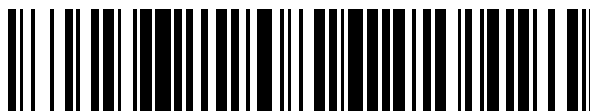


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 787 475**

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 18190058 (0)**

97 Fecha y número de publicación de la concesión europea: **01.04.2020** **EP 3470535**

54 Título: **Predicción de probabilidad de recurrencia del cáncer**

30 Prioridad:

24.06.2003 US 482339 P

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

16.10.2020

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:

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

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 787 475 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 probabilidad de recurrencia del cáncer

Antecedentes de la invención

Campo de la invención

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

Descripción de la técnica relacionada

- 10 Los oncólogos disponen de diversas opciones de tratamiento, incluyendo distintas combinaciones de fármacos quimioterapéuticos caracterizados como "protocolo de tratamiento", y diversos fármacos que no tienen indicación establecida para un cáncer en particular, pero para los que existe evidencia de eficacia en ese cáncer. Una mayor probabilidad de un buen resultado del tratamiento requiere que los pacientes sean asignados a un tratamiento del cáncer disponible óptimo, y que esa asignación se efectúe lo antes posible tras el diagnóstico.

- 15 Actualmente, los test de diagnóstico utilizados en la práctica clínica son de un solo analito, y por consiguiente no capturan el valor potencial de conocer las relaciones entre docenas de marcadores diferentes. Además, frecuentemente los tests de diagnóstico no son cuantitativos, basándose en la inmunohistoquímica. Este método a menudo aporta distintos resultados en diferentes laboratorios, en parte porque los reactivos no están estandarizados, y en parte porque las interpretaciones son subjetivas y no fácilmente cuantificadas. Los tests basados en ARN con frecuencia no se usan debido al problema de degradación del ARN con el paso del tiempo, y el hecho de que resulte difícil obtener muestras de tejido frescas de los pacientes para su análisis. El tejido embebido en parafina fijado es más fácilmente obtenible, y se han establecido métodos para detectar el ARN en tejido fijado. No obstante, típicamente estos métodos no permiten el estudio de grandes cifras de genes (ADN o ARN) de pequeñas cantidades de material. Por tanto, tradicionalmente se ha utilizado poco el tejido fijado, aparte de para la detección inmunohistoquímica de proteínas.

- 20 En los últimos años, varios grupos han publicado estudios sobre la clasificación de ciertos tipos de cáncer por análisis de expresión génica por microarrays (ver, p.ej., Golub et al., *Science* 286:531-537 (1999); Bhattacharjee et al., *Proc. Natl. Acad. Sci. USA* 98:13790-13795 (2001); Chen-Hsiang et al., *Bioinformatics* 17 (Suppl. 1):S316-S322 (2001); Ramaswamy et al., *Proc. Natl. Acad. Sci. USA* 98:15149-15154 (2001)). Se ha informado también de algunas clasificaciones de cánceres de mama humanos, basados en patrones de expresión génica (Martin et al., *Cancer Res.* 60:2232-2238 (2000); West et al., *Proc. Natl. Acad. Sci. USA* 98:11462-11467 (2001); Sorlie et al., *Proc. Natl. Acad. Sci. USA* 98:10869-10874 (2001); Yan et al., *Cancer Res.* 61:8375-8380 (2001)). No obstante, esos estudios se enfocan principalmente en mejorar y refinar la clasificación ya establecida de varios tipos de cáncer, incluyendo el cáncer de mama, y en general no proporcionan nueva información sobre las relaciones de los genes expresados diferencialmente, y no relacionan los hallazgos con estrategias de tratamiento, para mejorar los resultados clínicos de la terapia del cáncer.

- 35 Aunque la biología y bioquímica molecular modernas han revelado cientos de genes cuyas actividades influyen en el comportamiento de las células tumorales, el estado de su diferenciación y su sensibilidad o resistencia a determinados fármacos terapéuticos, con contadas excepciones, el estado de esos genes no ha sido explotado a efectos de la toma rutinaria de decisiones clínicas sobre tratamientos con fármacos. Una notable excepción es el uso de la expresión de la proteína del receptor de estrógenos (ER) en los carcinomas de mama para seleccionar pacientes para el tratamiento con fármacos antiestrogénicos, como tamoxifen. Otro ejemplo excepcional es el uso de la expresión de la 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 del cáncer sigue siendo objetivar pautas de tratamiento específicas para tipos de tumores patogénicamente distintos, y en última instancia personalizar el tratamiento del tumor para maximizar los resultados. Por consiguiente, se necesitan tests que proporcionen simultáneamente información predictiva acerca de la respuesta del paciente a las diversas opciones de tratamiento. Esto es especialmente cierto en el cáncer de mama, cuya biología es poco conocida. Es evidente que la clasificación del cáncer de mama en algunos subgrupos, como el subgrupo ErbB2+, y subgrupos caracterizados por una baja o ausente expresión génica del receptor de estrógenos (ER), y algunos factores transcripcionales adicionales (Perou et al., *Nature* 406:747-752 (2000)) no refleja la heterogeneidad celular y molecular del cáncer de mama, y no facilita el diseño de estrategias de tratamiento que maximicen la respuesta del paciente.

- 50 En particular, cuando a una paciente se le diagnostica un cáncer, como cáncer de mama u ovario, se necesitan en gran medida métodos que permitan al médico predecir el curso esperado de la enfermedad, incluyendo la probabilidad de recurrencia del cáncer, la supervivencia a largo plazo de la paciente, y similares, y seleccionar en consecuencia la opción de tratamiento más apropiada.

Bos et al. (Cancer, Vol. 97, No 6, 1573-1518, 2003) divulgan el uso de ARN HIF-a para predecir la probabilidad de supervivencia libre de enfermedad de una paciente con cáncer de mama nódulo negativo ductal invasivo. US2002/156263 divulga MELK (KIAA0175) como uno de un gran número de genes que se ha observado que se expresan diferencialmente en el cáncer de mama.

5 Resumen de la invención

La presente invención viene definida por las reivindicaciones adjuntas 1 a 8.

La presente invención considera el uso de material de biopsia embebido en parafina guardado, para el ensayo de todos los marcadores de la serie, y por consiguiente es compatible con el tipo de material de biopsia más ampliamente disponible. También es compatible con otros métodos distintos de obtención de tejido tumoral, por ejemplo, por biopsia con aguja gruesa, o aspiración con aguja fina. Además, para cada miembro de la serie de genes, la invención especifica las secuencias de oligonucleótidos que se pueden utilizar en el test.

En un aspecto, la presente invención se refiere a un método para predecir la probabilidad de supervivencia libre de enfermedad de una paciente con cáncer de mama nódulo negativo, ER positivo, ductal invasivo, comprendiendo la determinación del nivel de expresión de la transcripción de ARN de MELK en una célula o una muestra de tejido del cáncer de mama obtenidos de la paciente, normalizado contra el nivel de expresión de una serie de referencia de transcripciones de ARN, comparando el nivel de expresión normalizado de MELK de la paciente, con un nivel normalizado de expresión de MELK en una serie de referencia de tejido de cáncer de mama, comprendiendo pacientes que estaban (a) vivas sin recurrencia del cáncer de mama local, regional o distante, (b) vivas con recurrencia contralateral de cáncer de mama, (c) vivas con segundo cáncer primario no de mama, o (d) fallecieron antes de la recurrencia del cáncer de mama, donde un nivel aumentado de la expresión normalizada de MELK en la paciente indica una menor probabilidad de supervivencia libre de enfermedad.

En varias 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 transcripciones de ARN de pronóstico o sus productos de expresión.

En otra realización, el cáncer es cáncer de mama o cáncer de ovario.

El cáncer es cáncer de mama nódulo negativo, ER positivo. En otra realización, el ARN comprende ARN intrónico. Se determina el nivel de expresión de la transcripción de ARN de MELK, donde la expresión de MELK indica una menor probabilidad de supervivencia a largo plazo sin recurrencia del cáncer.

La invención consiste en un método de predicción de la probabilidad de supervivencia a largo plazo de una paciente con cáncer, sin recurrencia del cáncer, comprendiendo la determinación del nivel de expresión de transcripciones de ARN de MELK, en una célula cancerosa obtenida de dicha paciente, normalizado contra el nivel de expresión de todas las transcripciones de ARN o sus productos en la célula cancerosa, o de una serie de referencia de transcripciones de ARN o sus productos de expresión, donde el incremento en la expresión de MELK indica una menor probabilidad de supervivencia a largo plazo sin recurrencia del cáncer.

En una realización de este método, el ARN se aísla de una muestra de tejido del cáncer de mama de la paciente, se embebe en cera y se fija.

En otra realización, el ARN se aísla del tejido biopsiado con aguja gruesa, o células aspiradas con aguja fina.

En un aspecto diferente, la divulgación se refiere a un array que comprende la hibridación de polinucleótidos con dos o más de los siguientes genes: B_Catenina; 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, inmovilizado en una superficie sólida.

En un aspecto de la divulgación, el array comprende la hibridación de polinucleótidos en 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_Catenina, EstR1, CDH1, TS HER2, y cMYC.

En otro aspecto, el array comprende la hibridación de polinucleótidos en 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 otro aspecto, los arrays comprenden la hibridación de polinucleótidos en 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 listados.

En otro aspecto más, los arrays comprenden la hibridación de polinucleótidos en todos los genes listados.

Y aún en otro aspecto más, los arrays comprenden más de una hibridación de polinucleótidos en el mismo gen.

En un aspecto adicional, os arrays comprenden secuencias basadas en intrones.

En otro aspecto, los polinucleótidos son ADNc (ADN complementario), que pueden, por ejemplo, tener aproximadamente una longitud de 500 a 5000 bases.

5 En otro aspecto más, los polinucleótidos son oligonucleótidos, que pueden, por ejemplo, tener aproximadamente una longitud de 20 a 80 bases.

Los arrays pueden, por ejemplo, inmovilizarse sobre vidrio, y pueden contener cientos de miles, por ej., 330.000 oligonucleótidos.

10 En otro aspecto, la divulgación se refiere a un método de predicción de la probabilidad de supervivencia a largo plazo de una paciente diagnosticada de cáncer de mama invasivo, sin recurrencia del cáncer de mama, comprendiendo los pasos de

15 (a) determinar los niveles de expresión de las transcripciones de ARN, o los productos de expresión de genes de una serie de genes seleccionado del grupo compuesto por B_Catenina; BAG1; BIN1; BUB1; C20_orf1; CCNB1; CCNE2; CDC20; CDH1; CEGP1; CIAP1; cMYC; CTS2L2; DKFZp586M07; DR5; EpCAM; EstR1; FÖXM1; GRB7; GSTM1; GSTM3; HER2; HNRPA; ID1; IGF1R; ITGA7; Ki_67; KNSL2; LMNB1; MCM2; MELK; MMP12; MMP9; MYBL2; NEK2; NME1; N-PD009; PCNA; PR; PREP; PTTG1; RPLPO; Src; STK15; STMY3; SURV; TFRC; TOP2A; y TS en una célula de cáncer de mama obtenida de la paciente, normalizado contra los niveles de expresión de todas las transcripciones de ARN o sus productos de expresión en dicha célula de cáncer de mama, o de una serie de referencia de transcripciones de ARN o sus productos;

20 (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.

En otro aspecto más, la divulgación se refiere a un método de preparación de un perfil genómico personalizado para una paciente, comprendiendo los pasos de

25 (a) someter ARN extraído de un tejido de mama obtenido de la paciente al análisis de expresión génica;

(b) determinar el nivel de expresión en el tejido de uno o más genes seleccionados de la serie de genes del cáncer de mama listados en cualquiera de las Tablas 1 y 2, donde el nivel de expresión está normalizado con respecto a un gen o genes de control, y opcionalmente se compara con la cantidad hallada en un conjunto de tejidos de referencia del cáncer de mama; y

30 (c) crear un informe resumiendo los datos obtenidos en dicho análisis de expresión génica. El tejido mamario puede comprender células de cáncer de mama.

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

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

35 En otro aspecto, la divulgación se refiere a un método para medir los niveles de productos de RNAm de los genes listados en las Tablas 1 y 2 por reacción de la cadena de polimerasa en tiempo real (RT-PCR), utilizando un amplicón listado en la Tabla, y un conjunto y un conjunto de sonda-cebador indicado en las Tablas 4A - 4D. 25

En otro aspecto más, la divulgación se refiere a un conjunto de sonda-cebador de PCR listado en las Tablas 4A - 4D, y un amplicón de PCR listado en la Tabla 3.

40

A. Definiciones

45 Salvo que se defina de otro modo, todos los términos técnicos y científicos utilizados aquí tienen el mismo significado que el que comúnmente entienden los expertos en la técnica a la que pertenece esta invención. Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, NY 1994), y March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4th ed., John Wiley & Sons (New York, NY 1992), proporcionan a los expertos en la técnica una guía general sobre muchos de los términos utilizados en la presente solicitud.

Un experto en la técnica conocerá muchos métodos y materiales similares a los que se describen aquí, que podrían utilizarse en la práctica de la presente invención. La presente invención no se limita en modo alguno a los métodos y materiales descritos. A efectos de la presente invención, se definen a continuación los siguientes términos.

5 El término "microarray" se refiere a una disposición ordenada de elementos de array hibridables, de preferencia sondas de polinucleótidos, sobre un sustrato.

10 El término "polinucleótido", utilizado en singular o plural, se refiere en general a cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ARN o ADN no modificado, o ARN o ADN modificado. Así, por ejemplo, los polinucleótidos, como se definen aquí, incluyen entre otros, ADN de cadena simple y doble, ADN incluyendo regiones de cadena simple y doble, ARN de cadena simple y doble y ARN incluyendo regiones de cadena simple y doble, moléculas híbridas comprendiendo ADN y ARN que pueden ser de cadena simple o, más típicamente, de cadena doble, o incluir 15 regiones de cadena triple comprendiendo ARN o ADN, o ARN y ADN. Las cadenas de tales regiones pueden ser de la misma molécula o de moléculas distintas. Las regiones pueden incluir todas las de una o más de las moléculas, pero más típicamente incluyen solo una región de algunas de las moléculas. Una de las moléculas de una región de triple hélice es frecuentemente un oligonucleótido. El término "polinucleótido" incluye específicamente ADNcs. El término incluye ADNcs (incluyendo ADNcs) y ARNs que contienen una o más bases modificadas. Así, tal como se aplica el término aquí, ADNcs o ARNs con cadenas principales modificadas, por razones de estabilidad u otros motivos, son "polinucleótidos". Además, ADNcs o ARNs comprendiendo bases inusuales, como la inosina, o bases modificadas, como bases tritriadas, se incluyen en el término "polinucleótidos", como se define aquí. En general, el término "polinucleótidos" comprende todas las formas 20 modificadas química, enzimática y/o metabólicamente de polinucleótidos no modificados, así como las formas químicas de ADN y ARN características de los virus y las células, incluyendo las células simples y complejas.

25 El término "oligonucleótido" se refiere a un polinucleótido relativamente corto, incluyendo, entre otros, desoxirribonucleótidos de cadena simple, ribonucleótidos de cadena simple o doble, híbridos de ARN:ADN y ADNcs de cadena doble. Los oligonucleótidos, como los oligonucleótidos sonda de ADN de cadena simple, son frecuentemente sintetizados por métodos químicos, por ejemplo, sintetizadores de oligonucleótidos automáticos usmg, disponibles comercialmente. No obstante, los oligonucleótidos pueden hacerse por otros varios métodos, incluyendo técnicas in vitro mediadas por ADN recombinante, y por expresión de ADNcs en células y organismos.

30 Los términos "gen expresado diferencialmente", "expresión génica diferencial" y sus sinónimos, que se utilizan intercambiamente, se refieren a un gen cuya expresión es activada a un nivel superior o inferior, en un sujeto que padece una enfermedad, específicamente cáncer, como cáncer de mama, en relación con su expresión en un sujeto normal o de control. Los términos incluyen también genes cuya expresión es activada a un nivel superior o inferior en distintos estadios de la misma enfermedad. Se entiende también que un gen expresado diferencialmente puede ser activado o inhibido a nivel del ácido nucleico o a nivel de proteína, o puede ser sujeto a un empalme alternativo para dar un producto de polipéptido distinto. Tales diferencias pueden ponerse de manifiesto, por ejemplo, por un cambio en los 35 niveles de ARNm, expresión de superficie, secreción u otro fraccionamiento de un polipéptido. La expresión génica diferencial puede incluir una comparación de expresión entre dos o más genes, o sus productos génicos, o una comparación de los cocientes de la expresión entre dos o más genes, o sus productos génicos, o incluso una comparación de dos productos procesados diferentemente del mismo gen, que difieran entre sujetos normales y sujetos con una enfermedad, específicamente cáncer, o entre varios estadios de la misma enfermedad. La expresión diferencial incluye diferencias cuantitativas y cualitativas en el patrón de expresión temporal o celular de un gen, o sus productos de expresión 40 entre, por ejemplo, células normales o enfermas, o entre células que han pasado por distintos episodios de la enfermedad, o estadios de la misma. A los efectos de esta invención, la "expresión génica diferencial" se considera presente cuando hay una diferencia aproximada de al menos el doble, preferiblemente de al menos unas cuatro veces, más preferiblemente de al menos unas seis veces, y aún más preferiblemente de al menos unas diez veces entre la expresión de un gen determinado en sujetos normales y enfermos, o en varios estadios del desarrollo de la enfermedad en un sujeto enfermo. 45

El término "sobreexpresión" con respecto a una transcripción de ARN, se utiliza para referirse al nivel de la transcripción determinado por normalización respecto al nivel de referencia de ARNms, que pueden ser todas las transcripciones medidas en la muestra, o una serie de referencia de ARNms particular.

50 La frase "amplificación génica" se refiere a un proceso por el que se forman múltiples copias de un gen o fragmento de gen en una célula o línea celular particular. La región duplicada (un tramo de ADN amplificado) se denomina habitualmente "amplicón". Habitualmente, la cantidad de ARN mensajero (ARNm) producida, es decir, el nivel de expresión génica, aumenta también en proporción al número de copias hecho del gen particular expresado.

55 El término "pronóstico" se utiliza aquí con referencia a la predicción de probabilidad de progresión o muerte atribuible al cáncer, incluyendo recurrencia, diseminación metastásica y resistencia a fármacos, de una enfermedad neoplásica, como el cáncer de mama. El término "predicción" se usa aquí para referirse a la probabilidad de que un paciente responda de forma favorable o desfavorable a un fármaco o serie de fármacos, y también a la amplitud de esas respuestas, o de que un paciente sobreviva, tras la extracción quirúrgica del tumor primario y/o quimioterapia, durante determinado periodo de

tiempo, sin recurrencia del cáncer. Los métodos predictivos de la presente invención pueden ser utilizados clínicamente para adoptar decisiones de tratamiento, eligiendo las modalidades de tratamiento más apropiadas para cualquier paciente en particular. Los métodos predictivos de la presente invención son valiosos instrumentos para predecir si es probable que un paciente responda favorablemente a una pauta de tratamiento, como intervención quirúrgica, quimioterapia con un fármaco o combinación de fármacos determinados, y/o terapia de radiación, o si es probable una supervivencia a largo plazo del paciente, tras la cirugía y/o conclusión de la quimioterapia u otras modalidades de tratamiento.

El término "a largo plazo" se utiliza aquí para referirse a una supervivencia de al menos 3 años, más preferiblemente de al menos 8 años, aún más preferiblemente de al menos 10 años tras la cirugía u otro tratamiento.

El término "tumor", como se utiliza aquí, se refiere a todo crecimiento y proliferación celular neoplásica, maligna o benigna, y todas las células y tejidos precancerosos y cancerosos.

Los términos "cáncer" y "canceroso" se refieren o describen el estado fisiológico en mamíferos que se caracteriza típicamente por un crecimiento celular no regulado. Los ejemplos de cáncer incluyen, entre otros, cáncer de mama, cáncer de ovario, cáncer de colon, cáncer de pulmón, cáncer de próstata, cáncer hepatocelular, cáncer gástrico, cáncer de páncreas, cáncer cervical, cáncer hepático, cáncer de vejiga, cáncer del tracto urinario, cáncer de tiroides, cáncer renal, carcinoma, melanoma y cáncer cerebral.

La "patología" del cáncer incluye todos los fenómenos que afectan al bienestar del paciente. Esto incluye, entre otros, el crecimiento celular anormal o incontrolable, la metástasis, la interferencia con el normal funcionamiento de las células vecinas, la liberación de citoquinas u otros productos de secreción a niveles anormales, la supresión o agravamiento de la respuesta inflamatoria o inmunológica, neoplasia, premalignidad, malignidad, invasión de tejidos u órganos vecinos o distantes, como nódulos linfáticos, etc.

El "rigor" de las reacciones de hibridación puede ser determinado fácilmente por un experto en la técnica, y en general es un cálculo empírico dependiente de la longitud de la sonda, la temperatura de lavado y la concentración de sal. En general, sondas más largas requieren temperaturas superiores para una adecuada apareación, mientras que las sondas más cortas requieren temperaturas inferiores. La hibridación depende generalmente de la capacidad del ADN desnaturalizado para la apareación cuando están presentes cadenas complementarias en un entorno por debajo de su temperatura de fusión. Cuanto mayor sea el grado de homología deseado entre la sonda y la secuencia hibridable, tanto más alta es la temperatura relativa que puede utilizarse. La consecuencia es que temperaturas relativas superiores tenderán a hacer más estrictas las condiciones de reacción, mientras que temperaturas más bajas lo harán menos. Para detalles y explicaciones adicionales del rigor de las reacciones de hibridación, ver Ausubel et al., *Current Protocols in Molecular Biology*, Wiley

Las "condiciones rigurosas" o "condiciones muy rigurosas", como se definen aquí, típicamente: (1) emplean baja potencia iónica y alta temperatura para lavar, por ejemplo cloruro de sodio 0,015 M/ citrato sódico 0,0015 /dodecil sulfato sódico 0,1% a 50°C; (2) emplean durante la hibridación un agente desnaturalizante, como formamida, por ejemplo, 50% (v/v) formamida con albúmina de suero bovino 0,1% /Ficoll 01% /polivinilpirrolidona 0,1% /tampón fosfato sódico 50mM a pH 6.5 con cloruro sódico 750 mM, citrato sódico 75 mM a 42°C; o (3) emplean formamida 50%, 5 x SSC (NaCl 0,75 M, citrato sódico 0,075 M), fosfato sódico 50 mM (pH 6.8), pirofosfato sódico 0,1%, 5 x solución de Denhardt, ADN de esperma de salmón sonificado (50 µg/ml), SDS 0,1%, y sulfato de dextrano 10% a 42°C, con lavados a 42°C en 0,2 x SSC (cloruro sódico / citrato sódico) y formamida 50% a 55°C, seguido de un lavado muy riguroso consistente en 0,1 x SSC conteniendo EDTA a 55°C.

Las "condiciones moderadamente rigurosas" pueden ser identificadas como las describen Sambrook et al., *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Press, 1989, e incluyen el uso de una solución de lavado y condiciones de hibridación (por ej. temperatura, potencia iónica y % SDS) menos rigurosas que las descritas más arriba. Un ejemplo de condiciones moderadamente rigurosas es la incubación durante la noche a 37°C, en una solución comprendiendo: formamida 20%, 5 x SSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7.6), 5 x solución de Denhardt, sulfato de dextrano 10%, y ADN de esperma de salmón cortado desnaturalizado 20 mg/ml, seguido del lavado de filtros en 1 x SSC a aproximadamente 37-50°C. El experto sabrá cómo ajustar la temperatura, la potencia iónica, etc., lo necesario para ajustarlo a factores tales como la longitud de la sonda y similares.

En el contexto de la presente invención, la referencia a "al menos uno", "al menos dos", "al menos cinco", etc., de los genes listados en cualquier serie de genes particular significa cualquiera o cualquiera y todas las combinaciones de los genes listados.

El término cáncer "nódulo negativo", como en cáncer de mama "nódulo negativo", se utiliza aquí en referencia a un cáncer que no se ha diseminado a los nódulos linfáticos.

Los términos "empalme" y "empalme de ARN" son utilizados de forma intercambiable, y se refieren al procesado de ARN que elimina intrones y une exones, para producir ARNm maduro con secuencia de codificación continua que pasa al 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 transcrito pero eliminado de dentro de la transcripción empalmado juntos los exones de ambos lados. Operativamente, las secuencias de exón aparecen en la secuencia de ARNm de un gen como se define en los números de la Ref. SEC ID. Operativamente, 30 secuencias de intrón son las secuencias que intervienen en el ADN genómico de un gen, delimitadas por secuencias de exón, y con secuencias de consenso de empalme GT y AG en sus límites 5' y 3'.

B. Descripción detallada

En la práctica de la presente invención se emplearán, salvo que se indique lo contrario, técnicas convencionales de biología molecular (incluyendo técnicas recombinantes), microbiología, biología celular y bioquímica conocidas por los expertos en la técnica. Tales técnicas se explican en detalle en la literatura, como en "Molecular Cloning: A Laboratory Manual", 2nd edition (Sambrook et al., 1989); "Oligonucleotide Synthesis" (M.J. Gait, ed., 1984); "Animal Cell Culture" (R.I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.); "Handbook of Experimental Immunology", 4th edition (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 et al., eds., 1987); y "PCR: The Polymerase Chain Reaction", (Mullis et al., eds., 1994).

1. Perfil de expresión génica

Los métodos de perfilado de expresión génica incluyen métodos basados en el análisis de hibridación de polinucleótidos, métodos basados en la secuenciación de polinucleótidos y métodos basados en proteómica. Los métodos utilizados más comúnmente conocidos en la técnica de cuantificación de la expresión de ARNm en una muestra incluyen la hibridación *northern blotting* e in situ (Parker & Barnes, Methods in Molecular Biology 106:247-283 (1999)); ensayos de protección de ARNasa (Hod, Biotechniques so 13:852-854 (1992)); y métodos basados en PCR, como la reacción de cadena de polimerasa de transcripción inversa (RT-PCR) (Weis et al., Trends in Genetics 8:263-264 (1992)). Alternativamente, se pueden emplear anticuerpos que puedan reconocer dúplex específicos, incluyendo dúplex de ADN, dúplex de ARN, y dúplex híbridos de ADN-ARN o dúplex de proteínas de ADN. Los métodos representativos de análisis de expresión génica basados en secuenciación incluyen el Análisis en Serie de Expresión Génica (SAGE), y el análisis de expresión génica por secuenciación de firma paralela masiva (MPSS)

2. Métodos de perfilado de expresión génica basados en PCR

a PCR de transcriptasa inversa (RT-PCR)

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

El primer paso es aislar el ARNm de una muestra diana. El material de inicio es típicamente ARN total aislado de tumores humanos o líneas celulares tumorales, y tejidos o líneas celulares normales correspondientes, respectivamente. Por tanto, se puede aislar ARN de diversos tumores primarios, incluyendo tumores de mama, pulmón, colon, próstata, cerebro, hígado, riñón, páncreas, bazo, timo, testículos, ovario, útero, etc., o líneas de células tumorales, con ADN agrupado de donantes sanos. Si la fuente del ARNm es un tumor primario, se puede extraer ARNm, por ejemplo, de muestras de tejido congelado o guardadas y embebidas en parafina y fijadas (por ej. fijadas en formalina).

Los métodos generales de extracción de ARNm son bien conocidos por los expertos en la técnica, y se divulgan en libros de texto estándar de biología molecular, incluyendo Ausubel et al., Current Protocols of Molecular Biology, John Wiley and Sons (1997). Los métodos para la extracción de ARN de tejidos embebidos en parafina se divulgan, por ejemplo, en Rupp and Locker, Lab Invest. 56:A67 (1987), y De Andres et al., BioTechniques 18:42044 (1995). En particular, se puede aislar ARN utilizando un kit de purificación, conjunto de tampón y proteasa de fabricantes comerciales, como Qiagen, conforme con las instrucciones del fabricante. Por ejemplo, se puede aislar ARN total de células en cultivo utilizando minicolumnas Qiagen RNeasy. Otros kits de aislamiento de ARN disponibles comercialmente incluyen MasterPure™ Complete DNA y RNA Purification Kit (EPICENTRE®, Madison, WI), y Paraffin Block RNA Isolation Kit (Ambion, Inc.). Se puede aislar ARN total de muestras de tejido utilizando RNA Stat-60 (Tel-Test). Se puede aislar ARN de un tumor preparado, por ejemplo, por centrifugación de gradiente de densidad de cloruro de cesio.

Como el ARN no puede servir como plantilla para la PCR, el primer paso en el perfilado de expresión génica por RT-PCR es la transcripción inversa de la plantilla de ARN en ADNc, seguido de su amplificación exponencial en una reacción PCR. Las dos transcriptasas inversas más utilizadas comúnmente son la transcriptasa inversa de virus de mieloblastosis aviar (AMV-RT) y la transcriptasa inversa del virus de leucemia murina de Moloney (MMLV-RT). El paso de la transcripción inversa se ceba típicamente utilizando cebadores específicos, hexámeros aleatorios o cebadores oligo.dT, dependiendo de las circunstancias y el objetivo del perfilado de expresión. Por ejemplo, el ARN extraído pueden ser transcrito

inversamente utilizando un kit GeneAmp RNA PCR (Perkin Elmer, CA, USA), siguiendo las instrucciones del fabricante. El ADNc derivado puede ser utilizado entonces como plantilla en la posterior reacción de PCR.

Aunque en el paso de la PCR se puede utilizar diversas ADN polimerasas dependientes de ADN termoestable, se emplea típicamente la ADN polimerasa Taq, que tiene una actividad de nucleasa 5'-3', pero carece de actividad de endonucleasa correctora 3'-5'. Así, TaqMan® PCR utiliza típicamente la actividad de nucleasa 5' de polimerasa Taq o Tth para hidrolizar una sonda de hibridación ligada a su amplicón diana, pero se puede utilizar cualquier enzima con actividad de nucleasa 5' equivalente. Se utilizan dos cebadores de oligonucleótidos para generar un amplicón típico de una reacción PCR. Un tercer oligonucleótido, o sonda, está diseñado para detectar la secuencia de nucleótidos ubicada entre los dos cebadores de PCR. La sonda es no extensible por el enzima de la polimerasa ADN Taq, y está marcada con un colorante fluorescente reporter y un colorante fluorescente inhibidor. Cualquier emisión inducida por láser del colorante reporter es inhibida por el colorante inhibidor, cuando los dos colorantes están ubicados juntos como lo están en la sonda. Durante la reacción de amplificación, el enzima de la polimerasa de ADN Taq divide la sonda de una forma que depende de la plantilla. Los fragmentos de sonda resultantes se disocian en solución, y la señal del colorante reporter liberado queda libre del efecto de inhibición del segundo fluoróforo. Por cada nueva molécula sintetizada se libera una molécula de colorante reporter, y la detección del colorante reporter no inhibido proporciona la base para la interpretación cuantitativa de los datos.

Se puede realizar la ET-PCR TaqMan® utilizando los equipos disponibles comercialmente, como, por ejemplo, ABI PRISM 7700™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, CA, USA), o Lightcycler (Roche Molecular Biochemicals, Mannheim, Alemania). En una realización preferente, el procedimiento de la nucleasa 5' se realiza en un dispositivo de PCR cuantitativa de tiempo real, como ABI PRISM 7700™ Sequence Detection System™. El sistema está compuesto por un termociclador, láser, dispositivo de carga acoplada (CCD), cámara y ordenador. El sistema amplifica las muestras en un formato de 96 pocillos en un termociclador. Durante la amplificación, la señal fluorescente inducida por láser es recogida en tiempo real a través de cables de fibra óptica para los 96 pocillos, y detectada en el CCD. El sistema incluye software para la ejecución del instrumento y para analizar los datos.

Los datos del ensayo de 5'-nucleasa se expresan inicialmente como Ct, o el ciclo umbral. Como se ha comentado más arriba, los valores de fluorescencia se registran durante cada ciclo, y representan la cantidad de producto amplificado hasta ese punto de la reacción de amplificación. El punto donde se registra primero la señal fluorescente como estadísticamente significativa es el ciclo umbral (C_t).

Para minimizar los errores y el efecto de variación entre muestras, la RT-PCR se realiza habitualmente utilizando un estándar interno. El estándar interno ideal se expresa a un nivel constante entre distintos tejidos, y no es afectado por el tratamiento experimental. Los ARNs utilizados más frecuentemente para normalizar patrones de expresión génica son ARNm para los genes constitutivos gliceraldehido 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 la PCR mediante una sonda fluorogénica con marcaje dual (es decir, sonda TaqMan®). La PCR de tiempo real es compatible tanto con la PCR cuantitativa competitiva, donde se emplea para la normalización un competidor interno para cada secuencia diana, como con la PCR cuantitativa comparativa, utilizando un gen de normalización contenido dentro de la muestra, o un gen constitutivo RT-PCR. Para más detalles ver, por ejemplo, Held et al., Genome Research 6:986-994 (1996).

Los pasos de un protocolo representativo para el perfilado de la expresión génica, utilizando tejidos embebidos en parafina fijados como fuente de ARN, incluyen el aislamiento de ARNm, purificación, extensión y amplificación de cebador, se aportan en diversos artículos publicados en revistas (por ejemplo: T.E. Godfrey et al. J. Molec. Diagnostics 2: 84-91 [2000]; K. Specht et al., Am. J. Pathol. 158:419-29 [2001]). Resumiendo, un proceso representativo se inicia cortando secciones de un grosor de 10 μm aproximadamente de muestras de tejido tumoral embebidas en parafina. Se extrae entonces el ARN, y se elimina la proteína y el ADN. Tras el análisis de la concentración de ARN, se pueden incluir pasos de reparación y/o amplificación de ARN, si es necesario, y se transcribe inversamente el ARN utilizando promotores específicos de gen seguido de RT-PCR.

b. Sistema MassARRAY

En el método de perfilado de expresión génica basado en MasARRAY, desarrollado por Sequenom, Inc. (San Diego, CA), tras el aislamiento del ARN y la transcripción inversa, al ADNc obtenido se añade una molécula de ADN sintético (competidor), que coincide con la región diana de ADNc en todas las posiciones, excepto una base simple, y sirve como estándar interno. La mezcla de ADNc/competidor se amplifica por PCR y se somete a un tratamiento de enzima fosfatasa alcalina de camarón (SAP) post PCR, que produce la desfosforilación de los nucleótidos restantes. Tras la inactivación de la fosfatasa alcalina, los productos de la PCR del competidor del ADNc se someten a la extensión del cebador, que genera señales de masa distintas para los productos de la PCR derivados del competidor y el ADNc. Tras la purificación, esos productos se disponen en un array de chips, precargado con los componentes necesarios para el análisis con espectrometría de masas de tiempo de vuelo tipo desorción/ionización por láser asistido por matriz (MALDI-TOF MS). Se

cuantifica entonces el ADNc presente en la reacción analizando los cocientes de las áreas de pico en el espectro de masa generado. Para más detalles ver, ej. Ding and Cantor, Proc. Natl. Acad. Sci. USA 100:3059-3064 (2003).

c. Otros métodos basados en PCR

5 Otras técnicas basadas en PCR incluyen, por ejemplo, el display diferencial (Liang and Pardee, Science 257:967-971 (1992)); el polimorfismo de longitud de segmentos amplificados (iAFLP) (Kawamoto et al., Genome Res. 12:1305-1312 (1999)); tecnología BeadArray™ (Illumina, San Diego, CA; Oliphant et al., Discovery of Markers for Disease (Supplement to Biotechniques), June 2002; Ferguson et al., Analytical Chemistry 72:5618 (2000)); BeadsArray para la detección de la expresión gen 35 (BADGE), utilizando el sistema Luminex™ LabMAP disponible comercialmente y múltiples microesferas codificadas por colores (Luminex Corp., Austin, TX) en un ensayo rápido de expresión génica (Yang et al., Genome Res. 11:1888-1898 (2001)); y análisis de perfilado de expresión de alta cobertura (HiCEP) (Fukumura et al., Nucl. Acids. Res. 31(16):e94 (2003)).

3. Microarrays

15 La expresión génica diferencial puede ser también identificada, o confirmada utilizando la técnica de microarray. Por tanto, el perfil de expresión de los genes asociados al cáncer de mama puede ser medido en tejido tumoral fresco o embebido en parafina, utilizando la tecnología de microarray. En este método, secuencias de polinucleótidos de interés (incluyendo ADNcs y oligonucleótidos) son laminadas o dispuestas sobre un sustrato de microchip. Las secuencias dispuestas son entonces hibridadas con sondas de ADN específicas de células o tejidos de interés. Igual que en el método RT-PCR, la fuente de ARNm es típicamente ARN total aislado de tumores humanos o líneas celulares tumorales, y tejidos o líneas celulares correspondientes. Por tanto, el ARN puede aislarse de diversos tumores primarios o líneas celulares tumorales.

20 Si la fuente del ARNm es un tumor primario, se puede extraer ARNm, por ejemplo, de muestras de tejido congeladas o embebidas en parafina archivadas y fijadas (por ej. fijadas en formalina), preparadas de forma rutinaria y conservadas en la práctica clínica diaria.

En una realización específica de la técnica de microarray, se aplican insertos amplificados de PCR de clones de ADNc a un sustrato en un array denso. De preferencia, se aplican al sustrato al menos 10.000 secuencias de nucleótidos. Los genes microrordenados, inmovilizados en el microchip en 10.000 elementos cada uno, son aptos para la hibridación en condiciones rigurosas. Se pueden generar sondas de ADNc marcadas mediante incorporación de nucleótidos fluorescentes por transcripción inversa de ARN extraído de tejidos de interés. Las sondas de ADNc marcadas aplicadas sobre el chip se hibridan con especificidad para cada punto de ADN en el array. Tras un lavado estricto para eliminar las sondas no ligadas específicamente, el chip es escaneado por microscopía de láser confocal u otro método de detección, como cámara CCD. La cuantificación o hibridación de cada elemento ordenado permite la evaluación de la abundancia de ARNm correspondiente. Con fluorescencia de doble color, las sondas de ADNc marcadas de forma separada, generadas de dos fuentes de ARN, se hibridan por pares con el array. La abundancia relativa de las transcripciones de las dos fuentes correspondiendo a cada gen especificado se determina así simultáneamente. La escala miniaturizada de la hibridación permite la evaluación cómoda y rápida del patrón de expresión de grandes números de genes. Tales métodos han demostrado tener la sensibilidad requerida para detectar transcripciones raras, expresadas en unas pocas copias por célula, y detectar de forma reproducible, al menos aproximadamente, diferencias dobles en los niveles de expresión (Schen et al., Proc. Natl. Acad. Sci. USA 93(2): 106-149 (1996)). El análisis de microarray puede realizarse con los equipos disponibles comercialmente, siguiendo los protocolos del fabricante, como usar la tecnología Affymetrix GenChip, o la tecnología de microarray de Incyte.

40 El desarrollo de métodos de microarray, para el análisis a gran escala de expresión génica, hace posible la búsqueda sistemática de marcadores moleculares de clasificación del cáncer y predicción de resultados en diversos tipos de tumores.

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

45 El análisis en serie de la expresión génica (SAGE) es un método que permite el análisis simultáneo y cuantitativo de un gran número de transcripciones génicas, sin necesidad de proporcionar una sonda de hibridación individual para cada transcripción. Primero se genera una etiqueta de secuencia corta (aproximadamente 10-14 bp) que contiene suficiente información para identificar de forma inequívoca una transcripción, siempre que la etiqueta se obtenga de una posición única dentro de cada transcripción. A continuación se enlazan juntas muchas transcripciones para formar largas moléculas en serie, que pueden ser secuenciadas, revelando la identidad de las múltiples etiquetas simultáneamente. El patrón de expresión de cualquier población de transcripciones puede ser evaluado cuantitativamente determinando la abundancia de etiquetas individuales, e identificando el gen correspondiente a cada etiqueta. Para más detalles ver, p.ej., Velculescu et al., Science 270:484-487 (1995); y Velculescu et al., Cell 88:243-51 (1997).

5. Análisis de expresión génica por secuenciación de firma paralela masiva (MPSS)

Este método, descrito por Brenner et al., 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 plantillas sobre microperlas separadas de diámetro 5 µm. Primero se construye una biblioteca de microperlas de plantillas de ADN por clonación in vitro. A esto le sigue la creación de un array planar de la plantilla conteniendo microperlas en una celda de flujo a alta densidad (típicamente superior a 3 x 10⁶ microperlas/cm²). Los extremos libres de las plantillas clonadas en cada microperla se analizan simultáneamente, utilizando un método de secuenciación de firma basado en fluorescencia que no requiere la separación de fragmentos de ADN. Este método ha demostrado proporcionar de forma simultánea y precisa, en una sola operación, cientos de miles de secuencias de firma génica de una biblioteca de ADNc de levaduras.

6. Inmunohistoquímica

Los métodos de inmunohistoquímica son también adecuados para detectar los niveles de expresión de los marcadores de pronóstico de la presente invención. Así, para detectar la expresión se utilizan anticuerpos o antiseros, de preferencia antiseros policlonales, y aún mejor anticuerpos monoclonales específicos para cada marcador. Los anticuerpos pueden ser detectados por marcaje directo de los propios anticuerpos, por ejemplo, con marcadores radioactivos, marcadores fluorescentes, marcadores haptenos, como la biotina, o una enzima como la peroxidasa de rábano o la fosfatasa alcalina. Alternativamente, se utiliza un anticuerpo primario no marcado junto con un anticuerpo secundario marcado, comprendiendo antiseros, antiseros policlonales, o un anticuerpo monoclonal específico para el anticuerpo primario. Los protocolos y kits de inmunohistoquímica son bien conocidos en la técnica y están disponibles comercialmente.

7. Proteómica

El término "proteoma" se define como la totalidad de las proteínas presentes en una muestra (por ej. tejido, organismo o cultivo celular) en un momento determinado. La proteómica incluye, entre otros, el estudio de los cambios globales de la expresión proteica en una muestra (denominada también "proteómica de expresión"). La proteómica incluye típicamente los siguientes pasos: (1) separación de las proteínas individuales en una muestra por electroforesis en gel 2-D (2-D PAGE); (2) identificación de las proteínas individuales recuperadas del gel, por ej., por espectrometría de masas o secuenciación de terminal-N, y (3) análisis de los datos utilizando la bioinformática. Los métodos proteómicos son valiosos suplementos de otros métodos de perfilado de expresión génica, y se pueden usar 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, purificación y amplificación del ARNm

Los pasos de un protocolo representativo para el perfilado de la expresión génica, utilizando tejidos embebidos en parafina fijados, como fuente de ARN, incluyendo el aislamiento de ARNm, purificación, extensión y amplificación del cebador, se indican en diversos artículos publicados en revistas (por ejemplo: T.E. Godfrey et al., J. Molec. Diagnostics 2: 84-91 [2000]; K. Specht et al., Am. J. Pathol. 158:419-29 [2001]). Resumiendo, un proceso representativo se inicia cortando secciones de un grosor de 10 µm aproximadamente de muestras de tejido tumoral embebidas en parafina. Se extrae entonces el ARN, y se elimina la proteína y el ADN. Tras el análisis de la concentración de ARN, se pueden incluir pasos de reparación y/o amplificación de ARN, si es necesario, y se transcribe inversamente el ARN utilizando promotores específicos de gen seguido de RT-PCR. Finalmente, los datos son analizados para identificar la mejor opción(es) de tratamiento disponible para el paciente, en base al patrón de expresión génica característico identificado en la muestra tumoral examinada, dependiendo de la probabilidad prevista de recurrencia del cáncer.

9. Serie de genes de cáncer de mama, subsecuencias génicas ensayadas y aplicación clínica de los datos de expresión génica

Un aspecto importante de la presente invención es el uso de la expresión medida de determinados genes mediante tejido de cáncer de mama, para obtener información de pronóstico. Para este fin es necesario corregir (normalizar) diferencias en la cantidad de ARN ensayado y la variabilidad en la calidad de ARN usado. Por consiguiente, el ensayo típicamente mide e incorpora la expresión de determinados genes normalizadores, incluyendo genes constitutivos bien conocidos, como GAPDH y Cyp1. Alternativamente, la normalización puede basarse en la señal media o mediana (Ct) de todos los genes ensayados, o un amplio subconjunto de ellos (enfoque de normalización global). Sobre una base de gen-por-gen, la cantidad medida normalizada de ARNm de tumor de un paciente, se compara con la cantidad hallada en un conjunto de referencia de tejido de cáncer de mama. El número (N) de tejidos de cáncer de mama en esta serie de referencia debe ser suficientemente grande como para asegurar que series de referencia distintas (globalmente) se comporten esencialmente del mismo modo. Si se cumple esta condición, la identidad de los tejidos de cáncer de mama individuales presentes en una serie particular no tendrá un impacto significativo sobre las cantidades relativas de los genes ensayados. Habitualmente, la serie de referencia de tejido de cáncer de mama está compuesta como mínimo por unas 30 muestras FPE de tejido de cáncer de mama distintas, y preferiblemente al menos de unas 40. Salvo que se indique lo contrario, los niveles de expresión normalizada de cada ARNm/tumor estudiado/paciente, se expresarán como porcentaje del nivel de expresión medido en la serie de referencia. Más específicamente, la serie de referencia de un número de tumores suficientemente elevado (por ej., 40) proporciona una distribución de niveles normalizados de cada especie de ARNm. El nivel medido en una muestra particular de tumor a analizar cae en algún percentil dentro de este rango, que puede ser

determinado por métodos bien conocidos en la técnica. Más abajo, salvo que se indique lo contrario, la referencia a los niveles de expresión de un gen supone expresión normalizada en relación con la serie de referencia, aunque esto no se indique siempre explícitamente.

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

5 Según un aspecto de la presente invención, los cebadores y las sondas de PCR se diseñan basados en secuencias de intrones presentes en el gen a amplificar. En consecuencia, el primer paso en el diseño de cebadores/sondas es la delimitación de las secuencias de intrones en los genes. Esto puede hacerse mediante software de dominio público, como el software DNA BLAT desarrollado por Kent, W.J., *Genome Res.* 12(4):656-64 (2002), o el software BLAST, incluyendo sus variaciones. Los posteriores pasos siguen unos métodos bien establecidos de diseño de cebadores y sondas de PCR.

10 Para evitar señales no específicas, es importante enmascarar secuencias repetitivas en los intrones, al diseñar los cebadores y las sondas. Esto puede conseguirse fácilmente utilizando el programa Repeat Masker, obtenible online a través del Baylor College of Medicine, que criba secuencias de ADN contra una biblioteca de elementos repetitivos, y devuelve una secuencia de consulta en la que se enmascaran los elementos repetitivos. Las secuencias enmascaradas de intrones pueden ser usadas entonces para diseñar secuencias de cebador y sonda, utilizando cualquier paquete de
15 diseño de cebadores/sondas disponibles comercialmente o de otro modo, como Primer Express (Applied Biosystems); MGB assay-by-design (Applied Biosystems); Primer3 (Steve Rozen and Helen J. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. En: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386)

20 Los factores más importantes a considerar en el diseño de cebadores de PCR incluyen la longitud del cebador, la temperatura de fusión (T_m), y el contenido G/C, la especificidad, las secuencias de cebador complementarias y la secuencia de extremo 3'. En general, los cebadores de PCR óptimos son de 17-30 bases de longitud, y contienen aproximadamente un 20-80% de bases G+C, como por ejemplo, aproximadamente un 50-60%. Típicamente son preferibles temperaturas entre 50 y 80°C, p.ej., aproximadamente de 50 a 70°C.

25 Para más directrices para el diseño de cebadores y sondas de PCR ver, p.ej. Dieffenbach, C.W. et al., "General Concepts 50 for PCR Primer Design" in: *PCR Primer, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 133-155; Innis and Gelfand, "Optimization of PCRs" in: *PCR Protocols, A Guide to Methods and Applications*, CRC Press, London, 1994, pp. 5-11; y Plasterer, T.N. Primer select: Primer and probe design. *Methods Mol. Biol.* 70:520-527 (1997).

Se dan más detalles de la invención en el siguiente Ejemplo no limitativo.

30 **Ejemplo**

Un estudio de Fase II de expresión génica en 242 tumores de mama malignos

Se diseñó y llevó a cabo un estudio de expresión génica, con el objetivo primario de caracterizar molecularmente la expresión génica en muestras de tejido embebidas en parafina y fijadas, de carcinoma de mama ductal invasivo, y explorar la correlación entre tales perfiles moleculares y la supervivencia libre de enfermedad.

35 Diseño del estudio

Se realizaron ensayos moleculares en tejidos de tumor de mama primario, embebidos en parafina y fijados con formalina, obtenidos de 252 pacientes individuales diagnosticadas con cáncer de mama invasivo. Todas las pacientes que presentaban nódulos linfáticos negativos, eran ER positivas y estaban tratadas con Tamoxifen. La media de edad era de 52 años, y el tamaño medio del tumor clínico era de 2 cm. La media de seguimiento era de 10,9 años. A 1 de enero de 40 2003, 41 pacientes presentaban recurrencia local o distante de la enfermedad, o fallecimiento por cáncer de mama. Las pacientes eran incluidas en el estudio solo si la evaluación histopatológica, realizada como se describe en la sección Materiales y Métodos, indicaba cantidades adecuadas de tejido tumoral y patología homogénea.

Materiales y métodos

45 Cada bloque de tumor representativo se caracterizaba por histopatología estándar para el diagnóstico, evaluación semicuantitativa de la cantidad de tumor y grado del tumor. Si el área tumoral era inferior al 70% de la sección, el área tumoral se diseccionaba aproximadamente y se tomaba tejido de 6 secciones (10 micrones). De lo contrario se preparaba un total de 3 secciones (también de un grosor de 10 micrones cada una). Las secciones se colocaron en dos Tubos Costar Brand Microcentrifuge (polipropileno, tubos de 1,7 ml, transparentes). Si se obtenía más de un bloque tumoral como parte del procedimiento quirúrgico, se utilizaba para el análisis el bloque más representativo de la patología.

50

Análisis de la expresión génica

Se extrajo y se purificó ARNm de muestras de tejido embebidas en parafina y fijadas, y se prepararon para el análisis de expresión génica como se describe en el anterior capítulo 6.

5 Se realizaron ensayos moleculares de expresión génica cuantitativa por RT-PCR, utilizando el ABI PRISM 7900™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, CA, USA). El ABI PRISM 7900™ consiste en un termociclador, láser, dispositivo de carga acoplada (CCD), cámara y ordenador. El sistema amplifica muestras en un formato de 384 pocillos en un termociclador. Durante la amplificación, la señal fluorescente inducida por láser es recogida en tiempo real a través de cables de fibra óptica para los 384 pocillos, y detectada en el CCD. El sistema incluye software para la ejecución del instrumento y para analizar los datos.

Análisis y resultados

10 Se analizó tejido tumoral de 187 genes relacionados con cáncer y 5 genes de referencia. Se obtuvieron perfiles adecuados de RT-PCR de 242 de las 252 pacientes. Los valores de ciclo umbral (CT) de cada paciente se normalizaron en base a la media de los 7 genes de referencia para esa paciente particular. Se disponía de los datos del resultado clínico de todas las pacientes, de una revisión de los datos de registro y gráficos de las pacientes seleccionadas. Los resultados fueron clasificados como:

15 Evento: Viva con recurrencia local, regional o distante de cáncer de mama, o fallecimiento por cáncer de mama.

Sin evento: Viva sin recurrencia local, regional o distante de cáncer de mama, o viva con recurrencia de cáncer de mama contralateral, o viva con segundo cáncer primario no de mama, o fallecida antes de la recurrencia del cáncer de mama.

El análisis se realizó por:

20 A. Determinación de la relación entre la expresión génica normalizada y los resultados binarios de 0 o 1;

B. Análisis de la relación entre la expresión génica normalizada y el tiempo hasta el resultado (0 o 1 como se define más arriba), donde se descartaban las pacientes vivas sin recurrencia de cáncer de mama, o que habían fallecido debido a una causa distinta del cáncer de mama. Se utilizó este enfoque para evaluar el impacto en el pronóstico de genes individuales y también de series de genes múltiples.

25 Análisis de las pacientes con carcinoma de mama invasivo por enfoque binario

En el primer enfoque (binario), se realizó el análisis en las 242 pacientes con carcinoma de mama invasivo. Se realizó un test t en los grupos de pacientes clasificadas como sin recurrencia y fallecimiento no relacionado con cáncer de mama a los 10 años, versus recurrencia o muerte relacionada con cáncer de mama a los 10 años, y se calcularon los valores p para las diferencias entre grupos para cada gen.

30 En la Tabla 1 aparecen los 33 genes para los que el valor p para las diferencias entre los grupos era <0,05. La primera columna de valores de expresión medios corresponde a pacientes que presentaron recurrencia metastásica o fallecieron por cáncer de mama. La segunda columna de valores de expresión medios corresponde a pacientes que no tenían recurrencia metastásica o no fallecieron por cáncer de mama.

Tabla 1

Gen	Media grupo A		Media grupo B	
	Evento	Sin evento	Estadística T	Valor 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_orf1	-2,43	-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_Catenina	-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 t negativos indican una mayor expresión, asociada a mejores resultados, y a la inversa, valores superiores (positivos) de t indican una mayor expresión asociada a peores resultados. Así, por ejemplo, una expresión elevada del gen CCNB1 (valor t = 3,02, CT media viva <CT media fallecida) indica una menor probabilidad de supervivencia libre de enfermedad. De forma similar, una expresión elevada del gen GSTM1 (valor t = -3,56, CT media viva >CT media fallecida) indica una mayor probabilidad de supervivencia libre de enfermedad.

10 Así, en base a la serie de datos indicada en la Tabla 1, la expresión de cualquiera de los siguientes genes en el cáncer de mama indica una menor probabilidad de supervivencia sin recurrencia del cáncer: C20_orf1; CCNB1; CDC20; CDH1; CTSL2; EpCAM; GRB7; HER2; KNSL2; LMNB1; MCM2; MMP9; MYBL2; NEK2; PCNA; PREP; PTTG1; STMY3; SURV; TS; MELK.

En base a la serie de datos indicada en la Tabla 1, la expresión de cualquiera de los siguientes genes en el cáncer de mama indica un mejor pronóstico de supervivencia sin recurrencia del cáncer: BAG1; Bcatenin; CEGP1; CIAP1; cMYC; DKFZp586M07; EstR1; GSTM1; GSTM3; ID1; ITGA7; PR.

Análisis de genes múltiples e indicadores de resultados

15 Se adoptaron dos enfoques para determinar si el uso de genes múltiples proporcionaría una mejor discriminación entre resultados. Primero se realizó un análisis de discriminación utilizando un enfoque de avance progresivo. Se generaron modelos que clasificaron el resultado con mayor discriminación de la obtenida con cualquier gen individual solo. Según un segundo enfoque (enfoque de tiempo-hasta-el evento), se definió para cada gen un modelo Cox Proportional Hazards (ver, p.ej., Cox, D. R., and Oakes, D. (1984), Analysis of Survival Data, Chapman and Hall, London, New York), con el tiempo hasta la recurrencia o la muerte como variable dependiente, y el nivel de expresión del gen como variable independiente.

20 Se identificaron los genes con un valor p < 0,05 en el modelo Cox. Para cada gen, el modelo Cox proporciona el riesgo

5 relativo (RR) de recurrencia o muerte para el cambio de una unidad en la expresión del gen. Se puede optar por dividir las pacientes en subgrupos por cualquier valor umbral de la expresión medida (en la escala CT), donde todas las pacientes con valores de expresión superiores al umbral tienen un riesgo superior, y todas las pacientes con valores de expresión inferiores al umbral tienen menor riesgo, o al revés, dependiendo de si el gen es un indicador de mal (RR > 1,01) o buen (RR < 1,01) pronóstico. Así, cualquier valor umbral definirá los subgrupos de pacientes con riesgo aumentado o disminuido respectivamente. Los resultados se resumen en la Tabla 2, que muestra los 42 genes para los que el valor p para las diferencias entre los grupos era < 0,05.

Tabla 2

Gen	Riesgo Relativo	valor 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
PTTG1	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

(continuación)

Gen	Riesgo Relativo	valor p
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 En base a la serie de datos indicada en la Tabla 2, la expresión de cualquiera de los siguientes genes en el cáncer de mama indica una menor probabilidad de supervivencia sin recurrencia del cáncer: GRB7; SURV; LMNB1; MYBL2; HER2; MELK; C20_orf1; PTTG1; BUB1; CDC20; CCNB1; STMY3; KNSL2; CTSL2; MCM2; NEK2; Ki_67; CCNE2; TOP2A-4; PCNA; PRÉP; FOXM1; NME1; STK15; HNRPAB; MMP9; TS; Src; MMP12; TFRC.

En base a la serie de datos indicada en la Tabla 2, la expresión de cualquiera de los siguientes genes en el cáncer de mama indica un mejor pronóstico de supervivencia sin recurrencia del cáncer: PR; GSTM1; DR5; CEGP1; BAG1; EstR1; DKFZp586M07; BIN1; NPD009; RPLPO; GSTM3; IGF1R.

10 El análisis binario y de tiempo-hasta-el evento, con contadas 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 en los 15 genes superiores en ambas listas. Además, cuando ambos análisis identificaron el mismo gen a $[p < 0,10]$, lo que ocurrió para 26 genes, fueron siempre concordantes con respecto a la dirección (signo positivo o negativo) de la correlación con la supervivencia/recurrencia. Globalmente, estos resultados confirman la conclusión de que los marcadores
15 identificados tienen un valor de pronóstico significativo.

Análisis génico multivariado de 242 pacientes con carcinoma de mama invasivo

20 En los modelos Cox comprendiendo más de dos genes (modelos multivariados), se procede a la entrada gradual de cada gen individual en el modelo, donde el primer gen entrado se preselecciona entre los genes con valores p univariados significativos, y el gen seleccionado para la entrada en el modelo en cada paso posterior es el gen que mejora más el encaje del modelo con los datos. Este análisis puede ser realizado con cualquier número total de genes. En el análisis cuyos resultados se muestran más abajo, se llevó a cabo la entrada progresiva de hasta 10 genes.

Se realizó el análisis multivariado utilizando la siguiente ecuación:

Se realizó el análisis multivariado utilizando la siguiente ecuación:

$$R = \exp[\text{coef}(\text{genA}) \times \text{Ct}(\text{genA}) + \text{coef}(\text{genB}) \times \text{Ct}(\text{genB}) + \text{coef}(\text{genC}) \times \text{Ct}(\text{genC}) + \dots]$$

25 En esta ecuación, los coeficientes para genes que son predictores de resultado beneficioso son números positivos, y los coeficientes para genes que son predictores de resultado desfavorable son números negativos. Los valores "Ct" de la ecuación son Δ Cts, es decir, reflejan la diferencia entre el valor Ct normalizado de una población, y el valor Ct normalizado medido para la paciente en cuestión. La convención utilizada en el presente análisis ha sido que Δ Cts por debajo y por encima de la media de la población tienen signos positivos y signos negativos, respectivamente (reflejando una mayor o
30 menor abundancia de ARNm). El riesgo relativo (RR) calculado resolviendo esta ecuación indicará si la paciente tiene una probabilidad aumentada o reducida de supervivencia a largo plazo sin recurrencia del cáncer.

5 Se llevó a cabo un análisis gradual multivariado, utilizando el Cox Proportional Hazards Model, de los datos de expresión génica obtenidos de las 242 pacientes con carcinoma de mama invasivo. En este análisis se ha determinado que la siguiente serie de diez genes tiene un valor predictivo particularmente sólido de la supervivencia de las pacientes: GRB7; LMNB1; ER; STMY3; KLK10; PR; KRT5; FGFR1; MCM6; SNRPF. En esta serie 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.

Aunque la presente invención ha sido descrita con referencia a lo que se considera que son realizaciones específicas, hay que entender que la invención no se limita a tales realizaciones. Por el contrario, se pretende que la invención cubra diversas modificaciones incluidas en el alcance de las reivindicaciones adjuntas.

Gen	Acceso	Inicio	Parada	SEQ ID NO:	Secuencia
B-Catenin	NM_001904	1549	1629	SEQ ID NO: 1	GGCTCTTGTGCTACTGTCTTTCGGGCTGTGACAGCAACAGACATCACTGGAGCCTGGCATCTGTG CTGTCTGTGATCTGA
BAG1	NM_004323	673	754	SEQ ID NO: 2	CGTTGTGACGACTTCGAAACAGAGATGATGTTGGCCGGTTCATGTTAATTGSGAAGAAAGACAGTCCAG AGGAAAGAGGTTGAAG
BIN1	NM_004305	866	942	SEQ ID NO: 3	CGTGC AAAAGGAAACAAAGAGCCCTTCGGCTCCAGATAGCTGCGCTCCGCGCCGACGCGGAGTGAAGGCGGCAATG GAGTCAAGGDAAG
BUB1	NM_004336	1002	1070	SEQ ID NO: 4	CGDAGGTTAATCCAGCACGATATCGGCGGCAAGTGTAGGCTGCCACGACGAGGAGCTGAAGGCGGCAATG TCTT
C20orf1	NM_012112	2675	2740	SEQ ID NO: 5	TCAGCTGTGAGCTGGGATACCGCCCGCAATGGGACCTGCTCTTAACCTCAAACCTAGGACCGT
CCNB1	NM_031966	823	907	SEQ ID NO: 6	TTCAGGTTGTGGAGAGACAGTGTACATGACTGTCTCCATATTGATCGGTTGATGCGAGATAATT GTGTGCGCAAGAGATG
CCNE2	NM_057749	2026	2108	SEQ ID NO: 7	ATGTTGTGCGCTTCTTCCATCTAGCTGGGCGCTTCTTGACATGTAGGTTGCTTGGTATATAGCCTTTTGTGA TATGACAAATTTGGT
CDC20	NM_001255	679	747	SEQ ID NO: 8	TGGATTTGGAGTCTTCTGGGAATGACTGTGGCGGTGGCACTGACAGAGTGTGTACGTGTGGAATGCA AGC
CDH1	NM_004360	2499	2580	SEQ ID NO: 9	TGAGTGTGCGCGGATCTTCCGCGCGCTGGCAATCCCGGATGAAATGGAAATTTTATTGATGAAA ATGTGAAAGCGGCTG
CEBP1	NM_020974	563	640	SEQ ID NO: 10	EBACAACTGAGCAACCTGGATTCACCGCTGAGAAAGAGGCGCTTGAGCTGCAATUAATAGGATCAAG TCGTGATGTCACA
CIAP1	NM_001166	1822	1894	SEQ ID NO: 11	TGCGTGTGTGGCGAAGCTCAATGTAAGTGGAGCGAAGGATGATGCTATGTGACAGACACCGGAGGCG ATTTTTC
cMYC	NM_002467	1494	1578	SEQ ID NO: 12	TGCTTCCACTCGGAGGAGTATCTGCTGTGCAAGAGGAGGTTGAGTGGAGTGGAGTGTGAGAGTGTGAGAGTGTGAG GACAGATCAGCGAAGCG
CTSL2	NM_001333	671	738	SEQ ID NO: 13	TGTTCTCACTGAGGAGGAGGATCTGTTGAGTGTGAGTGTGCGGCTCCGAGGCGAATCAAGGCTGCAATG GT
DKFZp586	AL050227	559	633	SEQ ID NO: 14	TGCAATTTTACCTGTTAACTTCATCATATTTTGTGAGAGGCGCTGGAAAGCAAGAGAGAGAGAGGAGC CACTGGAT
DR5	NM_003842	1127	1211	SEQ ID NO: 15	CTCTGAGACAGTCTTCGATGACATTTTGGAGAGTGTGTTGCGGCTTTGACTCTGGAGGCGGCGTCAATGA GAAATTTGGCGCTCATGGA
EpcAM	NM_002354	435	510	SEQ ID NO: 16	GGCGGCTCCAGAACATGATGCGGCTTATGCTCCTCACTGCGATGAGAGGCGGCTCTTTAAGCGC AAGGAGTTCGA
EstR1	NM_000125	1956	2024	SEQ ID NO: 17	CGTGTGCGGCTCTATGACCTGCTGTGGAGATGCTGGAGCGCGCCACCGGCTACATGTCGCGCACTA GCG
FGFR1	NM_023109	2685	2759	SEQ ID NO: 18	CAGCGGACATTCAGCCACATGACACTATATAAAAAGACAAAGCGAGCGGCGGACTGCTGTGAGGTGG ATGGACCG
FOXM1	NM_021953	1898	1980	SEQ ID NO: 19	CGACCGCGAGGAAATCTGTGCTGCGCGAGACCGGCTGAATGCTGGAGGGCTCAAGCGCGGCGAGCGCAA GTAGCGGCGAGCTGGATTT
GRB7	NM_005310	1275	1342	SEQ ID NO: 20	CGAATGCAATGCAATCTTTTGTGGCTGGCGGCTGGAGAAAGTGGCTCAGATATAACCTGCTGGTGGC C
GSTM1	NM_000581	93	179	SEQ ID NO: 21	AAGGTATGAGGAAAGATACAGATGGGCGAGGCTGCTGATATGACAGAGGCGGAGTGGCTGA ATGAAATTCAGGCTGGGCG
GSTM3	NM_000849	248	324	SEQ ID NO: 22	CAATGGCAATCTTGGCCTACATCGTCCCAAGCACACATGTTGTGGTGAAGTCTGAGAGAGAAAGAT TGGAGTGGAG

TABLA 3

(continuación)

Gen	Acceso	Inicio	Parada	SEQ ID NO.	Secuencia
HER2	NM_004448	1138	1208	SEQ ID NO: 23	CGGTGTAGAGACTGCAGACAGCCCTGTGCCCCAGTGTGCTATGGTGTGGCCATGGACGACCTTCGAGAGG
HNRPAB	NM_004499	1086	1170	SEQ ID NO: 24	CAMGGAGCCGCAACTGATGCCACACATGCTTTGTTGGATATGGAGTGAACACATATATGTACC AAATTTAATTTGGCAAC
ID1	NM_002165	286	356	SEQ ID NO: 25	ADACCCCAAGGTGACCAAGGTGGAGATTCTCCAGCACTCATCGACTACATCAGGCACTTCAG TTGGA
IGF1R	NM_000875	3467	3550	SEQ ID NO: 26	CAGCTGTAGCCGAGATTTCAGAGTCAAAATCGGAGATTTTGGTATAGCGGAGATATATGACA CAGCTATTACGGGAAA
ITGA7	NM_002206	633	712	SEQ ID NO: 27	GATATGATGTGCTGCTGCTTTGTGCTCAGCCAGGACCTGCGCCATCCGGGATGATGATGATGCTGG GSAATGSAAGTTCT
KI-67	NM_002417	42	122	SEQ ID NO: 28	CGGAGTTTGGTGGACTTGGAGAGCGGTGTTGCAQAAATGGCCCTTCCGGGGGCAATGCTGCC AGTGGAAAGTTGTAA
KLK10	NM_002776	966	1044	SEQ ID NO: 29	GCCTGAGGCTCATGCTGCTCCTCTCTCCCAAGTCCGCTGMAACTCTCCCTGTGCTGCACT GTTCAAACTCTG
KNSL2	BC000712	1266	1343	SEQ ID NO: 30	CCAGCTGCCCATGATTTTCTTTGACGCGGTATTCCGACCAAGAAATGGAGGATGAAGTTT GAAGGATTGC
KRT5	NM_000424	1605	1674	SEQ ID NO: 31	TCAATGAGAAAGGATGGAGCAGTCAACATCTCTGTTGTCACAGCAAGTGTCTCTGCAATAG GCA
LMBN1	NM_005573	1500	1566	SEQ ID NO: 32	TGGAAAACCTGTGTTACACAGCCAGGCGGCACTGACCTCATCTGGAAAGAACAGACTGCTGGG G
MCM2	NM_004526	2442	2517	SEQ ID NO: 33	GAGTTTGGCCGCTAGCTTTTCATTCGCGCGTGAACAAATGAGCTGTGCTCTTCACTGAAAGCA GTTAGTGGC
MCM6	NM_005915	2669	2751	SEQ ID NO: 34	TGATGCTCTATGTCACATCATCAACAGTTTCAACCAACAGAGGCTTCAGCACTGCTTTGGT GTGTTGCTGTCGA
MELK	NM_014791	22	87	SEQ ID NO: 35	AACCCGGCGATCGAAAGATTCTAGAACCGCGTACCAGCCGCGTCTCTCAGGACAGCAGCCCC
MMIP12	NM_002426	816	894	SEQ ID NO: 36	CDAGCCTTGCMAATCTGTGACATTCAGAACCCAGCTCTCTGTGACCCCAATTTGATGTTTATGCT GTGATACGCT
MMP9	NM_004994	124	191	SEQ ID NO: 37	CAGAACCAATCTCACCGAGAGGCACTGGCCAGAGGAATACCTGTACCCGCTATGCTTACACTGGG TG
MYBL2	NM_002466	599	673	SEQ ID NO: 38	GCAGACTGCCCAAGATGTTGCCAGGAGGACAGCAATGCTGTGAGCAATCACTCGAACTGTAC CATCAAAG
NEK2	NM_002497	102	181	SEQ ID NO: 39	GTGAGGAGCCGCACTCTGGCCAGCTGGCCGCCATGCCCTTCCCGGGGCTGAGGACTATGAATGT TGTADGCAATGGCA
NME1	NM_000269	365	439	SEQ ID NO: 40	CCAACCTTCCAGACTGGAAAGGCTGGAGCCATCCGCTGGAGACTTCTGATACAAAGTTGGCAGGAG ATTATACAT
NPD009	NM_020686	589	662	SEQ ID NO: 41	GCTGTGGCTGAGGCTGTAGCATCTCTGCTGGAGGAGCACTCTGGGAACTGATTTGACCTGG AATGCTCC
PCNA	NM_002592	157	228	SEQ ID NO: 42	GAAAGTGTGGAGGCACTCAAGGAGCTCATCAACAGGCGCTGCTGGGATATTAGCTCAAGGGGTG TAAACC
PR	NM_000926	1895	1980	SEQ ID NO: 43	GCAACAGGCTGTGATTAATGCTGTCTTACTGTGGAGCTGTAGGCTCTCTTTTAAAGGGGCAATG GAAAGGCAAGCACTACT

(continuación)

Gen	Acceso	Inicio	Parada	SEQ ID NO.	Secuencia
PREP	NM_002726	859	965	SEQ ID NO. 44	GGACGGTGTACATTGAAGACGAAATGGCGAGTCTCCCAAGTATCGGGTGAATGACATTGACTTC TGGGATCTTG
PTTG1	NM_004219	48	122	SEQ ID NO. 45	GGCTAGTCTGATCTATGTTGATTAAGGMAAATGAGAACGAGCGACGGCTGTGCTTCTTAAGGATG GGCTGAAGCC
RPLPO	NM_C01002	791	866	SEQ ID NO. 46	CAATCTGATCATCAAGCGGGTACAAACGAGTCCCTGGCCCTTGTGTGTGAGACGGATTACACCTTCCG ACTTGGTGA
SNRPF	NM_003095	71	150	SEQ ID NO. 47	GGCTGGTGGGCGAGAGATAGCCCTGCAACATTGCGGCCGTGAGTTTACATGAGTTTACGCCCTCAATCC CAAAAGCTTTTCTCA
Src	NM_004383	979	1043	SEQ ID NO. 48	CCTGAACATGAAGGAGCTGAAGCTGCTGCAGACCCATCGGGGAGTTCCGGAGACGTGATG
STK15	NM_003600	1101	1170	SEQ ID NO. 49	CATCTTCGAGAGGAGACACTCTGCTGTGGGACCTGTGGACTGCTGCCCTCTGAAATGATTGAAAGGT CGGA
STMY3	NM_005940	2090	2180	SEQ ID NO. 50	CCTGGAGGCTGCAGACTACTGATCTGCTGTGCGAGCGCGGATCTCTGTAAGGCCCTTTTCGCAAG ACTGCTATCTCGAAGAGCGATTGTA
SURV	NM_001168	737	817	SEQ ID NO. 51	TGTTTTGATCCCGGCTTACCAAGTGTGAGAGTGTAGGAGAGGAGAGAGGAGTGTCCCTTTTGCTA GGCTGTGAGGCTTG
TFRC	NM_003234	2110	2178	SEQ ID NO. 52	GGCAACTGCTTTGATTTGTGAGGGATCTGAAAGCAATACAGAGGCAAGCATAAAGGAAATGGCCCTGA GT
TOP2A	NM_001067	4505	4577	SEQ ID NO. 53	AATCCAGGGGGAGAGTGTGATGACTTGCATATGACTTTTGACTTGAAGTGTGGCTCTGGGGDAAA TCTGTAG
TS	NM_001071	764	829	SEQ ID NO. 54	GCCTGGGTGGCCTTTCAACATGGCCAGCTACGCCCTGCTCACCGTACATGATTGGGCACATCACG

Gen	Acceso	Nombre	SEQ ID NO	Secuencia
B-Catenin	NM_001904	S2150/B-Cate.f3	SEQ ID NO:55	GGCTCTTGTGGTACTGTCTT
B-Catenin	NM_001904	S2151/B-Cate.r3	SEQ ID NO:56	TCAGATTGACGAAGACACAGATG
B-Catenin	NM_001904	S5046/B-Cate.p3	SEQ ID NO:57	AGGCTCAGTGATGCTTCCCTGTCCACCAG
BAG1	NM_004323	S1386/BAG1.f2	SEQ ID NO:58	CGTTGTCCAGCACTTGAATACAA
BAG1	NM_004323	S1387/BAG1.r2	SEQ ID NO:59	GTTCAACCTTCTCTGTGGACTGT
BAG1	NM_004323	S4731/BAG1.p2	SEQ ID NO:60	CCCAATTACATGACCCGGCAACCAT
BIN1	NM_004305	S2651/BIN1.f3	SEQ ID NO:61	CCTGCAAAAGGAAACAAGAG
BIN1	NM_004305	S2652/BIN1.r3	SEQ ID NO:62	CGTGGTTGACTCTGATCTCG
BIN1	NM_004305	S4954/BIN1.p3	SEQ ID NO:63	CTTCGCCTCAGATGGCTCCC
BUB1	NM_004336	S4234/BUB1.f1	SEQ ID NO:64	CCGAGGTTAATCCAGCACGTA
BUB1	NM_004336	S4235/BUB1.r1	SEQ ID NO:65	AAGACATGGCGCTCTCAGTTC
BUB1	NM_004336	S4236/BUB1.p1	SEQ ID NO:66	TGCTGGAGCCACACTTGGCCC
C20 orf1	NM_012112	S3560/C20 or.f1	SEQ ID NO:67	TCAGCTGTGAGCTGCGGATA
C20 orf1	NM_012112	S3561/C20 or.r1	SEQ ID NO:68	ACGGTCTAGGTTTGAGGTTAAGA
C20 orf1	NM_012112	S3562/C20 or.p1	SEQ ID NO:69	CAGTCCCATTTGCCGGGCG
CCNB1	NM_031966	S1720/CCNB1.f2	SEQ ID NO:70	TTCAGGTTTGGCAGGAGAC
CCNB1	NM_031966	S1721/CCNB1.r2	SEQ ID NO:71	CATCTTCTGGGCACACAAT
CCNB1	NM_031966	S4733/CCNB1.p2	SEQ ID NO:72	TGTCTCATTATTGATCGGTTGATGCA
CCNE2	NM_057749	S1458/CCNE2.f2	SEQ ID NO:73	ATGCTGTGGCTCTTCTTAAC
CCNE2	NM_057749	S1459/CCNE2.r2	SEQ ID NO:74	ACCCAAATTTGATATACAAAAGGTT
CCNE2	NM_057749	S4945/CCNE2.p2	SEQ ID NO:75	TACCAAGCAACCTACATGTCAAAGAAAGCCC
CDC20	NM_001255	S4447/CDC20.f1	SEQ ID NO:76	TGGATTGGAGTTCTGGGAATG
CDC20	NM_001255	S4448/CDC20.r1	SEQ ID NO:77	GCTTGCACTCCACAGGTACACA
CDC20	NM_001255	S4449/CDC20.p1	SEQ ID NO:78	ACTGGCCGTGGCACTGGACAACA
CDH1	NM_004360	S0073/CDH1.f3	SEQ ID NO:79	TGAGTGTCCCCCGGTATCTTC
CDH1	NM_004360	S0075/CDH1.r3	SEQ ID NO:80	CAGCCCGTTTCAGATTTTCAT
CDH1	NM_004360	S4990/CDH1.p3	SEQ ID NO:81	TGCCAATCCCGATGAAATTTGGAAATTT
CEGP1	NM_020974	S1494/CEGP1.f2	SEQ ID NO:82	TGACAATCAGCACACCTGCAT
CEGP1	NM_020974	S1495/CEGP1.r2	SEQ ID NO:83	TGTGACTACAGCCCGTGCCTTA
CEGP1	NM_020974	S4735/CEGP1.p2	SEQ ID NO:84	CAGGCCCTCTTCCGAGCGGT
CIAP1	NM_001166	S0764/CIAP1.f2	SEQ ID NO:85	TGCCTGTGGTGGGAAGCT
CIAP1	NM_001166	S0765/CIAP1.r2	SEQ ID NO:86	GGAAATGCCTCCCGGTGT
CIAP1	NM_001166	S4802/CIAP1.p2	SEQ ID NO:87	TGACATAGCATCATCTTTGGTCCCAGTT

TABLA 4

Gen	Acceso	Nombre	SEQ ID NO	Secuencia
cMYC	NM_002467	S0085/cMYC.i3	21	TCCCTCCACTCGGAAGGACTA
cMYC	NM_002467	S0087/cMYC.r3	22	CGGTTGTTGCTGATCTGTCTCA
cMYC	NM_002467	S4994/cMYC.p3	27	CTGACACTGTCCAACCTTGACCCCTCTT
CTSL2	NM_001333	S4354/CTSL2.f1	21	TGCTCACTGAGCGGAGGAGAA
CTSL2	NM_001333	S4355/CTSL2.r1	19	ACCAATTCAGCCCTGATTG
CTSL2	NM_001333	S4356/CTSL2.p1	24	CITGAGGACGGCAACAGTCCACCA
DKFzp596M0723	AL050227	S4396/DKFzP5.f1	27	TCCATTTTCTACCTGTGTTAACCTTCATC
DKFzp596M0723	AL050227	S4397/DKFzP5.r1	19	ATGCAGTCGGTCCCTTCCT
DKFzp596M0723	AL050227	S4398/DKFzP5.p1	23	TTGCTCCAGGGCCCTGCACAAA
DR5	NM_003842	S2551/DR5.i2	24	CTCTGACACAGTCTTCGATGACT
DR5	NM_003842	S2552/DR5.r2	19	CCATGAGGCCAACTTCCT
DR5	NM_003842	S4979/DR5.p2	23	CAGACTTGGTGCCCTTTGACTCC
EpcAM	NM_002354	S1807/EpcAM.f1	20	GGGCCCTCAGAAACAATGAT
EpcAM	NM_002354	S1808/EpcAM.r1	21	TGCACTGCTGGCCCTTAAAGA
EpcAM	NM_002354	S4984/EpcAM.p1	25	CCGCTCTCATCGCAGTCAGGATCAT
Esir1	NM_C00125	S0115/Esir1.f1	19	CGTGGTGCCCTCTATGAC
Esir1	NM_000125	S0117/Esir1.r1	19	GGCTAGTGGGCGCATGTAG
Esir1	NM_000125	S4737/Esir1.p1	19	CTGGAGATGCTGGACGCC
FGFR1	NM_023109	S0818/FGFR1.f3	20	CACGGGACATTCACCCACATC
FGFR1	NM_023109	S0819/FGFR1.r3	19	GGGTGCCATCCACTTACA
FGFR1	NM_023109	S4816/FGFR1.p3	27	ATAAAAAGACAAACCAACGGCCGACTGC
FOXM1	NM_021953	S2006/FOXM1.f1	19	CCACCCCGAGCAAATCTGT
FOXM1	NM_021953	S2007/FOXM1.r1	22	AAATCCAGTCCCTTACTTTGG
FOXM1	NM_021953	S4757/FOXM1.p1	23	CCTGAATCCTGGAGGCTCACGCC
GRB7	NM_005310	S0130/GRB7.f2	20	ccatctgcatccatctggt
GRB7	NM_005310	S0132/GRB7.r2	20	ggccaccagggtattatctg
GRB7	NM_005310	S4726/GRB7.p2	23	ctccccacccttgagaagtgcc
GSTM1	NM_000561	S2026/GSTM1.r1	20	GGCCGAGCTTGAATTTTTCA
GSTM1	NM_000561	S2027/GSTM1.f1	27	AAGCTATGAGGAAAAGAGTACACCGAT
GSTM1	NM_000561	S4739/GSTM1.p1	30	TCAAGCCACTGGCTTCTGTCAATAATCAGGAG
GSTM3	NM_000849	S2038/GSTM3.f2	21	CAATGCCATCTGGCGTACAT
GSTM3	NM_000849	S2039/GSTM3.r2	25	GTCCACTCGAATCTTTTCTTTCA
GSTM3	NM_000849	S5064/GSTM3.p2	27	CTCGAAAGCACAAACATGTGTGTGAGA

(continuación)

Gen	Acceso	Nombre	SEQ ID NO	Secuencia
HER2	NM_004448	S0142/HER2.f3	SEQ ID NO:121	CGGTGTGAGAAAGTGCAGCAA
HER2	NM_004448	S0144/HER2.f3	SEQ ID NO:122	CCTCTCGCAAGTGCTCCAT
HER2	NM_004448	S4729/HER2.p3	SEQ ID NO:123	CGACACATAGCACACTCGGGCAC
HNRPAB	NM_004499	S4510/HNRPAB.f3	SEQ ID NO:124	CAAGGGAGCGACCAACTGA
HNRPAB	NM_004499	S4511/HNRPAB.f3	SEQ ID NO:125	GTTTGCCAAAGTTAAATTTGGTACATAAT
HNRPAB	NM_004499	S4512/HNRPAB.p3	SEQ ID NO:126	CTCCATATCCAAACAAAGCATGTGTGCG
ID1	NM_002165	S0820/ID1.f1	SEQ ID NO:127	AGAACC CGCAAGGTGAGCAA
ID1	NM_002165	S0821/ID1.r1	SEQ ID NO:128	TCCAAGTGAAGGTCCCTGATG
ID1	NM_002165	S4832/ID1.p1	SEQ ID NO:129	TGGAGATTCTCCAGCACGTCATCGAC
IGF1R	NM_000875	S1249/IGF1R.f3	SEQ ID NO:130	GCATGGTAGCCGAAGATTCA
IGF1R	NM_000875	S1250/IGF1R.r3	SEQ ID NO:131	TTCCGGTAATAGTCTGCTCATAGATATC
IGF1R	NM_000875	S4895/IGF1R.p3	SEQ ID NO:132	CGCGTCATACAAAATCCCGATTTTGA
ITGA7	NM_002206	S0859/ITGA7.f1	SEQ ID NO:133	GATATGATTGGTCGCTGCTTTG
ITGA7	NM_002206	S0920/ITGA7.r1	SEQ ID NO:134	AGAACTTCCATTCCCAACCAT
ITGA7	NM_002206	S4795/ITGA7.p1	SEQ ID NO:135	CAGCCAGGACCTGGCCATCCG
KI-67	NM_002417	S0436/KI-67.f2	SEQ ID NO:136	CGGACTTTGGTGGGACTT
KI-67	NM_002417	S0437/KI-67.f2	SEQ ID NO:137	TTAACTCTCCACTGGGACGAT
KI-67	NM_002417	S4741/KI-67.p2	SEQ ID NO:138	CCACTTGTGCAACCAACCCGCTCGT
KLK10	NM_002776	S2624/KLK1 0.f3	SEQ ID NO:139	GCCCAGAGGCTCCATCGT
KLK10	NM_002776	S2625/KLK10.f3	SEQ ID NO:140	CAGAGTTTGAACAGTGCAGACA
KLK10	NM_002776	S4978/KLK10.p3	SEQ ID NO:141	CCTTCTCCTCCCAAGTCGGCTGA
KNSL2	BC000712	S4432/KNSL2.f2	SEQ ID NO:142	CCACCTCGCATGATTTTC
KNSL2	BC000712	S4433/KNSL2.r2	SEQ ID NO:143	GCAATCTTCAAACACTTCATCCT
KNSL2	BC000712	S4434/KNSL2.p2	SEQ ID NO:144	TTTGACCGGGTATCCCAACCAAGAA
KRT5	NM_000424	S0175/KRT5.f3	SEQ ID NO:145	tc:agtggaagaagggtlfgga
KRT5	NM_000424	S0177/KRT5.r3	SEQ ID NO:146	tgccatattccagaggaaaca
KRT5	NM_000424	S5015/KRT5.p3	SEQ ID NO:147	ccagticaaacatctctgtgticaagca
LMBN1	NM_005573	S4477/LMBN1.f1	SEQ ID NO:148	TGCAAACGCTGGTGTGACA
LMBN1	NM_005573	S4478/LMBN1.r1	SEQ ID NO:149	CCCCACAGGTTCTGGTCTTC
LMBN1	NM_005573	S4479/LMBN1.p1	SEQ ID NO:150	CAGCCCCCAACTGACCTCATC
MCM2	NM_004526	S1602/MCM2.f2	SEQ ID NO:151	GACTTTTCCCCGCTACCTTC
MCM2	NM_004526	S1603/MCM2.r2	SEQ ID NO:152	GCCACTAAGTCTTTCAGTATGAAG
MCM2	NM_004526	S4900/MCM2.p2	SEQ ID NO:153	ACAGCTCATTGTTGTACACGCCGGA

(continuación)

Gen	Acceso	Nombre	SEQ ID NO	Secuencia	24
MCM6	NM_005915	S1704/MCM6.f3	SEQ ID NO:154	TGATGGTCCTATGTGCACATTCA	24
MCM6	NM_005915	S1705/MCM6.r3	SEQ ID NO:155	TGGGACAGAAACACACACCAA	20
MCM6	NM_005915	S4919/MCM6.p3	SEQ ID NO:156	CAGGTTTCATACCAACACAGGCTTCAGCAC	30
MELK	NM_014791	S4318/MELK.f1	SEQ ID NO:157	AACCCGGCATCGAAAAG	18
MELK	NM_014791	S4319/MELK.r1	SEQ ID NO:158	GGGCTGCTGTCTGAGA	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	ACGGTAGTGACAGCATCAAAACTC	24
MMP12	NM_002426	S4383/MMP12.p2	SEQ ID NO:162	AACCAGCTCTGTGACCCCAATT	24
MMP9	NM_004994	S0656/MMP9.f1	SEQ ID NO:163	GAGAACAATCTACCCGACA	20
MMP9	NM_004994	S0657/MMP9.r1	SEQ ID NO:164	CACCCGAGTAAACCATAGC	20
MMP9	NM_004994	S4760/MMP9.p1	SEQ ID NO:165	ACAGGTATTCTCTGCCAGCTGCC	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	CAGCATTGTGTCTCCTCCCTGGCA	24
NEK2	NM_002497	S4327/NEK2.f1	SEQ ID NO:169	GTGAGCAGCGCGACTCT	18
NEK2	NM_002497	S4328/NEK2.r1	SEQ ID NO:170	TGCCAATGGTGTAACAACACTCA	23
NEK2	NM_002497	S4329/NEK2.p1	SEQ ID NO:171	TGCCCTCCCGGCTGAGGACT	21
NME1	NM_000269	S2526/NME1.f3	SEQ ID NO:172	CCAACCTGCAGACTCCAA	19
NME1	NM_000269	S2527/NME1.r3	SEQ ID NO:173	ATGTATAATGTTCTGCCAACTTGATG	28
NME1	NM_000269	S4949/NME1.p3	SEQ ID NO:174	CCTGGGACCCATCCGTGGAGACTTCT	25
NPD009	NM_020686	S4474/NPD009.f3	SEQ ID NO:175	GGCTGTGGCTGAGGCTGTAG	20
NPD009	NM_020686	S4475/NPD009.r3	SEQ ID NO:176	GGAGCATTCGAGGTCAAAATCA	21
NPD009	NM_020686	S4476/NPD009.p3	SEQ ID NO:177	TTCCAGAGTGTCTCACCTCCACGAG	28
PCNA	NM_002592	S0447/PCNA.f2	SEQ ID NO:178	GAAGGTGTTGGAGGCACTCAAG	22
PCNA	NM_002592	S0448/PCNA.r2	SEQ ID NO:179	GGTTTACCCGCTGGAGCTAA	21
PCNA	NM_002592	S4784/PCNA.p2	SEQ ID NO:180	ATCCGAGCAGGCCTGTTGATGAG	24
PR	NM_000926	S1336/PR.f6	SEQ ID NO:181	GCATCAGGCTGTCATTATGG	20
PR	NM_000926	S1337/PR.f6	SEQ ID NO:182	AGTAGTTGTGCTGCCCTTCC	20
PR	NM_000926	S4743/PR.p6	SEQ ID NO:183	TGCTCTACCTGTGGGAGCTGAAGGTC	28
PREP	NM_002726	S1771/PREP.f1	SEQ ID NO:184	GGGACGGTGTTCACATTCAAAG	21
PREP	NM_002726	S1772/PREP.r1	SEQ ID NO:185	CAGGATCCCAAGTCAATGTTG	23
PREP	NM_002726	S4929/PREP.p1	SEQ ID NO:186	TGCCAGTCTCCCAACTATCGCGT	24

(continuación)

Gen	Acceso	Nombre	SEQ ID NO	Secuencia	
PTTG1	NM_004219	S452/PTTG1.i2	SEQ ID NO:187	GGCTACTGTGATCTATGTTGATAAGGAA	28
PTTG1	NM_004219	S4526/PTTG1.i2	SEQ ID NO:188	GCTTCAGCCCATCCTTAGCA	20
PTTG1	NM_004219	S4527/PTTG1.p2	SEQ ID NO:189	CACCGGGTGCCTGGTTCTCCA	22
RPLPO	NM_001002	S0256/RPLPO.i2	SEQ ID NO:190	CCATTCTATCATCAACGGGTACAA	24
RPLPO	NM_001002	S0258/RPLPO.i2	SEQ ID NO:191	TCAGCAAGTGGGAAGGTGTAATC	23
RPLPO	NM_001002	S4744/RPLPO.p2	SEQ ID NO:192	TCTCCACAGACAAGGCCAGGACTCG	25
SNRPF	NM_003095	S4489/SNRPF.i2	SEQ ID NO:193	GGCTGGTCGGCAGAGAGTAG	20
SNRPF	NM_003095	S4490/SNRPF.i2	SEQ ID NO:194	TGAGGAAAGTTTGGGATTGA	21
SNRPF	NM_003095	S4491/SNRPF.p2	SEQ ID NO:195	AAACTCATGTAACCACCGCCGAATGTTG	29
Src	NM_004383	S1820/Src.i2	SEQ ID NO:196	CCTGAACATGAAGGAGCTGA	20
Src	NM_004383	S1821/Src.i2	SEQ ID NO:197	CATCAGTCTCCGAACTCC	19
Src	NM_004383	S5034/Src.p2	SEQ ID NO:198	TCCGATGGTGTGCAGCAGCT	21
STK15	NM_003600	S0794/STK15.i2	SEQ ID NO:199	CATCTCCAGGAGGACCACCT	20
STK15	NM_003600	S0795/STK15.i2	SEQ ID NO:200	TCCGACCTTCAATCAITTTCA	20
STK15	NM_003600	S4745/STK15.p2	SEQ ID NO:201	CTCTGTGGCACCCCTGGACTACCTG	24
STMY3	NM_005940	S2067/STMY3.i3	SEQ ID NO:202	CCTGGAGGCTGCAACATACC	20
STMY3	NM_005940	S2068/STMY3.i3	SEQ ID NO:203	TACAATGGCTTTGGAGGATAGCA	23
STMY3	NM_005940	S4746/STMY3.p3	SEQ ID NO:204	ATCCTCCTGAAGCCCTTTTCGCAGC	25
SURV	NM_001168	S02591/SURV.i2	SEQ ID NO:205	TGTTTTGATTCCTCCGGCTTA	20
SURV	NM_001168	S0261/SURV.i2	SEQ ID NO:206	CAAAGCTGCAGCTAGCAAAAAG	24
SURV	NM_001168	S4747/SURV.p2	SEQ ID NO:207	TGCCTTCTCCTCCCTCACCTCTCACCT	28
TFRC	NM_003234	S1352/TFRC.i3	SEQ ID NO:208	GCCAACTGCTTTCATTTGTG	20
TFRC	NM_003234	S1353/TFRC.i3	SEQ ID NO:209	ACTCAGGCCATTTCCCTTTA	20
TFRC	NM_003234	S4748/TFRC.p3	SEQ ID NO:210	AGGATCTGAACCAATACAGAGCAGACA	28
TOP2A	NM_001067	S0271/TOP2A.i4	SEQ ID NO:211	AATCCAAGGGGAGAGTGAT	20
TOP2A	NM_001067	S0273/TOP2A.i4	SEQ ID NO:212	GTACAGATTTTGCCCGGAGA	20
TOP2A	NM_001067	S4777/TOP2A.p4	SEQ ID NO:213	CATATGGACTTTGACTCAGCTGTGGC	26
TS	NM_001071	S0280/TS.i1	SEQ ID NO:214	GCCTCGGTGTCCTTTCA	18
TS	NM_001071	S0282/TS.i1	SEQ ID NO:215	CGTGATGTGCCAATCATG	19
TS	NM_001071	S4780/TS.p1	SEQ ID NO:216	CATCGCAGCTACGCCCTGCTC	22

LISTA DE SECUENCIAS

[0103]

- <110> Salud genómica
- <120> Predicción de la probabilidad de recurrencia del cáncer
- 5 <130> GRF/FP6349666

- <140> 04809450,2
- <141> 2004-06-17

- 10 <150> PCT/US2004/019567
- <151 > 2004-06-17

- <150> US 60/482,339
- <151 > 24/06/2003
- 15
- <160> 216

- <170> FastSEQ para Windows Versión 4.0

- 20 <210> 1 25
- <211> 80
- <212> ADN
- <213> Secuencia artificial

- 25 <220>
- <223> PCR Amplicon
- <400>
- ggctcttg^tg cgtactgtcc ttcgggctgg tgacagggaa gacatcactg agcctgccat 60
- ctgtgctctt cgtcatctga 80

- 30 <210>2
- <211> 81
- <212> ADN
- <213> Secuencia artificial
- <220>
- 35 <223> PCR Amplicon

- <400> 2
- cgttgtcagc acttgaata caagatggtt gccgggtcat gttaattggg aaaaagaaca 60
- gtccacagga agaggttgaa c 81

- <210> 3
- 40 <211> 76
- <212> ADN
- <213> Secuencia artificial

- <220>
- 45 <223> PCR Amplicon
- <400> 3

ES 2 787 475 T3

cctgcaaaag ggaacaagag cccttcgcct ccagatggct ccctgccc ccccccgag 60
atcagagtca accacg 76

<210>4

<211>68

<212> ADN

5 <213> Secuencia artificial

<220>

<223> PCR Amplicon

10 <400> 4

ccgaggttaa tccagcacgt atggggccaa gtgtaggctc ccagcaggaa ctgagagcgc 60
catgtctt 68

<210> 5

<211> 65

15 <212> ADN

<213> Secuencia artificial

<220>

<223> PCR Amplicon

20

<400> 5

tcagctgtga gctgcccgata ccgcccggca atgggacctg ctcttaacct caaacctagg 60
accgt 65

<210>6

25 <211> 84

<212> ADN

<213> Secuencia artificial

<220>

30 <223> PCR Amplicon

<400> 6

ttcaggttgt tgcaggagac catgtacatg actgtctcca ttattgatcg gttcatgcag 60
aataattgtg tgccaagaa gatg 84

<210> 7

35 <211> 82

<212> ADN

<213> Secuencia artificial

<220>

40 <223> PCR Amplicon

<400> 7

ES 2 787 475 T3

atgctgtggc tccttctaa ctggggcttt cttgacatgt aggttgcttg gtaataacct 60
ttttgtatat cacaatttgg gt 82

<210> 8

<211> 68

5 <212> ADN

<213> Secuencia artificial

<220>

<223> PCR Amplicon

10

<400> 8

tggattggag ttctgggaat gtactggccg tggcactgga caacagtgtg tacctgtgga 60
gtgcaagc 68

<210>9

15 <211> 81

<212> ADN

<213> Secuencia artificial

<220>

20 <223> PCR Amplicon

<400> 9

tgagtgtccc ccggtatctt ccccgccctg ccaatcccga tgaaattgga aattttattg 60
atgaaaatct gaaagcggct g 81

25 <210> 10

<211> 77

<212> ADN

<213> Secuencia artificial

30 <220>

<223> PCR Amplicon

<400> 10

tgacaatcag cacacctgca ttcaccgctc ggaagagggc ctgagctgca tgaataagga 60
tcacggctgt agtcaca 77

35

<210> 11

<211> 72

<212> ADN

<213> Secuencia artificial

40

<220>

<223> PCR Amplicon

ES 2 787 475 T3

<400> 11
tgcctgtggt ggggaagctca gtaactggga accaaaggat gatgctatgt cagaacaccg 60
gaggcatttt cc 72

5 <210> 12
<211> 84
<212> ADN
<213> Secuencia artificial

10 <220>
<223> PCR Amplicon

<400> 12
tccctccact cggaaggact atcctgctgc caagagggtc aagttggaca gtgtcagagt 60
cctgagacag atcagcaaca accg 84

15 <210> 13
<211> 67
<212> ADN
<213> Secuencia artificial

20 <220>
<223> PCR Amplicon

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

25 <210> 14
<211> 74
<212> ADN
<213> Secuencia artificial

30 <220>
<223> PCR Amplicon

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

35 tccattttct acctgttaac cttcatcatt ttgtgcaggc cctggaagca aagagaggaa 60
gggaccgact gcat

40 <210> 15
<211> 84
<212> ADN
<213> Secuencia artificial

ES 2 787 475 T3

<220>
<223> PCR Amplicon

5 <400> 15
ctctgagaca gtgcttcgat gactttgcag acttgggtgcc ctttgactcc tgggagccgc 60
tcatgaggaa gttgggcctc atgg 84

<210> 16
<211> 75
10 <212> ADN
<213> Secuencia artificial

<220>
<223> PCR Amplicon

15 <400> 16
gggccctcca gaacaatgat gggctttatg atcctgactg cgatgagagc gggctcttta 60
aggccaagca gtgca 75

<210> 17
20 <211> 68
<212> ADN
<213> Secuencia artificial

<220>
25 <223> PCR Amplicon

<400> 17
cgtgggtgcc ctctatgacc tgctgctgga gatgctggac gccaccgcc tacatgcgcc 60
cactagcc 68

30 <210> 18
<211> 74
<212> ADN
<213> Secuencia artificial

35 <220>
<223> PCR Amplicon

<400>
cacgggacat tcaccacatc gactactata aaaagacaac caacggccga ctgcctgtga 60
agtggatggc accc 74

40 <210> 19
<211> 82
<212> ADN
<213> Secuencia artificial

ES 2 787 475 T3

<220>

<223> PCR Amplicon

5 <400> 19 aaaagacaac caacggccga ctgcctgtga 60

ccaccccgag caaatctgtc ctccccagaa cccctgaatc ctggaggctc acgccccag 60
ccaaagtagg gggactggat tt 82

<210> 20

<211> 67

10 <212> ADN

<213> Secuencia artificial

<220>

<223> PCR Amplicon

15

<400> 20

ccatctgcat ccatcttgtt tgggctcccc acccttgaga agtgcctcag ataataccct 60
ggtggcc 67

<210> 21

20 <211> 86

<212> ADN

<213> Secuencia artificial

<220>

25 <223> PCR Amplicon

<400> 21

aagctatgag gaaaagaagt acacgatggg ggacgctcct gattatgaca gaagccagtg 60
gctgaatgaa aaattcaagc tgggcc 86

30 <210> 22

<211> 76

<212> ADN

<213> Secuencia artificial

35 <220>

<223> PCR Amplicon

<400> 22

caatgccatc ttgcgctaca tcgctcgcaa gcacaacatg tgtggtgaga ctgaagaaga 60
aaagattcga gtggac 76

40

<210> 23

<211> 70

<212> ADN

<213> Secuencia artificial

ES 2 787 475 T3

<220>

<223> PCR Amplicon

5 <400> 23

cggtgtgaga agtgcagcaa gccctgtgcc cgagtgtgct atggctctggg catggagcac 60
ttgcgagagg 70

<210>24

<211> 84

10 <212> ADN

<213> Secuencia artificial

<220>

<223> PCR Amplicon

15

<400> 24

caagggagcg accaactgat cgcacacatg ctttgtttgg atatggagtg aacacaatta 60
tgtaccaaat ttaacttggc aaac 84

<210> 25

20 <211> 70

<212> ADN

<213> Secuencia artificial

<220>

25 <223> PCR Amplicon

<400> 25

agaaccgcaa ggtgagcaag gtggagattc tccagcacgt catcgactac atcagggacc 60
ttcagttgga 70

30 <210> 26

<211> 83

<212> ADN

<213> Secuencia artificial

35 <220>

<223> PCR Amplicon

<400> 26

gcatggttagc cgaagatttc acagtcaaaa tcggagattt tggtatgacg cgagatatct 60
atgagacaga ctattaccgg aaa 83

40

<210> 27

<211> 79

<212> ADN

<213> Secuencia artificial

ES 2 787 475 T3

<220>
<223> PCR Amplicon

5 <400> 27
gatatgattg gtcgctgctt tgtgctcagc caggacctgg ccatccggga tgagttgat 60
ggtggggaat ggaagtct 79

<210> 28
<211> 80
10 <212> ADN
<213> Secuencia artificial

<220>
<223> PCR Amplicon

15 <400> 28
cggactttgg gtgcgacttg acgagcgggtg gttcgacaag tggccttgcg ggccggatcg 60
tcccagtgga agagttgtaa 80

<210> 29
20 <211> 78
<212> ADN
<213> Secuencia artificial

<220>
25 <223> PCR Amplicon

<400> 29
gccagaggc tccatcgtcc atcctcttcc tcccagtcg gctgaactct ccccttctct 60
gcactgttca aacctctg 78

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

35 <220>
<223> PCR Amplicon

<400> 30
ccacctcgcc atgatttttc ctttgaccgg gtattcccac caggaagtgg acaggatgaa 60
gtgtttgaag agattgc 77

40 <210> 31
<211> 69
<212> ADN
<213> Secuencia artificial

ES 2 787 475 T3

<220>
<223> PCR Amplicon

5 <400> 31
tcagtggaga aggagttgga ccagtcaaca tctctgttgt cacaagcagt gtttcctctg 60
gatatggca 69

<210> 32
<211> 66
10 <212> ADN
<213> Secuencia artificial

<220>
<223> PCR Amplicon

15 <400> 32
tgcaaacgct ggtgtcacag ccagcccccc aactgacctc atctggaaga accagaactc 60
gtgggg 66

<210> 33
20 <211> 75
<212> ADN
<213> Secuencia artificial

<220>
25 <223> PCR Amplicon

<400> 33
gacttttgcc cgctacctt cattccggcg tgacaacaat gagctggtgc ttttcatact 60
gaagcagtta gtggc 75

30 <210> 34
<211> 82
<212> ADN
<213> Secuencia artificial

35 <220>
<223> PCR Amplicon

<400> 34
tgatggtcct atgtgtcaca ttcatacag gtttcatacc aacacaggct tcagcacttc 60
ctttggtgtg tttcctgtcc ca 82

40 <210> 35
<211> 65
<212> ADN
<213> Secuencia artificial

45

ES 2 787 475 T3

<220>
<223> PCR Amplicon

<400> 35
5 aacccggcga tcgaaaagat tcttaggaac gccgtaccag ccgctctct caggacagca 60
ggccc 65

<210> 36
<211> 78
10 <212> ADN
<213> Secuencia artificial

<220>
<223> PCR Amplicon

15 <400> 36
ccaacgcttg ccaaatcctg acaattcaga accagctctc tgtgaccca atttgagttt 60
tgatgctgctc actaccgt 78

<210> 37
20 <211> 67
<212> ADN
<213> Secuencia artificial

<220>
<223> PCR Amplicon

25 <400> 37
gagaaccaat ctcaccgaca ggcagctggc agaggaatac ctgtaccgct atggttacac 60
tcgggtg 67

30 <210> 38
<211> 74
<212> ADN
<213> Secuencia artificial

35 <220>
<223> PCR Amplicon

<400> 38
gccgagatcg ccaagatggt gccagggagg acagacaatg ctgtgaagaa tcaactggaac 60
tctaccatca aaag 74

40 <210> 39
<211> 79
<212> ADN
<213> Secuencia artificial

ES 2 787 475 T3

<220>

<223> PCR Amplicon

5 <400> 39

gtgaggcagc ggcactctgg cgactggccg gccatgcctt cccgggctga ggactatgaa 60
gtgttgtaca ccattggca 79

<210> 40

<211> 74

10 <212> ADN

<213> Secuencia artificial

<220>

<223> PCR Amplicon

15

<400> 40

ccaaccctgc agactccaag cctgggacca tccgtggaga cttctgcata caagttggca 60
ggaacattat acat 74

<210> 41

20 <211> 73

<212> ADN

<213> Secuencia artificial

<220>

25 <223> PCR Amplicon

<400> 41

ggctgtggct gaggctgtag catctctgct ggaggtgaga cactctggga actgatttga 60
cctogaatgc tcc 73

30 <210> 42

<211> 71

<212> ADN

<213> Secuencia artificial

35 <220>

<223> PCR Amplicon

<400> 42

gaaggtggtg gaggcactca aggacctcat caacgaggcc tgctgggata ttagctccag 60
cggtgtaaac c 71

40

<210> 43

<211> 85

<212> ADN

<213> Secuencia artificial

ES 2 787 475 T3

<220>

<223> PCR Amplicon

5 <400> 43

gcatacaggct gtcattatgg tgccttacc tgtgggagct gtaaggctctt cttaagagg 60
gcaatggaag ggcagcacia ctact 85

<210>44

<211> 76

10 <212> ADN

<213> Secuencia artificial

}<220>

<223> PCR Amplicon

15

<400> 44

gggacgggtgt tcacattcaa gacgaatcgc cagtctccca actatcgcgt gatcaacatt 60
gacttctggg atcctg 76

<210> 45

20 <211> 74

<212> ADN

<213> Secuencia artificial

<220>

25 <223> PCR Amplicon

<400> 45

ggctactctg atctatggtg ataaggaaaa tggagaacca ggcacccgtg tggttgctaa 60
ggatgggctg aagc 74

30 <210> 46

<211> 75

<212> ADN

<213> Secuencia artificial

35 <220>

<223> PCR Amplicon

<400> 46

ccattctatc atcaacgggt acaaacgagt cctggccttg tctgtggaga cggattacac 60
cttcccactt gctga 75

40

<210> 47

<211> 79

<212> ADN

<213> Secuencia artificial

ES 2 787 475 T3

<220>

<223> PCR Amplicon

5 <400> 47

ggctggtcgg cagagagtag cctgcaacat tcggccgtgg ttacatgag ttaccacctc 60
aatcccaaac ctttcctca 79

<210> 48

<211> 64

10 <212> ADN

<213> Secuencia artificial

<220>

<223> PCR Amplicon

15

<400> 48

cctgaacatg aaggagctga agctgctgca gaccatcggg aagggggagt tcggagacgt 60
gatg 64

<210> 49

20 <211> 69

<212> ADN

<213> Secuencia artificial

<220>

25 <223> PCR Amplicon

<400> 49

catcttccag gaggaccact ctctgtggca cctggacta cctgccccct gaaatgattg 60
aaggtcgga 69

30 <210> 50

<211> 90

<212> ADN

<213> Secuencia artificial

35 <220>

<223> PCR Amplicon

<400> 50

cctggaggct gcaacatacc tcaatcctgt cccaggccgg atcctcctga agcccttttc 60
gcagcactgc tctcctcaa agccattgta 90

40

<210> 51

<211> 79

<212> ADN

<213> Secuencia artificial

ES 2 787 475 T3

<220>

<223> PCR Amplicon

5 <400> 51

tgttttgatt cccgggctta ccaggtgaga agtgagggag gaagaaggca gtgtcccttt 60
tgctagagct gacagcttg 79

<210> 52

<211> 68

10 <212> ADN

<213> Secuencia artificial

<220>

<223> PCR Amplicon

15

<400> 52

gccaaactgct ttcatttgtg agggatctga accaatacag agcagacata aaggaaatgg 60
gcctgagt 68

<210> 53

20 <211> 72

<212> ADN

<213> Secuencia artificial

<220>

25 <223> PCR Amplicon

<400> 53

aatccaaggg ggagagtgat gacttccata tggactttga ctcagctgtg gctcctcggg 60
caaaatctgt ac 72

30 <210> 54

<211> 65

<212> ADN

<213> Secuencia artificial

35 <220>

<223> PCR Amplicon

<400> 54

gcctcgggtgt gcctttcaac atcgccagct acgccctgct cacgtacatg attgcgacaca 60
tcacg 65

40

<210> 55

<211> 22

<212> ADN

<213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 5 <400> 55
 ggctctgtg cgtactgtcc tt 22

 <210> 56
 <211> 23
 10 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 15
 <400> 56
 tcagatgacg aagagcacag atg 23

 <210> 57
 20 <211> 29
 <212> ADN
 <213> Secuencia artificial

 <220>
 25 <223> PCR cebador-sonda

 <400> 57
 aggctcagtg atgtctccc tgcaccag 29

 30 <210> 58
 <211> 23
 <212> ADN
 <213> Secuencia artificial

 35 <220>
 <223> PCR cebador-sonda

 <400> 58
 cgtgtcagc acttgaata caa 23
 40
 <210> 59
 <211>24
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> PCR cebador-sonda

ES 2 787 475 T3

<400> 59
gttcaacctc ttctgtgga ctgt 24

5 <210> 60
<211> 26
<212> ADN 35
<213> Secuencia artificial

10 <220>
<223> PCR cebador-sonda

<400> 60
cccaattaac atgaccggc aacct 26

15 <210> 61
<211> 20
<212> ADN
<213> Secuencia artificial

20 <220>
<223> PCR cebador-sonda

<400> 61
cctgcaaaag ggaacaagag 20

25 <210> 62
<211> 20
<212> ADN
<213> Secuencia artificial

30 <220>
<223> PCR cebador-sonda

<220>
35 <223> PCR cebador-sonda

<400> 62
cgtggtgac tctgatctcg 20

40 <210> 63
<211> 21
<212> ADN 10
<213> Secuencia artificial

45 <220>
<223> PCR cebador-sonda

<400> 63

cttcgcctcc agatggctcc c 21

5 <210> 64
 <211> 21
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> PCR cebador-sonda

<400> 64
 ccgaggtaa tccagcacgt a 21

15 <210> 65
 <211>21
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> PCR cebador-sonda

<400> 65
 aagacatggc gctctcagtt c 21

25 <210>66
 <211> 23
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> PCR cebador-sonda

35 <400> 66
 tgctgggagc ctacactgg ccc 23

40 <210> 67
 <211> 20
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> PCR cebador-sonda

<400> 67
 tcagctgtga gctgcggata 20

<210> 68
 <211> 24
 <212> ADN

<213> Secuencia artificial
 <220>
 <223> PCR cebador-sonda
 5
 <400> 68
 acggtcctag gttgaggtt aaga 24
 <210>69
 10 <211> 19
 <212> ADN
 <213> Secuencia artificial
 <220>
 15 <223> PCR cebador-sonda
 <400> 69
 caggtcccat tgccgggcg 19
 20 <210> 70
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> PCR cebador-sonda
 <400> 70
 ttcaggttgt tgcaggagac 20
 30 <210> 71
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 35 <220>
 <223> PCR cebador-sonda
 <400> 71
 40 catcttcttg ggcacacaat 20
 <210> 72
 <211> 27
 <212> ADN
 45 <213> Secuencia artificial
 <220>
 <223> PCR cebador-sonda

<400> 72
 tgtctccatt attgatcggg tcatgca 27

5 <210> 73
 <211> 22
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> PCR cebador-sonda

<400> 73
 atgctgtggc tccttcctaa ct 22

15 <210> 74
 <211> 27
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> PCR cebador-sonda

<400> 74
 25 acccaaattg tgatatacaa aaaggtt 27

<210> 75
 <211> 30 25
 <212> ADN
 30 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

35 <400> 75
 taccaagcaa cctacatgtc aagaagccc 30

<210> 76
 <211 > 21
 40 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

45 <400> 76
 tggattggag ttctgggaat g 21

<210>77
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> PCR cebador-sonda

 <400> 77
 10 gcttgcactc cacaggtaca ca 22

 <210> 78
 <211> 23
 <212> ADN
 15 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 20 <400> 78
 actggccgtg gcaactggaca aca 23

 <210> 79
 <211>21
 25 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 30
 <400> 79
 tgagtgccc ccggtatctt c 21

 <210>80
 35 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 40 <223> PCR cebador-sonda

 <400> 80
 cagccgctt cagatttca t 21

 45 <210> 81
 <211> 27
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 5 <400> 81
 tgccaatccc gatgaaattg gaaattt 27

 <210> 82
 <211>21
 10 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 15
 <400> 82
 tgacaatcag cacacctgca t 21

 <210> 83
 20 <211> 23
 <212> ADN
 <213> Secuencia artificial

 <220>
 25 <223> PCR cebador-sonda

 <400> 83
 tgtgactaca gccgtgatcc tta 23

 30 <400> 88
 <220>
 <223> PCR cebador-sonda
 <210>84
 <211> 20
 35 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 40
 <400> 84
 caggccctct tccgagcgggt 20

 <210> 85
 45 <211> 18
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 <400> 85
 5 tgcctgtggt gggaagct 18

 <210> 86
 <211> 19
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 15 <400> 86
 ggaaaatgcc tccggttt 19

 <210> 87
 <211> 30
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 25
 <400> 87
 tgacatagca tcatccttg gttcccagtt 30

 <210> 88
 30 <211>21
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> PCR cebador-sonda

 <400> 88
 tccctccact cggaaggact a 21

 40 <210> 89
 <211> 22
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> PCR cebador-sonda

 <400> 89

cgggtgtgc tgatctgtct ca 22

 <210> 90
 <211> 27
 5 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 10
 <400> 90
 tctgacctg tccaactga ccctct 27

 <210> 91
 15 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 20 <223> PCR cebador-sonda

 <400> 91
 tgtctcactg agcgagcaga a 21

 25 <210> 92
 <211> 19
 <212> ADN
 <213> Secuencia artificial

 30 <220>
 <223> PCR cebador-sonda

 <400> 92
 accattgcag ccctgattg 19
 35
 <210> 93
 <211> 24
 <212> ADN
 <213> Secuencia artificial
 40
 <220>
 <223> PCR cebador-sonda

 <400> 93
 45 cttgaggacg cgaacagtcc acca 24

 <210> 94

<211> 27
 <212> ADN
 <213> Secuencia artificial

 5 <220>
 <223> PCR cebador-sonda

 <400> 94
 tccatttct acctgtaac cttcatc 27

 10 <210> 95
 <211> 19
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> PCR cebador-sonda

 <400> 95
 20 atgcagtgg tcccttct 19

 <210> 96
 <211> 23
 <212> ADN
 25 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 30 <400> 96
 ttgctccag ggctgcaca aaa 23

 <210> 97
 <211> 24
 35 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 40 <400> 97
 ctctgagaca gtgcttcgat gact 24

 <210> 98
 <211> 19
 45 <212> ADN
 <213> Secuencia artificial

 <220>

<223> PCR cebador-sonda

<400> 98
 ccatgaggcc caactcct 19

5

<210>99
 <211> 23
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> PCR cebador-sonda

<400> 99
 cagacttggg gcccttgac tcc 23

15

<210> 100
 <211> 20
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> PCR cebador-sonda

<400> 100
 gggccctcca gaacaatgat 20

25

<210> 101
 <211> 21
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <223> PCR cebador-sonda

35

<400> 101
 tgcactgctt ggccttaaag a 21

<210> 102
 <211>25
 <212> ADN
 <213> Secuencia artificial

40

<220>
 <223> PCR cebador-sonda

45

<400> 102
 ccgctctcat cgcagtcagg atcat 25

<210> 103
 <211> 19
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 10 <400> 103
 cgtggtgccc ctctatgac 19

 <210> 104
 <211> 19
 15 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 20
 <400> 104
 ggctagtggg cgcatgtag 19

 <210> 105
 25 <211> 19
 <212> ADN
 <213> Secuencia artificial

 <220>
 30 <223> PCR cebador-sonda

 <400> 105
 ctggagatgc tggacgccc 19

 35 <210> 106
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 40 <220>
 <223> PCR cebador-sonda

 <400> 106
 cacgggacat tcaccacatc 20
 45
 <210> 107
 <211> 19
 <212> ADN

<213> Secuencia artificial
 <220>
 <223> PCR cebador-sonda
 5
 <400> 107
 gggtgccatc cacttcaca 19
 <210> 108
 10 <211> 27
 <212> ADN
 <213> Secuencia artificial
 <220>
 15 <223> PCR cebador-sonda
 <400> 108
 ataaaaagac aaccaacggc cgactgc 27
 20 <210> 109
 <211> 19
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> PCR cebador-sonda
 <400> 109
 ccaccccgag caaatctgt 19
 30 <210> 110
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 35 <220>
 <223> PCR cebador-sonda
 <400> 110
 40 aatccagtc cccctacttt gg 22
 <210> 111
 <211>23
 <212> ADN 15
 45 <213> Secuencia artificial
 <220>
 <223> PCR cebador-sonda

<400> 111
cctgaatcct ggaggctcac gcc 23

5 <210> 112
<211> 20
<212> ADN
<213> Secuencia artificial

10 <220>
<223> PCR cebador-sonda

<400> 112
ccatctgcat ccatcttggt 20

15 <210> 113
<211> 20
<212> ADN
<213> Secuencia artificial

20 <220>
<223> PCR cebador-sonda

<400> 113
25 ggccaccagg gtattatctg 20

<210> 114
<211> 23
<212> ADN
30 <213> Secuencia artificial

<220>
<223> PCR cebador-sonda

35 <400> 114
ctccccacc ttgagaagtg cct 23

<210> 115
<211> 20
<212> ADN
40 <213> Secuencia artificial

<220>
<223> PCR cebador-sonda

45 <400> 115
ggcccagctt gaattttca 20

<210> 116

<211> 27
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> PCR cebador-sonda

<400> 116
 aagctatgag gaaaagaagt acacgat 27

10 <210> 117
 <211>30
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> PCR cebador-sonda

20 <400> 117
 tcagccactg gcttctgtca taatcaggag 30

<210> 118
 <211> 21
 <212> ADN
 25 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

30 <400> 118
 caatgccatc ttgcgctaca t 21

<210> 119
 <211> 25
 35 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

40 <400> 119
 gtccactcga atcttttctt ctca 25

<210> 120
 45 <211> 27 55
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 <400> 120
 5 ctcgcaagca caacatgtgt ggtgaga 27

 <210> 121
 <211> 27 55
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 15 <400> 121
 cgggtgaga agtcagcaa 20

 <210> 122
 <211> 19
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 25 <400> 122
 cctctcgcaa gtgctccat 19

 <210> 123
 30 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> PCR cebador-sonda

 <400> 123
 ccagaccata gcacactcgg gcac 24

 40 <210> 124
 <211>19
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> PCR cebador-sonda

 <400> 124

caagggagcg accaactga 19

<210> 125
 <211> 28
 5 <212> ADN
 <213> Secuencia artificial

<400> 125
 gttgccaag ttaaattgg tacataat 28

10 <210> 126
 <211> 28 5
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> PCR cebador-sonda

<400> 126
 20 ctccatatcc aaacaaagca tgtgtgcg 28

<210> 127
 <211> 19
 <212> ADN
 25 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

30 <400> 127
 agaaccgcaa ggtgagcaa 19

<210> 128
 <211> 21
 35 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

40 <400> 128
 tccaactgaa ggtccctgat g 21

<210> 129
 45 <211> 26
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 5 <400> 129
 tggagattct ccagcacgtc atcgac 26

 <210> 130
 <211> 21
 10 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 15
 <400> 130
 gcatgtagc cgaagatttc a 21

 <210> 131
 20 <211> 30
 <212> ADN
 <213> Secuencia artificial

 <220>
 25 <223> PCR cebador-sonda

 <400> 131
 tttccggtaa tagtctgtct catagatatc 30

 30 <210> 132
 <211> 28
 <212> ADN
 <213> Secuencia artificial

 35 <220>
 <223> PCR cebador-sonda

 <400> 132
 cgcgcatcac caaaatctcc gattttga 28
 40
 <210> 133
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> PCR cebador-sonda

ES 2 787 475 T3

<400> 133
gatatgattg gtcgctgctt tg 22

5 <210> 134
<211> 21
<212> ADN
<213> Secuencia artificial

10 <220>
<223> PCR cebador-sonda

<400> 134
agaacttcca ttccccacca t 21

15 <210> 135
<211> 21
<212> ADN
<213> Secuencia artificial

20 <220>
<223> PCR cebador-sonda

<400> 135
cagccaggac ctggccatcc g 21

25 <210> 136
<211> 19
<212> ADN
<213> Secuencia artificial

30 <220>
<223> PCR cebador-sonda

<400> 136
35 cggactttgg gtgcgactt 19

<210> 137
<211> 24
<212> ADN
40 <213> Secuencia artificial

<220>
<223> PCR cebador-sonda

45 <400> 137
ttacaactct tccactggga cgat 24

<210> 138
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> PCR cebador-sonda

 <400> 138
 10 ccacttgctg aaccaccgct cgt 23

 <210> 139
 <211>18
 <212> ADN
 15 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 20 <400> 139
 gccagaggc tccatcgt 18

 <210> 140
 <211> 23
 25 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 30
 <400> 140
 cagaggtttg aacagtcag aca 23

 <210> 141
 35 <211> 23
 <212> ADN
 <213> Secuencia artificial

 <220>
 40 <223> PCR cebador-sonda

 <400> 141
 cctcttctc cccagtcggc tga 23

 45 <210> 142
 <211>20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 5 <400> 142
 ccacctgcc atgattttc 20

 <210> 143
 <211> 25
 10 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 15
 <400> 143
 gcaatctctt caaacacttc atcct 25

 <210> 144
 20 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 25 <223> PCR cebador-sonda

 <400> 144
 ttgaccggg tattcccacc aggaa 25

 30 <210> 145
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 35 <220>
 <223> PCR cebador-sonda

 <400>145
 40 tcagtggaga aggagttgga 20

 <210> 146
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> PCR cebador-sonda

 <400>146

tgccatatcc agaggaaaca 20

 <210> 147
 <211> 20
 5 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 10
 <400>147
 ccagtcaaca tctctgtgt cacaagca 28

 <210> 148
 15 <211> 19
 <212> ADN
 <213> Secuencia artificial

 <220>
 20 <223> PCR cebador-sonda

 <400>148
 tgcaaacgct ggtgtcaca 19

 25 <210> 149
 <211> 21
 <212> ADN
 <213> Secuencia artificial

 30 <220>
 <223> PCR cebador-sonda

 <400>149
 ccccacgagt tctggttct c 21
 35
 <210> 150
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 40
 <220>
 <223> PCR cebador-sonda

 <400>150
 45 cagccccca actgacctca tc 22

 <210> 151
 <211> 21

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> PCR cebador-sonda

 <400>151
 gacttttgcc cgctaccttt c 21

 10 <210> 152
 <211> 26
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> PCR cebador-sonda

 <400> 152
 gccactaact gcttcagtat gaagag 26

 20 <210> 153
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 25 <220>
 <223> PCR cebador-sonda

 <400> 153
 30 acagctcatt gttgtcacgc cgga 24

 <210> 154
 <211> 24
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 40 <400> 154
 tgatggtcct atgtgcaca ttca 24

 <210> 155
 <211> 20
 45 <212> ADN
 <213> Secuencia artificial

 <220>

<223> PCR cebador-sonda

<400> 155
 tgggacagga aacacaccaa 20

5

<210> 156
 <211> 30
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> PCR cebador-sonda

<400> 156
 caggttcat accaacacag gcttcagcac 30

15

<210> 157
 <211> 18
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> PCR cebador-sonda

<400> 157
 aaccggcga tcgaaaag 18

25

<210> 158
 <211> 18
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <223> PCR cebador-sonda

35

<400> 158
 gggcctgctg tcctgaga 18

<210> 159
 <211> 24
 <212> ADN
 <213> Secuencia artificial

40

<220>
 <223> PCR cebador-sonda

45

<400> 159
 tottagaac gccgtaccag ccgc 24

<210> 160
 <211> 19
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 10 <400> 160
 ccaacgcttg ccaaatcct 19

 <210> 161
 <211> 24
 15 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 20
 <400> 161
 acggtagtga cagcatcaaa actc 24

 <210> 162
 25 <211> 24 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 30
 <400> 162
 aaccagctct ctgtgacccc aatt 24

 <210> 163
 35 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 40 <223> PCR cebador-sonda

 <400> 163
 gagaaccaat ctaccgaca 20

 45 <210> 164
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 <400> 164
 5 cacccgagtg taaccatagc 20

 <210> 165
 <211> 24
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 15 <400> 165
 acaggtattc ctctgccagc tgcc 24

 <210> 166
 <211> 18
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 25 <400> 166
 gccgagatcg ccaagatg 18

 <210> 167
 <211> 27
 30 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 35 <400> 167
 cttttgatgg tagagttcca gtgattc 27

 <210> 168
 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 45 <223> PCR cebador-sonda

 <400> 168
 cagcattgtc tgtctccct ggca 24

<210> 169
 <211> 18
 <212> ADN
 5 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 10 <400> 169
 gtgaggcagc gcgactct 18

 <210> 170
 <211> 23
 15 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 20
 <400> 170
 tgccaatggt gtacaacact tca 23

 <210> 171
 25 <211> 21
 <212> ADN
 <213> Secuencia artificial

 <220>
 30 <223> PCR cebador-sonda

 <400> 171
 tgcttcccg ggctgaggac t 21

 35 <210> 172
 <211> 19
 <212> ADN
 <213> Secuencia artificial

 40 <220>
 <223> PCR cebador-sonda

 <400> 172
 ccaaccctgc agactccaa 19
 45
 <210> 173
 <211> 28
 <212> ADN

<213> Secuencia artificial
 <220>
 <223> PCR cebador-sonda
 5
 <400> 173
 atgtataatg ttctgcca cttgatg 28

 <210> 174
 10 <211> 25
 <212> ADN
 <213> Secuencia artificial

 <220>
 15 <223> PCR cebador-sonda

 <400> 174
 cctgggacca tccgtggaga cttct 25

 20 <210> 175
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 25 <220>
 <223> PCR cebador-sonda

 <400> 175
 ggctgtggct gaggctgtag 20
 30
 <210> 176
 <211>21
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <223> PCR cebador-sonda

 <400> 176
 40 ggagcattcg aggtcaaac a 21

 <210> 177
 <211> 28
 <212> ADN
 45 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

<400> 177
 ttcccagagt gtctcacctc cagcagag 28

5 <210> 178
 <211> 22
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> PCR cebador-sonda

<400> 178
 gaaggtgttg gaggcactca ag 22

15 <210> 179
 <211> 21
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> PCR cebador-sonda
 <400> 179
 ggtttacacc gctggagcta a 21

25 <210> 180
 <211> 24
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> PCR cebador-sonda

<400> 180
 35 atcccagcag gcctcgttga tgag 24

<210> 181
 <211> 20
 <212> ADN
 40 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

45 <400> 181
 gcatcaggct gtcattatgg 20

<210> 182

<211> 20
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> PCR cebador-sonda

<400> 182
 agtagttgtg ctgcccttc 20

10 <210> 183
 <211> 28
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> PCR cebador-sonda

20 <400> 182
 tgccttacc tgtgggagct gtaaggct 28

<210> 183
 <211> 28
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> PCR cebador-sonda

30 <400> 183
 tgccttacc tgtgggagct gtaaggct 28

<210> 184
 <211> 21
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> PCR cebador-sonda

40 <400> 184
 gggacggtgt tcacattcaa g 21

45 <210> 185
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

 5 <400> 185
 caggatccca gaagtcaatg ttg 23

 <210> 186
 <211> 24
 10 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 15
 <400> 186
 tcgccagtct cccaactatc gcgt 24

 <210> 187
 20 <211> 28
 <212> ADN
 <213> Secuencia artificial

 <220>
 25 <223> PCR cebador-sonda

 <400> 187
 ggctactctg atctatgtg ataaggaa 28

 30 <210> 188
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 35 <220>
 <223> PCR cebador-sonda

 <400> 188
 gcttcagccc atccttagca 20
 40
 <210> 189
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <223> PCR cebador-sonda

<400> 189
 cacacgggtg cctggttctc ca 22

5 <210> 190
 <211> 24
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> PCR cebador-sonda

<400> 190
 ccattctatc atcaacgggt acaa 24

15 <210> 191
 <211> 23
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> PCR cebador-sonda

<400> 191
 tcagcaagtg ggaaggtgta atc 23

25 <210> 192
 <211> 25
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> PCR cebador-sonda

<400> 192
 35 tctccacaga caaggccagg actcg 25

<210> 193
 <211> 20
 <212> ADN
 40 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda

45 <400> 193
 ggctggtcgg cagagagtag 20

<210> 194

<211>21
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> PCR cebador-sonda

<400> 194
 tgaggaaagg ttgggattg a 21

10 <210> 195
 <211> 29
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> PCR cebador-sonda

20 <400> 195
 aaactcatgt aaaccacggc cgaatgtg 29

<210> 196
 <211> 20
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> PCR cebador-sonda

30 <400> 196
 cctgaacatg aaggagctga 20

<210> 197
 <211> 19
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> PCR cebador-sonda

40 <400> 197
 catcacgtct ccgaactcc 19

<210> 198
 <211>21
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <223> PCR cebador-sonda

 <400> 198
 5 tcccgatggt ctgcagcagc t 21

 <210> 199
 <211> 20
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 15 <400> 199
 catcttccag gaggaccact 20

 <210> 200
 <211> 20
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda
 25
 <400> 200
 tccgacctc aatcatttca 20

 <210> 201
 30 <211> 24
 <212> ADN
 <213> Secuencia artificial

 <220>
 35 <223> PCR cebador-sonda

 <400> 201
 ctctgtggca ccctggacta cctg 24

 40 <210>202
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <223> PCR cebador-sonda

 <400> 202

cctggaggct gcaacatacc 20

<210> 203
 <211> 23
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> PCR cebador-sonda
 10

<400> 203
 tacaatggct ttggaggata gca 23

<210> 204
 15 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 20 <223> PCR cebador-sonda

<400> 204
 atcctcctga agccctttc gcagc 25

<210> 205
 25 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 30 <223> PCR cebador-sonda

<400> 205
 tgtttgatt cccgggctta 20
 35

<210> 206
 <211> 24
 <212> ADN
 <213> Secuencia artificial
 40

<220>
 <223> PCR cebador-sonda

<400> 206
 45 caaagctgtc agctctagca aaag 24

<210> 207
 <211> 28

<212> ADN
 <213> Secuencia artificial

 <220>
 5 <223> PCR cebador-sonda

 <400> 207
 tgccctcttc ctcctcact tctcacct 28

 10 <210> 208
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <223> PCR cebador-sonda

 <400> 208
 gcccaactgct ttcattgtg 20

 20 <210> 209
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 25 <220>
 <223> PCR cebador-sonda

 <400> 209
 30 actcaggccc atttcctta 20

 <210>210
 <211> 28
 <212> ADN
 35 <213> Secuencia artificial

 <220>
 <223> PCR cebador-sonda

 40 <400> 210
 agggatctga accaatacag agcagaca 28

 <210> 211
 <211> 20
 45 <212> ADN 10
 <213> Secuencia artificial

 <220>

<223> PCR cebador-sonda

<400> 211
aatccaaggg ggagagtgat 20

5

<210> 212
<211> 20
<212> ADN
<213> Secuencia artificial

10

<220>
<223> PCR cebador-sonda

<400>212
gtacagatt tgcccgagga 20

15

<210> 213
<211> 26
<212> ADN
<213> Secuencia artificial

20

<220>
<223> PCR cebador-sonda

<400>213
catatggact ttgactcagc tgtggc 26

25

<210> 214
<211> 18
<212> ADN
<213> Secuencia artificial

30

<220>
<223> PCR cebador-sonda

<400> 214
gcctcgggtg gccttca 18

35

<210> 215
<211> 19
<212> ADN
<213> Secuencia artificial

40

<220>
<223> PCR cebador-sonda

<400> 215
cgtgatgtgc gcaatcatg 19

45

<210> 216
<211> 22
5 <212> ADN
<213> Secuencia artificial

<220>
<223> PCR cebador-sonda
10
<400> 216
catcgccagc tacgcctgc tc 22

REIVINDICACIONES

- 5 1. Un método para predecir la probabilidad de supervivencia libre de enfermedad de una paciente con cáncer de mama con nódulo negativo, ER positivo, invasivo ductal, comprendiendo la determinación del nivel de expresión de la transcripción de ARN de MELK en una célula o una muestra de tejido de cáncer de mama, obtenido de la paciente, normalizado contra el nivel de expresión de una serie de referencia de transcripciones de ARN, comparando el nivel de expresión normalizado de MELK de la paciente, con un nivel de expresión normalizado de MELK en un conjunto de referencia de tejido de cáncer de mama, comprendiendo pacientes que estaban (a) vivas sin recurrencia de cáncer 10 de mama local, regional o distante, (b) vivas con recurrencia contralateral de cáncer de mama, (c) vivas con segundo cáncer primario no de mama, o (d) fallecieron antes de la recurrencia del cáncer de mama, donde un nivel aumentado de expresión normalizado de MELK en la paciente indica una menor probabilidad de supervivencia libre de enfermedad.
2. Un método como se reivindica en la reivindicación 1, donde dicho ARN comprende ARN intrónico.
- 15 3. Un método como se reivindica en la reivindicación 1 o 2, donde dicho ARN es aislado de una muestra de tejido de cáncer de mama embebido en cera y fijado, de dicha paciente.
4. Un método como se reivindica en la reivindicación 1, 2 o 3, donde dicho ARN es aislado de tejido de biopsia de aguja gruesa.
- 20 5. Un método como se reivindica en cualquier reivindicación precedente, donde dicho ARN es aislado de células de aspirado de aguja fina.
6. Un método como se reivindica en cualquier reivindicación precedente, donde el nivel de expresión de dicha transcripción de ARN se determina utilizando un array que comprende un polinucleótido hibridando en el gen inmovilizado sobre una superficie sólida.
- 25 7. Un método como se reivindica en cualquier reivindicación precedente, donde el nivel de expresión de dicha transcripción de ARN se determina por reacción de la cadena de polimerasa de transcripción inversa (RT-PCR).
8. Un método como se reivindica en cualquier reivindicación precedente, donde el ARN está fragmentado.