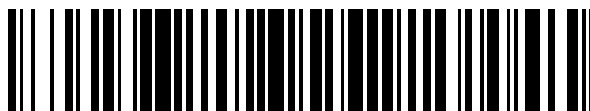


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 690 168**

51 Int. Cl.:

C12Q 1/6886 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.06.2004** **E 16196459 (8)**

97 Fecha y número de publicación de la concesión europea: **22.08.2018** **EP 3170906**

54 Título: **Predicción de la probabilidad de recidiva de cáncer**

30 Prioridad:

24.06.2003 US 482339 P

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

19.11.2018

73 Titular/es:

GENOMIC HEALTH, INC. (50.0%)
301 Penobscot Drive
Redwood City, CA 94063, US y
NSABP FOUNDATION, INC. (50.0%)

72 Inventor/es:

BAKER, JOFFRE;
SHAK, STEVE;
BRYANT, JOHN y
PAIK, SOONMYUNG

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 690 168 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

Predicción de la probabilidad de recidiva de cáncer

Antecedentes de la invención

5 La presente divulgación proporciona conjuntos de genes cuya expresión es importante en el diagnóstico y/o pronóstico del cáncer.

Descripción de la técnica relacionada

10 Los oncólogos disponen de varias opciones de tratamiento, incluyendo diferentes combinaciones de fármacos quimioterapéuticos que se caracterizan como "tratamiento de referencia", y varios fármacos que no tienen una declaración en la etiqueta para un cáncer particular, pero para los que hay pruebas de eficacia en ese cáncer. La mejor probabilidad de un buen desenlace del tratamiento requiere que se asignen los pacientes al tratamiento contra el cáncer disponible óptimo, y que esta asignación se realice lo más rápidamente posible tras el diagnóstico.

15 Actualmente, las pruebas de diagnóstico usadas en la práctica clínica son de un único analito, y por lo tanto no capturan el posible valor de conocer relaciones entre docenas de marcadores diferentes. Además, con frecuencia las pruebas de diagnóstico no son cuantitativas, basándose en la inmunohistoquímica. Con frecuencia este método proporciona resultados diferentes en diferentes laboratorios, en parte porque los reactivos no están normalizados, y en parte porque las interpretaciones son subjetivas y no pueden cuantificarse fácilmente. No se han usado pruebas basadas en ARN de manera frecuente debido al problema de la degradación del ARN a lo largo del tiempo y al hecho de que es difícil obtener muestras de tejido recientes de pacientes para su análisis. El tejido incluido en parafina, fijado, está más fácilmente disponible y se han establecido métodos para detectar ARN en tejido fijado. Sin embargo, estos métodos normalmente no permiten el estudio de grandes números de genes (ADN o ARN) a partir de pequeñas cantidades de material. Por lo tanto, tradicionalmente el tejido fijado se ha usado con poca frecuencia salvo para la detección mediante inmunohistoquímica de proteínas.

20 En los últimos años, varios grupos han publicado estudios relacionados con la clasificación de diversos tipos de cáncer mediante análisis de expresión génica en microalineamiento (véanse, por ejemplo Golub y col., *Science* 286: 531-537 (1999); Bhattacharjæ y col., *Proc. Natl. Acad. Sci. USA* 98: 13790-13795 (2001); Chen-Hsiang y col., *Bioinformatics* 17 (sup. 1): S316-S322 (2001); Ramaswamy y col., *Proc. Natl. Acad. Sci. USA* 98: 15149-15154 (2001)). También se han notificado determinadas clasificaciones de cánceres de mama humanos basándose en patrones de expresión génica (Martín y col., *Cancel Res.* 60: 2232-2238 (2000); West y col., *Proc. Natl. Acad. Sci. USA* 98: 11462-11467 (2001); Sorlie y col., *Proc. Natl. Acad. Sci. USA* 98: 10869-10874 (2001); Yan y col., *Cancer Res.* 61: 8375-8380 (2001)). Sin embargo, estos estudios se centran principalmente en mejorar y refinar la clasificación ya establecida de diversos tipos de cáncer, incluyendo cáncer de mama, y generalmente no proporcionan nuevos conocimientos sobre la relación de los genes expresados de manera diferencial, y no vinculan los hallazgos con estrategias de tratamiento con el fin de mejorar el desenlace clínico de la terapia contra el cáncer.

35 Aunque la biología molecular y la bioquímica modernas han revelado cientos de genes cuyas actividades influyen sobre el comportamiento de células tumorales, el estado de su diferenciación y su sensibilidad o resistencia a determinados fármacos terapéuticos, con pocas excepciones, no se ha aprovechado el estado de estos genes para el fin de tomar decisiones clínicas de manera rutinaria sobre tratamientos farmacológicos. Una excepción notable es el uso de la expresión de la proteína receptora de estrógenos (ER) en carcinomas de mama para seleccionar pacientes para un tratamiento con fármacos antiestrógenos, tales como tamoxifeno. Otro ejemplo excepcional es el uso de la expresión de proteína ErbB2 (Her2) en carcinomas de mama para seleccionar pacientes con el fármaco antagonista de Her2, Herceptin® (Genentech, Inc., South San Francisco, CA).

40 A pesar de los recientes avances, el desafío del tratamiento contra el cáncer sigue siendo dirigir regímenes de tratamiento específicos a tipos de tumores patógenamente diferenciados, y en última instancia personalizar el tratamiento contra el tumor con el fin de maximizar el desenlace. Por lo tanto, existe la necesidad de pruebas que proporcionen simultáneamente información predictiva sobre las respuestas del paciente a la diversidad de opciones de tratamiento. Esto es particularmente cierto para el cáncer de mama, cuya biología no se entiende mucho. Queda claro que la clasificación del cáncer de mama en unos pocos subgrupos, tales como subgrupo ErbB2+, y subgrupos caracterizados por una expresión génica de baja a ausente del receptor de estrógenos (ER) y unos pocos factores transcripcionales adicionales (Perou y col., *Nature* 406: 747-752 (2000)) no refleja la heterogeneidad celular y molecular del cáncer de mama, y no permite el diseño de estrategias de tratamiento que maximicen la respuesta del paciente.

45 En particular, una vez que a un paciente se le diagnostica cáncer, tal como cáncer de mama o de ovarios, existe una gran necesidad de métodos que permitan al médico predecir la evolución esperada de la enfermedad, incluyendo la probabilidad de recidiva de cáncer, la supervivencia a largo plazo del paciente, y similares, y seleccionar en consecuencia la opción de tratamiento más apropiada. <Reinhard Bos y col. (*Cancer*, Vol. 97, n.º 6, 1573-1581, 2003) desvelan el uso de ARN HIF-a para predecir la probabilidad de supervivencia libre de enfermedad de una

paciente con cáncer de mama ductal invasivo de ganglios negativos. WO 02/10436 divulga una sobreexpresión de CDC20 en los cánceres de mama de ganglios negativos clasificados por el Índice de Actividad Mitótica (MAI).

Resumen de la invención

5 La presente invención se define por las reivindicaciones. La expresión de CDC20 tiene valor pronóstico, específicamente con respecto a la supervivencia libre de enfermedad.

10 La presente invención se refiere al uso de material de biopsia incluido en parafina archivado para el ensayo de todos los marcadores en el conjunto y, por lo tanto, es compatible con el tipo más ampliamente disponible de material de biopsia. También es compatible con varios métodos diferentes de recogida de tejido tumoral, por ejemplo, a través de biopsia con aguja gruesa o aspiración con aguja fina. Además, para cada miembro del conjunto de genes, la invención especifica secuencias oligonucleotídicas que pueden usarse en el ensayo.

15 En un aspecto, la presente invención se refiere a un método para predecir la probabilidad de supervivencia a largo plazo en un paciente de cáncer de mama ductal ER positivo y ganglios negativos sin recidiva del cáncer, que consiste en determinar el nivel de expresión del transcrito de ARN de CDC20, o su producto de expresión, en una muestra de tejido o células de cáncer de mama obtenida del paciente, normalizada con respecto al nivel de expresión de una serie de transcritos de ARN de referencia, o sus productos de expresión, comparando el nivel de expresión normalizado de CDC20 del paciente con un nivel de expresión normalizado de CDC20 en una serie de referencia de tejidos de cáncer de mama que incluye a pacientes (a) vivos sin recidiva local, regional o distante del cáncer de mama, (b) vivos con recidiva contralateral del cáncer de mama, (c) vivos con un segundo cáncer primario distinto al de mama, o (d) murieron antes de la recidiva del cáncer de mama, donde un aumento del nivel de expresión normalizado de CDC20 del paciente indica una menor probabilidad de supervivencia a largo plazo sin recidiva del cáncer.

20 En un aspecto, la presente divulgación se refiere a un método para predecir la probabilidad de supervivencia a largo plazo en un paciente de cáncer sin recidiva del cáncer, que consiste en determinar el nivel de expresión de uno o más transcritos de ARN pronósticos, o sus productos de expresión, en una célula cancerígena obtenida del paciente, normalizada con respecto al nivel de expresión de todos los transcritos de ARN o sus productos en dicha célula cancerígena, o de una serie de referencia de transcritos de ARN, o sus productos de expresión, donde el transcrito de ARN pronóstico es el transcrito de uno o más genes seleccionados del grupo compuesto por B_Catenin; BAG1; BIN1; BUB1; C20_orf1; CCNB1; CCNE2; CDC20; CDH1; CEGP1; CIAP1; cMYC; CTSL2; DKFZp586M07; DR5; EpCAM; EstR1; FOXM1; GRB7; GSTM1; GSTM3; HER2; HNRPA; ID1; IGF1R; ITGA7; Ki_67; KNSL2; LMNB1; MCM2; MELK; MMP12; MMP9; MYBL2; NEK2; NME1; NP009; PCNA; PR; PREP; PTTG1; RPLPO; Src; STK15; STMY3; SURV; TFRC; TOP2A; y TS, inmovilizados sobre una superficie sólida,

30 donde la expresión de uno o más de los genes BUB1; C20_orf1; CCNB1; CCNE2; CDC20; CDH1; CTSL2; EpCAM; FOXM1; GRB7; HER2; HNRPA; Ki_67; KNSL2; LMNB1; MCM2; MELK; MMP12; MMP9; MYBL2; NEK2; NME1; PCNA; PREP; PTTG1; Src; STK15; STMY3; SURV; TFRC; TOP2A; y TS indica una disminución de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer; y

35 la expresión de uno o más de los genes BAG1; B_Catenin; BIN1; CEGP1; CIAP1; cMYC; DKFZp586M07; DR5; EstR1; GSTM1; GSTM3; ID1; IGF1R; ITGA7; NP009; PR; y RPLPO indica un aumento de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer.

40 En determinadas realizaciones, se determina el nivel de expresión de al menos 2, o al menos 5, o al menos 10, o al menos 15, o al menos 20, o al menos 25 transcritos de ARN pronósticos o sus productos de expresión.

El cáncer es un cáncer de mama de ganglios negativos, ER positivo.

En otra realización, el ARN comprende ARN intrónico.

45 En un aspecto de la presente divulgación, se determina el nivel de expresión de uno o más transcritos de ARN pronósticos, o sus productos de expresión, de uno o más genes seleccionados del grupo compuesto por MMP9, GSTM1, MELK, PR, DKFZp586M07, GSTM3, CDC20, CCNB1, STMY3, GRB7, MYBL2, CEGP1, SURV, LMNB1, CTSL2, PTTG1, BAG1, KNSL2, CIAP1, PREP, NEK2, EpCAM, PCNA, C20_orf1, ITGA7, ID1 B_Catenin, EstR1, CDH1, TS HER2, y cMYC,

50 donde la expresión de uno o más de los genes C20_orf1; CCNB1; CDC20; CDH1; CTSL2; EpCAM; GRB7; HER2; KNSL2; LMNB1; MCM2; MMP9; MYBL2; NEK2; PCNA; PREP; PTTG1; STMY3; SURV; TS; y MELK indica un descenso de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer; y

la expresión de uno o más de los genes BAG1; B_Catenin; CEGP1; CIAP1; cMYC; DKFZp586M07; EstR1; GSTM1; GSTM3; ID1; ITGA7; y PR indica un aumento de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer.

En otra realización de la presente divulgación, se determina el nivel de expresión de uno o más transcritos de ARN

pronósticos, o sus productos de expresión, de uno o más de los genes seleccionados del grupo compuesto por GRB7, SURV, PR, LMNB1, MYBL2, HER2, GSTM1, MELK, S20_orf1, PTTG1, BUB1, CDC20, CCNB1, STMY3, KNSL2, CTSL2, MCM2, NEK2; DR5, Ki_67, CCNE2, TOP2A, PCNA, PREP, FOXM1, NME1, CEGP1, BAG1, STK15, HNRPAB, EstR1, MMP9, DKFZp586M07, TS, Src, BIN1, NP009, RPLPO, GSTM3, MMP12, TFRC, y IGF1R.

5 donde la expresión de uno o más de los genes GRB7; SURV; LMNB1; MYBL2; HER2; MELK; C20_orf1; PTTG1; BUB1; CDC20; CCNB1; STMY3; KNSL2; CTSL2; MCM2; NEK2; Ki_67; CCNE2; TOP2A_4; PCNA; PREP; FOXM1; NME1; STK15; HNRPAB; MMP9; TS; Src; MMP12; y TFRC indica un descenso de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer; y

10 la expresión de uno o más de los genes PR; GSTM1; DR5; CEGP1; BAG1; EstR1; DKFZp586M07; BIN1; NPD009; RPLPO; GSTM3; e IGF1R indica un aumento de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer.

15 En otro aspecto, la divulgación se refiere a un método para predecir la probabilidad de supervivencia a largo plazo de un paciente de cáncer sin recidiva del cáncer, que consiste en determinar el nivel de expresión de uno o más transcritos de ARN pronósticos, o sus productos de expresión, en una célula cancerígena obtenida de dicho paciente, normalizada frente al nivel de expresión de todos los transcritos de ARN o sus productos de la célula cancerígena, o de una serie de referencia de transcritos de ARN, o sus productos de expresión, donde el transcrito de ARN pronóstico es el transcrito de uno o más de los genes seleccionados del grupo compuesto por GRB7; LMNB1; ER; STMY3; KLK10; PR; KRT5; FGFR1; MCM6; SNRPF,

20 donde la expresión de uno o más de los genes GRB7, LMNB1, STMY3, KLK10, FGFR1, y SNRPF indica un descenso de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer, y

la expresión de uno o más de los genes ER, PR, KRT5 y MCM6 ER, PR, KRT5 y MCM6 indica un aumento de la probabilidad de supervivencia a largo plazo sin recidiva del cáncer.

En una realización de este método, el ARN se aísla de una muestra de tejido de cáncer de mama embebida en cera.

25 En otra realización, el ARN se aísla del tejido extraído de biopsia con aguja gruesa o células aspiradas extraídas con aguja fina.

30 En un aspecto diferente, la divulgación se refiere al alineamiento que comprende polinucleótidos que se hibridan con dos o más de los siguientes genes: B_Catenin; BAG1; BIN1; BUB1; C20_orf1; CCNB1; CCNE2; CDC20; CDH1; CEGP1; CIAP1; cMYC; CTSL2; DKFZp586M07; DR5; EpCAM; EstR1; FOXM1; GRB7; GSTM1; GSTM3; HER2; HNRPAB; ID1; IGF1R; ITGA7; Ki_67; KNSL2; LMNB1; MCM2; MELK; MMP12; MMP9; MYBL2; NEK2; NME1; NPD009; PCNA; PR; PREP; PTTG1; RPLPO; Src; STK15; STMY3; SURV; TFRC; TOP2A; y TS, inmovilizados sobre una superficie sólida.

35 En una realización, el alineamiento comprende polinucleótidos que se hibridan con dos o más de los siguientes genes: MMP9, GSTM1, MELK, PR, DKFZp586M07, GSTM3, CDC20, CCNB1, STMY3, GRB7, MYBL2, CEGP1, SURV, LMNB1, CTSL2, PTTG1, BAG1, KNSL2, CIAP1, PREP, NEK2, EpCAM, PCNA, C20_orf1, ITGA7, ID1 B_Catenin, EstR1, CDH1, TS HER2, y cMYC.

40 En otra realización, el alineamiento comprende polinucleótidos que se hibridan con dos o más de los siguientes genes: GRB7, SURV, PR, LMNB1, MYBL2, HER2, GSTM1, MELK, S20_orf1, PTTG1, BUB1, CDC20, CCNB1, STMY3, KNSL2, CTSL2, MCM2, NEK2, DR5, Ki_67, CCNE2, TOP2A, PCNA, PREP, FOXM1, NME1, CEGP1, BAG1, STK15, HNRPAB, EstR1, MMP9, DKFZp586M07, TS, Src, BIN1, NP009, RPLPO, GSTM3, MMP12, TFRC, e IGF1R.

En una realización adicional, los alineamientos comprenden polinucleótidos que se hibridan con al menos 3, o al menos 5, o al menos 10, o al menos 15, o al menos 20, o al menos 25 de los genes enumerados.

En otra realización adicional, los alineamientos comprenden polinucleótidos que se hibridan con todos los genes enumerados.

45 En otra realización adicional, los alineamientos comprenden más de un polinucleótido que se hibrida con el mismo gen.

En una realización adicional, los alineamientos comprenden secuencias basadas en intrones.

En otra realización, los polinucleótidos son ADNc, que, por ejemplo, pueden ser de aproximadamente 500 a 5000 bases de longitud.

50 En otra realización más, los polinucleótidos son oligonucleótidos que, por ejemplo, pueden ser de aproximadamente 20 a 80 bases de longitud.

Por ejemplo, los alineamientos pueden inmovilizarse sobre vidrio, y pueden contener cientos de miles, por ejemplo,

330.000 oligonucleótidos.

En un aspecto adicional, la divulgación se refiere a un método para predecir la probabilidad de supervivencia a largo plazo de un paciente diagnosticado de cáncer de mama invasivo, sin recidiva de cáncer de mama, que comprende los pasos de

- 5 a) determinar los niveles de expresión de transcritos de ARN o productos de expresión de los genes de una serie de genes seleccionados del grupo compuesto por B_Catenin; BAG1; BIN1; BUB1; C20_orf1; CCNB1; CCNE2; CDC20; CDH1; CEGP1; CIAP1; cMYC; CTSL2; DKFZp586M07; DR5; EpCAM; EstRI; FOXM1; GRB7; GSTM1; GSTM3; HER2; HNRPAB; ID1; IGF1R; ITGA7; Ki_67; KNSL2; LMNB1; MCM2; MELK; MMP12; MMP9; MYBL2; NEK2; NME1; NPD009; PCNA; PR; PREP; PTTG1; RPLPO; Src; STK15; STMY3; SURV; TFRC; TOP2A; y TS en
 - 10 una célula de cáncer de mama obtenida del paciente, normalizada con respecto a los niveles de expresión de todos los transcritos de ARN o sus productos de expresión en dicha célula de cáncer de mama, o de una serie de referencia de transcritos de ARN o sus productos;
 - b) someter los datos obtenidos en el paso (a) a análisis estadístico; y
 - c) determinar si la probabilidad de dicha supervivencia a largo plazo ha aumentado o disminuido.

15 En un aspecto adicional, la divulgación se refiere a un método de preparación de un perfil genómico personalizado para un paciente, que comprende los pasos de

- a) someter el ARN extraído del tejido de mama obtenido del paciente a análisis de expresión génica;
- b) determinar el nivel de expresión en el tejido de uno o más genes seleccionados del grupo de genes de cáncer de mama enumerados en una cualquiera de las Tablas 1 y 2, en el que el nivel de expresión está normalizado con
 - 20 respecto a un gen o genes de control y se compara opcionalmente con el volumen hallado en un grupo de tejidos de referencia de cáncer de mama; y
 - c) crear un informe que resuma los datos obtenidos por dicho análisis de expresión génica.

El tejido de mama puede comprender células de cáncer de mama.

25 En otra realización, el tejido de mama se obtiene de una muestra de biopsia embebida en parafina, en la que el ARN puede estar fragmentado.

El informe puede incluir la predicción de probabilidad de supervivencia a largo plazo del paciente y/o la recomendación de una modalidad de tratamiento para dicho paciente.

30 En un aspecto adicional, la divulgación se refiere a un método para medir los niveles de productos de ARNm de los genes enumerados en las Tablas 1 y 2 por reacción en cadena de la polimerasa en tiempo real (RT-PCR), usando un amplicón enumerado en la Tabla y un grupo cebador-sonda enumerado en las Tablas 4A-4D.

En un aspecto adicional, la divulgación se refiere a un grupo cebador-sonda PCR enumerado en las Tablas 4A-4D, y un amplicón de PCR enumerado en la Tabla 3.

A. Definiciones

35 A menos que se defina otra cosa, los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el entendido comúnmente por un experto habitual en la técnica a la que pertenece esta invención. Singleton y col., Dictionary of Microbiology and Molecular Biology 2^a ed., J. Wiley & Sons (Nueva York, NY 1994), y March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4^a ed., John Wiley & Sons (Nueva York, NY 1992), proporcionan a un experto en la técnica orientación general para muchos de los términos usados en la presente solicitud.

40 Un experto en la técnica reconocerá muchos métodos y materiales similares o equivalentes a los descritos en el presente documento, que pueden usarse en la práctica de la presente invención. De hecho, la presente invención no se limita de ninguna manera a los métodos y materiales descritos. Para fines de la presente invención, a continuación se definen los siguientes términos.

45 El término "microalineamiento" se refiere a una disposición ordenada de elementos de alineamiento hibridables, preferiblemente sondas polinucleotídicas, sobre un sustrato.

El término "polinucleótido", cuando se usa en singular o en plural, se refiere de manera general a cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ARN o ADN sin modificar o ARN o ADN modificados. Por lo tanto, por ejemplo, los polinucleótidos como se definen en el presente documento incluyen, sin limitación, ADN mono y bicatenario, ADN que incluye regiones mono y bicatenarias, ARN mono y bicatenario, y ARN que

- 50 incluye regiones mono y bicatenarias, moléculas híbridas que comprenden ADN y ARN que pueden ser monocatenarios o, más normalmente, bicatenarios o incluyen regiones mono y bicatenarias. Además, el término "polinucleótido" como se usa en el presente documento se refiere a regiones tricatenarias que comprenden ARN o

ADN o tanto ARN como ADN. Las cadenas en tales regiones pueden ser de la misma molécula o de moléculas diferentes. Las regiones pueden incluir la totalidad de una o más de las moléculas, pero más normalmente implican sólo una región de algunas de las moléculas. Con frecuencia, una de las moléculas de una región de triple hélice es un oligonucleótido. El término "polinucleótido" incluye específicamente ADNc. El término incluye ADN (incluyendo ADNc) y ARN que contienen una o más bases modificadas. Por lo tanto, los ADN o ARN con estructuras principales modificadas por estabilidad o por otros motivos son "polinucleótidos" como está previsto ese término en el presente documento. Además, dentro del término "polinucleótidos" como se define en el presente documento se incluyen ADN o ARN que comprenden bases no habituales, tales como inosina, o bases modificadas, tales como bases tritiadas. En general, el término "polinucleótido" abarca todas las formas química, enzimática y/o metabólicamente modificadas de polinucleótidos no modificados; así como las formas químicas de ADN y ARN característicos de virus y células, incluyendo células simples y complejas.

El término "oligonucleótido" se refiere a un polinucleótido relativamente corto, incluyendo, sin limitación, desoxirribonucleótidos monocatenarios, ribonucleótidos mono o bicatenarios, híbridos ARN:ADN y ADN bicatenarios. Con frecuencia los oligonucleótidos, tales como oligonucleótidos de sondas de ADN monocatenarios, se sintetizan mediante métodos químicos, por ejemplo usando sintetizadores de oligonucleótidos automatizados que están disponibles en el mercado. Sin embargo, pueden prepararse oligonucleótidos mediante una variedad de otros métodos, incluyendo técnicas mediadas por ADN recombinante *in vitro* y mediante la expresión de ADN en células y organismos.

Los términos "gen expresado de manera diferencial", "expresión génica diferencial" y sus sinónimos, que se usan de manera intercambiable, se refieren a un gen cuya expresión está activada hasta un nivel superior o inferior en un sujeto que padece una enfermedad, específicamente cáncer, tal como cáncer de mama, con respecto a su expresión en un sujeto normal o control. Los términos también incluyen genes cuya expresión está activada hasta un nivel superior o inferior en diferentes estadios de la misma enfermedad. También se entiende que un gen expresado de manera diferencial puede estar o bien activado o bien inhibido a nivel de ácido nucleico o a nivel de proteína, o puede someterse a corte y empalme alternativo para dar como resultado un producto polipeptídico diferente. Tales diferencias pueden demostrarse mediante un cambio en los niveles de ARNm, expresión superficial, secreción u otro reparto de un polipéptido, por ejemplo. La expresión génica diferencial puede incluir una comparación de la expresión entre dos o más genes o sus productos génicos, o una comparación de las razones de la expresión entre dos o más genes o sus productos génicos, o incluso una comparación de dos productos procesados de manera diferente del mismo gen, que se diferencian entre sujetos normales y sujetos que padecen una enfermedad, específicamente cáncer, o entre diversos estadios de la misma enfermedad. La expresión diferencial incluye diferencias tanto cuantitativas como cualitativas, diferencias en el patrón de expresión temporal o celular en un gen o sus productos de expresión, por ejemplo, entre células normales y enfermas, o entre células que se han sometido a diferentes acontecimientos de enfermedad o estadios de enfermedad. Para el fin de esta invención, se considera que está presente "expresión génica diferencial" cuando hay una diferencia de al menos aproximadamente dos veces, preferiblemente al menos aproximadamente cuatro veces, más preferiblemente al menos aproximadamente seis veces, lo más preferiblemente al menos aproximadamente diez veces entre la expresión de un gen dado en sujetos normales y enfermos, o en diversos estadios del desarrollo de la enfermedad en un sujeto enfermo.

El término "sobrexpresión" con respecto a un transcrito de ARN se usa para hacer referencia al nivel del transcrito determinado mediante normalización con respecto al nivel de ARNm de referencia, que pueden ser todos transcritos medidos en la muestra o un conjunto de referencia particular de ARNm.

La expresión "amplificación génica" se refiere a un proceso mediante el cual se forman múltiples copias de un gen o fragmento génico en una célula o línea celular particular. Con frecuencia la región duplicada (un tramo de ADN amplificado) se denomina "amplión". Habitualmente, la cantidad del ARN mensajero (ARNm) producido, es decir, el nivel de expresión génica, también aumenta en la proporción del número de copias realizadas del gen particular expresado.

El término "pronóstico" se usa en el presente documento para hacer referencia a la predicción de la probabilidad de muerte atribuible al cáncer o progresión, incluyendo recidiva, propagación metastásica y resistencia a fármacos, de una enfermedad neoplásica, tal como cáncer de mama. El término "predicción" se usa en el presente documento para hacer referencia a la probabilidad de que un paciente responda o bien favorablemente o bien desfavorablemente a un fármaco o conjunto de fármacos, y también el grado de esas respuestas, o de que un paciente sobreviva, tras extirpación quirúrgica del tumor primario y/o quimioterapia durante un determinado periodo de tiempo sin recidiva de cáncer. Los métodos predictivos de la presente invención pueden usarse clínicamente para tomar decisiones de tratamiento eligiendo las modalidades de tratamiento más apropiadas para cualquier paciente particular. Los métodos predictivos de la presente invención son herramientas valiosas en la predicción de si es probable que un paciente responda favorablemente a un régimen de tratamiento, tal como intervención quirúrgica, quimioterapia con un fármaco o combinación de fármacos dados, y/o radioterapia, o si es probable la supervivencia a largo plazo del paciente, tras cirugía y/o terminación de quimioterapia u otras modalidades de tratamiento.

El término supervivencia "a largo plazo" se usa en el presente documento para hacer referencia a la supervivencia

durante al menos 3 años, más preferiblemente durante al menos 8 años, mucho más preferiblemente durante al menos 10 años tras la cirugía u otro tratamiento.

El término "tumor", como se usa en el presente documento, se refiere a todos los crecimientos y proliferaciones de células neoplásicas, ya sean malignos o benignos, y todas las células y tejidos precancerosos y cancerosos.

- 5 Los términos "cáncer" y "canceroso" se refieren a o describen la afección fisiológica en mamíferos que se caracteriza normalmente por un crecimiento celular no regulado. Los ejemplos de cáncer incluyen, pero sin limitación, cáncer de mama, cáncer de ovario, cáncer de colon, cáncer de pulmón, cáncer de próstata, cáncer hepatocelular, cáncer gástrico, cáncer pancreático, cáncer de cuello uterino, cáncer de hígado, cáncer de vejiga, cáncer de las vías urinarias, cáncer de tiroides, cáncer renal; carcinoma, melanoma y cáncer de cerebro.
- 10 La "patología" del cáncer incluye todos los fenómenos que ponen en peligro el bienestar del paciente. Esto incluye, sin limitación, crecimiento celular anómalo o no controlable, metástasis, interferencia con el funcionamiento normal de células vecinas, liberación de citocinas u otros productos de secreción a niveles anómalos, supresión o agravamiento de la respuesta inflamatoria o inmunológica, neoplasia, tumores premalignos, tumores malignos, invasión de tejidos u órganos circundantes o distantes, tales como ganglios linfáticos, etc.

- 15 La "rigurosidad" de reacciones de hibridación puede determinarlas fácilmente un experto habitual en la técnica, y generalmente es un cálculo empírico que depende de la longitud de la sonda, la temperatura de lavado y la concentración de sal. En general, las sondas más largas requieren temperaturas superiores para un apareamiento apropiado, mientras que sondas más cortas necesitan temperaturas inferiores. La hibridación depende generalmente de la capacidad del ADN desnaturalizado para volver a aparearse cuando están presentes cadenas complementarias en un entorno por debajo de su temperatura de fusión. Cuanto mayor es el grado de homología deseada entre la sonda y la secuencia hibridable, mayor es la temperatura relativa que puede usarse. Como resultado, se desprende que temperaturas relativas superiores tenderán a hacer que las condiciones de reacción sean más rigurosas, mientras que temperaturas inferiores las vuelven menos rigurosas. Para detalles adicionales y una explicación de la rigurosidad de reacciones de hibridación, véase Ausubel y col., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).
- 20
- 25

- Las "condiciones rigurosas" o "condiciones de alta rigurosidad", como se definen en el presente documento, normalmente: (1) emplean una baja fuerza iónica y una alta temperatura para el lavado, por ejemplo cloruro sódico 0,015 M/citrato sódico 0,0015 M/dodecilsulfato sódico al 0,1% a 50 °C; (2) emplean durante la hibridación un agente desnaturalizante, tal como formamida, por ejemplo, formamida al 50 % (v/v) con albúmina sérica bovina al 0,1%/Ficoll al 0,1%/polivinilpirrolidona al 0,1%/tampón fosfato sódico 50 mM a pH 6,5 con cloruro sódico 750 mM, citrato sódico 75 mM a 42 °C; o (3) emplean formamida al 50 %, 5 x SSC (NaCl 0,75 M, citrato sódico 0,075 M), fosfato sódico 50 mM (pH 6,8), pirofosfato sódico al 0,1%, solución de Denhardt 5 x, ADN de esperma de salmón sonificado (50 µg/ml), SDS al 0,1% y sulfato de dextrano al 10 % a 42 °C, con lavados a 42 °C en SSC 0,2 x (cloruro sódico/citrato sódico) y formamida al 50 %, seguido de un lavado de alta rigurosidad que consiste en SSC 0,1 x que contiene EDTA a 55 °C.
- 30
- 35

- Pueden identificarse "condiciones moderadamente rigurosas" como se describe por Sambrook y col., Molecular Cloning: A Laboratory Manual, Nueva York: Cold Spring Harbor Press, 1989, e incluyen el uso de una solución de lavado y condiciones de hibridación (por ejemplo, temperatura, fuerza iónica y % de SDS) menos rigurosas que las descritas anteriormente. Un ejemplo de condiciones moderadamente rigurosas es incubación durante la noche a 37 °C en una solución que comprende: formamida al 20 %, SSC 5 x (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7,6), solución de Denhardt 5 x, sulfato de dextrano al 10 % y 20 mg/ml de ADN de esperma de salmón fragmentado desnaturalizado, seguido de lavado de los filtros en SSC 1 x a aproximadamente 37-50 °C. El experto en la técnica reconocerá cómo ajustar la temperatura, la fuerza iónica, etc. según sea necesario para adaptarse a factores tales como la longitud de la sonda y similares.
- 40

- 45 En el contexto de la presente invención, la referencia a "al menos uno", "al menos dos", "al menos cinco", etc. de los genes enumerados en cualquier conjunto de genes particular significa uno cualquiera o todas y cada una de las combinaciones de los genes enumerados.

El término cáncer "de ganglios negativos", tal como cáncer de mama "de ganglios negativos", se usa en el presente documento para hacer referencia a cáncer que no se ha propagado a los ganglios linfáticos.

- 50 Los términos "corte y empalme" y "corte y empalme de ARN" se usan de manera intercambiable y se refieren a procesamiento de ARN que elimina intrones y une exones para producir ARNm maduro con secuencia codificante continua que se mueve en el citoplasma de una célula eucariota.

- En teoría, el término "exón" se refiere a cualquier segmento de un gen interrumpido que está representado en el producto de ARN maduro (B. Lewin. Genes IV Cell Press, Cambridge Mass. 1990). En teoría el término "intrón" se refiere a cualquier segmento de ADN que se transcribe pero se elimina del interior del transcrito mediante corte y empalme entre sí de los exones a ambos lados del mismo. Operacionalmente, se producen secuencias de exones
- 55

en la secuencia de ARNm de un gen como se define por los números de SEQ ID de referencia. Operacionalmente, las secuencias de intrones son las secuencias intermedias dentro del ADN genómico de un gen, intercaladas entre secuencias de exones y que tienen secuencias consenso de corte y empalme GT y AG en sus límites en 5' y 3'.

B. Descripción detallada

5 La práctica de la presente invención empleará, a menos que se indique lo contrario, técnicas convencionales de biología molecular (incluyendo técnicas recombinantes), microbiología, biología celular y bioquímica, que están dentro de la experiencia de la técnica. Tales técnicas se explican completamente en la bibliografía, tales como, "Molecular Cloning: A Laboratory Manual", 2ª edición (Sambrook y col., 1989); "Oligonucleotide Synthesis" (M.J. Gait, ed., 1984); "Animal Cell Culture" (R.I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.);
10 "Handbook of Experimental Immunology", 4ª edición (D.M. Weir & C.C. Blackwell, eds., Blackwell Science Inc., 1987); "Gene Transfer Vectors for Mammalian Cells" (J.M. Miller & M.P. Calos, eds., 1987); "Current Protocols in Molecular Biology" (F.M. Ausubel y col., eds., 1987); y "PCR: The Polymerase Chain Reaction", (Mullis y col., eds., 1994).

1. Obtención del perfil de expresión génica

15 Los métodos de obtención del perfil de expresión génica incluyen métodos basados en análisis de hibridación de polinucleótidos, métodos basados en secuenciación de polinucleótidos y métodos basados en proteómica. Los métodos más comúnmente usados conocidos en la técnica para la cuantificación de la expresión de ARNm en una muestra incluyen transferencia de tipo Northern e hibridación *in situ* (Parker & Barnes, Methods in Molecular Biology 106: 247-283 (1999)); ensayos de protección de ARNasa (Hod, Biotechniques 13: 852-854 (1992)); y métodos
20 basados en PCR, tales como reacción en cadena de la polimerasa con transcripción inversa (RT-PCR) (Weis y col., Trends in Genetics 8: 263-264 (1992)). Como alternativa, pueden emplearse anticuerpos que pueden reconocer dúplex específicos, incluyendo dúplex de ADN, dúplex de ARN y dúplex híbridos de ADN-ARN o dúplex de ADN-proteína. Los métodos representativos para el análisis de la expresión génica basado en secuenciación incluyen análisis en serie de la expresión génica (SAGE), y análisis de la expresión génica mediante secuenciación de firma
25 masiva en paralelo (MPSS).

2. Métodos de obtención del perfil de expresión génica basados en PCR

a. PCR con transcriptasa inversa (RT-PCR)

De las técnicas enumeradas anteriormente, el método cuantitativo más sensible y más flexible es la RT-PCR, que puede usarse para comparar niveles de ARNm en diferentes poblaciones de muestras, en tejidos normales y
30 tumorales, con o sin tratamiento farmacológico, para caracterizar patrones de expresión génica, para distinguir entre ARNm estrechamente relacionados y para analizar la estructura del ARN.

La primera etapa es el aislamiento de ARNm de una muestra diana. El material de partida es normalmente ARN total aislado de tumores o líneas celulares tumorales humanos, y tejidos o líneas celulares normales correspondientes, respectivamente. Por lo tanto, puede aislarse ARN de una variedad de tumores primarios, incluyendo tumor o líneas
35 celulares tumorales de mama, pulmón, colon, próstata, cerebro, hígado, riñón, páncreas, bazo, timo, testículos, ovario, útero, etc., con ADN combinado de donantes sanos. Si la fuente de ARNm es un tumor primario, puede extraerse ARNm, por ejemplo, de muestras de tejido incluidas en parafina congeladas o archivadas, y fijadas (por ejemplo fijadas en formalina).

En la técnica se conocen bien métodos generales para la extracción de ARNm y se desvelan en libros de texto convencionales de biología molecular, incluyendo Ausubel y col., Current Protocols of Molecular Biology, John Wiley and Sons (1997). Se desvelan métodos para la extracción de ARN a partir de tejidos incluidos en parafina, por
40 ejemplo, en Rupp y Locker, Lab Invest. 56: A67 (1987), y De Andrés y col., BioTechniques 18: 42044 (1995). En particular, el aislamiento de ARN puede realizarse usando un kit de purificación, conjunto de tampones y proteasa de fabricantes comerciales, tales como Qiagen, según las instrucciones del fabricante. Por ejemplo, puede aislarse ARN total a partir de células en cultivo usando mini-columnas RNeasy de Qiagen. Otros kits de aislamiento de ARN disponibles en el mercado incluyen el kit de purificación de ADN y ARN completo MasterPure™ (EPICENTRE®,
45 Madison, WI) y el kit de aislamiento de ARN de bloque de parafina (Ambion, Inc.). Puede aislarse el ARN total de muestras de tejido usando RNA Stat-60 (Tel-Test). El ARN preparado a partir del tumor puede aislarse, por ejemplo, mediante centrifugación en gradiente de densidad de cloruro de cesio.

50 Dado que el ARN no puede servir como molde para la PCR, la primera etapa en la obtención del perfil de expresión génica mediante RT-PCR es la transcripción inversa del molde de ARN para dar ADNc, seguido de su amplificación exponencial en una reacción de PCR. Las dos transcriptasas inversas más comúnmente usadas son transcriptasa inversa del virus de la mieloblastosis aviar (AMV-RT) y la transcriptasa inversa del virus de la leucemia murina de Moloney (MMLV-RT). La etapa de transcripción inversa se ceba normalmente usando cebadores específicos,
55 hexámeros al azar o cebadores de oligo-dT, dependiendo de las circunstancias y del objetivo de la obtención del perfil de expresión. Por ejemplo, ARN extraído puede someterse a transcripción inversa usando un kit de PCR de

ARN de GeneAmp (Perkin Elmer, CA, Estados Unidos), siguiendo las instrucciones del fabricante. Después, puede usarse el ADNc derivado como molde en la reacción de PCR posterior.

Aunque la etapa de PCR puede usar una variedad de ADN polimerasas dependientes de ADN termoestables, normalmente emplea la ADN polimerasa de Taq, que tiene una actividad nucleasa en 5'-3' pero carece de actividad endonucleasa de corrección en 3'-5'. Por lo tanto, la PCR TaqMan® utiliza normalmente la actividad nucleasa en 5' de polimerasa de Taq o Tth para hidrolizar una sonda de hibridación unida a su amplicón diana, pero puede usarse cualquier enzima con actividad nucleasa en 5' equivalente. Se usan dos cebadores oligonucleotídicos para generar un amplicón típico de una reacción de PCR. Se diseña un tercer oligonucleótido, o sonda, para detectar una secuencia de nucleótidos ubicada entre los dos cebadores de PCR. La sonda no puede extenderse mediante la enzima ADN polimerasa de Taq, y está marcada con un tinte fluorescente indicador y un tinte fluorescente extintor. Cualquier emisión inducida por láser del tinte indicador se extingue por el tinte extintor cuando los dos tintes están ubicados cerca uno del otro como lo están en la sonda. Durante la reacción de amplificación, la enzima ADN polimerasa de Taq escinde la sonda de una manera dependiente del molde. Los fragmentos de sonda resultantes se disocian en solución y la señal del tinte indicador liberado está libre del efecto extintor del segundo fluoróforo. Se libera una molécula de tinte indicador por cada nueva molécula sintetizada y la detección del tinte indicador no extinguido proporciona la base para la interpretación cuantitativa de los datos.

La RT-PCR TaqMan® puede realizarse usando equipo disponible en el mercado, tal como, por ejemplo, ABI PRISM 7700™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, CA, Estados Unidos), o Lightcycler (Roche Molecular Biochemicals, Mannheim, Alemania). En una realización preferida, el procedimiento de nucleasa en 5' se realiza en un dispositivo de PCR cuantitativa en tiempo real tal como el ABI PRISM 7700™ Sequence Detection System™. El sistema consiste en un termociclador, un láser, un dispositivo de carga acoplada (CCD), una cámara y un ordenador. El sistema amplifica muestras en un formato de 96 pocillos en un termociclador. Durante la amplificación, se recoge en tiempo real una señal fluorescente inducida por láser a través de cables de fibra óptica para los 96 pocillos y se detecta en el CCD. El sistema incluye software para hacer funcionar el instrumento y para analizar los datos.

Los datos del ensayo de nucleasa en 5' se expresan inicialmente como Ct, o el ciclo umbral. Como se ha analizado anteriormente, se registran los valores de fluorescencia durante cada ciclo y representan la cantidad de producto amplificado hasta ese punto en la reacción de amplificación. El punto en el que se registra por primera vez que la señal fluorescente es estadísticamente significativa es el ciclo umbral (Ct).

Para minimizar errores y el efecto de la variación de muestra a muestra, la RT-PCR se realiza habitualmente usando un patrón interno. El patrón interno ideal se expresa a un nivel constante entre diferentes tejidos y no se ve afectado por el tratamiento experimental. Los ARN usados con mayor frecuencia para normalizar patrones de expresión génica son ARNm para los genes de mantenimiento gliceraldehído-3-fosfato-deshidrogenasa (GAPDH) y β -actina.

Una variación más reciente de la técnica de RT-PCR es la PCR cuantitativa en tiempo real, que mide la acumulación de producto de PCR a través de una sonda fluorogénica con doble marcador (es decir, sonda TaqMan®). La PCR en tiempo real es compatible tanto con la PCR de competencia cuantitativa, en la que se usa un competidor interno para cada secuencia diana para la normalización, como con PCR comparativa cuantitativa que usa un gen de normalización contenido dentro de la muestra o un gen de mantenimiento para RT-PCR. Para más detalles, véase, por ejemplo Held y col., *Genome Research* 6: 986-994 (1996).

Las etapas de un protocolo representativo para obtener el perfil de expresión génica usando tejidos incluidos en parafina, fijados, como fuente de ARN, incluyendo aislamiento de ARNm, purificación, extensión por cebadores y amplificación, se facilitan en diversos artículos de revistas publicados {por ejemplo: T.E. Godfrey y col. *J. Molec. Diagnostics* 2: 84-91 [2000]; K. Specht y col., *Am. J. Pathol.* 158: 419-29 [2001]}. En resumen, un proceso representativo comienza con cortar secciones de aproximadamente 10 μ m de grosor de muestras de tejido tumoral incluidas en parafina. Después, se extrae el ARN y se retiran las proteínas y el ADN. Tras el análisis de la concentración de ARN, pueden incluirse etapas de reparación y/o amplificación de ARN, si es necesario, y se somete el ARN a transcripción inversa usando promotores específicos del gen seguido por RT-PCR.

b. Sistema MassARRAY

En el método de obtención del perfil de expresión génica basado en MassARRAY, desarrollado por Sequenom, Inc. (San Diego, CA) tras el aislamiento de ARN y transcripción inversa, al ADNc obtenido se le añaden cantidades conocidas de una molécula de ADN sintético (competidor), que se corresponde con la región de ADNc seleccionada como diana en todas las posiciones, excepto en una única base, y sirve como patrón interno. Se amplifica mediante PCR la mezcla de ADNc/competidor y se somete a un tratamiento con enzima fosfatasa alcalina de gamba (SAP) tras la PCR, que da como resultado la desfosforilación de los nucleótidos restantes. Tras la inactivación de la fosfatasa alcalina, los productos de PCR del competidor y el ADNc se someten a extensión por cebador, que genera señales de masa diferenciadas para los productos de PCR derivados de competidor y de ADNc. Después de la purificación, estos productos se dispensan en un alineamiento en chip, que está precargado con los componentes

necesarios para el análisis con análisis de espectrometría de masas de tiempo de vuelo con ionización por desorción láser asistida por matriz (MALDI-TOF MS). El ADNc presente en la reacción se cuantifica entonces analizando las relaciones de las áreas de pico en el espectro de masas generado. Para detalles adicionales véase, por ejemplo Ding y Cantor, Proc. Natl. Acad. Sci. USA 100: 3059-3064 (2003).

5 c. Otros métodos basados en PCR

Las técnicas basadas en PCR adicionales incluyen, por ejemplo, presentación diferencial (Liang y Pardee, Science 257: 967-971 (1992)); polimorfismo de longitud de fragmentos amplificados (iAFLP) (Kawamoto y col., Genome Res. 12: 1305-1312 (1999)); tecnología BeadArray™ (Illumina, San Diego, CA; Oliphant y col., Discovery of Markers for Disease (suplemento a Biotechniques), junio de 2002; Ferguson y col., Analytical Chemistry 72: 5618 (2000)); BeadsArray para la detección de expresión génica (BADGE), usando el sistema Luminex100 LabMAP disponible en el mercado y microesferas codificadas por múltiples colores (Luminex Corp., Austin, TX) en un ensayo rápido para determinar la expresión génica (Yang y col., Genome Res. 11: 1888-1898 (2001)); y análisis de obtención de perfiles de expresión de alta cobertura (HiCEP) (Fukumura y col., Nucl. Acids. Res. 31(16) e94 (2003)).

3. Microalineamientos

15 También puede identificarse o confirmarse la expresión génica diferencial usando la técnica de microalineamientos. Por lo tanto, puede medirse el perfil de expresión de genes asociados con cáncer de mama en tejido tumoral o bien reciente o bien incluido en parafina, usando tecnología de microalineamientos. En este método, se recubren o disponen en alineamiento secuencias de polinucleótido de interés (incluyendo ADNc y oligonucleótidos) sobre un sustrato de microchip. Después, las secuencias dispuestas en alineamiento se hibridan con sondas de ADN
20 específicas de células o tejidos de interés. Al igual que en el método de RT-PCR, la fuente de ARNm es normalmente ARN total aislado de tumores o líneas de células tumorales humanas, y tejidos o líneas celulares normales correspondientes. Por lo tanto, puede aislarse ARN de una variedad de tumores primarios o líneas de células tumorales. Si la fuente de ARNm es un tumor primario, puede extraerse ARNm, por ejemplo, de muestras de tejido incluidas en parafina congeladas o archivadas, y fijadas (por ejemplo fijadas en formalina), que se preparan y se conservan de manera rutinaria en la práctica clínica diaria.

En una realización específica de la técnica de microalineamientos, se aplican insertos amplificados por PCR de clones de ADNc a un sustrato en un alineamiento denso. Preferiblemente se aplican al menos 10.000 secuencias de nucleótidos al sustrato. Los genes dispuestos en microalineamiento, inmovilizados sobre el microchip a 10.000 elementos cada uno, son adecuados para su hibridación en condiciones rigurosas. Pueden generarse sondas de
30 ADNc marcadas con fluorescencia mediante incorporación de nucleótidos fluorescentes mediante transcripción inversa de ARN extraído a partir de tejidos de interés. Las sondas de ADNc marcadas aplicadas al chip se hibridan con especificidad a cada punto de ADN en el alineamiento. Tras un lavado riguroso para eliminar sondas unidas de manera no específica, se explora el chip mediante microscopía confocal láser o mediante otro método de detección, tal como una cámara CCD. La cuantificación de la hibridación de cada elemento dispuesto en alineamiento permite
35 la evaluación de la abundancia del ARNm correspondiente. Con fluorescencia de doble color, se hibridan por parejas con el alineamiento sondas de ADNc marcadas por separado generadas a partir de dos fuentes de ARN. Por lo tanto, se determina simultáneamente la abundancia relativa de los transcritos de las dos fuentes correspondientes a cada gen especificado. La escala miniaturizada de la hibridación proporciona una evaluación rápida y conveniente del patrón de expresión para grandes números de genes. Se ha mostrado que tales métodos tienen la sensibilidad
40 requerida para detectar transcritos poco frecuentes, que se expresan a pocas copias por célula, y para detectar de manera reproducible diferencias de al menos aproximadamente dos veces en los niveles de expresión (Schna y col., Proc. Natl. Acad. Sci. USA 93(2): 106-149 (1996)). Pueden realizarse análisis de microalineamiento mediante el equipo disponible en el mercado, siguiendo los protocolos del fabricante, tal como usando la tecnología GenChip de Affymetrix o la tecnología de microalineamiento de Incyte.

45 El desarrollo de métodos de microalineamiento para el análisis a gran escala de la expresión génica hace posible buscar de manera sistemática marcadores moleculares de clasificación del cáncer y predicción del desenlace en una variedad de tipos de tumores.

4. Análisis en serie de la expresión génica (SAGE)

El análisis en serie de la expresión génica (SAGE) es un método que permite el análisis simultáneo y cuantitativo de un gran número de transcritos génicos, sin necesidad de proporcionar una sonda de hibridación individual para cada
50 transcrito. En primer lugar, se genera una etiqueta de secuencia corta (aproximadamente 10-14 pb) que contiene información suficiente para identificar de manera única un transcrito, siempre que se obtenga la etiqueta de una posición única dentro de cada transcrito. Después, se unen entre sí muchos transcritos para formar moléculas en serie largas, que pueden secuenciarse, revelando simultáneamente la identidad de las etiquetas múltiples. El patrón de expresión de cualquier población de transcritos puede evaluarse cuantitativamente determinando la abundancia de etiquetas individuales e identificando el gen correspondiente para cada etiqueta. Para más detalles véase, por
55 ejemplo, Velculescu y col., Science 270: 484-487 (1995); y Velculescu y col., Cell 88: 243-51 (1997).

5. Análisis de la expresión génica mediante secuenciación de firma masiva en paralelo (MPSS)

Este método, descrito por Brenner y col., Nature Biotechnology 18: 630-634 (2000), es un enfoque de secuenciación que combina la secuenciación de firma no basada en gel con la clonación *in vitro* de millones de moldes sobre microperlas separadas de 5 µm de diámetro. En primer lugar, se construye una biblioteca de microperlas de moldes de ADN mediante clonación *in vitro*. Esto va seguido por el montaje de un alineamiento plano de las microperlas que contienen moldes en una célula de flujo a alta densidad (normalmente superior a 3×10^6 microperlas/cm²). Se analizan simultáneamente los extremos libres de los moldes clonados sobre cada microperla, usando un método de secuenciación de firma basado en fluorescencia que no requiere separación de fragmentos de ADN. Se ha mostrado que este método proporciona simultáneamente y con precisión, en una única operación, cientos de miles de secuencias de firma a partir de una biblioteca de ADNc de levadura.

6. Inmunohistoquímica

Los métodos inmunohistoquímicos también son adecuados para detectar los niveles de expresión de los marcadores de pronóstico de la presente invención. Por lo tanto, se usan anticuerpos o antisueros, preferiblemente anticuerpos policlonales, y mucho más preferiblemente anticuerpos monoclonales específicos para cada marcador, para detectar la expresión. Los anticuerpos pueden detectarse mediante marcaje directo de los propios anticuerpos, por ejemplo, con marcadores radiactivos, marcadores fluorescentes, marcadores de hapteno tales como biotina, o una enzima tal como peroxidasa de rábano picante o fosfatasa alcalina. Como alternativa, se usa un anticuerpo primario no marcado junto con un anticuerpo secundario marcado, comprendiendo antisueros, anticuerpos policlonales o un anticuerpo monoclonal específico para el anticuerpo primario. En la técnica se conocen bien protocolos y kits de inmunohistoquímica y están disponibles en el mercado.

7. Proteómica

El término "proteoma" se define como la totalidad de las proteínas presentes en una muestra (por ejemplo tejido, organismo o cultivo celular) en un determinado punto de tiempo. La proteómica incluye, entre otras cosas, el estudio de los cambios globales en la expresión de proteínas en una muestra (también denominada "proteómica de expresión"). La proteómica incluye normalmente las siguientes etapas: (1) separación de proteínas individuales en una muestra mediante electroforesis bidimensional (PAGE 2-D); (2) identificación de las proteínas individuales recuperadas del gel, por ejemplo mediante espectrometría de masas o secuenciación N-terminal y (3) análisis de los datos usando bioinformática. Los métodos de proteómica son complementos valiosos para otros métodos de obtención del perfil de expresión génica, y pueden usarse, solos o en combinación con otros métodos, para detectar los productos de los marcadores de pronóstico de la presente invención.

8. Descripción general del aislamiento, la purificación y la amplificación de ARNm

Las etapas de un protocolo representativo para obtener el perfil de expresión génica usando tejidos incluidos en parafina, fijados, como fuente de ARN, incluyendo aislamiento, purificación, extensión por cebadores y amplificación de ARNm, se proporcionan en diversos artículos de revistas publicados (por ejemplo: T.E. Godfrey et al., J. Molec. Diagnostics 2: 84-91 [2000]; K. Specht y col., Am. J. Pathol. 158: 419-29 [2001]). En resumen, un proceso representativo comienza con cortar secciones de aproximadamente 10 µm de grosor de muestras de tejido tumoral incluidas en parafina. Después, se extrae el ARN y se retiran las proteínas y el ADN. Tras el análisis de la concentración de ARN, pueden incluirse etapas de reparación y/o amplificación de ARN, si es necesario, y se somete el ARN a transcripción inversa usando promotores específicos del gen seguido por RT-PCR. Finalmente, se analizan los datos para identificar la mejor o mejores opciones de tratamiento disponibles para el paciente basándose en el patrón de expresión génica característico identificado en la muestra tumoral examinada, dependiendo de la probabilidad predicha de recidiva de cáncer.

9. Conjunto de genes de cáncer de mama, subsecuencias de genes ensayados y aplicación clínica de los datos de expresión génica

Un aspecto importante de la presente invención es usar la expresión medida de determinados genes por tejido de cáncer de mama para proporcionar información de pronóstico. Para ello, es necesario corregir (separar por normalización) tanto diferencias en la cantidad de ARN sometido a ensayo como variabilidad en la calidad del ARN usado. Por lo tanto, normalmente el ensayo mide e incorpora la expresión de determinados genes de normalización, incluyendo genes de mantenimiento bien conocidos, tales como GAPDH y Cyp1. Como alternativa, la normalización puede basarse en la media o la señal media (Ct) de todos los genes sometidos a ensayo o un gran subconjunto de los mismos (enfoque de normalización global). Basándose en cada gen, se compara la cantidad normalizada medida del ARNm tumoral de un paciente con la cantidad encontrada en un conjunto de referencia de tejido de cáncer de mama. El número (N) de tejidos de cáncer de mama en este conjunto de referencia debe ser lo suficientemente alto como para garantizar que diferentes conjuntos de referencia (en su totalidad) se comportan básicamente de la misma manera. Si se cumple esta condición, la identidad de los tejidos de cáncer de mama individuales presentes en un conjunto particular no tendrá ningún impacto significativo sobre las cantidades relativas de los genes sometidos a ensayo. Habitualmente, el conjunto de referencia de tejido de cáncer de mama consiste en al menos

aproximadamente 30, preferiblemente al menos aproximadamente 40 muestras de tejido de cáncer de mama FPE diferentes. A menos que se indique lo contrario, los niveles de expresión normalizados para cada ARNm/tumor sometido a prueba/paciente se expresarán como porcentaje del nivel de expresión medido en el conjunto de referencia. Más específicamente, el conjunto de referencia con un número suficientemente alto (por ejemplo 40) de tumores proporciona una distribución de niveles normalizados de cada especie de ARNm. El nivel medido en una muestra tumoral particular que va a analizarse se encuentra en algún percentil dentro de este intervalo, que puede determinarse mediante métodos bien conocidos en la técnica. A continuación, a menos que se indique otra cosa, la referencia a niveles de expresión de un gen supone la expresión normalizada con respecto al conjunto de referencia aunque no siempre se mencione esto de manera explícita.

10. Diseño de cebadores y sondas de PCR basados en intrones

De acuerdo con un aspecto de la presente invención, se diseñan cebadores y sondas de PCR basándose en secuencias de intrones presentes en el gen que va a amplificarse. Por consiguiente, la primera etapa en el diseño de cebadores/sondas es la delineación de secuencias de intrones dentro de los genes. Esto puede realizarse mediante software públicamente disponible, tal como el software DNA BLAT desarrollado por Kent, W.J., *Genome Res.* 12(4): 656-64 (2002), o mediante el software BLAST, incluyendo sus variaciones. Las etapas posteriores siguen métodos bien establecidos de diseño de cebadores y sondas de PCR.

Con el fin de evitar señales no específicas, es importante enmascarar secuencias de repetición dentro de los intrones cuando se diseñan los cebadores y las sondas. Esto puede lograrse fácilmente usando el programa Repeat Masker disponible en línea a través del Baylor College of Medicine, que examina secuencias de ADN frente a una biblioteca de elementos de repetición y devuelve una secuencia de consulta en la que los elementos de repetición están enmascarados. Después, pueden usarse las secuencias de intrones enmascaradas para diseñar secuencias de cebadores y sondas usando cualquier paquete de diseño de cebadores/sondas disponible en el mercado o públicamente de otro modo, tales como Primer Express (Applied Biosystems); MGB assay-by-design (Applied Biosystems); Primer3 (Steve Rozen y Helen J. Skaletsky (2000) Primer3 en Internet para usuarios generales y para programadores biólogos. En: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, págs. 365-386).

Los factores más importantes considerados en el diseño de cebadores de PCR incluyen longitud de cebador, temperatura de fusión (T_f) y contenido de G/C, especificidad, secuencias de cebador complementarias y secuencia en el extremo 3'. En general, los cebadores de PCR óptimos tienen generalmente 17-30 bases de longitud y contienen aproximadamente el 20-80 %, tal como, por ejemplo, aproximadamente el 50-60 % de bases G+C. Normalmente se prefieren T_f de entre 50 y 80 °C, por ejemplo de aproximadamente 50 a 70 °C.

Para directrices adicionales para el diseño de cebadores y sondas de PCR véase, por ejemplo Dieffenbach, C.W. y col., "General Concepts for PCR Primer Design" en: *PCR Primer, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Nueva York, 1995, págs. 133-155; Innis y Gelfand, "Optimization of PCRs" en: *PCR Protocols, A Guide to Methods and Applications*, CRC Press, Londres, 1994, págs. 5-11; y Plasterer, T.N. *Primerselect: Primer and probe design. Methods Mol. Biol.* 70: 520-527 (1997).

En el siguiente Ejemplo no limitante se describirán detalles adicionales de la invención.

Ejemplo

Un estudio de fase II de la expresión génica en 242 tumores de mama malignos

Se diseñó y se realizó un estudio de la expresión génica con el objetivo principal de caracterizar molecularmente la expresión génica en muestras de tejido incluidas en parafina, fijadas, de carcinoma ductal de mama invasivo, y explorar la correlación entre tales perfiles moleculares y la supervivencia libre de enfermedad.

Diseño del estudio

Se realizaron ensayos moleculares con tejidos de tumor de mama primario, incluidos en parafina, fijados en formalina, obtenidos de 252 pacientes individuales con diagnóstico de cáncer de mama invasivo. Todos los pacientes tenían ganglios linfáticos negativos, eran positivos para ER y se trataron con tamoxifeno. La edad media era de 52 años y el tamaño de tumor clínico medio era de 2 cm. El seguimiento medio fue de 1,9 años. A fecha del 1 de enero de 2003, 41 pacientes tenían recidiva de enfermedad local o distante o habían muerto por cáncer de mama. Los pacientes sólo se incluyeron en el estudio si la evaluación histopatológica, realizada como se describe en la sección de Materiales y métodos, indicó cantidades adecuadas de tejido tumoral y patología homogénea.

Materiales y métodos

Se caracterizó cada bloque de tumor representativo mediante histopatología convencional para el diagnóstico, evaluación semicuantitativa de la cantidad de tumor y clasificación tumoral. Cuando el área tumoral era inferior al 70 % de la sección, se sometió el área tumoral a disección macroscópica y se tomó tejido de 6 secciones (10 micrómetros). De lo contrario, se prepararon un total de 3 secciones (también de 10 micrómetros de grosor cada

una). Se colocaron las secciones en dos tubos de microcentrifugadora de marca Costar (tubos de polipropileno, de 1,7 ml, transparentes). Si se obtuvo más de un bloque de tumor como parte del procedimiento quirúrgico, se usó para el análisis el bloque más representativo de la patología.

Análisis de la expresión génica

5 Se extrajo ARNm y se purificó a partir de muestras de tejido incluidas en parafina, fijadas, y se preparó para el análisis de la expresión génica como se describió en el capítulo 6 anterior.

Se realizaron ensayos moleculares de la expresión génica cuantitativa mediante RT-PCR, usando el ABI PRISM 7900™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, CA, Estados Unidos). ABI PRISM 7900™ consiste en un termociclador, un láser, un dispositivo de carga acoplada (CCD), una cámara y un ordenador. El sistema amplifica muestras en un formato de 384 pocillos en un termociclador. Durante la amplificación, se recoge en tiempo real una señal fluorescente inducida por láser a través de cables de fibra óptica para los 384 pocillos y se detecta en el CCD. El sistema incluye software para hacer funcionar el instrumento y para analizar los datos.

Análisis y resultados

15 Se analizó tejido tumoral para detectar 187 genes relacionados con cáncer y 5 genes de referencia. Se obtuvieron perfiles de RT-PCR adecuados de 242 de los 252 pacientes. Se normalizaron los valores de ciclo umbral (CT) para cada paciente basándose en la media de los 7 genes de referencia para ese paciente particular. Había datos de desenlace clínico disponibles para todos los pacientes a partir de una revisión de datos de registro e historias clínicas de pacientes seleccionados. Los desenlaces se clasificaron como:

20 Acontecimiento: Vivo con recidiva de cáncer de mama local, regional o distante o muerte debido a cáncer de mama.
Sin acontecimiento: Vivo sin recidiva de cáncer de mama local, regional o distante o vivo con recidiva de cáncer de mama contralateral o vivo con segundo cáncer primario distinto de mama o muerte antes de la recidiva de cáncer de mama.

El análisis se realizó mediante:

25 A. Determinación de la relación entre la expresión génica normalizada y los desenlaces binarios de 0 o 1;
B. Análisis de la relación entre la expresión génica normalizada y el tiempo hasta el desenlace (0 o 1 como se definió anteriormente) en el que se censuraron los pacientes que estaban vivos sin recidiva de cáncer de mama o que murieron debido a una causa distinta al cáncer de mama. Este enfoque se usó para evaluar el impacto del pronóstico de genes individuales y también conjuntos de múltiples genes.

30 Análisis de pacientes con carcinomas de mama invasivos mediante enfoque binario

En el primer enfoque (binario) se realizó el análisis con los 242 pacientes con carcinoma de mama invasivo. Se realizó una prueba de la t con los grupos de pacientes clasificados como sin recidiva o sin muerte relacionada con cáncer de mama a los 10 años, frente a recidiva o muerte relacionada con cáncer de mama a los 10 años, y se calcularon los valores de p para las diferencias entre los grupos para cada gen.

35 La Tabla 1 enumera los 33 genes para los que el valor de p para las diferencias entre los grupos fue <0,05. La primera columna de valores de expresión medios se refiere a pacientes que tenían una recidiva metastásica de, o murieron por, cáncer de mama. La segunda columna de valores de expresión medios se refiere a pacientes que ni tenían recidiva metastásica ni murieron por cáncer de mama.

40 **Tabla 1**

Gen	Media del grupo A	Media del grupo B	Prueba estadística de la t	Valor de p
MMP9	-3,15	-4,27	3,75	0,00
GSTM1	-5,02	-4,03	-3,56	0,00
MELK	-3,89	-4,66	3,34	0,00
PR	-4,56	-3,18	-3,27	0,00
DKFZp586M07	-3,83	-2,94	-3,09	0,00
GSTM3	-2,56	-1,69	-3,06	0,00

MCM2	-3,51	-4,08	3,03	0,00
CDC20	-3,01	-3,75	3,01	0,00
CCNB1	-4,48	-5,17	3,02	0,00
STMY3	-0,58	-1,20	2,95	0,00
GRB7	-1,93	-3,01	2,98	0,00
MYBL2	-3,91	-4,78	2,91	0,01
CEGP1	-3,00	-1,85	-2,89	0,01
SURV	-4,23	-5,06	2,88	0,01
LMNB1	-2,40	-2,91	2,81	0,01
CTSL2	-5,74	-6,39	2,83	0,01
PTTG1	-3,49	-4,14	2,72	0,01
BAG1	-1,76	-1,30	-2,58	0,01
KNSL2	-3,35	-4,06	2,60	0,01
CIAP1	-4,44	-4,02	-2,58	0,01
PREP	-3,34	-3,74	2,56	0,01
NEK2	-5,25	-5,80	2,53	0,01
EpCAM	-1,95	-2,31	2,50	0,01
PCNA	-2,79	-3,13	2,42	0,02
C20_orf 1	-2,48	-3,09	2,39	0,02
ITGA7	-4,53	-3,87	-2,37	0,02
ID1	-2,58	-2,17	-2,30	0,02
B_Catenin	-1,32	-1,08	-2,28	0,03
EstR1	-0,78	-0,12	-2,28	0,03
CDH1	-2,76	-3,27	2,20	0,03
TS	-2,86	-3,29	2,18	0,03
HER2	0,53	-0,22	2,18	0,03
cMYC	-3,22	-2,85	-2,16	0,04

5 En la tabla 1 anterior, los valores de t negativos indican una mayor expresión, asociada a mejores desenlaces, y a la inversa, los valores de t superiores (positivos) indican una mayor expresión asociada a peores desenlaces. Por lo tanto, por ejemplo, una expresión elevada del gen CCNB1 (valor de t = 3,02; CT media vivos < CT media fallecidos) indica 5 una probabilidad reducida de supervivencia libre de enfermedad. De manera similar, la expresión elevada del gen GSTM1 (valor de t = -3,56; CT media vivos > CT media fallecidos) indica una probabilidad aumentada de supervivencia libre de enfermedad.

10 Por lo tanto, basándose en los datos expuestos en la Tabla 1, la expresión de cualquiera de los siguientes genes en cáncer de mama indica una probabilidad reducida de supervivencia sin recidiva de cáncer: C20_orf1; CCNB1; CDC20; CDH1; CTSL2; EpCAM; GRB7; HER2; KNSL2; LMNB1; MCM2; MMP9; MYBL2; NEK2; PCNA; PREP; PTTG1; STMY3; SURV; TS; MELK.

Basándose en los datos expuestos en la Tabla 1, la expresión de cualquiera de los siguientes genes en cáncer de mama indica un mejor pronóstico para la supervivencia sin recidiva de cáncer: BAG1; BCatenin; CEGP1; CIAP1; cMYC; DKFZp586M07; EstR1; GSTM1; GSTM3; ID1; ITGA7; PR.

15 Análisis de múltiples genes e indicadores del desenlace

Se tomaron dos enfoques con el fin de determinar si usar múltiples genes proporcionaba una mejor distinción entre desenlaces. En primer lugar, se realizó un análisis de distinción usando un enfoque de avance gradual. Se generaron modelos que clasificaron el desenlace con mayor distinción que la que se obtuvo con cualquier gen individual en solitario. De acuerdo con un segundo enfoque (enfoque de tiempo hasta el acontecimiento), para cada gen se definió un modelo de riesgos proporcionales de Cox (véase, por ejemplo Cox, D. R., y Oakes, D. (1984), Analysis of Survival Data, Chapman y Hall, Londres, Nueva York) con tiempo hasta la recidiva o muerte como la variable dependiente, y el nivel de expresión del gen como la variable independiente. Se identificaron los genes que tienen un valor de $p < 0,05$ en el modelo de Cox. Para cada gen, el modelo de Cox proporciona el riesgo relativo (RR) de recidiva o muerte para un cambio de una unidad en la expresión del gen. Puede elegirse repartir los pacientes en subgrupos a cualquier valor umbral de la expresión medida (en la escala de CT), en el que todos los pacientes con valores de expresión por encima del umbral tienen un riesgo mayor y todos los pacientes con valores de expresión por debajo del umbral tienen un riesgo menor, o viceversa, dependiendo de si el gen es un indicador de un pronóstico malo ($RR > 1,01$) o bueno ($RR < 1,01$). Por lo tanto, cualquier valor umbral definirá subgrupos de pacientes con un riesgo aumentado o reducido respectivamente. Los resultados se resumen en la Tabla 2, que enumera los 42 genes para los que el valor de p para las diferencias entre los grupos era $< 0,05$.

Tabla 2

Gen	Riesgo relativo	Valor de p
GRB7	1,52	0,000011
SURV	1,57	0,000090
PR	0,74	0,000129
LMNB1	1,92	0,000227
MYBL2	1,46	0,000264
HER2	1,46	0,000505
GSTM1	0,68	0,000543
MELK	1,59	0,000684
C20_orF1	1,59	0,000735
PTTG 1	1,63	0,001135
BUB1	1,58	0,001425
CDC20	1,54	0,001443
CCNB1	1,60	0,001975
STMY3	1,47	0,002337
KNSL2	1,48	0,002910
CTSL2	1,43	0,003877
MCM2	1,59	0,005203
NEK2	1,48	0,006533
DR5	0,62	0,006660
Ki_67	1,46	0,008188
CCNE2	1,38	0,009505
TOP2A	1,38	0,009551
PCNA	1,67	0,010237
PREP	1,69	0,012308
FOXM1	1,52	0,012837
NME1	1,46	0,013622

CEGP1	0,84	0,013754
BAG1	0,68	0,015422
STK15	1,46	0,017013
HNRPAB	1,96	0,017942
EstR1	0,80	0,018877
MMP9	1,19	0,019591
DKFZp586M07	0,79	0,020073
TS	1,44	0,025186
Src	1,70	0,037398
BIN1	0,75	0,038979
NPD009	0,80	0,039020
RPLPO	0,52	0,041575
GSTM3	0,84	0,041848
MMP12	1,27	0,042074
TFRC	1,57	0,046145
IGF1R	0,78	0,046745

5 Basándose en los datos expuestos en la Tabla 2, la expresión de cualquiera de los siguientes genes en el cáncer de mama indica una probabilidad reducida de supervivencia sin recidiva de cáncer: GRB7; SURV; LMNB1; MYBL2; HER2; MELK; C20_orf1; PTTG1; BUB1; CDC20; CCNB1; STMY3; KNSL2; CTSL2; MCM2; NEK2; Ki_67; CCNE2; TOP2A-4; PCNA; PPREP; FOXM1; NME1; STK15; HNRPAB; MMP9; TS; Src; MMP12; TFRC.

Basándose en los datos expuestos en la Tabla 2, la expresión de cualquiera de los siguientes genes en el cáncer de mama indica un mejor pronóstico para la supervivencia sin recidiva de cáncer: PR; GSTM11; DR5; CEGP1; BAG1; EstR1; DKFZp586M07; BIN1; NPD009; RPLPO; GSTM3; IGF1R.

10 Los análisis binario y de tiempo hasta el acontecimiento, con pocas excepciones, identificaron los mismos genes como marcadores de pronóstico. Por ejemplo, la comparación de las Tablas 1 y 2 muestra que 10 genes estaban representados entre los 15 genes más importantes en ambas listas. Además, cuando ambos análisis identificaron el mismo gen a $[p < 0,01]$, lo que sucedió para 26 genes, siempre concordaban con respecto al sentido (signo positivo o negativo) de la correlación con la supervivencia/recidiva. En conjunto, estos resultados refuerzan la conclusión de que los marcadores identificados tienen un valor de pronóstico significativo.

15 Análisis génico de múltiples variables de 242 pacientes con carcinoma de mama invasivo

20 Para modelos de Cox que comprenden más de dos genes (modelos de múltiples variables), se realiza la introducción gradual de cada gen individual en el modelo, en el que el primer gen introducido se selecciona previamente de los genes que tienen valores de p para una variable significativos, y el gen seleccionado para su introducción en el modelo en cada etapa posterior es el gen que mejora de la mejor manera el ajuste del modelo a los datos. Este análisis puede realizarse con cualquier número total de genes. En el análisis cuyos resultados se muestran a continuación, se realizó la introducción gradual para hasta 10 genes.

Se realizó un análisis de múltiples variables usando la siguiente ecuación: $RR = \exp[\text{coef}(\text{genA}) \times \text{Ct}(\text{genA}) + \text{coef}(\text{genB}) \times \text{Ct}(\text{genB}) + \text{coef}(\text{genC}) \times \text{Ct}(\text{genC}) + \dots]$.

25 En esta ecuación, los coeficientes para los genes que son factores pronóstico de un desenlace beneficioso son números positivos y los coeficientes para genes que son factores pronóstico de un desenlace desfavorable son números negativos. Los valores de "Ct" en la ecuación son ΔCt , es decir, reflejan la diferencia entre el valor de Ct normalizado promedio para una población y el Ct normalizado medido para el paciente en cuestión. El convenio usado en el presente análisis ha sido que ΔCt por debajo y por encima del promedio de la población tienen signos positivos y signos negativos, respectivamente (reflejando una mayor o menor abundancia de ARNm). El riesgo relativo (RR) calculado resolviendo esta ecuación indicará si el paciente tiene una probabilidad potenciada o
30 reducida de supervivencia a largo plazo sin recidiva de cáncer.

5 Se realizó un análisis gradual de múltiples variables, usando el modelo de riesgos proporcionales de Cox, sobre los datos de expresión génica obtenidos para los 242 pacientes con carcinoma de mama invasivo. Mediante este análisis se identificó que el siguiente conjunto de diez genes tenía un valor predictivo particularmente fuerte de la supervivencia de pacientes: GRB7; LMNB1; ER; STMY3; KLK10; PR; KRT5; FGFR1; MCM6; SNRPF. En este conjunto de genes, ER, PR, KRT5 y MCM6 contribuyen a un buen pronóstico, mientras que GRB7, LMNB1, STMY3, KLK10, FGFR1 y SNRPF contribuyen a un mal pronóstico.

10 Aunque se ha descrito la presente invención con referencia a lo que se considera que son las realizaciones específicas, debe entenderse que la invención no se limita a tales realizaciones. Por el contrario, se pretende que la invención cubra diversas modificaciones y equivalentes incluidos dentro del alcance de las reivindicaciones adjuntas.

TABLA 3

Gen	Acceso	Inicio	Terminación	SEQ ID NO.	Secuencia
B-Catenin	NM_001	1549	1629	SEQ NO: 1	ID GGCTCTTGTGCGTACTGTCTTCGGGCTGGTGACAGGGAAAGACATCACTGAGCCTGCCATCTGTGCTCTTCGTCATCTGA
BAG1	NM_004	673	754	SEQ NO: 2	ID CGTTGTGACGACTTGGAAATACAAGATGGTTGCCGGGTCATGTTAATTGGGAAAAAAGAACAGTCCACAGGAAGAGGTTGAAC
BIN1	NM_004	866	942	SEQ NO: 3	ID CCTGCAAAAGGGAAACAAGAGCCCTTCGCCTCCAGATGGCTCCCTGCCGCCACCCCGAGATCAAGTCAACCACG
BUB1	NM_004	1002	1070	SEQ NO: 4	ID CCGAGGTTAATCCAGCAGCATGTGGGGCCAAGTGTAGGCTCCAGCAGGAAGTGTGAGCGCCATGTCCT
C20orf1	NM_012	2675	2740	SEQ NO: 5	ID TCAGCTGTGAGCTGCGGATACCGCCCGCAATTGGACCTGCTCTAACCTCAAACCTAGGACCGT
CCNB1	NM_031	823	907	SEQ NO: 6	ID TTCAGGTTGTTGCGAGGACCATGTACATGACTGTCTCCATTATTGATCGGTTTCATGCAGAATAAATGTGTGCCAAGAAGATG
CCNE2	NM_057	2026	2108	SEQ NO: 7	ID ATGCTGTGGCTCCTTCTAACTGGGGCTTTCTTGACATGTAGGTTGCTTGGTAATAACCTTTTTGTATACAAATTTGGT
CDC20	NM_001	679	747	SEQ NO: 8	ID TGGATTGGAGTTCTGGGAATGTACTGGCCGTGGCACTGGACAACAGTGTGTACCTGTGGAGTGCAAGC
CDH1	NM_004	2499	2580	SEQ NO: 9	ID TGAGTGTCCCCCGGTATCTCCCCCGCCTGCCAATCCCGATGAAATTGGAAATTTATTGATGAAAATCTGAAAGCGGCTG
CEGP1	NM_020	563	640	SEQ NO: 10	ID TGACAATCAGCACACCTGCATTCACCGCTCGGAAGAGGGCTGAGCTGCATGAATAAGGATCAGCGCTGTAGTCACA
CIAP1	NM_001	1822	1894	SEQ NO: 11	ID TGCCCTGTGTGGGAAGCTCAGTAACCTGGAAACCAAGGATGATGTATGTCAGAACCCCGGAGGCCATTTTCC
cMYC	NM_002	1494	1578	SEQ NO: 12	ID TCCTCCACTCGGAAGGACTATCCTGCTGCCAAGAGGGTCAAGTTGGACAGTGTGACAGATCCTGAGACAGATCAGCAACCCG
CTSL2	NM_001	671	738	SEQ NO: 13	ID TGTCTCACTGAGCGAGCAGAATCTGGTGGACTGTTCCGCTCTCAAGGCAATCAGGGCTGCAATGT
DKFZp586	AL05022	559	633	SEQ NO: 14	ID TCCATTTTCTACCTGTTAACCTTCATATTTTGTGCAGGCCCTGGAAGCAAAAGAGAGGAAAGGGACCCTGACTGCAT
DR5	NM_003	1127	1211	SEQ NO: 15	ID CTCTGAGACAGTGTCTCGATGACTTTGCAGACTTGGTGCCCTTTGACTCCTGGGAGCCGCTCATGAGGAAGTTGGCCCTCATGG
EpCAM	NM_002	435	510	SEQ NO: 16	ID GGGCCCTCCAGAACAATGATGGGCTTTATGATCCTGACTGCGATGAGAGCGGGCTCTTTAAGGCCAAGCAGTGCA
EstR1	NM_000	1956	2024	SEQ NO: 17	ID CGTGGTCCCTCTATGACCTGCTGCTGGAGATGCTGGACGCCACCCCTACATGCGCCCACTAGCC
FGFR1	NM_023	2685	2759	SEQ NO: 18	ID CACGGGACATTCACCACATCGACTACTATAAAAAGACAACCAACCGCCGACTGCCTGTGAAGTGGATGGCACCC
FOXM1	NM_021	1898	1980	SEQ NO: 19	ID CCACCCCGAGCAAAATCTGCTCCCGAGAACCCTGAATCCTGGAGGCTCACGCCCCAGCCAAAAGTGGGGGACTGGATT
GRB7	NM_005	1275	1342	SEQ NO: 20	ID CCATCTGCATCCATCTTGTGGTCCCGCCCTTGAGAAGTGCCTCAGATAATACCCTGGTGGCC
GSTM1	NM_000	93	179	SEQ NO: 21	ID AAGCTATGAGGAAAAGAGTACACGATGGGGGACGCTCCTGATTATGACAGAAGCCAGTGGCTGATGAAAAATTCAGCTGGCC
GSTM3	NM_000	248	324	SEQ NO: 22	ID CAATGCCATCTTGCCTACATCGCTCGCAAGCACAACATGTGTGGTGTGAGACTGAAGAAGAAAAGATCGAGTGGAC
HER2	NM_004	1138	1208	SEQ NO: 23	ID CGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGAGTGTGCTATGGTCTGGGCATGGAGCACTTGCAGAGG
HNRPA B	NM_004	1086	1170	SEQ NO: 24	ID CAAGGGAGCGACCAACTGATCGCACACATGCTTTGTTGGATGAGAGTGAACACAATTATGTACCATAATTAACCTGGCAAAC
ID1	NM_002	286	356	SEQ NO: 25	ID AGAACCGCAAGGTGAGCAAGGTGGAGATTCTCCAGCAGCTCATCGACTACATCAGGGACCTTCAGTTGGA
IGF1R	NM_000	3467	3550	SEQ NO: 26	ID GCATGGTAGCCGAAGATTTACAGTCAAATCGGAGATTTGGTATGACCGGAGATATCTATGAGACAGACTATTACCGGAAA
ITGA7	NM_002	633	712	SEQ NO: 27	ID GATATGATTGGTCTGCTTTGTGCTCAGCCAGGACCTGGCCATCCGGGATGAGTTGGATGGTGGGAATGGAAGTTCT
KI-67	NM_002	42	122	SEQ NO: 28	ID CGGACTTTGGGTGCGACTTGACGAGCGGTGGTTCGACAAGTGGCCTTCCGGGCGGATCGTCCCAGTGGAAAGATTGTA
KLK10	NM_002	966	1044	SEQ NO: 29	ID GCCAGAGGCTCCATCGTCCATCCTCTTCTCCCGAGTCCGCTGAACTCTCCCTGTCTGCACTGTTCAAACCTCTG
KNSL2	BC00071	1266	1343	SEQ NO: 30	ID CCACCTCGCCATGATTTTCTTTGACCGGATTTCCACCAGGAAGTGGACAGGATGAAGTGTGGAAGAGATTGC
KRT5	NM_000	1605	1674	SEQ NO: 31	ID TCAGTGGAAGGAGTGGACCAGTCAACATCTCTGTTGCACAAGCAGTGTCTCTGGATATGCA
LMNB1	NM_005	1500	1566	SEQ NO: 32	ID TGCAAAACGCTGGTGTACAGCCAGCCCCCAACTGACCTCATCTGGAAGAACCAGAAGTCTGGGG
MCM2	NM_004	2442	2517	SEQ NO: 33	ID GACTTTTCCCGCTACCTTTCATTCCGGCGTGACAACAATGAGCTGTTGCTCTTACTACTGAAGCAGTTAGTGGC
MCM6	NM_005	2669	2751	SEQ NO: 34	ID TGATGGTCTATGTGTACATTATCACAGTTTCATACCAACACAGGCTTCCAGCACTTCTTTGGTGTGTTTCTGTCCCA
MELK	NM_014	22	87	SEQ NO: 35	ID AACCCGCGGATCGAAAAGATTCTTAGGAACCGCTACCAGCCGCTCTCAGGACAGCAGGCC
MMP12	NM_002	816	894	SEQ NO: 36	ID CCAACGCTTGCCAAATCTGACAATTGACAACAGCTCTCTGTGACCCCAATTTGAGTTTTGATGCTGCACTACCGT
MMP9	NM_004	124	191	SEQ NO: 37	ID GAGAACCAATCTCACCGACAGGCGAGTGGCAGAGGAATACCTGTACCGCTATGGTTACACTCGGGTG
MYBL2	NM_002	599	673	SEQ NO: 38	ID GCCGAGATCGCAAGATGTTGCCAGGAGGACAGACAATGCTGTGAAGAATCACTGGAACCTACCATCAAAAG
NEK2	NM_002	102	181	SEQ NO: 39	ID GTGAGGACGCGGACTCTGGCGACTGGCCGGCCATGCCTTCCCGGGCTGAGGACTATGAAGTGTGTACACCAATTGGCA
NME1	NM_000	365	439	SEQ NO: 40	ID CCAACCCTGCAGACTCAAAGCCTGGGACCATCCGTGGAGACTCTGCATACAAGTTGGCAGGAACATTATACAT
NPD009	NM_020	589	662	SEQ	ID GGCTGTGGCTGAGGCTGTAGCATCTCTGCTGGAGGTGAGACTCTGGAACTGATTTGACCTCG

ES 2 690 168 T3

	686			NO: 41	AATGCTCC
PCNA	NM_002	157	228	SEQ	ID GAAGGTGTTGGAGGCACTCAAGGACCTCATCAACGAGGCCTGCTGGGATATTAGCTCCAGCGGTG
	592			NO: 42	TAAACC
PR	NM_000	1895	1980	SEQ	ID GCATCAGGCTGTCATTATGGTGCTCCTTACCTGTGGGAGCTGTAAGGTCTTCTTTAAGAGGGCAATG
	926			NO: 43	GAAGGGCAGCACAACTACT
PREP	NM_002	889	965	SEQ	ID GGGACGGTGTTCACATTCAAGACGAATCGCCAGTCTCCCAACTATCGCGTGATCAACATTGACTTC
	726			NO: 44	TGGGATCCTG
PTTG1	NM_004	48	122	SEQ	ID GGCTACTCTGATCTATGTTGATAAGGAAAATGGAGAACCAGGCACCCGTGTGGTTGCTAAGGATG
	219			NO: 45	GGCTGAAGC
RPLPO	NM_001	791	866	SEQ	ID CCATTCTATCATCAACGGGTACAAACGAGTCCTGGCCTTGTCTGTGGAGACGGATTACACCTTCCC
	002			NO: 46	ACTTGCTGA
SNRPF	NM_003	71	150	SEQ	ID GGCTGGTCGGCAGAGAGTAGCCTGCAACATTGGCCGGTGGTTACATGAGTTTACCCCTCAATCC
	095			NO: 47	CAAACCTTTCCTCA
Src	NM_004	979	1043	SEQ	ID CCTGAACATGAAGGAGCTGAAGCTGCTGCAGACCATCGGGAAGGGGGAGTTCGGAGACGTGATG
	383			NO: 48	
STK15	NM_003	1101	1170	SEQ	ID CATCTTCCAGGAGGACCACTCTCTGTGGCACCCCTGGACTACCTGCCCCCTGAAATGATTGAAGGT
	600			NO: 49	CGGA
STMY3	NM_005	2090	2180	SEQ	ID CCTGGAGGCTGCAACATACCTCAATCCTGTCCCAGGCCGGATCCTCCTGAAGCCCTTTTCGCAGC
	940			NO: 50	ACTGCTATCCTCAAAGCCATTGTA
SURV	NM_001	737	817	SEQ	ID TGTTTTGATTCCCAGGCTTACCAGGTGAGAAGTGAGGGAGGAAGAAGGCAGTGTCCCTTTTGCTA
	168			NO: 51	GAGCTGACAGCTTG
TFRC	NM_003	2110	2178	SEQ	ID GCCAACTGCTTTTCATTTGTGAGGGATCTGAACCAATACAGAGCAGACATAAAGGAAATGGGCCTGA
	234			NO: 52	GT
TOP2A	NM_001	4505	4577	SEQ	ID AATCCAAGGGGAGAGTGATGACTTCCATATGGACTTTGACTCAGCTGTGGCTCCTCGGGCAAAA
	067			NO: 53	TCTGTAC
TS	NM_001	764	829	SEQ	ID GCCTCGGTGTGCCTTTCAACATCGCCAGCTACGCCCTGCTCACGTACATGATTGCGCACATCAGC
	071			NO: 54	

TABLA 4

Gen	Acceso	Nombre	SEQ ID NO	Secuencia	
B-Catenin	NM_001904	S2150/B-Cate.f3	SEQ ID NO:55	GGCTCTTGTGCGTACTGTCTT	22
B-Catenin	NM_001904	S2151/B-Cate.r3	SEQ ID NO:56	TCAGATGACGAAGAGCACAGATG	23
B-Catenin	NM_001904	S5046/B-Cate.p3	SEQ ID NO:57	AGGCTCAGTGATGTCTTCCCTGTACCAG	29
BAG1	NM_004323	S1386/BAG1.f2	SEQ ID NO:58	CGTTGTCAGCACTTGAATACAA	23
BAG1	NM_004323	S1387/BAG1.r2	SEQ ID NO:59	GTTCAACCTCTTCTGTGGACTGT	24
BAG1	NM_004323	S4731/BAG1.p2	SEQ ID NO:60	CCCAATTAACATGACCCGGCAACCAT	26
BIN1	NM_004305	S2651/BIN1.f3	SEQ ID NO:61	CCTGCAAAGGGAAACAAGAG	20
BIN1	NM_004305	S2652/BIN1.r3	SEQ ID NO:62	CGTGGTTGACTCTGATCTCG	20
BIN1	NM_004305	S4954/BIN1.p3	SEQ ID NO:63	CTTCGCCTCCAGATGGCTCCC	21
BUB1	NM_004336	S4294/BUB1.f1	SEQ ID NO:64	CCGAGGTTAATCCAGCACGTA	21
BUB1	NM_004336	S4295/BUB1.r1	SEQ ID NO:65	AAGACATGGCGCTCTCAGTTC	21
BUB1	NM_004336	S4296/BUB1.p1	SEQ ID NO:66	TGCTGGGAGCCTACACTTGGCCC	23
C20 orf1	NM_012112	S3560/C20 or.f1	SEQ ID NO:67	TCAGCTGTGAGCTGCGGATA	20
C20 orf1	NM_012112	S3561/C20 or.r1	SEQ ID NO:68	ACGGTCCTAGGTTTGAGGTTAAGA	24
C20 orf1	NM_012112	S3562/C20 or.p1	SEQ ID NO:69	CAGGTCCCATTGCCGGGCG	19
CCNB1	NM_031966	S1720/CCNB1.f2	SEQ ID NO:70	TTCAGGTTGTTGCAGGAGAC	20
CCNB1	NM_031966	S1721/CCNB1.r2	SEQ ID NO:71	CATCTTCTGGGCACACAAT	20
CCNB1	NM_031966	S4733/CCNB1.p2	SEQ ID NO:72	TGTCTCCATTATTGATCGGTTATGCA	27
CCNE2	NM_057749	S1458/CCNE2.f2	SEQ ID NO:73	ATGCTGTGGCTCCTCCTAACT	22
CCNE2	NM_057749	S1459/CCNE2.r2	SEQ ID NO:74	ACCCAAATTGTGATATACAAAAGGTT	27
CCNE2	NM_057749	S4945/CCNE2.p2	SEQ ID NO:75	TACCAAGCAACCTACATGTCAAGAAAGCCC	30
CDC20	NM_001255	S4447/CDC20.f1	SEQ ID NO:76	TGGATTGGAGTTCTGGGAATG	21
CDC20	NM_001255	S4448/CDC20.r1	SEQ ID NO:77	GCTTGCACTCCACAGGTACACA	22
CDC20	NM_001255	S4449/CDC20.p1	SEQ ID NO:78	ACTGGCCGTGGCACTGGACAACA	23
CDH1	NM_004360	S0073/CDH1.f3	SEQ ID NO:79	TGAGTGTCCCCCGGTATCTTC	21
CDH1	NM_004360	S0075/CDH1.r3	SEQ ID NO:80	CAGCCGCTTTCAGATTTTCAT	21
CDH1	NM_004360	S4990/CDH1.p3	SEQ ID NO:81	TGCCAATCCCGATGAAATGGAAATTT	27
CEGP1	NM_020974	S1494/CEGP1.f2	SEQ ID NO:82	TGACAATCAGCACACCTGCAT	21
CEGP1	NM_020974	S1495/CEGP1.r2	SEQ ID NO:83	TGTGACTACAGCCGTGATCCTTA	23
CEGP1	NM_020974	S4735/CEGP1.p2	SEQ ID NO:84	CAGGCCCTCTCCGAGCGGT	20
CIAP1	NM_001166	S0764/CIAP1.f2	SEQ ID NO:85	TGCCTGTGGTGGGAAGCT	18
CIAP1	NM_001166	S0765/CIAP1.r2	SEQ ID NO:86	GGAAAATGCCTCCGGTGTT	19
CIAP1	NM_001166	S4802/CIAP1.p2	SEQ ID NO:87	TGACATAGCATCATCCTTTGGTTCCAGTT	30
cMYC	NM_002467	S0085/cMYC.f3	SEQ ID NO:88	TCCTCCACTCGGAAGGACTA	21
cMYC	NM_002467	S0087/cMYC.r3	SEQ ID NO:89	CGGTTGTTGCTGATCTGTCTCA	22
cMYC	NM_002467	S4994/cMYC.p3	SEQ ID NO:90	TCTGACTGTCCAACCTTGACCCCTCT	27
CTSL2	NM_001333	S4354/CTSL2.f1	SEQ ID NO:91	TGTCTCACTGAGCGAGCAGAA	21
CTSL2	NM_001333	S4355/CTSL2.r1	SEQ ID NO:92	ACCATTGCAGCCCTGATTG	19
CTSL2	NM_001333	S4356/CTSL2.p1	SEQ ID NO:93	CTTGAGGACGCGAACAGTCCACCA	24
DKFZp586M0723	AL050227	S4396/DKFZp5.f1	SEQ ID NO:94	TCCATTTTCTACCTGTTAACCTTCATC	27
DKFZp586M0723	AL050227	S4397/DKFZp5.r1	SEQ ID NO:95	ATGCAGTCGGTCCCTTCCT	19
DKFZp586M0723	AL050227	S4398/DKFZp5.p1	SEQ ID NO:96	TTGCTTCCAGGGCCTGCACAAAA	23
DR5	NM_003842	S2551/DR5.f2	SEQ ID NO:97	CTCTGAGACAGTGCTTCGATGACT	24
DR5	NM_003842	S2552/DR5.r2	SEQ ID NO:98	CCATGAGGCCCAACTTCCT	19
DR5	NM_003842	S4979/DR5.p2	SEQ ID NO:99	CAGACTTGGTGCCCTTTGACTCC	23
EpCAM	NM_002354	S1807/EpCAM.f1	SEQ ID NO:100	GGGCCCTCCAGAACAATGAT	20
EpCAM	NM_002354	S1808/EpCAM.r1	SEQ ID NO:101	TGCACTGCTTGGCCTTAAGA	21
EpCAM	NM_002354	S4984/EpCAM.p1	SEQ ID NO:102	CCGCTCTCATCGCAGTCAGGATCAT	25

ES 2 690 168 T3

EstR1	NM_000125	S0115/EstR1.f1	SEQ ID NO:103	CGTGGTGCCCTCTATGAC	19
EstR1	NM_000125	S0117/EstR1.r1	SEQ ID NO:104	GGCTAGTGGGCGCATGTAG	19
EstR1	NM_000125	S4737/EstR1.p1	SEQ ID NO:105	CTGGAGATGCTGGACGCCC	19
FGFR1	NM_023109	S0818/FGFR1.f3	SEQ ID NO:106	CACGGGACATTCCACACATC	20
FGFR1	NM_023109	S0819/FGFR1.r3	SEQ ID NO:107	GGGTGCCATCCACTTCACA	19
FGFR1	NM_023109	S4816/FGFR1.p3	SEQ ID NO:108	ATAAAAAGACAACCAACGGCCGACTGC	27
FOXM1	NM_021953	S2006/FOXM1.f1	SEQ ID NO:109	CCACCCCGAGCAAATCTGT	19
FOXM1	NM_021953	S2007/FOXM1.r1	SEQ ID NO:110	AAATCCAGTCCCCCTACTTTGG	22
FOXM1	NM_021953	S4757/FOXM1.p1	SEQ ID NO:111	CCTGAATCCTGGAGGCTCACGCC	23
GRB7	NM_005310	S0130/GRB7.f2	SEQ ID NO:112	ccatctgcatccatctgft	20
GRB7	NM_005310	S0132/GRB7.r2	SEQ ID NO:113	ggccaccagggtattatctg	20
GRB7	NM_005310	S4726/GRB7.p2	SEQ ID NO:114	ctccccacccttgagaagtgct	23
GSTM1	NM_000561	S2026/GSTM1.r1	SEQ ID NO:115	GGCCGAGCTTGAATTTTTCA	20
GSTM1	NM_000561	S2027/GSTM1.f1	SEQ ID NO:116	AAGCTATGAGGAAAAGAAGTACACGAT	27
GSTM1	NM_000561	S4739/GSTM1.p1	SEQ ID NO:117	TCAGCCACTGGCTTCTGTACATAATCAGGAG	30
GSTM3	NM_000849	S2038/GSTM3.f2	SEQ ID NO:118	CAATGCCATCTTGGCTACAT	21
GSTM3	NM_000849	S2039/GSTM3.r2	SEQ ID NO:119	GTCCACTCGAATCTTTCTTCTTCA	25
GSTM3	NM_000849	S5064/GSTM3.p2	SEQ ID NO:120	CTCGCAAGCACACATGTGTGGTGAGA	27
HER2	NM_004448	S0142/HER2.f3	SEQ ID NO:121	CGGTGTGAGAAGTGACAGCAA	20
HER2	NM_004448	S0144/HER2.r3	SEQ ID NO:122	CCTCTCGCAAGTGCTCCAT	19
HER2	NM_004448	S4729/HER2.p3	SEQ ID NO:123	CCAGACCATAGCACACTCGGGCAC	24
HNRPAB	NM_004499	S4510/HNRPAB.f3	SEQ ID NO:124	CAAGGGAGCGACCAACTGA	19
HNRPAB	NM_004499	S4511/HNRPAB.r3	SEQ ID NO:125	GTTTGCCAAGTTAAATTTGGTACATAAT	28
HNRPAB	NM_004499	S4512/HNRPAB.p3	SEQ ID NO:126	CTCCATATCCAAACAAAGCATGTGTGCG	28
ID1	NM_002165	S0820/ID1.f1	SEQ ID NO:127	AGAACCGCAAGGTGAGCAA	19
ID1	NM_002165	S0821/ID1.r1	SEQ ID NO:128	TCCAACCTGAAGTCCCTGATG	21
ID1	NM_002165	S4832/ID1.p1	SEQ ID NO:129	TGGAGATTCTCCAGCACGTCATCGAC	26
IGF1R	NM_000875	S1249/IGF1R.f3	SEQ ID NO:130	GCATGGTAGCCGAAGATTCA	21
IGF1R	NM_000875	S1250/IGF1R.r3	SEQ ID NO:131	TTTCCGGTAATAGTCTGTCTCATAGATATC	30
IGF1R	NM_000875	S4895/IGF1R.p3	SEQ ID NO:132	CGCGTCATACCAAATCTCCGATTTTGA	28
ITGA7	NM_002206	S0859/ITGA7.f1	SEQ ID NO:133	GATATGATTGGTGCCTGCTTTG	22
ITGA7	NM_002206	S0920/ITGA7.r1	SEQ ID NO:134	AGAACTTCCATTCCCCACCAT	21
ITGA7	NM_002206	S4795/ITGA7.p1	SEQ ID NO:135	CAGCCAGGACCTGGCCATCCG	21
KI-67	NM_002417	S0436/KI-67.f2	SEQ ID NO:136	CGGACTTTGGGTGCGACTT	19
KI-67	NM_002417	S0437/KI-67.r2	SEQ ID NO:137	TTACAACTCTTCCACTGGGACGAT	24
KI-67	NM_002417	S4741/KI-67.p2	SEQ ID NO:138	CCACTTGTGCAACCACCGCTCGT	23
KLK10	NM_002776	S2624/KLK10.f3	SEQ ID NO:139	GCCCAGAGGCTCCATCGT	18
KLK10	NM_002776	S2625/KLK10.r3	SEQ ID NO:140	CAGAGGTTTGAACAGTGACAGACA	23
KLK10	NM_002776	S4978/KLK10.p3	SEQ ID NO:141	CCTCTTCTCCCCAGTCGGCTGA	23
KNSL2	BC000712	S4432/KNSL2.f2	SEQ ID NO:142	CCACCTCGCCATGATTTTTC	20
KNSL2	BC000712	S4433/KNSL2.r2	SEQ ID NO:143	GCAATCTCTCAAACACTTCATCCT	25
KNSL2	BC000712	S4434/KNSL2.p2	SEQ ID NO:144	TTTGACCGGGTATTCCCACCAGGAA	25
KRT5	NM_000424	S0175/KRT5.f3	SEQ ID NO:145	tcagtgaggagaaggagtgga	20
KRT5	NM_000424	S0177/KRT5.r3	SEQ ID NO:146	tgccatatccagaggaaaca	20
KRT5	NM_000424	S5015/KRT5.p3	SEQ ID NO:147	ccagtcacacatctgtgtgcacaagca	28
LMNB1	NM_005573	S4477/LMNB1.f1	SEQ ID NO:148	TGCAAACGCTGGTGTACACA	19
LMNB1	NM_005573	S4478/LMNB1.r1	SEQ ID NO:149	CCCCACGAGTTCTGGTTCTTC	21
LMNB1	NM_005573	S4479/LMNB1.p1	SEQ ID NO:150	CAGCCCCCAACTGACCTCATC	22
MCM2	NM_004526	S1602/MCM2.f2	SEQ ID NO:151	GACTTTTGCCCGTACCTTTTC	21

ES 2 690 168 T3

MCM2	NM_004526	S1603/MCM2.r2	SEQ ID NO:152	GCCACTAACTGCTTCAGTATGAAGAG	26
MCM2	NM_004526	S4900/MCM2.p2	SEQ ID NO:153	ACAGCTCATTGTTGTGCACGCCGGA	24
MCM6	NM_005915	S1704/MCM6.f3	SEQ ID NO:154	TGATGGTCCTATGTGTACATTCA	24
MCM6	NM_005915	S1705/MCM6.r3	SEQ ID NO:155	TGGGACAGGAAACACACCAA	20
MCM6	NM_005915	S4919/MCM6.p3	SEQ ID NO:156	CAGGTTTCATACCAACACAGGCTTCAGCAC	30
MELK	NM_014791	S4318/MELK.f1	SEQ ID NO:157	AACCCGGCGATCGAAAAG	18
MELK	NM_014791	S4319/MELK.r1	SEQ ID NO:158	GGGCCTGCTGTCCTGAGA	18
MELK	NM_014791	S4320/MELK.p1	SEQ ID NO:159	TCTTAGGAACGCCGTACCAGCCGC	24
MMP12	NM_002426	S4381/MMP12.f2	SEQ ID NO:160	CCAACGCTTGCCAAATCCT	19
MMP12	NM_002426	S4382/MMP12.r2	SEQ ID NO:161	ACGGTAGTGACAGCATCAAACTC	24
MMP12	NM_002426	S4383/MMP12.p2	SEQ ID NO:162	AACCAGCTCTCTGTGACCCCAATT	24
MMP9	NM_004994	S0656/MMP9.f1	SEQ ID NO:163	GAGAACCAATCTCACCGACA	20
MMP9	NM_004994	S0657/MMP9.r1	SEQ ID NO:164	CACCCGAGTGTAAACCATAGC	20
MMP9	NM_004994	S4760/MMP9.p1	SEQ ID NO:165	ACAGGTATTCCTCTGCCAGCTGCC	24
MYBL2	NM_002466	S3270/MYBL2.f1	SEQ ID NO:166	GCCGAGATCGCCAAGATG	18
MYBL2	NM_002466	S3271/MYBL2.r1	SEQ ID NO:167	CTTTTGATGGTAGAGTTCAGTGATTC	27
MYBL2	NM_002466	S4742/MYBL2.p1	SEQ ID NO:168	CAGCATTGTCTGTCTCCCTGGCA	24
NEK2	NM_002497	S4327/NEK2.f1	SEQ ID NO:169	GTGAGGCAGCGCGACTCT	18
NEK2	NM_002497	S4328/NEK2.r1	SEQ ID NO:170	TGCCAATGGTGTACAACACTTCA	23
NEK2	NM_002497	S4329/NEK2.p1	SEQ ID NO:171	TGCCTTCCCGGGCTGAGGACT	21
NME1	NM_000269	S2526/NME1.f3	SEQ ID NO:172	CCAACCCTGCAGACTCCAA	19
NME1	NM_000269	S2527/NME1.r3	SEQ ID NO:173	ATGTATAATGTTCTGCCAACTGTATG	28
NME1	NM_000269	S4949/NME1.p3	SEQ ID NO:174	CCTGGGACCATCCGTGGAGACTTCT	25
NPD009	NM_020686	S4474/NPD009.f3	SEQ ID NO:175	GGCTGTGGCTGAGGCTGTAG	20
NPD009	NM_020686	S4475/NPD009.r3	SEQ ID NO:176	GGAGCATTGAGGTCAAATCA	21
NPD009	NM_020686	S4476/NPD009.p3	SEQ ID NO:177	TTCCCAGAGTGTCTCACCTCCAGCAGAG	28
PCNA	NM_002592	S0447/PCNA.f2	SEQ ID NO:178	GAAGGTGTTGGAGGCACTCAAG	22
PCNA	NM_002592	S0448/PCNA.r2	SEQ ID NO:179	GGTTTACACCGCTGGAGCTAA	21
PCNA	NM_002592	S4784/PCNA.p2	SEQ ID NO:180	ATCCCAGCAGGCCTCGTTGATGAG	24
PR	NM_000926	S1336/PR.f6	SEQ ID NO:181	GCATCAGGCTGTCATTATGG	20
PR	NM_000926	S1337/PR.r6	SEQ ID NO:182	AGTAGTTGTGCTGCCCTTCC	20
PR	NM_000926	S4743/PR.p6	SEQ ID NO:183	TGTCTTACCTGTGGGAGCTGTAAGGTC	28
PREP	NM_002726	S1771/PREP.f1	SEQ ID NO:184	GGGACGGTGTTCACATTCAAG	21
PREP	NM_002726	S1772/PREP.r1	SEQ ID NO:185	CAGGATCCAGAAGTCAATGTTG	23
PREP	NM_002726	S4929/PREP.p1	SEQ ID NO:186	TCGCCAGTCTCCCAACTATCGCGT	24
PTTG1	NM_004219	S4525/PTTG1.f2	SEQ ID NO:187	GGCTACTCTGATCTATGTTGATAAGGAA	28
PTTG1	NM_004219	S4526/PTTG1.r2	SEQ ID NO:188	GCTTCAGCCCATCCTTAGCA	20
PTTG1	NM_004219	S4527/PTTG1.p2	SEQ ID NO:189	CACACGGGTGCCTGGTTCTCCA	22
RPLPO	NM_001002	S0256/RPLPO.f2	SEQ ID NO:190	CCATTCTATCATCAACGGGTACAA	24
RPLPO	NM_001002	S0258/RPLPO.r2	SEQ ID NO:191	TCAGCAAGTGGGAAGGTGTAATC	23
RPLPO	NM_001002	S4744/RPLPO.p2	SEQ ID NO:192	TCTCCACAGACAAGGCCAGGACTCG	25
SNRPF	NM_003095	S4489/SNRPF.f2	SEQ ID NO:193	GGCTGGTCCGCAGAGAGTAG	20
SNRPF	NM_003095	S4490/SNRPF.r2	SEQ ID NO:194	TGAGGAAAGGTTTGGGATTGA	21
SNRPF	NM_003095	S4491/SNRPF.p2	SEQ ID NO:195	AAACTCATGTAACCACGCCGAATGTTG	29
Src	NM_004383	S1820/Src.f2	SEQ ID NO:196	CCTGAACATGAAGGAGCTGA	20
Src	NM_004383	S1821/Src.r2	SEQ ID NO:197	CATCAGTCTCCGAACTCC	19
Src	NM_004383	S5034/Src.p2	SEQ ID NO:198	TCCCGATGGTCTGCAGCAGCT	21
STK15	NM_003600	S0794/STK1 5.f2	SEQ ID NO:199	CATCTTCCAGGAGGACCACT	20
STK15	NM_003600	S0795/STK15.r2	SEQ ID NO:200	TCCGACCTTCAATCAITTC	20

ES 2 690 168 T3

STK15	NM_003600	S4745/STK1 5.p2	SEQ ID NO:201	CTCTGTGGCACCCCTGGACTACCTG	24
STMY3	NM_005940	S2067/STMY3.f3	SEQ ID NO:202	CCTGGAGGCTGCAACATACC	20
STMY3	NM_005940	S2068/STMY3.r3	SEQ ID NO:203	TACAATGGCTTTGGAGGATAGCA	23
STMY3	NM_005940	S4746/STMY3.p3	SEQ ID NO:204	ATCCTCCTGAAGCCCTTTTCGCAGC	25
SURV	NM_001168	S02591/SURV.f2	SEQ ID NO:205	TGTTTTGATTCCCGGGCTTA	20
SURV	NM_001168	S0261/SURV.r2	SEQ ID NO:206	CAAAGCTGTCAGCTCTAGCAAAAG	24
SURV	NM_001168	S4747/SURV.p2	SEQ ID NO:207	TGCCTTCTCCTCCCTCACTTCTCACCT	28
TFRC	NM_003234	S1352/TFRC.f3	SEQ ID NO:208	GCCAACTGCTTTTCATTTGTG	20
TFRC	NM_003234	S1353/TFRC.r3	SEQ ID NO:209	ACTCAGGCCCATTTCCCTTA	20
TFRC	NM_003234	S4748/TFRC.p3	SEQ ID NO:210	AGGGATCTGAACCAATACAGAGCAGACA	28
TOP2A	NM_001067	S0271/TOP2A.f4	SEQ ID NO:211	AATCCAAGGGGGAGAGTGAT	20
TOP2A	NM_001067	S0273/TOP2A.r4	SEQ ID NO:212	GTACAGATTTTGCCCGAGGA	20
TOP2A	NM_001067	S4777/TOP2A.p4	SEQ ID NO:213	CATATGGACTTTGACTCAGCTGTGGC	26
TS	NM_001071	S0280/TS.f1	SEQ ID NO:214	GCCTCGGTGTGCCTTTCA	18
TS	NM_001071	S0282/TS.r1	SEQ ID NO:215	CGTGATGTGCGCAATCATG	19
TS	NM_001071	S4780/TS.p1	SEQ ID NO:216	CATCGCCAGCTACGCCCTGCTC	22

LISTA DE SECUENCIAS

- <110> Salud genómica

- 5 <120> Predicción de la probabilidad de recidiva de cáncer

- <130> GRF/FP6349666

- <140> 04809450.2
- 10 <141> 17-06-2004

- <150> PCT/US2004/019567
- <151> 17-06-2004

- 15 <150> US 60/482,339
- <151> 24-06-2003

- <160> 216

- 20 <170> FastSEQ para Windows Versión 4.0

- <210> 1
- <211> 80
- <212> ADN
- 25 <213> Secuencia artificial

- <220>
- <223> Ampliación de PCR

- 30 <400> 1
- ggctcttgtg cgtactgtcc ttcgggctgg tgacagggaa gacatcactg agcctgccat 60
- ctgtgctctt cgtcatctga 80

- <210> 2
- <211> 81
- 35 <212> ADN
- <213> Secuencia artificial

- <220>
- <223> Ampliación de PCR

ES 2 690 168 T3

<400> 2

cgttgtcagc acttgaata caagatggtt gccgggtcat gttaattggg aaaaagaaca 60
gtccacagga agaggttga c 81

<210> 3

5

<211> 76

<212> ADN

<213> Secuencia artificial

<220>

10

<223> Ampliación de PCR

<400> 3

cctgcaaaag ggaacaagag cccttcgctt ccagatggct ccctgcccgc ccccccgag 60
atcagagtca accacg 76

15

<210> 4

<211> 68

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Ampliación de PCR

<400> 4

ccgaggttaa tccagcacgt atggggccaa gtgtaggctc ccagcaggaa ctgagagcgc 60
catgtctt 68

25

<210> 5

<211> 65

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Ampliación de PCR

<400> 5

35

tcagctgtga gctgcccgata ccgcccggca atgggacctg ctcttaacct caaacctagg 60
accgt 65

<210> 6

ES 2 690 168 T3

<211> 84

<212> ADN

<213> Secuencia artificial

5 <220>

<223> Amplificación de PCR

<400> 6

ttcaggttgt tgcaggagac catgtacatg actgtctcca ttattgatcg gttcatgcag 60
aataattgtg tgccaagaa gatg 84

10

<210> 7

<211> 82

<212> ADN

<213> Secuencia artificial

15

<220>

<223> Amplificación de PCR

<400> 7

atgctgtggc tccttcctaa ctggggcttt cttgacatgt aggttgcttg gtaataacct 60
tttgtatat cacaatttgg gt 82

20

<210> 8

<211> 68

<212> ADN

25

<213> Secuencia artificial

<220>

<223> Amplificación de PCR

30 <400> 8

tggattggag ttctgggaat gtactggccg tggcactgga caacagtgtg tacctgtgga 60
gtgcaagc 68

<210> 9

<211> 81

35

<212> ADN

<213> Secuencia artificial

<220>

ES 2 690 168 T3

<223> Amplificación de PCR

<400> 9

tgagtgtccc ccggtatctt ccccgccctg ccaatcccga tgaattgga aattttattg 60
atgaaaatct gaaagcggct g 81

5

<210> 10

<211> 77

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Amplificación de PCR

<400> 10

tgacaatcag cacacctgca ttcaccgctc ggaagagggc ctgagctgca tgaataagga 60
tcacggctgt agtcaca 77

15

<210> 11

<211> 72

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Amplificación de PCR

<400> 11

tgccctgtggt gggaagctca gtaactggga accaaaggat gatgctatgt cagaacaccg 60
gaggcatttt cc 72

25

<210> 12

<211> 84

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Amplificación de PCR

35

<400> 12

tccctccact cggaaggact atcctgctgc caagagggtc aagttggaca gtgtcagagt 60
cctgagacag atcagcaaca accg 84

ES 2 690 168 T3

<210> 13
<211> 67
<212> ADN
5 <213> Secuencia artificial

<220>
<223> Amplificación de PCR

10 <400> 13
tgtctcactg agcgagcaga atctggtgga ctgttcgctg cctcaaggca atcagggctg 60
caatggt 67

<210> 14
<211> 74
15 <212> ADN
<213> Secuencia artificial

<220>
<223> Amplificación de PCR

20 <400> 14
tccattttct acctgttaac cttcatcatt ttgtgcaggc cctggaagca aagagaggaa 60
gggaccgact gcat 74

<210> 15
25 <211> 84
<212> ADN
<213> Secuencia artificial

<220>
30 <223> Amplificación de PCR

<400> 15
ctctgagaca gtgcttcgat gactttgcag acttggtgcc ctttgactcc tgggagccgc 60
tcatgaggaa gttgggcctc atgg 84

35 <210> 16
<211> 75
<212> ADN

ES 2 690 168 T3

<213> Secuencia artificial

<220>

<223> Amplificación de PCR

5

<400> 16

gggccctcca gaacaatgat gggctttatg atcctgactg cgatgagagc gggctcttta 60
aggccaagca gtgca 75

<210> 17

10 <211> 68

<212> ADN

<213> Secuencia artificial

<220>

15 <223> Amplificación de PCR

<400> 17

cgtaggtgcc ctctatgacc tgctgctgga gatgctggac gccaccgcc tacatgagcc 60
cactagcc 68

20 <210> 18

<211> 74

<212> ADN

<213> Secuencia artificial

25 <220>

<223> Amplificación de PCR

<400> 18

cacgggacat tcaccacatc gactactata aaaagacaac caacggccga ctgcctgtga 60
agtggatggc accc 74

30

<210> 19

<211> 82

<212> ADN

<213> Secuencia artificial

35

<220>

<223> Amplificación de PCR

ES 2 690 168 T3

<400> 19

ccaccccgag caaatctgtc ctccccagaa cccctgaatc ctggaggctc acgccccag 60
ccaagtagg gggactggat tt 82

<210> 20

5 <211> 67

<212> ADN

<213> Secuencia artificial

<220>

10 <223> Ampliación de PCR

<400> 20

ccatctgcat ccatcttgtt tgggctcccc acccttgaga agtgcctcag ataataccct 60
ggtggcc 67

15 <210> 21

<211> 86

<212> ADN

<213> Secuencia artificial

20 <220>

<223> Ampliación de PCR

<400> 21

aagctatgag gaaaagaagt acacgatggg ggacgctcct gattatgaca gaagccagtg 60
gctgaatgaa aaattcaagc tgggcc 86

25 <210> 22

<211> 76

<212> ADN

<213> Secuencia artificial

30 <220>

<223> Ampliación de PCR

<400> 22

35 caatgccatc ttgcgtaca tcgctcgcaa gcacaacatg tgtggtgaga ctgaagaaga 60
aaagattcga gtggac 76

<210> 23

ES 2 690 168 T3

<211> 70
<212> ADN
<213> Secuencia artificial

5 <220>
<223> Ampliación de PCR

<400> 23
cggtgtgaga agtgcagcaa gccctgtgcc cgagtgtgct atggctctggg catggagcac 60
ttgcgagagg 70

10 <210> 24
<211> 84
<212> ADN
<213> Secuencia artificial

15 <220>
<223> Ampliación de PCR

<400> 24
caagggagcg accaactgat cgcacacatg ctttgtttgg atatggagtg aacacaatta 60
tgtaccaaat ttaacttggc aaac 84

20 <210> 25
<211> 70
<212> ADN

25 <213> Secuencia artificial

<220>
<223> Ampliación de PCR

30 <400> 25
agaaccgcaa ggtgagcaag gtggagattc tccagcacgt catcgactac atcagggacc 60
ttcagttgga 70

<210> 26
<211> 83

35 <212> ADN
<213> Secuencia artificial

<220>

<223> Amplificación de PCR

<400> 26

gcatggtagc cgaagatttc acagtcaaaa tcggagattt tggtagacg cgagatatct 60
atgagacaga ctattaccgg aaa 83

5

<210> 27

<211> 79

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Amplificación de PCR

<400> 27

gatatgattg gtcgctgctt tgtgctcagc caggacctgg ccatccggga tgagttggat 60
ggtggggaat ggaagtct 79

15

<210> 28

<211> 80

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Amplificación de PCR

<400> 28

cggactttgg gtcgacttg acgagcggtg gttcgacaag tggccttgcg ggccggatcg 60
tcccagtgga agagttgtaa 80

25

<210> 29

<211> 78

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Amplificación de PCR

35

<400> 29

gcccagaggc tccatcgccc atcctcttcc tcccagtcg gctgaactct ccccttgctc 60
gcactgttca aacctctg 78

ES 2 690 168 T3

<210> 30
<211> 77
<212> ADN
<213> Secuencia artificial

5

<220>
<223> Amplicón de PCR

<400> 30

10 ccacctgcc atgatttttc ctttgaccgg gtattcccac caggaagtgg acaggatgaa 60
gtgtttgaag agattgc 77

<210> 31
<211> 69
<212> ADN
<213> Secuencia artificial

15

<220>
<223> Amplicón de PCR

<400> 31

20 tcagtggaga aggagtggga ccagtcaaca tctctgttgt cacaagcagt gtttcctctg 60
gatatggca 69

<210> 32
<211> 66
<212> ADN
<213> Secuencia artificial

25

<220>
<223> Amplicón de PCR

<400> 32

30 tgcaaacgct ggtgtcacag ccagccccc aactgacctc atctggaaga accagaactc 60
gtgggg 66

<210> 33
<211> 75
<212> ADN
<213> Secuencia artificial

35

ES 2 690 168 T3

<220>

<223> Ampliación de PCR

<400> 33

5 gacttttgcc cgctacctt cattccggcg tgacaacaat gagctggtgc tcttcatact 60
gaagcagtta gtggc 75

<210> 34

<211> 82

<212> ADN

10 <213> Secuencia artificial

<220>

<223> Ampliación de PCR

15 <400> 34

tgatggtcct atgtgtcaca ttcacacag gttcatacc aacacaggct tcagcacttc 60
ctttggtg tttcctgtcc ca 82

<210> 35

<211> 65

20 <212> ADN

<213> Secuencia artificial

<220>

<223> Ampliación de PCR

25

<400> 35

aaccggcgga tcgaaaagat tcttaggaac gccgtaccag ccgctctct caggacagca 60
ggccc 65

<210> 36

30 <211> 78

<212> ADN

<213> Secuencia artificial

<220>

35 <223> Ampliación de PCR

ES 2 690 168 T3

<400> 36

ccaacgcttg ccaaatcctg acaattcaga accagctctc tgtgaccca atttgagttt 60
tgatgctgtc actaccgt 78

<210> 37

5

<211> 67

<212> ADN

<213> Secuencia artificial

<220>

10

<223> Amplificación de PCR

<400> 37

gagaaccaat ctcaccgaca ggcagctggc agaggaatac ctgtaccgct atggttacac 60
tcgggtg 67

15

<210> 38

<211> 74

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Amplificación de PCR

<400> 38

gccgagatcg ccaagatggt gccagggagg acagacaatg ctgtgaagaa tcaactggaac 60
tctaccatca aaag 74

25

<210> 39

<211> 79

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Amplificación de PCR

<400> 39

35

gtgaggcagc gcgactctgg cgactggccg gccatgcctt cccgggctga ggactatgaa 60
gtgttgata ccattggca 79

<210> 40

<211> 74
<212> ADN
<213> Secuencia artificial

5 <220>
<223> Ampliación de PCR

<400> 40

ccaaccctgc agactccaag cctgggacca tccgtggaga cttctgcata caagttggca 60
ggaacattat acat 74

10

<210> 41
<211> 73
<212> ADN
<213> Secuencia artificial

15

<220>
<223> Ampliación de PCR

<400> 41

ggctgtggct gaggctgtag catctctgct ggaggtgaga cactctggga actgatttga 60
cctcgaatgc tcc 73

20

<210> 42
<211> 71
<212> ADN
<213> Secuencia artificial

25

<220>
<223> Ampliación de PCR

<400> 42

gaaggtgttg gaggcactca aggacctcat caacgaggcc tgctgggata ttagctccag 60
cgggtgtaaac c 71

30

<210> 43
<211> 85
<212> ADN
<213> Secuencia artificial

35

<220>

ES 2 690 168 T3

<223> Amplificación de PCR

<400> 43

gcatcaggct gtcattatgg tgccttacc tgtgggagct gtaaggctct cttaagagg 60
gcaatggaag ggcagcacia ctact 85

5

<210> 44

<211> 76

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Amplificación de PCR

<400> 44

gggacgggtg tcacattcaa gacgaatcgc cagtctccca actatcgcgt gatcaacatt 60
gacttctggg atcctg 76

15

<210> 45

<211> 74

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Amplificación de PCR

<400> 45

ggctactctg atctatggtg ataaggaaaa tggagaacca ggcacccgtg tggttgctaa 60
ggatgggctg aagc 74

25

<210> 46

<211> 75

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Amplificación de PCR

<400> 46

ccattctatc atcaacgggt acaaacgagt cctggccttg tctgtggaga cggattacac 60
cttcccactt gctga 75

35

<210> 47
 <211> 79
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> Ampliación de PCR

 10 <400> 47
 ggctggtcgg cagagagtag cctgcaacat tcggccgtgg ttacatgag ttaccacctc 60
 aatcccaaac ctttctca 79

 <210> 48
 <211> 64
 15 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Ampliación de PCR

 20 <400> 48
 cctgaacatg aaggagctga agctgctgca gaccatcggg aagggggagt tcggagacgt 60
 gatg 64

 <210> 49
 25 <211> 69
 <212> ADN
 <213> Secuencia artificial

 <220>
 30 <223> Ampliación de PCR

 <400> 49
 catcttccag gaggaccact ctctgtggca cctggacta cctgccccct gaaatgattg 60
 aaggtcggg 69

 35 <210> 50
 <211> 90
 <212> ADN
 <213> Secuencia artificial

ES 2 690 168 T3

<220>

<223> Ampliación de PCR

5

<400> 50

cctggaggct gcaacatacc tcaatcctgt cccaggccgg atcctcctga agcccttttc 60
gcagcactgc taccctcaa agccattgta 90

<210> 51

<211> 79

10

<212> ADN

<213> Secuencia artificial

<220>

<223> Ampliación de PCR

15

<400> 51

tgttttgatt cccgggctta ccaggtgaga agtgagggag gaagaaggca gtgtcccttt 60
tgctagagct gacagcttg 79

<210> 52

20

<211> 68

<212> ADN

<213> Secuencia artificial

<220>

25

<223> Ampliación de PCR

<400> 52

gccactgct ttcatttggt agggatctga accaatacag agcagacata aaggaaatgg 60
gcctgagt 68

30

<210> 53

<211> 72

<212> ADN

<213> Secuencia artificial

35

<220>

<223> Ampliación de PCR

<400> 53

ES 2 690 168 T3

aatccaaggg ggagagtgat gacttccata tggactttga ctcagctgtg gctcctcggg 60
caaatctgt ac 72

<210> 54

<211> 65

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Amplicón de PCR

10

<400> 54

gcctcgggtgt gcctttcaac atcgccagct acgccctgct cacgtacatg attgcgcaca 60
tcacg 65

<210> 55

15 <211> 22

<212> ADN

<213> Secuencia artificial

<220>

20 <223> Cebador-sonda de PCR

<400> 55

ggctcttgtg cgtactgtcc tt

22

<210> 56

<211> 23

25 <212> ADN

<213> Secuencia artificial

<220>

<223> Cebador-sonda de PCR

30

<400> 56

tcagatgacg aagagcacag atg

23

<210> 57

<211> 29

<212> ADN

35 <213> Secuencia artificial

<220>

<223> Cebador-sonda de PCR

<400> 57

aggctcagtg atgtcttccc tgtcaccag

29

5

<210> 58

<211> 23

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Cebador-sonda de PCR

<400> 58

cgttgtcagc acttggaata caa

23

15

<210> 59

<211> 24

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Cebador-sonda de PCR

<400> 59

gttcaacctc ttctgtgga ctgt

24

25

<210> 60

<211> 26

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Cebador-sonda de PCR

<400> 60

cccaattaac atgacccggc aaccat

26

<210> 61

<211> 20

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador-sonda de PCR	
	<400> 61	
	cctgcaaaag ggaacaagag	20
	<210> 62	
	<211> 20	
10	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15		
	<400> 62	
	cgtggttgac tctgatctcg	20
	<210> 63	
	<211> 21	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 63	
25	cttcgcctcc agatggctcc c	21
	<210> 64	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
30		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 64	
	ccgaggttaa tccagcacgt a	21

	<210> 65	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Cebador-sonda de PCR	
	<400> 65	
	aagacatggc gctctcagtt c	21
10	<210> 66	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 66	
	tgctgggagc ctacacttgg ccc	23
20	<210> 67	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 67	
	tcagctgtga gctgcggata	20
30	<210> 68	
	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> Cebador-sonda de PCR	

	<400> 68	
	acggtcctag gtttgaggtt aaga	24
	<210> 69	
	<211> 19	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 69	
10	caggtcccat tgccgggcg	19
	<210> 70	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 70	
	ttcaggttgt tgcaggagac	20
20	<210> 71	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 71	
	catcttcttg ggcacacaat	20
	<210> 72	
30	<211> 27	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	

<223> Cebador-sonda de PCR

<400> 72
tgtctccatt attgatcggc tcatgca **27**

<210> 73

5 <211> 22

<212> ADN

<213> Secuencia artificial

<220>

10 <223> Cebador-sonda de PCR

<400> 73
atgctgtggc tccttcctaa ct **22**

<210> 74

15 <211> 27

<212> ADN

<213> Secuencia artificial

<220>

20 <223> Cebador-sonda de PCR

<400> 74
accCAAattg tgatatacaa aaagggtt **27**

<210> 75

25 <211> 30

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador-sonda de PCR

<400> 75
taccaagcaa cctacatgtc aagaaagccc **30**

30 <210> 76

<211> 21

<212> ADN

<213> Secuencia artificial

	<220>	
	<223> Cebador-sonda de PCR	
5	<400> 76 tggattggag ttctgggaat g	21
	<210> 77	
	<211> 22	
	<212> ADN	
	<213> Secuencia artificial	
10	<220>	
	<223> Cebador-sonda de PCR	
	<400> 77 gcttgcactc cacaggtaca ca	22
15	<210> 78	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> Cebador-sonda de PCR	
	<400> 78 actggccgtg gcactggaca aca	23
	<210> 79	
25	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
30	<223> Cebador-sonda de PCR	
	<400> 79 tgagtgtccc ccggtatctt c	21
	<210> 80	
	<211> 21	
35	<212> ADN	

<213> Secuencia artificial

<220>

<223> Cebador-sonda de PCR

5

<400> 80
cagccgcttt cagattttca t **21**

<210> 81

<211> 27

<212> ADN

10 <213> Secuencia artificial

<220>

<223> Cebador-sonda de PCR

15 **<400> 81**
tgccaatccc gatgaaattg gaaattt **27**

<210> 82

<211> 21

<212> ADN

<213> Secuencia artificial

20 <220>

<223> Cebador-sonda de PCR

25 **<400> 82**
tgacaatcag cacacctgca t **21**

<210> 83

<211> 23

<212> ADN

<213> Secuencia artificial

30 <220>

<223> Cebador-sonda de PCR

35 **<400> 83**
tgtgactaca gccgtgatcc tta **23**

<210> 84

<211> 20

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador-sonda de PCR	
	<400> 84	
	caggccctct tccgagcgg	20
	<210> 85	
	<211> 18	
10	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 85	
15	tgccctgtggt gggaagct	18
	<210> 86	
	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> Cebador-sonda de PCR	
	<400> 86	
25	ggaaaatgcc tccggtgtt	19
	<210> 87	
	<211> 30	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Cebador-sonda de PCR	
	<400> 87	
	tgacatagca tcataccttg gttcccagtt	30
	<210> 88	
35	<211> 21	

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador-sonda de PCR	
	<400> 88	
	tccctccact cggaaggact a	21
	<210> 89	
	<211> 22	
10	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15	<400> 89	
	cggttggtgc tgatctgtct ca	22
	<210> 90	
	<211> 27	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 90	
25	tctgacactg tccaacttga ccctctt	27
	<210> 91	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Cebador-sonda de PCR	
	<400> 91	
	tgtctcactg agcgagcaga a	21
35	<210> 92	

	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Cebador-sonda de PCR	
	<400> 92	
	accattgcag ccctgattg	19
	<210> 93	
10	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> Cebador-sonda de PCR	
	<400> 93	
	cttgaggacg cgaacagtcc acca	24
	<210> 94	
	<211> 27	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
25	<400> 94	
	tccattttct acctgttaac cttcatc	27
	<210> 95	
	<211> 19	
	<212> ADN	
30	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 95	
35	atgcagtcgg tcccttcct	19

	<210> 96	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Cebador-sonda de PCR	
	<400> 96	
	ttgcttccag ggcctgcaca aaa	23
10	<210> 97	
	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> PCR primer-prope	
	<400> 97	
	ctctgagaca gtgcttcgat gact	24
20	<210> 98	
	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 98	
	ccatgaggcc caacttcct	19
30	<210> 99	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> Cebador-sonda de PCR	

	<400> 99	
	cagacttggt gccctttgac tcc	23
	<210> 100	
	<211> 20	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 100	
10	gggccctcca gaacaatgat	20
	<210> 101	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 101	
20	tgcaactgctt ggccttaaag a	21
	<210> 102	
	<211> 25	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 102	
	ccgctctcat cgcagtcagg atcat	25
30	<210> 103	
	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	

	<223> Cebador-sonda de PCR	
	<400> 103	
	cgtggtgccc ctctatgac	19
	<210> 104	
5	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
10	<223> Cebador-sonda de PCR	
	<400> 104	
	ggctagtggg cgcatgtag	19
	<210> 105	
	<211> 19	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
20		
	<400> 105	
	ctggagatgc tggacgccc	19
	<210> 106	
	<211> 20	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 106	
	cacgggacat tcaccacatc	20
30	<210> 107	
	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
35		

	<220>	
	<223> Cebador-sonda de PCR	
	<400> 107	
	gggtgccatc cacttcaca	19
5	<210> 108	
	<211> 27	
	<212> ADN	
	<213> Secuencia artificial	
10	<220>	
	<223> Cebador-sonda de PCR	
	<400> 108	
	ataaaaagac aaccaacggc cgactgc	27
	<210> 109	
15	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> Cebador-sonda de PCR	
	<400> 109	
	ccacccccgag caaatctgt	19
	<210> 110	
	<211> 22	
25	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
30	<400> 110	
	aatccagtc cccctacttt gg	22
	<210> 111	
	<211> 23	
	<212> ADN	

	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
5		
	<400> 111	
	cctgaatcct ggaggctcac gcc	23
	<210> 112	
	<211> 20	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15		
	<400> 112	
	ccatctgcat ccatcttggt	20
	<210> 113	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
20		
	<220>	
	<223> Cebador-sonda de PCR	
25		
	<400> 113	
	ggccaccagg gtattatctg	20
	<210> 114	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
30		
	<220>	
	<223> Cebador-sonda de PCR	
35		
	<400> 114	
	ctccccaccc ttgagaagtg cct	23
	<210> 115	
	<211> 20	

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador-sonda de PCR	
	<400> 115	
	ggcccagctt gaatttttca	20
	<210> 116	
	<211> 27	
10	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15		
	<400> 116	
	aagctatgag gaaaagaagt acacgat	27
	<210> 117	
	<211> 30	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
25		
	<400> 117	
	tcagccactg gcttctgtca taatcaggag	30
	<210> 118	
	<211> 21	
	<212> ADN	
30	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	

	<400> 118 caatgccatc ttgcgctaca t	21
	<210> 119 <211> 25 <212> ADN <213> Secuencia artificial	
5	<220> <223> Cebador-sonda de PCR	
	<400> 119 gtccactcga atcttttctt cttca	25
10	<210> 120 <211> 27 <212> ADN <213> Secuencia artificial	
15	<220> <223> Cebador-sonda de PCR	
	<400> 120 ctcgcaagca caacatgtgt ggtgaga	27
20	<210> 121 <211> 20 <212> ADN <213> Artificial sequence	
25	<220> <223> Cebador-sonda de PCR	
	<400> 121 cggtgtgaga agtgcagcaa	20
30	<210> 122 <211> 19 <212> ADN <213> Secuencia artificial	
35	<220>	

<223> Cebador-sonda de PCR

<400> 122

cctctcgcaa gtgctccat

19

5

<210> 123

<211> 24

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Cebador-sonda de PCR

<400> 123

ccagaccata gcacactcgg gcac

24

15

<210> 124

<211> 19

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Cebador-sonda de PCR

<400> 124

caagggagcg accaactga

19

25

<210> 125

<211> 28

<212> ADN

<213> Secuencia artificial

<220>

30

<223> Cebador-sonda de PCR

<400> 125

gtttgccaaag ttaaatttgg tacataat

28

<210> 126

<211> 28

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador-sonda de PCR	
	<400> 126	
	ctccatatcc aaacaaagca tgtgtgcg	28
	<210> 127	
	<211> 19	
10	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15		
	<400> 127	
	agaaccgcaa ggtgagcaa	19
	<210> 128	
	<211> 21	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 128	
	tccaactgaa ggtccctgat g	21
25	<210> 129	
	<211> 26	
	<212> ADN	
	<213> Secuencia artificial	
30		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 129	
	tggagattct ccagcacgtc atcgac	26
35		

	<210> 130	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
5		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 130	
	gcatggtagc cgaagatttc a	21
10		
	<210> 131	
	<211> 30	
	<212> ADN	
	<213> Secuencia artificial	
15		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 131	
	tttccggtaa tagtctgtct catagatatc	30
20		
	<210> 132	
	<211> 28	
	<212> ADN	
	<213> Secuencia artificial	
25		
	<220>	
	<223> Cebador-sonda de PCRs	
	<400> 132	
	cgcgatcatc caaaatctcc gattttga	28
30		
	<210> 133	
	<211> 22	
	<212> ADN	
	<213> Secuencia artificial	
35		
	<220>	

<223> Cebador-sonda de PCR

<400> 133

gatatgattg gtcgctgctt tg

22

5 <210> 134

<211> 21

<212> ADN

<213> Secuencia artificial

10 <220>

<223> Cebador-sonda de PCR

<400> 134

agaacttcca ttccccacca t

21

15 <210> 135

<211> 21

<212> ADN

<213> Secuencia artificial

20 <220>

<223> Cebador-sonda de PCR

<400> 135

cagccaggac ctggccatcc g

21

25 <210> 136

<211> 19

<212> ADN

<213> Secuencia artificial

30 <220>

<223> Cebador-sonda de PCR

<400> 136

cggactttgg gtgcgactt

19

35 <210> 137

	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Cebador-sonda de PCR	
	<400> 137	
	ttacaactct tccactggga cgat	24
10	<210> 138	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 138	
	ccacttgtcg aaccaccgct cgt	23
20	<210> 139	
	<211> 18	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 139	
	gcccagaggc tccatcgt	18
30	<210> 140	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> Cebador-sonda de PCR	

	<400> 140 cagaggtttg aacagtgcag aca	23
	<210> 141	
	<211> 23	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 141 cctcttcctc cccagtcggc tga	23
10	<210> 142	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 142 ccacctcgcc atgatttttc	20
20	<210> 143	
	<211> 25	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 143 gcaatctctt caaacacttc atcct	25
	<210> 144	
30	<211> 25	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	

	<223> Cebador-sonda de PCR	
	<400> 144	
	tttgaccggg tattcccacc aggaa	25
	<210> 145	
5	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
10	<223> Cebador-sonda de PCR	
	<400> 145	
	tcagtggaga aggagttgga	20
	<210> 146	
	<211> 20	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
20		
	<400> 146	
	tgccatatcc agaggaaaca	20
	<210> 147	
	<211> 28	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 147	
30	ccagtcaaca tctctgttgt cacaagca	28
	<210> 148	
	<211> 19	
	<212> ADN	
35	<213> Secuencia artificial	

	<220>	
	<223> Cebador-sonda de PCR	
5	<400> 148 tgcaaacgct ggtgtcaca	19
	<210> 149	
	<211> 21	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15	<400> 149 ccccacgagt tctggttctt c	21
	<210> 150	
	<211> 22	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
25	<400> 150 cagccccca actgacctca tc	22
	<210> 151	
	<211> 21	
	<212> ADN	
30	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
35	<400> 151 gacttttgcc cgctaccttt c	21

<210> 152
 <211> 26
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Cebador-sonda de PCR

<400> 152
gccactaact gcttcagtat gaagag **26**
 10
 <210> 153
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 15
 <220>
 <223> Cebador-sonda de PCR

<400> 153
acagctcatt gttgtcacgc cgga **24**
 20
 <210> 154
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 25
 <220>
 <223> Cebador-sonda de PCR

<400> 154
tgatggctct atgtgtcaca ttca **24**
 30
 <210> 155
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 35
 <220>
 <223> Cebador-sonda de PCR ,

	<400> 155 tgggacagga aacacaccaa	20
	<210> 156 <211> 30 <212> ADN <213> Secuencia artificial	
5		
	<220> <223> Cebador-sonda de PCR	
10		
	<400> 156 caggtttcat accaacacag gcttcagcac	30
	<210> 157 <211> 18 <212> ADN <213> Secuencia artificial	
15		
	<220> <223> Cebador-sonda de PCR	
	<400> 157 aaccggcgga tcgaaaag	18
20		
	<210> 158 <211> 18 <212> ADN <213> Secuencia artificial	
25		
	<220> <223> Cebador-sonda de PCR	
	<400> 158 gggcctgctg tcctgaga	18
30		
	<210> 159 <211> 24 <212> ADN <213> Secuencia artificial	
35		
	<220>	

<223> Cebador-sonda de PCR

<400> 159

tcttaggaac gccgtaccag ccgc

24

5

<210> 160

<211> 19

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Cebador-sonda de PCR

<400> 160

ccaacgcttg ccaaatcct

19

15

<210> 161

<211> 24

<212> ADN

<213> Artificial sequence

<220>

20

<223> Cebador-sonda de PCR

<400> 161

acggtagtga cagcatcaaa actc

24

25

<210> 162

<211> 24

<212> ADN

<213> Secuencia artificial

<220>

30

<223> Cebador-sonda de PCR

<400> 162

aaccagctct ctgtgacccc aatt

24

35

<210> 163

<211> 20

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador-sonda de PCR	
	<400> 163	
	gagaaccaat ctcaccgaca	20
	<210> 164	
10	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> Cebador-sonda de PCR	
	<400> 164	
	cacccgagtg taaccatagc	20
	<210> 165	
	<211> 24	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
25		
	<400> 165	
	acaggtattc ctctgccagc tgcc	24
	<210> 166	
	<211> 18	
30	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
35		

	<400> 166	
	gccgagatcg ccaagatg	18
	<210> 167	
	<211> 27	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 167	
10	cttttgatgg tagagttcca gtgattc	27
	<210> 168	
	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 168	
20	cagcattgtc tgtcctccct ggca	24
	<210> 169	
	<211> 18	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 169	
30	gtgaggcagc gcgactct	18
	<210> 170	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	

	<223> Cebador-sonda de PCR	
	<400> 170	
	tgccaatggt gtacaacact tca	23
	<210> 171	
5	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
10	<223> Cebador-sonda de PCR	
	<400> 171	
	tgccttcccg ggctgaggac t	21
	<210> 172	
	<211> 19	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
20		
	<400> 172	
	ccaaccctgc agactccaa	19
	<210> 173	
	<211> 28	
25	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
30		
	<400> 173	
	atgtataatg ttctgccaa cttgtatg	28
	<210> 174	
	<211> 25	
35	<212> ADN	

	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
5		
	<400> 174	
	cctgggacca tccgtggaga cttct	25
	<210> 175	
	<211> 20	
10	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15		
	<400> 175	
	ggctgtggct gaggctgtag	20
	<210> 176	
	<211> 21	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 176	
25	ggagcattcg aggtcaaatc a	21
	<210> 177	
	<211> 28	
	<212> ADN	
	<213> Secuencia artificial	
30		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 177	
	ttcccagagt gtctcacctc cagcagag	28

	<210> 178	
	<211> 22	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> Cebador-sonda de PCR	
	<400> 178	
	gaaggtgttg gaggcactca ag	22
10	<210> 179	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 179	
	ggtttacacc gctggagcta a	21
20	<210> 180	
	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 180	
	atcccagcag gcctcgttga tgag	24
30	<210> 181	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> Cebador-sonda de PCR	

	<400> 181 gcatcaggct gtcattatgg	20
	<210> 182 <211> 20 <212> ADN <213> Secuencia artificial	
5		
	<220> <223> Cebador-sonda de PCR	
10		
	<400> 182 agtagttgtg ctgcccttcc	20
	<210> 183 <211> 28 <212> ADN <213> Secuencia artificial	
15		
	<220> <223> Cebador-sonda de PCR	
20		
	<400> 183 tgtccttacc tgtgggagct gtaaggtc	28
	<210> 184 <211> 21 <212> ADN <213> Secuencia artificial	
25		
	<220> <223> Cebador-sonda de PCR	
	<400> 184 gggacggtgt tcacattcaa g	21
30		
	<210> 185 <211> 23 <212> ADN <213> Secuencia artificial	
35		

	<220>	
	<223> Cebador-sonda de PCR	
	<400> 185	
	caggatccca gaagtcaatg ttg	23
5	<210> 186	
	<211> 24	
	<212> ADN	
	<213> Secuencia artificial	
10	<220>	
	<223> Cebador-sonda de PCR	
	<400> 186	
	tcgccagtct cccaactatc gcgt	24
15	<210> 187	
	<211> 28	
	<212> ADN	
	<213> Secuencia artificial	
20	<220>	
	<223> Cebador-sonda de PCR	
	<400> 187	
	ggctactctg atctatgttg ataaggaa	28
25	<210> 188	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Cebador-sonda de PCR	
	<400> 188	
	gcttcagccc atccttagca	20
	<210> 189	
	<211> 22	
	<212> ADN	

	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
5		
	<400> 189	
	cacacgggtg cctggttctc ca	22
	<210> 190	
	<211> 24	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15		
	<400> 190	
	ccattctatc atcaacgggt acaa	24
	<210> 191	
	<211> 23	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
25		
	<400> 191	
	tcagcaagtg ggaaggtgta atc	23
	<210> 192	
	<211> 25	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Cebador-sonda de PCR	
35		
	<400> 192	
	tctccacaga caaggccagg actcg	25
	<210> 193	
	<211> 20	

	<212> ADN	
	<213> Secuencia artificial	
	<220>	
5	<223> Cebador-sonda de PCR	
	<400> 193	
	ggctggtcgg cagagagtag	20
	<210> 194	
10	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> Cebador-sonda de PCR	
	<400> 194	
	tgaggaaagg tttgggattg a	21
	<210> 195	
	<211> 29	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
25		
	<400> 195	
	aaactcatgt aaaccacggc cgaatggtg	29
	<210> 196	
	<211> 20	
	<212> ADN	
30	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 196	
35	cctgaacatg aaggagctga	20

	<210> 197	
	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
5		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 197	
	catcacgtct ccgaactcc	19
10		
	<210> 198	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
15		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 198	
	tcccgatggt ctgcagcagc t	21
20		
	<210> 199	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
25		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 199	
	catcttccag gaggaccact	20
30		
	<210> 200	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
35		
	<220>	
	<223> Cebador-sonda de PCR	

	<400> 200	
	tccgaccttc aatcatttca	20
	<210> 201	
	<211> 24	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 201	
10	ctctgtggca ccctggacta cctg	24
	<210> 202	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> Cebador-sonda de PCR	
	<400> 202	
	cctggaggct gcaacatacc	20
20	<210> 203	
	<211> 23	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> Cebador-sonda de PCR	
	<400> 203	
	tacaatggct ttggaggata gca	23
	<210> 204	
30	<211> 25	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
35	<223> Cebador-sonda de PCR	

	<400> 204	
	atcctcctga agcccttttc gcagc	25
	<210> 205	
	<211> 20	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
10		
	<400> 205	
	tgttttgatt cccgggctta	20
	<210> 206	
	<211> 24	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 206	
20	caaagctgtc agctctagca aaag	24
	<210> 207	
	<211> 28	
	<212> ADN	
	<213> Secuencia artificial	
25		
	<220>	
	<223> Cebador-sonda de PCR	
	<400> 207	
	tgctttcttc ctccctcact tctcacct	28
30	<210> 208	
	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	

	<223> Cebador-sonda de PCR	
	<400> 208	
	gccaaactgct ttcatttgtg	20
	<210> 209	
5	<211> 20	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
10	<223> Cebador-sonda de PCR	
	<400> 209	
	actcaggccc atttccttta	20
	<210> 210	
15	<211> 28	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> Cebador-sonda de PCR	
	<400> 210	
	agggatctga accaatacag agcagaca	28
	<210> 211	
	<211> 20	
25	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
30		
	<400> 211	
	aatccaaggg ggagagtgat	20
	<210> 212	
	<211> 20	
	<212> ADN	
35	<213> Secuencia artificial	

	<220>	
	<223> Cebador-sonda de PCR	
5	<400> 212 gtacagattt tgccccgagga	20
	<210> 213	
	<211> 26	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
15	<400> 213 catatggact ttgactcagc tgtggc	26
	<210> 214	
	<211> 18	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
25	<400> 214 gcctcgggtgt gcctttca	18
	<210> 215	
	<211> 19	
	<212> ADN	
30	<213> Secuencia artificial	
	<220>	
	<223> Cebador-sonda de PCR	
35	<400> 215 cgtgatgtgc gcaatcatg	19
	<210> 216	
	<211> 22	

<212> ADN

<213> Secuencia artificial

<220>

5

<223> Cebador-sonda de PCR

<400> 216

catcgccagc tacgccctgc tc

22

REIVINDICACIONES

1. Un método de predicción de la probabilidad de supervivencia a largo plazo de un paciente con cáncer de mama ductal con ganglios negativos, ER positivo, sin recidiva del cáncer, que consiste en determinar el nivel de expresión del transcrito de ARN de CDC20, o su producto de expresión, en una muestra de célula o tejido de cáncer de mama obtenida del paciente, normalizado frente al nivel de expresión de un conjunto de referencia de transcritos de ARN, o sus productos de expresión, comparando el nivel de expresión normalizado de CDC20 del paciente con un nivel de expresión normalizado de CDC20 en un grupo de referencia de tejido de cáncer de mama que comprende pacientes (a) vivos sin recidiva de cáncer de mama local, regional o distante, (b) vivos con recidiva de cáncer de mama contralateral, (c) vivos con un segundo cáncer primario distinto del de mama, o (d) murieron antes de la recidiva de cáncer de mama, en el que un nivel de expresión normalizado aumentado de CDC20 del paciente indica una probabilidad reducida de supervivencia a largo plazo sin recidiva del cáncer.
5
2. Un método como se ha indicado en la reivindicación 1, en el que dicho ARN comprende ARN intrónico.
3. Un método como se ha indicado en la reivindicación 1, en el que dicho ARN se aísla de una muestra de tejido de cáncer de mama embebida en cera de dicho paciente.
10
4. Un método como se ha indicado en la reivindicación 1, en el que dicho ARN se aísla del tejido biopsiado con aguja gruesa.
15
5. Un método como se ha indicado en la reivindicación 1, en el que dicho ARN se aísla de células aspiradas con aguja fina.
6. Un método como se ha indicado en cualquier reivindicación anterior, en el que el nivel de expresión del mencionado transcrito de ARN se determina mediante reacción en cadena de la polimerasa con transcripción inversa (RT-PCR).
20