(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 548 299**

(21) Número de solicitud: 201430349

(51) Int. Cl.:

C12Q 1/68 (2006.01)

(12) SOLICITUD DE PATENTE A1

(22) Fecha de presentación:

13.03.2014

(43) Fecha de publicación de la solicitud:

15.10.2015

(56) Se remite a la solicitud internacional:

PCT/ES2015/070179

(71) Solicitantes:

UNIVERSIDAD DE MÁLAGA (60.0%)
Plaza de El Ejido s/n
29071 Málaga ES;
SERVICIO ANDALUZ DE SALUD (30.0%) y
FUNDACIÓN PÚBLICA ANDALUZA PARA LA INVESTIGACIÓN DE MÁLAGA EN BIOMEDICINA Y SALUD (FIMABIS) (10.0%)

(72) Inventor/es:

LOZANO CASTRO, José;
ALBA CONEJO, Emilio;
PÉREZ RIVAS, Luís Gustavo;
JEREZ ARAGONÉS, José y
RIBELLES ENTRENA, Nuria

(74) Agente/Representante:

FÚSTER OLAGUIBEL, Gustavo Nicolás

(54) Título: **Firma de microARN como indicador del riesgo de recurrencia temprana en pacientes con cáncer de mama**

(57) Resumen:

La presente invención se sitúa en el campo de la oncología y el tratamiento del cáncer. La presente invención incluye métodos para pronosticar el riesgo de recurrencia de tumores de mama empleando la firma de expresión de determinados miARNs. Específicamente, la presente invención se refiere a un procedimiento para determinar el riesgo de recurrencia de cáncer de mama que comprende medir los niveles de expresión de al menos un miARN seleccionado de entre el grupo que consiste en miR-149-5p (SEQ ID NO: 1), miR-10a-5p, (SEQ ID NO: 2), miR-20b-5p, (SEQ ID NO: 3), miR-30a-3p (SEQ ID NO: 4); y miR-342-5p, (SEQ ID NO: 5) en una muestra del tumor, en el que una alteración del nivel de expresión de al menos un miARN en el tumor, con respecto al nivel de expresión en una muestra control, es indicativa de un elevado riesgo de recurrencia del tumor. Además, la presente invención se refiere a herramientas y kits para llevar a cabo el procedimiento de la presente invención.

ES 2 548 299 A1

DESCRIPCIÓN

FIRMA DE MICROARN COMO INDICADOR DEL RIESGO DE RECURRENCIA TEMPRANA EN PACIENTES CON CÁNCER DE MAMA

Campo técnico

La presente invención se sitúa en el campo de la oncología y el tratamiento y pronóstico del cáncer. La presente invención incluye métodos para pronosticar el riesgo de recurrencia de tumores de mama empleando la firma de expresión de determinados miARNs.

Introducción

El cáncer de mama constituye un grupo de enfermedades heterogéneas que se pueden clasificar en base tanto a sus características clínicas como moleculares [1-5]. A nivel mundial, el cáncer de mama es el cáncer invasivo más común en las mujeres. Las mejoras en la detección precoz de los tumores primarios y el desarrollo de nuevas terapias, junto con el uso sistemático de la quimioterapia adyuvante, han reducido drásticamente las tasas de mortalidad y aumentado la supervivencia libre de enfermedad (DFS) en el cáncer de mama. Sin embargo, aproximadamente un tercio de los pacientes sometidos a la extirpación del tumor de mama desarrollará metástasis, la mayor amenaza para la supervivencia asociada al tumor, estrechamente asociada con un peor pronóstico [6,7].

Al cáncer se le llama recurrente cuando reaparece después del tratamiento. La recurrencia puede ser local (en el mismo seno o en la cicatriz de la mastectomía) o en un área distante.

El riesgo de recaída (o recurrencia) después de la resección del tumor no es constante en el tiempo. Un detallado examen de los estudios de seguimiento a largo plazo en las últimas dos décadas revela una función de riesgo bimodal con dos picos de recurrencia temprana y tardía que ocurren al cabo de 1,5 y 5 años, respectivamente, seguido de una meseta casi plana en la que el riesgo de recurrencia tiende a cero [8-10].

Algunos investigadores han propuesto una relación de causalidad entre la cirugía del tumor y el patrón bimodal de recurrencia (es decir, un efecto iatrogénico) [11]. De acuerdo con ese modelo, la extirpación quirúrgica del tumor primario de mama aceleraría el crecimiento de focos metastásicos latentes, alterando el equilibrio entre factores circulantes pro- y anti-angiogénicos [9, 11-14]. Esta hipótesis se ve apoyada por el hecho de que los dos picos de recaída se observan independientemente de otros factores, a parte de la cirugía, como el estado de los ganglios axilares, el tipo de cirugía o la administración de terapia adyuvante.

Aunque los tumores negativos para los receptores de estrógeno (ER) se asocian con un mayor riesgo de recurrencia temprana [15], el patrón de distribución bimodal se observa con independencia del estado de los receptores hormonales [16].

Por lo tanto, dde acuerdo con este modelo de función de riesgo bimodal, la recurrencia precoz se puede explicar como un efecto iatrogénico de la extirpación quirúrgica del tumor primario. Las recurrencias tardías, por el contrario, no están sincronizados y, por tanto, se cree que no están influenciadas por el procedimiento quirúrgico. En su lugar, se propone que son el resultado de la expansión repentina de micrometástasis de una sola célula durante la evolución natural de la enfermedad [11]. Siguiendo esta hipótesis, informes anteriores se han centrado en focos de metástasis latentes, el microambiente circundante del tumor o incluso señales presentes en el suero para encontrar factores que podrían explicar los diferentes riesgos de recurrencia.

Otros estudios también sugieren que la dinámica de la recaída del tumor puede ser una consecuencia de la intervención quirúrgica para extirpar el tumor primario, que pudieran alterar los niveles circulantes de VEGF, TNFα y otras citoquinas inflamatorias [17-19]. Sin embargo, aún se requiere evidencia empírica que demuestre un vínculo molecular entre la cirugía del tumor primario de mama y el patrón de recurrencia bimodal.

En general, se ha informado que algunos factores biológicos están implicados en diversos tipos particulares de cánceres, incluyendo las mutaciones somáticas de genes concretos, la expresión alterada de proteínas, la activación alterada de proteínas, y modelos alterados de expresión génica.

La identificación de patrones de expresión distintivos basados en el perfil de expresión génica mediante procedimientos de perfilado génico, tal como una micromatriz ("*microarray*", en inglés) ha dado lugar a una clasificación de los tumores de mama en cinco subtipos diferentes: luminal A, luminal B, tumores que sobreexpresan HER2 (HER2+), tipo basal y tipo normal [3,4]. Esta clasificación se ha adoptado en la rutina clínica, definiéndose subtipos intrínsecos con características histológicas distintas, respuesta al tratamiento farmacológico y resultados clínicos [3,20-23]. Los subtipos HER2+ y tipo basal se asocian comúnmente con un riesgo de recurrencia más elevado, mientras que los tumores luminales frecuentemente se relacionan con una supervivencia libre de tumor más duradera [24-26].

Un estudio reciente del Consorcio Internacional de la Taxonomía Molecular del Cáncer de

Mama (METABRIC, por sus siglas en inglés, *Molecular Taxonomy of Breast Cancer International Consortium*) ha propuesto una nueva clasificación basada en el genoma, mediante la integración tanto de datos genómicos como transcriptómicos. Esta nueva estratificación molecular se basa en el impacto del número de copias somáticas aberrantes (CNAs, por sus siglas en inglés, "*copy number aberrations*") en el transcriptoma y clasifica los tumores de mama en 10 agrupaciones integradoras (del inglés "*integrative clusters*", IntClusts 1-10), cada uno asociado con distintos resultados clínicos [27,28].

Los microARN (miARN o "*miRNA*, en inglés) son pequeñas moléculas de ARN de cadena sencilla que desempeñan un importante papel en la regulación de la expresión génica [29,30]. Se transcriben como grandes precursores de ARN (PRI-miARNs o "*PRI-miRNAs*", en inglés) que se procesan secuencialmente en el núcleo para producir una horquilla de ARN ("*RNA hairpin*", en inglés) de 65 nucleótidos (nt), denominada el precursor-miARN (pre-miARN) y en el citoplasma para producir un miARN de 19-23 nt, maduro y activo [31-33].

En general, los miARNs actúan como moduladores negativos de la expresión génica. Los miARNs se unen a una secuencia parcialmente complementaria que generalmente se encuentra en la región no traducida 3' (3'-UTR, por sus siglas en inglés, "*untraslated region*") de su ARNm diana e inhiben su traducción [34]. Debido a esta complementariedad parcial, un único miARN puede dirigirse a múltiples transcritos, por lo tanto, regulando a la baja la expresión de muchas proteínas en la misma o diferente vía [29]. Actualmente se han identificado 1.872 precursores de miARNs y 2.578 miARNs maduros en el genoma humano (miRBase 20, www.mirbase.org) [35], aunque no se conoce en detalle el papel biológico de la mayor parte de ellos.

Como sucede con la expresión de ARNm, el perfil transcripcional de los miARNs puede variar entre los diferentes tejidos y etapas de desarrollo. Alteraciones de los patrones de expresión del miARN y sus secuencias son comunes en varias enfermedades, incluyendo el cáncer [36,37]. Los miARNs están involucrados en muchas rutas desreguladas en las células tumorales, especialmente las relacionadas con los marcadores tumorales [38,39] y a menudo se encuentran regiones de punto de ruptura ("*breakpoint regions*", en inglés) que en cáncer se encuentran amplificadas, eliminadas o tanslocadas [36]. Varios miARNs poseen actividad oncogénica (oncomiRs) o actividad supresora de tumores (TS-MIR) y, por lo tanto, pueden contribuir a la tumorogénesis, a la progresión tumoral y a la metástasis [40-42].

Además, los perfiles de expresión del miARN pueden proporcionar información molecular de

relevancia clínica en el cáncer [43]. Así, los tumores de diferente origen se pueden clasificar de acuerdo con patrones de expresión de miARN específico (también conocidos como "firmas de miARN" o "*miRNA signatures*", en inglés).

Diferentes subtipos de un mismo tumor pueden también diferenciarse por el patrón de expresión de miARN, y en algunos casos, este patrón de expresión proporciona también un valor predictivo en la evaluación clínica [37,44-46]. De esta manera, mediante dos tipos de enfoque en el cáncer de mama, es decir, el perfilado y el análisis funcional del miARN, ha sido posible mejorar nuestro conocimiento acerca de algunos marcadores moleculares asociados con el cáncer de mama.

Los microARNs se encuentran generalmente bien conservados en una amplia gama de tipos de muestras, incluyendo los fluidos corporales y tejidos fijados en formalina incluidos en parafina (FFPE, por sus siglas en inglés "*formalin-fixed paraffin-embedded*") [47]. Varios miARNs se han relacionado con la metástasis del cáncer de mama [48,49].

Como se ha señalado anteriormente, cualquier miARN dado tiene una gran variedad de dianas diferentes (en algunas estimaciones aproximadamente 200 dianas por miARN en promedio), y un miARN dado está implicado normalmente en la multiplicidad de diferentes rutas celulares. Las estimaciones del riesgo de recurrencia, decisiones de tratamiento y respuesta al tratamiento se han basado hasta el momento principalmente en la etapa del tumor Hasta el presente momento no ha sido posible predecir qué miARN podría estar asociado con la recurrencia <u>temprana</u> (<u>en los primeros dos años, que es la que tiene un peor pronóstico</u>) de un tumor concreto.

La predicción de recurrencia temprana en tumores de mama representa un reto importante en la práctica clínica, debido a que una recurrencia temprana a menudo se corresponde con un tumor más agresivo, con menos opciones terapéuticas y un pronóstico más desalentador. De hecho, hasta la mitad de todas las recaídas residen en el pico temprano de recurrencia descrito por Demicheli y colegas [16]. Existe, por tanto, la necesidad de identificar biomarcadores fiables en pacientes que padecen/han padecido cáncer de mama, lo que permitiría predecir el riesgo de recurrencia del tumor tras la cirugía o un tratamiento específico, también ayudaría en la elección de la terapia apropiada o el seguimiento adecuado para estos individuos y de esta manera, en último extremo, ayudaría a la prolongación de la vida o a la mejora de la calidad de vida del paciente. Además, sería necesario identificar nuevas dianas terapéuticas así como nuevos medicamentos para

prevenir o tratar el cáncer de mama recurrente.

Términos y abreviaturas

ADN	ácido desoxirribonucleico
ARNm	ARN mensajero
ER	receptores de estrógeno
PRI-miARNs	precursores de ARN
nt	nucleótidos
pre-miARN	precursor-miARN
3'-UTR	región no traducida 3’
FFPE	fijados en formalina incluidos en parafina
miARN	microARN
miR	microARN
ARN	ácido ribonucleico
RQ-PCR	PCR cuantitativa en tiempo real
SEQ	secuencia
cADN	ADN complementario
RFS	Supervivencia libre de recurrencia
	ROC “*Receiver Operating Characteristic*”, o Característica Operativa del Receptor

Breve descripción de la invención

Tal como se detalla adicionalmente a continuación, los microARN (miARN) son moléculas cortas de ARN no codificante que funcionan como reguladores post-traduccionales de la expresión génica.

Los inventores han identificado cinco miARNs como biomarcadores para pronosticar el riesgo de recurrencia de cáncer de mama en pacientes con cáncer de mama.

Un primer aspecto de la presente invención se refiere por lo tanto a un procedimiento para pronosticar el riesgo de recurrencia de cáncer de mama en un sujeto que comprende medir los niveles de expresión de al menos un miARN seleccionado de entre el grupo que consiste en:

miR-149-5p (SEQ ID NO: 1)

miR-10a-5p, (SEQ ID NO: 2)
miR-20b-5p, (SEQ ID NO: 3)
miR-30a-3p (SEQ ID NO: 4); y
miR-342-5p, (SEQ ID NO: 5)

en una muestra del tumor, en el que una reducción del nivel de expresión de al menos un miARN en el tumor, con respecto a un control, es indicativo de un elevado riesgo de recurrencia del tumor.

En un segundo aspecto, la invención proporciona un procedimiento para clasificar un sujeto humano que padece cáncer de mama en uno de dos grupos, en el que el grupo 1 comprende los sujetos identificables como con alto riesgo de recurrencia temprana mediante el procedimiento de la invención descrito anteriormente, y en el que el grupo 2 representa el resto de sujetos.

La invención proporciona además una composiciones farmacéuticas, medicamentos, anticuerpos y, en general, cualquier tipo de terapia adecuada para tratar un sujeto humano del grupo 1 identificable mediante el procedimiento descrito anteriormente.

En un tercer aspecto, la presente invención proporciona un procedimiento para pronosticar la supervivencia de un sujeto con cáncer de mama que comprende medir los niveles de expresión de al menos un miARN seleccionado de entre el grupo que consiste en:

miR-149-5p (SEQ ID NO: 1)
miR-10a-5p, (SEQ ID NO: 2)
miR-20b-5p, (SEQ ID NO: 3)
miR-30a-3p (SEQ ID NO: 4); y
miR-342-5p, (SEQ ID NO: 5)

en una muestra del tumor, en el que una reducción del nivel de expresión de al menos un miARN en el tumor, con respecto a un control, es indicativa de una baja supervivencia.

En un cuarto aspecto, la invención proporciona un kit que comprende al menos cinco oligonucleótidos, caracterizado por que los cinco oligonucleótidos se seleccionan de entre el grupo que consiste en cinco oligonucleótidos capaces de hibridarse con uno cualquiera de dos o más, y preferiblemente todos los miARN o sus cDNAs tal como se ha definido en las SEQ ID NOs: 1 a 5 en condiciones de rigor.

En un quinto aspecto la invención proporciona además un miARN según se define en las

siguientes secuencias: SEQ ID NO: 1 (miR-149-5p -5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) para su uso en un método de tratamiento para el cáncer de mama o, alternativamente, para su uso en un método de prevención de la recurrencia del cáncer de mama en un sujeto que sufre o ha sufrido cáncer de mama.

Breve descripción de las figuras

Figura 1. La Figura 1 muestra la estabilidad de los niveles de expresión de RNU48 y miR-16-5p en muestras de tumores recurrentes (B) y no recurrentes (A). Para este cálculo, tras la lectura de cada microarray según el Ejemplo 2 de la presente invención, todas las señales se normalizan teniéndose en cuenta todas las sondas del microarray mediante un procedimiento matemático estándar, el RMA (Robust Multi-Array Average, cita: ^ Irizarry, RA; Hobbs, B; Collin, F; Beazer-Barclay, YD; Antonellis, KJ; Scherf, U; Speed, TP (2003). "Exploration, normalization, and summaries of high density oligonucleotide array probe level data". Biostatistics 4 (2): 249-64. PMID 12925520.). Una vez que todos los microarrays han sido corregidos mediante RMA se comparan los datos de expresión. Aquellos que varíen poco en la población se emplean como control. Así, en esta Figura 1, se muestran los valores de dos ARNs, RNU48 y miR16 que son relativamente constantes en la población de pacientes (1 paciente= 1 microarray).

Figura 2. Perfiles de expresión de los miARN en tumores primarios de pacientes con distintas prognosis. Se obtuvo el ARN total a partir de 71 tumores de mama, se convirtió a cADN y se hibridó al chip micromatriz de miARN Affymetrix 2.0. Después de la normalización, se analizó la diferencia de los datos de expresión de miARN empleando un algoritmo de agrupamiento jerárquico no supervisado (unsupervised hierarchical clustering). Las barras de color en la parte superior del mapa de calor se refieren al grupo de pronóstico y al subtipo intrínseco de cada tumor. El Grupo A incluye los tumores de los pacientes libres de la enfermedad ≥ 60 meses después de la cirugía, el grupo B incluye los tumores de pacientes con recurrencia temprana (≤ 24 meses) y el grupo C incluye los tumores de pacientes con recurrencia tardía (50 a 60 meses después de la cirugía). Los tumores se agrupan en dos grupos principales (cluster 1 y cluster 2 en la Figura), mostrándose perfiles de expresión opuestos y fuertemente asociados con los grupos de pronóstico.

En el grupo 1 se incluyen la mayor parte de tumores luminales y / o
no recurrentes, mientras que en el grupo 2 se incluyen en su mayoría tumores de tipo basal

y / o de recurrencia temprana.

Figura 3. El perfil de expresión de miARN se asocia con la recurrencia precoz en cáncer de mama. Agrupación jerárquica de 71 muestras de tumores basada en el perfil de expresión de 5 miARNs. Téngase en cuenta que los niveles más bajos de expresión de los 5 miARN definen un grupo distinto, cluster 2b en la Figura, que incluye principalmente los tumores de pacientes de "alto riesgo" (grupo B). Por el contrario, la mayoría de los pacientes con buen pronóstico (grupo A) presentan tumores con niveles de expresión de los 5-miARNs normales o superiores a los niveles normales de expresión, definiendo un grupo diferente, cluster 1b en la Figura ("bajo riesgo").

Figura 4. El perfil de expresión de los 5 miARN discrimina pacientes con distinta supervivencia libre de recurrencia (RFS). **A)** Gráfico de Kaplan-Meier para toda la cohorte de pacientes incluidos en este estudio. **B)** Aquellos pacientes cuyos tumores mostraron una reducción de los niveles de expresión de los 5 miARN (es decir, los de cluster 2b, Fig. 3) se clasificaron como de "alto riesgo" (línea gris claro) y se calculó su RFS acumulativa (línea gris claro). La RFS también se calculó para los demás pacientes de la cohorte ("de bajo riesgo", la línea gris oscuro). El gráfico de Kaplan-Meier muestra que la expresión de los 5 miARN discrimina específicamente tumores con un riesgo global de recurrencia temprana más alto.

Figura 5. Curva ROC (acrónimo de “Receiver Operating Characteristic”, o Característica Operativa del Receptor) para la recurrencia temprana del cáncer de mama según los niveles de expresión de los 5 miARN. Las curvas ROC generadas utilizando la información del pronóstico y de los niveles de expresión de los 5 miARN pueden discriminar entre los pacientes que desarrollarán recurrencia temprana y los que van a permanecer libres de la enfermedad. Téngase en cuenta que, a pesar de que miR-30-3p y miR10a-5p presentan un elevado valor de área bajo la curva (AUC) los niveles de expresión de los 5 miARN presentan el valor predictivo más sólido (AUC = 0,993) para discriminar aquellos pacientes con elevado riesgo de presentar recurrencia precoz (grupo B en nuestra cohorte).

Figura 6. Predicción de dianas de mARN susceptibles de ser reguladas por la firma de los 5 miARN. Redes biológicas fueron creadas utilizando el software Cytoscape. Cada red incluye dos tipos de nodos: los cinco miARNs individuales incluidos en la firma 5-miARN y su pronosticado mARN objetivos (círculos amarillos), obtenidos a partir de dos bases de datos públicas diferentes (miRTarBase y miRecords). El número de bases de datos incluidas en el

análisis define la red de interacción regulatoria (RIN) del umbral. Por lo tanto, en Rin = 1 la red incluye todos los objetivos de mARN que aparecen en, al menos, una base de datos. Las bases de datos incluidas en el RIN se identifican por el color de las flechas que conectan: miRTarBase y miRecords. Aunque muchos mARNs son objetivos potenciales de miR-149-5p y miR-342-5p, las versiones miRTarBase y miRecords incluidos en este estudio no revelaron ningún objetivo experimentalmente validados para los dos miARNs.

Figura 7. Términos de Ontología Génica (GO) relacionados con los ARNm predichos como dianas de la firma 5-miARN. Se realizó un análisis GO utilizando términos del vocabulario "proceso biológico". Se muestra el número de identificación GO (GO ID), el nombre del proceso biológico (GO term) y los mRNA diana asociados a cada término GO. Únicamente se consideraron términos GO con evidencia experimental y p-valor corregido <0.01

Figura 8. -. Los pacientes con un mayor riesgo de recaída tienen tumores de mama con mayor capacidad proliferativa. La capacidad angiogénica, proliferativa y la expresión del receptor de estrógenos (ER) en los tumores primarios se determinó mediante inmunohistoquímica con anticuerpos específicos para el factor de crecimiento endotelial vascular (VEGF), la proteína Ki67 y el receptor hormonal ER, respectivamente. Los tumores se clasificaron de acuerdo con el estado de VEGF (High VEGF/Low VEGF, indicando expresión elevada o expresión baja/ausencia de expresión, respectivamente), Ki67 (Ki67+ o Ki67-, indicando expresión o ausencia de expresión, respectivamente) y ER (ER+ o ER-, indicando expresión o ausencia de expresión, respectivamente) y se calculó la de RFS acumulada de los pacientes. Los gráficos de Kaplan-Meier muestran una reducción de RFS en los pacientes con tumores con una elevada expresión de VEGF, positivos para Ki67 y negativos para ER, aunque las diferencias sólo fueron estadísticamente significativas para Ki67 (Log-rank p = 0,044). Los 71 tumores incluidos en este estudio fueron procesadas para la tinción Ki67 y ER, mientras que sólo 67 pudieron ser procesados para tinción VEGF.

Figura 9. Resumen de las potenciales funciones biológicas afectadas por la expresión de la firma 5-miRNA en tumores de mama recurrentes (< 2 años). El conjunto de los 30 mARNs diana predichas para la firma 5-miRNA (mostradas en la Fig. 6) se integraron en el servidor KEGG (Kyoto Encyclopedia of Genes and Genomes) para generar un mapa de las proteínas clave (indicadas con estrellas) y las vías más probables asociadas a los mARNs diana. Obsérvese que un aumento de los mARN diana (debido a la expresión reducida de la firma 5-miRNA en los tumores con recidiva precoz) daría lugar a un aumento neto en la proliferación y la angiogénesis.

Descripción detallada de la invención

La descripción detallada da a conocer variantes específicas y/o preferidas de las características individuales de la invención. La presente invención contempla también como realizaciones particularmente preferidas aquellas realizaciones que se generan combinando dos o más de las variantes específicas y/o preferidas descritas para dos o más de las características de la presente invención.

Se proporciona en la presente invención un procedimiento para la predicción de la recurrencia temprana del cáncer de mama, en el que el sujeto padece o ha padecido cáncer de mama.

Según estudios recientes, para las mujeres con cáncer de mama en fase temprana sometidas al tratamiento conservador de la mama, la recidiva (recurrencia) locorregional sigue siendo el factor pronóstico más importante al cabo de cinco años sin enfermedad. Además, después de un intervalo de diez años sin enfermedad, la recidiva locorregional sigue siendo el único factor pronóstico independiente que queda. La recidiva locorregional, el tamaño del tumor, la extensión del cáncer a los ganglios linfáticos, la edad temprana, la presencia de receptores de estrógenos y el tratamiento con quimioterapia son factores pronósticos independientes con impacto significativo sobre el resultado a largo plazo, siendo la recidiva locorregional el factor más potente.

El modelo de recurrencia bimodal se basa en el concepto de la latencia del tumor y de que la homeostasis metastásica se perturba al eliminar el tumor [11,12]. Las células tumorales dejan el sitio primario del tumor y se instalan en tejidos diferentes, distantes donde permanecen en estado latente durante un período variable de tiempo, ya sea como células individuales o como micrometástasis [102-104]. La mayoría de micrometástasis no proliferan activamente y sólo una pequeña proporción (<10 %) presenta un fenotipo angiogénico [11].

Las micrometástasis no angiogénicas permanecen en reposo en ausencia de un cambio angiogénico, e incluso las micrometástasis angiogénicos no pueden pasar de ser focos avasculares sin la neovascularización adecuada. El letargo resulta entonces del equilibrio entre las señales pro- y anti- angiogénicos que afectan a las micrometástasis. Alteraciones genéticas adquiridas en el curso natural de la enfermedad podrán producir a su vez un

desequilibrio entre los factores pro-y anti- angiogénicos que favorezcan la neovascularización y el crecimiento de los focos micrometastásica (el “interruptor angiogénico”) [103]. El modelo de recurrencia bimodal supone que el tumor primario contribuye a la homeostasis de la metástasis a distancia mediante la liberación de factores anti- angiogénicos que mantienen a las células metastásicos angiogénicos en un estado avascular latente. La extirpación quirúrgica del tumor aboliría la represión angiogénica y favorecería el proceso metastásicos. Además, diversos factores de crecimiento y citocinas liberadas como consecuencia de la herida del tejido durante la cirugía añadirían estímulos angiogénicos, impulsando el proceso metastásico [9, 11, 13]. Los inventores han descrito la presencia de características intrínsecas del tumor que favorecen el crecimiento metastásico.

Los sujetos cuya recurrencia se predice en la presente invención son sujetos humanos que padecen o han padecido cáncer de mama. Los términos “sujeto humano”, “sujeto” y “paciente” se usan por tanto de manera indistinta en esta memoria descriptiva.

“Uno o más” tal como se usa también en la presente memoria descriptiva incluye uno y la especificación individualizada de cualquier número que sea más de uno, tal como dos, tres, cuatro, cinco, seis, etc. “Más de uno” o “algunos” tal como se usa en la presente memoria descriptiva incluye la especificación individualizada de cualquier número que sea más de uno, tal como dos, tres, cuatro, cinco, seis, etc.

A no ser que se especifique expresamente de otra forma, el término “que comprende” se usa en el contexto de este documento para indicar que pueden estar opcionalmente presentes miembros adicionales además de los miembros de la lista introducida por el término “que comprende”. Se contempla, sin embargo, como una realización específica de la presente invención que el término “que comprende” abarque la posibilidad de que no se encuentren presentes miembros adicionales, es decir, a los fines de esta realización “que comprende” debe entenderse que tiene el significado de “que consiste en”.

En el contexto de esta memoria descriptiva, el término “tratamiento” o “tratar” significan la administración de un agente para evitar, aliviar o eliminar el cáncer de mama o de uno o más síntomas asociados con dicha enfermedad. “Tratamiento” incluye también evitar, aliviar o eliminar las secuelas fisiológicas de la enfermedad. En el contexto de esta memoria, el término “tratamiento” o “tratamiento antitumoral” también incluye la extirpación quirúrgica del tumor. En el contexto de esta invención, el término “aliviar” se entiende que significa cualquier mejora de la situación del paciente tratado, de manera subjetiva (los sentimientos

de o acerca del paciente) y de manera objetiva (los parámetros medidos).

En el contexto de esta memoria descriptiva, el término “firma miARN”, (o términos semejantes como por ejemplo “firma de los 5 miARNs”) se refiere al perfil de expresión de al menos uno, preferiblemente 2, más preferiblemente 3, aún más preferiblemente 4 y todavía más preferiblemente los 5 miARNs de la invención (según se describen en las SEQ ID NO.:1-5).

Procedimientos para pronosticar el riesgo de recurrencia de cáncer de mama

La presente invención se basa, en parte, en la identificación de microARN particulares en células de cáncer de mama y la asociación de estos microARN con características pronósticas particulares.

Los microARN son ARN monocatenarios de ~22-nucleótidos (aprox. 18–25 nucleótidos) que regulan negativamente (inhiben) la expresión génica mediante la inhibición de la traducción o la escisión del ARNm. De esta manera, los miARN son reguladores postranscripcionales que se unen a secuencias complementarias en las transcripciones de ARN mensajeros diana (ARNm), dando como resultado la represión de la traducción o la degradación de la diana y el silenciamiento del gen. La mayoría de genes de miARN conocidos se encuentran en regiones intergénicas o están orientados en sentido contrario a los genes adyacentes y se piensa, por tanto, que se transcriben como unidades independientes. Sus genes se transcriben usualmente mediante la ARN polimerasa II y los transcritos procesados, exportados desde el núcleo y procesados adicionalmente mediante maquinarias específicas, como se conoce bien en la técnica (véase por ejemplo. He *y col.*, Nat. Rev. Genet. Jul 2004; 5(7):522-31). Se puede acceder a las secuencias de miARN en http://www.mirbase.org.

En un aspecto particular de la invención, la comprensión de los detalles moleculares de la acción de los miARN de la invención no es crítica, debido a que los niveles detectados de los miARN indicadores solo permiten llevar a cabo el procedimiento de la invención. Se entiende que la molécula activa de ARN de 18-25 nucleótidos también se puede producir directamente por síntesis química o biológica, sin tener que procesarse a partir del miR precursor.

En un sistema de nomenclatura normalizado, los nombres se asignan a miARN experimentalmente confirmados como sigue: el prefijo "mir" está seguido (por un guion y) un

número, en donde el último puede indicar el orden de nomenclatura. El "mir-" en minúscula se refiere al pre-miARN, mientras que un "miR-" en mayúscula se refiere a la forma madura. Los miARN con secuencias casi idénticas, excepto uno o dos nucleótidos, se indican letras minúsculas adicionales, por ejemplo miR-10a-5p. Los pre-miARN que conducen a miARN maduros 100% idénticos pero que se localizan en diferentes lugares en el genoma se indican con un sufijo numérico adicional separado por guion. Se puede designar la especie original con un prefijo de tres letras, por ejemplo, miR-10a-5p es un miARN humano (Homo sapiens). Debido a que en el contexto de este documento, todos los miARN individualizados son miARN humanos, el prefijo “hsa-“ se omite algunas veces. Cuando dos microARN se originan de brazos opuestos del mismo pre-miARN, se denotan con un sufijo -3p o -5p, tal como por ejemplo miR-10a-5p. Cuando se conocen niveles de expresión relativos, un asterisco tras el nombre indica un miARN expresado a bajos niveles con respecto al miARN en el brazo opuesto de una horquilla.

Como se usa en el presente documento de manera intercambiable, un “producto génico del miARN”, “miARN”, “microARN” o “miR” se refiere al ARN procesado o sin procesar transcrito a partir de un gen miR. Como los productos génicos del miR no se traducen en proteínas, la expresión “productos génicos del miR” no incluye proteínas. Además, la expresión “niveles de expresión del miARN” se refiere a los niveles de miARN, sin incluir por lo tanto proteínas.

La recurrencia del tumor, o la recurrencia del cáncer de mama, puede definirse como la reaparición del tumor (o del cáncer) después del tratamiento. La recurrencia puede ser local (en el mismo seno o en la cicatriz de la mastectomía) o en un área distante. En el contexto de la presente invención, la recurrencia es preferentemente en un área distante (metástasis).

Los presentes inventores han investigado la asociación entre el perfil del miARN y la recurrencia del cáncer de mama en pacientes que padecen o han padecido cáncer de mama, preferiblemente en pacientes que han sido tratados para combatir dicho cáncer de mama. En una realización particular, los pacientes se han sometido a cirugía para extirpar el tumor. En una realización preferida, el tumor que se extirpa durante la cirugía es un tumor primario.

En un primer aspecto, la invención proporciona un procedimiento para estimar o pronosticar el riesgo de recurrencia del cáncer de mama. Dicho procedimiento comprende la detección en una muestra procedente de un sujeto humano, de los niveles de uno o más miARN

particulares. Ejemplos ilustrativos no limitantes de dichas muestras incluyen diferentes tipos de muestras de tejidos, así como de fluidos biológicos, tales como sangre, suero, plasma, fluido cerebroespinal, fluido peritoneal, heces. Preferiblemente, dichas muestras son muestras de tejidos y lo más preferible, dichas muestras de tejidos se originan de tejido del tumor del individuo cuya recurrencia se va a predecir, y se pueden originar a partir de biopsias, preferiblemente provenientes de resección quirúrgica.

Los niveles de uno o más miARN particulares que se miden en el procedimiento de la presente invención se seleccionan preferiblemente de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p).

Tabla 1: Secuencias de miARN

MicroARN	SEQ ID NO	Número de acceso (miRBase)	Secuencia Madura
hsa-miR-149-5p	1	MIMAT0000450	ucuggcuccgugucuucacuccc
hsa-miR-10a-5p	2	MIMAT0000253	uacccuguagauccgaauuugug
hsa-miR-20b-5p	3	MIMAT0001413	caaagugcucauagugcagguag
hsa-miR-30a-3p	4	MIMAT0000088	cuuucagucggauguuugcagc
hsa-miR-342-5p	5	MIMAT0004694	aggggugcuaucugugauuga

Los inventores han proporcionado evidencias de que estos miARN son indicadores del riesgo de recurrencia del tumor en un sujeto (véase por ejemplo la Figura 2).

Los niveles de estos miARNs particulares se comparan con los niveles de esos mismos miARNs en una muestra de referencia o con un valor control. En el contexto de la presente invención, se entiende “muestra de referencia” como la muestra de referencia que se usa para determinar la variación de los niveles de los miARNs de la presente invención. En una realización, el valor de referencia se obtiene a partir de la señal proporcionada usando una muestra de tejido obtenida de un individuo que no ha presentado recurrencia. Preferiblemente, las muestras se toman del mismo tejido de varios individuos que no han presentado recurrencia y se combinan, de tal manera que el valor de referencia refleja el valor promedio de dichas moléculas en la población sujetos que no presentan recurrencia. El “valor de referencia” puede definirse como el nivel de un miARN de la invención en la muestra de referencia. En una realización, el valor de referencia indicativo de no recurrencia

para cada miARN concreto debe conocerse antes de llevar a cabo el procedimiento de la presente invención.

En otra realización, el valor de referencia se obtiene a partir de la señal proporcionada por un ácido nucleico cuyos niveles se mantienen estables independientemente de factores variables tales como la edad, el tipo de célula, la enfermedad (por ejemplo recurrencia-no recurrencia del tumor), el sexo, el estado fisiológico, o la respuesta a condiciones externas, o similares.

Por ejemplo, el valor de referencia puede obtenerse a partir de la señal proporcionada por el RNU6B (SEQ ID No: 6), y/o RNU48 (SEQ ID No: 7), y/o miR-16-5p (SEQ ID No: 8). Preferentemente, el valor de referencia puede obtenerse a partir de la señal proporcionada por RNU48, y/o miR-16-5p. La Figura 1 muestra la estabilidad de los niveles de expresión de RNU48 y miR-16-5p en muestras de tumores recurrentes (B) y no recurrentes (A). Esta gráfica únicamente muestra los valores de expresión (tras normalización RMA) de RNU48 y miR16, demostrando que su expresión es estable.
Se prefiere que la muestra de referencia sea del mismo tejido y/u obtenida mediante el mismo procedimiento que la muestra a partir de la cual se va a predecir la el riesgo de recurrencia del tumor en el sujeto.

De esta manera, mediante el procedimiento de la invención, se puede predecir si paciente individual posee (i) un elevado riesgo de recurrencia del tumor o (ii) un bajo riesgo de recurrencia del tumor. Los inventores han demostrado que, cuando el riesgo de recurrencia es elevado, los niveles de expresión de uno o más miARN, preferentemente de todos los miARNs tal y como se definen en las SEQ ID NOs: 1, 2, 3, 4 y 5 han disminuido en comparación con un valor de referencia.

La cantidad de miARN diana, normalizada a una referencia endógena y con respecto a una muestra del control es un indicativo del riesgo de recurrencia del tumor. Por ejemplo: si el nivel de expresión de uno o varios (preferiblemente todos) los miARNs de la presente invención (la firma de los 5-miARNs) en una muestra de un tumor (preferiblemente primario) de un sujeto es menor que el nivel de expresión de uno (o varios) ácido(s) nucleico(s) control de ese mismo sujeto (en ese mismo tumor o tejidos adyacentes, o tejidos alejados del tumor), entonces el riesgo de recurrencia tal y como se define en la presente invención es alto. Se considera que los niveles de expresión de uno o varios (preferiblemente todos) miARNs de la presente invención (la firma de los 5-miARNs) en una muestra de un tumor

(preferiblemente primario) son menores que el nivel de expresión de uno (o varios) ácido(s) nucleico(s) control de ese mismo sujeto si la relación entre el valor que los niveles de expresión normalizados de los uno o varios miARNs y los niveles de expresión normalizados de uno (o varios) ácido(s) nucleico(s) control es menor que 1, preferiblemente menor que 0,95, más preferiblemente menor que 0,9, como por ejemplo menor que 0,85, como por ejemplo menor que 0,75, como por ejemplo menor que 0,7, como por ejemplo menor que 0,5, como por ejemplo menor que 0,3, como por ejemplo 0,1, como por ejemplo menor que 0,05, como por ejemplo menor que 0,01.

Por ejemplo si la relación entre los niveles de expresión de uno o varios (preferiblemente todos) miARNs de la presente invención (la firma de los 5-miARNs) en una muestra de un tumor (Rx) y los niveles de expresión normalizados de uno (o varios) ácido(s) nucleico(s) control (Rc) es menor que 1 (por ejemplo 0,9, o 0,85, o 0,75, o 0,7, o 0,65, o 0,6, o 0,5, o 0,4, o 0,3, o 0,2, o 0,1, o 0,05, 0,01, etc.), el riesgo de recurrencia tal y como se define en la presente invención es alto.

Riesgo de recurrencia alto: $Rx/Rc \leq 1$

El <u>riesgo de recurrencia</u> según la presente invención se puede definir como la probabilidad de que el tumor reaparezca al cabo de un determinado periodo de tiempo tras un tratamiento antitumoral, preferiblemente tras la extirpación quirúrgica del tumor, preferiblemente un tumor primario. Por ejemplo, el tumor puede reaparecer de 3 meses a 10 años después del tratamiento antitumoral, preferiblemente de 3 meses a 5 años después del tratamiento antitumoral, más preferiblemente de 1 año a 4 años después del tratamiento antitumoral, aún más preferiblemente de 1,5 años a 2 años después del tratamiento antitumoral. Por ejemplo, el tumor puede reaparecer 3 meses después del tratamiento antitumoral. Por ejemplo, el tumor puede reaparecer 6 meses después del tratamiento antitumoral. Por ejemplo, el tumor puede reaparecer 1 año después del tratamiento antitumoral. Por ejemplo, el tumor puede reaparecer 2 años después del tratamiento antitumoral. Por ejemplo, el tumor puede reaparecer de 0 a 2 años después del tratamiento antitumoral. Por ejemplo, el tumor puede reaparecer 5 años después del tratamiento antitumoral. Preferiblemente, el riesgo de recurrencia según la presente invención se puede definir como la probabilidad de que el tumor reaparezca al cabo de 1 a 2 años después del tratamiento antitumoral (preferentemente cirugía). El riesgo de recurrencia definido como la probabilidad de que el tumor reaparezca al cabo de 0 a 2 años después del tratamiento antitumoral (preferentemente cirugía) puede referirse también a lo largo de la presente

memoria como "*riesgo de recurrencia temprana*".

Por lo tanto, preferiblemente, el riesgo de recurrencia de acuerdo con la presente invención se refiere al riesgo aparición de metástasis en un periodo no superior a dos años después del tratamiento antitumoral (preferentemente cirugía).

Un elevado (o, indistintamente, alto) riesgo de recurrencia puede definirse como una probabilidad de más del 50% de que el tumor reaparezca en dicho sujeto. Preferiblemente, un elevado riesgo de recurrencia se define como una probabilidad de más del 55% de que el tumor reaparezca, como por ejemplo una probabilidad del 60% de que el tumor reaparezca. Preferiblemente, un elevado riesgo de recurrencia se define como una probabilidad de más del 63% de que el tumor reaparezca, como por ejemplo una probabilidad de más del 70%, del 75%, del 80%, del 85% o del 90%. Preferiblemente, un elevado riesgo de recurrencia se define como una probabilidad de más del 99% de que el tumor reaparezca.

Un bajo riesgo de recurrencia puede definirse como una probabilidad de menos del 50% de que el tumor reaparezca en dicho sujeto. Preferiblemente, un bajo riesgo de recurrencia se define como una probabilidad de menos del 45% de que el tumor reaparezca, como por ejemplo una probabilidad del 40% de que el tumor reaparezca. Preferiblemente, un bajo riesgo de recurrencia se define como una probabilidad de menos del 37% de que el tumor reaparezca, como por ejemplo una probabilidad de más del 30%, del 25%, del 20%, del 15%, del 10% o del 1 %. Preferiblemente, un elevado riesgo de recurrencia se define como una probabilidad de menos del 1% de que el tumor reaparezca

Por lo tanto, en una realización de la presente invención, un alto riesgo de recurrencia se puede definir como una probabilidad de más del 50%, y/o de más del 55% y/o de más del 60% y/o de más del 63% de que el tumor reaparezca, como por ejemplo una probabilidad de más del 70%, del 75%, del 80%, del 85%, del 90%, o más del 99% de que el tumor reaparezca en el sujeto en cuestión al cabo de 3 meses a 10 años después del tratamiento antitumoral, preferiblemente de 3 meses a 5 años después del tratamiento antitumoral, más preferiblemente de 1 año a 5 años después del tratamiento antitumoral, aún más preferiblemente de 1 a 2 años después del tratamiento antitumoral (preferentemente cirugía).

Por lo tanto, en otra realización de la presente invención, un alto riesgo de recurrencia se puede definir como una probabilidad de más del 50%, y/o de más del 55% y/o de más del

60% y/o de más del 63% de que el tumor reaparezca, como por ejemplo una probabilidad de más del 70%, del 75%, del 80%, del 85%, del 90%, o más del 99% de que el tumor reaparezca en el sujeto en cuestión al cabo de 1 a 2 años después del tratamiento antitumoral.

Por lo tanto, en otra realización de la presente invención, un alto riesgo de recurrencia se puede definir como una probabilidad de más del 60% de que el tumor reaparezca en el sujeto en cuestión al cabo de 1 a 2 años después de la cirugía para extirpar el tumor.

Por lo tanto, en otra realización de la presente invención, un alto riesgo de recurrencia se puede definir como una probabilidad de más del 63% de que el tumor reaparezca en el sujeto en cuestión al cabo de 1 a 2 años después de la cirugía para extirpar el tumor.

Por lo tanto, en otra realización de la presente invención, un alto riesgo de recurrencia se puede definir como una probabilidad de más del 70% de que el tumor reaparezca en el sujeto en cuestión al cabo de 1 a 2 años después de la cirugía para extirpar el tumor.

Por lo tanto, en otra realización de la presente invención, un alto riesgo de recurrencia se puede definir como una probabilidad de más del 85% de que el tumor reaparezca en el sujeto en cuestión al cabo de 1 a 2 años después de la cirugía para extirpar el tumor.

La probabilidad de recurrencia aumenta con el número de miARNs (seleccionados de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) cuyos niveles se encuentran reducidos en la muestra de tumor con respecto al control, como se ha descrito anteriormente. Por ejemplo, si un sujeto presenta niveles de un único miARN (seleccionados de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) reducidos con respecto al control, el riesgo de recurrencia puede ser de aproximadamente 55% - 65%, preferiblemente de aproximadamente 62%.

Por ejemplo, si un sujeto presenta niveles de al menos dos miARNs (seleccionados de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) reducidos con respecto al control, el riesgo de recurrencia puede ser de aproximadamente 60% - 65%, preferiblemente de aproximadamente 63%.

Por ejemplo, si un sujeto presenta niveles de al menos tres miARNs (seleccionados de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) reducidos con respecto al control, el riesgo de recurrencia puede ser de aproximadamente 65% - 72%, preferiblemente de aproximadamente 71%.

Por ejemplo, si un sujeto presenta niveles de al menos cuatro miARNs (seleccionados de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) reducidos con respecto al control, el riesgo de recurrencia puede ser de aproximadamente 70% - 775%, preferiblemente de aproximadamente 72%.

Por ejemplo, si un sujeto presenta niveles de los cinco miARNs (firma 5 miARNs) (seleccionados de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) reducidos con respecto al control, el riesgo de recurrencia puede ser de aproximadamente 85% - 90%, preferiblemente de aproximadamente 87%.

Se prefiere que el sujeto al cual se le va a predecir la recurrencia haya experimentado la cirugía del tumor, es decir, la resección quirúrgica del cáncer antes de la aplicación del procedimiento de la invención. Las opciones para la cirugía incluyen, pero no se limitan a, (a) mastectomía parcial (se elimina solo parte de la mama) b) mastectomía (se extirpa el seno por completo y conlleva la extirpación de todo el tejido del seno, y algunas veces, también los tejidos adyacentes). Por lo tanto, los niveles de al menos uno, preferiblemente dos, más preferiblemente tres, aún más preferiblemente cuatro y todavía más preferiblemente cinco miARN de la invención se determinan a partir del tumor primario obtenido por cirugía. Además, preferiblemente, los niveles de ARN que se usan como muestra de referencia, tal y como se indica anteriormente, se obtienen preferiblemente a partir de tejido del mismo paciente. Por ejemplo, se pueden obtener partir del mismo tumor o a partir de tejido sano adyacente.

No se excluye que se usen otros tratamientos tales como la radiación o la quimioterapia junto con la cirugía. En el contexto de la presente invención se entiende la quimioterapia como tratamiento del cáncer con un fármaco antineoplásico o con una combinación de dichos fármacos. Sin pretender estar vinculado a ninguna teoría particular, se entiende que

la quimioterapia actúa normalmente eliminando células que se dividen rápidamente, lo que es una de las propiedades principales de la mayor parte de células cancerosas.

Es posible que en el momento de tomar la muestra del sujeto humano, el sujeto humano esté (i) siendo tratado con quimioterapia o (ii) no esté siendo tratado con quimioterapia. Se prefiere particularmente la siguiente secuencia: (a) pacientes con cáncer de mama que experimentan la cirugía del tumor (preferentemente primario) de mama, (b) a partir de la cirugía, se extrae una muestra del tumor para el análisis del miARN, (c) (opcional) se tratan los pacientes mediante quimioterapia. La muestra del tumor es preferiblemente una muestra reciente del tumor, pero también puede ser por ejemplo una muestra no reciente del tumor, como por ejemplo una muestra de un tumor guardado en parafina.

El aislamiento de la muestra del sujeto vivo como tal no es parte de la invención. La cirugía de un cuerpo humano vivo como tal no es parte de la invención. Lo que es una parte de la invención es el procedimiento *in vitro* para determinar el nivel de expresión del miARN de la invención en una muestra que se ha obtenido mediante las etapas (a) y (b). De esta manera, el procedimiento para predecir la recurrencia del tumor de acuerdo con la presente invención es preferiblemente un procedimiento *in vitro*.

En la práctica, la etapa (b) comprende normalmente: (b1) la extracción del ARN (total) o del miARN (total) procedentes de la muestra del tumor, y los procedimientos adecuados para este fin no están particularmente limitados. Esta comprende también normalmente la etapa (b2), es decir, la detección de los niveles de los miARN de interés entre el ARN extraído en (b1). Los procedimientos a modo de ejemplo adecuados no limitantes se proporcionan en materiales y procedimientos.

El procedimiento de la presente invención se puede aplicar con muestras de individuos a cualquier edad. Por ejemplo, el procedimiento de la presente invención se puede aplicar con muestras de individuos de menos de 50 años. Por ejemplo, el procedimiento de la presente invención se puede aplicar con muestras de individuos de más de 50 años. Preferentemente, el procedimiento de la presente invención se aplica en mujeres.

En la invención, el procedimiento de determinar el resultado, es decir, el nivel de expresión del miARN, no necesita limitarse de manera particular y se puede seleccionar a partir de un procedimiento de perfilado génico, tal como una micromatriz y/o un procedimiento que comprenda la PCR, tal como la PCR en tiempo real; y/o la transferencia Northern ("*Northern*

Blot", en inglés).

Una micromatriz ("*microarray*", en inglés) es una matriz sobre un sustrato sólido (normalmente un porta de vidrio o una celda con una película fina de silicio) que evalúa grandes cantidades de material biológico, en el presente caso, una gran cantidad de miARN diferentes o, preferiblemente, sus transcritos de ADN inverso, que son detectables mediante sondas inmovilizadas específicas sobre el sustrato sólido.

Una transferencia Northern implica el uso de electroforesis para separar muestras de ARN por tamaño y la posterior detección con una sonda complementaria a (parte de) la secuencia diana del ARN de interés.

La PCR cuantitativa en tiempo real (RQ-PCR) es una técnica de cuantificación de la expresión génica sensible y reproducible que se puede usar de manera particular para la expresión del perfil del miARN en células y tejidos. Se puede usar cualquier procedimiento para la evaluación de los resultados de la RT-PCR y se puede preferir el procedimiento $\Delta\Delta C_t$. El procedimiento $\Delta\Delta C_t$ se describe en detalle en Livak *y col*. (Methods 2001, 25:402-408). (C_t = Valores umbral del ciclo). Cuando se lleva la presente invención a la práctica, se deberá usar preferiblemente el procedimiento $\Delta\Delta C_t$ tal como describen Livak *y col*. (Methods 2001, 25:402-408). El procedimiento $\Delta\Delta C_t$ implicará una "muestra del control" y una "muestra del sujeto".

La "muestra del sujeto" es una muestra procedente del sujeto que se va a analizar. Por cada muestra, se incluyen un gen diana (aquí: el miARN de interés) y un gen endógeno del control (tal como se describe a continuación) para la amplificación de la PCR a partir de alícuotas (normalmente diluciones en serie). Normalmente se usan varias réplicas de cada concentración diluida para derivar la eficacia de la amplificación. La eficacia de la amplificación de la PCR se puede definir como el porcentaje de amplificación (de 0 a 1). Durante la reacción de la qPCR, un software mide normalmente el número de ciclos de cada muestra en el cual la fluorescencia cruza una línea arbitraria (indicadora de la amplificación de la PCR), el umbral. Este punto de cruce es el valor C_t. Muestras más diluidas cruzarán a valores C_t posteriores.

En el procedimiento de la presente invención, se puede normalizar la expresión del miARN en relación con un control endógeno. El control endógeno es preferiblemente la expresión endógena (es decir, en el mismo individuo) de otro ácido nucleico, preferiblemente un ARN,

tal como uno cualquiera seleccionado entre ARNm, miARN, ARN nucleolar, ARNr y similares. Se prefiere en gran manera que dicho ARN se sepa que se expresa de manera estable. La expresión estable significa cualquier expresión que se sabe que se expresa de manera independiente de factores variables tales como la edad, el tipo de célula, la enfermedad, el sexo, el estado fisiológico, o la respuesta a condiciones externas, o similares. En una realización particular, se puede escoger este ácido nucleico entre ARN nucleolares pequeños, y es preferiblemente RNU6b y/o RNU48. Se pueden usar también para la normalización otros ARN nucleolares tales como RNU44, RNU48, RNU24, RNU43, RNU6B, RNU19, y Z30, así como el miARN de genes. RNU44 y RNU48 pueden ser particularmente útiles. Preferiblemente, se puede usar adicionalmente u opcionalmente miR-16.

Preferiblemente, para normalizar la expresión del(los) miARN(s) se emplea la expresión endógena de RNU6B, y/o RNU48, y/o miR-16-5p. Preferiblemente se emplea la expresión endógena de miR-16-5p y RNU48.

Se puede acceder a todas las secuencias de los nucleótidos de los miARN en línea en la miRBase, que es el principal depositario en línea de todas las secuencias y anotaciones del microARN (Kozormara *y col*., Nucl. Acids Res. (2011) 39 (suppl 1)). Las secuencias de nucleótidos de los otros ARN humanos, tales como los ARN nucleolares pequeños se pueden encontrar en línea en http://www.ncbi.nlm.nih.gov/ (ensamblaje GRCh37 (GCA_000001405.6) procedente del Genome Reference Consortium). Se describen moléculas adicionales adecuadas para la normalización en Davoren *y col*., BMC Mol Biol. 2008; 9: 76, y se pueden usar en el procedimiento de la invención.

RNU6B NR_002752 (SEQ ID NO.:6) secuencia:
GUGCUCGCUUCGGCAGCACAUAUACUAAAAUUGGAACGAUACAGAGAAGAUUAGCAUGGCCCCUGCGCAAGGAUGACACGCAAAUUCGUGAAGCGUUCCAUAUUUUU

RNU48 NR_002745 (SEQ ID NO.:7) secuencia:
AGUGAUGAUGACCCCAGGUAACUCUGAGUGUGUCGCUGAUGCCAUCACCGCAGCGCUCUGACC

miR-16-5p MIMAT0000069 (SEQ ID NO.:8) secuencia:
UAGCAGCACGUAAAUAUUGGCG

Es también posible que se use más de un ARN diferente para la normalización, tal como dos o más, tres o más y/o hasta cinco ARN diferentes. Preferiblemente, expresión del miARN se normaliza la en relación con la expresión endógena (es decir, en el mismo individuo) de Preferiblemente se emplea la expresión endógena de miR-16-5p y RNU48.

En cualquier caso, se obtiene el valor de referencia por normalización con respecto al mismo ácido nucleico que el valor del cual se va a predecir la respuesta del paciente. Únicamente a fines de ilustración: si se obtiene el valor de referencia por normalización de la expresión de un(os) miARN concreto(s) de la invención para la expresión de miR-16-5p y/o RNU48, entonces, en el paciente del cual se va a predecir la respuesta, se obtiene también el valor de la expresión por normalización de la expresión de un(os) miARN concreto(s) de la invención para la expresión de miR-16-5p y/o RNU48.

Para cuantificar la expresión génica del miARN, se puede dividir el C_t de un ácido nucleico procedente del miARN del gen de interés por el C_t del ácido nucleico procedente del control endógeno en la misma muestra para normalizar la variación en la cantidad y calidad del ARN entre diferentes muestras y obtener la expresión relativa (con respecto al control endógeno) de cada una de la "muestra del sujeto" y de la "muestra control". Opcionalmente, esto se lleva a cabo por duplicado, triplicado, cuadriplicado y de manera similar, respectivamente. Opcionalmente, se puede obtener de manera adecuada un valor ΔC_t del control calculando el promedio de los valores ΔC_t obtenidos a partir de muestras de un grupo del control de varios individuos con los cuales se van a comparar los valores de la "muestra del sujeto".

Preferiblemente, para cuantificar la expresión génica del miARN, el valor de expresión normalizado de uno o más miARNs de la presente invención se calcula de la siguiente manera. Este modo de cálculo aplica también a casos en los que los ácidos nucleicos normalizadores son diferentes, o cuando existe un único o varios ácido(s) nucleico(s) normalizador(es).

En primer lugar, es necesario calcular la eficiencia de amplificación (e) y el ciclo umbral (Ct) para cada muestra y cada miR (incluyendo normalizadores). Posteriormente, definimos la eficiencia media (E) de un miR x como el valor promedio de todas las eficiencias para ese miR.

$$E=(e1+e2+e3....en)/n,$$

siendo n el número de pacientes.

El valor de expresión (R) de un miR dado (x) para una muestra (n) se obtiene de la siguiente fórmula:

Rxn= (1+Ex)^-Ctx,

siendo Ctx el ciclo umbral del miR x para la muestra n. Por ejemplo para el ácido nucleico normalizador RNU6b:

R6bn= (1+E6b)^-Ct6b,

siendo E6b la eficiencia de amplificación de RNU6b y Ct6b su ciclo umbral para la muestra n.

Por ejemplo, para el para el ácido nucleico normalizador miR-16, sería:

R16n= (1+E16)^-Ct16,

siendo E16 la eficiencia de amplificación del miR x y Ct16 su ciclo umbral para la muestra n.

Finalmente se calculael valor de expresión normalizado para el miR x en la muestra n de la siguiente manera:

Rn=Rxn/(Rn1+Rn2)

siendo el denominador (Rn1+Rn2 en este caso) la suma de los valores de expresión (R) de uno o más nucleótidos control ((R6bn+R16n) en este caso).

Este valor de expresión normalizado es el que preferentemente se utiliza. Los cálculos para los valores de eficiencia y ciclo umbral para cada muestra se realizan preferentemente según se indica en Zhao and Fernald. J. Comput. Biol. 2005 Oct;12(8):1045-62.

Como cebadores ("*primers*") para la amplificación de la PCR se puede usar cualquier secuencia que el experto en la materia sabe identificar. Los cebadores poseen una parte complementaria a la secuencia de cada miRNA que quiere amplificarse. Por ejemplo, pueden emplearse sondas (cebadores) disponibles comercialmente, como por ejemplo sondas TaqMan® de Applied Biosystems. Las sondas pueden estar marcadas (por ejemplo con un fluróforo) para facilitar la detección.

Sólo a modo de ejemplo se detallan a continuación secuencias que pueden emplearse como

cebadores para la amplificación de los miARN de la invención por RT-PCR:

Para amplificar el miR-149-5p: 5-GGGAGTGAAGACACGGAGCCAGA-3 (SEQ ID NO.: 9)
Para amplificar el miR-10a-5p: 5-CACAAATTCGGATCTACAGGGTA-3 (SEQ ID NO.: 10)
Para amplificar el miR-20b-5p: 5-CTACCTGCACTATGAGCACTTTG-3 (SEQ ID NO.: 11)
Para amplificar el miR-30a-3p: 5-GCTGCAAACATCCGACTGAAAG-3 (SEQ ID NO.: 12)
Para amplificar el miR-342-5p: 5-TCAATCACAGATAGCACCCCT-3 (SEQ ID NO.: 13)
Para amplificar el RNU6B: 5-AAAAATATGGAACGCTTCACG-3 (SEQ ID NO.: 14)
Para amplificar el RNU48: 5-GGTCAGAGCGCTGCGGTGATG-3 (SEQ ID NO.: 15)
Para amplificar el miR-16-5p: 5-CGCCAATATTTACGTGCTGCTA-3 (SEQ ID NO.: 16)

En el procedimiento de la invención, se pueden usar uno o más miARN como control negativo. Para los miARN del control negativo, se prefiere principalmente un origen no humano.

Es también posible aislar los miARN de la invención mediante cualquier procedimiento conocido adecuado a dichos fines.

En la presente invención, independientemente del método que se use para la cuantificación de los miARNs (o para la cuantificación de las muestras control) la "muestra del sujeto" se compara con la "muestra del control" (o el promedio de las muestras del control, tal y como se ha definido anteriormente). El valor de la relación de los valores de expresión de ambas muestras (muestra de sujeto/muestra control) indica el riesgo de recurrencia, tal y como se ha descrito anteriormente.

El procedimiento de la presente invención puede conllevar además una etapa de proporcionar un resultado en forma de datos y/o información y guardar ese resultado en un soporte adecuado para contener o guardar datos y/o información.

Procedimiento para clasificar un sujeto humano que padece cáncer de mama en uno de dos grupos

En un segundo aspecto, la invención proporciona también un procedimiento para clasificar un sujeto humano que padece cáncer de mama en uno de dos grupos. A no ser que se especifique de manera implícita o explícita de otra forma, los detalles de la invención tal como se ha detallado anteriormente se aplican también al segundo aspecto de la invención.

El grupo 1 comprende los sujetos identificables mediante el procedimiento de la invención tal como se ha detallado anteriormente que poseen un alto riesgo de sufrir una recurrencia del tumor; y en el que el grupo 2 representa el resto de sujetos.

El procedimiento según este segundo aspecto de la invención puede conllevar además una etapa de proporcionar un resultado en forma de datos y/o información y guardar ese resultado en un soporte adecuado para contener o guardar datos y/o información.

Procedimiento para pronosticar la supervivencia de un sujeto con cáncer de mama

Un tercer aspecto de la presente invención proporciona un procedimiento para pronosticar la supervivencia de un sujeto que padece o ha padecido cáncer de mama.

La supervivencia de pacientes de cáncer se expresa de manera adecuada generalmente mediante las curvas de Kaplan-Meier, que han recibido su denominación de Edward L. Kaplan y Paul Meier que fueron los que describieron éstas en primer lugar (Kaplan, Meier: Amer. Statist. Assn. 53:457–481). El estimador Kaplan–Meier se conoce también como el estimador límite del producto. Sirve para estimar la función de supervivencia a partir de los datos del tiempo de vida. Una representación gráfica de la estimación de Kaplan–Meier de la función de supervivencia es una serie de etapas horizontales de magnitud decreciente que, cuando se toma una muestra suficientemente grande, se aproxima a la función de supervivencia verdadera de esta población. El valor de la función de supervivencia entre sucesivas observaciones muestreadas distintas se supone que es constante. La significancia de las diferencias estadísticas puede estimarse empleando por ejemplo, un tipo de procedimientos de prueba de los rangos para datos de supervivencia censurados (log-rank, pruebas Tarone-Ware y Peto-Peto).

En la estadística en la que se basa la presente invención, los pacientes de cáncer de mama se agrupan en categorías, por ejemplo, aquellos con un perfil de expresión concreto tal como el perfil de expresión de al menos un miARN particular (y preferiblemente los cinco miARNs de la presente invención) y aquellos sin dicho perfil. Con respecto a la presente invención, se puede usar el estimador de Kaplan-Meier para medir la fracción de pacientes vivos sin recurrencia del tumor (*relapse-free survival,* definida como el periodo de tiempo transcurrido desde la cirugía hasta la recurrencia del tumor) a lo largo de un espacio de tiempo tras ser sometidos a cirugía para extirpar el tumor (*relapse-free survival,* definida como el periodo de tiempo transcurrido desde la cirugía hasta la recurrencia del tumor).

Un cuarto aspecto de la presente invención proporciona un kit que comprende al menos un(os) oligonucleótido(s) capaz(ces) de hibridarse con uno cualquiera de dos o más, y preferiblemente todos los miARN (o con sus respectivos cDNAs, en el caso de su amplificación por RT-qPCR). La persona experta en la materia puede diseñar estos oligonucleótidos. A modo de ejemplo, se incluyen unas secuencias de oligonucleótidos que pueden usarse para la amplificación de los miRNAs de la presente invención (que se definen como las SEQ ID NOs: 1 a 5). En la Tabla 2 se proporcionan las realizaciones particulares de las secuencias de dichos oligonucleótidos, descritos anteriormente.

Siempre que se refiera a la hibridación del(de los) oligonucleótido(s), se prefiere que dicho(s) oligonucleótido(s) sea(n) capaz(ces) de hacerlo en condiciones rigurosas.

El rigor es un término usado en los experimentos de hibridación. El rigor refleja el grado de complementariedad entre el oligonucleótido y el ácido nucleico (que es en este caso el del miARN que se va a detectar), el de mayor rigor, el mayor porcentaje de identidad entre la sonda y el ácido nucleico que se une al filtro. La persona experta en la materia sabe bien que la temperatura y las concentraciones salinas tienen un efecto directo en los resultados que se obtienen. Se reconoce que los resultados de la hibridación están relacionados con el número de grados por debajo de la T_f (temperatura de fusión) del ADN a la cual el experimento se lleva a cabo. A menudo, se definen las condiciones de rigor como un lavado con 0,1X SSC (tampón de solución salina-citrato de sodio (SSC) a 65°C. (SSC se proporciona normalmente como una disolución madre 20X, que consta de cloruro de sodio 3 M y citrato trisódico 300 mM (ajustado a pH 7,0 con HCl)).

El kit se basa en el poder predictivo del procedimiento de la presente invención.

Tal como se ha mencionado anteriormente, el valor indicador de referencia para el bajo riesgo de recurrencia de cada miARN concreto se puede determinar antes de llevar a cabo el procedimiento de la presente invención, y también se puede determinar durante el procedimiento de la presente invención, calculando los niveles de expresión de uno o más ácido(s) nucleico(s) cuyos niveles se mantienen estables independientemente de factores variables tales como la edad, el tipo de célula, la enfermedad (por ejemplo recurrencia-no recurrencia del tumor), el sexo, el estado fisiológico, o la respuesta a condiciones externas, o similares y comparando los niveles de expresión de uno o más miARN de la firma de los 5

miARN con los niveles de expresión de la muestra control o valor de referencia. Si los niveles de expresión de uno o más miARN de la firma de los 5 miARN son inferiores a los niveles de la muestra control o valor de referencia, entonces, el riesgo de recurrencia tal y como se describe en la presente invención es alto.

En el caso particular del kit, alternativamente, se pueden proporcionar con el kit el valor indicador de referencia para el bajo riesgo de recurrencia (y/o un valor de referencia indicador para el alto riesgo de recurrencia).

Tabla 2: Ejemplos de secuencias de oligonucleótidos

Secuencia del oligonucleótido	SEQ ID NO	Para la amplificar la SEQ ID NO:
5-GGGAGTGAAG ACACGGAGCC AGA-3	9	1
5-CACAAATTCGGATCTACAGGGTA-3	10	2
5-CTACCTGCACTATGAGCACTTTG-3	11	3
5-GCTGCAAACATCCGACTGAAAG-3	12	4
5- TCAATCACAGATAGCACCCCT-3	13	5
5-AAAAATATGGAACGCTTCACG-3	14	6
5-GGTCAGAGCGCTGCGGTGATG-3	15	7
5- CGCCAATATTTACGTGCTGCTA-3	16	8

En realizaciones particulares, el kit se selecciona entre (a) un kit adecuado para la PCR, (b) un kit adecuado para la transferencia Northern (*Northern Blot*) y (c) un kit adecuado para los análisis de micromatriz (*microarray*). Se pueden combinar cualquiera de dos o más de estas realizaciones, de tal manera que el kit puede comprender, por ejemplo, ambos de (a) y (c).

En el caso de (a) un kit adecuado para la PCR, esta PCR es normalmente la PCR cuantitativa en tiempo real (RQ-PCR o RT-qPCR), una técnica de cuantificación de la expresión génica sensible y reproducible. En este caso se desea que el kit comprenda adicionalmente un cebador del oligonucleótido poliT además del (de los) oligonucleótido(s) del kit (ver Tabla 2 para ejemplo de dichos cebadores). El cebador del oligonucleótido poliT se puede usar junto con el(los) oligonucleótido(s) de la invención para el cebado de la PCR, tras la poliadenilación de los miARN aislados mediante los procedimientos conocidos por la persona experta, tales como usando la polimerasa poli(A) y ATP. Estos reactivos pueden

estar comprendidos opcionalmente en el kit.

Una Transferencia Northern implica el uso de electroforesis para separar las muestras de ARN por tamaño y la posterior detección con un(os) oligonucleótido(s) (sonda de hibridación) complementaria con (parte de) la secuencia diana del ARN de interés.

Es también posible que el(los) oligonucleótido(s) estén inmovilizados en manchas sobre una superficie (preferiblemente sólida). En una de sus realizaciones, el kit comprende una micromatriz. Una micromatriz de ARN es una matriz sobre un sustrato sólido (normalmente un porta de vidrio o una celda de una película fina de silicio) que evalúa grandes cantidades de diferentes ARN (aquí miARN) que son detectables mediante sondas específicas inmovilizadas sobre manchas sobre un sustrato sólido. Cada mancha contiene una secuencia específica de ácido nucleico, normalmente una secuencia de ADN, como sondas (o indicadores). Aunque el número de manchas no está limitado de manera alguna, existe una realización preferida en la que la micromatriz se personaliza para los procedimientos de la invención. En una realización, dicha matriz personalizada comprende cincuenta manchas o menos, tal como treinta manchas o menos, incluyendo veinte manchas o menos.

Preferiblemente, el kit de la invención es un kit adecuado para la PCR, preferiblemente para la PCR cuantitativa en tiempo real (RT-qPCR).

El kit de la presente invención puede incluir soporte adecuado para contener o guardar los datos y/o información generados con la utilización del kit de la invención.

El kit se puede usar y el uso no está particularmente limitado, aunque se prefiere el uso en el procedimiento de la invención en cualquiera de sus realizaciones. Preferentemente, la invención comprende el uso del kit de la invención en la predicción del riesgo de recurrencia del cáncer de mama en sujetos que padecen/han padecido cáncer de mama, y/o en la clasificación de un sujeto en uno de dos grupos, tal y como se describe en las realizaciones de la presente invención.

Método de tratamiento y de prevención de la recurrencia del cáncer de mama empleando los miARNs de la invención

Un quinto aspecto de la presente invención proporciona un método para tratar o prevenir la recurrencia del cáncer de mama en un sujeto con cáncer de mama.

El método de tratamiento del cáncer de mama comprende administrar al sujeto que padece cáncer de mama al menos un miARN seleccionado preferiblemente de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p).

El método de tratamiento del cáncer de mama comprende administrar al sujeto que padece cáncer de mama una composición farmacéutica que comprenda al menos un miARN seleccionado preferiblemente de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p).

Cuando se contemplan aplicaciones clínicas, las composiciones farmacéuticas se pueden preparar en una forma adecuada para la aplicación prevista. Generalmente, esto implicará la preparación de composiciones que están esencialmente libres de pirógenos, así como otras impurezas que podrían ser perjudiciales para los seres humanos. Los sistemas de dispersión coloidal, tales como complejos de macromoléculas, nanocápsulas, microesferas, perlas y sistemas basados en lípidos, incluyendo emulsiones de aceite en agua, micelas, micelas mixtas y liposomas, se pueden usar como vehículos de administración de los microARN. La preparación y uso de tales sistemas son bien conocidos en la técnica.

Generalmente se emplearán sales y tampones apropiados para que los vectores de transporte sean estables y accedan a las células diana (las células del tumor). Las composiciones de la presente invención comprenden una cantidad eficaz del vector de transporte (que incluye el miARN, o el miARN desnudo), disuelto o dispersado en un transportador farmacéuticamente aceptable o en un medio acuoso. Los términos "farmacéuticamente aceptables" o "farmacológicamente aceptables" se refieren a entidades moleculares y composiciones que no producen reacciones adversas, alérgicas, etc. cuando son administradas al ser humano. Tal como se utiliza en la presente memoria, "vehículo farmacéuticamente aceptable" incluye disolventes, tampones, soluciones, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, y similares aceptables para uso en productos farmacéuticos, tales como productos farmacéuticos adecuados para la administración a seres humanos. El uso de tales medios y agentes para sustancias farmacéuticamente activas se conoce bien en la técnica. Excepto en la medida en que cualquier medio o agente convencional sea incompatible con los ingredientes activos de la presente invención (los miARNs), se contempla su uso en las composiciones terapéuticas.

También pueden incorporarse en las composiciones de la presente invención ingredientes activos suplementarios, siempre que no inactivan los ácidos nucleicos (miARN). Las composiciones activas de la presente invención pueden incluir preparaciones farmacéuticas clásicas.

La administración de estas composiciones de acuerdo con la presente invención puede ser a través de cualquier ruta común siempre que el tejido diana esté disponible a través de esa ruta. Esto incluye administración oral, nasal, o bucal. Alternativamente, la administración puede ser mediante inyección intradérmica, transdérmica, subcutánea, intramuscular, intraperitoneal o intravenosa, o mediante inyección directa en el tumor de mama. Tales composiciones se administrarían normalmente como composiciones farmacéuticamente aceptables, como se ha descrito anteriormente. Los compuestos activos también se pueden administrar parenteralmente o intraperitonealmente.

La cantidad de principio activo (miARN(s)) que se administra al sujeto en cuestión se puede determinar empíricamente

La presente invención proporciona por lo tanto al menos un miARN seleccionado preferiblemente de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) para su uso en un método de tratamiento para el cáncer de mama o, alternativamente, para su uso en un método de prevención de la recurrencia del cáncer de mama.

Por lo tanto, la presente invención proporciona el uso al menos un miARN seleccionado preferiblemente de entre los miARNs siguientes: SEQ ID NO: 1 (miR-149-5p); SEQ ID NO: 2 (miR-10a-5p); SEQ ID NO: 3 (miR-20b-5p); SEQ ID NO: 4 (miR-30a-3p); SEQ ID NO: 5 (miR-342-5p) en la elaboración de un medicamento para tratar el cáncer de mama o, alternativamente, para prevenir la recurrencia del cáncer de mama.

Ejemplos

Ejemplo 1: Materiales y métodos

Muestras de pacientes

Las 75 pacientes con cáncer de mama incluidas en el estudio se sometieron a cirugía para extirpar el tumor primario en el Hospital Universitario Virgen de la Victoria (HUVV, Málaga, España) en algún momento durante el período 1998-2005. Todas las pacientes dieron su consentimiento informado por escrito para su inclusión en el estudio, que fue aprobado por el Comité Ético de Investigación Clínica (CEIC) del hospital. Los pacientes fueron tratados de manera uniforme y seguidos de acuerdo con los protocolos establecidos por el Departamento de Oncología Clínica, basada en evidencia científica y las recomendaciones internacionales. Todas las investigaciones clínicas se llevaron a cabo de acuerdo con los principios expresados en la Declaración de Helsinki. Ninguno de los pacientes recibió terapia neoadyuvante. Los datos clínico-patológicos y la información de seguimiento se obtuvieron para cada paciente mediante revisión de sus historias clínicas.

Inmunohistoquímica

Los tumores (n = 75), fijados en formalina e incluidos en parafina (en inglés, formalin-fixed and paraffin-embedded, FFPE), se obtuvieron del archivo del hospital. Las zonas más representativas de cada tumor fueron seleccionadas por un patólogo y se construyeron microarrays de tejidos en los que las muestras se depositaron por triplicado (cortes circulares de 0,6 mm de diámetro). Los tumores se clasificaron previamente mediante inmunohistoquímica, según los denominados subtipos intrínsecos (luminal A, luminal B, basal-like y Her-2+). Para ello, se utilizaron anticuerpos específicos contra el receptor de estrógeno (ER, clona SP1), receptor de progesterona (PR, clona Y85), Ki-67 (clona SP6) receptor del factor de crecimiento epidérmico 1 (EGFR1, clona EP38Y), factor de crecimiento endotelial vascular (VEGF, clona EP1176Y) y citoqueratina 5/6 (CK5/6, clona D5/16B4). Los anticuerpos mencionados fueron todos adquiridos a través del proveedor Maestro Diagnostica (España). El nivel de expresión de HER2 se determinó también mediante inmunotinción con el ensayo comercial HercepTestTM (Dako, Dinamarca). La interpretación de los datos de inmunohistoquímica se según los criterios definidos previamente [Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, et al. (2004) Immunohistochemical and clinical characterization of the basal-like subtype of invasive

breast carcinoma. Clin Cancer Res 10: 5367-5374; Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, et al. (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 25: 118-145; Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, et al. (2008) Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. Clin Cancer Res 14: 1368-1376] por dos patólogos que desconocían las características clínico-patológicas y la evolución clínica de cada paciente.

Extracción de ARN e hibridación de microarrays

A partir de los bloques de parafina (FFPE), se seleccionaron áreas tumorales con más de un 90% de células tumorales. Las células se microdiseccionaron manualmente a partir de 3 cortes de 10 micras por cada tumor. El ARN total fue extraído utilizando el sistema de aislamiento de ácidos nucleicos comercial: RecoverAll Total Nucleic Acid Isolation kit (Life Technologies, Grand Island, NY, EE.UU.). El ARN se convirtió en ADN complementario (ADNc) mediante una reacción in vitro de retrotranscripción y el ADNc se hibridó sobre una micromatriz Affymetrix miARN chip 2.0 (Affymetrix, Santa Clara, CA, EE.UU.). Tanto la hibridación y la detección de la señal de hibridación se realizaron en el centro de Genómica Funcional (Instituto de Investigación Biomédica, el IRB Barcelona, España), como servicio externo y siguiendo las recomendaciones del fabricante Affymetrix.

Análisis de datos de microarrays

Todos los análisis estadísticos se realizaron utilizando el entorno de programación R de código abierto, junto con el paquete de algoritmos Bioconductor [Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80]. En primer lugar, se comprobó la calidad de los datos de lectura de los chips de microARN y, posteriormente, la expresión de cada miARN se normalizó en relación a moléculas de miARN control. Aunque los miARN chip 2.0 (Affymetrix) contienen sondas representativas de 131 organismos diferentes, para nuestro análisis sólo se tuvieron en cuenta los datos correspondientes a las sondas humanas. Para aumentar la potencia estadística de los análisis, se eliminaron aquellos miARNs cuya variabilidad de expresión estaba por debajo de un umbral de la desviación estándar del 66%. La expresión diferencial de los datos normalizados se evaluó mediante dos paquetes R: limma, que es un test estadístico moderado basado en un

enfoque empírico de Bayes [Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article 3] y RankProd, un método estadístico no paramétrico sencillo basado en un ordenamiento según los valores de cambio de expresión (fold change) de los miARNs [Breitling R, Armengaud P, Amtmann A, Herzyk P (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 573: 83-92]. El efecto multi-prueba fue corregido mediante el ajuste de los valores de p (p-valor) por el método de Benjamini y Hochberg. En nuestro estudio consideramos que la expresión de un miARN estaba alterada significativamente siempre que se cumplieran los dos criterios siguientes: i) p-valor ajustado ≤ 0,05 y ii) fold change ≥ 2. Los mejores miARNs candidatos son los que aparecieron en ambas pruebas estadísticas. Se llevaron a cabo las siguientes comparaciones independientes: B vs A (B/A) , BC vs A (BC/A) y B vs AC (B/AC). El grupo C por sí solo no podría compararse con A o B ya que los datos de los chips de miARN correspondientes al grupo C no proporcionaron diferencias estadísticamente significativas. Los datos MIAME (minimun information about a microarray experiment) obtenidos a partir de los microarrays (formato MIAME) han sido depositados en el repositorio público ArrayExpress, del European Bioinformatics Institute (EBI, Reino Unido), con el número de acceso E- MTAB-1989 .

Validación de los miARNs candidatos mediante RT-qPCR

Diez nanogramos (10 ng) de ARN total de cada muestra de tumor se utilizaron para obtener ADNc mediante transcripción inversa con cebadores específicos y reactivos del sistema comercial TaqMan MicroARN Reverse Transcription kit (Life Technologies, Grand Island, NY, EE.UU.). A continuación, se amplificaron los productos de PCR mediante utilizando el sistema comercial TaqMan microARN Assay, utilizando para ello el reactivo comercial TaqMan PCR Universal Master Mix. Todos los ensayos se realizaron por triplicado de acuerdo con las instrucciones del fabricante. La expresión relativa de los miARNs se calculó utilizando el método ΔΔCt. Los ARNs de pequeño tamaño (sARNs) RNU6b y miR-16 se utilizaron como referencia para la normalización.

Análisis de supervivencia

El programa de análisis de supervivencia en R se utilizó para calcular las estimaciones de supervivencia y llevar a cabo el análisis de regresión multivariante [Lumley T (2007) The survival package. R help guide]. Se analizaron las variables clínico-patológicas

y de expresión de miARN y se consideró la supervivencia acumulada libre de recaída (relapse-free survival, RFS), definida como el tiempo desde la cirugía hasta la recurrencia, como criterio de valoración clínica en los resultados de supervivencia. Los pacientes sin recaída o perdidos durante el seguimiento fueron censurados para el análisis. El cálculo de la tasa de supervivencia actuarial se realizó por el método de Kaplan-Meier y la significación de las diferencias estadísticas se evaluó mediante los procedimientos log-rank, pruebas Tarone-Ware y Peto-Peto. Se utilizó un modelo de regresión de riesgos proporcionales de Cox [Cox DR, Oakes D (1984) Analysis of survival data. London ; New York: Chapman and Hall. viii, 201 p. p.] para examinar las relaciones entre RFS y los factores pronósticos. Todas las posibles combinaciones de covariables fueron probados para identificar el mejor modelo de acuerdo con el valor AIC (Akaike Information Criterium) [Sakamoto Y, Ishiguro M, Kitagawa G (1986) Akaike information criterion statistics. Tokyo Dordrecht; Boston Hingham, MA: KTK Scientific Publishers; D. Reidel; Sold and distributed in the U.S.A. and Canada by Kluwer Academic Publishers. xix, 290 p. p], que es una medida de la calidad relativa de un modelo estadístico. El supuesto de proporcionalidad de riesgo de los modelos de Cox se comprobó mediante la prueba de regresión lineal generalizada de los residuos de Schoenfeld escalados en función del tiempo [Sakamoto Y, Ishiguro M, Kitagawa G (1986) Akaike information criterion statistics. Tokyo Dordrecht; Boston Hingham, MA: KTK Scientific Publishers; D. Reidel; Sold and distributed in the U.S.A. and Canada by Kluwer Academic Publishers. xix, 290 p. p]. Una pendiente diferente de cero indica una violación de la hipótesis de riesgo proporcional y, por tanto, su exclusión del análisis.

Modelo de predicción de la firma 5-miARN.

Tres son los pasos involucrados en la estimación de los perfiles de expresión para generar un predictor adecuado a futuras observaciones: el selección del modelo, evaluación de la predicción y selección de la función [Molinaro AM, Simon R, Pfeiffer RM (2005) Prediction error estimation: a comparison of resampling methods. Bioinformatics 21: 3301-3307]. El clasificador Naïve Bayes se utilizó en este trabajo para predecir la clase de observaciones futuras. Se trata de un modelo estándar basado en el teorema de Bayes sin supuestos específicos de dominio. Con este modelo, cada nueva muestra se clasifica en la clase más probable, según su probabilidad a posteriori y se calcula de acuerdo con el teorema de Bayes . Este clasificador se utiliza para estimar los modelos predictivos en las comparaciones entre grupos. El índice C es la medida más aceptada de la capacidad de discriminación de un modelo predictivo. En casos binarios, esta métrica es equivalente al área bajo la curva característica (Area Under the Curve, AUC) de funcionamiento del

receptor, que se utiliza comúnmente para medir la capacidad predictiva de modelos de regresión logística. Para evaluar la precisión de la predicción de los clasificadores Naïve Bayes en las curvas TPR, FPR y ROC, se utilizó una generalización de la estimación bootstrap del error de clasificación 0.632+ [Efron B, Tibshirani R (1993) An introduction to the bootstrap. New York: Chapman & Hall. xvi, 436 p. p]. El estimador bootstrap se obtiene mediante la toma de muestras de rutina de carga B de tamaño N con reemplazo. Las observaciones en las muestras de arranque se utilizan para la formación, mientras que las observaciones restantes (la muestra fuera de bolsa) se usan para la prueba. Estimaciones de rendimiento se promediaron para cada predicción sobre todas las muestras fuera de la bolsa, y la estimación de arranque de la curva de TPR, FPR y ROC se define de forma análoga al error de arranque. Los e1071 [Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80] y Daim (http://CRAN.R-project.org/package=Daim) paquetes fueron utilizados en R para realizar estos análisis. La selección de características se refiere a decidir qué miARNs para incluir en la predicción, y es un paso crucial en el desarrollo de una clase predictor . Incluyendo demasiadas características podría reducir la precisión del modelo y puede conducir a un sobreajuste de datos [Ransohoff DF (2004) Rules of evidence for cancer molecular-marker discovery and validation. Nat Rev Cancer 4: 309-314]. Para evitarlo, todas las combinaciones de miARNs se pusieron a prueba para identificar el modelo que contiene el patrón de expresión de miARN que predice con mayor precisión entre los grupos en situación de riesgo.

Predicción de genes diana de los miARNs

Las bases de datos mirTarBase (http://mirtarbase.mbc.nctu.edu.tw/) [Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, et al. (2011) miRTarBase: a database curates experimentally validated microARN-target interactions. Nucleic Acids Res 39: D163-169] y miRecords (http://mirecords.biolead.org/) [Xiao F, Zuo Z, Cai G, Kang S, Gao X, et al. (2009) miRecords: an integrated resource for microARN-target interactions. Nucleic Acids Res 37: D105-110], se utilizaron para identificar los las moléculas de ARN mensajero (mARNs) validados experimentalmente como diana de cada uno de los miARNs de la firma 5-miARN. Ambas bases de datos contienen datos validados experimentalmente de interacciones funcionales miARN/mARN diana. Para visualizar de manera gráfica las relaciones miARN/mARN, se utilizó el plugin CyTargetLinker, en el programa de visualización Cytoscape [Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome

Res 13: 2498-2504]. Los plugin ClueGo y CluePedia plugins [Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, et al. (2009) ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25: 1091-1093; Bindea G, Galon J, Mlecnik B (2013) CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics 29: 661-663] se utilizaron para recuperar la Ontología Génica (Gene Ontology, GO) asociada a los mARNs diana identificados con Cytargetlinker [Kutmon M, Kelder T, Mandaviya P, Evelo CT, Coort SL (2013) CyTargetLinker: A Cytoscape App to Integrate Regulatory Interactions in Network Analysis. PLoS One 8: e82160]. Solo se consideraron los términos GO con p-valor <0,01, corregido por el método de Benferroni.

Ejemplo 2: Análisis de microarrays.

Para identificar los miARNs asociados con recurrencia temprana y tardía, la abundancia de 1105 miARNs se analizó en una cohorte de 75 tumores primarios de mama por el la tecnología de microarrays. Los tumores se clasificaron en tres grupos de pronóstico de acuerdo con la evolución clínica de los pacientes de la siguiente manera: Grupo A, pacientes que estaban libres de enfermedad 60 meses después de la extirpación del tumor; Grupo B, pacientes que desarrollaron recurrencia temprana (24 meses después de la cirugía); y Grupo C, pacientes que desarrollaron recurrencia tardía (50-60 meses después de la cirugía). La Tabla 3 resume los datos clínicos y patológicos de la población de estudio. Excepto para el grupo C, la cohorte fue equilibrado para los 4 subtipos intrínsecos: tipo luminal A, tipo luminal B, tipo basal y HER2 +. Los dos últimos subtipos están asociados con un fenotipo más agresivo y más alto riesgo de recaída [24,26], lo cual explica por qué el grupo C (recidiva tardía) se compone principalmente de tumores luminales (Tabla 3).

Tabla 3. Características clínicas y patológicas de la población de estudio

		Grupo A*		**Grupo B***		**Grupo C***		**p-valor**
		n	**(%)**	**n**	**(%)**	**n**	**(%)**	
Número de pacientes		36	(48.0)	27	(36.0)	12	(16.0)	
Edad	≤ 50	15	(45.5)	10	(37.0)	4	(36.4)	0. 8034
	> 50	18	(54.5)	17	(63.0)	7	(63.6)	
Estado hormonal	Preperim. **	15	(41.7)	10	(37.0)	6	(50.0)	

	Postmen.**	21	(58.3)	15	(55.6)	6	(50.0)	
	Desconocido	0	(0.0)	2	(7.4)	0	(0.0)	
Tamaño del tumor (cm)	<2	5	(15.2)	2	(7.4)	3	(27.3)	0. 7487
	2-5	22	(66.7)	17	(63.0)	7	(63.6)	
	>5	5	(15.2)	4	(14.8)	1	(9.1)	
	Desconocido	1	(3.0)	4	(14.8)	0	(0.0)	
Estadío del tumor	I	3	(8.3)	4	(14.8)	1	(8.3)	
	II	19	(52.8)	8	(29.6)	8	(66.7)	
	III	14	(38.9)	15	(55.6)	3	(25.0)	
Grado histológico	1	4	(11.1)	0	(0.0)	0	(0.0)	
	2	14	(38.9)	16	(59.3)	9	(75.0)	
	3	16	(44.4)	8	(29.6)	2	(16.7)	
	Desconocido	2	(5.6)	3	(11.1)	1	(8.3)	
Subtipo histológico	Lobulilar	4	(11.1)	0	(0.0)	0	(0.0)	
	Ductal	29	(80.6)	24	(88.9)	11	(91.7)	
	Medular	0	(0.0)	1	(3.7)	0	(0.0)	
	Carcinoma	1	(2.8)	2	(7.4)	0	(0.0)	
	Mixto	2	(5.6)	0	(0.0)	1	(8.3)	
Subtipo intrínseco	Luminal A	9	(25.0)	3	(11.1)	7	(58.3)	
	Luminal B	9	(25.0)	6	(22.2)	3	(25.0)	
	Tipo Basal	9	(25.0)	9	(33.3)	1	(8.3)	
	HER2-enriched	9	(25.0)	9	(33.3)	1	(8.3)	
Tipo de cirugía	Conservativa	22	(61.1)	9	(33.3)	5	(41.7)	
	Radical	14	(38.9)	18	(66.7)	7	(58.3)	
Nódulo linfático afectado	Negativo	14	(42.4)	18	(48.6)	5	(45.5)	0.0292
	1-3	13	(39.4)	5	(13.5)	5	(45.5)	
	≥ 4	6	(12.8)	14	(37.8)	1	(9.1)	

Terapia	Quim.***	28	(77.8)	23	(85.2)	7	(58.3)	
	Horm.***	20	(55.6)	15	(55.6)	10	(83.3)	
	Rad.***	25	(69.4)	13	(48.1)	6	(50.0)	

Los microarrays se analizaron por primera vez para el control de calidad y se normalizaron para la expresión de los genes miARN (véase Ejemplo 1). Se excluyeron del estudio cuatro muestras con una calidad de la señal pobre, dejando una cohorte de 71 tumores de mama para su posterior análisis.

La agrupación jerárquica sin supervisión de los datos de los microarrays mostró que los perfiles de transcripción de los miARNs discriminan los grupos pronóstico en dos clusters (grupos) distintos (Figura 2). El cluster 1 incluye el 70% de todos los tumores del grupo A y el 26% de todos los tumores del grupo B, mientras que el cluster 2 incluye el 74% de todos los tumores del grupo B y el 30% de todos los tumores del grupo A. En general, los tumores del grupo C se distribuyeron dentro de los cluster 1 y 2, con una proporción ligeramente mayor (63%) de los tumores agrupados en el cluster 2. Debido a que el grupo C representa un estado clínico intermedio entre la no recurrencia (grupo A) y la recurrencia temprana (grupo B), la amplia distribución de los tumores del grupo C dentro de los grupos 1 y 2 podría reflejar la misma variación a nivel molecular. Una explicación alternativa podría ser que el grupo C no puede ser identificado por una firma de miARN única, ya sea debido a su naturaleza molecular intrínseca o a un menor tamaño de la muestra. Es de destacar que los tumores tienden a agruparse en función de su condición de ER ("ER+" o "ER-") y, por lo tanto, la mayor parte de los tumores luminales ("ER+") se agruparon en el cluster 1, mientras que el cluster 2 agrupó principalmente tumores HER2+ y tumores de tipo basal, que son ambos "ER-" (Figura 2). Pruebas "*multiple pairwise comparison test*" (pruebas aparejadas de comparación múltiple) mostraron que las mayores diferencias de expresión se produjeron entre tumores tipo luminal A y tumores de tipo basal. En consecuencia, la mayor lista de posibles miARNs candidatos se obtuvo al comparar tumores tipo luminal A *versus* tipo basal o tipo basal *versus* otros subtipos (Tabla 4). Estos resultados sugieren que los tres grupos de tumores (A, B y C) representan entidades biológicas distintas.

Tabla 4. miARNs más significativamente desregulados en los tumores de mama de los diferentes subtipos intrínsecos

	Luminal A vs Basal-like		**Lumina A vs HER2**		**Luminal B vs Basal-like**		**Luminal B vs HER2**		**Luminal A vs others**		**Basal-like vs others**	
ID	logFC	adj.P.Val	logFC	adj.P.Val	logFC	adj.P.Val	logFC	adj.P.Val	logFC	adj.P.Val	logFC	adj.P.Val
hsa-miR-342-3p	2,01	1,29E-07	1,89	8,31E-07	1,72	1,76E-05	1,60	0,00	1,42	1,02E-05	-1,27	1,45E-04
hsa-miR-342-5p	2,22	2,52E-04	2,01	2,21E-03	1,95	2,38E-03	1,74	0,01	1,52	8,54E-03	-1,45	4,47E-03
hsa-miR-29b-2-5p	1,45	7,39E-03	1,27	3,98E-02	1,92	4,73E-04						
hsa-miR-18a	-2,15	4,54E-04			-1,61	2,81E-02					1,62	1,96E-03
hsa-miR-934	-1,51	1,00E-03			-1,37	6,97E-03					1,37	4,19E-04
hsa-miR-193b	1,27	7,80E-03			1,36	7,97E-03						
hsa-miR-375	3,10	4,38E-03			3,60	1,33E-03					-3,01	6,65E-04
hsa-miR-146a	-1,60	2,26E-02									1,34	1,92E-02
hsa-miR-362-5p	-1,38	4,61E-03									0,90	4,01E-02
hsa-miR-20a	-1,27	1,55E-03									1,01	2,89E-03
hsa-miR-17	-1,32	2,52E-04										
hsa-miR-106a	-1,30	4,32E-04										
hsa-miR-505-5p	-1,29	4,38E-03										
hsa-miR-17-3p	-1,08	3,33E-02										
hsa-miR-30a			1,86	4,91E-03			1,51	0,02				
hsa-miR-612			-1,16	4,33E-02								
hsa-miR-1308			-1,03	4,91E-03								
hsa-miR-30c			1,14	4,12E-02								
hsa-miR-141			1,41	4,09E-02								
hsa-miR-30a-3p			1,98	3,11E-02								
hsa-miR-224-3p							-1,29	0,03				
hsa-miR-125a-5p							1,03	0,02				
hsa-miR-193b							1,09	0,03				
hsa-miR-29b-2-5p							1,74	0,00			-1,17	9,67E-03
hsa-miR-149											-1,33	3,46E-02
hsa-miR-425											-0,77	4,01E-02
hsa-miR-99b											-0,60	2,19E-02
hsa-miR-1975											-0,41	4,01E-02
hsa-miR-18a-3p											0,54	4,01E-02
hsa-miR-18b											0,73	1,92E-02
hsa-miR-19b											0,94	4,16E-02

Para seleccionar los miARNs cuya expresión es estadísticamente significativa y diferencial de la Figura 2, se realizaron comparaciones múltiples y emparejamientos entre los grupos de pronóstico A, B y C. Para ello se emplearon dos enfoques diferentes, "*limma*" y "*RankProd Bioconductor*". Sólo los candidatos con un cambio (FC, del inglés "*fold change*") > 2 (ya sea regulado positivamente (incremento) o regulado negativamente (disminución)) y un p-valor ajustado (del inglés "*adj-pval*") < 0,05 fueron seleccionados (Tabla 5).

Tabla 5. Los miARN regulados negativamente más significativos en tumores de pacientes con recurrencia

		limma F*		**RankProd****		**RT-qPCR*****	
Comparación#	**miARN**	**logFC**	**adj-pval**	**logFC**	**adj-pval**	**logFC**	**SE**
B/A	hsa-miR-149	-1.410	0.0016	-1.615	<0.00001	-2.646	0.724
	hsa-miR-20b	-1.048	0.0071	-1.237	<0.00001	-1.542	0.521
	hsa-miR-30a-3p	-1.359	0.0078	-1.521	<0.00001	-1.001	0.514
	hsa-miR-625	-1.149	0.0014	-1.377	<0.00001	-0.347	0.282
	hsa-miR-10a	-1.235	0.0168	-1.547	<0.00001	-1.108	0.404
BC/A	hsa-miR-149	-1.120	0.0117	-1.329	<0.00001	-2.555	0.681
	hsa-miR-20b	-1.016	0.0076	-1.155	<0.00001	-1.470	0.536
	hsa-miR-30a-3p	-1.124	0.0256	-1.326	<0.00001	-0.994	0.458
	hsa-miR-625	-1.003	0.0049	-1.223	<0.00001	-0.266	0.237
B/AC	hsa-miR-149	-1.294	0.0052	-1.446	<0.00001	-2.340	0.698
	hsa-miR-10a	-1.397	0.0093	-1.647	<0.00001	-1.241	0.404
	hsa-miR-342-5p	-1.123	0.0159	-1.254	<0.00001	-1.194	0.627

#Grupo A = no recurrencia, Group B = recurrencia temprana (≤24 meses tras la cirugía), Grupo C = recurrencia tardía (50-60 meses tras la cirugía)

**limma F*, análisis de datos filtrados (sd> 70%) usando limma

***RankProd*, análisis de datos no filtrados usando el algoritmo RankProduct

****RT-qPCR*, la expresión relativa de los miARN se calculó empleando el método $\Delta\Delta C_t$. El error estándard (SE) se calculó en base a la teoría de la propagación del error (theory of error propagation, en inlgés) (Quackenbush (2002) Microarray data normalization and transformation. Nat Genet 32 Suppl: 496-501).

Por lo tanto, la comparación de los p-valores obtenidos con bibliotecas tanto "limma" y "RankProd logFC" condujo a la identificación de miR -149, miR- 20b, miR- 30a- 3p,

miR- 342- 5p , miR- 625 y miR- 10a como los miARNs cuya expresión cambia más significativamente cuando se comparan tumores de pacientes libres de recurrencia con tumores de pacientes con recurrencia, es decir, el grupo B *vs* A o BC *vs* A (Tabla 5). Tal y como se había observado en la agrupación jerárquica (Figura 2), las diferencias más grandes en la expresión de los seis miARNs se detectaron de nuevo al comparar el grupo B *vs* el grupo A (Tabla 5). En contraste, las comparaciones pareadas de cualquiera de los grupos A o B con el grupo C no resultaron en ningún miARN estadísticamente significativo. En particular, los niveles relativos de todos los miARNs candidatos fueron más bajos en las muestras del grupo B en comparación con los otros, lo que sugiere que estos miARNs podrían actuar, directa o indirectamente, como supresores de la metástasis.

En cuanto a los subtipos intrínsecos, se encontraron niveles bajos de miR-149, miR-30a-3p y miR-342-5p en tumores “ER-“ (Tabla 6).

La expresión diferencial de los seis miARNs también se determinó por RT- qPCR en los tres grupos de pronóstico (Tabla 5). Con la excepción de miR- 625, que no pudo ser validada, se confirmó que miR- 149-5p , miR- 20b-5p, miR10a-5p , miR-30a-3p y miR-342-5p (la "*firma 5-miARN* " o “*firma de los 5 miARNs*”, de ahora en adelante) estaban regulados negativamente en los tumores de pacientes con recurrencia (grupos B o C) en comparación con la expresión en tumores de pacientes sin recurrencia (grupo A, Tabla 5). El miR- 625 fue excluido de cualquier estudio adicional ya que los datos RT- qPCR mostraron una variación mínima entre los grupos (FC < 2). A continuación, se reagruparon los 71 tumores en función de la firma 5-miARN. Como se muestra en la Figura 3, los tumores de los grupos A y B se separaron claramente en dos grupos distintos, que incluían la mayor parte de las muestras que se espera en cada categoría: 78,8% de los tumores de tipo A en el cluster 1b (riesgo bajo) y el 70,4 % de los tumores grupo B en el cluster 2b (alto riesgo). Es de destacar que el análisis supervisado incluye la mayoría de los tumores del grupo C (72,8 %), en el cluster 1b, lo que indica que la firma 5-miARN discrimina específicamente tumores con un riesgo global mayor de recurrencia temprana.

<u>Ejemplo 3</u>: El valor pronóstico de la firma de los 5 miARN

Se examinó la relación entre la expresión de la firma 5-miARN y la supervivencia libre de recurrencia (RFS) mediante un análisis de supervivencia. La Figura 4A muestra un gráfico de Kaplan-Meier para toda la serie de pacientes incluidos en el estudio. Debido a las características intrínsecas de la cohorte, el descenso en el RFS sólo se observa en los intervalos de 0-24 y 50-60 meses (correspondientes a los grupos B y C, respectivamente). A continuación, se agruparon los tumores de acuerdo a su firma 5-miARN en dos grupos diferentes. Un primer grupo incluye los tumores con los cinco miARNs regulados negativamente simultáneamente, (FC> 2 y p <0,05) y un segundo grupo incluye aquellos tumores que no tenían los cinco miARNs regulados negativamente simultáneamente. Se realizó un análisis de supervivencia utilizando los datos clínicos de los pacientes correspondientes. Como se muestra en la Figura 4B, las gráficas de Kaplan-Meier para ambos grupos demuestran que la firma 5-miARN define un grupo de "alto riesgo" de los pacientes con RFS cortos (test Peto-Peto con p-valor = 0,02, al comparar la bajo frente a los grupos de alto riesgo).

Utilizando un modelo de regresión de riesgos proporcionales de Cox ("*Cox proportional hazard regression model*", en inglés), también se pusieron a prueba todas las posibles combinaciones de diferentes covariables (subtipo de tumor, edad del paciente, tamaño del tumor, número de ganglios linfáticos afectados y la firma 5-miARNs) con recurrencia temprana (≤ 24 meses) para identificar los mejores factores pronósticos. El mejor modelo de acuerdo con el criterio AIC incluye el tamaño del tumor y la expresión de la firma 5-miARN (datos no mostrados). Sólo la firma 5-miARN (los cinco miARNs regulados negativamente) resultó estadísticamente significativa en el modelo de Cox para el grupo de alto riesgo (p-valor = 0.02 con HR = 2.73, 95% CI: 1.17-6.36). Los datos de expresión de 5-miARN también se utilizaron para desarrollar un modelo de predicción a través de "bootstrapping" en un clasificador "Naive Bayes" (B = 200 con N = 71, véase Ejemplo 1, métodos). La precisión de pronóstico de los modelos se evaluó por el test ROC ("*receiver operating charateristic*", en inglés) (Figura 5). Considerados individualmente, miR-30a-3p y miR-10a-5p mostraron un área bajo la curva (AUC) sorprendentemente alta (0,890 y 0,875, respectivamente). Este resultado sugiere que los mARN diana regulados por miR-30a-3p y miR-10a-5p podrían tener una potencial mayor contribución para el resultado final de la enfermedad. Sin embargo, la firma de

los 5 miARN tuvo el valor predictivo más fuerte a la hora de discriminar los tumores de pacientes que desarrollarán recurrencia precoz (grupo B) de los que no desarrollan recurrencia (libres de enfermedad) (grupo A), con un AUC = 0,993 (Figura 5). En resumen, la firma de los 5 miARN posee un elevado valor como predictor del riesgo de recurrencia temprana del cáncer de mama.

Ejemplo 4: Posibles mARN diana de la firma 5 - miARN

Empleando bases de datos públicas existentes (MTI). En particular, los objetivos validados se obtuvieron de miTarBase y miRecords (ver Ejemplo 1, métodos). En primer lugar, se creó una red biológica en Cytoscape (Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498-2504), que contiene todos los miARNs individuales incluidos en la firma 5 - miARN: miR-149-5p, miR-10a-5p, miR-20b-5p, miR-30a-3p y miR-342-5p). A continuación, se amplió la red mediante la adición de datos de MTI de *H. sapiens* recuperados de los depósitos indicados y, por último, se generaronlas redes de interacción de regulación extendidas (RIN, del inglés "*regulatory interaction networks*") y se visualizaron en Cytoscape. Cada interacción regulatoria en la red consta de dos nodos, un componente normativo (miARN) y una biomolécula diana (ARNm) conectados a través de una arista dirigida. La Figura 6 muestra la red extendida cuando el umbral RIN se establece en 1 (es decir, cada diana predicha aparece en, al menos, un RIN). Así, con RIN = 1 la red incluyó 14 dianas validadas asignadas a miR-20b-5p ((VEGFA, BAMBI, EFNB2, MYLIP, CRIM1, ARID4B, HIF1A, HIPK3, CDKN1A, PPARG, STAT3, MUC17, EPHB4, y ESR1) , 7 dianas validadas asignados a miR-10a-5p (HOXA1, NCOR2, SRSF1, SRSF10/TRA2B, MAP3K7, USF2 y BTRC) y 9 dianas validadas asignados a miR-30a-3p (THBS1, VEZT, TUBA1A, CDK6, WDR82, TMEM2, KRT7, CYR61 y SLC7A6) (Figura 5). Teniendo en cuenta estos resultados y teniendo en cuenta que i) la red extendida se construyó considerando la firma 5-miARN como los nodos de la red y ii) todos los MTI representados en la Figura 6 se han verificado experimentalmente, se sugiere que al menos algunos de los 30 mARNs (Figura 6) pueden regularse *in vivo* mediante la firma 5 - miARN en tumores con recurrencia temprana.

Para obtener una mayor comprensión de las bases moleculares del valor pronóstico de la firma 5 - miARN, se investigaron los procesos biológicos asociados con las 30 dianas experimentalmente verificadas de la Figura 6. Para ello, se realizaron búsquedas de términos de Ontología Genética (GO) y búsquedas de proteínas clave en el servidor KEGG (Kyoto Encyclopedia of Genes and Genomes) asociados con los 30 mARN dians en conjunto. Cabe señalar, sin embargo, que se aplicó un enfoque restrictivo, incluyendo sólo dianas de los miARN validadas experimentalmente - miR-149 y miR-342-5p quedaron fuera del análisis GO - y, por tanto, vías biológicas adicionales podrían verse afectados por la regulación negativa de la firma 5-miARN. Para aumentar el valor predictivo del análisis GO, se consideraron sólo términos de ontología con evidencia experimental y un p-valor ≤ 0,01. Curiosamente, la mayoría de las dianas en el conjunto se asociaron con términos GO relacionados con la angiogénesis y la migración celular (GO: 0001954, GO: 0002040 , GO: 0002042 , GO: 0043534 y GO : 0043536), además de los términos de GO " respuesta al estímulo estradiol" (GO: 0032355), "diferenciación de monocitos" (GO : 0030224) y "vía de señalización del receptor “ephrin”" (GO: 0048013) (Figura 7). Otros términos GO de particular relevancia para el estudio fueron: "regulación positiva de la proliferación de fibroblastos" (GO: 0048146) "regulación de la quimiotaxis " (GO: 0050920), " regulación de la respuesta celular a la estimulación del factor de crecimiento" (GO: 0090287) y "regulación positiva de las especies reactivas del oxígeno proceso metabólico" (GO: 2000379). En su conjunto, el análisis computacional de las supuestas dianas de ARNm experimentalmente verificadas para la firma 5-miARN y sus términos GO asociados (p-valor ≤ 0.01) sugieren que la recurrencia precoz en el cáncer de mama es una consecuencia del potencial angiogénico, invasivo y proliferativo superior de un subgrupo de tumores con niveles inferiores de, al menos, miR-20b-5p, miR-10a-5p y miR- 30a-3p (Figura 6). De hecho, la integración de los términos GO en los mapas de vías KEGG, apoya esta hipótesis, ya que el efecto neto de los cambios en las vías de regulación afectadas por un aumento de las dianas previstas sería un aumento tanto de la proliferación como de la angiogénesis (Figura 9).

En un esfuerzo para validar esta hipótesis, se recogieron datos retrospectivos de inmunohistoquímica para la expresión de marcadores de angiogénesis (VEGF) y

proliferación (Ki67) en el conjunto de tumores primarios (Tabla 5), siempre que fuese posible. Como control positivo, se incluyeron los datos de expresión de los receptores de estrógeno (ER), ya que a menudo se asocian con el pronóstico [3,4]. Dado que la baja expresión de los genes miARN debe dar lugar a un aumento de la estabilidad de sus mARNs diana, se prevé un aumento de la expresión de VEGF y Ki67 en aquellos tumores identificados por la firma (grupo de alto riesgo) de los 5 miARN. La cuantificación de VEGF, Ki67 y la inmunotinción de ER se realizó como se ha descrito previamente [Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, et al. (2004) Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 10: 5367-5374; Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, et al. (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol 25: 118-145; Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, et al. (2008) Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. Clin Cancer Res 14: 1368-1376) y se calculó el porcentaje de tumores que muestran baja o alta expresión de cada marcador para cada grupo de pronóstico (A, B o C) o el estado de la firma 5 - miARN (riesgo bajo o alto). Los resultados del análisis se resumen en la Tabla 6.

Tabla 6. Niveles de expresión de VEGF, Ki67 y ER en tumores

		Grupo				**Nivel de riesgo**		
		A	**B**	**C**		**Bajo**	**Alto**	
		n (%)	**n (%)**	**n (%)**	**p-val**	**n (%)**	**n (%)**	**p-val**
VEGF	Bajo	6 (20.7)	4 (15.4)	1 (9.1)	NS	14 (37.8)	5 (17.2)	NS
	Alto	23 (79.3)	22 (84.6)	10 (90.9)		23 (62.2)	24 (82.8)	
Ki67	Negativo	15 (45.4)	4 (14.8)	6 (54.5)	0.012	17 (40.5)	8 (25.6)	NS
	Positivo	18 (54.6)	23 (85.2)	5 (45.5)		25 (59.5)	21 (72.4)	
ER	Negativo	17 (51.5)	18 (66.7)	2 (18.2)	0.025	14 (33.3)	23 (79.3)	< 0.0001
	Positivo	16 (48.5)	9 (33.3)	9 (81.8)		28 (66.7)	6 (20.7)	

Sólo se observó una asociación estadísticamente significativa al comparar Ki67 *vs*

grupos de pronóstico (p -valor = 0,012), ER vs grupos de pronóstico (p-valor = 0,025) o ER vs los grupos de riesgo (p -valor < 0,0001). Por el contrario, no se observó una relación significativa de la expresión de VEGF con cualquiera de los grupos de pronóstico o la firma 5 miARN (Tabla 6). A pesar de eso, se observó un ligero aumento en los niveles de VEGF y Ki67 en tumores con recurrencia temprana (grupo B) y en el grupo de "alto riesgo" (Tabla 6). Un análisis de supervivencia también mostró una reducción de RFS en esos pacientes con tumores positivos para Ki67, negativos para ER y con un aumento de la expresión de VEGF (Figura 8). Una vez más, sin embargo, sólo los niveles de Ki67 se asociaron significantemente con RFS (P = 0,044, Figura 8, panel central).

Se ha demostrado aquí que los tumores primarios de pacientes con recurrencia tempranas son diferentes a los tumores de pacientes sin recurrencia, por lo menos en relación a su perfil de miARN. Estas diferencias en la expresión de miARN-que también deben tener un impacto en el tumor transcriptoma-reflejan dos entidades biológicas distintas con, al menos, diferente potencial proliferativo (Figura 8). El presente análisis de redes predijo varios dianas, que también podrían conferir una mayor capacidad angiogénica e invasiva a los tumores con recurrencia temprana. Estos resultados demuestran que los tumores de mama con diferente riesgo de recurrencia pueden diferenciarse en el nivel de expresión de miARN (de la firma de los 5 miARNs, incluyendo miR-149, miR-30a-3p, miR-20b, miR-10a y miR-342-5p). Estos miARN poseen un elevado valor pronóstico (AUC = 0,993, valor de p <0.05) y están regulados negativamente en los tumores primarios de mama de pacientes que desarrollan recurrencia temprana. Además, se hna identificado un conjunto de 30 mARNs reguladss por la firma 5-miRNA en tumores de que muestran recurrencia temprana. Cabe destacar que el conjunto incluye los ARNm que codifican para proteínas implicadas principalmente en la angiogénesis y la proliferación (VEGFA, THBS1, EPHB4 CDK6 y DCKN1, entre otros). Se ha demostrado que los tumores primarios recurrentes tienen un mayor potencial proliferativo significativo, medida por inmunotinción de Ki67.

Referencias:

1. Fisher ER, Palekar AS, Gregorio RM, Redmond C, Fisher B (1978) Pathological findings from the national surgical adjuvant breast project (Protocol No. 4). IV. Significance of tumor necrosis. Hum Pathol 9: 523-530.

2. Elston CW, Ellis IO (1991) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 19: 403-410.

3. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, et al. (2000) Molecular portraits of human breast tumours. Nature 406: 747-752.

4. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, et al. (2001) Gene expression 12 patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98: 10869-10874.

5. Spitale A, Mazzola P, Soldini D, Mazzucchelli L, Bordoni A (2009) Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland. Ann Oncol 20: 628-635.

6. Steeg PS (2006) Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 12: 895-904.

7. Chiang AC, Massague J (2008) Molecular basis of metastasis. N Engl J Med 359: 2814-2823.

8. Saphner T, Tormey DC, Gray R (1996) Annual hazard rates of recurrence for breast cancer after primary therapy. J Clin Oncol 14: 2738-2746.

9. Demicheli R, Abbattista A, Miceli R, Valagussa P, Bonadonna G (1996) Time distribution of the recurrence risk for breast cancer patients undergoing mastectomy: further support about the concept of tumor dormancy. Breast Cancer Res Treat 41: 177-185.

10. Jatoi I, Tsimelzon A, Weiss H, Clark GM, Hilsenbeck SG (2005) Hazard rates of recurrence following diagnosis of primary breast cancer. Breast Cancer Res Treat 89: 173-178.

11. Demicheli R, Retsky MW, Hrushesky WJ, Baum M (2007) Tumor dormancy and surgery-driven interruption of dormancy in breast cancer: learning from failures. Nat Clin Pract Oncol 4: 699-710.

12. Demicheli R, Hrushesky WJ, Retsky MW, Bonadonna G, Valagussa P (2007) Comment to "Effect of primary tumor extirpation in breast cancer patients who present with stage IV disease and intact primary tumor" by G V Babiera, R Rao, L Feng, F Meric-Bernstam, H M Kuerer, S E Singletary, K K Hunt, M I Ross, K M Gwyn, B W Feig, F C Ames, G N Horthobagyi. Ann Surg Oncol 2006;13:776-782. Ann Surg Oncol 14: 1519-1520.

13. Demicheli R, Retsky MW, Hrushesky WJ, Baum M, Gukas ID (2008) The effects of surgery on tumor growth: a century of investigations. Ann Oncol 19: 1821- 1828.

14. Demicheli R, Valagussa P, Bonadonna G (2001) Does surgery modify growth kinetics of breast cancer micrometastases? Br J Cancer 85: 490-492.

15. Hess KR, Pusztai L, Buzdar AU, Hortobagyi GN (2003) Estrogen receptors and distinct patterns of breast cancer relapse. Breast Cancer Res Treat 78: 105-118.

16. Demicheli R, Biganzoli E, Ardoino I, Boracchi P, Coradini D, et al. (2010) Recurrence and mortality dynamics for breast cancer patients undergoing mastectomy according to estrogen receptor status: different mortality but similar recurrence. Cancer Sci 101: 826-830.

17. Perez-Rivas LG, Jerez JM, Fernandez-De Sousa CE, de Luque V, Quero C, et al. (2012) Serum protein levels following surgery in breast cancer patients: a protein microarray approach. Int J Oncol 41: 2200-2206.

18. Carlsson A, Wingren C, Kristensson M, Rose C, Ferno M, et al. (2011) Molecular serum portraits in patients with primary breast cancer predict the development of distant metastases. Proc Natl Acad Sci U S A 108: 14252-14257.

19. Pietrowska M, Polanska J, Marczak L, Behrendt K, Nowicka E, et al. (2010) Mass spectrometry-based analysis of therapy-related changes in serum proteome patterns of patients with early-stage breast cancer. J Transl Med 8: 66.

20. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, et al. (2005) Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 11: 5678-5685.

21. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, et al. (2010) A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 16: 5222-5232.

22. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, et al. (2009) Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 101: 736-750.

23. Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, et al. (2008) Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. Clin Cancer Res 14: 1368-1376.

24. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, et al. (2010) Metastatic behavior of breast cancer subtypes. J Clin Oncol 28: 3271-3277.

25. Parker JS, Mullins M, Cheang MC, Leung S, Voduc 1 D, et al. (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 27:1160-1167.

26. Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO, et al. (2010) Breast cancer subtypes and the risk of local and regional relapse. J Clin Oncol 28: 1684-1691.

27. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, et al. (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 486: 346-352.

28. Dawson SJ, Rueda OM, Aparicio S, Caldas C (2013) A new genome-driven integrated classification of breast cancer and its implications. EMBO J 32: 617-628.

29. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136: 215-233.

30. Djuranovic S, Nahvi A, Green R (2011) A parsimonious model for gene regulation by miRNAs. Science 331: 550-553.

31. Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U (2004) Nuclear export of microRNA precursors. Science 303: 95-98.

32. Brodersen P, Voinnet O (2009) Revisiting the principles of microRNA target recognition and mode of action. Nat Rev Mol Cell Biol 10: 141-148.

33. Carthew RW, Sontheimer EJ (2009) Origins and Mechanisms of miRNAs and siRNAs. Cell 136: 642-655.

34. Huntzinger E, Izaurralde E (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet 12: 99-110.

35. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36: D154-158.

36. Croce CM (2009) Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet 10: 704-714.

37. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, et al. (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 103: 2257-2261.

38. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144: 646-674.

39. Ruan K, Fang X, Ouyang G (2009) MicroRNAs: novel regulators in the hallmarks of human cancer. Cancer Lett 285: 116-126.

40. Grammatikakis I, Gorospe M, Abdelmohsen K (2013) Modulation of Cancer Traits by Tumor Suppressor microRNAs. Int J Mol Sci 14: 1822-1842.

41. White NM, Fatoohi E, Metias M, Jung K, Stephan C, et al. (2011) Metastamirs: a stepping stone towards improved cancer management. Nat Rev Clin Oncol 8: 75-84.

42. Cho WC (2007) OncomiRs: the discovery and progress of microRNAs in cancers. Mol Cancer 6: 60.

43. Lee YS, Dutta A (2009) MicroRNAs in cancer. Annu Rev Pathol 4: 199-227.

44. Foekens JA, Sieuwerts AM, Smid M, Look MP, de Weerd V, et al. (2008) Four miRNAs associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer. Proc Natl Acad Sci U S A 105: 13021-13026.

45. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, et al. (2005) MicroRNA expression profiles classify human cancers. Nature 435: 834-838.

46. Volinia S, Croce CM (2013) Prognostic microRNA/mRNA signature from the integrated analysis of patients with invasive breast cancer. Proc Natl Acad Sci U S A 110: 7413-7417.

REIVINDICACIONES

1. Procedimiento para pronosticar el riesgo de recurrencia de cáncer de mama que comprende medir los niveles todos los miARN del grupo que consiste en:

miR-149-5p (SEQ ID NO: 1)
miR-10a-5p, (SEQ ID NO: 2)
miR-20b-5p, (SEQ ID NO: 3)
miR-30a-3p (SEQ ID NO: 4); y
miR-342-5p, (SEQ ID NO: 5)

en una muestra del tumor, en el que una reducción del nivel de todos los miARN en el tumor, con respecto a un control, es indicativo de un elevado riesgo de recurrencia del tumor.

2. Procedimiento para pronosticar la supervivencia de un sujeto con cáncer de mama que comprende medir los niveles de todos los miARN del grupo que consiste en:

miR-149-5p (SEQ ID NO: 1)
miR-10a-5p, (SEQ ID NO: 2)
miR-20b-5p, (SEQ ID NO: 3)
miR-30a-3p (SEQ ID NO: 4); y
miR-342-5p, (SEQ ID NO: 5)

en una muestra del tumor, en el que una reducción del nivel de todos los miARN en el tumor, con respecto a un control, es indicativa de una baja supervivencia.

3. El procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores caracterizado por que el nivel de expresión todos los miARN se puede obtener mediante:

a. Un procedimiento de perfilado génico, tal como una micromatriz; y/o
b. Un procedimiento que comprende la PCR, tal como la PCR en tiempo real; y/o
c. Transferencia Northern.

4. El procedimiento de acuerdo con la reivindicación 3, caracterizado por que el nivel de expresión de todos los miARN se puede obtener mediante la PCR cuantitativa en tiempo real.

5. El procedimiento de acuerdo con una cualquiera de las reivindicaciones anteriores, caracterizado por que el procedimiento se lleva a cabo *in vitro* usando una muestra originaria del sujeto humano.

6. El procedimiento de una cualquiera o más de las reivindicaciones anteriores, en el que la expresión de los miARN está normalizada, preferiblemente en relación con la expresión de RNU48 y/o miR-16.

7. Un procedimiento para clasificar un sujeto humano que padece cáncer de mama en uno de dos grupos, en el que el grupo 1 comprende los sujetos identificables como con alto riesgo de recurrencia mediante el procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, y en el que el grupo 2 representa el resto de sujetos.

8. Un kit que comprende al menos cinco oligonucleótidos, caracterizado por que los cinco oligonucleótidos se seleccionan de entre el grupo que consiste en cinco oligonucleótidos capaces de hibridarse con los miARN o su cADN tal como se ha definido en las SEQ ID NOs: 1 a 5 en condiciones de rigor.

9. Un kit de acuerdo con la reivindicación 8, caracterizado por que los oligonucleótidos están inmovilizados en manchas sobre una superficie, que es preferiblemente la superficie de una micromatriz.

10. Uso del kit de acuerdo con la reivindicación 8 en un procedimiento de acuerdo con la reivindicación 1.

Figura 1.

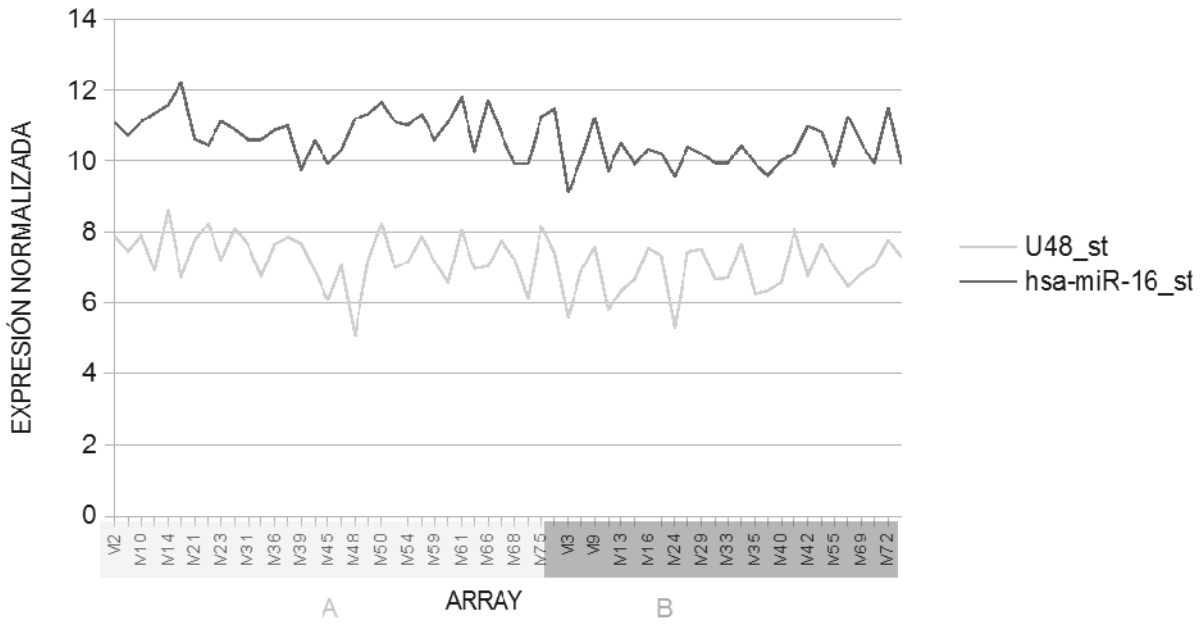

14
12
10
8
6
4
2
0
EXPRESIÓN NORMALIZADA
V2
V10
V14
V21
V23
V31
V36
V39
V45
V48
V50
V54
V59
V61
V66
V68
V75
V3
V9
V13
V16
V24
V29
V33
V35
V40
V42
V55
V69
V72
A
ARRAY
B
U48_st
hsa-miR-16_st

Figura 2

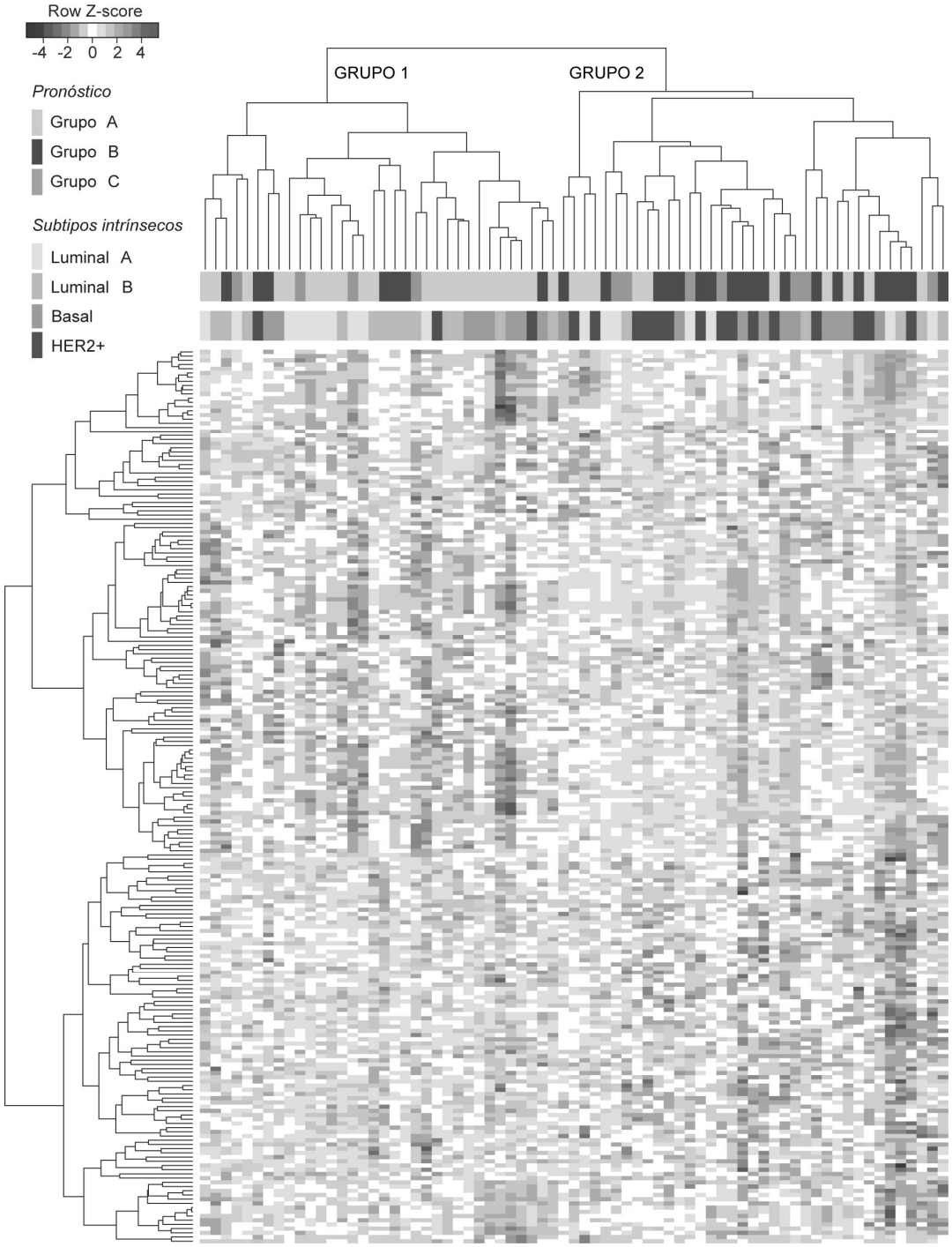

Row Z-score
-4 -2 0 2 4
Pronóstico
Grupo A
Grupo B
Grupo C
Subtipos intrínsecos
Luminal A
Luminal B
Basal
HER2+
GRUPO 1
GRUPO 2

Figura 3

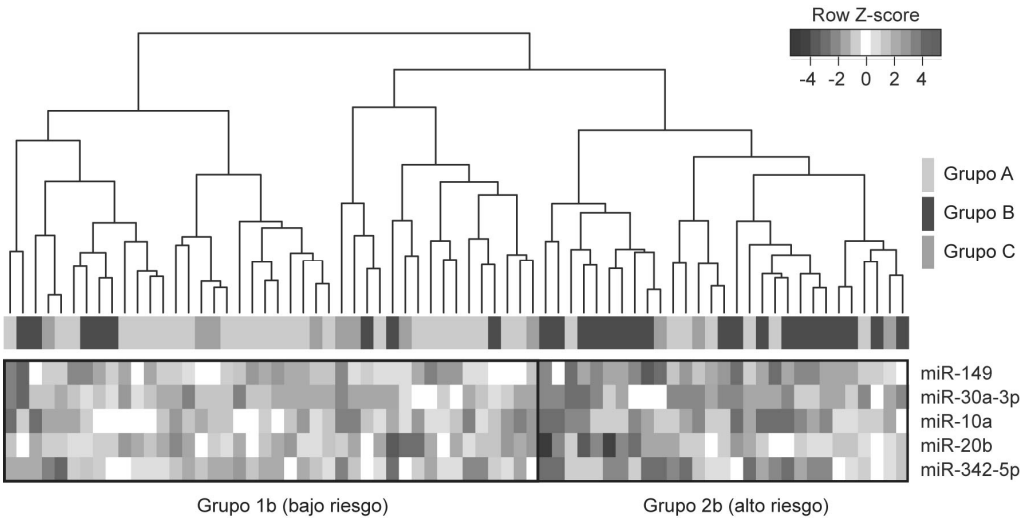


Figura 4

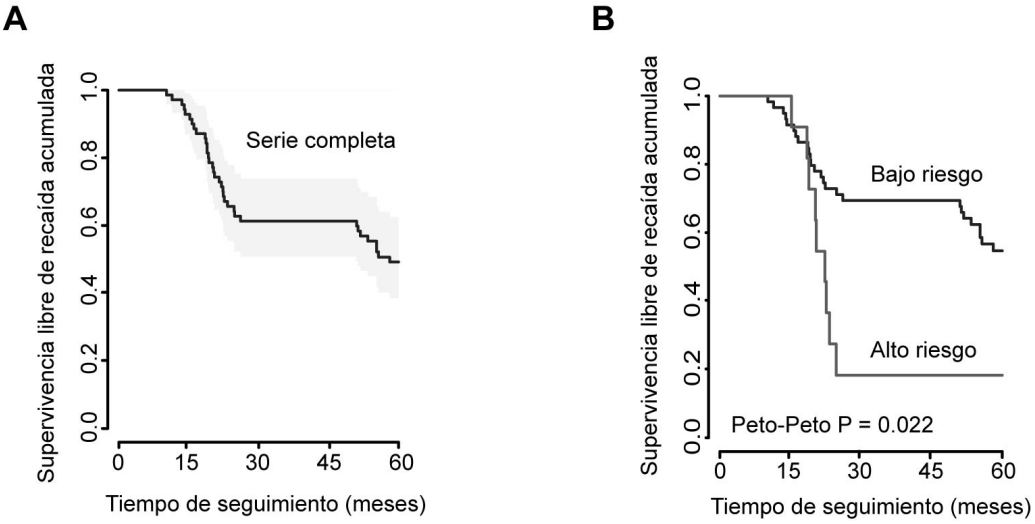


Figura 5

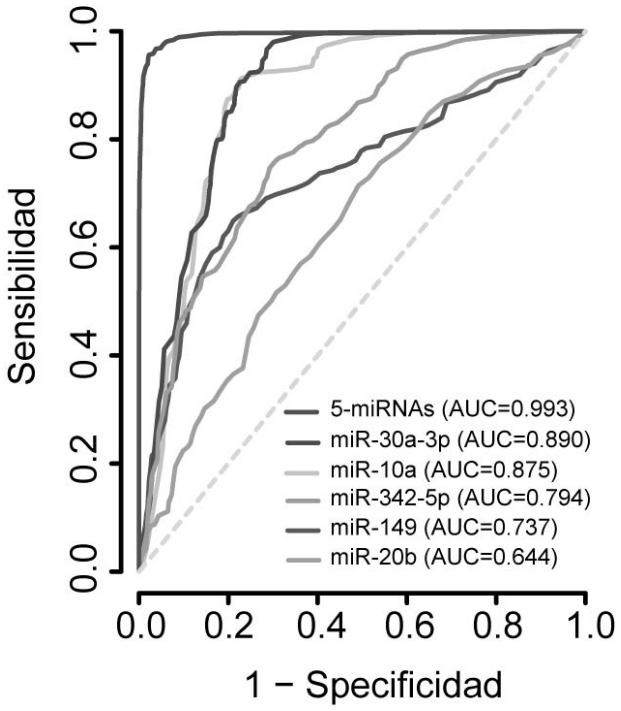


Figura 6

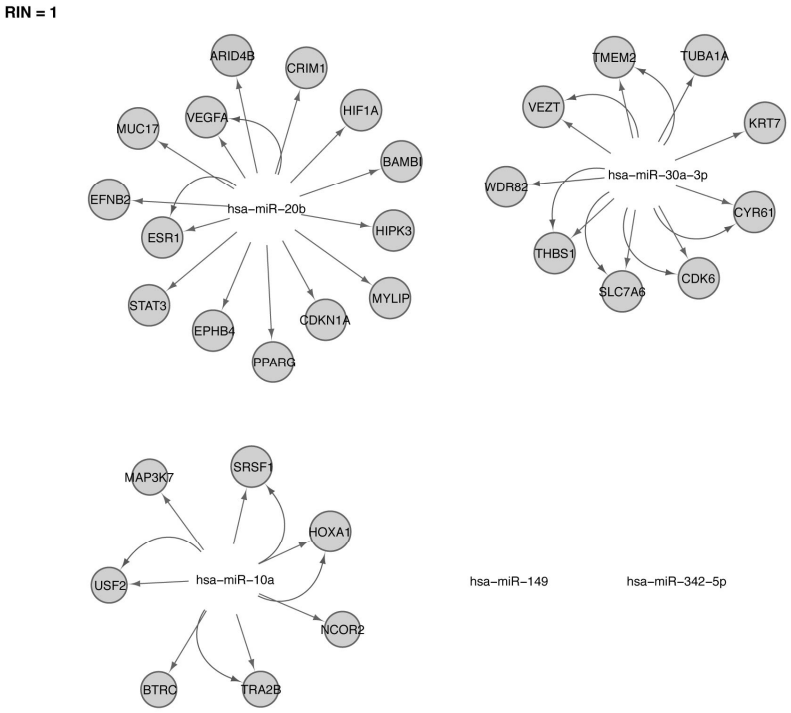


Figura 7

GO ID	Término GO	Genes/término GO
GO:0002040	sprouting angiogenesis	[EFNB2, EPHB4, THBS1, VEGFA]
GO:0043534	blood vessel endothelial cell migration	[EFNB2, EPHB4, THBS1, VEGFA]
GO:0002042	cell migration involved in sprouting angiogenesis	[EFNB2, EPHB4, VEGFA]
GO:0050795	regulation of behavior	[EFNB2, HOXA1, THBS1, VEGFA]
GO:0001952	regulation of cell-matrix adhesion	[CDK6, THBS1, VEGFA]
GO:0043542	endothelial cell migration	[EFNB2, EPHB4, THBS1, VEGFA]
GO:0010810	regulation of cell-substrate adhesion	[CDK6, THBS1, VEGFA]
GO:0007160	cell-matrix adhesion	[CDK6, THBS1, VEGFA]
GO:0030224	monocyte differentiation	[PPARG, VEGFA]
GO:0048013	ephrin receptor signaling pathway	[EFNB2, EPHB4]
GO:0090287	regulation of cellular response to growth factor stimulus	[BAMBI, THBS1, VEGFA]
GO:0050920	regulation of chemotaxis	[EFNB2, THBS1, VEGFA]
GO:0001954	positive regulation of cell-matrix adhesion	[CDK6, VEGFA]
GO:0032355	response to estradiol stimulus	[ESR1, STAT3]
GO:0043536	positive regulation of blood vessel endothelial cell migration	[THBS1, VEGFA]
GO:2000379	positive regulation of reactive oxygen species metabolic process	[CDKN1A, THBS1]
GO:0010811	positive regulation of cell-substrate adhesion	[CDK6, VEGFA]
GO:0048146	positive regulation of fibroblast proliferation	[CDK6, CDKN1A]
GO:0090288	negative regulation of cellular response to growth factor stimulus	[BAMBI, THBS1]
GO:0010595	positive regulation of endothelial cell migration	[THBS1, VEGFA]
GO:0045667	regulation of osteoblast differentiation	[CDK6, CYR61]
GO:0043627	response to estrogen stimulus	[ESR1, STAT3]
GO:0090100	positive regulation of transmembrane receptor protein S/T kinase signaling pathway	[CYR61, THBS1]
GO:0017015	regulation of transforming growth factor beta receptor signaling pathway	[BAMBI, THBS1]
GO:0035239	tube morphogenesis	[HOXA1, VEGFA]
GO:0043154	negative regulation of cysteine-type endopeptidase activity involved in apoptotic process	[THBS1, VEGFA]
GO:0008360	regulation of cell shape	[BAMBI, VEGFA]

p-valor
0.00 0.01

Figura 8

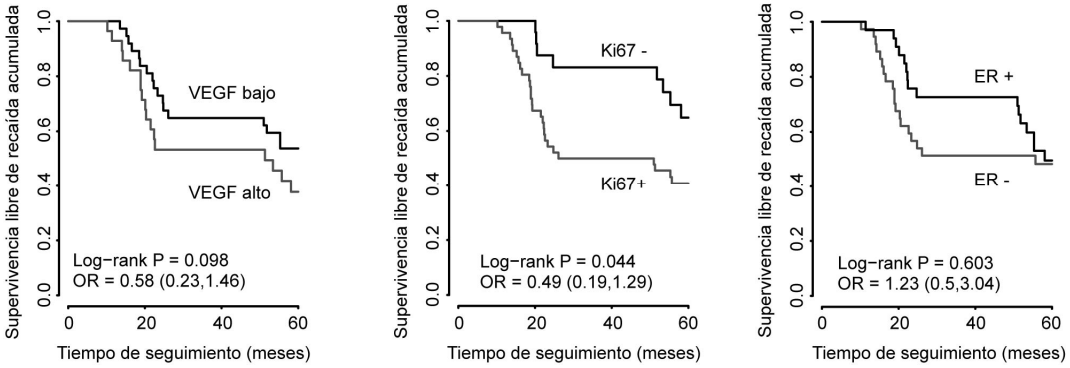
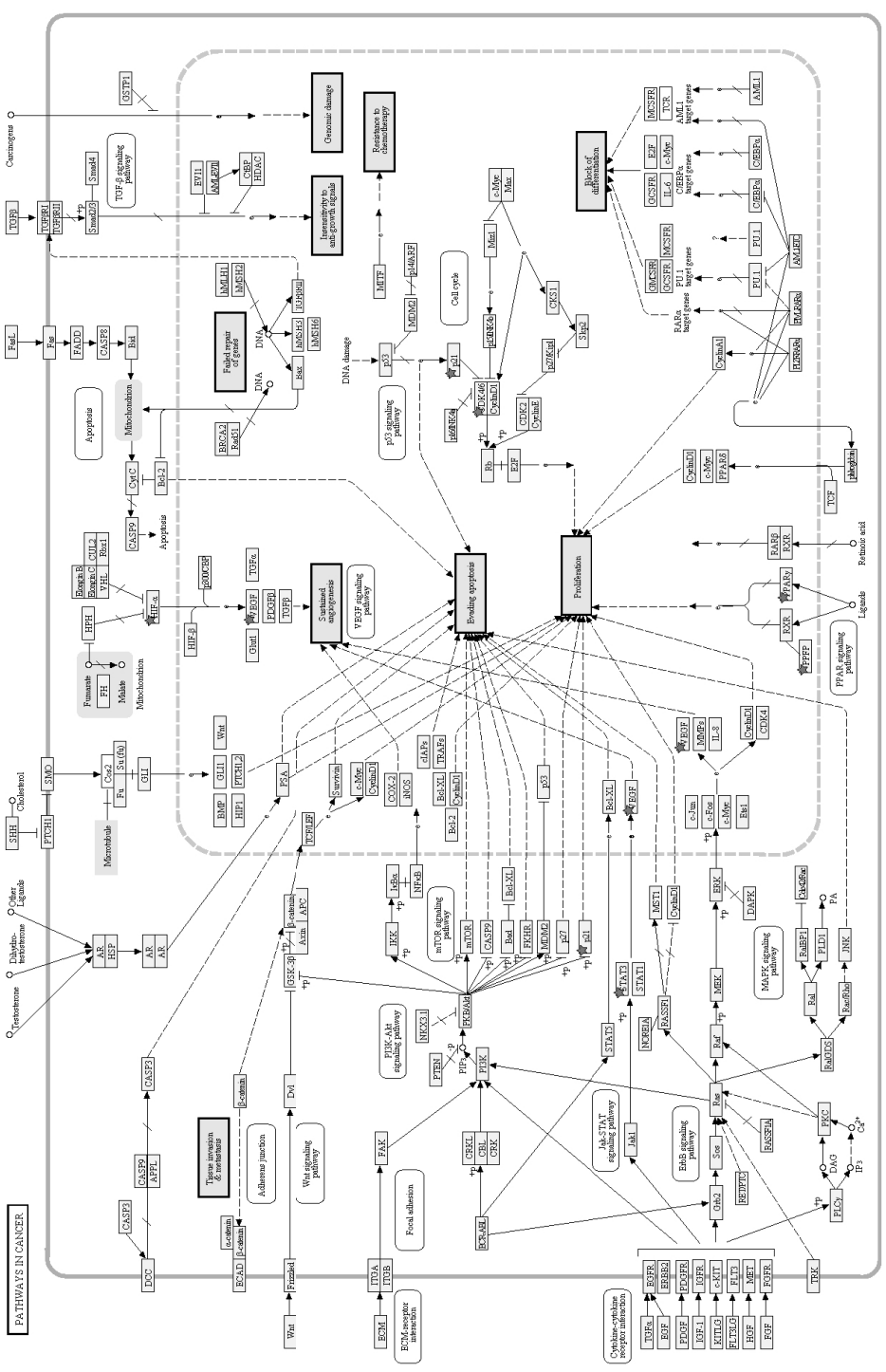


Figura 9


PATHWAYS IN CANCER
Carcinogens
TGFβ
TGFβR1
TGFβR2
Smad4
Smad2/3
TGF-β signaling pathway
GSTP1
Genomic damage
Resistance to chemotherapy
Insensitivity to anti-growth signals
EVI1
AML/ETO
CBP/HDAC
Fas L
Fas
FADD
CASP8
Bid
Apoptosis
Mitochondrion
Cyt C
Bcl-2
CASP9
Apoptosis
Failed repair of genes
DNA
hMLH1
hMSH2
hMSH3
hMSH6
UNG/MU
Ku86
Ku70
BRCA2
Rad51
DNA damage
p53
p21
MDM2
MDTP
p14ARF
p16INK4a
p27Kip1
CDK2
Cyclin E
CDK6
Cyclin D1
Rb
E2F
p38MAPK
Mad1
c-Myc
Max
CKS1
Skp2
Cell cycle
Block of differentiation
M-CSFR
TCR
AML1 target genes
E2F
c-Myc
GCSFR
IL-6
C/EBPα target genes
C/EBPε
C/EBPβ
GM-CSFR
G-CSFR
M-CSFR
PU.1 target genes
PU.1
RARα target genes
RARα
RXRα
PPARγ
Cyclin A
Cyclin D1
c-Myc
PPARδ
TGF
Retinoic acid
Ligands
PPAR signaling pathway
Evading apoptosis
Proliferation
Sustained angiogenesis
VEGF signaling pathway
HIF-α
HPH
Fumarate
FH
Malate
Mitochondrion
Boxes B
Boxes C
CUL2
VHL
Rbx1
HIF-β
p300/CBP
TGFα
EGF
FGF
FGFR
TGFβ
Grb1
Cytochrome
SHH
PTCH1
SMO
Cholesterol
Cos2
Fu
Su(fu)
GLI
Microtubule
GLI1
Wnt
BMP
HIP1
PTCH1,2
FRA
Survivin
c-Myc
Cyclin D1
COX-2
iNOS
c-IAPs
TRAFs
Bcl-XL
Bcl-2
Cyclin D1
TCF/LEF
Other Ligands
Dihydro-testosterone
Testosterone
AR
HSP
AR
AR
GSK-3β
β-catenin
Axin
APC
IKK
IκBα
NFκB
mTOR signaling pathway
mTOR
CASP9
Bad
Bcl-XL
FKHR
MDM2
p27
p21
p53
Bcl-XL
EGF
c-Jun
c-Fos
c-Myc
Elk1
MMPs
IL-8
Cyclin D1
CDK4
MST1
Cyclin D1
ERK
DUSP6
MAPK signaling pathway
Ras-GRas
PLD1
JNK
Rac/Rho
Rac
RaMDS
Raf
MEK
Ras
RASSF1
RASFR
PKC
DAG
IP3
Ca2+
PLCγ
Grb2
Sos
RENTS
ERK signaling pathway
PI3K-Akt signaling pathway
NCK3.1
PTEN
PIP2
PIP3
PI3K
PKB/Akt
STAT5
STAT1
NF-κB/IκB
RASSF
Jak-STAT signaling pathway
Jak1
CASP3
CASP9
APPL
CASP3
DCC
Tissue invasion & metastasis
β-catenin
E-cadherin
E-CAD
Adhesion junction
Dvl
Wnt signaling pathway
Frizzled
Wnt
FAK
Focal adhesion
CRKL
CRK
CRK
BCR-ABL
ITGA
ITGB
ECM
ECM-receptor interaction
Cytokine-cytokine receptor interaction
TGFα
EGF
PDGF
IGF-1
KITLG
c-KIT
FLT3
FLT3LG
HGF
HGF
PGF
EGFR
ERBB2
PDGFRB
IGF1R
FLT3
MET
FGFR
TPK

Listado de Secuencias

<110> UNIVERSIDAD DE MALAGA
SERVICIO ANDALUZ DE SALUD

<120> FIRMA DE MICROARN COMO INDICADOR DEL RIESGO DE RECURRENCIA TEMPRANA EN PACIENTES CON CÁNCER DE MAMA

<130> 900 322

<160> 16

<170> PatentIn version 3.5

<210> 1
<211> 23
<212> RNA
<213> Homo sapiens

<400> 1
ucuggcuccg ugucuucacu ccc 23

<210> 2
<211> 23
<212> RNA
<213> Homo sapiens

<400> 2
uacccuguag auccgaauuu gug 23

<210> 3
<211> 23
<212> RNA
<213> Homo sapiens

<400> 3
caaagugcuc auagugcagg uag 23

<210> 4
<211> 22
<212> RNA
<213> Homo sapiens

<400> 4
cuuucagucg gauguuugca gc 22

<210> 5
<211> 21
<212> RNA
<213> Homo sapiens

<400> 5
aggggugcua ucugugauug a 21

<210> 6
<211> 107
<212> RNA
<213> Homo sapiens

<400> 6
gugcucgcuu cggcagcaca uauacuaaaa uuggaacgau acagagaaga uuagcauggc 60

cccugcgcaa ggaugacacg caaauucgug aagcguucca uauuuuu 107

<210> 7
<211> 63
<212> RNA
<213> Homo sapiens

<400> 7
agugaugaug accccaggua acucugagug ugucgcugau gccaucaccg cagcgcucug 60

acc 63

<210> 8
<211> 22
<212> RNA
<213> Homo sapiens

<400> 8
uagcagcacg uaaauauugg cg 22

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Para amplificar miR-149-5p

<400> 9
gggagtgaag acacggagcc aga 23

<210> 10
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Para amplificar miR-10a-5p

<400> 10
cacaaattcg gatctacagg gta 23

<210> 11
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Para amplificar miR-20b-5p

<400> 11
ctacctgcac tatgagcact ttg 23

<210> 12
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Para amplificar miR-30a-3p

<400> 12
gctgcaaaca tccgactgaa ag 22

<210> 13
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Para amplificar miR-342-5p

<400> 13
tcaatcacag atagcacccc t 21

<210> 14
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Para amplificar RNU6B

<400> 14
aaaaatatgg aacgcttcac g 21

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Para amplificar RNU48

<400> 15
ggtcagagcg ctgcggtgat g 21

<210> 16
<211> 22
<212> DNA

<213> Artificial Sequence

<220>
<223> Para amplificar el miR-16-5p

<400> 16
cgccaatatt tacgtgctgc ta 22