significativamente el riesgo de la reaparición del tumor. Aparentemente, son aun más efectivos los inhibidores de aromatasa, que pertenecen a una nueva clase de fármacos endocrinos. En contraste con el Tamoxifén, que es un inhibidor competitivo de la unión del estrógeno, los inhibidores de aromatasa bloquean la producción de estrógeno y, de ese modo, logran reducir el estímulo del crecimiento para las células tumorales con receptor de estrógeno positivo. Los ensayos clínicos recientes han demostrado un mejor resultado de la enfermedad en las pacientes tratadas con estos agentes en comparación con las tratadas con Tamoxifén. No obstante, algunas pacientes todavía experimentan una recaída a pesar del tratamiento endocrino y, en particular, estas pacientes podrían beneficiarse de fármacos terapéuticos adicionales. Se ha demostrado que la quimioterapia con antraciclinas, taxanos y otros agentes es eficiente para reducir la recidiva de la enfermedad en pacientes tanto con receptor de estrógeno positivo como con receptor de estrógeno negativo. El estudio NSABP-20 comparó el Tamoxifén solo con el Tamoxifén más la quimioterapia en pacientes con receptor de estrógeno positivo sin afectación ganglionar y mostró que el tratamiento combinado fue más efectivo que el Tamoxifén solo.

35 Un nuevo enfoque es el tratamiento de pacientes con cáncer de mama con anticuerpos monoclonales dirigidos a moléculas de la superficie celular asociadas con el tumor o a moléculas de la superficie celular en células endoteliales de la vasculatura tumoral.

Recientemente, se ha demostrado que el anticuerpo monoclonal administrado sistémicamente Trastuzumab (Herceptin; Genentech, South San Francisco, CA, Estados Unidos) dirigido al antígeno HER2 en la superficie de las células tumorales reduce el riesgo de recidiva en pacientes con tumores con sobreexpresión de HER2. En la reunión de la Sociedad Americana de Oncología Clínica de 2005 se presentaron dos informes que describían una reducción del 50% en el riesgo de recaída durante el tratamiento con Trastuzumab en pruebas que comprendían alrededor de 8.000 pacientes con cáncer de mama. El antígeno HER2 se codifica mediante el gen ERBB2 que se ubica en el cromosoma 17q12-q21. Varios estudios realizados en diversos laboratorios han demostrado que entre el 25% y el 30% de todos los tumores malignos de mama y ovario muestran la amplificación y sobreexpresión de este gen [4, 5]. La amplificación del gen ERBB2 se encuentra en más del 90% de los casos que tienen sobreexpresión de proteína ERBB2 [6, 7], pero en una mama normal la expresión de ERBB2 se debe a una activación transcripcional [8]. La sobreexpresión de ERBB2 en mujeres con cáncer de mama con afectación ganglionar [4,5] y sin afectación ganglionar [9] se asocia con un mal pronóstico, y varios estudios han encontrado una correlación entre la sobreexpresión de ERBB2 y un período libre de enfermedad más corto y una supervivencia general más corta [10, 11].

Si bien en la técnica se dispone de varios procedimientos para evaluar el estado de ERBB2 [12], los más utilizados son la inmunohistoquímica (IHC; para sobreexpresión de proteínas) y la hibridación fluorescente in situ (FISH; para la amplificación de genes). Los informes de casos de falsos positivos de Herceptest llevaron a sugerir que los Herceptests que proporcionan un puntaje 2+ también deberían estudiarse mediante FISH [13]. Tubbs y sus colegas [14] solicitaron que la Administración de Alimentos y Fármacos de los Estados Unidos (FDA) ordenara la retractación de los criterios aceptados previamente para la terapia con Trastuzumab, precisamente aquella con Herceptests que proporcionan un puntaje 2+, salvo que esos casos también fueran confirmados mediante FISH.

El ensayo de FISH aprobado por la FDA, PathVysion (Vysis, Inc., Downers Grove, IL, Estados Unidos), es un sistema de sonda dual para la enumeración simultánea del gen ERBB2 y la región centromérica del cromosoma 17, definiendo la amplificación de ERBB2 como una relación de copias del gen ERBB2 por centrómero del cromosoma 17 [15]. Otro ensayo de FISH aprobado por la FDA, INFORM (Ventana Medical Systems, Tucson, AZ, Estados Unidos), define la amplificación de ERBB2 como un número absoluto medio de copia del gen ERBB2 de más de cuatro puntos por núcleo, sin corrección de centrómero 17. El FISH con la sonda PathVysion y el ensayo inmunohistoquímico con Herceptest son altamente concordantes para los casos que muestran puntajes de IHC 3+ y para casos negativos, si bien el grupo de IHC de puntaje 2+ incluye tumores ERBB2 amplificados y no amplificados, mostrando una tasa de discordancia relativamente alta [16,17]. Los mecanismos para la expresión de ERBB2 en tumores no amplificados puntuados 2+ mediante IHC no son claros y pueden implicar una mayor dosificación de genes mediante la polisomía del cromosoma 17 [18].

La administración de Trastuzumab puede considerarse como un tipo de inmunoterapia pasiva. La inmunoterapia activa, por ejemplo la vacunación, no ha sido desarrollada con éxito debido, en parte, a la comprensión limitada del impacto funcional del sistema inmune en el cáncer de mama. Los estudios histológicos y de citometría de flujo han documentado que los cánceres de mama humanos son infiltrados por una población heterogénea de células inmunes, que consiste en diferentes proporciones de células T, células B, linfocitos citolíticos naturales y macrófagos. Aparentemente, el sistema inmune es capaz de responder al cáncer de mama como se evidencia por la activación sistémica, regional e intratumoral de leucocitos y la presencia de leucocitos que infiltran los tumores de mama ha sido asociada con un pronóstico favorable en algunos estudios. Sin embargo, la importancia de pronóstico de infiltrados linfoides en el sitio del tumor y la importancia relativa de estos infiltrados como un indicador del estado inmune del huésped sigue siendo controversial [14, 15, 16, 17]. Hasta ahora, para las pacientes individuales, no existe una relación predecible entre la composición de leucocitos y el pronóstico de la enfermedad.

Puesto que la terapia con Trastuzumab es costosa, la provisión de este fármaco a todas las pacientes con cáncer de mama con sobreexpresión de ERBB2 resultaría en una enorme carga económica para el sistema sanitario. En consecuencia, una prueba de diagnóstico capaz de distinguir entre los tumores de mama con sobreexpresión de ERBB2 con pronósticos buenos o malos podría ayudar a priorizar la urgencia del tratamiento con Trastuzumab.

Slamon et al. [9] mostraron que la amplificación del gen ERBB2 está asociada con un mayor riesgo de recaída y muerte en las pacientes con cáncer de mama en etapa inicial. Sin embargo, no se divulgó que el nivel de expresión de anticuerpos puede utilizarse como marcador de pronóstico en pacientes con cáncer ERBB2 positivo (ERBB2(+)).

El documento WO05/001138 describe la identificación de los patrones de expresión génica que pueden utilizarse para predecir la supervivencia de las personas afectadas con cáncer de mama y la probabilidad de recidiva del cáncer de mama, en particular en el subgrupo de pacientes de receptor de estrógeno (ER) positivo. Los procedimientos divulgados en WO05/001138 utilizan el nivel de expresión de genes marcadores para el pronóstico del cáncer de mama solamente en el subgrupo de pacientes de ER positivo. Sin embargo, no se divulga ni se sugiere una categorización alternativa de las pacientes en un grupo "ERBB2 positivo" y "ERBB2 negativo".

Disis et al. [18] encontraron que la sobreexpresión de proteína HER-2/neu se asocia con una enfermedad agresiva y es un indicador independiente de un mal pronóstico en varios subconjuntos de pacientes. Sin embargo, no se divulga que el nivel de anticuerpos presentes en el tejido tumoral puede utilizarse como indicador para el resultado de la enfermedad en un subconjunto de pacientes con cáncer que tienen tumores ERBB2(+).

Aaltomaa et al. [15] investigaron el valor de pronóstico de los infiltrados de linfocitos (LI) para la supervivencia sin recidiva (RFS) y la supervivencia relacionada con el cáncer de mama en pacientes con cáncer de mama. Se mostró que el valor de pronóstico de los LI podía mejorarse significativamente si las pacientes se clasificaban en grupos de pacientes con tumores "de proliferación rápida" y "de proliferación lenta", respectivamente. Tanto la determinación de la velocidad de proliferación como de la extensión de la infiltración de linfocitos son procedimientos experimentales difíciles de realizar, que requieren gran experiencia del personal de laboratorio.

Por consiguiente, existe la necesidad en la técnica de proporcionar procedimientos para la identificación de un resultado probable de cáncer de mama en una paciente que sean confiables y fáciles de llevar a cabo de manera experimental.

Sumario de la invención

La presente invención se basa en el hallazgo inesperado de que la clasificación de las muestras de tejido de cáncer de mama de ERBB2(+) en grupos de "buenos" y "malos resultados" puede lograrse mediante la medición del nivel de expresión de genes de inmunoglobulina o sus productos de proteínas en dichas muestras de tejido de cáncer de mama, por ejemplo, genes de anticuerpos de cadena pesada o liviana, preferiblemente del subtipo IgM o IgG. En procedimientos de la invención, el estado de ERBB2 se determina por procedimientos conocidos en la técnica, como la inmunohistoquímica (IHC), la hibridación fluorescente in situ (FISH), la hibridación cromogénica in situ (CISH), el análisis de expresión de ARNm o cualquier otro procedimiento apropiado. Pueden aplicarse técnicas experimentales similares para determinar la abundancia relativa de los transcritos de ARNm de inmunoglobulina o sus productos de proteínas en el tejido de cáncer de mama. Si el tumor se clasifica como ERBB2(+), el pronóstico de la paciente se

deduce posteriormente de la abundancia relativa de al menos un anticuerpo o inmunoglobulina en el tejido o la muestra del cáncer de mama.

La presente invención se relaciona, por lo tanto, con el procedimiento de acuerdo con la reivindicación 1 para la identificación de un resultado probable de cáncer de mama en una paciente con cáncer de mama que comprende las siguientes etapas: (a) determinar el estado de ERBB2 de un tumor de dicha paciente y (b) determinar en una muestra de tejido de cáncer de mama de dicha paciente el nivel de expresión de al menos una inmunoglobulina en dicho tumor de dicha paciente al menos si dicho estado de ERBB2 de dicho tumor es positivo, en el que dicha identificación de un resultado probable se basa en dicho nivel de expresión si el estado de ERBB2 de dicho tumor es positivo. Si se determina que el tumor es ERBB2(-), no se deducirá la identificación de un resultado probable del nivel de expresión de dicha inmunoglobulina en dicho tumor.

El uso de kits y aparatos para llevar a cabo los procedimientos anteriores son aspectos adicionales de la invención.

Breve descripción de las figuras

Figura 1: Esquema de clasificación para el pronóstico de tejidos de cáncer de mama de acuerdo con los procedimientos de la invención.

Figura 2: Clasificación de los tejidos de cáncer de mama en una clase positiva o negativa de ERBB2 sobre la base de las mediciones de expresión de ARN en la Cohorte I.

Figura 3: Curva de ROC para la determinación del corte de ARN de inmunoglobulina en la Cohorte I.

Figura 4: Curva de supervivencia de Kaplan-Meier para pacientes con cáncer de mama ERBB2 positivo clasificadas en el grupo de “buenos” o “malos resultados” en la Cohorte I.

Figura 5: Clasificación de tejidos de cáncer de mama en una clase de ERBB2 positivo o negativo sobre la base de las mediciones de la expresión de ARN en la Cohorte II.

Figura 6: Curva de Característica del Operador del Receptor (ROC) para la determinación del corte de ARN de inmunoglobulina en la Cohorte II.

Figura 7: Relación de singularidades de supervivencia libre de enfermedad a 60 meses (DFS) para la porción de ERBB2 positivo de la Cohorte II.

Descripción detallada de la invención

“Tejido de cáncer de mama”, de conformidad con la invención, se refiere al tejido obtenido del cuerpo humano mediante la resección o biopsia, que contiene células de cáncer de mama. Este tejido puede originarse a partir de un carcinoma in situ, un tumor primario invasivo, un tumor recurrente, ganglios linfáticos infiltrados mediante células tumorales o una lesión metastásica. El significado de “tejido de cáncer de mama” es independiente del “estado de ERBB2”, dentro del significado de la invención, se relaciona con la clasificación del tejido de cáncer de mama en una clase ERBB2 positivo (ERBB2(+)) o ERBB2 negativo (ERBB2(-)) basado en la determinación del nivel de expresión del gen ERBB2 según se mide por procedimientos como la inmunohistoquímica (IHC), la hibridación fluorescente in situ (FISH), la hibridación cromosómica in situ (CISH), el genochip u otros procedimientos conocidos por la persona entendida en la técnica. Dichos procedimientos se relacionan con la determinación de la proteína ERBB2, ADN o ARNm de un tejido de cáncer de mama. La clasificación inmunohistoquímica puede realizarse mediante el uso del HercepTest™ (DakoCytomation, Glostrup, Dinamarca). La lectura de los resultados de la reacción inmunohistoquímica de HER2 puede realizarse utilizando un sistema de puntuación en el cual el grado de amplificación génica corresponde a la reacción de tinción del HercepTest™. Como lo recomienda el fabricante de la prueba, la tinción debería estar solamente en la membrana y se evalúa en una escala de 0 a 3+, dependiendo de la intensidad de la tinción en más del 10% de las células tumorales invasivas. Se ignora cualquier tinción citoplasmática. 3+ es positivo (fuerte), 2+ es positivo (débil) y 1+ así como 0 es negativo. EL HercepTest™ se encuentra disponible en el mercado y está aprobado por la FDA para la predicción de la sensibilidad al Trastuzumab, al igual que las pruebas Pathway™ (Ventana Medical Systems, Tucson, AZ) e InSite™ (Biogenex, San Ramón, CA, Estados Unidos). Como alternativa a la detección inmunohistoquímica de HER2, se puede utilizar una prueba FISH para determinar en el nivel del gen si el ERBB2 se amplifica y el estado de ERBB2 es negativo o positivo de forma correspondiente. Entre las pruebas FISH disponibles en el mercado se incluyen la prueba HER2 FISH de PathVysion™ (Abbott-Vysis, Downers Grove, IL, Estados Unidos), la HER2 FISH de INFORM® (Ventana Medical Systems) y la HER2 FISH de PharmDX™ (Dako-Cytomation). Otra posibilidad para la determinación del estado de ERBB2 es el uso de PCR, en particular la PCR en tiempo real con cebadores y sondas específicos para el ADN de ERBB2, o alternativamente específicos para el ARNm que codifica el ERBB2.

“Inmunoglobulina”, de conformidad con la invención, se refiere a polipéptidos o transcritos de ARNm codificados por genes de inmunoglobulina. Las inmunoglobulinas pueden también ser partes o fragmentos de dichos polipéptidos o transcritos de ARNm. Los genes de inmunoglobulina codifican las moléculas de anticuerpo que son las moléculas de reconocimiento de antígenos producidas por las células B. Los transcritos de ARN de un gen de inmunoglobulina

pueden ser moléculas que codifican cadenas pesadas o livianas de anticuerpos. Los transcritos de cadena liviana pueden codificar la isoforma kappa o lambda de un anticuerpo, los transcritos de cadena pesada pueden codificar la isoforma μ o γ de un anticuerpo. Las regiones particularmente adecuadas para la detección y cuantificación de transcritos de inmunoglobulina son las secuencias de ácido nucleico que codifican los dominios constantes de cadenas pesadas o livianas, respectivamente. En la Tabla 1 se presenta una lista no exclusiva de secuencias de transcritos de inmunoglobulina ejemplares. La abundancia relativa de inmunoglobulinas en el tejido de cáncer de mama también puede determinarse midiendo la abundancia de péptidos codificados por genes de inmunoglobulina. Se puede aplicar la inmunohistoquímica para este propósito, utilizando moléculas de anticuerpos específicas para los transcritos de gen de inmunoglobulina como herramienta de detección. Las moléculas de anticuerpos particularmente adecuadas son específicas para las regiones constantes de las cadenas pesadas o livianas de inmunoglobulina.

“Abundancia relativa”, de conformidad con la invención, define la relación de la expresión total de una inmunoglobulina particular en una muestra de tejido tumoral en comparación con un valor umbral dado que es específico para la inmunoglobulina seleccionada y el procedimiento general, pero no para una muestra de tejido tumoral específica. Una expresión que es más alta que el valor umbral dado se indicará como “alta expresión”, mientras que una expresión que es más baja que el umbral dado se indicará como “baja expresión”.

“Determinación de un nivel de expresión” de un gen en un tejido, de conformidad con la invención, se entenderá como cualquier determinación de la cantidad de ARNm que codifica dicho gen, o una parte de dicho gen, en dicho tejido; o cualquier determinación de la cantidad del producto de la proteína codificado por dicho gen en dicho tejido. Se conocen varios procedimientos en la técnica para determinar el nivel de expresión de un gen en un tejido. Estos procedimientos comprenden, sin limitación, procedimientos PCR, procedimientos PCR en tiempo real, experimentos con genochip, inmunohistoquímica (IHC), hibridación fluorescente in situ (FISH), hibridación cromosómica in situ (CISH) y otros procedimientos conocidos por la persona entendida en la técnica.

“Genochip” o “micromatriz”, de conformidad con la invención, se entenderá como cualquier tipo de material de soporte sólido que comprende diversas características locales, cada una de las cuales comprende sondas de ácido nucleico inmovilizadas. Estas sondas de ácido nucleico son capaces de unirse a ácidos nucleicos libres en una muestra, en la que dicha unión puede detectarse mediante procedimientos adecuados. Varias implementaciones técnicas de genochips son conocidas por las personas entendidas en la técnica y se encuentran disponibles en el mercado. Un ejemplo muy conocido de un genochip es el GeneChip™ de Affymetrix, Inc. (Santa Clara, CA). La tecnología de micromatriz o genochip permite la medición del nivel de ARNm de estado estacionario de miles de genes simultáneamente, presentando de ese modo una herramienta poderosa para identificar efectos como el inicio, la detención o la modulación de la proliferación celular descontrolada. Actualmente hay dos tecnologías de micromatriz que son ampliamente utilizadas. La primera está compuesta por matrices de ADNc y la segunda se compone de matrices de oligonucleótidos. Si bien existen diferencias en la construcción de estos chips, esencialmente todos los análisis de datos de fase posterior y de salida son los mismos. El producto de estos análisis son, generalmente, mediciones de la intensidad de la señal recibida de una sonda marcada utilizada para detectar una secuencia de ADNc de la muestra que se hibrida con una secuencia de ácido nucleico en una ubicación conocida en la micromatriz. Por lo general, la intensidad de la señal es proporcional a la cantidad de ADNc y, por lo tanto el ARNm, expresado en las células de la muestra. Un gran número de dichas técnicas son conocidas por la persona entendida en la técnica y son útiles en el contexto de la presente invención.

“Clasificación”, de conformidad con la invención, denota el mapeo de una muestra desconocida a uno de al menos dos grupos predefinidos. En el contexto de la invención, la clasificación se lleva a cabo en base al nivel de expresión de ciertas inmunoglobulinas o en base al nivel de expresión del gen ERBB2.

“Buenos resultados” y “malos resultados”, de conformidad con la invención, se refieren a la clasificación de las pacientes con cáncer de mama en al menos dos grupos diferentes de acuerdo con la probabilidad de que el cáncer de mama tenga un resultado favorable o desfavorable. Las pacientes en un grupo de “buenos resultados” con frecuencia tendrán, por ejemplo, mejor supervivencia general, menor posibilidad de aparición de metástasis dentro de los 60 meses, etc.

La presente divulgación se refiere a procedimientos para la identificación de un resultado probable de cáncer de mama en una paciente con cáncer de mama que comprenden las siguientes etapas:

- (a) determinar el estado de ERBB2 de un tumor de dicha paciente, y
- (b) determinar el nivel de expresión de al menos una inmunoglobulina en dicho tumor de dicha paciente al menos si dicho estado de ERBB2 de dicho tumor es positivo, en los que dicha identificación de un resultado probable se basa en dicho nivel de expresión si el estado de ERBB2 de dicho tumor es positivo. Se entenderá que la denominación de las etapas del procedimiento como (a) y (b) no indicará que estas etapas tengan que ser realizadas en una cierta secuencia. Por el contrario, las etapas (a) y (b) pueden realizarse en cualquier orden cronológico o, preferiblemente, pueden realizarse al mismo tiempo, por ejemplo, en paralelo, como en un genochip,

basándose dicha identificación de un resultado probable en dicho nivel de expresión solamente si el estado de ERBB2 es positivo. Si se determina que el estado de ERBB2 en el tumor es negativo, el nivel de expresión de la inmunoglobulina puede descartarse y/o se pueden utilizar otros procedimientos para la identificación de un resultado probable.

- 5 De acuerdo con otra realización de la invención, la etapa de determinar el estado de ERBB2 de dicho tumor es anterior a la etapa de determinar el nivel de expresión de al menos una inmunoglobulina.

En realizaciones preferidas de la invención, en la etapa (b) se determina el nivel de expresión de más de una inmunoglobulina. Esto puede mejorar la calidad estadística y la fiabilidad del procedimiento de identificación.

En procedimientos de la divulgación, la identificación de un resultado probable se relaciona preferiblemente con

- 10 (a) una determinación de la probabilidad de recidiva de un tumor en dicha paciente, y/o
 (b) la predicción de la supervivencia a largo plazo de dicha paciente, y/o
 (c) la identificación de un resultado probable de cáncer de mama en una paciente con cáncer de mama luego de la cirugía, y/o
 (d) el pronóstico de cáncer de mama.

- 15 En procedimientos de la invención, la identificación del resultado probable comprende una clasificación de la paciente en uno de al menos dos grupos, correspondiendo dichos dos grupos a un grupo de "buenos resultados" y un grupo de "malos resultados", respectivamente. Un grupo de "buenos resultados" también puede ser visto como un grupo que tiene un "bajo riesgo" de recidiva de cáncer y un grupo de "malos resultados" puede, del mismo modo, considerarse como un grupo que tiene un "alto riesgo" de recidiva de cáncer. Estas expresiones pueden utilizarse de forma intercambiable.

- 20 En procedimientos de la invención, la determinación del estado de ERBB2 puede comprender una comparación de un nivel de expresión de ERBB2 con un valor de umbral predeterminado. Los tumores que tienen niveles de expresión de ERBB2 con medidas por encima de dicho nivel de umbral pueden categorizarse como "ERBB2 positivos" y los tumores que tienen niveles de expresión de ERBB2 con medidas iguales a, o por debajo de, dicho nivel de umbral pueden categorizarse como "ERBB2 negativos". El nivel de umbral preferiblemente está predeterminado, es decir, se da un valor fijo a dicho valor de umbral antes de la evaluación del nivel de expresión de ERBB2. Dicho valor predeterminado puede obtenerse de experimentos previos o de la literatura.

- 25 En procedimientos preferidos de la invención, dicho valor de umbral está determinado de un modelo de mezcla gaussiano de un histograma determinado de una cohorte adecuado de pacientes con cáncer de mama, en el que el histograma comprende los niveles de expresión de ERBB2 en tumores de dicha cohorte. En las Figs. 2 y 5 se muestra un ejemplo de cómo lograr esto. Un histograma se construye utilizando los niveles de expresión de ERBB2 en tumores de pacientes con cáncer de mama en una cohorte representativa de pacientes. El número relativo de pacientes en "intervalos" del histograma se representa gráficamente sobre el logaritmo natural del nivel de expresión de ERBB2. Se utilizó un modelo de mezcla gaussiano (GMM) para separar los dos modos de distribución bimodal. Los parámetros para las distintas distribuciones y el valor de umbral se estimaron con un algoritmo esperanza-maximización (EM), que es muy conocido por la persona entendida en la técnica. El valor de umbral puede obtenerse alternativamente utilizando otros procedimientos conocidos por el entendido en la técnica.

- 30 En realizaciones preferidas de la invención, dicha determinación de dicho nivel de expresión de al menos una inmunoglobulina comprende la determinación de la cantidad de ARNm que codifica dicha inmunoglobulina. Se puede emplear cualquier procedimiento adecuado para la determinación de un nivel de expresión conocido por el entendido en la técnica.

- 35 En otras realizaciones preferidas de la invención, dicha determinación de dicho nivel de expresión de al menos una inmunoglobulina comprende la determinación de la cantidad de ARNm que codifica al menos una parte de dicha inmunoglobulina.

- 40 En realizaciones preferidas adicionales de la invención, el estado de ERBB2 se determina mediante reacción en cadena de la polimerasa (PCR), o mediante PCR en tiempo real, o mediante inmunohistoquímica, o mediante hibridación fluorescente in situ (FISH), o mediante hibridación cromosómica in situ (CISH) o en un genochip. Si se desea, pueden emplearse varios de los procedimientos anteriores para la determinación.

- 45 En otras realizaciones preferidas de la invención, el nivel de expresión de dicha inmunoglobulina se determina mediante PCR, o mediante PCR en tiempo real, o mediante inmunohistoquímica o en un genochip.

Especialmente se prefieren los procedimientos en los cuales el estado de ERBB2 y el nivel de expresión de dicha inmunoglobulina se determinan al mismo tiempo, por ejemplo, en el mismo genochip. Esto ayuda a reducir trabajo y costos.

En una realización preferida de la invención, un cáncer se clasifica en el grupo de “malos resultados” si el nivel de expresión de dicha al menos una inmunoglobulina está por debajo de un nivel de umbral predeterminado.

5 En otros procedimientos preferidos, un cáncer se clasifica en el grupo de “malos resultados” si el valor combinado de los niveles de expresión (o el nivel de expresión medio) de al menos dos genes de inmunoglobulina está por debajo de un nivel de umbral predeterminado.

En procedimientos preferidos de la invención, la inmunoglobulina es un anticuerpo.

En realizaciones preferidas adicionales de la invención, dicha inmunoglobulina se selecciona del grupo que consiste en IGH@, IGHG1, IGHG4, IGHM, IGHD, IGHG3, IGJ, IGKC, IGLC2, IGLV@, IGLJ3, IGSF2, IGSF3, IGSF4B e IGSF4C, más preferiblemente del grupo que consiste en IGCK, IGHG1 e IGH@.

10 Otro aspecto de la invención se refiere al uso de un kit que comprende

(a) medios para la determinación del estado de ERBB2 de una paciente con cáncer de mama a partir de una muestra de tumor de dicha paciente,

(b) medios para la determinación del nivel de expresión de al menos una inmunoglobulina en una muestra de tumor de dicha paciente, y

15 (c) material impreso o un portador de datos que comprenden instrucciones para realizar un procedimiento de acuerdo con cualquiera de las reivindicaciones 1-13.

Dichos kits pueden utilizarse para realizar los procedimientos de la invención mencionados anteriormente. Preferiblemente, dichos medios para la determinación del estado de ERBB2 y dichos medios para la determinación del nivel de expresión de al menos una inmunoglobulina son los mismos. Dichos medios pueden comprender, por ejemplo, un genochip.

20 Otro aspecto de la invención se refiere al uso de un aparato para la identificación de un probable resultado de cáncer de mama en una paciente con cáncer de mama, comprendiendo dicho dispositivo

(a) medios para la determinación del estado de ERBB2 de una paciente con cáncer de mama a partir de una muestra de tumor de dicha paciente,

25 (b) medios para la determinación del nivel de expresión de al menos una inmunoglobulina en una muestra de tumor de dicha paciente, y

(c) medios informáticos capaces de identificar dicho resultado probable, en los que dicha identificación se basa en el nivel de expresión de dicha al menos una inmunoglobulina si dicho estado de ERBB2 es positivo.

30 Los medios informáticos se conectan preferiblemente con los medios para la determinación del estado de ERBB2 y con los medios para la determinación del nivel de expresión de al menos una inmunoglobulina mediante medios de transmisión de señal, de modo que el resultado de las respectivas determinaciones esté disponible fácilmente para los medios informáticos para la identificación de dicho resultado probable de dicho cáncer de mama.

35 En procedimientos preferidos de la invención, la determinación del estado de ERBB2 del tumor y la determinación del nivel de expresión de la al menos una inmunoglobulina se realizan en muestras de dichas pacientes. Se prefiere que estas muestras estén disponibles previamente.

Se prefiere que la determinación del estado de ERBB2 del tumor y la determinación del nivel de expresión de la al menos una inmunoglobulina sean in vitro.

40 La identificación de un resultado probable del cáncer de mama está determinada por separado a partir del diagnóstico de la enfermedad, por ejemplo luego de que el cáncer de mama haya sido diagnosticado en dicha paciente. En realizaciones preferidas de la invención, la identificación no implica el diagnóstico.

La invención se ilustra adicionalmente por medio de los siguientes ejemplos.

Ejemplos

Ejemplo 1: Pronóstico del cáncer de mama ERBB2 positivo – Cohorte I

45 Selección de la paciente y aislamiento de ARN a partir de secciones de tejido tumoral congeladas y medición de la expresión génica utilizando micromatrices HG-U133A de Affymetrix

Se tomaron las secciones congeladas para la histología y la presencia del cáncer de mama se confirmó en muestras de 205 pacientes con cáncer de mama. El contenido de células tumorales excedió el 30% en todos los casos y estuvo por encima del 50% en la mayoría de los casos. Se molieron aproximadamente 50 mg de tejido tumoral de mama congelado al instante en nitrógeno líquido. Se agregó la solución amortiguadora RLT (QIAGEN, Hilden,

Alemania) y se hizo girar el homogeneizado a través de una columna QIAshredder (QIAGEN, Hilden, Alemania). A partir del eluido, se aisló el ARN total mediante el Kit RNeasy (QIAGEN, Hilden, Alemania) de acuerdo con las instrucciones del fabricante. El rendimiento de ARN se determinó mediante la absorbancia de UV y la calidad de ARN se evaluó mediante análisis de integridad de la banda de ARN ribosómico en el Agilent Bioanalyzer (Palo Alto, CA, Estados Unidos).

Comenzando con 5 µg de ARN total se preparó ARNc marcado para las 205 muestras de tumor utilizando los kits de Roche de acuerdo con las instrucciones del fabricante: Microarray cDNA Synthesis, Microarray RNA Target Synthesis (T7) y Microarray Target Purification. Brevemente, la síntesis de la primera hebra de ADNc fue realizada por un cebador oligo-dT enlazado a T7, seguida por la síntesis de la segunda hebra. El producto de ADNc de doble hebra se purificó y luego se utilizó como plantilla para una reacción de transcripción in vitro (IVT) en presencia de UTP biotinilado. Se hibridó el ARNc marcado con matrices HG-U133A (Santa Clara, CA, Estados Unidos) a 45° por 16 h en un horno de hibridación a una rotación constante (60 r.p.m.) y luego se lavó y se tiñó con un conjugado de estreptavidina-ficoeritrina utilizando la estación de fluidos GeneChip. Las matrices se escanearon a 560 nm utilizando el escáner GeneArray G2500A de Hewlett Packard. Las lecturas del escaneo cuantitativo se analizaron utilizando la Microarray Analysis Suit 5.0 (MAS 5.0) de Affymetrix. En las configuraciones de análisis se eligió el procedimiento de escala global que multiplicó las intensidades de señal de salida de cada matriz hasta una intensidad objetivo media de 500. Las imágenes de matriz fueron inspeccionadas visualmente para detectar defectos y su calidad se controló utilizando el software Refiner de GeneData. De forma rutinaria se obtuvieron más del 50% de llamadas presentes por chip como fue calculado mediante MAS 5.0.

Determinación del estado de ERBB2 sobre la base de la medición de ARNm

En las 205 muestras de esta cohorte se midieron los niveles de expresión génica. Como en el ejemplo anterior, se utilizó un modelo de mezcla gaussiano en la expresión absoluta logaritimizada (natural) de ERBB2 (juego de sondas ID 216836_s_at de Affymetrix) para separar los dos modos de la distribución. De los 205 valores de expresión, se encontró que un valor de umbral de 8,49 en una escala logarítmica (natural) (aproximadamente 4880 en una escala lineal) era el más probable.

Las muestras que exhibieron una expresión igual o superior a dicho valor se marcaron como “ERBB2 positivo”, las muestras restantes se marcaron como “ERBB2 negativo”. De las 205 muestras, 24 muestras fueron ERBB2(+) y las 191 muestras que habían sido marcadas como ERBB2 negativo no se investigaron adicionalmente.

Clasificación de muestras de cáncer de mama ERBB2 positivas en grupos de “buenos resultados” y “malos resultados”

Para las muestras de ERBB2(+), se midieron los perfiles de expresión de genes relacionados con el sistema inmune mediante HG_U133A GeneChip™. Los genes medidos se listan en la Tabla 1, junto con un identificador único para el juego de sondas respectivo en el Affymetrix GeneChip™.

Tabla 1:

ID del juego de sondas	Símbolo del gen	Nombre(s) del gen
214973_x_at	IGHD	delta constante pesada de inmunoglobulina
216892_at	IGH@	sitio pesado de inmunoglobulina
216541_x_at	IGHG1	gamma 1 constante pesada de inmunoglobulina (marcador GIm)
212592_at	IGJ	polipéptido J de inmunoglobulina, proteína enlazante para polipéptidos de inmunoglobulina alfa y mu
215621_s_at	IGHD	delta constante pesada de inmunoglobulina
211648_at	IGHM	mu constante pesada de inmunoglobulina /// mu constante pesada de inmunoglobulina
217198_x_at	IGH@ ///	sitio pesado de inmunoglobulina /// gamma 1 constante pesada de inmunoglobulina (marcador GIm)
217442_at	IGSF4B	superfamilia de inmunoglobulina, miembro 4B
214677_x_at	IGLC2	constante 2 lambda de inmunoglobulina (marcador Kern-Oz)
211640_x_at	IGHG1	gamma 1 constante pesada de inmunoglobulina (marcador GIm) /// gamma 1 constante pesada de inmunoglobulina (marcador GIm)
211634_x_at	IGHM	mu constante pesada de inmunoglobulina /// mu constante pesada de inmunoglobulina
209374_s_at	IGHM	mu constante pesada de inmunoglobulina
215035_at	IGLC2	lambda 2 constante de inmunoglobulina (marcador Kern-Oz) /// región variable SAR de cadena liviana
221651_x_at	IGKC	lambda 6 amiloide constante kappa de inmunoglobulina

ES 2 375 998 T3

221671_x_at	IGKC	constante kappa de inmunoglobulina
215121_x_at	IGLC2	lambda 2 constante de inmunoglobulina (marcador Kern-Oz)
2H430_s_at	IGHG1	gamma 1 constante pesada de inmunoglobulina (marcador G1m)
211647_x_at	IGHG1	gamma 1 constante pesada de inmunoglobulina (marcador G1m) /// gamma 1 constante pesada de inmunoglobulina (marcador G1m)
202421_at	IGSF3	superfamilia de inmunoglobulina, miembro 3
222293_at	IGSF4C	superfamilia de inmunoglobulina, miembro 4C
207167_at	IGSF2	superfamilia de inmunoglobulina, miembro 2
217369_at	IGH@ ///	región V-D-J de ARNm de cadena pesada redispuesta de IgM /// sitio pesado de inmunoglobulina
	IGHG1 ///	/// gamma 1 constante pesada de inmunoglobulina (marcador G1m) /// gamma 4 constante pesada de inmunoglobulina (marcador G4m) /// VH3 de cadena pesada de inmunoglobulina (H11)
	IGHG4 ///	/// cadena mu Ig redispuesta (V6-1/Dxp4-DA5/JH6b) /// Similar a cadena M de inmunoglobulina /// sitio transcrito, moderadamente similar a XP_234649.1 similar a región variable de cadena pesada de inmunoglobulina [Rattus norvegicus] /// mu constante pesada de inmunoglobulina
	IGHM	/// Similar a cadena M de inmunoglobulina /// sitio transcrito, moderadamente similar a XP_234649.1 similar a región variable de cadena pesada de inmunoglobulina [Rattus norvegicus] /// mu constante pesada de inmunoglobulina
211146_at	IGKC	constante kappa de inmunoglobulina
215721_at	IGHG1	gamma 1 constante pesada de inmunoglobulina (marcador G1m)
211798_x_at	IGLJ3	lambda de inmunoglobulina de unión 3
216846_at	IGLC2 ///	región v de cadena liviana lambda de inmunoglobulina (Humla203) /// sitio transcrito,
	IGLV@ ///	levemente similar a XP_213585.2 similar a cadena liviana M de inmunoglobulina anti-idiotipo [Rattus norvegicus] /// cadena liviana de lambda E51 de inmunoglobulina Anti-HIV-1gp120 /// constante 2 lambda de inmunoglobulina (marcador Kern-Oz) /// ARNm de cadena lambda de inmunoglobulina, regiones VJ, hibridoma JB32 /// grupo variable lambda de inmunoglobulina /// dominio variable de anticuerpo de antígeno de superficie de hepatitis B/// región variable de cadena liviana de inmunoglobulina TASS aislada /// lambda de inmunoglobulina de unión 3 /// región V de cadena lambda Ig (VL-HUC)
211644_x_at	IGKC	HRV Fab 027-VL /// HRV Fab 027-VL /// Similar a secuencia de ADNC BC012256 /// Similar a secuencia de ADNC BC012256 /// región V de cadena kappa redispuesta de inmunoglobulina del antígeno de la rabia /// región V de cadena kappa redispuesta de inmunoglobulina del antígeno de la rabia /// HRV Fab 026-VL /// HRV Fab 026-VL /// constante kappa de inmunoglobulina /// constante kappa de inmunoglobulina /// región variable de cadena kappa de anticuerpo monoclonal /// región variable de cadena kappa de anticuerpo monoclonal /// región V de cadena kappa Ig (VL-COL) /// región V de cadena kappa Ig (VL-COL)
211881_x_at	IGLJ3	lambda de inmunoglobulina de unión 3
217169_at	IGHG3	región VDJ de IgM /// gamma 3 constante pesada de inmunoglobulina (marcador G3m)
216535_at	IGSF4B	superfamilia de inmunoglobulina, miembro 4B

Un análisis adicional se basó en el nivel de expresión de la cadena kappa de inmunoglobulina (IGKC; ID de juego de sondas de Affymetrix 221671_x_at). De las 24 muestras de ERBB2(+), solo 21 pudieron utilizarse debido a que las 4 pacientes restantes tuvieron un seguimiento de menos de 60 meses, por lo que la información de metástasis a los 60 meses del seguimiento permaneció desconocida.

5

Para un valor de umbral elegido para la expresión de IGKC, las muestras que tenían un valor de expresión más bajo

con respecto a este umbral fueron marcadas como "IGKC bajo" y las que tenían valores de expresión más altos se marcaron como "IGKC alto". Se asumió una especificidad de al menos un 90% (que, en este caso, permitió que ningún caso sea clasificado erróneamente como "buenos resultados") y se eligió el valor de umbral que permitió la mayor sensibilidad (bajo la restricción de especificidad del 90%).

- 5 En el caso de las 21 muestras de ERBB2(+), se determinó que un valor de umbral era 14424 y se obtuvo una sensibilidad máxima del 75% (con el grupo de "malos resultados" integrado por las muestras con una expresión de IGKC menor a 14424 y el grupo de "buenos resultados" compuesto por las muestras con una expresión de IGKC mayor a 14424). Dado que no se permitieron falsos negativos, no pudo computarse una relación de probabilidades de manera significativa porque su valor sería infinito. Las curvas Kaplan-Meier se presentan en la Figura 4.
- 10 El pronóstico general, por lo tanto, se realiza de la siguiente manera: primero se toman las medidas de ERBB2 e IGKC. Si el valor de expresión de ERBB2 está por encima de 4880, la muestra se marca como "ERBB2(+)", de lo contrario, se marca como "ERBB2(-)". Si el estado de ERBB2(+) es positivo, la expresión de IGKC se compara con el valor de umbral de 14424. Las muestras que exhiben una expresión más alta que este valor se consideran "buenos resultados", las muestras que exhiben un nivel de expresión igual o menor a 14424 se consideran "malos resultados".

Mientras que el análisis anterior se basa en el gen IGKC, también pueden utilizarse para el análisis otros genes que codifican inmunoglobulinas (por ejemplo, genes listados en la Tabla 1). La Tabla 2 proporciona el "área bajo la curva" en los diagramas de Característica del Operador (cf. Figura 3), la máxima sensibilidad a la especificidad del 90% y el valor de corte correspondiente para las mediciones del nivel de expresión para los genes listados en la Tabla 1. También se proporciona en la Tabla 2 la especificidad máxima a una sensibilidad del 90% y el corte correspondiente para el nivel de expresión, como se determina en este ejemplo. Los genes que proporcionan una sensibilidad a una especificidad del 90% o mayor al 0,3 se clasifican del 1 al 6, en donde un rango bajo indica una alta sensibilidad. Se prefieren los genes que proporcionan una alta sensibilidad, es decir, en el rango 1-6. Como lo apreciará la persona entendida en la técnica, en el análisis podrán utilizarse más de uno de dichos genes a la vez.

25

Tabla 2: Cohorte I (24 Muestras)

ID del juego de sondas	Símbolo del gen	Rango	Área bajo la curva (AUC)	Sensibilidad máxima (especificidad del 90%)	Corte	Sensibilidad máxima (especificidad del 90%)	Corte
214973_x_at	IGHD	1	0,820513	0,692308	312,9	0,222222	156,7
216892_at	IGH@	2	0,820513	0,538462	81,4	0,333333	16,6
216541_x_at	IGHG1	4	0,820513	0,384615	92,3	0,555556	15,3
212592_at	IGJ	3	0,803419	0,461538	1254,0	0,444444	272,4
215621_s_at	IGHD	3	0,786325	0,461538	47,3	0,444444	6,9
211648_at	IGHM	3	0,709402	0,461538	28,6	0,222222	6,1
217198_x_at	IGH@ /// IGHG1	3	0,649573	0,461538	97,8	0,222222	19,7
217442_at	IGSF4B		0,632479	0,153846	27,2	0,222222	5,2
214677_x_at	IGLC2		0,726496	0,153846	24897,0	0,333333	5089,0
211640_x_at	IGHG1	4	0,863248	0,384615	130,9	0,555556	16,1
211634_x_at	IGHM	4	0,820513	0,384615	432,7	0,666667	75,2
209374_s_at	IGHM	4	0,760684	0,384615	4354,0	0,111111	299,3
215035_at	IGLC2		0,683761	0	65,0	0	4,1
221651_x_at	IGKC		0,794872	0,153846	25414,0	0,555556	10630,0
221671_x_at	IGKC		0,794872	0,153846	26122,0	0,555556	10704,0
215121_x_at	IGLC2		0,760684	0,153846	32860,0	0,444444	4391,0
211430_s_at	IGHG1		0,769231	0,230769	22246,0	0,444444	6221,0
211647_x_at	IGHG1	6	0,752137	0,307692	93,2	0,111111	10,0
202421_at	IGSF3	5	0,726496	0,333333	1009,0	0,230769	707,4
222293_at	IGSF4C	5	0,717949	0,333333	80,0	0,153846	14,6
207167_at	IGSF2	6	0,803419	0,307692	128,8	0,555556	49,0

(Cont.)							
217369_at	IGH@ /// IGHG1 /// IGHG4 /// IGHM	6	0,803419	0,307692	474,9	0,555556	213,5
211146_at	IGKC	6	0,615385	0,307692	8,6	0,222222	4,1
215721_at	IGHG1	6	0,521368	0,307692	21,3	0	5,1
211798_x_at	IgLJ3		0,726496	0,153846	2100,0	0,444444	190,7
216846_at	IGLC2 /// IGLV@ /// IGLJ3		0,606838	0	115,8	0	9,4
211644_x_at	IGKC		0,820513	0,153846	5594,0	0,666667	445,8
211881_x_at	IgLJ3		0,717949	0,153846	1035,0	0,444444	188,7
217169_at	IGHG3		0,649573	0,111111	57,1	0,153846	5,0
216535_at	IGSF4B		0,598291	0	159,4	0	27,9

Ejemplo 2: Pronóstico de muestras de cáncer de mama ERBB2 positivo – Cohorte II

5 Selección de pacientes, aislamiento de ARN a partir de secciones de tejido tumoral congeladas y medición de la expresión génica utilizando micromatrices HG-U133A de Affymetrix

Las 286 muestras congeladas de tejido de cáncer de mama se seleccionaron del banco de tumores del Centro Médico Erasmus (Róterdam, Países Bajos) y se procesaron como se describió en Wang et al. (2005. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer, Lancet 365: 671-679). Brevemente, el ARN total se aisló a partir de 20-40 secciones de criostato de 30 µm de espesor (50-100 mg) con RNazol B (Campro Scientific, Veenendaal, Países Bajos). Los objetivos biotinilados se prepararon mediante procedimientos publicados (Affymetrix, Santa Clara, CA, Estados Unidos) y se hibridaron con la micromatriz de oligonucleótidos U133a GeneChip de Affymetrix. Las matrices se escanearon mediante protocolos convencionales de Affymetrix. Cada juego de sondas se trató como un gen separado. Los valores de expresión se calcularon mediante el uso del software de análisis MAS 5.0 GeneChip de Affymetrix. Se rechazaron los chips con una intensidad media menor a 40 o una señal de fondo de más de 100. Para la normalización del chip, los juegos de sondas se escalaron a una intensidad de objetivo de 600 y no se seleccionaron archivos de máscara de escala.

Determinación del estado de ERBB2 sobre la base de la medición de ARNm

Se tomaron datos de las 286 muestras de tejido. Se aplicó un modelo de mezcla gaussiano para el logaritmo de la expresión absoluta de ERBB2 (ID de sonda de Affymetrix: 216836_s_at) para separar los dos modos de la distribución bimodal. De acuerdo con este modelo, un valor de umbral óptimo para los datos en consideración era una expresión de ERBB2 absoluta de aproximadamente 7254. Las muestras que exhibieron un nivel de expresión de dicho gen particular que fue al menos tan alto como el umbral elegido fueron marcadas como “ERBB2 positivo”, las muestras restantes se marcaron en consecuencia como “ERBB2 negativo”. De las 286 muestras proporcionadas, se encontró que 51 estaban en el grupo ERBB2(+). Las 235 muestras restantes estaban en ERBB2(-), por lo que no se consideraron adicionalmente.

Clasificación de muestras de cáncer de mama ERBB2 positivas en grupos de “buenos resultados” y “malos resultados”

Para las muestras de ERBB2(+), se midieron los niveles de expresión de los genes listados en la Tabla 1.

30 Para cada uno de los genes, se definió un operador de clasificación mediante el marcado de los genes con respecto a su expresión relativa a un umbral elegido. El umbral óptimo para cada gen se determinó sobre la base de una característica de operador derivada de elegir diferentes valores de umbral y computar la sensibilidad y especificidad con respecto a la información de metástasis para cada muestra. Un ejemplo de IGKC (cadena kappa de inmunoglobulina) se muestra en la Figura 6.

35 Se permitió una especificidad del 90% o mayor y se eligió el valor de umbral que proporciona la mejor sensibilidad. En el caso de IGKC, la mejor sensibilidad fue de aproximadamente el 30%. Si bien este valor puede parecer bajo, la estadística total de caso/testigo proporcionada en la Figura 7 permite concluir que de hecho se tiene una relación de probabilidades del clasificador de 8,5 si identificamos BAJO (es decir el grupo que tiene una expresión de IGKC que está por debajo del valor de umbral de 37485,4) como un “grupo de malos resultados” y en consecuencia ALTO como un grupo de “buenos resultados” (que contiene pacientes con una expresión de IGKC de más de 37485,4).

ES 2 375 998 T3

- 5 Por lo tanto, para el pronóstico general se procedería de la siguiente manera: primero se toman las medidas de ERBB2 e IGKC. Si el valor de expresión de ERBB2 está por encima del valor de umbral de 7254, la muestra se marca como "ERBB2(+)", de lo contrario se la marca como "ERBB2(-)". Para las muestras positivas de ERBB2, la expresión de IGKC se compara con el umbral de 37485,4. Las muestras que exhiben una expresión más alta que este valor se marcan como muestras de "buenos resultados" y las muestras con un nivel de expresión más bajo que 37485,4 se marcan como de "malos resultados".
- 10 Mientras que el análisis anterior se basa en el gen IGKC, para el análisis también se pueden utilizar otros genes que codifican inmunoglobulinas (por ejemplo, genes listados en la Tabla 1). La Tabla 3 proporciona el "área bajo la curva" en los diagramas de Característica del Operador, la máxima sensibilidad a la especificidad del 90% y el valor de corte correspondiente para las mediciones del nivel de expresión para los genes listados en la Tabla 1. También se proporcionan en la Tabla 3 la máxima especificidad al 90% de sensibilidad y el corte correspondiente para el nivel de expresión, como se determina en este ejemplo. Los genes que proporcionan una sensibilidad a una especificidad del 90% o mayor a 0,3 se clasifican del 1 al 7, en donde un rango bajo indica una alta sensibilidad. En el contexto de la invención se prefieren los genes que proporcionan una alta sensibilidad, por ejemplo, en el rango 1-7.
- 15 IGCK, IGHG1 e IGH@ proporcionan una alta sensibilidad a una especificidad del 90% en los ejemplos 1 y 2 (es decir, cohortes I y II). Por lo tanto, en los procedimientos de la invención se prefieren estos genes.

Tabla 3: Cohorte II (51 muestras)

ID del juego de sondas	Símbolo del gen	Rango	Área bajo la curva (AUC)	Sensibilidad máxima (especificidad del 90%)	Corte	Sensibilidad máxima (especificidad del 90%)	Corte
214973_x_at	IGHD		0,752525	0,242424	687,8	0,333333	94,6
216892_at	IGH@		0,641414	0,181818	304,9	0,277778	13,2
216541_x_at	IGHG1	1	0,732323	0,484848	120,9	0,222222	15,0
212592_at	IGJ		0,712121	0,242424	2773,6	0,444444	371,8
215621_s_at	IGHD		0,60101	0,212121	74,7	0,166667	9,3
211648_at	IGHM		0,540404	0,166667	72,9	0,060606	7,3
217198_x_at	IGH@ /// IGHG1	6	0,624579	0,30303	170,3	0,166667	17,5
217442_at	IGSF4B	2	0,737374	0,454545	17,5	0,166667	9,6
214677_x_at	IGLC2	3	0,759259	0,393939	37438,0	0,444444	9551,5
211640_x_at	IGHG1		0,66835	0,212121	164,2	0,333333	13,1
211634_x_at	IGHM		0,646465	0,212121	487,1	0,333333	26,1
209374_s_at	IGHM		0,685185	0,212121	7064,2	0,388889	618,2
215035_at	IGLC2	4	0,671717	0,363636	23,1	0,166667	5,6
221651_x_at	IGKC	5	0,777778	0,333333	37254,1	0,444444	11564,1
221671_x_at	IGKC	5	0,774411	0,333333	35208,7	0,444444	10851,4
215121_x_at	IGLC2	5	0,740741	0,333333	35894,3	0,388889	3257,7
211430_s_at	IGHG1	5	0,703704	0,333333	39837,2	0,111111	796,2
211647_x_at	IGHG1	5	0,648148	0,333333	112,8	0,055556	12,5
202421_at	IGSF3		0,587542	0,222222	2285,0	0,333333	1205,7
222293_at	IGSF4C		0,654882	0,151515	153,3	0,277778	24,6
207167_at	IGSF2		0,550505	0,030303	195,4	0,111111	42,4
207369_at	IGH@ /// IGHG1 /// IGHG4 /// IGHM		0,572391	0,121212	493,1	0,222222	142,4
211146_at	IGKC	7	0,680135	0,272727	15,7	0,111111	5,4
215721_at	IGHG1		0,510101	0,166667	66,9	0,060606	8,8
211798_x_at	IGLJ3	6	0,757576	0,30303	974,6	0,5	137,5
216846_at	IGLC2 /// IGLV@ ///	6	0,700337	0,30303	32,2	0,166667	14,4

(Cont.)

	IGLJ3						
211644_x_at	IGKC	7	0,744108	0,272727	2760,2	0,277778	65,6
211881_x_at	IGLJ3	7	0,740741	0,272727	989,7	0,388889	159,3
217169_at	IGHG3	7	0,739057	0,272727	69,2	0,222222	14,3
216535_at	IGSF4B	7	0,617845	0,272727	191,5	0,222222	46,1

Bibliografía citada

- 5 (1) Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thulimann B, Senn HJ. Meeting Highlights: updated international expert consensus on the primary therapy of early breast cancer. *J Clin Oncol* 21: 3357-3365, 2003
- (2) Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* 352: 930-942, 1998
- 10 (3) Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 351: 1451-1467, 1998
- (4) Shapiro CL and Recht A. Side effects of adjuvant treatment of breast cancer. *N Engl J Med* 344: 1997-2008, 2001
- (5) Altman DG and Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer. *Br Cancer Res Treat* 52: 289-303, 1998
- 15 (6) Jatoli I, Hilsenbeck SG, Clark GM, Osborne CK. Significance of axillary lymph node metastasis in primary breast cancer. *J Clin Oncol* 17: 2334-2340, 1999
- (7) Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415: 530-536, 2002
- 20 (8) Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA et al. Molecular portraits of human breast tumours. *Nature* 406: 747-752, 2000
- (9) Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL: Human breast cancer: correlation of relapse and survival with amplification of the ERBB2 oncogene. *Science* 1987, 235: 177-181 1
- (10) Reese D, Slamon DJ: ERBB2 signal transduction in human breast and ovarian cancer. *Stem Cells* 1997, 15:1-8
- (11) Reese D, Slamon DJ: ERBB2 signal transduction in human breast and ovarian cancer. *Stem Cells* 1997, 15:1-8
- 25 (12) Salido M, Solé F, Tusquets I, Coromines JM, Espinet B, Marinosa ML, T Baró, MC Vela, X Fabregat, S Serrano, et al.: A comparative study of ERBB2 amplification and over expression using fluorescence in situ hybridization (FISH) and immunohistochemistry (IHC) in 101 breast cancer patients. *Oncologia* 2002, 4:255-259
- 30 (13) Bose S, Mohammed M, Shintaku P, Rao PN: ERBB2 gene amplification in low to moderately expressing breast cancers: possible role of chromosome 17/ERBB2 polysomy. *Breast J* 2001, 7:337-344
- (14) An T, Sood U, Pietruk T, Cummings G, Hashimoto K, Crissman JD: In situ quantitation of inflammatory mononuclear cells in ductal infiltrating breast carcinoma. *Am J Pathology* 1987, 128:52-60
- 35 (15) Aaltomaa S, Lipponen P, Eskelinen M, Kosma VM, Marin S, Alhava E, Syrjänen K: Lymphocyte infiltrates as a prognostic variable in female breast cancer. *Eur J Cancer* 1992, 28:859-864
- (16) Menard S, Tomasic G, Casalini P, Balsari A, Pilotti S, Cascinelli N, Salvadori B, Colnaghi MI, Rilke F: Lymphoid infiltration as a prognostic variable for early-onset breast carcinomas. *Clin Can Res* 1997, 3:817-819
- 40 (17) Lespagnard L, Gancberg D, Rouas G, Leclercq G, Somerhausen N De S, Di Leo A, Piccart M, Verhest A, Larsimont D: Tumour-infiltrating dendritic cells in adenocarcinomas of the breast: A study of 143 neoplasms with a correlation to usual prognostic factors and to clinical outcome. *Int J Cancer* 1999, 84:309-314
- (18) Disis et al: Existent T-cell and antibody immunity to HER-2/neu protein in patients with breast cancer. *Cancer Res* 1994, 54:16-20

REIVINDICACIONES

1. Un procedimiento para la identificación de un resultado probable de cáncer de mama en una paciente con cáncer de mama que comprende las etapas de
- (a) determinar el estado de ERBB2 de un tumor de dicha paciente, y
- 5 (b) determinar en una muestra de tejido de cáncer de mama de dicha paciente el nivel de expresión de al menos una inmunoglobulina en dicho tumor de dicha paciente al menos si dicho estado de ERBB2 de dicho tumor es positivo,
- en el que dicha identificación de un resultado probable se basa en dicho nivel de expresión si el estado de ERBB2 de dicho tumor es positivo,
- 10 en el que la identificación del resultado probable comprende una clasificación de la paciente en uno de al menos dos grupos, correspondiendo dichos dos grupos a un grupo de "buenos resultados" y un grupo de "malos resultados", respectivamente, y en el que un cáncer se clasifica como perteneciente al grupo de "malos resultados" si el nivel de expresión de dicha al menos una inmunoglobulina está por debajo de un nivel de umbral predeterminado.
2. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que el etapa de determinar el estado de ERBB2 de dicho tumor es anterior a la etapa de determinar el nivel de expresión de al menos una inmunoglobulina.
- 15 3. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que en la etapa (b) se determina el nivel de expresión de más de una inmunoglobulina.
4. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que la identificación del resultado probable es
- 20 (a) una determinación de la probabilidad de recidiva de un tumor en dicha paciente, y/o
- (b) la predicción de la supervivencia a largo plazo de dicha paciente, y/o
 - (c) luego de la cirugía, y/o es
 - (d) el pronóstico de cáncer de mama.
- 25 5. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que la determinación del estado de ERBB2 comprende una comparación de un nivel de expresión de ERBB2 con un valor de umbral predeterminado.
6. El procedimiento de la reivindicación 5, en el que dicho valor de umbral se determina a partir de un modelo de mezcla gaussiano de un histograma determinado a partir de una cohorte adecuada de pacientes con cáncer de mama, en el que el histograma comprende los niveles de expresión de ERBB2 en tumores de dicha cohorte.
- 30 7. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que dicha determinación de dicho nivel de expresión de al menos una inmunoglobulina comprende la determinación de la cantidad de ARNm que codifica dicha inmunoglobulina.
8. Un procedimiento de acuerdo con la reivindicación 7, en el que dicha determinación de dicho nivel de expresión de al menos una inmunoglobulina comprende la determinación de la cantidad de ARNm que codifica al menos partes de dicha inmunoglobulina.
- 35 9. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que el estado de ERBB2 se determina mediante PCR, o mediante PCR en tiempo real, o mediante inmunohistoquímica, o mediante hibridación fluorescente in situ (FISH), o mediante hibridación cromosómica in situ (CISH) o en un genochip.
- 40 10. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que el nivel de expresión de dicha inmunoglobulina se determina mediante PCR, o PCR en tiempo real, o mediante inmunohistoquímica o en un genochip.
11. Un procedimiento de cualquiera de las reivindicaciones precedentes, en el que el estado de ERBB2 y el nivel de expresión de dicha inmunoglobulina se determinan en el mismo genochip.
- 45 12. Un procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que dicha inmunoglobulina se selecciona del grupo que consiste en IGH@, IGHG1, IGHG4, IGHM, IGHD, IGHG3, IGJ, IGKC, IGLC2, IGLV@, IGLJ3, IGSF2, IGSF3, IGSF4B e IGSF4C.
13. Un procedimiento de acuerdo con la reivindicación 12, en el que la inmunoglobulina es IGCK o IGHG1 o IGH@.

14. El uso de un kit para realizar el procedimiento de acuerdo con cualquiera de las reivindicaciones 1-13, comprendiendo dicho kit
- (a) medios para la determinación del estado de ERBB2 de una paciente con cáncer de mama a partir de una muestra de tumor de dicha paciente,
- 5
- (b) medios para la determinación del nivel de expresión de al menos una inmunoglobulina en una muestra de tumor de dicha paciente, y
 - (c) material impreso o un portador de datos que comprenden instrucciones para realizar un procedimiento de acuerdo con cualquiera de las reivindicaciones 1-13.
- 10
15. El uso de un aparato para realizar el procedimiento de acuerdo con cualquiera de las reivindicaciones 1-13, comprendiendo dicho aparato
- (a) medios para la determinación del estado de ERBB2 de una paciente con cáncer de mama a partir de una muestra de tumor de dicha paciente,
 - (b) medios para la determinación del nivel de expresión de al menos una inmunoglobulina en una muestra de tejido de cáncer de mama de dicha paciente, y
- 15
- (c) medios informáticos adaptados para la identificación de dicho resultado probable, en donde dicha identificación se basa en el nivel de expresión de al menos una inmunoglobulina si dicho estado de ERBB2 es positivo,
- 20
- en el que la identificación del resultado probable comprende una clasificación de la paciente en uno de al menos dos grupos, correspondiendo dichos dos grupos a un grupo de "buenos resultados" y un grupo de "malos resultados", respectivamente, y en el que un cáncer se clasifica como perteneciente al grupo de "malos resultados" si el nivel de expresión de dicha al menos una inmunoglobulina está por debajo de un nivel de umbral predeterminado.

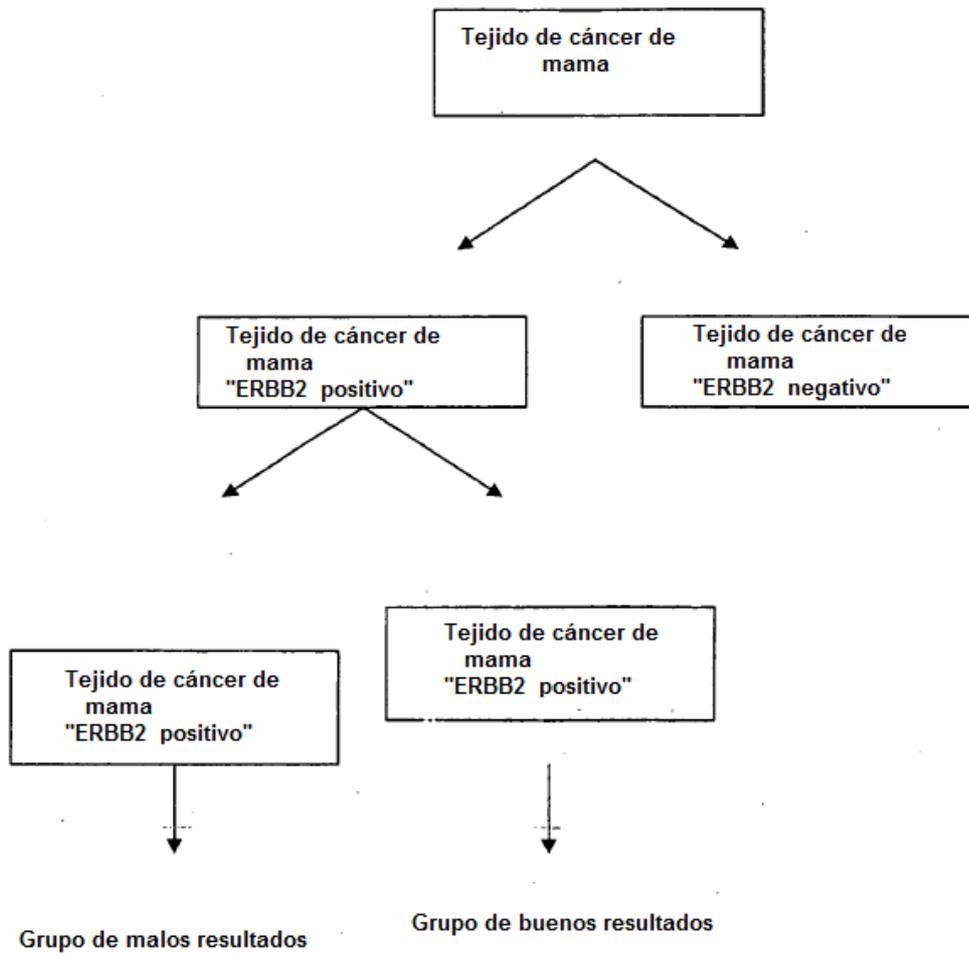


Figura 1

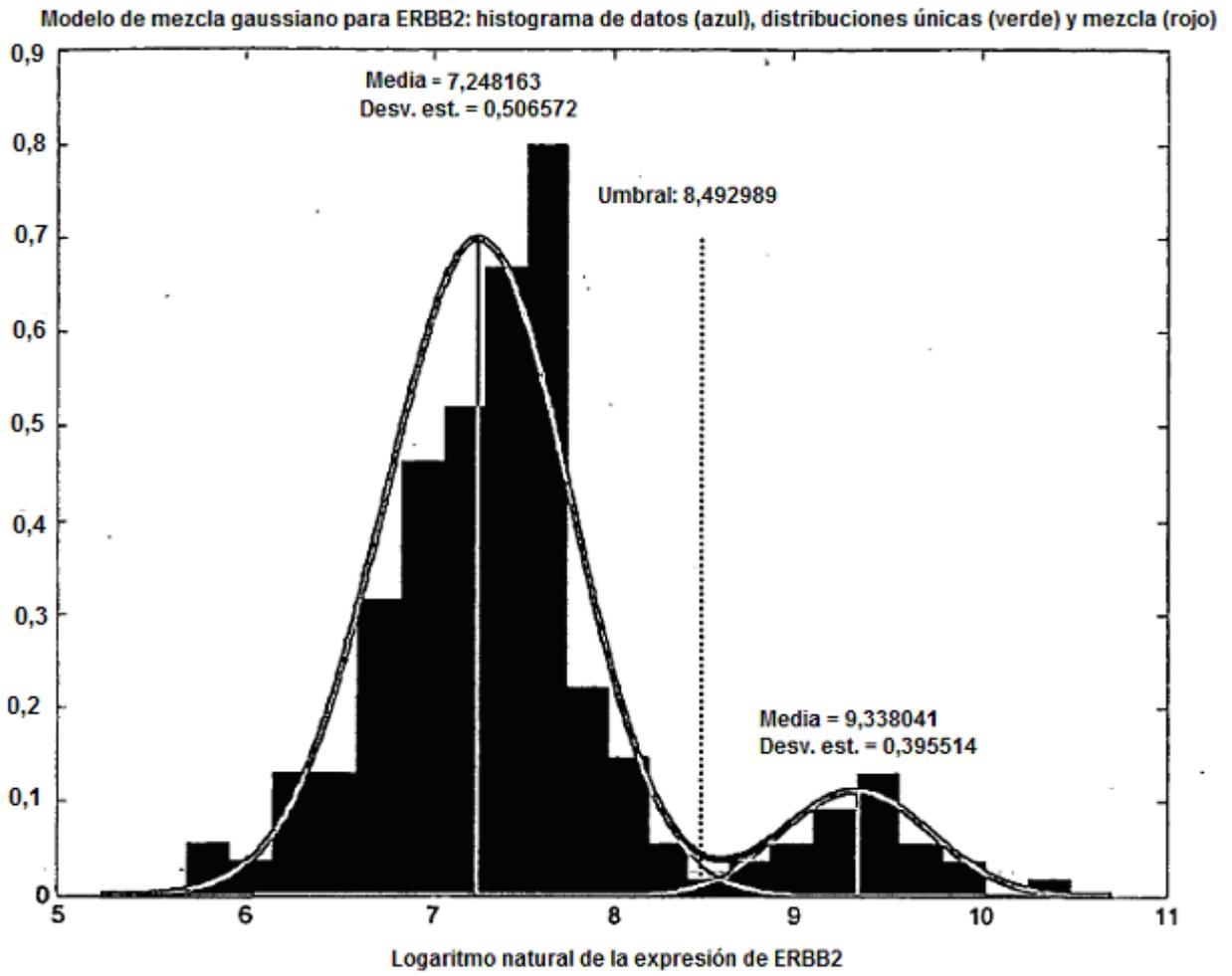


Figura 2

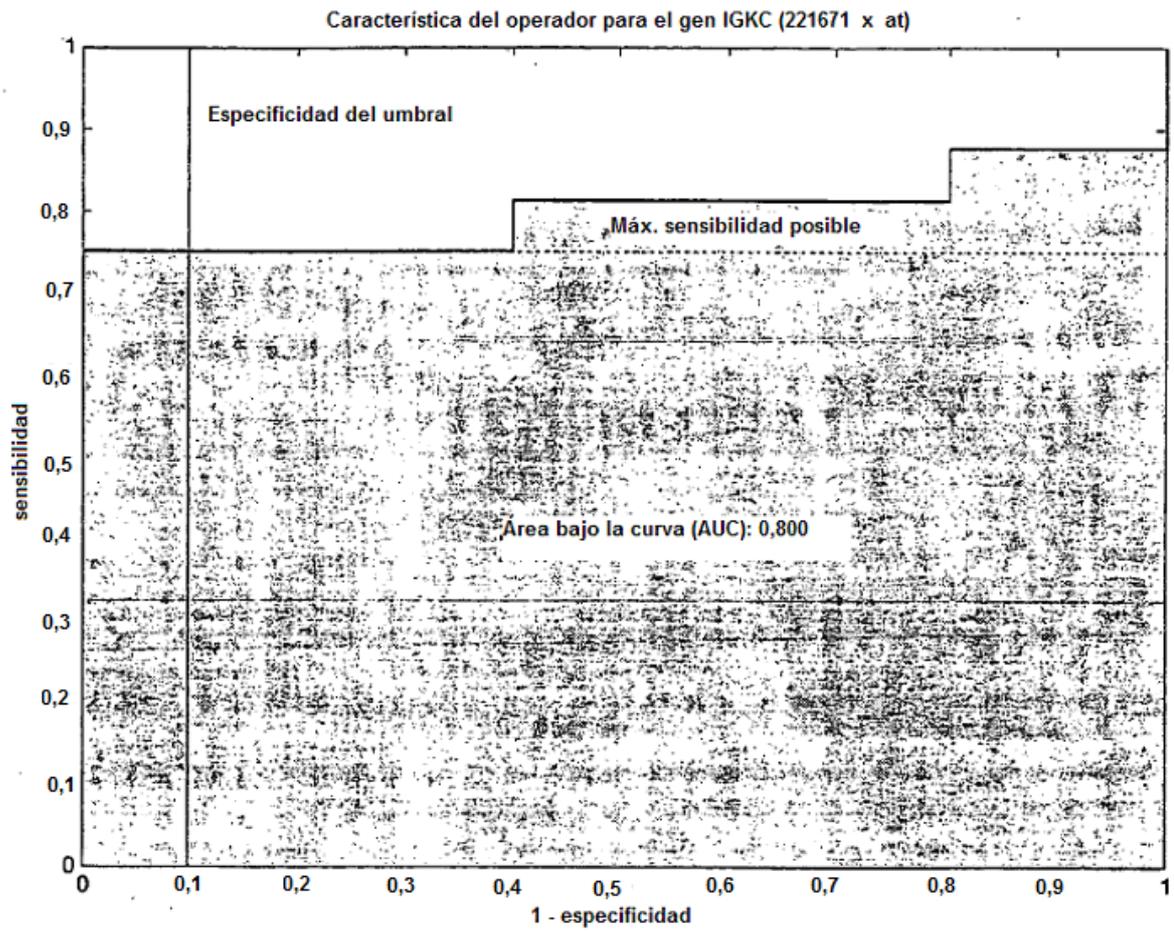


Figura 3

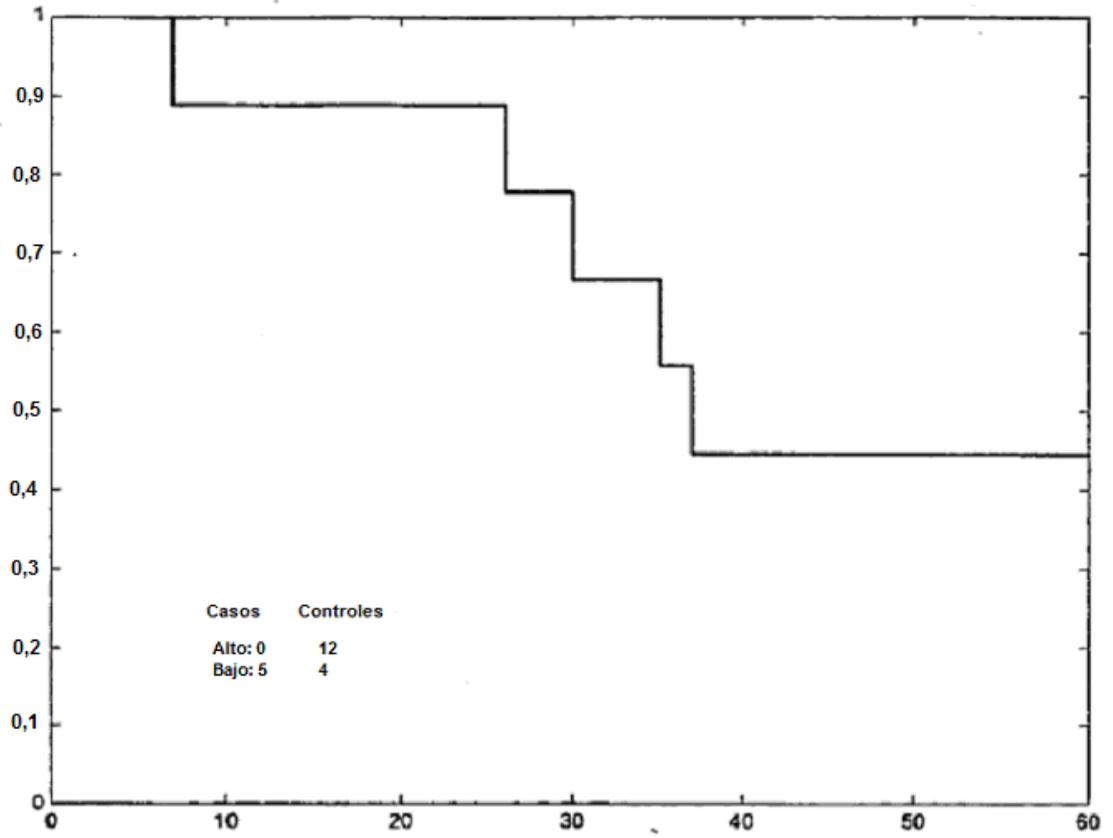


Figura 4

Modelo de mezcla gaussiano para ERBB2: histograma de datos (azul), distribuciones únicas (verde) y mezcla (rojo)

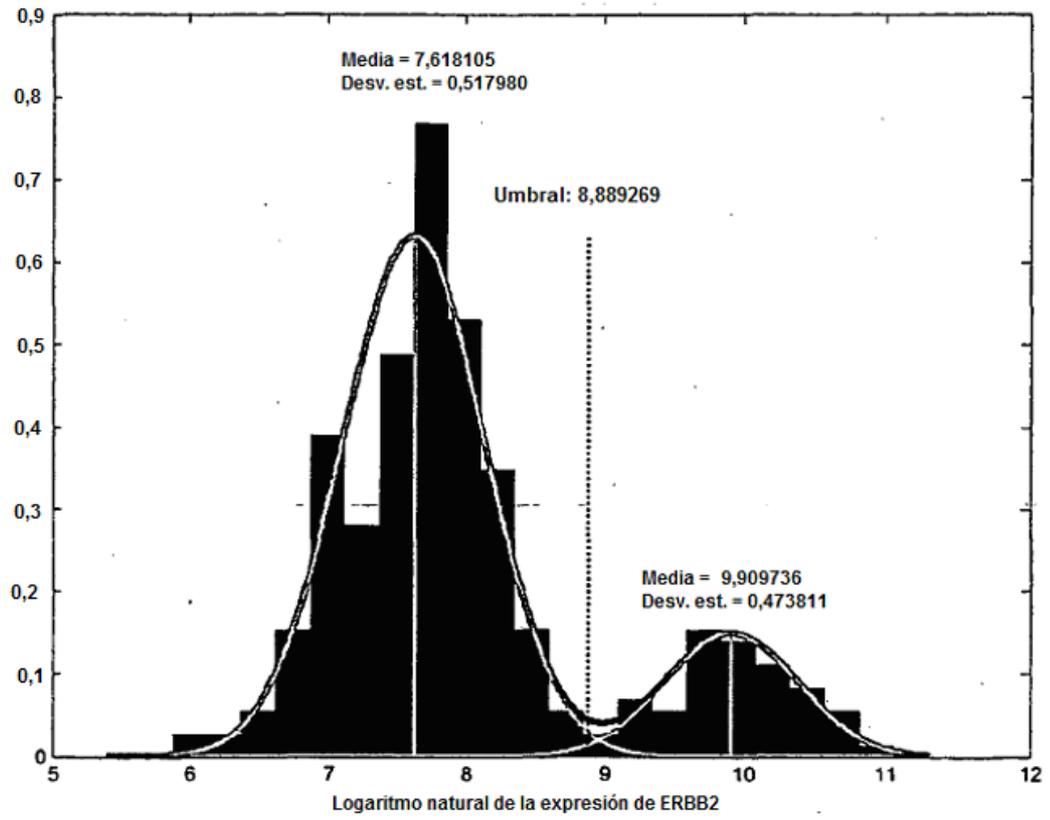


Figura 5

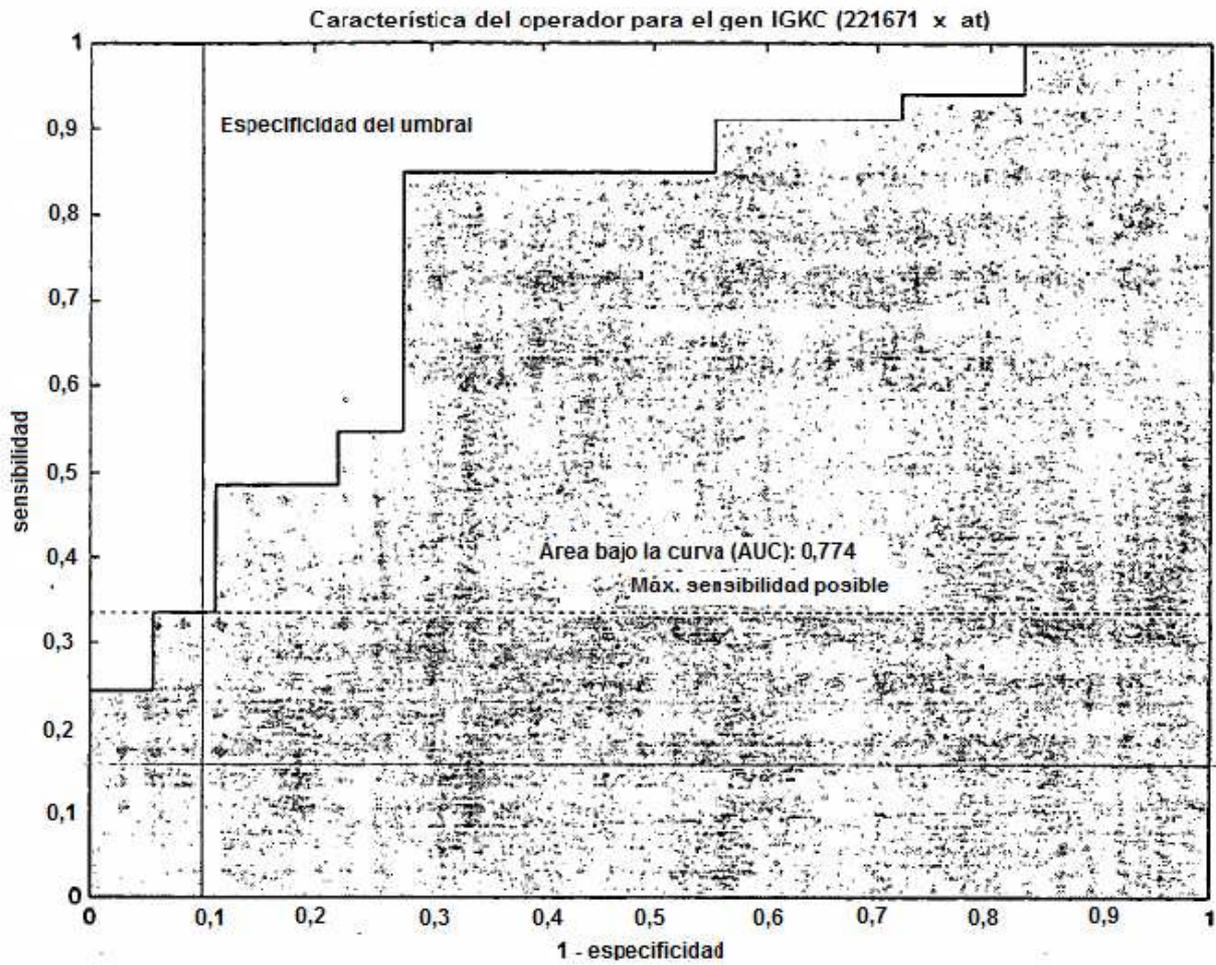


Figura 6

UMBRAL PARA EL GEN IGKC (221671_X_AT): 39781,3

5 BAJO: CASOS: 17 CONTROLES: 22 SINGULARIDADES: 0,77
 ALTO: CASOS: 1 CONTROLES: 11 SINGULARIDADES: 0,091

RELACIÓN DE SINGULARIDADES: 8,5

10 **Figura7**