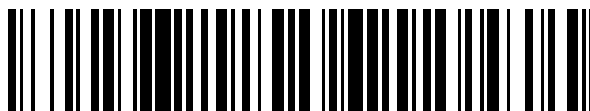


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 447 566**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.10.2003** **E 09009372 (5)**

97 Fecha y número de publicación de la concesión europea: **25.12.2013** **EP 2157191**

54 Título: **Uso de ácidos nucleicos de PITX2 para mejorar el tratamiento de trastornos proliferativos de células mamarias**

30 Prioridad:

01.10.2002 DE 10245779
07.01.2003 DE 10300096
17.04.2003 DE 10317955

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
12.03.2014

73 Titular/es:

EPIGENOMICS AG (100.0%)
INTELLECTUAL PROPERTY KLEINE
PRASIDENTENSTRASSE 1
10178 BERLIN, DE

72 Inventor/es:

FOEKENS, JOHN;
HARBECK, NADIA;
KÖNIG, THOMAS;
MAIER, SABINE;
MARTENS, JOHN;
MODEL, FABIAN;
NIMMRICH, INKO;
RUJAN, TAMAS;
SCHMITT, ARMIN;
SCHMITT, MANFRED;
LOOK, MAXIME, P.;
MARX, ALMUTH y
HÖFLER, HEINZ

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 447 566 T3

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

DESCRIPCIÓN

Uso de ácidos nucleicos de PITX2 para mejorar el tratamiento de trastornos proliferativos de células mamarias

Campo de la invención

Los niveles de observación que se han estudiado mediante los desarrollos metodológicos de los últimos años en biología molecular son los propios genes, la traducción de estos genes en ARN y las proteínas resultantes. La cuestión de qué gen está activado en qué punto en el transcurso del desarrollo de un individuo y cómo se controlan la activación e inhibición de genes específicos en células y tejidos específicos se correlaciona con el grado y el carácter de la metilación de los genes o del genoma. A este respecto, los propios estados patogénicos pueden manifestarse en un patrón de metilación modificado de genes individuales o del genoma.

La metilación del ADN desempeña un papel, por ejemplo, en la regulación de la transcripción, en la huella genética y en la tumorigénesis. Por tanto, la identificación de 5-metilcitosina como componente de la información genética es de considerable interés. Sin embargo, las posiciones de 5-metilcitosina no pueden identificarse mediante secuenciación puesto que la 5-metilcitosina tiene el mismo comportamiento de apareamiento de bases que la citosina. Además, la información epigenética portada por la 5-metilcitosina se pierde completamente durante la amplificación por PCR.

Un procedimiento relativamente nuevo y actualmente el más frecuentemente usado para analizar el ADN para determinar la 5-metilcitosina se basa en la reacción específica de bisulfito con citosina que, tras hidrólisis alcalina posterior, se convierte en uracilo que se corresponde con timidina en su comportamiento de apareamiento de bases. Sin embargo, la 5-metilcitosina permanece sin modificar en estas condiciones. Por consiguiente, el ADN original se convierte de una manera tal que la metilcitosina, que originalmente no podía distinguirse de la citosina por su comportamiento de hibridación, puede detectarse ahora como la única citosina restante usando técnicas de biología molecular "normales", por ejemplo, mediante amplificación e hibridación o secuenciación. Todas estas técnicas se basan en el apareamiento de bases que puede ahora explotarse totalmente. En cuanto a la sensibilidad, la técnica anterior se define mediante un procedimiento que encierra el ADN que va a analizarse en una matriz de agarosa, evitando así la difusión y renaturalización del ADN (el bisulfito sólo reacciona con ADN monocatenario), y que sustituye todas las etapas de precipitación y purificación por diálisis rápida (Olek A, Oswald J, Walter J. A modified and improved method for bisulphite based cytosine methylation analysis. *Nucleic Acids Res.* 15 de diciembre de 1996; 24(24):5064-6). Usando este procedimiento, es posible analizar células individuales, lo que ilustra el potencial del procedimiento. Sin embargo, actualmente sólo se analizan regiones individuales de una longitud de hasta aproximadamente 3000 pares de bases, no es posible un análisis global de células para miles de posibles acontecimientos de metilación. Sin embargo, este procedimiento tampoco puede analizar de manera fiable fragmentos muy pequeños de cantidades de muestra pequeñas. Estas se pierden por la matriz a pesar de la protección de la difusión.

Una visión general de los procedimientos conocidos adicionales de detección de 5-metilcitosina puede recogerse del siguiente artículo de revisión: Rein, T., DePamphilis, M. L., Zorbas, H., *Nucleic Acids Res.* 1998, 26, 2255.

Hasta la fecha, excluyendo pocas excepciones (por ejemplo, Zeschnigk M, Lich C, Buiting K, Doerfler W, Horsthemke B. A single-tube PCR test for the diagnosis of Angelman and Prader-Willi syndrome based on allelic methylation differences at the SNRPN locus. *Eur J Hum Genet.* marzo – abril de 1997; 5(2):94-8), la técnica de bisulfito se usa sólo en investigación. Siempre, sin embargo, se amplifican fragmentos cortos específicos de un gen conocido después de un tratamiento con bisulfito y o bien se secuencian completamente (Olek A, Walter J. The pre-implantation ontogeny of the H19 methylation imprint. *Nat Genet.* Noviembre de 1997; 17 (3):275-6) o bien se detectan las posiciones de citosina individuales mediante una reacción de extensión de cebador (Gonzalzo ML, Jones PA. Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (Ms-SNuPE). *Nucleic Acids Res.* 15 de junio de 1997; 25(12):2529-31, documento WO 95/00669) o mediante digestión enzimática (Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* 15 de junio de 1997; 25(12):2532-4). Además, también se ha descrito la detección mediante hibridación (Olek *et al.*, documento WO 99/28498).

Otras publicaciones que tratan el uso de la técnica de bisulfito para la detección de la metilación en genes individuales son: Grigg G, Clark S. Sequencing 5-methylcytosine residues in genomic DNA. *Bioessays.* Junio de 1994; 16(6):431-6, 431; Zeschnigk M, Schmitz B, Ditttrich B, Buiting K, Horsthemke B, Doerfler W. Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. *Hum Mol Genet.* Marzo de 1997; 6(3):387-95; Feil R, Charlton J, Bird AP, Walter J, Reik W. Methylation analysis on individual chromosomes: improved protocol for bisulphite genomic sequencing. *Nucleic Acids Res.* 25 de febrero de 1994; 22(4):695-6; Martin V, Ribieras S, Song-Wang X, Rio MC, Dante R. Genomic sequencing indicates a correlation between DNA hypomethylation in the 5' region of the pS2 gene and its expression in human breast cancer cell lines. *Gene.* 19 de mayo de 1995; 157(1-2):261-4; documentos WO 97/46705, WO 95/15373 y WO 97/45560.

Una visión general de la técnica anterior en la fabricación de series de oligómeros puede recogerse de una edición especial de *Nature Genetics* (*Nature Genetics Supplement*, volumen 21, enero de 1999), publicado en enero de

1999, y de la bibliografía citada en ese documento.

Con frecuencia se usan sondas marcadas con fluorescencia para la exploración de series de ADN inmovilizado. La simple unión de colorantes Cy3 y Cy5 al OH en 5' de la sonda específica es particularmente adecuada para marcadores de fluorescencia. La detección de la fluorescencia de las sondas hibridadas puede llevarse a cabo, por ejemplo, por medio de un microscopio confocal. Los colorantes Cy3 y Cy5, además de muchos otros, están comercialmente disponibles.

La espectrometría de masas de desorción – ionización por láser asistida por matriz (MALDI-TOF) es un desarrollo muy eficaz para el análisis de biomoléculas (Karas M, Hillenkamp F. Laser desorption ionisation of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 15 de octubre de 1988; 60(20):2299-301). Se incrusta un analito en una matriz de absorción de luz. La matriz se evapora mediante un pulso de láser corto transportando así la molécula de analito hacia la fase de vapor de una manera sin fragmentar. Se ioniza el analito mediante colisiones con moléculas de la matriz. Un voltaje aplicado acelera los iones en un tubo de vuelo sin campo. Debido a sus diferentes masas, los iones se aceleran a diferentes velocidades. Iones más pequeños alcanzan el detector antes que iones más grandes.

La espectrometría de MALDI-TOF se adecua excelentemente al análisis de péptidos y proteínas. El análisis de ácidos nucleicos es un tanto más difícil (Gut I G, Beck S. DNA and Matrix Assisted Laser Desorption Ionization Mass Spectrometry. Current Innovations and Future Trends. 1995, 1; 147-57). La sensibilidad a ácidos nucleicos es aproximadamente 100 veces peor que a péptidos y disminuye de manera no proporcional con el aumento del tamaño de los fragmentos. Para ácidos nucleicos que tienen una estructura principal con múltiples cargas negativas, el proceso de ionización por medio de la matriz es considerablemente menos eficaz. En espectrometría de MALDI-TOF, la selección de la matriz desempeña un papel sumamente importante. Para la desorción de péptidos, se han encontrados varias matrices muy eficaces que producen una cristalización muy fina. Ahora hay varias matrices sensibles para ADN, sin embargo, no se ha reducido la diferencia en sensibilidad. La diferencia en sensibilidad puede reducirse modificando químicamente el ADN de tal manera que se vuelva más similar a un péptido. Los ácidos nucleicos de fosforotioato, en los que los fosfatos habituales de la estructura principal se sustituyen por tiofosfatos, pueden convertirse en un ADN de carga neutra usando química de alquilación simple (Gut IG, Beck S. A procedure for selective DNA alkylation and detection by mass spectrometry. Nucleic Acids Res. 25 de abril de 1995; 23(8):1367-73). El acoplamiento de una etiqueta de carga a este ADN modificado da como resultado un aumento en la sensibilidad al mismo nivel que se encontró para péptidos. Una ventaja adicional del etiquetado de carga es el aumento de estabilidad del análisis frente a impurezas que hacen considerablemente más difícil la detección de sustratos sin modificar.

Se obtiene ADN genómico de ADN de muestras de prueba de células, tejido u otras usando procedimientos convencionales. Esta metodología convencional se encuentra en referencias tales como Fritsch and Maniatis eds., Molecular Cloning: A Laboratory Manual, 1989.

El cáncer de mama es actualmente el segundo tipo más común de cáncer entre las mujeres. En 2001 se diagnosticaron más de 190.000 nuevos casos de cáncer de mama invasivo y más de 47.000 casos adicionales de cáncer de mama *in situ*. Las tasas de incidencia y muerte aumentan con la edad, durante el periodo de 1994 – 1998, la incidencia de cáncer de mama entre mujeres entre las edades de 20 y 24 años era sólo de 1,5 por 100.000 individuos. Sin embargo, el riesgo aumenta hasta 489,7 con el grupo de edad de 75 - 79. Las tasas de mortalidad han disminuido en aproximadamente el 5% a lo largo de la última década y los factores que afectan a las tasas de supervivencia a 5 años incluyen la edad, el estadio del cáncer, los factores socioeconómicos y la raza.

El cáncer de mama se define como la proliferación incontrolada de células en los tejidos de mama. Las mamas están compuestas por de 15 a 20 lóbulos unidos entre sí mediante conductos. El cáncer surge lo más comúnmente en el conducto, pero también se encuentra en los lóbulos, denominándose el tipo más raro de cáncer cáncer de mama inflamatorio.

Los expertos en la técnica apreciarán que existe una necesidad continua de mejorar los procedimientos de detección temprana, clasificación y tratamiento de cánceres de mama. A diferencia de la detección de algunos cánceres comunes distintos tales como cáncer de cuello uterino y dérmico, existen dificultades inherentes en la clasificación y detección de cánceres de mama.

La primera etapa de cualquier tratamiento es la evaluación del estado del paciente en comparación con clasificaciones definidas de la enfermedad. Sin embargo, el valor de un sistema de este tipo depende de manera inherente de la calidad de la clasificación. Los cánceres de mama se clasifican según su tamaño, ubicación y aparición de metástasis. Los procedimientos de tratamiento incluyen el uso de cirugía, radioterapia, quimioterapia y terapia endocrina, que se usan también como tratamientos complementarios a la cirugía.

Se han desarrollado terapias endocrinas con el fin de bloquear los efectos del estrógeno sobre las células cancerosas o de reducir los niveles de estrógeno en suero. El tamoxifeno (TAM) es el fármaco antiestrogénico más ampliamente usado para pacientes con cáncer de mama. Actúa bloqueando la estimulación de estrógenos de las células de cáncer de mama, inhibiendo tanto el traslado como la unión nuclear del receptor de estrógeno. TAM se

usa en entornos complementarios para pacientes con cáncer de mama primario y para el tratamiento de enfermedad metastásica, y su eficacia se ha demostrado en varios ensayos clínicos. El tratamiento durante cinco años reduce la recaída de enfermedad anual en un 47% y las muertes anuales en un 26%, y esta reducción es similar en grupos de edades diferentes. Reduce también la incidencia de desarrollo de tumores de mama contralaterales. TAM está entre los agentes antineoplásicos menos tóxicos pero, ya que tiene propiedades estrogénicas en algunos tejidos, aumenta en 3 veces el riesgo de cáncer de endometrio. Otras terapias endocrinas incluyen inhibidores de aromatasa, que bloquean la enzima aromatasa en el tejido adiposo (la fuente principal de estrógenos en mujeres postmenopáusicas) y conducen a reducir o suprimir la producción de estrógenos en el tejido adiposo. También se han desarrollado nuevos agentes que bloquean o modulan el receptor de estrógeno, se resumen generalmente como SERM = *selective estrogen receptor modulators* (moduladores de receptores de estrógeno selectivos). Aún otro modo para reducir la cantidad de estrógeno disponible para el tumor es aplicar agentes que interrumpen el bucle de retroalimentación en la regulación de hormonas sexuales (análogos de la LH-RH) a un nivel más alto.

En general, los efectos secundarios del tratamiento endocrino son inofensivos en comparación con la quimioterapia. Sin embargo, ya que las terapias endocrinas eliminan o reducen el estrógeno como factor de crecimiento para el cáncer, funcionan sólo para tumores que se basan en estrógeno como factor de crecimiento. El crecimiento de cánceres que no tienen una ruta de ER funcional no puede modularse mediante terapias endocrinas.

El modo actual para determinar si un tumor responderá a terapias endocrinas es el análisis de la expresión de receptores hormonales. El estado de receptores hormonales debe someterse a prueba antes de cualquier tratamiento endocrino ya que sólo un grupo de pacientes se beneficiará del tratamiento. Para este fin se somete a prueba el estado de ER (receptor de estrógeno) y PR (receptor de progesterona) de manera rutinaria en todos los pacientes y se consideran marcadores de predicción de la respuesta (un factor o marcador de predicción es cualquier medición asociada con la respuesta o falta de respuesta a un tratamiento particular). El estado de ER es predictivo de la respuesta en entornos de terapia endocrina complementaria y también en enfermedad metastásica avanzada.

Actualmente, los pacientes positivos para ER y/o PR reciben TAM y el 10% restante de los pacientes reciben quimioterapia. El problema es que el tratamiento endocrino es sólo eficaz en un subgrupo de pacientes positivos para receptores hormonales y también, que un pequeño subgrupo de pacientes negativos para ER parece responder inicialmente a TAM. Entonces, el subgrupo de pacientes que no responden no sólo no se beneficiará del tratamiento endocrino sino que tienen los efectos adversos del mismo y no tendrán la oportunidad de recibir en su lugar quimioterapia. Por otro lado, el subgrupo negativo para ER que podría beneficiarse del tratamiento endocrino recibirá en su lugar quimioterapia. Se encuentran disponibles varios ensayos diferentes para medir PR y ER, incluyendo análisis bioquímicos e IHC (inmunohistoquímicos), pero existe una falta de normalización de los procedimientos de tinción y variabilidad entre laboratorios (Clin Cancer Res 2000, 6:616-621).

Actualmente se están evaluando varios marcadores de predicción. Ya que hasta ahora la mayoría de los pacientes han recibido tamoxifeno como tratamiento endocrino, la mayoría de los marcadores han mostrado estar asociados con una respuesta o resistencia a tamoxifeno. Sin embargo, generalmente se supone que existe un gran solapamiento entre pacientes que responden a uno o el otro tratamiento endocrino. De hecho, se usa la expresión de ER y PR para seleccionar pacientes para cualquier tratamiento endocrino. Entre los marcadores que se han asociado con la respuesta a TAM está bcl-2. Altos niveles de bcl-2 mostraron una correlación prometedora con la respuesta al tratamiento con TAM en pacientes con enfermedad metastásica y la supervivencia prolongada e información valiosa añadida para el subgrupo de pacientes negativos para ER (J Clin Oncology 1997, 15 5:1916-1922; Endocrine, 2000, 13(1):1-10). Existen pruebas contradictorias referentes al valor de predicción independiente de c-erbB2 (Her2/neu) con respecto a la expresión en pacientes con cáncer de mama avanzado que requieren evaluación y verificación adicionales (British J of Cancer 1999, 79 (7/8):1220-1226; J Natl Cancer Inst, 1998, 90 (21): 1601-1608).

Otros marcadores de predicción incluyen SRC-1 (coactivador 1 del receptor de esteroides), sobreexpresión del gen CGA, cinética celular y ensayos de fracción de estadio S (Breast Cancer Res and Treat, 1998, 48:87-92; Oncogene 2001, 20:6955-6959). Recientemente, uPA (activador del plasminógeno de tipo uroquinasa) y PAI-1 (inhibidor de activador del plasminógeno tipo 1) juntos mostraron ser útiles para definir un subgrupo de pacientes que tienen peor pronóstico y que se beneficiarían del tratamiento sistémico complementario (J Clinical Oncology, 2002, 20 n.º 4). Todos estos marcadores necesitan evaluaciones adicionales en ensayos prospectivos ya que ninguno de ellos es todavía un marcador validado de la respuesta.

YAN P S et al. (Dissecting complex epigenetic alterations in breast cancer using CpG Island microarrays", CANCER RESEARCH, vol. 61, nº 23, 1 December 2001, páginas 8375-8380) divulgan una micromatriz de marcadores cortos ricos en GC para estudiar 11 tejidos pareados de tumores de mama y de controles normales. Se detectó hipermetilación de múltiples loci de islotes de CpG. Agrupaciones jerárquicas segregaron estos tumores en base a sus perfiles de metilación e identificaron un grupo de loci de islotes de CpG que corresponde al estado del receptor hormonal del cáncer de mama.

Se ha demostrado que varios genes asociados al cáncer se inactivan mediante hipermetilación de islas de CpG durante la oncogenia de mama. La disminución de la expresión de la proteína de unión a calcio S100A2 (número de

acceso NM_005978) se ha asociado con el desarrollo de cánceres de mama. La hipermetilación de la región promotora de este gen se ha observado en células neoplásicas proporcionando así pruebas de que la represión de S100A2 en células tumorales está mediada por metilación específica de sitio.

El gen SYK (número de acceso NM_003177) codifica una proteína tirosina quinasa, Syk (tirosina quinasa de bazo), que se expresa altamente en células hematopoyéticas. Syk se expresa en células epiteliales ductales de mama normales pero no en un subconjunto de carcinoma de mama invasivo. Además, la pérdida de expresión de Syk parece estar asociada con fenotipos malignos tales como aumento en la motilidad e invasión. La pérdida de expresión se produce a nivel de la transcripción, y, tal como se indica por Yuan Y, Mendez R, Sahin A y Dai JL (Hypermethylation leads to silencing of the SYK gene in human breast cancer. *Cancer Res.* 15 de julio de 2001; 61(14):5558-61.), como resultado de la hipermetilación del ADN.

El receptor del TGF- β tipo 2 (codificado por el gen TGFBR2, NM_003242) desempeña un papel en rutas de señalización transmembrana por medio de un complejo de serina/treonina quinasas. Se han detectado mutaciones en el gen en algunos tumores primarios y en varios tipos de líneas celulares derivadas de tumor, incluyendo de mama (Lucke CD, Philpott A, Metcalfe JC, Thompson AM, Hughes-Davies L, Kemp PR, Hesketh R. "Inhibiting mutations in the transforming growth factor beta type 2 receptor in recurrent human breast cancer". *Cancer Res.* 15 de enero de 2001; 61(2):482-5.).

Los genes COX7A2L y GRIN2D se identificaron como elementos sensibles a estrógeno por Watanabe *et. al.* (Isolation of estrogen-responsive genes with a CpG island library. *Molec. Cell. Biol.* 18: 442-449, 1998.) usando el procedimiento de CpG-GBS (sitio de unión genómico). El gen COX7A2L (número de acceso NM_004718 codifica una subunidad VIIA de la citocromo C oxidasa similar al polipéptido 2. El análisis de transferencia tipo Northern detectó una regulación por incremento de COX7A2L tras el tratamiento con estrógeno de una línea celular de cáncer de mama. El gen GRIN2D (número de acceso NM_000836) codifica el N-metil-D-aspartato, receptor de D-glutamato de subunidad 2, ionotrópico, una subunidad de los canales de receptores de NMDA asociados con la señalización neuronal, además se ha observado expresión del ADNc en una línea celular de osteosarcoma. El gen VTN (conocido también como vitronectina, número de acceso NM_000638) codifica una glicoproteína de 75 kD (denominada también factor de propagación en suero o proteína S del complemento) que promueve la unión y la propagación de células animales *in vitro*, inhibe la citólisis mediante el complejo C5b-9 del complemento y modula la acción de la antitrombina III – trombina en la coagulación sanguínea. Además, la expresión de este gen se ha relacionado con la progresión y la capacidad de invasión de las células cancerosas.

El gen SFN (conocido también como estratífina) codifica un polipéptido de la familia 14-3-3, 14-3-3 sigma. La familia de proteínas 14-3-3 media la transducción de señales mediante la unión a proteínas que contienen fosfoserina. La expresión del gen SFN se pierde en carcinomas de mama, esto se debe probablemente a la hipermetilación durante las fases tempranas de la transformación neoplásica (véase Umbricht CB, Evron E, Gabrielson E, Ferguson A, Marks J, Sukumar S. Hypermethylation of 14-3-3 sigma (stratifin) is an early event in breast cancer. *Oncogene.* 7 de junio de 2001; 20(26):3348-53).

El gen PSA (número de acceso NM_021154), conocido también como PSA-T, no ha de confundirse con el gen denominado también popularmente PSA (número de acceso NM_001648) que codifica antígeno específico de próstata y cuyo nombre técnicamente correcto es calicreína 3. El gen PSA-T codifica la proteína fosfoserina aminotransferasa que es la enzima catalizadora de segunda etapa en la ruta de la biosíntesis de serina. Los cambios en los niveles de expresión génica se han monitorizado mediante análisis de expresión de ARNm y se ha identificado la regulación por incremento del gen en carcinoma de colon en un estudio de 6 muestras (Electrophoresis, junio de 2002; 23(11):1667-76 mRNA differential display of gene expression in colonic carcinoma. Ojala P, Sundstrom J, Gronroos JM, Virtanen E, Talvinen K, Nevalainen TJ.).

El gen estatmina (NM_005563) codifica una oncoproteína 18, conocida también como estatmina, una fosfoproteína citosólica conservada que regula la dinámica de microtúbulos. La proteína se expresa altamente en una variedad de tumores malignos humanos. En cánceres de mama humanos se ha demostrado que el gen estatmina se regula por incremento en un subconjunto de los tumores.

El gen PRKCD codifica un miembro de la familia de enzimas proteína quinasas c, y está implicado en la señalización de células B y en la regulación del crecimiento, la apoptosis y la diferenciación de una variedad de tipos de células.

Algunas de estas moléculas interaccionan de manera similar a una cascada. La actividad de PRKCD que se dirige a STMN1 se modula mediante la unión a SFN y la fosforilación de SYK. Junto a esto influye la polimerización de tubulina que se requiere para la división celular.

El gen MSMB (número de acceso NM_002443) se ha mapeado en 10q11.2. Codifica la beta-micro-seminoproteína (MSP) que es una de las principales proteínas secretadas por la próstata. Además, puede ser útil como marcador de diagnóstico para el cáncer de próstata. Usando análisis de ARNm, se han relacionado bajos niveles de expresión de ARNm de beta-MSP y proteína con la progresión bajo terapia endocrina, y se ha postulado que puede ser indicativo de cáncer de próstata potencialmente agresivo (véase Sakai H, Tsurusaki T, Kanda S, Koji T, Xuan JW, Saito Y. "Prognostic significance of beta-microseminoprotein mRNA expression in prostate cancer". *Prostate.* 1 de marzo de

1999; 38(4):278-84.).

El gen TP53 (número de acceso NM_000546) codifica la proteína p53, una de las proteínas supresoras de tumores mejor caracterizadas. La proteína p53 actúa como factor de transcripción factor y sirve como regulador clave del ciclo celular. La inactivación de este gen a través de mutación perturba el ciclo celular, que a su vez, ayuda en la formación de tumor. Se ha notificado que los cambios en la metilación asociados con este gen son significativos en el cáncer de mama. Saraswati *et. al.* (Nature 405, 974 - 978 (22 de junio de 2000) "Compromised HOXA5 function can limit p53 expression in human breast tumours" notificaron que bajos niveles de ARNm de p53 en tumores de mama estaba correlacionado con la metilación del gen HOXA5. El producto del gen HOXA5 se une a la región promotora del p53 y media la expresión del gen. Se ha notificado la propia metilación de la región promotora del gen p53 (Kang JH, Kim SJ, Noh DY, Park IA, Choe KJ, Yoo OJ, Kang HS. "Methylation in the p53 promoter is a supplementary route to breast carcinogenesis: correlation between CpG methylation in the p53 promoter and the mutation of the p53 gene in the progression from carcinoma ductal *in situ* to invasive ductal carcinoma". Lab Invest. Abril de 2001; 81(4):573-9.). En ese documento se demostró que la metilación de CpG en la región promotora de p53 se encuentra en el cáncer de mama y se propuso la hipótesis de que la metilación en la región promotora de p53 podría ser una ruta alternativa a la progresión neoplásica en tumores de mama. Se ha observado que el tratamiento con tamoxifeno reduce el nivel de expresión del gen p53 (Farczadi E, Kaszas I, Baki M, Szende B. "Changes in apoptosis, mitosis, Her-2, p53 y Bcl2 expression in breast carcinomas after short-term tamoxifen treatment". Neoplasma. 2002; 49(2):101-3.)

El gen CYP2D6 (número de acceso: NM_000106) es un miembro de la superfamilia citocromo P450 (CYP) humano. Muchos miembros de esta familia están implicados en el metabolismo de fármacos (véase por ejemplo Curr Drug Metab. Junio de 2002; 3(3):289-309. Rodrigues AD, Rushmore TH.), de éstos CYP2D6 de citocromo P450 es uno de los más exhaustivamente caracterizados. Es altamente polimórfico (se han descrito más de 70 variaciones del gen), y la variación alélica puede dar como resultado tanto un aumento como una disminución de la actividad enzimática. La enzima CYP2D6 cataliza el metabolismo de un gran número de fármacos clínicamente importantes incluyendo antidepresivos, neurolépticos, algunos antiaritmicos (Nature, 25 de octubre de 1990; 347(6295):773-6 Identification of the primary gene defect at the cytochrome P450 CYP2D locus. Gough AC, Miles JS, Spurr NK, Moss JE, Gaedigk A, Eichelbaum M, Wolf CR.).

El gen PTGS2 (número de acceso NM_000963) codifica una isozima inducible de prostaglandina-endoperoxido sintasa (prostaglandina-endoperoxido sintasa 2). Se ha identificado la metilación anómala de este gen en carcinomas de pulmón (Cancer Epidemiol Biomarkers Prev, marzo de 2002; 11(3):291-7 Hierarchical clustering of lung cancer cell lines using DNA methylation markers. Virmani AK, Tsou JA, Siegmund KD, Shen LY, Long TI, Laird PW, Gazdar AF, Laird-Offringa IA.).

El gen CGA (número de acceso NM_000735) codifica el polipéptido alfa de hormonas de glicoproteína. Además, se ha identificado como un gen sensible al receptor alfa de estrógeno (ER alfa), y la sobreexpresión del gen se ha relacionado con la positividad para ER en tumores de mama. Bieche *et. al.* examinaron los niveles de ARNm de dicho gen en 125 pacientes con cáncer de mama postmenopáusicas positivas para ER alfa tratadas con cirugía primaria seguida de tratamiento complementario con tamoxifeno. Los resultados iniciales indicaron relaciones significativas entre la sobreexpresión del gen CGA y el grado histopatológico I+II de Scarff-Bloom-Richardson y la positividad para receptores de estrógeno y progesterona, lo que sugirió que CGA es un marcador de baja agresividad tumoral ("Identification of CGA as a Novel Estrogen Receptor-responsive Gene in Breast Cancer: An Outstanding Candidate Marker to predict the Response to Endocrine Therapy Cancer Research" 61, 1652-1658, 15 de febrero de 2001. Ivan Bièche, Béatrice Parfait, Vivianne Le Doussal, Martine Olivi, Marie-Christine Rio, Rosette Lidereau y Michel Vidaud). El análisis de la expresión de ARNm adicional relacionó los niveles de expresión de CGA con la respuesta a tamoxifeno, se postuló que cuando se combina con el análisis del marcador ERBB2 (un marcador de escasa respuesta) el gen puede ser útil como marcador de predicción de la capacidad de respuesta a tamoxifeno en el cáncer de mama (Oncogene, 18 de octubre de 2001; 20(47):6955-9. The CGA gene as new predictor of the response to endocrine therapy in ER alpha-positive postmenopausal breast cancer patients. Bièche I, Parfait B, Nogues C, Andrieu C, Vidaud D, Spyatos F, Lidereau R, Vidaud M.). Los autores proporcionaron datos significativos que asociaban la expresión del gen CGA con la respuesta al tratamiento con tamoxifeno. Sin embargo, dichos análisis se han centrado todos en el análisis de los niveles relativos de expresión de ARNm. Ésta no es una metodología que sea adecuada para un rendimiento medio o alto, ni es una base adecuada para el desarrollo de un ensayo clínico.

El gen PITX-2 (NM_000325) codifica un factor de transcripción (PITX-2) que se sabe que se expresa durante el desarrollo de estructuras anteriores tales como el ojo, los dientes y la adenohipófisis. Aunque la expresión de este gen está asociada con la diferenciación y proliferación celular, hasta ahora no tiene un papel reconocido en la carcinogénesis o capacidad de respuesta al tratamiento endocrino. Además, hasta la fecha no se ha notificado ningún análisis conocido del estado de metilación de este gen.

El gen RASSF1A (familia 1A del dominio de asociación de Ras) es un gen supresor de tumores candidato en 3p21.3. Las GTPasas Ras son una superfamilia de interruptores moleculares que regulan la proliferación y apoptosis celulares en respuesta a señales extracelulares. Se afirma que RASSF1A es un gen supresor de tumores, y se han observado alteraciones epigenéticas de este gen en una variedad de cánceres. La metilación de RASSF1A se ha

- asociado con el mal pronóstico en el cáncer de pulmón de células no pequeñas primario (Kim DH, Kim JS, Ji YI, Shim YM, Kim H, Han J, Park J., "Hypermethylation of RASSF1A promoter is associated with the age at starting smoking and a poor prognosis in primary non-small cell lung cancer". *Cancer Res.* 1 de julio de 2003; 63(13):3743-6.). Se ha asociado también con el desarrollo de cáncer pancreático (Kuzmin I, Liu L, Dammann R, Geil L, Stanbridge EJ, Wilczynski SP, Lerman MI, Pfeifer GP. "Inactivation of RAS association domain family 1A gene in cervical carcinomas and the role of human papillomavirus infection". *Cancer Res.* 15 de abril de 2003; 63(8):1888-93.), así como tumores testiculares y carcinoma de próstata entre otros. La aplicación de la metilación de este gen como marcador de diagnóstico del cáncer se ha descrito en la patente estadounidense 6.596.488, sin embargo no describe su aplicación en la selección de regímenes de tratamiento apropiados para los pacientes.
- 5 Dentro de 3p21 también está ubicado el gen precursor de distroglicano (glicoproteína 1 asociada a distrofina) (NM_004393). El distroglicano (DG, conocido también como DAG1) es una molécula de adhesión que comprende dos subunidades, concretamente alfa-DG y beta-DG. La molécula es responsable de interacciones cruciales entre la matriz extracelular y el compartimento citoplasmático, y se ha propuesto la hipótesis de que como tal puede contribuir a la progresión hasta enfermedad metastásica. La expresión disminuida de este gen se ha asociado con
- 10 un estadio y grado tumoral mayor correlacionado en tumores de colon, próstata y mama.
- 15 El gen del factor de transcripción oncut-2 (NM_004852) está ubicado en 18q21.31 es un regulador de factor de transcripción tipo homeodominio de la expresión génica del hígado en adultos y durante el desarrollo.
- 20 El gen del factor trébol (TFF) 2 (NM_003225) es un miembro de la familia trébol de proteínas. Se expresan normalmente a los niveles más altos en la mucosa del tubo digestivo, sin embargo, con frecuencia se expresan de manera ectópica en tumores primarios de otros tejidos, incluyendo la mama. La expresión de TFF 1 se regula por estrógeno en células de cáncer de mama sensibles a estrógeno en cultivo, su expresión está asociada con la del receptor de estrógeno y TFF 1 es un marcador de la capacidad de respuesta a hormonas en tumores. La metilación del promotor de TFF1 se ha observado en tejidos y líneas celulares derivadas de carcinoma gástrico de no expresión.
- 25 TMEFF2 (NM_016192) codifica una proteína transmembrana que contiene un motivo similar al factor de crecimiento epidérmico (EGF) y dos dominios de folistatina. Se ha demostrado que se sobreexpresa en tejidos de próstata y cerebro y se ha sugerido que es un gen regulado por andrógenos que muestra efectos antiproliferativos en células de cáncer de próstata.
- 30 La metilación del gen ESR1 (NM_000125), que codifica el receptor de estrógeno, se ha relacionado con varios tipos de cáncer incluyendo de pulmón, de esófago, de cerebro y colorrectal. El receptor de estrógeno (ESR) es un factor de transcripción activado por ligando compuesto por varios dominios importantes para la unión a hormonas, unión a ADN y activación de la transcripción.
- 35 El gen PCAF (NM_003884) gene codifica el factor asociado a p300/CBP (PCAF). CBP y p300 son proteínas nucleares grandes que se unen a muchos factores específicos de secuencia implicados en el crecimiento y/o la diferenciación celulares. El factor asociado a p300/CBP presenta actividad de unión *in vivo* con CBP y p300. La proteína tiene actividad histona acetil transferasa con histonas núcleo y partículas núcleo del nucleosoma, lo que indica que desempeña un papel directo en la regulación de la transcripción. El factor asociado a p300/CBP también se asocia a p65 de NF-kappa-B. Se ha demostrado que esta proteína regula la expresión del gen p53 mediante la acetilación de Lys320 en la parte C-terminal de p53.
- 40 El gen WBP11 (NM_016312) codifica una proteína nuclear, que se ubica conjuntamente con factores de corte y empalme de ARNm y redes perinucleares que contienen filamentos intermedios. Contiene dos regiones ricas en prolina que se unen al dominio WW de Npw38, una proteína nuclear y, por tanto, esta proteína se denomina también proteína de unión a Npw38, NpwBP.

Descripción

- 45 La presente divulgación proporciona procedimientos para predecir la respuesta de un paciente con un trastorno proliferativo de células de los tejidos de mama a tratamiento endocrino. Con el uso de los procedimientos y ácidos nucleicos descritos en el presente documento, pueden desarrollarse modelos estadísticamente significativos de la capacidad de respuesta del paciente a dicho régimen de tratamiento y utilizarse para ayudar a pacientes y médicos a determinar las opciones de tratamiento. El procedimiento descrito va a usarse para evaluar la idoneidad del
- 50 tratamiento endocrino como tratamiento para pacientes que padecen un trastorno proliferativo de células de los tejidos de mama. Por tanto, la presente divulgación se considerará para reducir los problemas asociados con los procedimientos de predicción de la respuesta a tratamiento actuales.

- 55 Con el uso de los procedimientos y ácidos nucleicos tal como se describe en el presente documento, puede evaluarse la capacidad de respuesta del paciente antes o durante el tratamiento endocrino para un trastorno proliferativo de células de los tejidos de mama, con el fin de proporcionar información crítica al paciente y al médico tales como los riesgos, las cargas y los beneficios asociados al tratamiento endocrino. Se apreciará, por tanto, que los procedimientos y ácidos nucleicos mostrados a modo de ejemplo en el presente documento pueden servir para mejorar la calidad de vida de un paciente y las probabilidades de éxito del tratamiento permitiendo tanto al paciente

como al médico una evaluación más precisa de las opciones de tratamiento del paciente.

El objetivo de la descripción se consigue de la manera más preferible por medio del análisis de los patrones de metilación de uno o una combinación de genes tomados del grupo tomados del grupo xxSTMN1, SFN, S100A2, TGFR2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF (véase la tabla 1) y/o sus regiones reguladoras. La invención **se caracteriza porque** el ácido nucleico de uno o una combinación de genes tomados del grupo STMN1, SFN, S100A2, TGFR2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF se ponen en contacto con un reactivo o serie de reactivos que puede distinguir entre dinucleótidos CpG metilados y no metilados dentro de la secuencia genómica de interés.

El objetivo de la divulgación puede conseguirse también mediante el análisis de la metilación de CpG de uno o una pluralidad de cualquier subconjunto del grupo de genes STMN1, SFN, S100A2, TGFR2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF, en particular se prefieren los siguientes subconjuntos:

- TP53, PTGS2, FGFR1, PSA, CGA, CYP2D6 y MSMB
- STMN 1, PITX2, PSA y CGA
- STUN1, SFN, S100A2, TGFR2, SYK, GRIN2D, PSA, COX7A2L, VTN y PRKCD
- ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF

La presente divulgación pone a disposición un procedimiento para determinar parámetros genéticos y/o epigenéticos de ADN genómico. El procedimiento es para su uso en el tratamiento y el control mejorados de trastornos proliferativos de células de mama, permitiendo la predicción exacta de la respuesta de un paciente al tratamiento con un tratamiento que comprende uno o más fármacos que se dirigen a la ruta del receptor de estrógeno o están implicados en el metabolismo, la producción o la secreción de estrógeno. En una realización particularmente preferida, el procedimiento dado a conocer permite la diferenciación entre pacientes que responden a dicho tratamiento y aquéllos que no.

El procedimiento de acuerdo con la divulgación puede usarse para el análisis de una amplia variedad de trastornos proliferativos de células de los tejidos de mama incluyendo, pero sin limitarse a, carcinoma ductal *in situ*, carcinoma lobular, carcinoma coloide, carcinoma tubular, carcinoma medular, carcinoma metaplásico, carcinoma intraductal *in situ*, carcinoma lobular *in situ* y carcinoma papilar *in situ*.

El procedimiento de acuerdo con la divulgación es particularmente adecuado para la predicción de la respuesta al tratamiento mencionado anteriormente en dos entornos de tratamiento. En una realización, el procedimiento se aplica a pacientes que reciben tratamiento dirigido a la ruta endocrina como tratamiento secundario a un tratamiento no quimioterápico inicial, por ejemplo cirugía (a continuación en el presente documento denominado el entorno complementario) tal como se ilustra en la figura 1. Un tratamiento de este tipo se recomienda con frecuencia a pacientes que padecen carcinomas de mama de estadio 1 a 3. En esta realización, los pacientes que responden se definen como aquéllos que no tienen una recidiva detectable del cáncer de mama en un periodo de tiempo especificado, los pacientes que no responden son aquéllos que experimentan una recidiva dentro de dicho periodo de tiempo.

En una realización adicionalmente preferida, dicho procedimiento se aplica a pacientes que padecen una recidiva de cáncer de mama tras el tratamiento mediante un medio primario (preferentemente cirugía) seguido de un periodo sin enfermedad, y en los que se ha recomendado el tratamiento que se dirige a la ruta endocrina en respuesta a una detección de una recidiva del carcinoma. Un tratamiento de este tipo se recomienda con frecuencia a pacientes que padecen carcinomas en estadio tardío, particularmente en los que ha aparecido metástasis. Por tanto, este entorno clínico debe denominarse también a continuación en el presente documento el "entorno metastático". En esta realización, los pacientes que responden son aquéllos que entran en remisión parcial o completa, es decir, sujetos cuyo cáncer retrocede hasta niveles indetectables a diferencia de aquéllos cuyas enfermedades metastatizan adicionalmente o permanecen por encima de niveles detectables. Esta metodología presenta mejoras adicionales con respecto al estado de la técnica porque el procedimiento puede aplicarse a cualquier sujeto, independientemente del estado de los receptores de estrógeno y/o de progesterona. Por tanto, en una realización preferida, no se requiere que el sujeto se haya sometido a prueba para determinar el estado de los receptores de estrógeno o de progesterona.

Además, el procedimiento permite el análisis de metilaciones de citosina y polimorfismos de nucleótidos individuales dentro de dichos genes.

El objeto de la divulgación se consigue por medio del análisis de los patrones de metilación de uno o más de los genes STMN1, SFN, S100A2, TGFB2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF y/o sus regiones reguladoras. En una realización particularmente preferida las secuencias de dichos genes comprenden las SEC ID N° 27, 40, 122, 43, 131, 50, 74, 127, 135, 86, 90, 128, 129, 99, 105, 115, 121, 126, 137, 129, 125, 132, 122, 123, 131, 133, 134, 127, 130, 135, 124, 128 y 136 y secuencias complementarias a las mismas.

El objeto de la divulgación puede conseguirse también analizando los patrones de metilación de uno o más genes tomados de los siguientes subconjuntos de dicho grupo de genes mencionado anteriormente. En una realización, el objetivo dado a conocer se consigue mediante el análisis de los patrones de metilación de uno o más genes tomados del grupo constituido por TP53, PTGS2, FGFR1, PSA, CGA, CYP2D6 y MSMB, y prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 68, 50, 74, 90, 91, 92 y 99. En una realización adicional, el objetivo dado a conocer se consigue mediante el análisis de los patrones de metilación de uno o más genes tomados del grupo constituido por STMN1, PITX2, PSA y CGA, y prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 27, 83, 90 y 91. El objetivo dado a conocer puede conseguirse también mediante el análisis de los patrones de metilación de uno o más genes tomados del grupo constituido por STMN1, SFN, S100A2, TGFB2, SYK, GRIN2D, PSA, COX7A2L, VTN y PRKCD, y prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 27, 40, 41, 43, 78, 86, 90, 105, 115, 121. También es una realización que el objetivo se consiga mediante el análisis de los patrones de metilación de uno o más genes tomados del grupo constituido por ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF, y prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 126, 137, 129, 125, 132, 122, 123, 131, 133, 134, 127, 130, 135, 124, 128 y 136.

En el caso de que el paciente reciba el tratamiento como tratamiento complementario, se prefiere que dichos genes se seleccionen del grupo constituido por TP53, PTGS2, PITX2, CYP2D6, MSMB, WBP11, TMEFF2, ESR1, PITX2 y PCAF, y prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 131, 50, 135, 129, 99, 137, 133, 134, 135 y 136.

También se prefieren los siguientes subconjuntos de este grupo de genes:

TP53, PTGS2, CYP2D6 y MSMB, prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 68, 50, 92 y 99.

PITX2, en el que se prefiere adicionalmente que la secuencia de dicho gen comprenda la SEC ID N° 83.

WBP11, TMEFF2, ERBB2, ESR1, PITX2 y PCAF, prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 137, 132, 133, 134, 135 y 136.

En el caso de que el paciente reciba el tratamiento en respuesta en un entorno metastásico se prefiere que dichos genes se seleccionen del grupo constituido por STMN1, SFN, S100A2, TGFB2, FGFR1, SYK, GRIN2D, PSA, CGA, COX7A2L, VTN, PRKCD, ONECUT2, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, SYK, RASSF1, PSAT1 y CGA, y prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 27, 40, 122, 43, 74, 127, 86, 90, 128, 105, 115, 121, 126, 129, 125, 132, 122, 123, 131, 127, 130, 124 y 128.

También se prefieren los siguientes subconjuntos de este grupo de genes:

FGFR1, PSA y CGA, prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 74, 90 y 91.

STMN1, PSA y CGA, prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 27, 90 y 91.

STMN1, SFN, S100A2, TGFB2, SYK, GRIN2D, PSA, COX7A2L, VTN y PRKCD, prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 27, 40, 41, 43, 78, 86, 90, 105, 115, 121.

ONECUT2, CYP2D6, DAG1, S100A2, TFF1, TP53, SYK, RASSF1, PSAT1 y CGA, prefiriéndose adicionalmente que la secuencia de dichos genes comprendan las SEC ID N° 126, 129, 125, 122, 123, 131, 127, 130, 124 y 128.

En una realización preferida, dicho procedimiento se consigue poniendo en contacto dichas secuencias de ácido nucleico en una muestra biológica obtenida de un sujeto con al menos un reactivo o una serie de reactivos, en el que dicho reactivo o serie de reactivos, distingue entre dinucleótidos CpG metilados y no metilados dentro del ácido nucleico diana.

En una realización preferida, el procedimiento comprende las siguientes etapas:

en la primera etapa del procedimiento, la muestra de ADN genómico debe aislarse de fuentes tales como

células o componentes celulares que contienen ADN, fuentes de ADN que comprenden, por ejemplo, líneas celulares, cortes histológicos, biopsias, tejido incrustado en parafina, tejidos de mama, sangre, plasma, líquido linfático, tejido linfático, células ductales, líquido de lavado ductal, líquido de aspiración del pezón, médula ósea y combinaciones de los mismos. La extracción puede ser por medios que son convencionales para un experto en la técnica, éstos incluyen el uso de lisados con detergente, sonicación y agitación con vórtex con perlas de vidrio. Una vez se han extraído los ácidos nucleicos se usa en el análisis el ADN bicatenario genómico.

En una realización preferida, el ADN puede escindirse antes de la siguiente etapa del procedimiento, esto puede ser mediante cualquier medio convencional en el estado de la técnica, en particular, pero sin limitarse a, con endonucleasas de restricción.

En la segunda etapa del procedimiento, la muestra de ADN genómico se trata de manera que las bases de citosina que no están metiladas en la posición 5' se convierten en uracilo, timina u otra base que es diferente de citosina en cuanto al comportamiento de hibridación. Esto se entenderá como "pretratamiento" a continuación en el presente documento.

El tratamiento descrito anteriormente del ADN genómico se lleva a cabo preferentemente con bisulfito (sulfito, disulfito) y posterior hidrólisis alcalina que da como resultado una conversión de nucleobases de citosina no metiladas en uracilo o en otra base que es diferente a citosina en cuanto al comportamiento de apareamiento de bases. Si se usa disolución de bisulfito para la reacción, entonces tiene lugar una adición en las bases de citosina no metiladas. Además, debe estar presente un disolvente o reactivo de desnaturalización así como un interceptor de radicales. Entonces, una hidrólisis alcalina posterior da lugar a la conversión de nucleobases de citosina no metiladas en uracilo. El ADN convertido se usa entonces para la detección de citosinas metiladas.

Se amplifican fragmentos del ADN tratado previamente, usando conjuntos de oligonucleótidos de cebador, y una polimerasa, preferentemente termoestable. Debido a consideraciones estadísticas y prácticas, se amplifican preferentemente más de seis fragmentos diferentes que tienen una longitud de 100 - 2000 pares de bases. La amplificación de varios segmentos de ADN puede llevarse a cabo simultáneamente en uno y el mismo recipiente de reacción. Habitualmente, la amplificación se lleva a cabo por medio de una reacción en cadena de la polimerasa (PCR).

El diseño de tales cebadores es obvio para un experto en la técnica. Éstos deben incluir al menos dos oligonucleótidos cuyas secuencias son cada una complementaria inversa o idéntica a un segmento de al menos 18 pares de bases de longitud de las siguientes secuencias de bases especificadas en el apéndice: SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794. Dichos oligonucleótidos de cebador **se caracterizan** preferentemente **porque** no contienen ningún dinucleótido CpG. En una realización particularmente preferida del procedimiento, la secuencia de dichos oligonucleótidos de cebador se diseña de modo que se aparee selectivamente a y amplifica, sólo el ADN específico de células de mama de interés, minimizando de ese modo la amplificación de ADN de fondo o no relevante. En el contexto de la presente descripción, se considera que ADN de fondo significa ADN genómico que no tiene un patrón de metilación específico de tejido relevante, siendo en este caso el tejido relevante tejidos de mama.

Según la presente divulgación, se prefiere que al menos un oligonucleótido de cebador esté unido a una fase sólida durante la amplificación. Las diferentes secuencias de oligonucleótidos y/o de oligómero de PNA pueden disponerse sobre una fase sólida plana en forma de retícula rectangular o hexagonal, estando compuesta la superficie de la fase sólida preferentemente por silicio, vidrio, poliestireno, aluminio, acero, hierro, cobre, níquel, plata u oro, siendo posible usar también otros materiales tales como nitrocelulosa o plásticos.

Los fragmentos obtenidos por medio de la amplificación pueden portar una etiqueta directa o indirectamente detectable. Se prefieren etiquetas en forma de etiquetas de fluorescencia, radionúclidos o fragmentos de moléculas separables que tienen una masa típica que puede detectarse en un espectrómetro de masas, prefiriéndose que los fragmentos que se producen tengan una única carga neta positiva o negativa para una mejor capacidad de detección en el espectrómetro de masas. La detección puede llevarse a cabo y visualizarse por medio de espectrometría de masas de desorción – ionización por láser asistida por matriz (MALDI) o usando espectrometría de masas por electropulverización (ESI).

En la siguiente etapa, se analizan los amplificadores de ácido nucleico con el fin de determinar el estado de metilación del ADN genómico antes del tratamiento.

El análisis posterior al tratamiento de los ácidos nucleicos puede llevarse a cabo usando procedimientos alternativos. A continuación se describen varios procedimientos para el análisis específico del estado de metilación de los ácidos nucleicos tratados, otros procedimientos alternativos serán obvios para un experto en la técnica.

El análisis puede llevarse a cabo durante la etapa de amplificación del procedimiento. En una realización de este tipo, el estado de metilación de posiciones de CpG seleccionadas previamente dentro de los genes xxSTMN1, SFN, S100A2, TGFBR2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT22, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF y/o sus regiones reguladoras puede detectarse mediante el uso de oligonucleótidos de cebador específicos de metilación. El término "MSP" (PCR específica de metilación) se refiere al ensayo de metilación reconocido en la técnica descrito por Herman *et al.* Proc. Natl. Acad. Sci. USA 93: 9821-9826, 1996, y dado a conocer también en las patentes estadounidenses n.º 5.786.146 y n.º 6.265.171. El uso de cebadores específicos del estado de metilación para la amplificación de ADN tratado con bisulfito permite la diferenciación entre ácidos nucleicos metilados y no metilados. Las parejas de cebadores de MSP contienen al menos un cebador que hibrida con un dinucleótido CpG tratado con bisulfito. Por tanto, la secuencia de dichos cebadores comprende al menos un dinucleótido CG, TG o CA. Los cebadores de MSP específicos para ADN no metilado contienen una "T" en la posición 3' de la posición C en el CpG. Según la presente invención, se prefiere por tanto que se requiera que la secuencia de bases de dichos cebadores comprenda una secuencia que tenga una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratada previamente según las SEC ID N.º: 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794 y secuencias complementarias a las mismas, comprendiendo la secuencia de bases de dichos oligómeros al menos un dinucleótido CG, TG o CA.

En una realización del procedimiento, el estado de metilación de las posiciones de CpG puede determinarse por medio de análisis de hibridación. En esta realización del procedimiento, los amplificados obtenidos en la segunda etapa del procedimiento se hibridan con una serie o un conjunto de oligonucleótidos y/o sondas de PNA. En este contexto, la hibridación tiene lugar de la manera descrita tal como sigue. El conjunto de sondas usadas durante la hibridación está compuesto preferentemente por al menos 4 oligonucleótidos u oligómeros de PNA. En el procedimiento, los amplificados sirven como sondas que hibridan con oligonucleótidos unidos previamente a una fase sólida. Los fragmentos no hibridados se eliminan posteriormente. Dichos oligonucleótidos contienen al menos una secuencia de bases que tiene una longitud de 10 nucleótidos que es complementaria inversa o idéntica a un segmento de las secuencias de bases especificadas en el apéndice, conteniendo el segmento al menos un dinucleótido CpG o TpG. En una realización adicionalmente preferida la citosina del dinucleótido CpG, o en el caso de TpG, la tiamina, es el 5º a 9º nucleótido desde el extremo 5' del 10-mero. Existe un oligonucleótido para cada dinucleótido CpG o TpG.

Entonces se eliminan los amplificados no hibridados. En la etapa final del procedimiento, se detectan los amplificados hibridados. En este contexto, se prefiere que las etiquetas unidas a los amplificados puedan identificarse en cada posición de la fase sólida en la que se ubica una secuencia de oligonucleótidos.

En una realización adicional del procedimiento, el estado de metilación de las posiciones de CpG puede determinarse por medio de sondas de oligonucleótido que se hibridan con el ADN tratado simultáneamente con los cebadores de amplificación por PCR (pudiendo ser dichos cebadores o bien específicos de metilación o bien convencionales).

Una realización particularmente preferida de este procedimiento es el uso de PCR cuantitativa en tiempo real basada en fluorescencia (Heid *et al.*, Genome Res. 6:986-994, 1996) empleando una sonda de oligonucleótido fluorescente marcada dual (PCR de TaqMan™, usando un sistema de detección de secuencias ABI Prism 7700, Perkin Elmer Applied Biosystems, Foster City, California). La reacción de PCR de TaqMan™ emplea el uso de un oligonucleótido interrogante no extensible, denominado una sonda de TaqMan™, que está diseñado para hibridar con una secuencia rica en GpC ubicada entre los cebadores de amplificación directo e inverso. La sonda de TaqMan™ comprende además un "resto indicador" fluorescente y un "resto extintor" covalentemente unidos a restos de ligador (por ejemplo, fosforamidas) unidos a los nucleótidos del oligonucleótido de TaqMan™. Para el análisis de la metilación dentro de ácidos nucleicos posterior al tratamiento con bisulfito se requiere que la sonda sea específica de metilación, tal como se describe en el documento U.S. 6.331.393, (incorporado por el presente documento por referencia) conocido también como ensayo Methyl Light. Las variaciones en la metodología de detección TaqMan™ que también son adecuadas para su uso con los procedimientos descritos incluyen el uso de tecnología de sondas duales (Lightcycler™) o cebadores de amplificación fluorescentes (tecnología Sunrise™). Estas dos técnicas pueden adaptarse de manera adecuada para su uso con ADN tratado con bisulfito, y además para el análisis de la metilación dentro de dinucleótidos CpG.

Un procedimiento adecuado adicional para el uso de oligonucleótidos de sonda para la evaluación de la metilación mediante el análisis de ácidos nucleicos tratados con bisulfito es el uso de oligonucleótidos bloqueantes. El uso de tales oligonucleótidos se ha descrito en BioTechniques 23(4), 1997, 714-720 D. Yu, M. Mukai, Q. Liu, C. Steinman. Se hibridan oligonucleótidos de sonda de bloqueo con el ácido nucleico tratado con bisulfito simultáneamente con los cebadores de PCR. La amplificación por PCR del ácido nucleico se termina en la posición 5' de la sonda de bloqueo, suprimiéndose de ese modo la amplificación de un ácido nucleico en el que está presente la secuencia

complementaria a la sonda de bloqueo. Las sondas pueden diseñarse para hibridar con el ácido nucleico tratado con bisulfito de una manera específica del estado de metilación. Por ejemplo, para la detección de ácidos nucleicos metilados dentro de una población de ácidos nucleicos no metilados, se llevaría a cabo la supresión de la amplificación de los ácidos nucleicos que no están metilados en la posición en cuestión mediante el uso de sondas de bloqueo que comprenden un "CG" en la posición en cuestión, a diferencia de un "CA".

Para procedimientos de PCR usando oligonucleótidos bloqueantes, la perturbación eficaz de la amplificación mediada por polimerasa requiere que los oligonucleótidos bloqueantes no se alarguen mediante la polimerasa. Preferentemente, esto se consigue a través del uso de bloqueantes que son 3'-desoxioligonucleótidos, u oligonucleótidos derivatizados en la posición 3' con un grupo distinto de un hidroxilo "libre". Por ejemplo, 3'-O-acetiloligonucleótidos son representativos de una clase preferida de molécula bloqueante.

Adicionalmente, debe impedirse la descomposición mediada por la polimerasa de los oligonucleótidos bloqueantes. Preferentemente, tal impedimento comprende o bien el uso de una polimerasa que carece de actividad exonucleasa en 5'-3', o bien el uso de oligonucleótidos bloqueantes modificados que tienen, por ejemplo, puentes de tioato en los extremos 5' de los mismos que hacen a la molécula bloqueante resistente a la nucleasa. Aplicaciones particulares pueden no requerir tales modificaciones en 5' del bloqueante. Por ejemplo, si los sitios de unión del bloqueante y del cebador solapan, impidiendo de ese modo la unión del cebador (por ejemplo, con bloqueante en exceso), se impedirá sustancialmente la degradación del oligonucleótido bloqueante. Esto se debe a que la polimerasa no extenderá el cebador hacia, y a través (en la dirección 5'-3') del bloqueante, un proceso que normalmente da como resultado la degradación del oligonucleótido bloqueante hibridado.

Una realización de bloqueante/PCR particularmente preferida, para fines de la presente descripción y tal como se pone en práctica en el presente documento, comprende el uso de oligómeros de ácido nucleico peptídico (PNA) como oligonucleótidos de bloqueo. Tales oligómeros bloqueantes de PNA son adecuados de manera ideal, debido a que no se descomponen ni extienden por la polimerasa.

Preferentemente, por tanto, se requiere que la secuencia de bases de dichos oligonucleótidos de bloqueo comprenda una secuencia que tenga una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratada previamente según una de las SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794 y secuencias complementarias a las mismas, comprendiendo la secuencia de bases de dichos oligonucleótidos al menos un dinucleótido CpG, TpG o CpA.

En una realización adicionalmente preferida del procedimiento, la determinación del estado de metilación de las posiciones de CpG se lleva a cabo mediante el uso de extensión de oligonucleótidos dirigida por molde, tal como MS SNUPE tal como se describe por Gonzalgo y Jones (Nucleic Acids Res. 25:2529-2531).

En una realización adicional del procedimiento, se permite la determinación del estado de metilación de las posiciones de CpG mediante la secuenciación y el análisis de secuencia posterior del amplificado generado en la segunda etapa del procedimiento (Sanger F., *et al.*, 1977 PNAS USA 74: 5463-5467).

El procedimiento de acuerdo con la divulgación puede permitirse mediante cualquier combinación de los medios anteriores. En un modo particularmente preferido, se combina simultáneamente el uso de sondas de detección en tiempo real con MSP y/o oligonucleótidos bloqueantes.

Una realización adicional es un procedimiento para el análisis del estado de metilación del ADN genómico sin la necesidad de tratamiento previo. En las etapas primera y segunda del procedimiento, la muestra de ADN genómico debe obtenerse y aislarse de fuentes celulares o de tejido. Tales fuentes pueden incluir líneas celulares, cortes histológicos, fluidos corporales o tejido incrustado en parafina. La extracción puede ser por medios que son convencionales para un experto en la técnica, éstos incluyen el uso de lisados con detergente, sonicación y agitación con vórtex con perlas de vidrio. Una vez se han extraído los ácidos nucleicos, se usa en el análisis el ADN bicatenario genómico.

En una realización preferida, el ADN puede escindir-se antes del tratamiento, esto puede ser mediante cualquier medio convencional en el estado de la técnica, en particular con endonucleasas de restricción. En la tercera etapa, se digiere entonces el ADN con una o más enzimas de restricción sensibles a la metilación. La digestión se lleva a cabo de modo que la hidrólisis del ADN en el sitio de restricción informe del estado de metilación de un dinucleótido CpG específico.

En una realización preferida se amplifican los fragmentos de restricción. En una realización adicionalmente preferida, esto se lleva a cabo usando la reacción en cadena de la polimerasa.

En la etapa final se detectan los amplificados. La detección puede ser mediante cualquier medio convencional en la

técnica, por ejemplo, pero sin limitarse a, análisis mediante electroforesis en gel, análisis de hibridación, incorporación de etiquetas detectables dentro de los productos de PCR, análisis de series de ADN, análisis mediante MALDI o ESI.

5 El procedimiento mencionado anteriormente se usa preferentemente para determinar parámetros genéticos y/o epigenéticos del ADN genómico.

10 Con el fin de permitir este procedimiento, la descripción proporciona además el ADN modificado de uno o una combinación de genes tomados del grupo STMN1, SFN, S100A2, TGFB2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF así como oligonucleótidos y/u oligómeros de PNA para detectar metilaciones de citosina dentro de dichos genes. La presente descripción se basa en el descubrimiento de que los parámetros genéticos y epigenéticos y, en particular, los patrones de metilación de citosina de dichos ADN genómicos, son particularmente adecuados para el tratamiento y el control mejorados de trastornos proliferativos de células de mama.

15 Los ácidos nucleicos de acuerdo con la presente divulgación pueden usarse para el análisis de parámetros genéticos y/o epigenéticos del ADN genómico.

20 Este objetivo de acuerdo con la presente divulgación se consigue usando un ácido nucleico que contiene una secuencia de al menos 18 bases de longitud del ADN genómico tratado previamente según una de las SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794 y secuencias complementarias a las mismas.

25 El objetivo de la divulgación puede conseguirse también usando un ácido nucleico que contiene una secuencia de al menos 18 bases de longitud del ADN genómico tratado previamente según una de las SEC ID N° 345, 346, 381, 382, 393, 394, 425, 426, 427, 428, 429, 430, 443, 444, 619, 620, 655, 656, 667, 668, 699, 700, 701, 702, 703, 704, 717 y 718 y secuencias complementarias a las mismas.

30 El objetivo de la divulgación puede conseguirse también usando un ácido nucleico que contiene una secuencia de al menos 18 bases de longitud del ADN genómico tratado previamente según una de las SEC ID N° 299, 300, 411, 412, 425, 426, 427, 428, 573, 574, 685, 686, 699, 700, 701 y 702 y secuencias complementarias a las mismas.

35 El objetivo de la divulgación puede conseguirse también usando un ácido nucleico que contiene una secuencia de al menos 18 bases de longitud del ADN genómico tratado previamente según una de las SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 401, 402, 417, 418, 425, 426, 455, 456, 475, 476, 487, 488, 573, 574, 599, 600, 601, 602, 605, 606, 675, 676, 691, 692, 699, 700, 729, 730, 749, 750, 761, 762 y secuencias complementarias a las mismas.

40 El objetivo de la divulgación puede conseguirse también usando un ácido nucleico que contiene una secuencia de al menos 18 bases de longitud del ADN genómico tratado previamente según una de las SEC ID N° 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794 y secuencias complementarias a las mismas.

Los ácidos nucleicos modificados no podían conectarse hasta ahora con la determinación de parámetros genéticos y epigenéticos relevantes de enfermedad.

45 El objetivo de la presente divulgación se consigue además mediante un oligonucleótido u oligómero para el análisis de ADN tratado previamente, para detectar el estado de metilación de citosina genómica, conteniendo dicho oligonucleótido al menos una secuencia de bases que tiene una longitud de al menos 10 nucleótidos que hibrida con un ADN genómico tratado previamente según las SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794. Las sondas de oligómero dadas a conocer constituyen herramientas importantes y eficaces que, por primera vez, hacen posible determinar parámetros genéticos y epigenéticos específicos durante el análisis de muestras biológicas para determinar características asociadas con la respuesta de un paciente al tratamiento endocrino. Dichos oligonucleótidos permiten el tratamiento y el control mejorados de trastornos proliferativos de células de mama. La secuencia de bases de los oligómeros contiene preferentemente al menos un dinucleótido CpG o TpG. Las sondas pueden existir también en forma de un PNA (ácido nucleico peptídico) que

tiene propiedades de apareamiento particularmente preferidas. Se prefieren particularmente oligonucleótidos en los que la citosina del dinucleótido CpG esté dentro del tercio medio de dicho oligonucleótido, por ejemplo el 5° - 9° nucleótido desde el extremo 5' de un oligonucleótido de 13 meros; o en el caso de oligómeros de PNA, se prefiere que la citosina del dinucleótido CpG sea el 4° - 6° nucleótido desde el extremo 5' del 9-mero.

- 5 Los oligómeros de acuerdo con la presente divulgación se usan normalmente en los denominados "conjuntos" que contienen al menos dos oligómeros y hasta un oligómero para cada uno de los dinucleótidos CpG dentro de las SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794.

- 10 En una realización particularmente preferida del procedimiento, los oligómeros según las SEC ID N° 1733-1736, 1967-1994, 2003-2024, 2031 - 2034, 2045-2134 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno metastásico.

En una realización adicional del procedimiento, los oligómeros según las SEC ID N° 2011, 2012, 2017 - 2024 y 2031 - 2034 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno metastásico.

En una realización adicional del procedimiento, los oligómeros según las SEC ID N° 2003 - 2024 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno metastásico.

- 20 En una realización adicional del procedimiento, los oligómeros según las SEC ID N° 2003- 2020 y 2045- 2112 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno metastásico.

En una realización adicional del procedimiento, los oligómeros según las SEC ID N° 1733-1736, 1967-1994, 2011-2025, 2045-2052, 2069-2078 y 2127-2134 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno metastásico.

- 25 En una realización particularmente preferida del procedimiento, los oligómeros según las SEC ID N° 1691-1692, 1925-1932, 1941-1954, 1965-1966, 1995-2002, 2025-2030, 2035 - 2044 y 2135 - 2142 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno complementario.

En una realización adicional del procedimiento, los oligómeros según las SEC ID N° 2035- 2044 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno complementario.

- 30 En una realización adicional del procedimiento, los oligómeros según las SEC ID N° 2025-2030 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno complementario.

En una realización adicional del procedimiento, los oligómeros según las SEC ID N° 1691-1692, 1925-1932, 1941-1954, 1965-1966, 1995-2002 y 2135 -2142 se usan para predecir la respuesta de un sujeto al tratamiento endocrino en un entorno complementario.

- 35 En el caso de los conjuntos de oligonucleótidos dados a conocer, se prefiere que al menos un oligonucleótido esté unido a una fase sólida. Se prefiere adicionalmente que todos los oligonucleótidos de un conjunto estén unidos a una fase sólida.

La presente divulgación se refiere además a un conjunto de al menos 2 n (oligonucleótidos y/u oligómeros de PNA) usado para detectar el estado de metilación de citosina del ADN genómico, mediante el análisis de dicha secuencia o versiones tratadas de dicha secuencia (de los genes STMN1, SFN, S100A2, TGFB2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF, tal como se detalla en el listado de secuencias y en la tabla 1) y secuencias complementarias a las mismas). Estas sondas permiten el tratamiento y el control mejorados de trastornos proliferativos de células de mama.

- 40 El conjunto de oligómeros puede usarse también para detectar polimorfismos de nucleótidos individuales (SNP) mediante el análisis de dicha secuencia o versiones tratadas de dicha secuencia de los genes STMN1, SFN, S100A2, TGFB2, TP53, PTGS2, FGFR1, SYK, PITX2, GRIN2D, PSA, CGA, CYP2D6, MSMB, COX7A2L, VTN, PRKCD, ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF.

- 50 Será obvio para un experto en la técnica que el procedimiento dado a conocer se mejorará y complementará mediante la incorporación de marcadores e indicadores clínicos conocidos el estado de la técnica y actualmente usados como factor de predicción del resultado de tratamientos que se dirigen a rutas endocrinas o endocrino-asociadas. Más preferentemente, dichos marcadores incluyen el estado de los ganglios, la edad, el estado menopáusico, el grado, receptores de estrógeno y de progesterona.

Los genes que forman la base de la presente descripción pueden usarse para formar un “panel de genes”, es decir una colección que comprende las secuencias genéticas particulares de la presente invención y/o sus sitios de metilación de información respectivos. La formación de paneles de genes permite un análisis rápido y específico de aspectos específicos del tratamiento del cáncer de mama. El/los panel(es) de genes tal como se describe(n) y emplea(n) en el presente documento puede(n) usarse con una eficacia sorprendentemente alta para el tratamiento de trastornos proliferativos de células de mama mediante la predicción del resultado del tratamiento con un tratamiento que comprende uno o más fármacos que se dirigen a la ruta del receptor de estrógeno o que están implicados en el metabolismo, la producción o la secreción de estrógeno. El análisis de cada gen del panel contribuye a la evaluación de la capacidad de respuesta del paciente, sin embargo, en una realización menos preferida, la evaluación del paciente puede conseguirse mediante el análisis de sólo un único gen. El análisis de un único miembro del “panel de genes” permitiría un medio barato pero menos preciso de evaluación de la capacidad de respuesta del paciente, el análisis de múltiples miembros del panel proporcionaría un medio bastante más caro de llevar a cabo el procedimiento, pero con una mayor precisión (la solución técnicamente preferida).

La eficacia del procedimiento dado a conocer se mejora cuando se aplica a pacientes que no se han tratado con quimioterapia. En consecuencia, es una realización particularmente preferida del procedimiento, en el caso de que el procedimiento se use para la evaluación de sujetos que no se han sometido a quimioterapia.

Según la presente divulgación, se prefiere que una disposición de diferentes oligonucleótidos y/u oligómeros de PNA (una denominada “serie”) puesta a disposición por la presente descripción esté presente de una manera que esté probablemente unida a una fase sólida. Esta serie de diferentes secuencias de oligonucleótidos y/u oligómeros de PNA puede caracterizarse porque está dispuesta sobre la fase sólida en forma de una redícula rectangular o hexagonal. La superficie de la fase sólida está compuesta preferentemente por silicio, vidrio, poliestireno, aluminio, acero, hierro, cobre, níquel, plata u oro. Sin embargo, son alternativas adecuadas nitrocelulosa así como plásticos tales como nailon que puede existir en forma de gránulos o también como matrices de resina.

Por tanto, una materia objeto adicional de la presente divulgación es un procedimiento para la fabricación de una serie fijada a un material portador para el tratamiento y el control mejorados de trastornos proliferativos de células de mama. En dicho procedimiento, al menos un oligómero según la presente descripción está acoplado a una fase sólida. Se conocen procedimientos para la fabricación de tales series, por ejemplo, de la patente estadounidense 5.744.305 por medio de la química de fase sólida y grupos protectores fotolábiles.

Una materia objeto adicional de la presente divulgación se refiere a un chip de ADN para el tratamiento y el control mejorados de trastornos proliferativos de células de mama. El chip de ADN contiene al menos un ácido nucleico según la presente descripción. Se conocen chips de ADN, por ejemplo, en la patente estadounidense 5.837.832.

Además, una materia objeto de la presente divulgación es un kit que puede estar compuesto, por ejemplo, por un reactivo que contiene bisulfito, un conjunto de oligonucleótidos de cebador que contienen al menos dos oligonucleótidos cuyas secuencias corresponden en cada caso a o son complementarias a un segmento de 18 bases de longitud de las secuencias de bases especificadas en las SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794 y/u oligómeros de PNA así como instrucciones para llevar a cabo y evaluar el procedimiento descrito.

En una realización adicionalmente preferida, dicho kit puede comprender además reactivos convencionales para realizar un análisis de metilación específico de posición de CpG, comprendiendo dicho análisis una o más de las siguientes técnicas: MS-SNuPE, MSP, Methyl Light, Heavy Methyl, y secuenciación de ácidos nucleicos. Sin embargo, un kit del tipo de la presente descripción puede contener también sólo parte de los componentes mencionados anteriormente.

Los reactivos típicos (por ejemplo, tal como podrían encontrarse en un kit basado en MethyLight típico) para el análisis MethyLight® pueden incluir, pero no se limitan a: cebadores de PCR para un gen específico (o isla de CpG o secuencia de ADN alterada mediante metilación); sondas de TaqMan®; tampones de PCR optimizados y desoxinucleótidos; y Taq polimerasa.

Los reactivos típicos (por ejemplo, tal como se encontrarían en un kit basado en Ms-SNuPE típico) para análisis de Ms-SNuPE pueden incluir, pero no se limitan a: cebadores de PCR para un gen específico (o isla de CpG o secuencia de ADN alterada mediante metilación); tampones de PCR optimizados y desoxinucleótidos; kit de extracción en gel; cebadores control positivos; cebadores de Ms-SNuPE para un gen específico; tampón de reacción (para la reacción de Ms-SNuPE); y nucleótidos radiactivos. Adicionalmente, los reactivos de conversión con bisulfito pueden incluir: tampón de desnaturalización de ADN; tampón de sulfonación; kit o reactivos de recuperación de ADN (por ejemplo, columna de afinidad, ultrafiltración, precipitación); tampón de desulfonación; y componentes de recuperación de ADN.

Los reactivos típicos (por ejemplo, tal como podrían encontrarse en un kit basado en MSP típico) para el análisis de MSP pueden incluir, pero no se limitan a: cebadores de PCR metilados y no metilados para un gen específico (o isla de CpG o secuencia de ADN alterada mediante metilación), tampones de PCR optimizados y desoxinucleótidos, y sondas específicas.

- 5 Los oligómeros según la presente divulgación o series de los mismos así como un kit según la presente invención están previstos para usarse para el tratamiento y el control mejorados de trastornos proliferativos de células de mama. Según la presente descripción, el procedimiento se usa preferentemente para el análisis de parámetros genéticos y/o epigenéticos importantes dentro del ADN genómico, en particular para su uso en el tratamiento y el control mejorados de trastornos proliferativos de células de mama.

- 10 Los procedimientos de la presente divulgación se usan para el tratamiento y el control mejorados del trastorno proliferativo de células de mama permitiendo regímenes de tratamiento más informados.

La presente divulgación se refiere además al diagnóstico y/o pronóstico de acontecimientos que son desfavorables o relevantes para pacientes o individuos en los que parámetros genéticos y/o epigenéticos importantes dentro del ADN genómico, dichos parámetros obtenidos por medio de la presente descripción pueden compararse con otro conjunto de parámetros genéticos y/o epigenéticos, sirviendo las diferencias como la base para el diagnóstico y/o pronóstico de acontecimientos que son desfavorables o relevantes para pacientes o individuos.

- 15 En el contexto de la presente divulgación, el término “hibridación” ha de entenderse como un enlace de un oligonucleótido con una secuencia completamente complementaria del tipo de los apareamientos de base de Watson-Crick en el ADN de muestra, formando una estructura doble.

- 20 En el contexto de la presente divulgación, “parámetros genéticos” son mutaciones y polimorfismos de ADN genómico y secuencias requeridas adicionalmente para su regulación. Han de designarse como mutaciones, en particular, inserciones, deleciones, mutaciones puntuales, inversiones y polimorfismos y, se prefieren particularmente, SNP (polimorfismos de nucleótidos individuales).

- 25 En el contexto de la presente divulgación, se considera que el término “estado de metilación” significa el grado de metilación presente en un ácido nucleico de interés, éste puede expresarse en términos absolutos o relativos, es decir, como un porcentaje u otro valor numérico o mediante comparación con otro tejido y describirse en ese caso como hipermetilado, hipometilado o como que tiene un estado de metilación significativamente similar o idéntico.

- 30 En el contexto de la presente descripción, se considera que la expresión “región reguladora” de un gen significa secuencias de nucleótidos que afectan a la expresión de un gen. Dichas regiones reguladoras pueden ubicarse dentro de, de manera proximal o distal con respecto a dicho gen. Dichas regiones reguladoras incluyen, pero no se limitan a promotores constitutivos, promotores específicos de tejido, promotores específicos de desarrollo, promotores inducibles y similares. Los elementos reguladores de promotor pueden incluir también ciertos elementos de secuencia potenciadores que controlan la eficacia de transcripción o traducción del gen.

- 35 En el contexto de la presente divulgación, se considera que el término “quimioterapia” significa el uso de fármacos o sustancias químicas para tratar el cáncer. Esta definición excluye la radioterapia (tratamiento con rayos o partículas de alta energía), terapia hormonal (tratamiento con hormonas o análogos de hormonas (sustitutos sintéticos) y tratamiento quirúrgico.

- 40 En el contexto de la presente descripción, “parámetros epigenéticos” son, en particular, metilaciones de citosina y modificaciones adicionales de bases de ADN de ADN genómico y secuencias requeridas adicionalmente para su regulación. Los parámetros epigenéticos adicionales incluyen, por ejemplo, la acetilación de histonas que, no pueden analizarse directamente usando el procedimiento descrito pero que, a su vez, se correlaciona con la metilación de ADN.

- 45 En el contexto de la presente divulgación, se considera que la expresión “tratamiento complementario” significa un tratamiento de un paciente con cáncer inmediatamente después de un tratamiento inicial no quimioterápico, por ejemplo cirugía. En general, el fin de un tratamiento complementario es proporcionar un riesgo significativamente más pequeño de recaídas en comparación a sin el tratamiento complementario.

En el contexto de la presente divulgación, se considera que la expresión “positivo para receptores de estrógeno y/o de progesterona” significa células que expresan en su superficie receptores que son sensibles a la unión de estrógenos y/o progesteronas.

- 50 En lo siguiente, se explicará en mayor detalle la presente descripción basándose en las secuencias, figuras y ejemplos sin limitarse a los mismos.

- 55 La figura 1 muestra una aplicación preferida del procedimiento dado a conocer. El eje X muestra la masa del/de los tumor(es), mostrando la línea “3” el límite de detectabilidad. El eje Y muestra el tiempo. En consecuencia, dicha figura ilustra un modelo simplificado de tratamiento endocrino de un tumor de mama de estadio 1-3 en el que el tratamiento primario fue cirugía (en el punto 1), seguido de tratamiento complementario con tamoxifeno. En un

primer caso, se muestra que un paciente que responde al tratamiento (4) permanece por debajo del límite de detectabilidad durante la duración de la observación. Un paciente que no responde al tratamiento (5) tiene un periodo de supervivencia libre de enfermedad (2) seguido de recidiva cuando la masa del carcinoma alcanza el nivel de detectabilidad.

5 La figura 2 muestra otra aplicación preferida del procedimiento dado a conocer. El eje X muestra la masa del/de los tumor(es), mostrando la línea "3" el límite de detectabilidad. El eje Y muestra el tiempo. En consecuencia, dicha figura ilustra un modelo simplificado de tratamiento endocrino de un tumor de mama de estadio tardío en el que el tratamiento primario fue cirugía (en el punto 1), seguido de recidiva que se trata mediante tamoxifeno (2). En un primer caso se muestra que un paciente que responde al tratamiento (4) permanece por debajo del límite de detectabilidad durante la duración de la observación. Un paciente que no responde al tratamiento (5) no se recupera de la recidiva.

10 La figura 3 muestra el análisis de metilación de islas de CpG según el ejemplo 1. Se agruparon las islas de CpG por gen y se determinó su correlación con la respuesta objetivo mediante estadísticos T^2 de Hotelling. Los puntos negros indican el valor de P del gen indicado. Se muestran los 20 genes más informativos, clasificados de izquierda a derecha con valor de P creciente. La línea de puntos superior marca el valor de significación sin corregir ($P < 0,05$). La línea de puntos inferior marca la significación tras una corrección de la tasa de falsos positivos del 25%. Todos los genes con un valor de P inferior o igual al gen con el mayor valor de P que está por debajo de la línea inferior (en este caso COX7A2L) se consideran significativos. La corrección FDR seleccionada garantiza que los genes identificados tienen un 75% de posibilidades de ser verdaderos descubrimientos.

20 La figura 4 muestra los datos sin procesar de metilación de dinucleótido CpG de los 10 genes que se correlacionan significativamente con la tasa de respuesta global tal como se determina en la figura 3 (ejemplo 1). Tumores de pacientes que responden objetivos (CR + PR) se muestran a la izquierda; tumores de pacientes con enfermedad progresiva (PD) se muestran a la derecha. Las sombras en gris representan la distancia relativa del estado de metilación de dinucleótido CpG con respecto al valor medio. Se agrupan los dinucleótidos CpG por gen. Los genes indicados están ordenados de arriba a abajo con significación creciente. El color gris claro representa un dinucleótido CpG hipometilado mientras que el color gris oscuro indica un dinucleótido CpG hipermetilado. Los datos de metilación se centran y normalizan con respecto a una desviación estándar para cada dinucleótido CpG individual.

30 Las figuras 5 a 11 muestran una matriz de clasificación de los datos adicionales obtenidos según el ejemplo 1 (conjunto de datos 1) según diferencias de metilación de CpG entre las dos clases de tejidos usando un algoritmo. Las figuras 5, 7, 9 y 11 se muestran en escala de grises, estando las posiciones de CpG más significativas en la parte inferior de la matriz disminuyendo la significación hacia la parte superior. El color negro indica la metilación total en una posición de CpG dada, el blanco representa ausencia de metilación en la posición particular, con los grados de metilación representados en gris, desde claro (baja proporción de metilación) hasta oscuro (alta proporción de metilación). Cada fila representa una posición de CpG específica dentro de un gen y cada columna muestra el perfil de metilación para los diferentes CpG para una muestra. En el lado izquierdo, se muestra el nombre del gen, se muestran en el lado derecho los valores de p para las posiciones de CpG individuales. Los valores de p son las probabilidades de que la distribución observada se produzca por casualidad en el conjunto de datos. Las Figuras 6, 8, 10 y 12 son las versiones en rojo-verde de las figuras anteriores (es decir, las figuras 5, 7, 9 y 11, respectivamente). El rojo indica metilación total en una posición de CpG dada, el verde representa ausencia de metilación en la posición concreta.

45 Las Figuras 12 y 13 muestran una matriz de clasificación de los datos obtenidos según el ejemplo 1 según diferencias de metilación de CpG entre las dos clases de tejidos (paciente que responde y paciente que no responde al tratamiento endocrino en un entorno metastásico), usando un algoritmo. Las figuras se muestran en escala de grises, estando las posiciones de CpG más significativas en la parte inferior de la matriz disminuyendo la significación hacia la parte superior. El color negro indica la metilación total en una posición de CpG dada, el blanco representa ausencia de metilación en la posición particular, con los grados de metilación representados en gris, desde claro (baja proporción de metilación) hasta oscuro (alta proporción de metilación). Cada fila representa una posición de CpG específica dentro de un gen y cada columna muestra el perfil de metilación para los diferentes CpG para una muestra. En el lado izquierdo, se muestra el nombre del gen, se muestran en el lado derecho los valores de p para las posiciones de CpG individuales. Los valores de p son las probabilidades de que la distribución observada se produzca por casualidad en el conjunto de datos.

55 La figura 14 muestra las curvas de supervivencia libre de progresión (PFS) de puntuaciones de combinación según el ejemplo 1 (conjunto de datos 3). Se muestran las curvas de PFS tras el comienzo del tratamiento con tamoxifeno de los pacientes. Se agrupan usando una puntuación de predicción ($\chi^2=14,0$; 3 grados de libertad) basada en los marcadores de predicción tradicionales edad, sitio dominante de recidiva, intervalo libre de enfermedad y nivel de ER (A). Se presenta una puntuación basada en la metilación del ADN ($\chi^2=27,5$; 3 grados de libertad) que comprende los genes de predicción estadísticamente independientes fosfoserina amino transferasa (PSA-T), estatmina (STMN1), proteína A2 de S100 (S100A2, receptor II del factor de crecimiento transformante β (TGFB2), y receptor 2D de glutamato (GRIN2D) (B). Se generaron curvas de supervivencia y se calculó la significación tal como se describe en el ejemplo 1. Los valores de predicción calculados se dividieron en 4 grupos (q1 - q4) con 50 pacientes

cada uno y usando el percentil 25, 50 y 75 de los valores. Los números por debajo del eje X representan el número de pacientes que corren riesgo en cada grupo en los puntos de tiempo indicados de 0, 12, 24 y 36 meses. La supervivencia libre de progresión (en porcentaje) está representada en el eje Y y el tiempo hasta la progresión (en meses) está representado en el eje X.

- 5 La figura 15 muestra los valores de p sin corregir en una escala logarítmica. Se calcularon valores de P a partir de pruebas de cocientes de probabilidades (LR) de modelos de regresión logística multivariable. Cada región genómica individual de interés se representa como un punto, la línea de puntos superior representa el punto de corte para la tasa de falso descubrimiento del 25%, la línea de puntos inferior muestra el límite del 5% corregido de Bonferroni.

- 10 La figura 16 muestra una matriz de clasificación de los 10 mejores amplificados de los datos obtenidos según el ejemplo 1 (conjunto de datos 4). Se calcularon valores de P a partir de pruebas de cocientes de probabilidades (LR) de modelos de regresión logística multivariable. La figura se muestra en escala de grises, estando las posiciones de CpG más significativas en la parte inferior de la matriz disminuyendo la significación hacia la parte superior. El color negro indica la metilación total en una posición de CpG dada, el color blanco representa ausencia de metilación en la posición particular, con los grados de metilación representados en gris, desde claro (baja proporción de metilación) hasta oscuro (alta proporción de metilación). Cada fila representa una posición de CpG específica dentro de un gen y cada columna muestra el perfil de metilación para los diferentes CpG para una muestra. Los valores de p para las posiciones de CpG individuales se muestran en el lado derecho. Los valores de p son las probabilidades de que la distribución observada se produzca por casualidad en el conjunto de datos.

- 20 La figura 17 muestra las curvas de supervivencia libre de enfermedad estimadas de Kaplan-Meier para el gen ESR1, la línea de puntos (curva inferior) muestra pacientes que no responden mientras que la línea continua (curva superior) muestra pacientes que responden.

La figura 18 muestra las curvas de supervivencia libre de enfermedad estimadas de Kaplan-Meier para el gen PCAF, la línea de puntos (curva superior) muestra pacientes que no responden mientras que la línea continua (curva inferior) muestra pacientes que responden.

- 25 La figura 19 muestra las curvas de supervivencia libre de enfermedad estimadas de Kaplan-Meier para el gen PITX2, la línea de puntos (curva superior) muestra pacientes que no responden mientras que la línea continua (curva inferior) muestra pacientes que responden.

- 30 La figura 20 muestra las curvas de supervivencia libre de enfermedad estimadas de Kaplan-Meier para el gen TMEFF2, la línea de puntos (curva superior) muestra pacientes que no responden mientras que la línea continua (curva inferior) muestra pacientes que responden.

La figura 21 muestra las curvas de supervivencia libre de enfermedad estimadas de Kaplan-Meier para el gen WBP11, la línea de puntos (curva inferior) muestra pacientes que no responden mientras que la línea continua (curva superior) muestra pacientes que responden.

- 35 La figura 22 muestra las curvas de supervivencia libre de enfermedad estimadas de Kaplan-Meier para el gen ERBB2, la línea de puntos (curva inferior) muestra pacientes que no responden mientras que la línea continua (curva superior) muestra pacientes que responden.

- 40 Las SEC ID N° 27, 40, 122, 43, 131, 50, 74.127.135, 86, 90.128,129, 99,105, 115, 121, 126, 137, 129, 125, 132, 122, 123, 131, 133, 134, 127, 130, 135, 124, 128 y 136 representan regiones en 5' y/o reguladoras y/o regiones ricas en CpG de los genes según la tabla 1. Estas secuencias se derivan de Genbank y se considerará que incluyen todas las variaciones menores del material de secuencia que actualmente no se prevén, por ejemplo, pero sin limitarse a, SNP y deleciones menores.

- 45 Las SEC ID N° 299, 300, 325, 326, 327, 328, 331, 332, 345, 346, 381, 382, 393, 394, 401, 402, 411, 412, 417, 418, 425, 426, 427, 428, 429, 430, 443, 444, 455, 456, 475, 476, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 05, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 573, 574, 599, 600, 601, 602, 605, 606, 619, 620, 655, 656, 667, 668, 675, 676, 685, 686, 691, 692, 699, 700, 701, 702, 703, 704, 717, 718, 729, 730, 749, 750, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793 y 794 muestran la secuencia de ADN tratada previamente derivada de la secuencia genómica según la tabla 1. Se considerará que estas secuencias incluyen todas las variaciones menores del material de secuencia que actualmente no se prevén, por ejemplo, pero sin limitarse a, SNP y deleciones menores.

- 50 Las SEC ID N° 2003, 2003, 2003, 2003, 2004, 2004, 2004, 2004, 2005, 2005, 2005, 2005, 2006, 2006, 2006, 2006, 2007, 2007, 2007, 2007, 2008, 2008, 2008, 2008, 2009, 2009, 2009, 2009, 2010, 2010, 2010, 2010, 2011, 2011, 2011, 2012, 2012, 2012, 2013, 2013, 2014, 2014, 2015, 2015, 2016, 2016, 2017, 2017, 2017, 2018, 2018, 2018, 2019, 2019, 2019, 2020, 2020, 2020, 2021, 2021, 2022, 2022, 2023, 2023, 2024, 2024, 2025, 2025, 2026, 2026, 2027, 2027, 2028, 2028, 2029, 2029, 2030, 2030 - 2035, 2035, 2036, 2036, 2037, 2037, 2038, 2038 - 2045, 2045, 2046, 2046, 2047, 2047, 2048, 2048, 2049, 2049, 2050, 2050, 2051, 2051, 2052, 2052, 2053, 2053, 2054, 2054, 2055, 2055, 2056, 2056, 2057, 2057, 2058, 2058, 2059, 2059, 2060, 2060 - 2073, 2073, 2074, 2074 - 2077, 2077,

- 2078, 2078 - 2081, 2081, 2082, 2082 - 2095, 2095, 2096, 2096, 2097, 2097, 2098, 2098, 2099, 2099, 2100, 2100, 2101, 2101, 2102, 2102, 2103, 2103, 2104, 2104 - 2109, 2109, 2110, 2110, 2111, 2111, 2112, 2112, 2113, 2113, 2114, 2114 - 2123, 2123, 2124, 2124, 2125, 2125, 2126 y 2126 - 2142 muestran la secuencia de oligómeros que son útiles para el análisis de posiciones de CpG dentro del ADN genómico según la SEC ID N° 1 a SEC ID N° 63 según las tablas 2 y 3.
- 5 La SEC ID N° 2001 a SEC ID N° 2140 muestran la secuencia de oligómeros que son particularmente útiles para el análisis del estado de metilación de posiciones de CpG de ADN genómico según las SEC ID N° 27, 40, 122, 43, 131, 50, 74, 127, 135, 86, 90, 128, 129, 99, 105, 115, 121, 126, 137, 129, 125, 132, 122, 123, 131, 133, 134, 127, 130, 135, 124, 128 y 136.
- 10 Las SEC ID N° 54, SEC ID N° 56 y SEC ID N° 133 representan secuencias genómicas del gen TMEFF2 y las SEC ID N° 349, 350, 353, 354, 511, 623, 624, 627, 628, 785 y 786 representan la secuencia tratada previamente de las mismas.
- Las SEC ID N° 56, 70 y 136 representan secuencias genómicas del gen ESR1 y las SEC ID N° 357, 358, 385, 386, 515, 514, 631, 632, 659, 660, 787 y 788 representan la secuencia tratada previamente de las mismas.
- 15 Las SEC ID N° 135 y SEC ID N° 83 representan la secuencia genómica del gen PITX2 y las SEC ID N° 411, 412, 515, 516, 685, 686, 789 y 790 representan la secuencia tratada previamente de las mismas.
- Las SEC ID N° 119 y SEC ID N° 136 representan secuencias genómicas del gen PCAF y las SEC ID N° 483, 484, 517, 518, 757, 758, 791, 792 representan la secuencia tratada previamente de las mismas.
- 20 La SEC ID N° 137 representa la secuencia genómica del gen WBP11 y las SEC ID N° 519, 520, 793 y 794 representan la secuencia tratada previamente de la misma.
- La SEC ID N° 27 representa la secuencia genómica del gen STMN1 y las SEC ID N° 299, 300, 573 y 574 representan la secuencia tratada previamente de la misma.
- 25 Las SEC ID N° 90 y SEC ID N° 124 representan la secuencia genómica del gen PSA conocido también como PSAT1 y las SEC ID N° 425, 426, 699, 700, 493, 494, 767 y 768 representan la secuencia tratada previamente de las mismas.
- Las SEC ID N° 91 y SEC ID N° 128 representan la secuencia genómica del gen CGA y las SEC ID N° 427, 428, 501, 502, 701, 702, 775 y 776 representan la secuencia tratada previamente de las mismas.
- La SEC ID N° 74 representa la secuencia genómica del gen FGFR1 y las SEC ID N° 393, 394, 667 y 668 representan la secuencia tratada previamente de la misma.
- 30 La SEC ID N° 50 representa la secuencia genómica del gen PTGS2 y las SEC ID N° 345, 346, 619 y 620 representan la secuencia tratada previamente de la misma.
- La SEC ID N° 90 representa la secuencia genómica del gen MSMB y las SEC ID N° 443, 444, 717 y 718 representan la secuencia tratada previamente de la misma.
- 35 Las SEC ID N° 68 y SEC ID N° 131 representan la secuencia genómica del gen TP53 y las SEC ID N° 381, 382, 507, 508, 655, 656, 781 y 782 representan la secuencia tratada previamente de las mismas.
- La SEC ID N° 92 representa la secuencia genómica del gen CYP2D6 y las SEC ID N° 429, 430, 703 y 704 representan la secuencia tratada previamente de la misma.
- Las SEC ID N° 41 y SEC ID N° 122 representan la secuencia genómica del gen S100A2 y las SEC ID N° 327, 328, 489, 490, 601, 601, 763 y 764 representan la secuencia tratada previamente de las mismas.
- 40 La SEC ID N° 121 representa la secuencia genómica del gen PRKCD y las SEC ID N° 487, 488, 761 y 762 representan la secuencia tratada previamente de la misma.
- Las SEC ID N° 78 y SEC ID N° 127 representan la secuencia genómica del gen SYK y las SEC ID N° 401, 402, 499, 500, 675, 676, 773 y 774 representan la secuencia tratada previamente de las mismas.
- 45 La SEC ID N° 115 representa la secuencia genómica del gen VTN y las SEC ID N° 475, 476, 749, 750 representan la secuencia tratada previamente de la misma.
- La SEC ID N° 86 representa la secuencia genómica del gen GRIN2D y las SEC ID N° 417, 418, 691 y 692 representan la secuencia tratada previamente de la misma.
- La SEC ID N° 43 representa la secuencia genómica del gen TGFB2 y las SEC ID N° 331, 332, 605 y 606 representan la secuencia tratada previamente de la misma.

La SEC ID N° 105 representa la secuencia genómica del gen COX7A2L y las SEC ID N° 455, 456, 729 y 730 representan la secuencia tratada previamente de la misma.

Las SEC ID N° 42 y SEC ID N° 123 representan la secuencia genómica del gen TFF1 y las SEC ID N° 329, 330, 491, 492, 603, 604, 765 y 766 representan la secuencia tratada previamente de las mismas.

- 5 La SEC ID N° 125 representa la secuencia genómica del gen DAG1 y las SEC ID N° 495, 496, 769 y 770 representan la secuencia tratada previamente de la misma.

La SEC ID N° 126 representa la secuencia genómica del gen oncut y las SEC ID N° 497, 498, 771 y 772 representan la secuencia tratada previamente de la misma.

- 10 La SEC ID N° 129 representa la secuencia genómica del gen CYP2D6 y las SEC ID N° 503, 504, 777 y 778 representan la secuencia tratada previamente de la misma.

Las SEC ID N° 59 y SEC ID N° 130 representan la secuencia genómica del gen RASSF1 y las SEC ID N° 363, 364, 505, 506, 637, 638, 779, 780 representan la secuencia tratada previamente de las mismas.

La SEC ID N° 132 representa la secuencia genómica del gen ERBB2 y las SEC ID N° 509, 510, 783 y 784 representan la secuencia tratada previamente de las mismas.

15 Ejemplo 1

Se extrajeron muestras de ADN usando el kit Wizzard (Promega), se analizaron muestras de 200 pacientes, se llevaron a cabo cuatro análisis de datos en una selección de marcadores candidatos.

Tratamiento con bisulfito y mPCR

- 20 Se trató con bisulfito el ADN genómico total de todas las muestras convirtiendo las citosinas no metiladas en uracilo. Las citosinas metiladas permanecieron conservadas. Se realizó el tratamiento con bisulfito con modificaciones menores según el protocolo descrito en Olek *et al.* (1996). Tras la bisulfitación, se usaron 10 ng de cada muestra de ADN en reacciones de mPCR posteriores que contenían 6-8 pares de cebadores.

Cada reacción contenía lo siguiente:

- 25 dNTP 400 μ M
2 pmoles de cada cebador
1 U de Hot Star Taq (Qiagen)
10 ng de ADN (tratado con bisulfito)

Se muestran detalles adicionales de los cebadores en la TABLA 1.

- 30 Se llevaron a cabo cuarenta ciclos tal como sigue: desnaturalización a 95°C durante 15 min., seguido de reasociación a 55°C durante 45 s, elongación del cebador a 65°C durante 2 min. Se llevó a cabo una elongación final a 65°C durante 10 min.

Hibridación

- 35 Se hibridaron entonces todos los productos de PCR de cada muestra individual en portaobjetos de vidrio que portaban un par de oligonucleótidos inmovilizados para cada posición de CpG bajo análisis. Se diseñó cada uno de estos oligonucleótidos de detección para que hibridase con la secuencia convertida mediante bisulfito alrededor de un sitio de CpG que originalmente estaba o bien no metilado (TG) o bien metilado (CG). Véase la tabla 2 para detalles adicionales de los oligonucleótidos de hibridación usados. Se seleccionaron las condiciones de hibridación para permitir la detección de las diferencias de un único nucleótido entre las variantes TG y CG.

- 40 Se diluyó un volumen de 5 μ l de cada producto de PCR multiplex en 10 x tampón Ssarc (10 x Ssarc: 230 ml de 20 x SSC, 180 ml de disolución lauroilsarcosinato de sodio al 20%, se diluye hasta 1000 ml con dH₂O). Se hibridó entonces la mezcla de reacción con los oligonucleótidos de detección tal como sigue. Desnaturalización a 95°C, enfriamiento hasta 10°C, hibridación a 42°C durante la noche seguido de lavado con 10 x Ssarc y dH₂O a 42°C. Se muestran detalles adicionales de los oligonucleótidos de hibridación en la TABLA 2.

- 45 Se detectaron señales fluorescentes de cada oligonucleótido hibridado usando software y escáner genepix. Se calcularon las razones para las dos señales (del oligonucleótido CG y el oligonucleótido TG usados para analizar cada posición de CpG) basándose en la comparación de la intensidad de las señales fluorescentes.

Procedimientos de análisis de datos

Análisis de los datos del chip: desde las intensidades de hibridación sin procesar hasta las razones de metilación; se determina la razón de metilación logarítmica ($\log(CG/TG)$) en cada posición de CpG según un canal de procesamiento normalizado que incluye las siguientes etapas: para cada punto se resta la intensidad de píxel de fondo media de la intensidad de píxel de primer plano media (esto proporciona una buena estimación de las intensidades de hibridación corregidas para el fondo); para los oligonucleótidos de detección tanto CG como TG de cada posición de CpG, se toma la mediana corregida para el fondo de las intensidades de 4 puntos redundantes; para cada chip y cada posición de CpG se calcula la razón de $\log(CG/TG)$; para cada muestra se toma la mediana de las intensidades de $\log(CG/TG)$ con respecto a las repeticiones de chip redundantes. Esta razón tiene la propiedad de que el ruido de hibridación tiene una varianza aproximadamente constante a lo largo del intervalo completo de posibles tasas de metilación (Huber *et al.*, 2002).

Análisis de componentes principales

El análisis de componentes principales (ACP) proyecta vectores de medición (por ejemplo, datos del chip, perfiles de metilación en varios CpG etc.) sobre un nuevo sistema de coordenadas. Los nuevos ejes de coordenadas se denominan componentes principales. La primera componente principal abarca la dirección de la varianza más grande de los datos. Las componentes posteriores se ordenan por varianza decreciente y son ortogonales entre sí. Posiciones de CpG diferentes contribuyen con diferentes pesos a la extensión de la nube de datos a lo largo de diferentes componentes. El ACP es una técnica no supervisada, es decir, no tiene en cuenta las etiquetas de los puntos de datos (para detalles adicionales véase por ejemplo Ripley (1996)).

El ACP se usa normalmente para proyectar datos con muchas dimensiones (en este caso datos de series de metilación) sobre subespacios con menos dimensiones con el fin de visualizar o extraer características con alta varianza a partir de los datos. En el presente informe se usan proyecciones en 2 dimensiones para el control de calidad estadístico de los datos. Se investiga el efecto de diferentes parámetros de procesamiento sobre los datos del chip y se excluye que el cambio de los parámetros de procesamiento provoque grandes alteraciones en los valores de medición. Se usa una versión sólida del ACP para detectar chips individuales con valores atípicos y excluirlos del análisis adicional (Model *et al.*, 2002).

Pruebas de hipótesis

La tarea principal es identificar marcadores que muestran diferencias significativas en el grado de metilación promedio entre dos clases. Se detecta una diferencia significativa cuando puede rechazarse la hipótesis nula de que la metilación promedio de las dos clases es idéntica con $p < 0,05$. Dado que se aplica esta prueba a un conjunto completo de marcadores potenciales, hay que corregir los valores de p para pruebas múltiples. Se realizó esto aplicando el procedimiento de la tasa de falsos positivos (TFP) (Dudoit *et al.*, 2002).

Para someter a prueba la hipótesis nula de que los niveles de metilación en las dos clases son idénticos, se usó la prueba de cocientes de probabilidades para modelos de regresión logística (Venables y Ripley, 2002). El modelo de regresión logística para un único marcador es una combinación lineal de mediciones de metilación a partir de todas las posiciones de CpG en la región genómica de interés (ROI) respectiva. Un valor de p significativo para un marcador significa que esta ROI tiene cierta correlación sistemática con la cuestión de interés tal como se facilita mediante las dos clases. Sin embargo, al menos formalmente no realiza ninguna afirmación sobre el poder de predicción real del marcador.

Predicción de la clase mediante aprendizaje supervisado

Con el fin de facilitar una estimación fiable de en qué grado el conjunto de CpG de un marcador seleccionado puede diferenciar entre diferentes clases de tejido, puede determinarse su precisión de predicción mediante clasificación. Para ese fin, se calcula una función de predicción basada en el perfil de metilación usando un determinado conjunto de muestras de tejido con su etiqueta de clase. Esta etapa se denomina entrenamiento y se aprovecha del conocimiento anterior representado por las etiquetas de datos. Se somete a prueba entonces la precisión de predicción de esa función mediante validación cruzada o en un conjunto de muestras independientes. Como procedimiento de elección, se usa el algoritmo de máquinas de vectores soporte (SVM, “*support vector machine*” (Duda (2001), Christiannini (2000)) para aprender la función de predicción. Si no se establece otra cosa, para este informe se fija que el riesgo asociado con clasificaciones de falsos positivos o falsos negativos es igual en relación con los tamaños de clase respectivos. Se entiende que el algoritmo de aprendizaje obtiene una función de predicción de la clase con el objetivo de optimizar la precisión en un conjunto de muestras de prueba independientes. Por tanto, puede esperarse que la sensibilidad y especificidad del clasificador resultante sea aproximadamente igual.

Análisis de datos

Conjuntos de datos 1 y 2 (figuras 5 a 13)

Se clasifican entonces los datos en una matriz de clasificación según las diferencias de metilación de CpG entre las

dos clases de tejidos, usando un algoritmo. Las figuras 5, 7, 9 y 11 a 13 se muestran en escala de grises, estando las posiciones de CpG más significativas en la parte inferior de la matriz, disminuyendo la significación hacia la parte superior. El color negro indica la metilación total en una posición de CpG dada, el color blanco representa ausencia de metilación en la posición particular, con los grados de metilación representados en gris, desde claro (baja proporción de metilación) hasta oscuro (alta proporción de metilación). Cada fila representa una posición de CpG específica dentro de un gen y cada columna muestra el perfil de metilación para los diferentes CpG para una muestra. En el lado izquierdo, se muestra un CpG y un identificador génico; puede hacerse referencia cruzada al mismo con la tabla adjunta (tabla 5) con el fin de determinar el gen en cuestión y el oligómero de detección usados. En el lado derecho, se muestran valores de p para las posiciones de CpG individuales. Los valores de p son las probabilidades de que la distribución observada se produzca por casualidad en el conjunto de datos. Las figuras 6, 8, 10 y 12 son las versiones en rojo-verde originales de las figuras anteriores (es decir, figuras 5, 7, 9 y 11 respectivamente). El color gris oscuro indica la metilación total en una posición de CpG dada, el color gris claro representa ausencia de metilación en la posición particular.

Conjunto de datos 1: Entorno complementario

- 15 Se muestra el análisis de los patrones de metilación de muestras de pacientes tratados con tamoxifeno como tratamiento complementario inmediatamente tras la cirugía (véase la figura 1) en las matrices según las figuras 5 a 8. En este análisis puede observarse que los genes PTGS2, MSMB, TP53 y CYP2D6 estaban metilados de manera significativamente diferente entre las dos clases de tejidos (pacientes que responden al tratamiento y pacientes que no responden al tratamiento).
- 20 En la clasificación mostrada en las figuras 3 y 4, los genes TP53 y MSMB estaban significativamente más metilados en pacientes que no respondían al fármaco tamoxifeno, clasificándose los sujetos con una supervivencia libre de enfermedad inferior a 36 meses como pacientes que no respondían y los sujetos con una supervivencia libre de enfermedad superior a 36 meses como pacientes que respondían. Se llevó a cabo esta clasificación con una sensibilidad de 0,84 y una especificidad de 0,16.
- 25 En la clasificación mostrada en las figuras 7 y 8, el gen PTGS2 estaba significativamente más metilado en pacientes que no responden al fármaco tamoxifeno, clasificándose los sujetos con una supervivencia libre de enfermedad inferior a 24 meses como pacientes que no respondían y los sujetos con una supervivencia libre de enfermedad de 100 meses como pacientes que respondían. Se llevó a cabo esta clasificación con una sensibilidad de 0,89 y una especificidad de 0,48.

Conjunto de datos 1: Entorno metastásico

- Se muestra el análisis de los patrones de metilación de muestras de pacientes tratados con tamoxifeno en un entorno metastásico (véase la figura 2) en las matrices según las figuras 9 a 12. Los sujetos analizados en esta clasificación han experimentado una recidiva tras un tratamiento inicial, tratándose la metástasis posterior mediante tamoxifeno.
- 35 En la clasificación mostrada en las figuras 9 y 10, los genes FGFR1 y PSA estaban significativamente menos metilados en pacientes que no respondían al fármaco tamoxifeno. Los sujetos con una enfermedad progresiva se clasificaron como pacientes que no respondían y los sujetos que lograron una remisión parcial o completa se clasificaron como pacientes que respondían. Se llevó a cabo esta clasificación con una sensibilidad de 0,45 y una especificidad de 0,81.
- 40 La sensibilidad de esta clasificación pudo mejorarse analizando sólo los pacientes que no se habían sometido a quimioterapia. En esta clasificación, mostrada en las figuras 11 y 12, los genes FGFR1, PSA y CGA estaban significativamente menos metilados en pacientes que no respondían al fármaco tamoxifeno. La sensibilidad de la clasificación se mejoró de ese modo hasta 0,54 y la especificidad hasta 0,85.

Conjunto de datos 2: Entorno metastásico

- 45 La figura 12 muestra el análisis del conjunto de muestras total en el que en los pacientes que respondían y en comparación con los pacientes que no respondían, se observaron diferencias significativas en la metilación entre las dos clases de tejido en los genes STMN 1 y PSA. En el caso de que pacientes que habían recibido quimioterapia como terapia adicional se excluyeran del conjunto de datos, se observó un marcador adicional CGA (véase la figura 13).

Conjunto de datos 2: Entorno complementario

Se puso cada CpG en un modelo de riesgo proporcional de Cox junto con los marcadores de predicción conocidos estadio N y tamaño tumoral. El mejor marcador era el gen PITX 2.

El modelo de Cox para el mejor marcador (número de oligonucleótido 3522:2087) es:

	coef.	Exp(coef.)	Ee(coef.)	z	P
Estadio n	1,677	5,35	0,5390	3,11	0,00190
3522:2087	2,885	17,90	0,8211	3,51	0,00044
Tamaño tumoral	0,120	1,13	0,0926	1,30	0,19000

Esto muestra que 3522:2087 facilita información sobre el tiempo de supervivencia independiente de estadio n. El tamaño tumoral no tiene poder de predicción significativo para el tiempo de supervivencia esperado.

5 Conjunto de datos 4: Entorno metastásico

Con el fin de determinar la capacidad de cada promotor génico para predecir el éxito o el fallo del tratamiento con tamoxifeno, se combinaron los CpG individuales medidos por gen usando estadísticos T^2 de Hotelling. Varios genes estaban significativamente asociados con la respuesta a tamoxifeno tras la corrección para comparaciones múltiples con una tasa de falsos positivos conservadora moderada del 25% (véase la figura 3). Los genes eran ONECUT2, WBP11, CYP2D6, DAG1, ERBB2, S100A2, TFF1, TP53, TMEFF2, ESR1, SYK, RASSF1, PITX2, PSAT1, CGA y PCAF.

La figura 15 muestra los valores de p no corregidos en una escala logarítmica. Se calcularon los valores de P a partir de pruebas del coeficiente de probabilidades (CP) a partir de modelos de regresión logística multivariable. Cada región genómica individual de interés está representada como un punto, la línea de puntos superior representa el punto de corte para la tasa de falsos positivos del 25%, la línea de puntos inferior muestra el límite del 5% corregido de Bonferroni.

La figura 16 muestra una matriz de clasificación de los mejores 11 amplificadores de datos obtenidos según el ejemplo 1 (conjunto de datos 4). Se calcularon los valores de P a partir de pruebas del coeficiente de probabilidades (CP) a partir de modelos de regresión logística multivariable. Se muestra la figura en escala de grises, estando las posiciones de CpG más significativas en la parte inferior de la matriz disminuyendo la significación hacia la parte superior. El color negro indica la metilación total en una posición de CpG dada, el color blanco representa ausencia de metilación en la posición particular, con los grados de metilación representados en gris, desde claro (baja proporción de metilación) hasta oscuro (alta proporción de metilación). Cada fila representa una posición de CpG específica dentro de un gen y cada columna muestra el perfil de metilación para los diferentes CpG de una muestra. Se muestran los valores de p para las posiciones de CpG individuales en el lado derecho. Los valores de p son las probabilidades de que la distribución observada se produzca por casualidad en el conjunto de datos.

Conjunto de datos 4: Entorno complementario

Para cada amplificador, se calculó la metilación media a lo largo de todos los pares de oligonucleótidos para ese amplificador y se dividió la población en grupos de igual tamaño según sus valores de metilación medios. Se muestran los resultados en las figuras 17 a 21, como curvas de supervivencia libre de enfermedad estimadas de Kaplan-Meier. Los valores de p eran:

ESR1:	p = 0,0371
PCAF:	p = 0,0105
PITX2:	p = 3e-04
TMEFF2:	p = 0,0106
WBP 11:	p = 0,0366
ERBB2:	p = 0,018

Conjunto de datos 3

Con el fin de determinar la capacidad de cada promotor génico para predecir el éxito o el fallo del tratamiento con tamoxifeno, se combinaron los CpG individuales medidos por gen usando estadísticos T^2 de Hotelling. Para identificar marcadores potenciales, los inventores realizaron el análisis inicial en los 123 pacientes que mostraron los

tipos extremos de respuesta; o bien una respuesta objetivo (CR+PR, 45 pacientes) o bien una enfermedad progresiva justo a partir del comienzo del tratamiento (PD, 78 pacientes). Varios genes estaban significativamente asociados con la respuesta a tamoxifeno tras la corrección para comparaciones múltiples con una tasa de falsos positivos conservadora moderada del 25% (véase la figura 3). Los genes eran STMN1, SFN, S100A2, TGFB2, SYK, GRIN2D, PSA, COX7A2L, VTN y PRKCD.

A partir de las matrices de datos de metilación del ADN, se concluyó que para la mayoría de los genes la hipermetilación de las islas de CpG estaba correlacionada con una respuesta favorable al tratamiento con tamoxifeno, mientras que para estatmina 1/oncoproteína 18 (STMN1) y estratifina/14-3-3 proteína σ (SFN) la hipometilación estaba asociada con la remisión.

Posteriormente, se analizaron los 200 pacientes incluyendo aquéllos con un tipo de respuesta intermedia (59 pacientes con enfermedad estable (SD, sin cambios en la enfermedad durante más de 6 meses) y 18 pacientes con enfermedad estable durante 6 meses o menos) en análisis de regresión logística. Para cada uno de los genes seleccionados, se combinaron por gen las estimaciones de la regresión logística para la respuesta global (= respuesta objetivo + SD) de los diferentes dinucleótidos de CpG medidos. En la cohorte completa, para los diez genes, la tasa de respuesta al tratamiento con tamoxifeno era de aproximadamente el 15% al 20% superior (cocientes de probabilidades (CP) que oscilaban entre 1,5 y 3,5) para pacientes cuyos tumores tenían un estado de metilación del ADN favorable que aquéllos con un estado desfavorable. En concordancia, para los genes marcadores, la PFS media en el primer grupo de pacientes era en promedio 1,8 veces más larga que en el segundo grupo (razones de riesgo (HR) que oscilaban entre 0,54 y 0,77). Por tanto, también en la cohorte completa, estos genes marcadores estaban asociados con el desenlace de la enfermedad tras el tratamiento con tamoxifeno.

La predicción de la respuesta al tratamiento con tamoxifeno en pacientes con cáncer de mama avanzado se basa tradicionalmente en las características del tumor, la enfermedad y el paciente, es decir, la edad, el primer sitio dominante de recidiva, el intervalo libre de enfermedad y el estado de ER. Se comparó una puntuación de predicción calculada con las estimaciones multivariantes de los factores de predicción tradicionales (puntuación basada en factores tradicionales) con una puntuación derivada de los datos de metilación del ADN sólo (puntuación basada en la metilación del ADN). Para una puntuación basada en la metilación del ADN, se incluyeron sólo genes cuyo estado de metilación del ADN predecía una respuesta independiente entre sí. Para establecer esto, se añadió la información del estado de metilación del ADN de cada uno de los genes marcadores simultáneamente a un modelo de regresión logística multivariable. Tras la eliminación hacia atrás de los genes no significativos ($P < 0,05$), los genes STMN1, SFN, S100A2, TGFB2, SYK, GRIN2D, PSA, COX7A2L, VTN y PRKCD eran estadísticamente significativos y pronosticaban una respuesta, independientemente entre sí. La puntuación basada en la metilación del ADN que se comparó posteriormente con la puntuación basada en factores tradicionales se derivó de las puntuaciones multivariantes de estos genes.

Se usó la puntuación para cada tumor basada o bien en factores de predicción tradicionales o bien en el estado de metilación del ADN para dividir los pacientes en 4 grupos por sus valores de percentil 25, 50 y 75 (tabla 3). La puntuación basada en la metilación del ADN mostró una fuerza discriminatoria mucho mejor con la respuesta global que la puntuación basada en factores tradicionales (tabla 3). Las tasas de respuesta global en la puntuación basada en la metilación del ADN aumentó desde el 26% para los pacientes en el primer cuarto (q1) (CP fijado a 1,0), hasta el 34% para el segundo cuarto (q2) (CP=1,5), el 66% en el tercero para q3 (CP=5,5) y el 82% y el último cuarto para q4 (CP=13,0) ($P < 0,001$) mientras que para la puntuación basada en factores tradicionales aumentó sólo desde el 36% (CP, fijado a 1,0) para q1 de esta puntuación, a través del 46% (CP=1,5) (q2) hasta el 64% (CP=3,2) y el 62% (CP=2,9) (q3) (q4), respectivamente ($P = 0,013$) (tabla 3).

De manera similar, la PFS media en los cuartos extremos para la puntuación basada en la metilación del ADN difirió 4,2 veces (desde 3,8 meses hasta 16,1 meses) mientras que para la puntuación basada en factores tradicionales la diferencia entre los dos cuartos extremos era sólo la mitad de la misma (2,1 veces; desde 4,7 meses hasta 9,7 meses) (compárense las curvas de la figura 14A con la figura 14B).

Por tanto, los inventores han demostrado por primera vez que un perfil epigenético basado en el estado de metilación del ADN en islas de CpG de regiones promotoras de sólo cinco genes puede predecir la probabilidad de respuesta al tratamiento en pacientes con cáncer de mama avanzado positivo para ER tratados con tratamiento con tamoxifeno. Además, parece que esta predicción era mucho más precisa que la predicción basada en factores de predicción tradicionales comúnmente usados, lo que sugiere que el análisis de la metilación del ADN puede ser útil para decisiones de tratamiento clínico. La obtención del perfil epigenético de ADN de tejido tumoral puede ser por tanto una alternativa competitiva y sencilla a la tediosa obtención del perfil de ARNm que es altamente dependiente del material de tejido tumoral bien conservado. Además de que los perfiles epigenéticos de ADN son muy estables, se localizan en sitios específicos (la región promotora de un gen) y pueden amplificarse fácilmente mediante PCR. Más allá, puede realizarse la determinación del estado de metilación del ADN en material de tejido tumoral incrustado en parafina congelado y archivado o fácilmente accesible, lo que hace que pueda aplicarse en el entorno clínico de rutina.

Ejemplo 2: Ensayo LightCycler

En el siguiente ejemplo, se verificó el estado de metilación relevante para la respuesta a tamoxifeno del gen PSAT1 mediante el análisis de muestras de 24 pacientes usando un ensayo en tiempo real que comprendía un oligonucleótido bloqueante y sondas LightCycler.

5

Cebador: TAGGTTGGTAGGGGTAG	(SEC ID N° 2143)
Cebador: AAAACTCACATCCCCATTAA	(SEC ID N° 2144)
Sonda: TTTTTTGTATTGATTAAAAATGGGGGTTGAAATAGTA	(SEC ID N° 2145)
Sonda específica de metilación: CGCGAGGAGGAGTAATTGTTTC	(SEC ID N° 2146)
Sonda específica de metilación: TGTGAGGAGGAGTAATTGtTTTGATT	(SEC ID N° 2147)

Se ejecutaron las reacciones por triplicado en cada muestra de ADN con las siguientes condiciones de ensayo:

Disolución de la reacción;

HotStarTaq de Qiagen	5 U
10 x tampón de PCR	1x
MgCl ₂	3 mM
Cebadores	300 nM cada uno
Sondas	250 nM cada una
dNTP	200 nM cada uno
BSA	0,25 mg/ml
ADN molde	10 ng

10 *Perfil de los ciclos:*

Desnaturalización a 95°C durante 15 minutos

55 ciclos: 95°C 10 segundos, 55°C 20 segundos, 57°C 3 segundos, 72°C 20 segundos. Los resultados se muestran en la tabla 4.

15

Tabla 1

N°	Gen:	Cebador:	Longitud amplificado: del
1	ABCB1 (SEC ID N° 1)	TAAGTATGTTGAAGAAAGATTATTGTAG (SEC ID N° 795) TAAAACTATCCCATAATAACTCCCAAC (SEC ID N° 796)	633
2	APOC2 (SEC ID N° 2)	ATGAGTAGAAGAGGTGATAT (SEC ID N° 797) CCCTAAATCCCTTTCTTACC (SEC ID N° 798)	533

Nº	Gen:	Cebador:	Longitud del amplificado:
3	CACNA1G (SEC ID Nº 3)	GGGATTTAAGAGAAATTGAGGTA (SEC NO: 799) AAACCCCAAACATCCTTTAT (SEC ID Nº 800)	707
4	EGR4 (SEC ID Nº 4)	AGGGGGATTGAGTGTTAAGT (SEC ID Nº 802) CCCAAACATAAACACAAAAT (SEC ID Nº 801)	293
5	AR (SEC ID Nº 5)	GTAGTAGTAGTAGTAAGAGA (SEC ID Nº 803) ACCCCCTAAATAATTATCCT (SEC ID Nº 804)	460
6	RB1 (SEC ID Nº 6)	TTTAAGTTTGTGTTTGTGTTTGGT (SEC ID Nº 805) TCCTACTCTAAATCCTCCTCAA (SEC ID Nº 806)	718
7	GP1BB (SEC ID Nº 7)	GGTGATAGGAGAATAATGTTGG (SEC ID Nº 807) TCTCCCACTACAACCAAAC (SEC ID Nº 808)	379
8	WT1 (SEC ID Nº 8)	AAAGGGAAATTAAGTGTTGT (SEC ID Nº 810) TAACTACCCTCAACTTCCC (SEC ID Nº 809)	747
9	HLA-F (SEC ID Nº 9)	TTGTTGTTTTTAGGGGTTTTGG (SEC ID Nº 811) TCCTTCCCATTCTCCAAATATC (SEC ID Nº 812)	946
10	ELK1 (SEC ID Nº 10)	AAGTGTTTTAGTTTTTAATGGGTA (SEC ID Nº 813) CAAACCCAAAACCTACCTAT (SEC ID Nº 814)	966
11	ARHI (SEC ID Nº 12)	GTGAGTTTTTGGGGTGTTTA (SEC ID Nº 815) TCAATCTTACTTTCACACTACATAA (SEC ID Nº 816)	442
12	BCL2 (SEC ID Nº 13)	GTATTTTATGTTAAGGGGGAAA (SEC ID Nº 817) AAAAACCACAATCCTCCC (SEC ID Nº 818)	640
13	BRCA1 (SEC ID Nº 14)	TGGATGGGAATTGTAGTTTT (SEC ID Nº 819) TTAACCACCCAATCTACCC (SEC ID Nº 820)	537
14	CALCA (SEC ID Nº 15)	GTTTTGGAAGTATGAGGGTG (SEC ID Nº 821) CCAAATTCTAAACCAATTTC (SEC ID Nº 822)	614
15	CCND2 (SEC ID Nº 16)	TTTTGGTATGTAGGTTGGATG (SEC ID Nº 823) CCTAACCTCCTTCCTTAACT (SEC ID Nº 824)	426
16	CDH1 (SEC ID Nº 17)	GAGGTTGGGGTTAGAGGAT (SEC ID Nº 825) CAAACTCACAAATACTTTACAATTC (SEC ID Nº 826)	478
17	CDKN1B (SEC ID Nº 18)	GTGGGGAGGTAGTTGAAGA (SEC ID Nº 827) ATACACCCCTAACCCAAAAT (SEC ID Nº 828)	478

Nº	Gen:	Cebador:	Longitud del amplificado:
18	CDKN2A (SEC ID Nº 19)	TTGAAAATTAAGGGTTGAGG (SEC ID Nº 829) CACCCTCTAATAACCAACCA (SEC ID Nº 830)	598
19	CDKN2A (SEC ID Nº 19)	GGGGTTGGTTGGTTATTAGA (SEC ID Nº 831) AACCTCTACCCACCTAAAT (SEC ID Nº 832)	256
20	CDKN2B (SEC ID Nº 20)	GGTTGGTTGAAGGAATAGAAAT (SEC ID Nº 833) CCCACTAAACATACCCTTATTC (SEC ID Nº 834)	708
21	CD44 (SEC ID Nº 21)	GAAAGGAGAGGTTAAAGGTTG (SEC ID Nº 835) AACTCACTTAACCTCAATCCC (SEC ID Nº 836)	696
22	CSPG2 (SEC ID Nº 22)	GGATAGGAGTTGGGATTAAGAT (SEC ID Nº 837) AAATCTTTTTCAACACCAAAAT (SEC ID Nº 838)	414
23	DAPK1 (SEC ID Nº 23)	AACCCTTTCTTCAAATTACAAA (SEC ID Nº 839) TGATTGGGTTTTAGGGAAATA (SEC ID Nº 840)	348
24	GGT1 (SEC ID Nº 24)	GTGAAGGGTGTGAGTTGTTTA (SEC ID Nº 841) CACAATCAATTTCCACAA (SEC ID Nº 842)	562
25	GSTP1 (SEC ID Nº 25)	ATTTGGGAAAGAGGGAAAG (SEC ID Nº 843) TAAAACTCTAAACCCCATCC (SEC ID Nº 844)	300
26	HIC1 (SEC ID Nº 26)	TGGGTTGGAGAAGAAGTTTA (SEC ID Nº 845) TCATATTTCCAAAAACACACC (SEC ID Nº 846)	280
27	STMN1 (SEC ID Nº 27)	GAGTTTGTATTTAAGTTGAGTGGTT (SEC ID Nº 847) AACAAAACAATACCCCTTCTAA (SEC ID Nº 848)	334
28	STK11 (SEC ID Nº 28)	TAAAAGAAGGATTTTTGATTGG (SEC ID Nº 850) CATCTTATTTACCTCCCTCCC (SEC ID Nº 849)	528
29	ING4 (SEC ID Nº 29)	ATTAGGGATGAGAGGATTTGTA (SEC ID Nº 851) TCTTCCTAACCATACACACTAACC (SEC ID Nº 852)	212
30	MGMT (SEC ID Nº 30)	AAGGTTTTAGGGAAGAGTGTTT (SEC ID Nº 853) ACCTTTTCTATCACAAAAATAA (SEC ID Nº 854)	636
31	MLH1 (SEC ID Nº 31)	TAAGGGGAGAGGAGGAGTTT (SEC ID Nº 855) ACCAATTCTCAATCATCTCTTT (SEC ID Nº 856)	545

Nº	Gen:	Cebador:	Longitud del amplificado:
32	MYC (SEC ID Nº 33)	AGAGGGAGTAAAAGAAAATGGT (SEC ID Nº 857) CCAAATAAACAAAATAACCTCC (SEC ID Nº 858)	712
33	N33 (SEC ID Nº 35)	TTTAGATTGAGGTTTTAGGGT (SEC ID Nº 859) ATCCATTCTACCTCCTTTTTCT (SEC ID Nº 860)	497
34	PAX6 (SEC ID Nº 36)	GGAGGGGAGAGGGTTATG (SEC ID Nº 861) TACTATACACACCCCAAAACAA (SEC ID Nº 862)	374
35	X51730 PGR (SEC ID Nº 67)	TTTGGGAATGGGTTGTAT (SEC ID Nº 921) CTACCCTTAACCTCCATCCTA (SEC ID Nº 922)	369
36	PTEN (SEC ID Nº 38)	TTTAGGTAGTTATATTGGGTATGTT (SEC ID Nº 865) TCAACTCTCAAACCTCCATCA (SEC ID Nº 866)	346
37	RARB (SEC ID Nº 39)	TTGTTGGGAGTTTTTAAGTTTT (SEC ID Nº 867) CAAATTCTCCTTCCAAATAAAT (SEC ID Nº 868)	353
38	SFN (SEC ID Nº 40)	GAAGAGAGGAGAGGGAGGTA (SEC ID Nº 870) CTATCCAACAAACCCAACA (SEC ID Nº 869)	489
39	S100A2 (SEC ID Nº 41)	GTTTTTAAGTTGGAGAAGAGGA (SEC ID Nº 1031) ACCTATAAATCACAACCCACTC (SEC ID Nº 1032)	460
40	TGFBR2 (SEC ID Nº 43)	GTAATTTGAAGAAAGTTGAGGG (SEC ID Nº 873) CCAACAATAAACAAAACCTCT (SEC ID Nº 874)	296
41	TIMP3 (SEC ID Nº 44)	TGAGAAAATTGTTGTTTGAAGT (SEC ID Nº 875) CAAAATACCCTAAAAACCACTC (SEC ID Nº 876)	306
42	VHL (SEC ID Nº 45)	TGTAAATGAATAAAGTTAATGAGTG (SEC ID Nº 877) TCCTAAATTCAAATAATCCTCCT (SEC ID Nº 878)	362
43	CDKN1C (SEC ID Nº 46)	GGGGAGGTAGATATTTGGATAA (SEC ID Nº 879) AACTACACCATTTATATTCCCAC (SEC ID Nº 880)	300
44	CAV1 (SEC ID Nº 47)	GTTAGTATGTTTGGGGGTAAAT (SEC ID Nº 882) ATAAATAACACCTTCCACCCTA (SEC ID Nº 881)	435
45	CDH13 (SEC ID Nº 48)	TTTGATTAGGTTGGAAGTGGT (SEC ID Nº 883) CCCAAATAAATCAACAACAACA (SEC ID Nº 884)	286

Nº	Gen:	Cebador:	Longitud del amplificado:
46	NDRG1 (SEC ID Nº 49)	GGTTTTGGGTTTAGTGGTAAAT (SEC ID Nº 886) AACTTTCATAACTCACCTTTC (SEC ID Nº 885)	416
47	PTGS2 (SEC ID Nº 50)	GATTTTTGGAGAGGAAGTTAAG (SEC ID Nº 888) AAAACTAAAAACCAAACCCATA (SEC ID Nº 887)	381
48	THBS 1 (SEC ID Nº 51)	TGGGGTTAGTTTAGGATAGG (SEC ID Nº 889) CTTAAAAACACTAAAACTTCTCAA (SEC ID Nº 890)	398
49	TMEFF2 (SEC ID Nº 52)	TTGTTTGGGTTAATAAATGGA (SEC ID Nº 891) CTTCTCTCTTCTCCCCTCTC (SEC ID Nº 892)	295
50	PLAU (SEC ID Nº 53)	TATTATAGGAGGATTGAGGAGG (SEC ID Nº 893) CCCATAAAATCATACCACTTCT (SEC ID Nº 894)	499
51	TMEFF2 (SEC ID Nº 54)	TGTTGGTTGTTGTTGTTGTT (SEC ID Nº 1037) CTTTCTACCCATCCCAAAA (SEC ID Nº 1038)	319
52	DNMT1 (SEC ID Nº 55)	TCCCCATCACACCTAAAA (SEC ID Nº 897) GGGAGGAGGGGATGTATT (SEC ID Nº 898)	210
53	ESR1 (SEC ID Nº 70)	AGGGGGAATTAATAGAAAGAG (SEC ID Nº 927) CAATAAAACCATCCCAAATACT (SEC ID Nº 928)	662
54	APAF 1 (SEC ID Nº 57)	AGATATGTTTGGAGATTTTAGGA (SEC ID Nº 901) AACTCCCCACCTCTAATTCTAT (SEC ID Nº 902)	674
55	HOXA5 (SEC ID Nº 58)	AAACCCCAAACAACCTCTAT (SEC ID Nº 904) GAAGGGGGAAAGTTATTAGTTA (SEC ID Nº 903)	392
56	RASSF1 (SEC ID Nº 59)	ACCTCTCTACAAATTACAAATTCA (SEC ID Nº 905) AGTTTGGGTTAGTTTGGGTT (SEC ID Nº 906)	347
57	BRCA2 (SEC ID Nº 60)	ACCCACCCAAACCTAACT (SEC ID Nº 908) GGTTGGTAGAGATAAAAGGGTA (SEC ID Nº 907)	388
58	CCND1 (SEC ID Nº 61)	GATTATAGGGGAGTTTTGTTGA (SEC ID Nº 909) CACCTCCAACATCCAAATA (SEC ID Nº 910)	404
59	EDNRB (SEC ID Nº 62)	TCAAAACATCCTCTATCTCTCC (SEC ID Nº 912) ATAATTGGGGGTTGTATGTATT (SEC ID Nº 911)	484

ES 2 447 566 T3

Nº	Gen:	Cebador:	Longitud del amplificado:
60	EGFR (SEC ID Nº 63)	GGGTTTGGTTGTAATATGGATT (SEC ID Nº 913) CCCAACACTACCCCTCTAA (SEC ID Nº 914)	732
61	ERBB2 (SEC ID Nº 64)	GAGGTAGAGGTTGTGGTGAGT (SEC ID Nº 915) TCCCAACTTCACTTTCTCC (SEC ID Nº 916)	528
62	FOS (SEC ID Nº 65)	ATCCTCCACTTTCTACTTCCA (SEC ID Nº 918) TTTTAGGGTTATAGGGAAAGGT (SEC ID Nº 917)	500
63	NFKB1 (SEC ID Nº 66)	TTTGTAGAATGAAAAGTAGAGTGTG (SEC ID Nº 919) ACCTTAAAAACCCCAACAAT (SEC ID Nº 920)	485
64	X51730 PGR (SEC ID Nº 67)	TTTTGGGAATGGGTTGTAT (SEC ID Nº 921) CTACCCTTAACCTCCATCCTA (SEC ID Nº 922)	369
65	TP53 (SEC ID Nº 68)	GAGTAGGTAGTTGTTGGGTTTC (SEC ID Nº 923) ACCCCTAATTTAACACTTCTCA (SEC ID Nº 924)	702
66	TP73 (SEC ID Nº 69)	AGTAAATAGTGGGTGAGTTATGAA (SEC ID Nº 925) GAAAAACCTCTAAAACTACTCTCC (SEC ID Nº 926)	607
67	ESR1 (SEC ID Nº 70)	AGGGGGAATTAAATAGAAAGAG (SEC ID Nº 927) CAATAAAACCATCCCAAATACT (SEC ID Nº 928)	662
68	SERPINE1 (SEC ID Nº 71)	GGGGTATAGAGAGAGTTTGGAT (SEC ID Nº 929) ACAATTAAACAAACCCCAATAA (SEC ID Nº 930)	492
69	CALM1 (SEC ID Nº 72)	AAAACTCTAACCCCTTCTCAA (SEC ID Nº 932) TATTTTGTGTTTGGGGTGTGT (SEC ID Nº 931)	414
70	CSNK2B (SEC ID Nº 73)	TACCCCTCACCATTACTCTAAC (SEC ID Nº 934) TAGTTTGTGTTTATTGGGTGA (SEC ID Nº 933)	437
71	FGFR1 (SEC ID Nº 74)	AGGGAGTTAGTGGTGTGGTAT (SEC ID Nº 935) CCTTACCCTTCTCAAATCTAA (SEC ID Nº 936)	367
72	MKI67 (SEC ID Nº 75)	CCAATACTCTACAACCATCAAA (SEC ID Nº 938) GGGAAGTTGAAGTAGGAAGAT (SEC ID Nº 937)	499
73	NPM1 (SEC ID Nº 76)	AAGGAAGGAGGAAGTAATTTGT (SEC ID Nº 939) TTACACCAACCCCTAACTAAC (SEC ID Nº 940)	454

ES 2 447 566 T3

Nº	Gen:	Cebador:	Longitud del amplificado:
74	MAPK1 (SEC ID Nº 77)	TTTAGATAATTTTAGGATGGGG (SEC ID Nº 941) TTCTCATTACAAAAACAAAA (SEC ID Nº 942)	743
75	SYK (SEC ID 78)	NO: GTGGGTTTTGGGTAGTTATAGA (SEC ID Nº 1041) TAACCTCCTCTCCTTACCAA (SEC ID Nº 1042)	485
76	TK2 (SEC ID Nº 79)	ATACAACCTCAAATCCTATCCA (SEC ID Nº 946) AGGGAGAAGGAAGTTATTTGTT (SEC ID Nº 945)	485
77	HSPB1 (SEC ID Nº 80)	CCTACCTCTACCACTTCTCAAT (SEC ID Nº 948) AAGAGGGTTTAGTTTTTATTGG (SEC ID Nº 947)	216
78	TES (SEC ID Nº 81)	AGGTTGGGGATTTTAGTTTTT (SEC ID Nº 949) ACCTTCTTCACTTTATTTTCCA (SEC ID Nº 950)	448
79	SDC4 (SEC ID Nº 82)	CCTAACTACCCTCATTCCTTT (SEC ID Nº 952) AGTTGGGGAAATTAAGGTTTAG (SEC ID Nº 951)	269
80	PITX2 (SEC ID Nº 83)	TCCTCAACTCTACAAACCTAAAA (SEC ID Nº 1056) GTAGGGGAGGGAAGTAGATGT (SEC ID Nº 1055)	408
81	GPR37 (SEC ID Nº 84)	ACTTATTTTTCTTTTCTCTAAAAAC (SEC ID Nº 956) TATGGTTTGGTGAGGGTATATT (SEC ID Nº 955)	489
82	FGF1 (SEC ID Nº 85)	AGTTGTGTTTAATTGGGAAGAG (SEC ID Nº 957) CTTATCCCATCCACTATACCAT (SEC ID Nº 958)	420
83	GRIN2D (SEC ID Nº 86)	ATAGTTTGTGGTTTGGATTTTT (SEC ID Nº 959) AAAACCTTCCCTAACTCAAT (SEC ID Nº 960)	435
84	CTSB (SEC ID Nº 87)	AAAATTCCATCAAATAACCATAA (SEC ID Nº 962) AAAAAGGAAGGTAGTAGGATTGT (SEC ID Nº 961)	450
85	CTSD (SEC ID Nº 88)	ATACAACCCTCCAACCTTCTAC (SEC ID Nº 964) AAGGGGTTTTTAAGGAAATG (SEC ID Nº 963)	498
86	PLAUR (SEC ID Nº 89)	TGGTTAAAATGGAGGGTTTAAT (SEC ID Nº 965) CCCCAAATTACCTAAATACAAA (SEC ID Nº 966)	348

Nº	Gen:	Cebador:	Longitud del amplificado:
87	PSA (SEC ID Nº 90)	TAAGAGAGAGGAGTTGAGGTTT (SEC ID Nº 967) CCAAAATTAACCCACCTACCTAA (SEC ID Nº 968)	478
88	CGA (SEC ID Nº 91)	TAGTGGTATAAGTTTGAAATGTT (SEC ID Nº 1047) TCCACCTACATCTAAACCCTAA (SEC ID Nº 1048)	364
89	CYP2D6 (SEC ID Nº 92)	TAAGGGTTTGGAGTAGGAAGTA (SEC ID Nº 971) CACATACAACAAAATTACCCAA (SEC ID Nº 972)	403
90	CYP3A4 (SEC ID Nº 93)	TATCACCACCTTCCCATATTTA (SEC ID Nº 974) GTTTGATGAATGGATTGTATGA (SEC ID Nº 973)	484
91	TK1 (SEC ID Nº 94)	ACCTCTACAAACATCTTATTCCA (SEC ID Nº 976) TTGGGGGAGTTAGGTAGTATAG (SEC ID Nº 975)	487
92	RENBP (SEC ID Nº 95)	TTTGGTAGGGTTAAGGTTTTTA (SEC ID Nº 977) CTTACTCATCCCTCCTACTCC (SEC ID Nº 978)	350
93	F12 (SEC ID Nº 96)	TAGGTTTAGGAGGGTAGTTTGA (SEC ID Nº 979) CTCTCACAACCCAAAAATACA (SEC ID Nº 980)	450
94	REN (SEC ID Nº 97)	ACCTACTCCAAAAATCACAAAA (SEC ID Nº 982) TATGTGGAAAAGTTAGGGTGTT (SEC ID Nº 981)	489
95	EBAG9 (SEC ID Nº 98)	CCAAAACCTATTAACCTCCCA (SEC ID Nº 984) AATGTTTTAGAGGTTAGGGTTG (SEC ID Nº 983)	463
96	MSMB (SEC ID Nº 99)	GTTTTGTAGGATGGTTTGATTT (SEC ID Nº 985) TATATTTACCTTATCCCCACCC (SEC ID Nº 986)	324
97	Exón 1 y región en 5' del gen angiotensinógeno X15323 (SEC ID Nº 100)	AAACTCTCCCCTACCCTCTAC (SEC ID Nº 988) GATGGAGTTGTTTTAGGTTGT (SEC ID Nº 987)	374
98	ZNF147 (SEC ID Nº 101)	TTTGTGTAATAAGATGTGGGA (SEC ID Nº 989) TAAACCCCTACAAAACCTACCAA (SEC ID Nº 990)	484
99	EBBP (SEC ID Nº 102)	GTATTTGTTTTTGGTGAGGGT (SEC ID Nº 991) ATCATCTTCCTAAACATTCCAA (SEC ID Nº 992)	482
100	CALR (SEC ID Nº 103)	TAAATCACAACCATTAACCAAA (SEC ID Nº 994) ATAAGAGGGGAGGAAGGTTTA (SEC ID Nº 993)	490

Nº	Gen:	Cebador:	Longitud del amplificado:
101	BCAR1 (SEC ID Nº 104)	AATTCTTCCTTCTATCTCCCTC (SEC ID Nº 996) TTTATTTTGGGAAGGTTGTT (SEC ID Nº 995)	499
102	COX7A2L (SEC ID Nº 105)	AATCCTAAAAACCCTAACTTTTAAT (SEC ID Nº 998) GGAGGTGTAAGGAGAATAGAGA (SEC ID Nº 997)	398
103	Gen de calicreína glandular AF174646, región promotora y secuencia parcial (SEC ID Nº 106)	CCAATTCATCATTCAACATCTA (SEC ID Nº 1000) ATTTATTTGGGAGGATAGTGG (SEC ID Nº 999)	325
104	KLK3 (SEC ID Nº 107)	TTGGAGTGTAAGGATTTAGGT (SEC ID Nº 1001) AACCCACATAATAACACAACCTCT (SEC ID Nº 1002)	387
105	AKR1B1 (SEC ID Nº 108)	CAACAAAAACATTCTTCTAAACTC (SEC ID Nº 1004) AGGTATTTAATTTTAGGATGGGT (SEC ID Nº 1003)	446
106	TGM4 (SEC ID Nº 109)	AATCCTAACTTTTAATCACCCA (SEC ID Nº 1006) GAGAGGGTAATGGTTTTAGGTA (SEC ID Nº 1005)	435
107	AR (SEC ID Nº 110)	AATATAGGGAGGTTTAGGGTTT (SEC ID Nº 1007) TAACCATACATTTCTCATCCAA (SEC ID Nº 1008)	424
108	HSPA1A (SEC ID Nº 111)	AACCTTTCAAATTCACAATCA (SEC ID Nº 1010) GGATTTATTGGAGGGGATAG (SEC ID Nº 1009)	495
109	FKBP4 (SEC ID Nº 112)	TTTTTAAGTAGGGAAGGGTTT (SEC ID Nº 1011) TCCTTCTAACTACCTACCCCC (SEC ID Nº 1012)	308
110	ESR2 (SEC ID Nº 113)	AAACCTTCCCAATAACCTCTTA (SEC ID Nº 1014) TAGAGGGGAGTAGTGTTTGAGT (SEC ID Nº 1013)	471
111	IGF1 (SEC ID Nº 114)	TACCCTTCTCCCAAAATAATAA (SEC ID Nº 1016) GTATTAAAGGAATATGGGGGAT (SEC ID Nº 1015)	402
112	VTN (SEC ID Nº 115)	GTTATTTGGGTAAATGTAGGGA (SEC ID Nº 1017) TCTATCCCCTCAAACCTTAAAAA (SEC ID Nº 1018)	492

ES 2 447 566 T3

Nº	Gen:	Cebador:	Longitud del amplificado:
113	CTSL (SEC ID Nº 116)	CTACACCCACCCTTAAATAAAA (SEC ID Nº 1020) TTAGTGGAATTTGGAGGAAGTAG (SEC ID Nº 1019)	328
114	TGFB3 (SEC ID Nº 117)	CCTACTAAAAATCAAAACCCAA (SEC ID Nº 1022) AAGGTGGTGTAAGTGGATAGAG (SEC ID Nº 1021)	369
115	MAPKAPK5 (SEC ID Nº 118)	AAACCTACCTCCCCAACTAA (SEC ID Nº 1024) ATTTTGGTTTTAGGGTTGTAA (SEC ID Nº 1023)	495
116	PCAF (SEC ID Nº 119)	GGATAAATGATTGAGAGGTTGT (SEC ID Nº 1057) CCTCCCTTAATTCTCCTACC (SEC ID Nº 1058)	369
117	NCOA3 (SEC ID Nº 120)	AAGGGGGTGTGTTGTTAGATT (SEC ID Nº 1027) CCTAACCCCTACCCTTAATTTTT (SEC ID Nº 1028)	330
118	PRKCD (SEC ID Nº 121)	CTTAACCCATCCCAATCA (SEC ID Nº 1030) GATAGAAGGATTTTAGTTTTATTGTT (SEC ID Nº 1029)	322
119	S100A2 (SEC ID Nº 41)	GTTTTTAAGTTGGAGAAGAGGA (SEC ID Nº 1031) ACCTATAAATCACAACCCACTC (SEC ID Nº 1032)	460
120	PSA (SEC ID Nº 90)	GTAGGTGGTTAATTTTGGGTT (SEC ID Nº 1033) CTCATTCACACTATATCCATTCA (SEC ID Nº 1034)	500
121	ESR1 (SEC ID Nº 70)	CTATCAATTCCCCCAACTACT (SEC ID Nº 1036) TTGTTGGATAGAGGTTGAGTTT (SEC ID Nº 1035)	349
122	TMEFF2 (SEC ID Nº 54)	TGTTGGTTGTTGTTGTTGTT (SEC ID Nº 1037) CTTTCTACCCATCCCCAAA (SEC ID Nº 1038)	319
123	TP53 (SEC ID Nº 68)	TTGATGAGAAGAAAGGATTTAGT (SEC ID Nº 1039) TCAAATTCAATCAAAAACTTACC (SEC ID Nº 1040)	496
124	SYK (SEC ID Nº 78)	GTGGGTTTTGGGTAGTTATAGA (SEC ID Nº 1041) TAACCTCCTCTCCTTACCAA (SEC ID Nº 1042)	485

Nº	Gen:	Cebador:	Longitud del amplificado:
125	DAG1 (SEC ID Nº 125)	AATACCAACCCAAACATCTACC (SEC ID Nº 1044) TTTGGTTATGTGGAGTTTATTGT (SEC ID Nº 1043)	315
126	ONECUT2 (SEC ID Nº 126)	TTTGTTGGGATTTGTTAGGAT (SEC ID Nº 1045) AAACATTTTACCCCTCTAAACC (SEC ID Nº 1046)	467
127	CGA (SEC ID Nº 91)	TAGTGGTATAAGTTTGGAAATGTT (SEC ID Nº 1047) TCCACCTACATCTAAACCCTAA (SEC ID Nº 1048)	364
128	CYP2D6 (SEC ID Nº 129)	AATTCCTAACCCACTATCCTC (SEC ID Nº 1050) ATTTGTAGTTTGGGGTGATTT (SEC ID Nº 1049)	379
129	RASSF1 (SEC ID Nº 59)	AGTGGGTAGGTTAAGTGTGTTG (SEC ID Nº 1051) CCCCCAAATCCAACTAAA (SEC ID Nº 1052)	319
130	ERBB2 (SEC ID Nº 64)	GGAGGGGGTAGAGTTATTAGTT (SEC ID Nº 1053) TATACTTCCTCAAACAACCTC (SEC ID Nº 1054)	258
131	PITX2 (SEC ID Nº 83)	TCCTCAACTCTACAAACCTAAAA (SEC ID Nº 1056) GTAGGGGAGGGAAGTAGATGT (SEC ID Nº 1055)	408
132	PCAF (SEC ID Nº 119)	GGATAAATGATTGAGAGGTTGT (SEC ID Nº 1057) CCTCCCTTAATTCTCCTACC (SEC ID Nº 1058)	369
133	WBP11 (SEC ID Nº 137)	AAGAGGTGAGGAAGAGTAGTAAAT (SEC ID Nº 1059) CTCCCAACAATAATCAAAAT (SEC ID Nº 1060)	437

Tabla 2

Nº:	Gen	Oligonucleótido:
1	STMN1 (SEC ID Nº 27)	GTTATCGGTTTCGGAATT (SEC ID Nº 2001)
2	STMN1 (SEC ID Nº 27)	GTTATTGGTTTGGGAATT (SEC ID Nº 2002)
3	STMN1 (SEC ID Nº 27)	GTTATCGGTTTCGGAATT (SEC ID Nº 2001)
4	STMN1 (SEC ID Nº 27)	GTTATTGGTTTGGGAATT (SEC ID Nº 2002)
5	STMN1 (SEC ID Nº 27)	GTAAGAACGTATATAGTGAG (SEC ID Nº 2003)
6	STMN1 (SEC ID Nº 27)	GTAAGAATGTATATAGTGAG (SEC ID Nº 2004)

Nº:	Gen	Oligonucleótido:
7	STMN1 (SEC ID Nº 27)	GTAAGAACGTATATAGTGAG (SEC ID Nº 2003)
8	STMN1 (SEC ID Nº 27)	GTAAGAATGTATATAGTGAG (SEC ID Nº 2004)
9	STMN1 (SEC ID Nº 27)	AGAAATTACGATGATGTTAT (SEC ID Nº 2005)
10	STMN1 (SEC ID Nº 27)	AGAAATTATGATGATGTTAT (SEC ID Nº 2006)
11	STMN1 (SEC ID Nº 27)	AGAAATTACGATGATGTTAT (SEC ID Nº 2005)
12	STMN1 (SEC ID Nº 27)	AGAAATTATGATGATGTTAT (SEC ID Nº 2006)
13	STMN1 (SEC ID Nº 27)	ATGGGTGATACGTCGGTG (SEC ID Nº 2007)
14	STMN1 (SEC ID Nº 27)	ATGGGTGATATGTTGGTG (SEC ID Nº 2008)
15	STMN1 (SEC ID Nº 27)	ATGGGTGATACGTCGGTG (SEC ID Nº 2007)
16	STMN1 (SEC ID Nº 27)	ATGGGTGATATGTTGGTG (SEC ID Nº 2008)
17	PSA (SEC ID Nº 90)	TTTCGATTCGGTTTAGA (SEC ID Nº 2009)
18	PSA (SEC ID Nº 90)	AATTGTTTTGATTTGGTT (SEC ID Nº 2010)
19	PSA (SEC ID Nº 90)	TAATGGGGCGTCGATT (SEC ID Nº 2011)
20	PSA (SEC ID Nº 90)	TTAATGGGGTGTTGATT (SEC ID Nº 2012)
21	PSA (SEC ID Nº 90)	TATCGTAGCGGTTAGG (SEC ID Nº 2013)
22	PSA (SEC ID Nº 90)	TATTGTAGTGGTTAGGAA (SEC ID Nº 2014)
23	PSA (SEC ID Nº 90)	AGGAACGTTAGTCGTT (SEC ID Nº 2015)
24	PSA (SEC ID Nº 90)	TAGGAATGTTAGTTGTTT (SEC ID Nº 2016)
25	PSA (SEC ID Nº 90)	GGTCGTCGTATTATGGA (SEC ID Nº 2017)
26	PSA (SEC ID Nº 90)	TGGTTGTTGTATTATGGA (SEC ID Nº 2018)
27	CGA (SEC ID Nº 91)	TAAATTGACGTTATGGTA (SEC ID Nº 2019)
28	CGA (SEC ID Nº 91)	AAATTGATGTTATGGTAAA (SEC ID Nº 2020)
29	CGA (SEC ID Nº 91)	AATTGACGTTATGGTAAT (SEC ID Nº 2021)
30	CGA (SEC ID Nº 91)	TAAAAATTGATGTTATGGT (SEC ID Nº 2022)
31	PITX2 (SEC ID Nº 83)	AGTCGGGAGAGCGAAA (SEC ID Nº 2023)
32	PITX2 (SEC ID Nº 83)	AGTTGGGAGAGTGAAA (SEC ID Nº 2024)
33	PITX2 (SEC ID Nº 83)	AGTCGGGAGAGCGAAA (SEC ID Nº 2023)
34	PITX2 (SEC ID Nº 83)	AGTTGGGAGAGTGAAA (SEC ID Nº 2024)
35	PITX2 (SEC ID Nº 83)	AAGAGTCGGGAGTCGGA (SEC ID Nº 2025)
36	PITX2 (SEC ID Nº 83)	AAGAGTTGGGAGTTGGA (SEC ID Nº 2026)
37	PITX2 (SEC ID Nº 83)	AAGAGTCGGGAGTCGGA (SEC ID Nº 2025)
38	PITX2 (SEC ID Nº 83)	AAGAGTTGGGAGTTGGA (SEC ID Nº 2026)
39	PITX2 (SEC ID Nº 83)	GGTCGAAGAGTCGGGA (SEC ID Nº 2027)
40	PITX2 (SEC ID Nº 83)	GGTTGAAGAGTTGGGA (SEC ID Nº 2028)

Nº:	Gen	Oligonucleótido:
41	PITX2 (SEC ID Nº 83)	GGTCGAAGAGTCGGGA (SEC ID Nº 2027)
42	PITX2 (SEC ID Nº 83)	GGTTGAAGAGTTGGGA (SEC ID Nº 2028)
43	FGFR1 (SEC ID Nº 74)	GTATTTTCGTTGGTTAAGT (SEC ID Nº 2029)
44	FGFR1 (SEC ID Nº 74)	GTGTATTTTGTGTTAA (SEC ID Nº 2030)
45	FGFR1 (SEC ID Nº 74)	ATGTGAACGAAGTTAAG (SEC ID Nº 2031)
46	FGFR1 (SEC ID Nº 74)	ATGTGAATGAAGTTAAGA (SEC ID Nº 2032)
47	PSA (SEC ID Nº 90)	TTTCGATTCGGTTTAGA (SEC ID Nº 2009)
48	PSA (SEC ID Nº 90)	AATTGTTTTGATTTGGTT (SEC ID Nº 2010)
49	PSA (SEC ID Nº 90)	AGGAACGTTAGTCGTT (SEC ID Nº 2015)
50	PSA (SEC ID Nº 90)	TAGGAATGTTAGTTGTTT (SEC ID Nº 2016)
51	PSA (SEC ID Nº 90)	GGTCGTCGTATTATGGA (SEC ID Nº 2017)
52	PSA (SEC ID Nº 90)	TGGTTGTTGTATTATGGA (SEC ID Nº 2018)
53	CGA (SEC ID Nº 91)	TAAATTGACGTTATGGTA (SEC ID Nº 2019)
54	CGA (SEC ID Nº 91)	AAATTGATGTTATGGTAAA (SEC ID Nº 2020)
55	CGA (SEC ID Nº 91)	AATTGACGTTATGGTAAT (SEC ID Nº 2021)
56	CGA (SEC ID Nº 91)	TAAAAATTGATGTTATGGT (SEC ID Nº 2022)
57	PTGS2 (SEC ID Nº 50)	TTTATCGGGTTTACGTAATT (SEC ID Nº 2033)
58	PTGS2 (SEC ID Nº 50)	TTTATTGGGTTTATGTAATT (SEC ID Nº 2034)
59	PTGS2 (SEC ID Nº 50)	TTTATCGGGTTTACGTAATT (SEC ID Nº 2033)
60	PTGS2 (SEC ID Nº 50)	TTTATTGGGTTTATGTAATT (SEC ID Nº 2034)
61	PTGS2 (SEC ID Nº 50)	GTACGAAAAGGCGGAAAG (SEC ID Nº 2035)
62	PTGS2 (SEC ID Nº 50)	GTATGAAAAGGTGGAAAG (SEC ID Nº 2036)
63	PTGS2 (SEC ID Nº 50)	GTACGAAAAGGCGGAAAG (SEC ID Nº 2035)
64	PTGS2 (SEC ID Nº 50)	GTATGAAAAGGTGGAAAG (SEC ID Nº 2036)
65	MSMB (SEC ID Nº 99)	ATAGGGCGAAGGTTTA (SEC ID Nº 2037)
66	MSMB (SEC ID Nº 99)	ATAGGGTGAAGGTTTAG (SEC ID Nº 2038)
67	TP53 (SEC ID Nº 68)	TTTTTACGACGGTGAT (SEC ID Nº 2039)
68	TP53 (SEC ID Nº 68)	TGTTTTTTATGATGGTGA (SEC ID Nº 2040)
69	CYP2D6 (SEC ID Nº 92)	AAGTAGCGGTAAGGAT (SEC ID Nº 2041)
70	CYP2D6 (SEC ID Nº 92)	GAAGTAGTGGAAGGAT (SEC ID Nº 2042)
71	STMN1 (SEC ID Nº 27)	GTTATCGGTTTCGGGAATT (SEC ID Nº 2001)
72	STMN1 (SEC ID Nº 27)	GTTATTGGTTTGGGAATT (SEC ID Nº 2002)
73	STMN1 (SEC ID Nº 27)	GTTATCGGTTTCGGGAATT (SEC ID Nº 2001)
74	STMN1 (SEC ID Nº 27)	GTTATTGGTTTGGGAATT (SEC ID Nº 2002)

Nº:	Gen	Oligonucleótido:
75	STMN1 (SEC ID Nº 27)	GTAAGAACGTATATAGTGAG (SEC ID Nº 2003)
76	STMN1 (SEC ID Nº 27)	GTAAGAATGTATATAGTGAG (SEC ID Nº 2004)
77	STMN1 (SEC ID Nº 27)	GTAAGAACGTATATAGTGAG (SEC ID Nº 2003)
78	STMN1 (SEC ID Nº 27)	GTAAGAATGTATATAGTGAG (SEC ID Nº 2004)
79	STMN1 (SEC ID Nº 27)	AGAAATTACGATGATGTTAT (SEC ID Nº 2005)
80	STMN1 (SEC ID Nº 27)	AGAAATTATGATGATGTTAT (SEC ID Nº 2006)
81	STMN1 (SEC ID Nº 27)	AGAAATTACGATGATGTTAT (SEC ID Nº 2005)
82	STMN1 (SEC ID Nº 27)	AGAAATTATGATGATGTTAT (SEC ID Nº 2006)
83	STMN1 (SEC ID Nº 27)	ATGGGTGATACGTCGGTG (SEC ID Nº 2007)
84	STMN1 (SEC ID Nº 27)	ATGGGTGATATGTTGGTG (SEC ID Nº 2008)
85	STMN1 (SEC ID Nº 27)	ATGGGTGATACGTCGGTG (SEC ID Nº 2007)
86	STMN1 (SEC ID Nº 27)	ATGGGTGATATGTTGGTG (SEC ID Nº 2008)
87	PSA (SEC ID Nº 90)	TTTCGATTCGGTTTAGA (SEC ID Nº 2009)
88	PSA (SEC ID Nº 90)	AATTGTTTTGATTTGGTT (SEC ID Nº 2010)
89	PSA (SEC ID Nº 90)	TAATGGGGCGTCGATT (SEC ID Nº 2011)
90	PSA (SEC ID Nº 90)	TTAATGGGGTGTTGATT (SEC ID Nº 2012)
91	PSA (SEC ID Nº 90)	TATCGTAGCGGTTAGG (SEC ID Nº 2013)
92	PSA (SEC ID Nº 90)	TATTGTAGTGGTTAGGAA (SEC ID Nº 2014)
93	PSA (SEC ID Nº 90)	AGGAACGTTAGTCGTT (SEC ID Nº 2015)
94	PSA (SEC ID Nº 90)	TAGGAATGTTAGTTGTTT (SEC ID Nº 2016)
95	PSA (SEC ID Nº 90)	GGTCGTCGTATTATGGA (SEC ID Nº 2017)
96	PSA (SEC ID Nº 90)	TGGTTGTTGTATTATGGA (SEC ID Nº 2018)
97	S 100A2 (SEC ID Nº 41)	TTTAATTGCGGTTGTGTG (SEC ID Nº 2043)
98	S 1 00A2 (SEC ID Nº 41)	TTTAATTGTGTTGTGTG (SEC ID Nº 2044)
99	S 100A2 (SEC ID Nº 41)	TTTAATTGCGGTTGTGTG (SEC ID Nº 2043)
100	S100A2 (SEC ID Nº 41)	TTTAATTGTGTTGTGTG (SEC ID Nº 2044)
101	S100A2 (SEC ID Nº 41)	TATATAGGCGTATGTATG (SEC ID Nº 2045)
102	S 100A2 (SEC ID Nº 41)	TATATAGGTGTATGTATG (SEC ID Nº 2046)
103	S100A2 (SEC ID Nº 41)	TATATAGGCGTATGTATG (SEC ID Nº 2045)
104	S100A2 (SEC ID Nº 41)	TATATAGGTGTATGTATG (SEC ID Nº 2046)
105	S 1 00A2 (SEC ID Nº 41)	TATGTATACGAGTATTGGAT (SEC ID Nº 2047)
106	S 1 00A2 (SEC ID Nº 41)	TATGTATATGAGTATTGGAT (SEC ID Nº 2048)
107	S 100A2 (SEC ID Nº 41)	TATGTATACGAGTATTGGAT (SEC ID Nº 2047)
108	S100A2 (SEC ID Nº 41)	TATGTATATGAGTATTGGAT (SEC ID Nº 2048)

Nº:	Gen	Oligonucleótido:
109	S100A2 (SEC ID Nº 41)	AGTTTTAGCGTGTGTTTA (SEC ID Nº 2049)
110	S100A2 (SEC ID Nº 41)	AGTTTTAGTGTGTGTTTA (SEC ID Nº 2050)
111	S100A2 (SEC ID Nº 41)	AGTTTTAGCGTGTGTTTA (SEC ID Nº 2049)
112	S100A2 (SEC ID Nº 41)	AGTTTTAGTGTGTGTTTA (SEC ID Nº 2050)
113	SFN (SEC ID Nº 40)	GAGTAGGTCTGAACGTTAT (SEC ID Nº 2051)
114	SFN (SEC ID Nº 40)	GAGTAGGTTGAATGTTAT (SEC ID Nº 2052)
115	SFN (SEC ID Nº 40)	GAGTAGGTCTGAACGTTAT (SEC ID Nº 2051)
116	SFN (SEC ID Nº 40)	GAGTAGGTTGAATGTTAT (SEC ID Nº 2052)
117	SFN (SEC ID Nº 40)	TTTGCGAAGAGCGAAATT (SEC ID Nº 2053)
118	SFN (SEC ID Nº 40)	TTTGTGAAGAGTGAAATT (SEC ID Nº 2054)
119	SFN (SEC ID Nº 40)	TTTGCGAAGAGCGAAATT (SEC ID Nº 2053)
120	SFN (SEC ID Nº 40)	TTTGTGAAGAGTGAAATT (SEC ID Nº 2054)
121	SFN (SEC ID Nº 40)	TAACGAGGAGGGTTTCGGA (SEC ID Nº 2055)
122	SFN (SEC ID Nº 40)	TAATGAGGAGGGTTTGGA (SEC ID Nº 2056)
123	SFN (SEC ID Nº 40)	TAACGAGGAGGGTTTCGGA (SEC ID Nº 2055)
124	SFN (SEC ID Nº 40)	TAATGAGGAGGGTTTGGA (SEC ID Nº 2056)
125	SFN (SEC ID Nº 40)	GTTTCGAGGTGCGTGAGTA (SEC ID Nº 2057)
126	SFN (SEC ID Nº 40)	GTTTGAGGTGTGTGAGTA (SEC ID Nº 2058)
127	SFN (SEC ID Nº 40)	GTTTCGAGGTGCGTGAGTA (SEC ID Nº 2057)
128	SFN (SEC ID Nº 40)	GTTTGAGGTGTGTGAGTA (SEC ID Nº 2058)
129	PRKCD (SEC ID Nº 121)	ATTTATTTTTCGTTGTAGG (SEC ID Nº 2059)
130	PRKCD (SEC ID Nº 121)	TATTTATTTTTGTTGTAGG (SEC ID Nº 2060)
131	PRKCD (SEC ID Nº 121)	TTTCGGAACGGGAAT (SEC ID Nº 2061)
132	PRKCD (SEC ID Nº 121)	TAGTTTTGGAAATGGGA (SEC ID Nº 2062)
133	PRKCD (SEC ID Nº 121)	GGACGGAGTTATCGGT (SEC ID Nº 2063)
134	PRKCD (SEC ID Nº 121)	GGATGGAGTTATTGGTA (SEC ID Nº 2064)
135	PRKCD (SEC ID Nº 121)	GTTTAGCGGAGGGATA (SEC ID Nº 2065)
136	PRKCD (SEC ID Nº 121)	TGTTTAGTGAGGGAT (SEC ID Nº 2066)
137	SYK (SEC ID Nº 78)	GAAGTTATCGCGTTGG (SEC ID Nº 2067)
138	SYK (SEC ID Nº 78)	AGAAGTTATTGTGTTGG (SEC ID Nº 2068)
139	SYK (SEC ID Nº 78)	GATCGATGCGGTTTAT (SEC ID Nº 2069)
140	SYK (SEC ID Nº 78)	GGGATTGATGTGGTTTA (SEC ID Nº 2070)
141	SYK (SEC ID Nº 78)	GTTTCGGCGGAGGAGA (SEC ID Nº 2071)
142	SYK (SEC ID Nº 78)	GTTTGGTGGGAGGAGA (SEC ID Nº 2072)

Nº:	Gen	Oligonucleótido:
143	SYK (SEC ID Nº 78)	GTTCGGCGGGAGGAGA (SEC ID Nº 2071)
144	SYK (SEC ID Nº 78)	GTTTGGTGGGAGGAGA (SEC ID Nº 2072)
145	SYK (SEC ID Nº 78)	AGTCGATTTTCGTTTAG (SEC ID Nº 2073)
146	SYK (SEC ID Nº 78)	TAGTTGATTTTGTTTAGT (SEC ID Nº 2074)
147	SYK (SEC ID Nº 78)	GGAAGAGTCGCGGGTT (SEC ID Nº 2075)
148	SYK (SEC ID Nº 78)	GGAAGAGTTGTGGGTT (SEC ID Nº 2076)
149	SYK (SEC ID Nº 78)	GGAAGAGTCGCGGGTT (SEC ID Nº 2075)
150	SYK (SEC ID Nº 78)	GGAAGAGTTGTGGGTT (SEC ID Nº 2076)
151	VTN (SEC ID Nº 115)	GGTGGTATCGATTGAT (SEC ID Nº 2077)
152	VTN (SEC ID Nº 115)	TGGTGGTATTGATTGAT (SEC ID Nº 2078)
153	VTN (SEC ID Nº 115)	TAGTGATTCGCGGGGA (SEC ID Nº 2079)
154	VTN (SEC ID Nº 115)	TAGTGATTTGTGGGGA (SEC ID Nº 2080)
155	VTN (SEC ID Nº 115)	TAGTGATTCGCGGGGA (SEC ID Nº 2079)
156	VTN (SEC ID Nº 115)	TAGTGATTTGTGGGGA (SEC ID Nº 2080)
157	VTN (SEC ID Nº 115)	TTATGTCGGAGGATGA (SEC ID Nº 2081)
158	VTN (SEC ID Nº 115)	ATTATGTTGGAGGATGA (SEC ID Nº 2082)
159	VTN (SEC ID Nº 115)	ATACGTTTATGACGAT (SEC ID Nº 2083)
160	VTN (SEC ID Nº 115)	ATATGGTTTATGATGATGG (SEC ID Nº 2084)
161	GRIN2D (SEC ID Nº 86)	GAGAGTCGGGATGATT (SEC ID Nº 2085)
162	GRIN2D (SEC ID Nº 86)	GGAGAGTTGGGATGAT (SEC ID Nº 2086)
163	GRIN2D (SEC ID Nº 86)	AGAGATTTTCGATTTGGA (SEC ID Nº 2087)
164	GRIN2D (SEC ID Nº 86)	AAGAGATTTTGATTTGGA (SEC ID Nº 2088)
165	GRIN2D (SEC ID Nº 86)	TAGGGTCGAGATTTGG (SEC ID Nº 2089)
166	GRIN2D (SEC ID Nº 86)	TTAGGGTTGAGATTTGG (SEC ID Nº 2090)
167	GRIN2D (SEC ID Nº 86)	AGTGTGGCGAATATTG (SEC ID Nº 2091)
168	GRIN2D (SEC ID Nº 86)	GTGTGGTGAATATTGAA (SEC ID Nº 2092)
169	TGFBR2 (SEC ID Nº 43)	GAAAACGTGGACGTTTTT (SEC ID Nº 2093)
170	TGFBR2 (SEC ID Nº 43)	GAAAATGTGGATGTTTTT (SEC ID Nº 2094)
171	TGFBR2 (SEC ID Nº 43)	GAAAACGTGGACGTTTTT (SEC ID Nº 2093)
172	TGFBR2 (SEC ID Nº 43)	GAAAATGTGGATGTTTTT (SEC ID Nº 2094)
173	TGFBR2 (SEC ID Nº 43)	ATTTGGAGCGAGGAATTT (SEC ID Nº 2095)
174	TGFBR2 (SEC ID Nº 43)	ATTTGGAGTGAGGAATTT (SEC ID Nº 2096)
175	TGFBR2 (SEC ID Nº 43)	ATTTGGAGCGAGGAATTT (SEC ID Nº 2095)
176	TGFBR2 (SEC ID Nº 43)	ATTTGGAGTGAGGAATTT (SEC ID Nº 2096)

Nº:	Gen	Oligonucleótido:
177	TGFBR2 (SEC ID N° 43)	AGTTGAAAGTCGGTTAAA (SEC ID N° 2097)
178	TGFBR2 (SEC ID N° 43)	AGTTGAAAGTTGGTTAAA (SEC ID N° 2098)
179	TGFBR2 (SEC ID N° 43)	AGTTGAAAGTCGGTTAAA (SEC ID N° 2097)
180	TGFBR2 (SEC ID N° 43)	AGTTGAAAGTTGGTTAAA (SEC ID N° 2098)
181	TGFBR2 (SEC ID N° 43)	AAAGTTTTTCGAGGGGTT (SEC ID N° 2099)
182	TGFBR2 (SEC ID N° 43)	AAAGTTTTTCGAGGGGTT (SEC ID N° 2100)
183	TGFBR2 (SEC ID N° 43)	AAAGTTTTTCGAGGGGTT (SEC ID N° 2099)
184	TGFBR2 (SEC ID N° 43)	AAAGTTTTTCGAGGGGTT (SEC ID N° 2100)
185	TGFBR2 (SEC ID N° 43)	GGTAGTTACGAGAGAGTT (SEC ID N° 2101)
186	TGFBR2 (SEC ID N° 43)	GGTAGTTATGAGAGAGTT (SEC ID N° 2102)
187	TGFBR2 (SEC ID N° 43)	GGTAGTTACGAGAGAGTT (SEC ID N° 2101)
188	TGFBR2 (SEC ID N° 43)	GGTAGTTATGAGAGAGTT (SEC ID N° 2102)
189	COX7A2L (SEC ID N° 105)	TTGTTCTGAAGATCGTT (SEC ID N° 2103)
190	COX7A2L (SEC ID N° 105)	GTTGTTTGAAGATTGTTT (SEC ID N° 2104)
191	COX7A2L (SEC ID N° 105)	TAGCGTAAGGATTCGGT (SEC ID N° 2105)
192	COX7A2L (SEC ID N° 105)	TTAGTGTAAGGATTTGGT (SEC ID N° 2106)
193	COX7A2L (SEC ID N° 105)	AGAGTTCGGTTTTTCGTA (SEC ID N° 2107)
194	COX7A2L (SEC ID N° 105)	AGAGTTTGGTTTTTTGTA (SEC ID N° 2108)
195	COX7A2L (SEC ID N° 105)	AGAGTTCGGTTTTTCGTA (SEC ID N° 2107)
196	COX7A2L (SEC ID N° 105)	AGAGTTTGGTTTTTTGTA (SEC ID N° 2108)
197	COX7A2L (SEC ID N° 105)	ATTCGTATTTGCGGGTTA (SEC ID N° 2109)
198	COX7A2L (SEC ID N° 105)	ATTTGTATTTGTGGGTTA (SEC ID N° 2110)
199	COX7A2L (SEC ID N° 105)	ATTCGTATTTGCGGGTTA (SEC ID N° 2109)
200	COX7A2L (SEC ID N° 105)	ATTTGTATTTGTGGGTTA (SEC ID N° 2110)
201	DAG1 (SEC ID N° 125)	TTTCGTGGCGGAGAAT (SEC ID N° 2111)
202	DAG1 (SEC ID N° 125)	TTTTGTGGTGGAGAAT (SEC ID N° 2112)
203	DAG1 (SEC ID N° 125)	TTTCGTGGCGGAGAAT (SEC ID N° 2111)
204	DAG1 (SEC ID N° 125)	TTTTGTGGTGGAGAAT (SEC ID N° 2112)
205	DAG1 (SEC ID N° 125)	TACGGATATTTTCGGTT (SEC ID N° 2113)
206	DAG1 (SEC ID N° 125)	AATTATGGATATTTTGGTT (SEC ID N° 2114)
207	DAG1 (SEC ID N° 125)	TTACGATTCGTAGGTT (SEC ID N° 2115)
208	DAG1 (SEC ID N° 125)	TATTATTATGATTTGTAGGT (SEC ID N° 2116)
209	ONECUT2 (SEC ID N° 126)	TACGTAGTTGCGCGTT (SEC ID N° 2117)
210	ONECUT2 (SEC ID N° 126)	GTATGTAGTTGTGTGTT (SEC ID N° 2118)

Nº:	Gen	Oligonucleótido:
211	ONECUT2 (SEC ID Nº 126)	TTTTGTGCGTACGGAT (SEC ID Nº 2119)
212	ONECUT2 (SEC ID Nº 126)	TTTTGTGTGTATGGAT (SEC ID Nº 2120)
213	ONECUT2 (SEC ID Nº 126)	TTAAGCGGGCGTTGAT (SEC ID Nº 2121)
214	ONECUT2 (SEC ID Nº 126)	TTAAGTGGGTGTTGAT (SEC ID Nº 2122)
215	ONECUT2 (SEC ID Nº 126)	TTAAGCGGGCGTTGAT (SEC ID Nº 2121)
216	ONECUT2 (SEC ID Nº 126)	TTAAGTGGGTGTTGAT (SEC ID Nº 2122)
217	ONECUT2 (SEC ID Nº 126)	TAGAGGCGCGGGTTAT (SEC ID Nº 2123)
218	ONECUT2 (SEC ID Nº 126)	TAGAGGTGTGGTTAT (SEC ID Nº 2124)
219	ONECUT2 (SEC ID Nº 126)	TAGAGGCGCGGGTTAT (SEC ID Nº 2123)
220	ONECUT2 (SEC ID Nº 126)	TAGAGGTGTGGTTAT (SEC ID Nº 2124)
221	CYP2D6 (SEC ID Nº 129)	GAGATCGCGTTTCGT (SEC ID Nº 2125)
222	CYP2D6 (SEC ID Nº 129)	AGAGATTGTGTTTTGT (SEC ID Nº 2126)
223	CYP2D6 (SEC ID Nº 129)	ATTCGCGGCGAGGATA (SEC ID Nº 2127)
224	CYP2D6 (SEC ID Nº 129)	GATTTGTGGTGAGGAT (SEC ID Nº 2128)
225	CYP2D6 (SEC ID Nº 129)	GTCGTTTCGGGGACGT (SEC ID Nº 2129)
226	CYP2D6 (SEC ID Nº 129)	GTTGTTTTGGGGATGTG (SEC ID Nº 2130)
227	CYP2D6 (SEC ID Nº 129)	TAAGTAGCGTCGATAG (SEC ID Nº 2131)
228	CYP2D6 (SEC ID Nº 129)	AAGTAGTGTGATAGGG (SEC ID Nº 2132)
229	WBP11 (SEC ID Nº 137)	TTACGAGAAGCGGGTA (SEC ID Nº 2133)
230	WBP11 (SEC ID Nº 137)	ATTATGAGAAGTGGGTA (SEC ID Nº 2134)
231	WBP11 (SEC ID Nº 137)	AGGGGGCGATTTTCGG (SEC ID Nº 2135)
232	WBP11 (SEC ID Nº 137)	TAGGGGGTGATTTTGG (SEC ID Nº 2136)
233	WBP11 (SEC ID Nº 137)	TTAGCGTCGTTTGATT (SEC ID Nº 2137)
234	WBP11 (SEC ID Nº 137)	TTTGTAGTGTGTTGATT (SEC ID Nº 2138)
235	WBP11 (SEC ID Nº 137)	AGTTCGTTTTATTGCGT (SEC ID Nº 2139)
236	WBP11 (SEC ID Nº 137)	GAGTTTGTGTTTATTGTGT (SEC ID Nº 2140)

Tabla 3

Tabla 3. Puntuación de predicción para la respuesta ^a .										
Factor		<i>n</i>		<i>P</i>		OR ^b		(IC del 95%) ^c		
Puntuación basada en factores tradicionales, (χ^2 =10,9, 3 grados de libertad)										
q1	q2	q3	q4	50	50	50	50	0,013	1 1,51 3,16 2,90	(0,68-3,38) (1,40-7,15) (1,29-6,53)
Puntuación basada en la metilación, (χ^2 =44,3, 3 grados de libertad)										
q1	q2	q3	q4	50	50	50	50	<0,001	1 1,47 5,52 13,0	(0,62-3,47) (2,34-13,1) (4,97-33,8)

^a, Se usó regresión logística para someter a prueba la fuerza de una relación de un factor con el tipo de respuesta frente al tratamiento con tamoxifeno; ^b, Cociente de probabilidades; ^c, Intervalo de confianza del 95%; ^d, Puntuación basada en el análisis de variables múltiples de los factores tradicionales edad (>70 años, 55-70 años, 41-55 años frente a ≤ 40 años) sitio dominante de recidiva (recidiva en víscera o hueso con respecto a tejido blando), intervalo libre de enfermedad (>12 meses frente a ≤ 12 meses), y valores de log ER. La puntuación se dividió en grupos (q1 - q4) basados en los valores de los percentiles 25, 50 y 75; ^e, Puntuación basada en el estado de metilación del ADN de los 5 genes de predicción independientes PSA-T, STMN1, GRIN2D, TGFBR2 y S100A2. La puntuación se dividió en grupos, q1 - q4, igual que para la puntuación basada en factores tradicionales.

Tabla 4

Muestra	Respuesta	% de metilación mediante ensayo de CL	Chip
1	Remisión	80	+
2	Remisión	27	+
3	Remisión (resultado atípico)	3	-
4	Remisión	88	+
5	Remisión	68	+
6	Remisión	62	+
7	Remisión	60	+
8	Remisión	83	+
9	Remisión	56	+
10	Remisión (resultado atípico)	0	-
11	Remisión	0	+
12	Remisión	0	+
13	Enfermedad progresiva	0	-
14	Enfermedad progresiva	0	-
15	Enfermedad progresiva	4	-
16	Enfermedad progresiva	7	-
17	Enfermedad progresiva (resultado atípico)	0	+
18	Enfermedad progresiva	0	-
19	Enfermedad progresiva	0	-
20	Enfermedad progresiva	0	-
21	Enfermedad progresiva	0	-
22	Enfermedad progresiva	0	-
23	Enfermedad progresiva (resultado atípico)	0	+
24	Enfermedad progresiva	0	-

REIVINDICACIONES

1. Uso de

- i) una molécula de ácido nucleico que consiste en una secuencia de al menos 18 bases de longitud de una de las secuencias seleccionadas del grupo que consiste en las SEC ID N° 411, 412, 515, 516, 685, 686, 789 y 790, y secuencias complementarias de las mismas,
- ii) un oligonucleótido o PNA-oligómero que consiste en al menos una secuencia de bases que tiene una longitud de al menos 10 nucleótidos que es complementaria o idéntica a una de las secuencias de ácido nucleico de acuerdo con la SEC ID N° 515, 686 y 789,
- iii) un conjunto de oligonucleótidos de acuerdo con ii), o
- iv) un kit que comprende un reactivo bisulfito(= bisulfito, hdrogeno sulfito) así como oligonucleótidos y/o PNA-oligómeros de acuerdo con ii) o iii)

para el análisis del patrón de mutilación de PITX2 para predecir la respuesta de los pacientes con trastornos proliferativos de células mamarias a una terapia complementaria, que comprende uno o más fármacos dirigidos a la vía del receptor de estrógenos o están implicados en el metabolismo, producción o secreción de estrógenos, en el que dicho uno o más fármacos se seleccionan de tamoxifeno, inhibidores de la aromatasa, moduladores selectivos de los receptores de estrógenos y análogos de LH-RH y en el que dichos fármacos bloquean los efectos de los estrógenos sobre las células cancerosas o reducen los niveles de estrógenos en suero.

2. El uso de acuerdo con la reivindicación 1, **que se caracteriza porque** al menos un oligonucleótido está unido a una fase sólida.

3. El uso de acuerdo con la reivindicación 1, en el que dicho al menos un oligonucleótido se usa para la amplificación de secuencias de ácido nucleico que comprenden una de SEC ID N° 515, 686, 789 y secuencias complementarias de las mismas.

4. El uso de acuerdo con la reivindicación 1, en el que dicho kit comprende reactivos estándar para realizar un ensayo de mutilación del grupo que consiste en MS-SNuPE, MSP, metilo ligero, metilo pesado, secuenciación de ácidos nucleicos y combinaciones de los mismos.

Figura 1

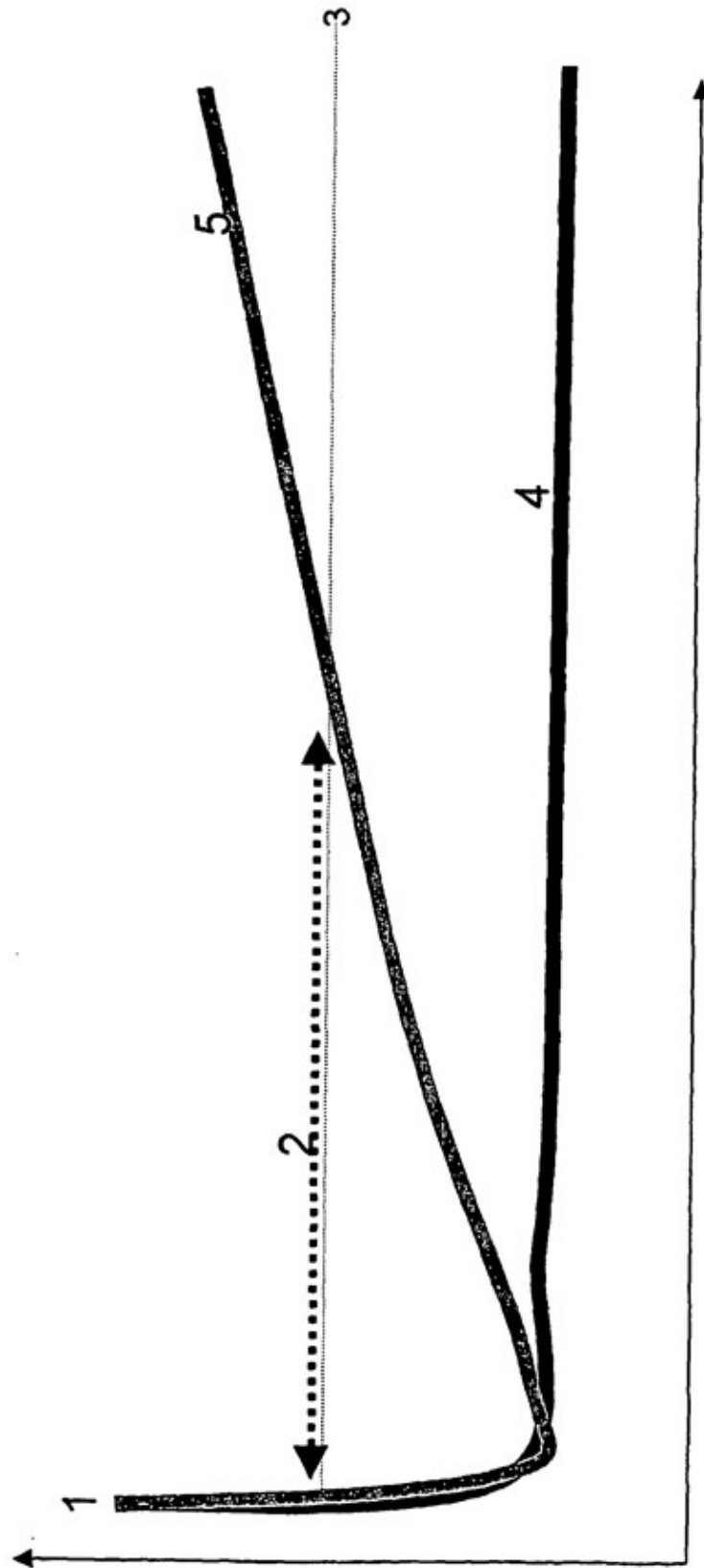
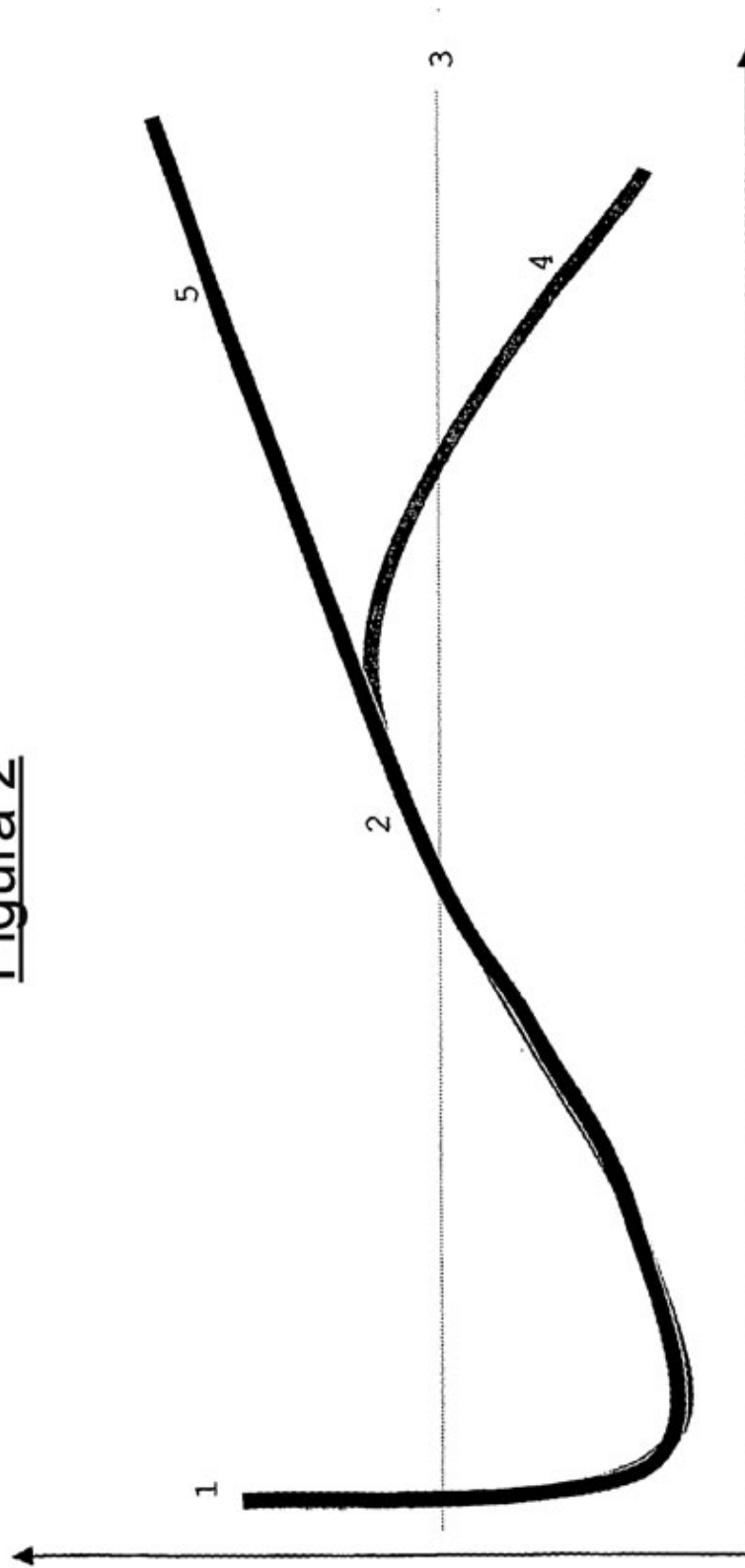


Figura 2



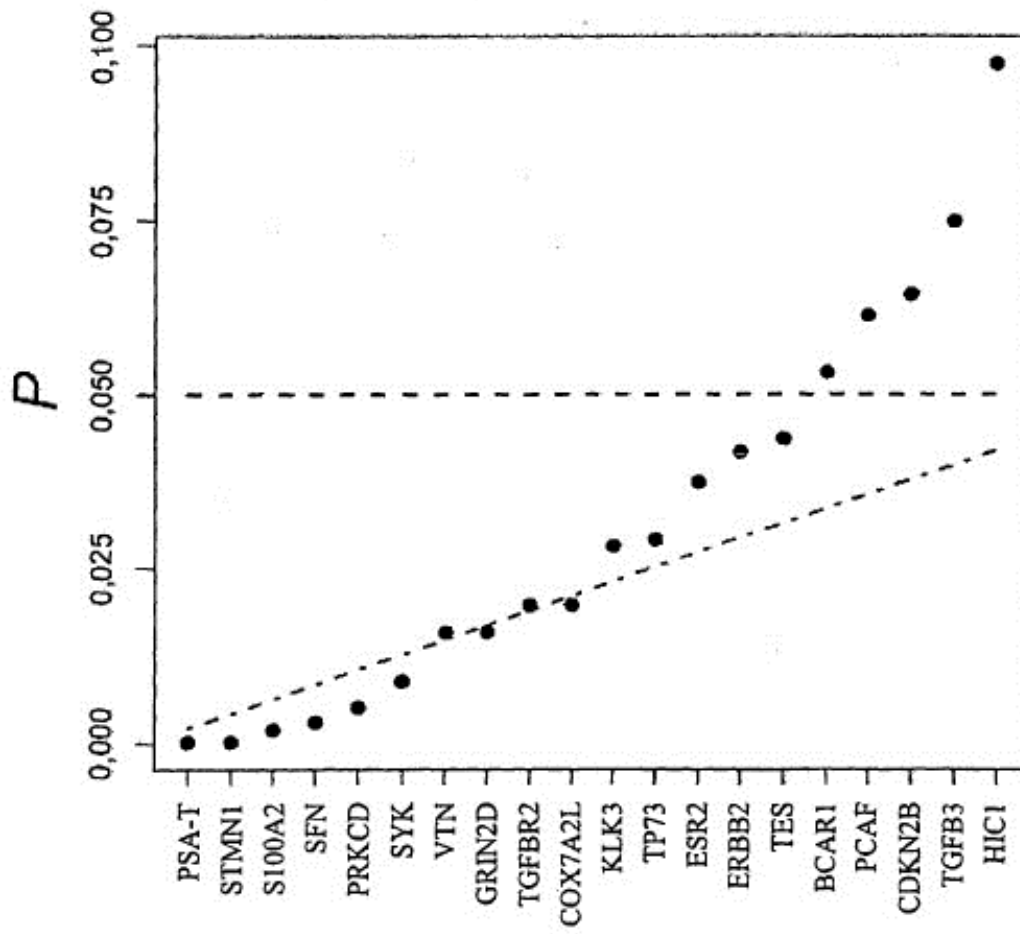


Figura 3

Figura 4

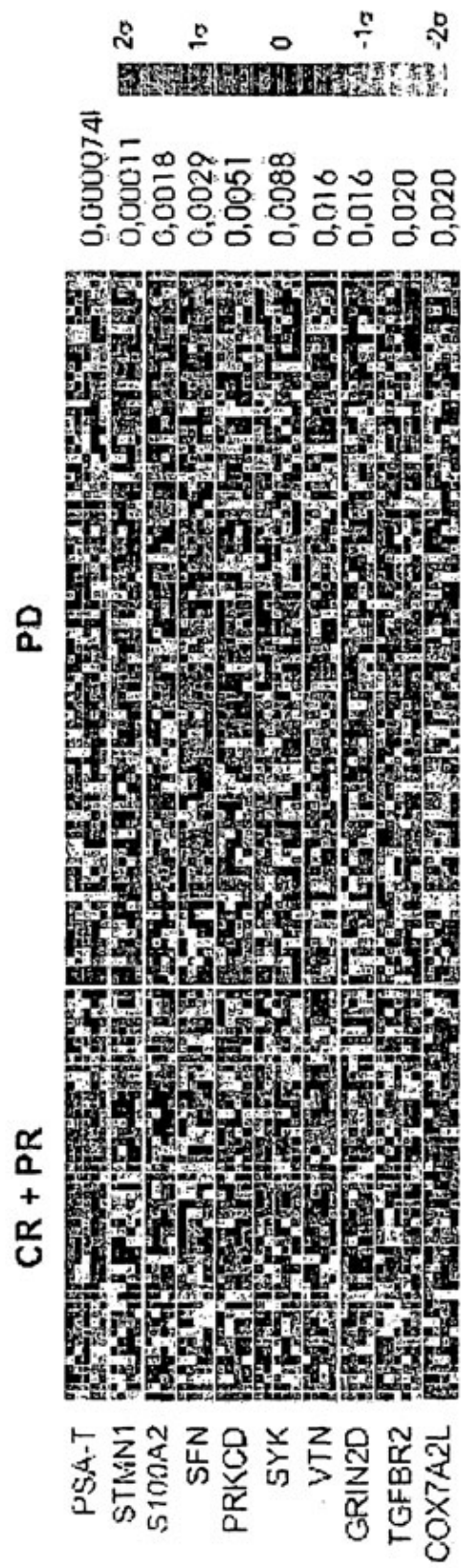


Figura 5

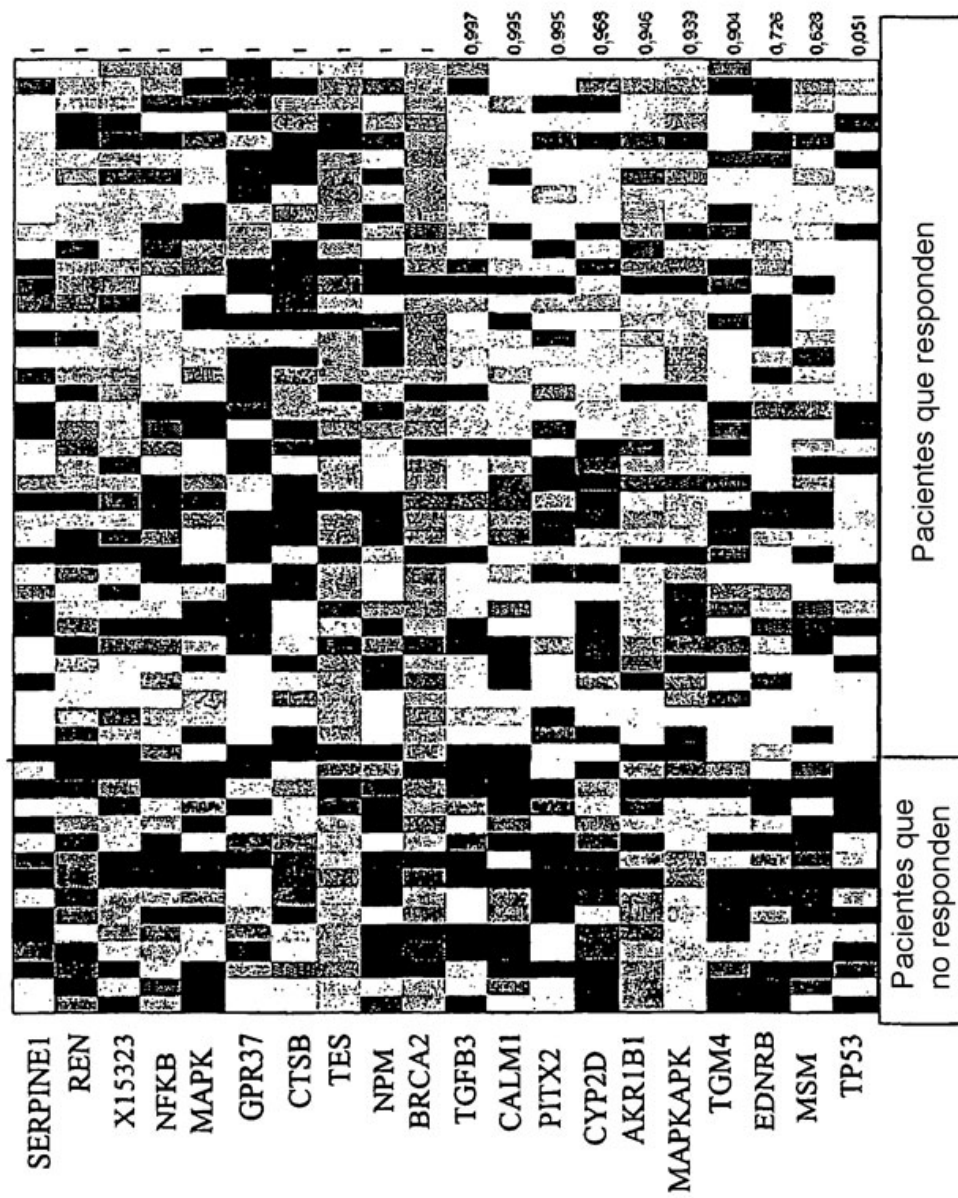


Figura 6

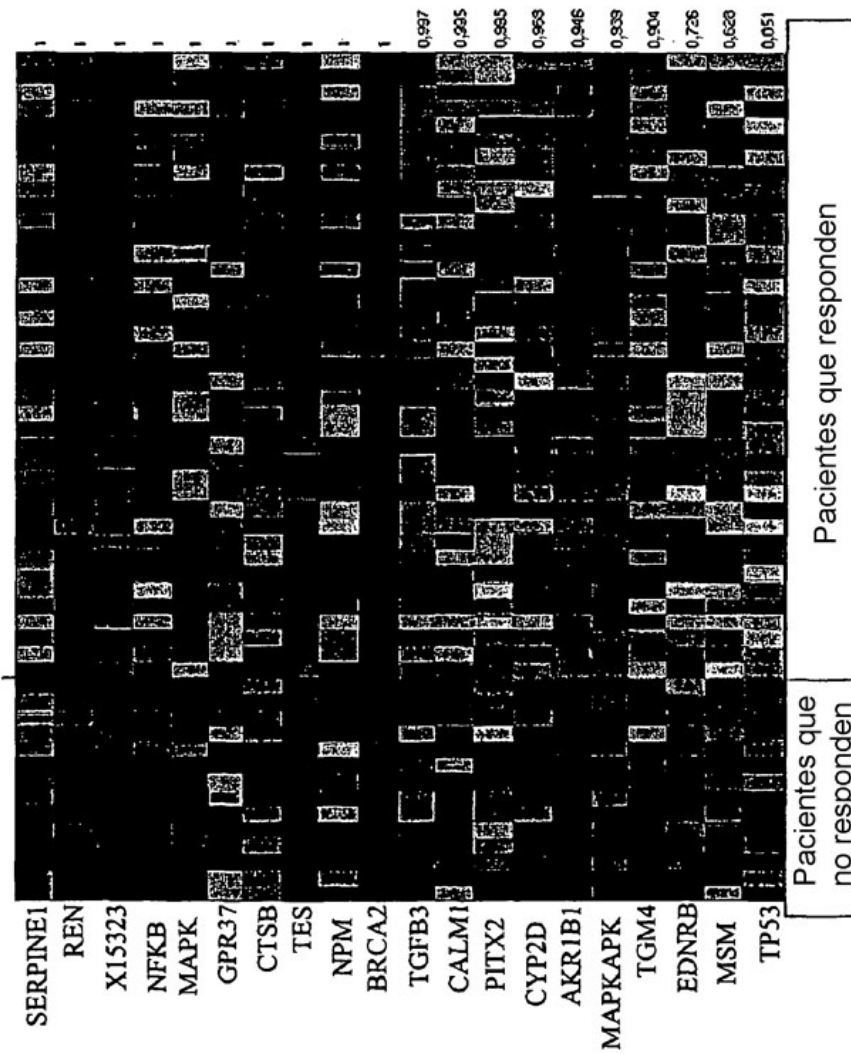


Figura 7

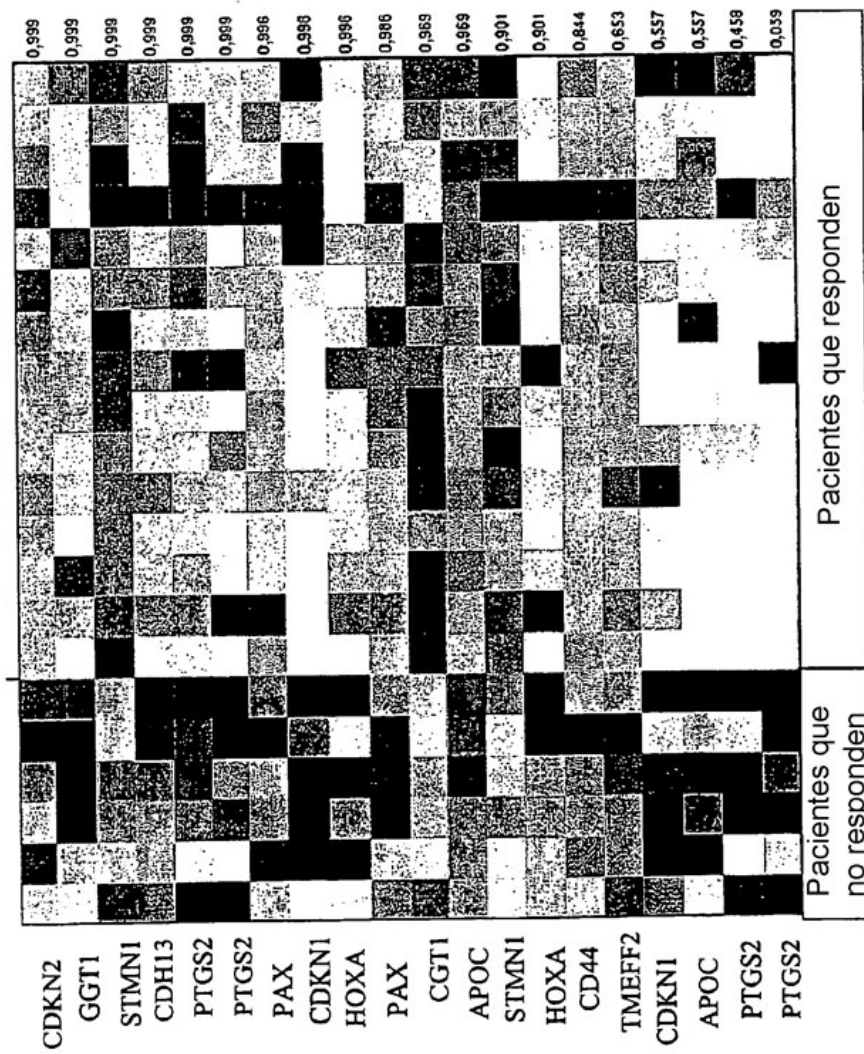


Figura 8

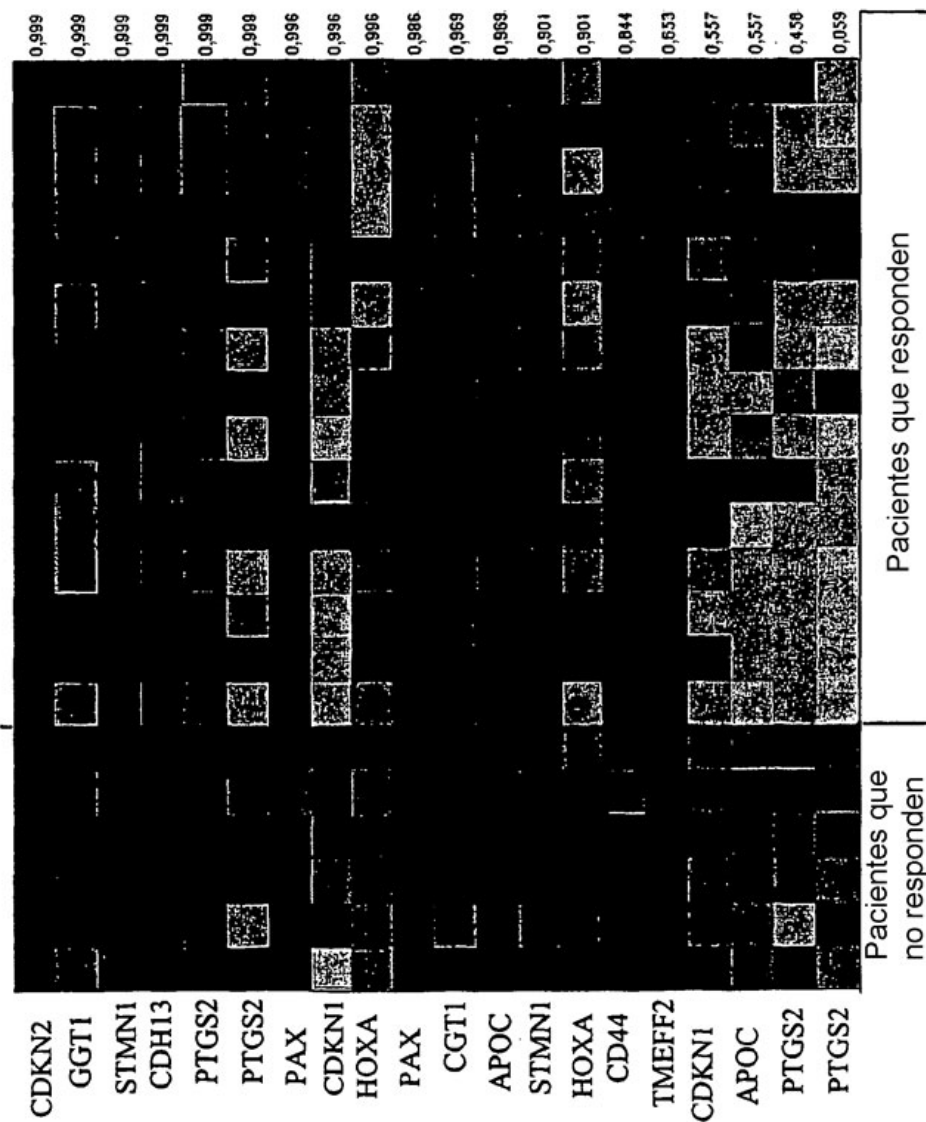


Figura 9

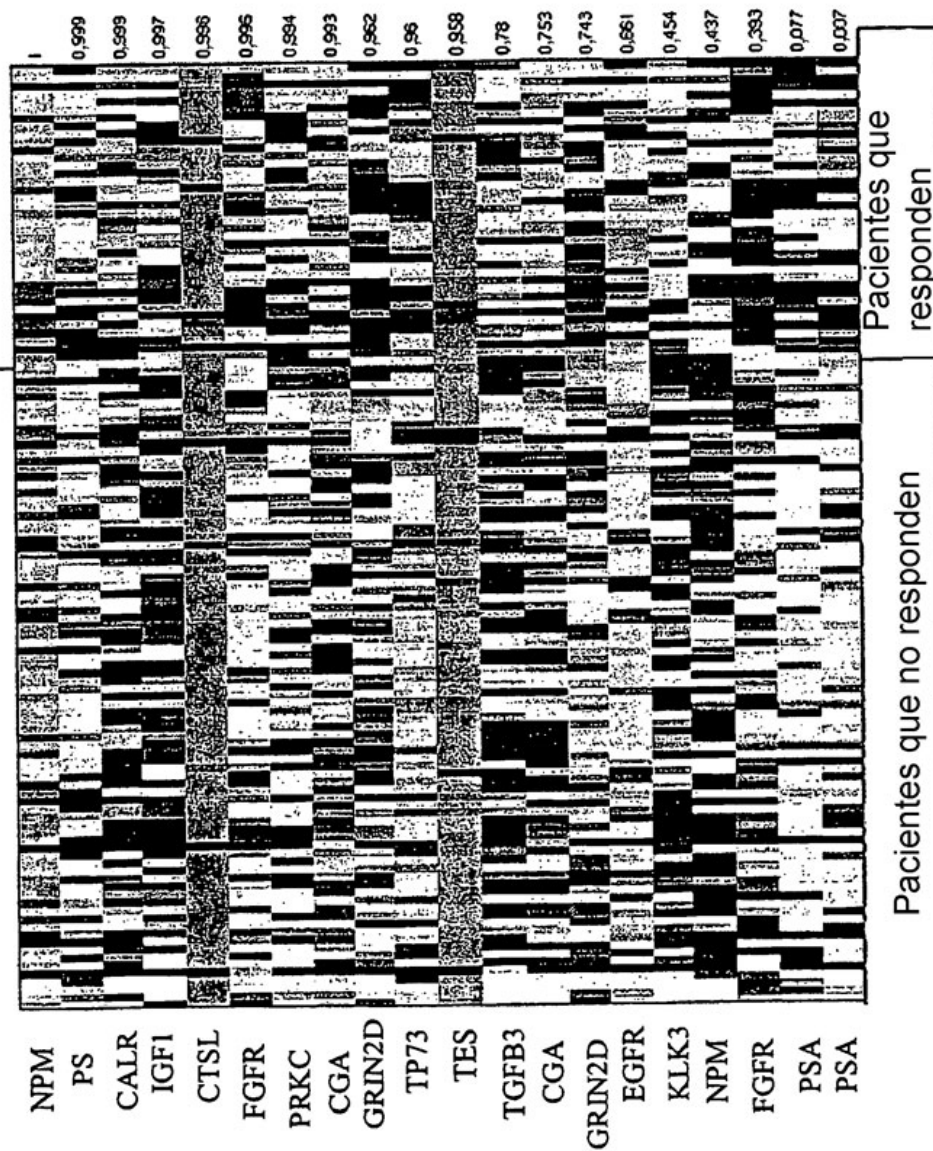


Figura 10

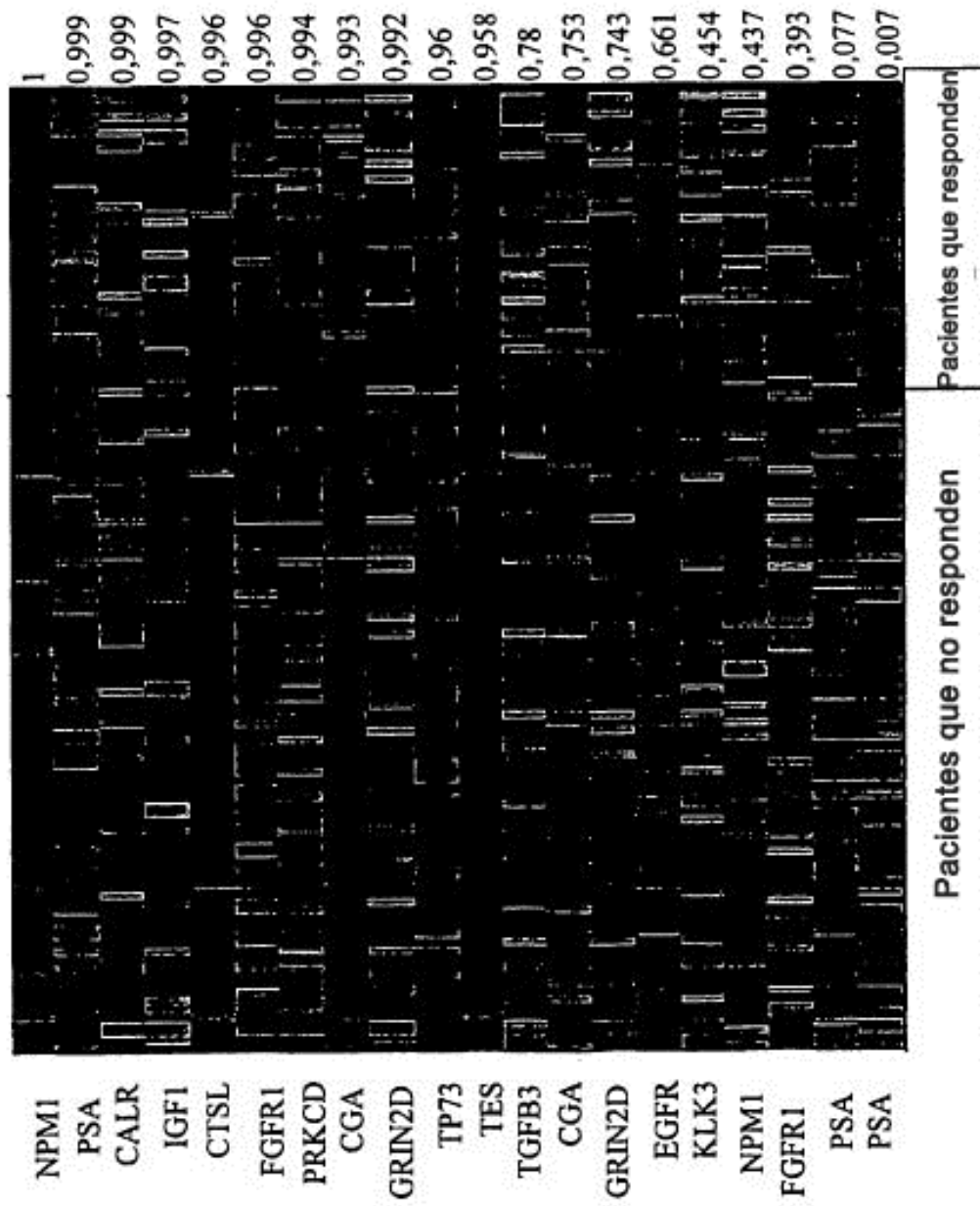


Figura 11

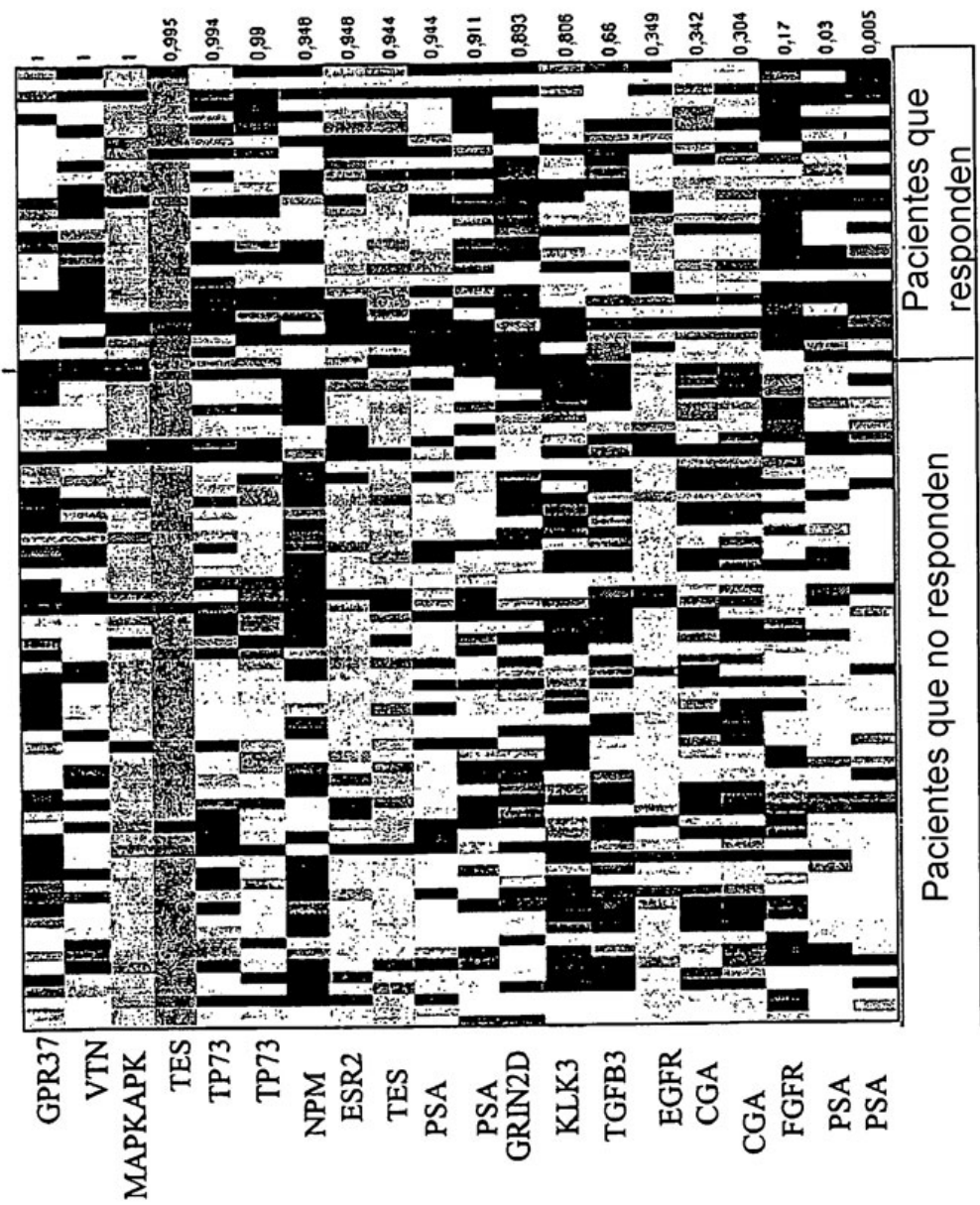


Figure 12

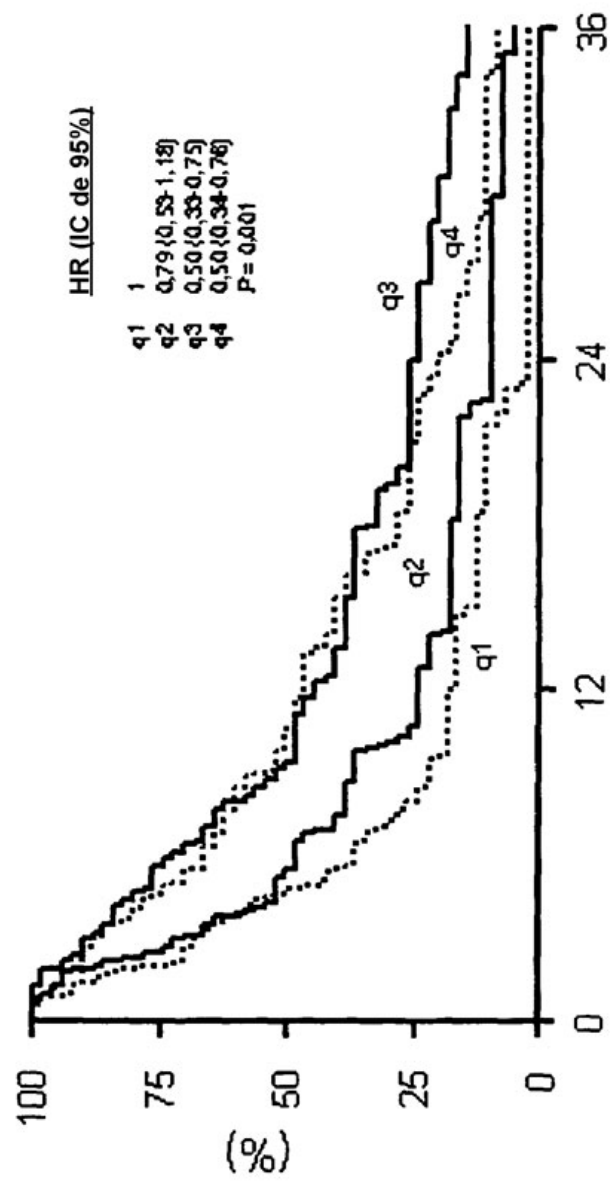


Figura 13



Figura 14

A



B

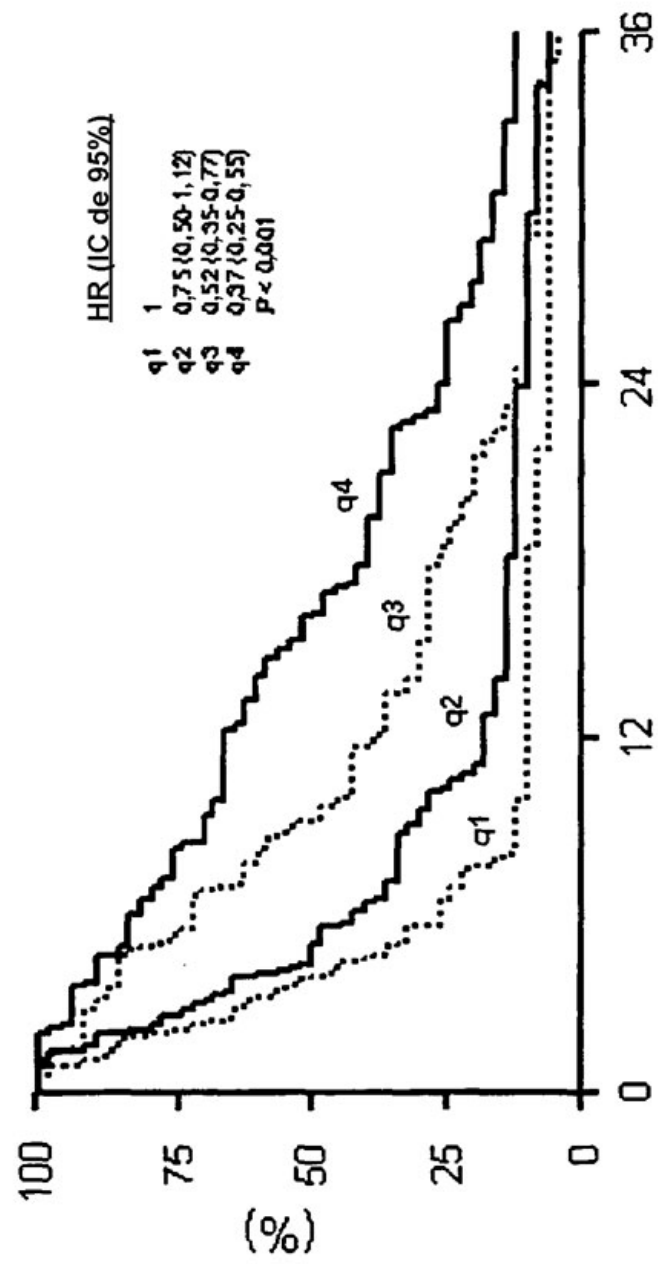


Figura 15

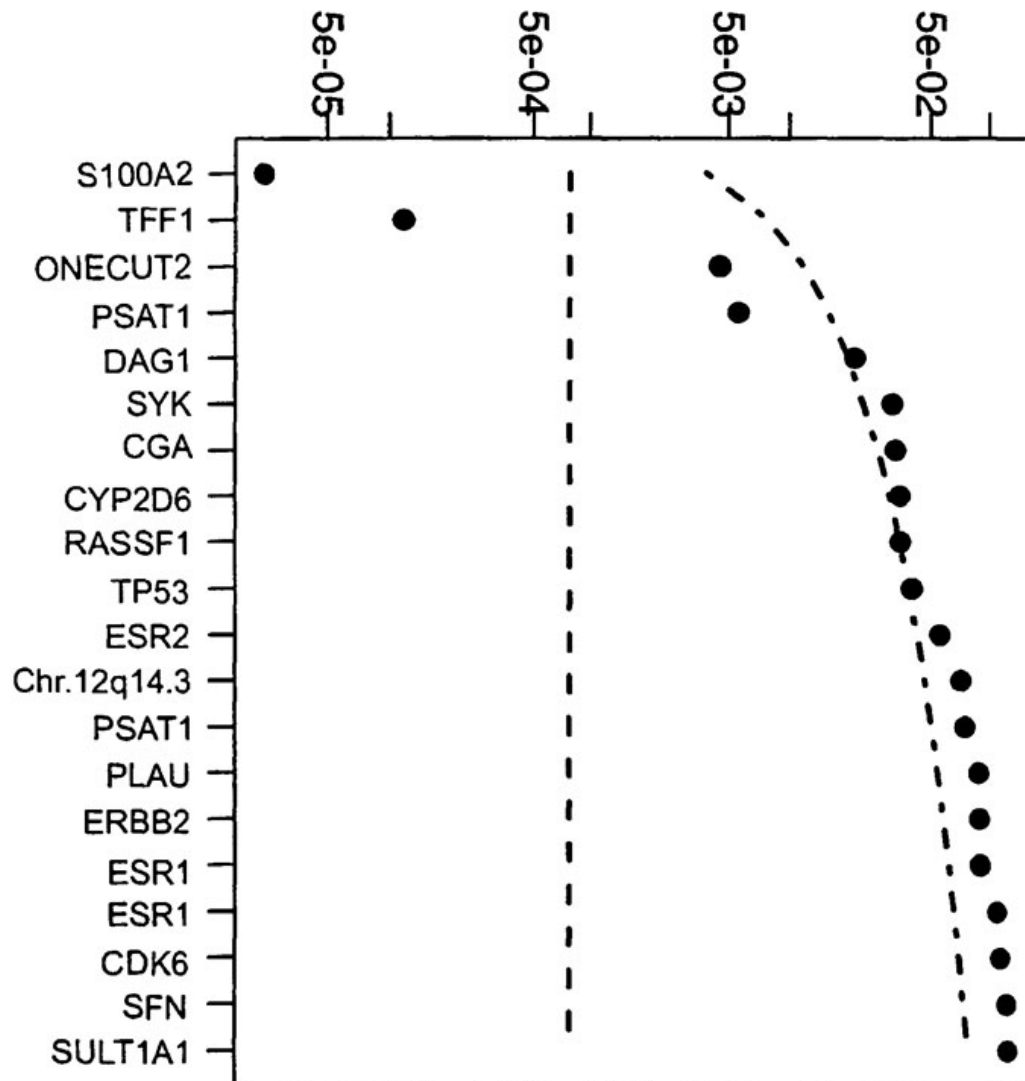


Figura 16

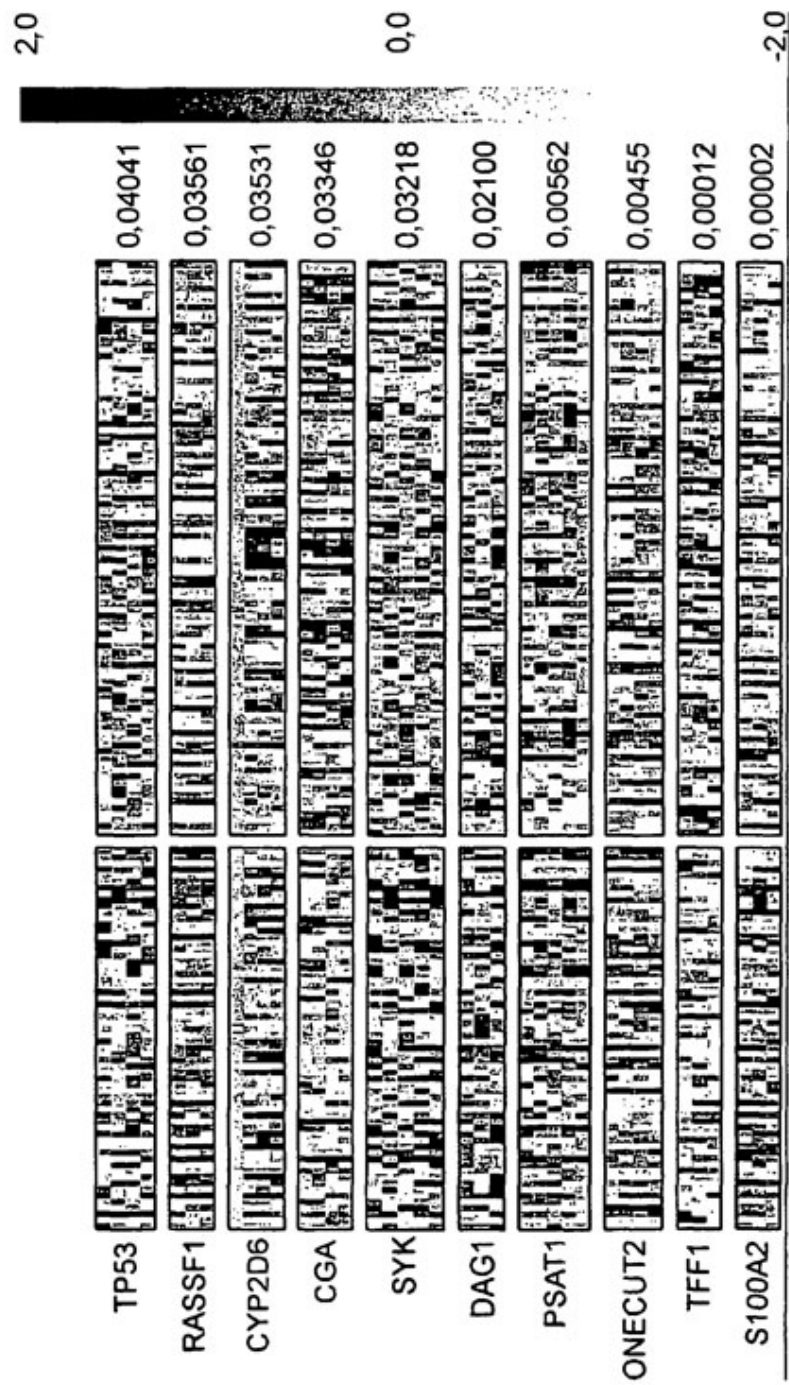


Figura 17

ESR1

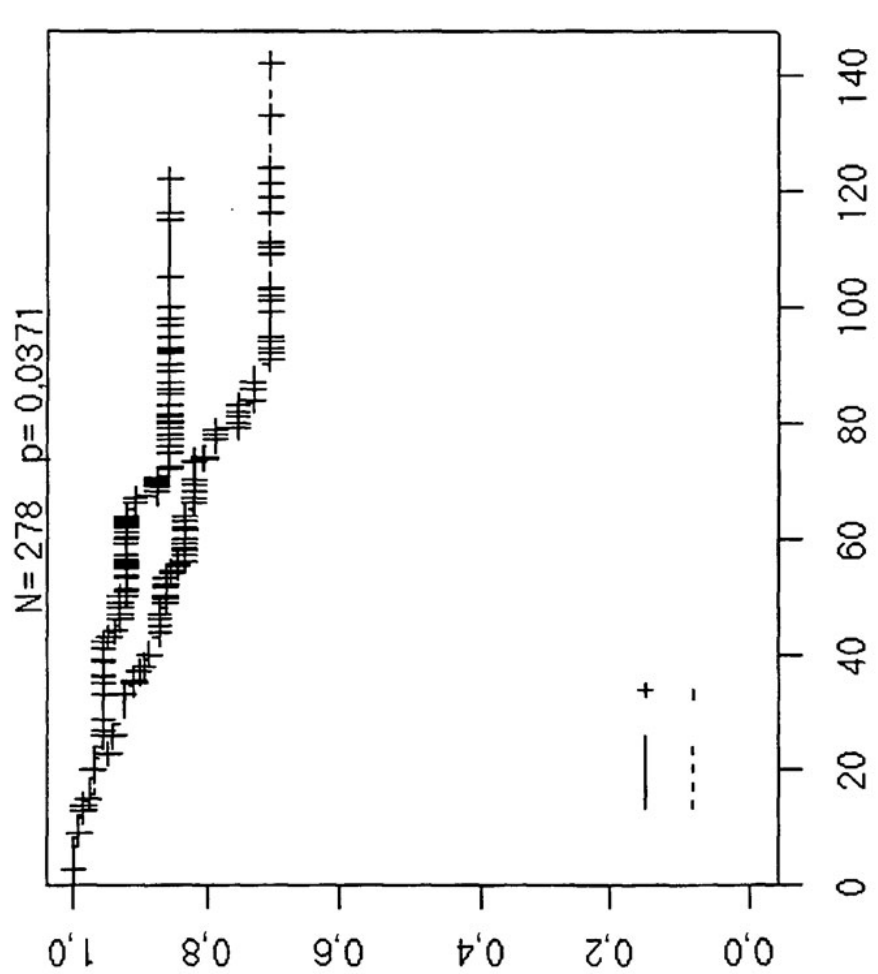


Figura 18

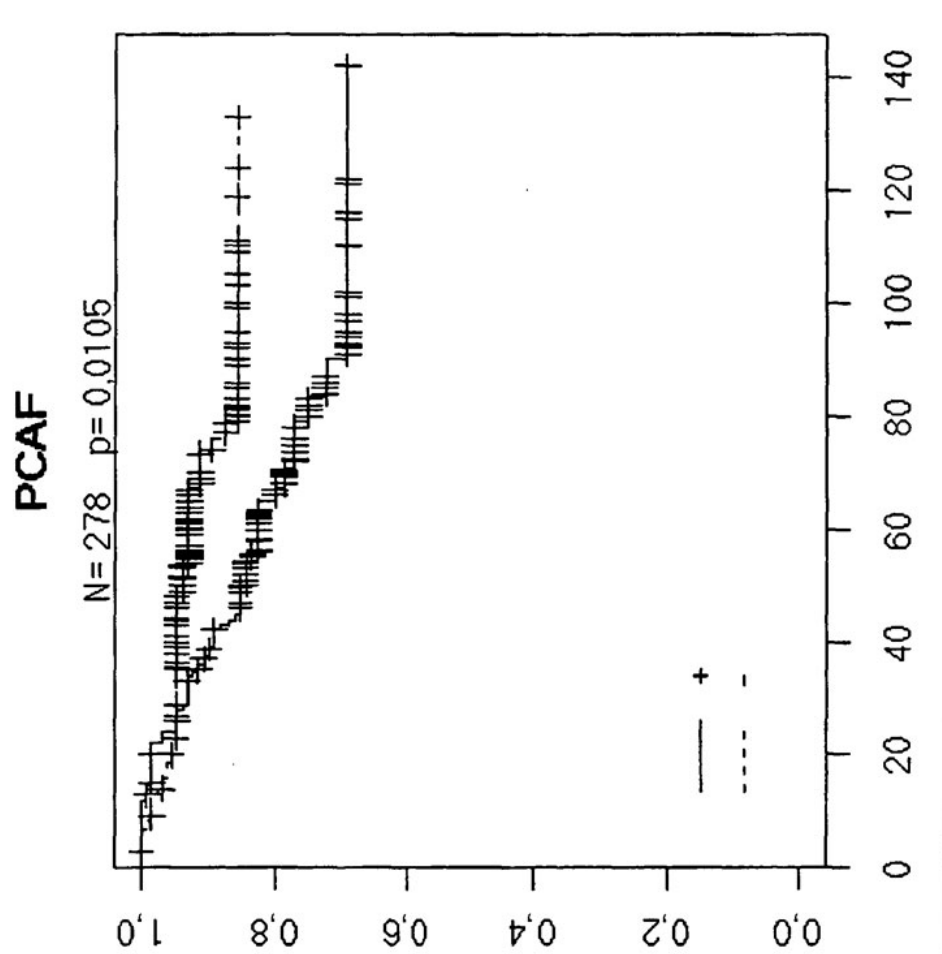


Figura 19

PITX2

N= 278 p= 3e-04

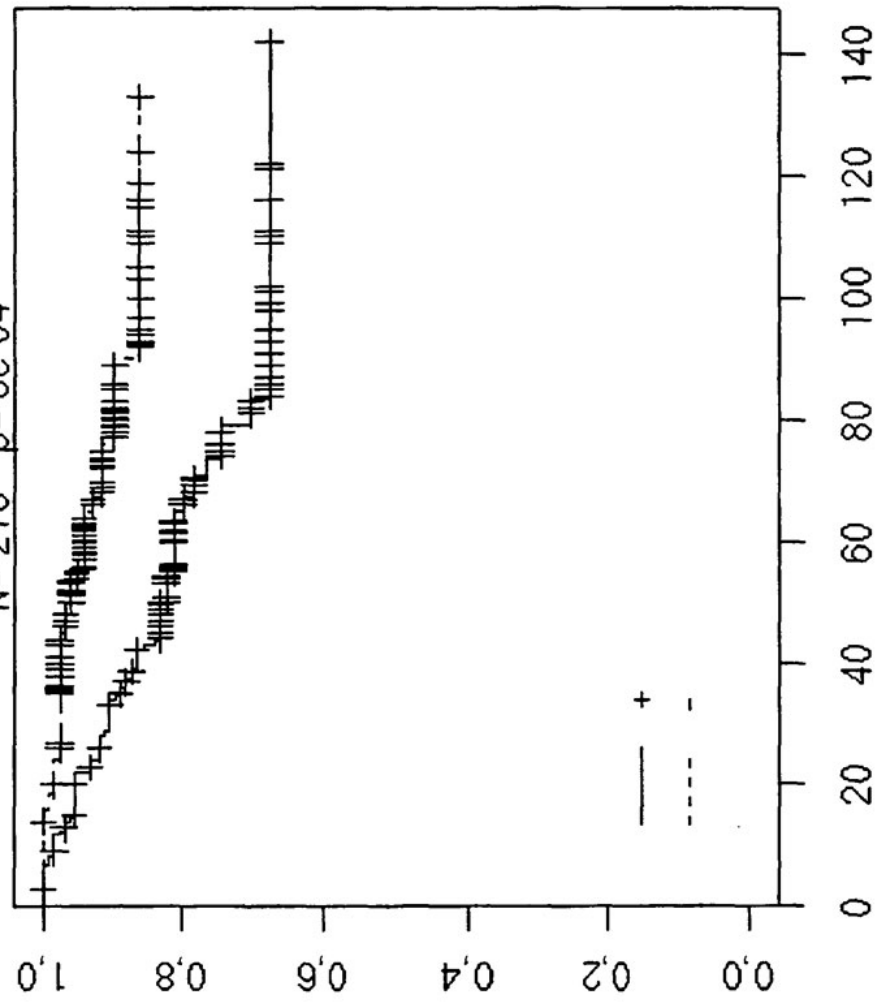


Figura 20

TMEFF2

N = 278 p = 0,0106

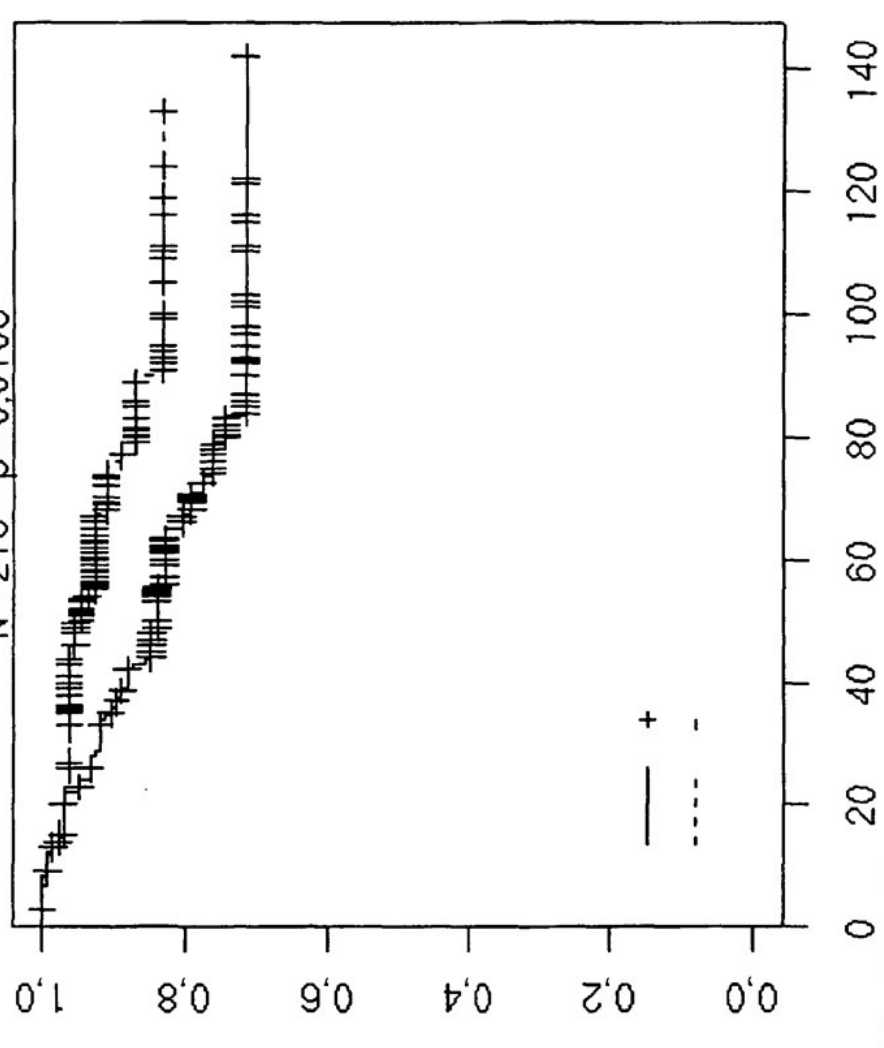


Figura 21

WBP11

N= 278 p= 0,0366

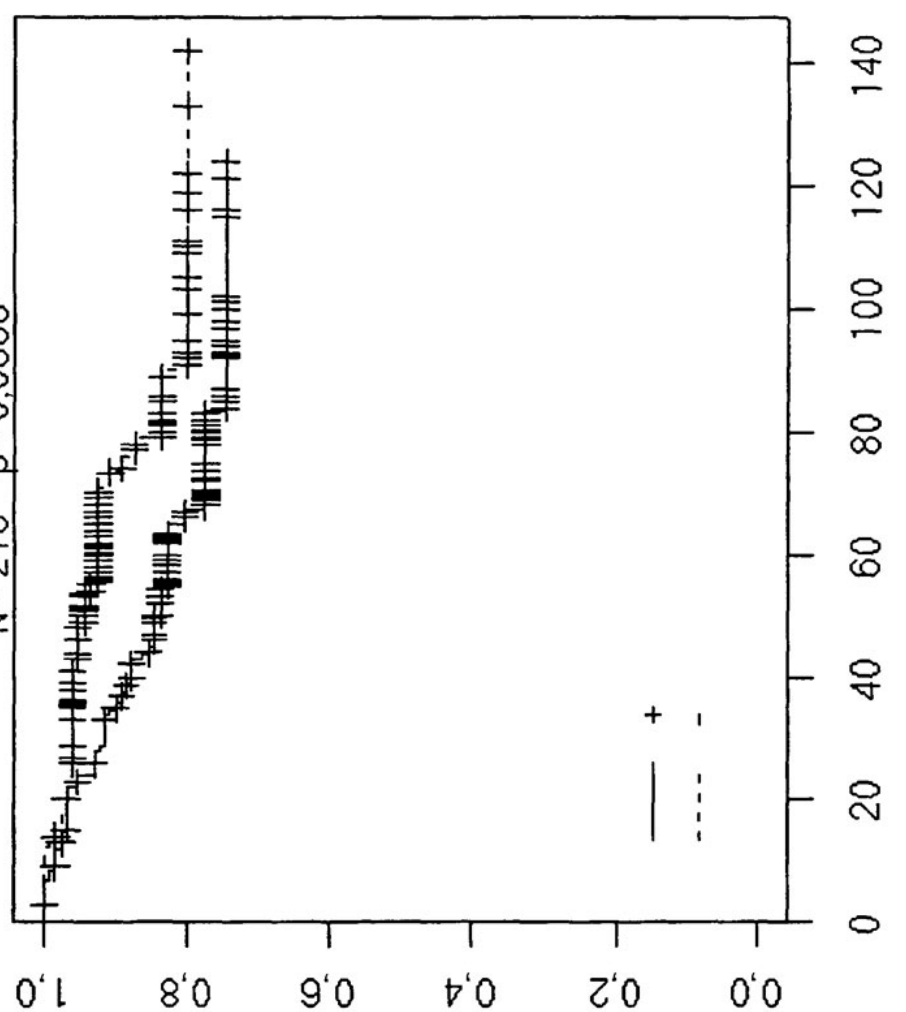


Figura 22

