

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 724 404**

51 Int. Cl.:

**C12Q 1/6886** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.02.2016 PCT/DK2016/050053**

87 Fecha y número de publicación internacional: **01.09.2016 WO16134727**

96 Fecha de presentación y número de la solicitud europea: **26.02.2016 E 16707624 (9)**

97 Fecha y número de publicación de la concesión europea: **03.04.2019 EP 3262186**

54 Título: **Un procedimiento basado en microARN para evaluar el pronóstico de un paciente de cáncer de próstata**

30 Prioridad:

**27.02.2015 DK 201500113**

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

**10.09.2019**

73 Titular/es:

**QIAGEN GMBH (33.3%)  
Qiagen Strasse 1  
40724 Hilden, DE;  
AARHUS UNIVERSITET (33.3%) y  
REGION MIDTJYLLAND (33.3%)**

72 Inventor/es:

**THOMSEN RØNFELDT, ANNI;  
KRISTENSEN, HELLE;  
SØRENSEN DALSGAARD, KARINA;  
KONGSBAK, LARS;  
MOURITZEN, PETER y  
ØRNTOFT, TORBEN**

74 Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

**ES 2 724 404 T3**

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

**DESCRIPCIÓN**

Un procedimiento basado en microARN para evaluar el pronóstico de un paciente de cáncer de próstata

**5 CAMPO DE LA INVENCION:**

La presente invención se refiere a la predicción de la malignidad del cáncer de próstata (pronóstico) en el momento del diagnóstico (o más tarde) midiendo el nivel de unos cuantos biomarcadores de microARN característicos en el tejido prostático, para guiar las decisiones sobre el tratamiento (por ejemplo, conducta expectante, cirugía, radiación).

10

**ANTECEDENTES DE LA INVENCION:**

El cáncer de próstata es el cáncer en hombres diagnosticado con más frecuencia y la quinta causa principal de mortalidad asociada a cáncer en los países industrializados (1). El cáncer de próstata se diagnostica típicamente en base a los niveles incrementados del antígeno prostático específico (PSA) en suero, seguido del estudio histopatológico de las biopsias por punción.

15

Sin embargo, el uso del PSA para la detección del cáncer de próstata se asocia con considerables tasas de falsos positivos y no distingue bien entre los tumores de escasa malignidad y de gran malignidad. Durante las últimas décadas, el uso incrementado de las pruebas de PSA y la detección basada en PSA ha dado como resultado incidencias más altas, así como un descenso del estadio de la enfermedad.

20

Sin embargo, el PSA, así como los demás indicadores pronóstico disponibles actualmente (principalmente el número de biopsias positivas, el estadio TNM clínico y la puntuación de Gleason), no puede predecir con precisión los pacientes con un cáncer de próstata de gran malignidad que requieren un tratamiento inmediato. Esto da lugar a un marcado sobretreatmento, y muchos pacientes se someten a PR o radioterapia innecesarias, lo que se asocia con efectos secundarios peores que vivir con el cáncer de próstata no mortal no tratado.

25

De ahí que exista una importante necesidad no satisfecha en el diagnóstico del cáncer de próstata de desarrollar procedimientos que puedan mejorar la evaluación pronóstica distinguiendo correctamente entre los cánceres de baja malignidad, que se pueden controlar de forma segura mediante una conducta expectante, y los cánceres de gran malignidad, que se beneficiarán de una intervención precoz.

30

Una reciente nueva clase de potenciales biomarcadores para el cáncer de próstata es el microARN.

35

Los microARN comprenden una clase de pequeños ARN reguladores no codificantes endógenos (22 nt), que controlan la expresión génica a nivel postranscripcional en diversos organismos, incluyendo mamíferos (2). Los microARN se transcriben como transcritos de microARN primarios de bucle-tallo emparejados imperfectos largos (pri-microARN) mediante la ARN polimerasa II, y se procesan adicionalmente en microARN precursores de horquilla (pre-microARN) mediante la endonucleasa RNasa III nuclear, Drosha (3). Después de la exportación al citoplasma mediante exportina-5-Ran-GTP, otra endonucleasa RNasa III, Dicer, escinde el pre-microARN en un dúplex de microARN de 22 nt maduro (3). Los microARN maduros median su función mientras se incorporan en el complejo de silenciamiento inducido por microARN (miRISC). El microARN guía este complejo a los ARNm diana perfectamente/casi perfectamente complementarios, dando lugar a la inhibición traduccional o bien a la degradación del ARNm (4).

40

45

Los microARN son una de las clases más abundantes de moléculas reguladoras de genes y la última publicación de la miRBase (versión 21) contiene 2588 microARN humanos maduros (1881 precursores), <http://www.mirbase.org/> (5). En conjunto, se ha estimado que los microARN regulan hasta dos tercios de todos los ARNm humanos. En consecuencia, los microARN influyen en numerosos procesos en la célula, por ejemplo, la diferenciación celular, la progresión del ciclo celular y la apoptosis, y la desregulación de los microARN a menudo están conectadas con patologías humanas, incluyendo el cáncer (6). Adicionalmente, algunos microARN parecen ser de tipo celular y específicos de enfermedad y la expresión de los microARN desregulados se ha asociado con tanto el desarrollo como la progresión del cáncer (7). Por tanto, se ha investigado la expresión de los microARN anómala como una fuente potencial prometedora de biomarcadores novedosos para el diagnóstico precoz del cáncer (7). Además, los microARN tienen el potencial de usarse como dianas de los tratamientos basados en microARN para el cáncer (8). Varios estudios de identificación de microARN también han informado de microARN expresados de forma anómala en el desarrollo y/o progresión del cáncer de próstata (9). Sin embargo, la mayoría de los estudios de biomarcadores de microARN en el cáncer de próstata publicados hasta la fecha han usado números de muestras de pacientes relativamente bajos y, a menudo, carecen de una validación clínica independiente rigurosa para confirmar el potencial de los biomarcadores de los microARN identificados candidatos.

50

55

60

El documento WO 2014/085906 divulga biomarcadores de microARN para detectar un alto riesgo de recidiva y/o recaída del cáncer de próstata. Sin embargo, el documento WO 2014/085906 no divulga la combinación específica de miR-10b-5p y miR-374b-5p.

65

Aquí se ha realizado la identificación del miRnoma de más de 750 de los microARN más abundantes e identificado los microARN regulados significativamente anómalos en muestras FFPE de tejido tumoral prostático de pacientes con y sin recidiva bioquímica (RBQ) después de una prostatectomía radical (PR). Se identificaron cinco clasificadores pronóstico en la cohorte 1 y se evaluó su precisión en el pronóstico como factores pronóstico del tiempo hasta la recidiva, observada como recidiva bioquímica (PSA) después de la extirpación de la próstata (prostatectomía radical (PR)) (ejemplo 1). A continuación, se validó la precisión en el pronóstico de los clasificadores en dos cohortes de prostatectomía radical independientes (cohorte 2 y cohorte 3) (ejemplo 2-6). A pesar del hecho de que las muestras de tumores prostáticos en la cohorte 3 fueron de diferente origen nacional (EE. UU.), se muestrearon de una manera diferente (ultracongeladas), se sometieron a diferentes procedimientos de extracción de ARN, se analizaron mediante una diferente plataforma de detección de la expresión de los microARN y diferentes características de cohorte (la cohorte 3 fue, en general, de menor malignidad y tuvo menos acontecimientos de recidiva que las cohortes 1 y 2), cuatro de los clasificadores pronóstico de microARN funcionaron igualmente bien en la cohorte externa, lo que subraya la solidez de estos clasificadores.

Los cinco clasificadores de microARN pronóstico mostraron valor pronóstico independiente significativo para la predicción del tiempo hasta la RBQ después de la PR, más allá de las variables clinicopatológicas habituales.

#### SUMARIO DE LA INVENCION:

Como el procedimiento del antígeno prostático específico (PSA) se asocia con considerables tasas de falsos negativos y no distingue bien entre los tumores de escasa malignidad y de gran malignidad clínicas, existe una necesidad de obtener marcadores novedosos del cáncer de próstata que se puedan usar por sí solos o en combinación con los marcadores existentes. La presente invención presenta un conjunto de marcadores y un procedimiento para aplicarlos para la evaluación del pronóstico (malignidad del tumor/riesgo de recidiva).

En el primer aspecto, la invención hace referencia, por tanto, a un procedimiento *in vitro* para la predicción del pronóstico del cáncer de próstata, que comprende medir el nivel de expresión de al menos dos miR seleccionados de un grupo de 13 miR que consiste en: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p, en una muestra de ARN de células de una biopsia tomada de un paciente de cáncer de próstata, en el que un nivel de expresión distinto de dichos al menos 2 miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.

Un segundo aspecto de la invención se refiere a un kit para la evaluación *in vitro* del pronóstico del cáncer de próstata.

Un tercer aspecto de la invención se relaciona con un procedimiento de tratamiento de un paciente que necesita tratamiento para el cáncer de próstata, comprendiendo el procedimiento realizar el procedimiento *in vitro* del primer aspecto, caracterizar a los pacientes con respecto a su pronóstico y seleccionar un tratamiento apropiado para el paciente en base a esta información. Obviamente, el pronóstico es un factor fundamental a considerar cuando se planea el tratamiento adicional del paciente, puesto que una evaluación de pronóstico muy malo perjudicaría a los procedimientos estresantes y angustiosos.

#### DEFINICIONES

El índice de concordancia de Harrell, o índice c, se refiere a un área generalizada debajo de la curva de rendimiento diagnóstico (ABC) para las observaciones objeto de censura estadística y es igual a la probabilidad de concordancia entre la probabilidad predicha de recaída y el desenlace de la recaída (Harrell (2001) Regression modelling strategies: with applications to linear models, logistic regression, and survival analysis. Nueva York: Springer-Verlag).

Las expresiones "microARN", "miARN" y "miR" se usan como sinónimos para hacer referencia a ARN no codificantes de aproximadamente 18-25 nucleótidos (nt) de longitud derivados de genes endógenos. Los microARN se procesan a partir de precursores en forma de horquilla más largos (aprox. 75 nt) denominados pre-miR. Los microARN se ensamblan en complejos denominados miRISC y reconocen a sus dianas mediante complementariedad antisentido. Si los microARN coinciden en un 100 % con su diana, es decir, la complementariedad es completa, el ARNm diana se escinde y el miR actúa como un ARNiP. Si la coincidencia es incompleta, es decir, la complementariedad es parcial, entonces, se bloquea la traducción del ARNm diana.

El término "expresión", como se usa en el presente documento, se refiere a la transcripción y/o acumulación de moléculas de ARN en una célula o una muestra de tejido.

En el presente contexto, los términos "nivel de expresión de un miR", "nivel de expresión de miR" y "nivel de un miR" se usan como sinónimos para medir la "cantidad de un miR específico" que se detecta en la muestra. La "cantidad de un miR específico" se puede expresar en mediciones absolutas, relativas o bien normalizadas y se refiere a los valores obtenidos mediante procedimientos cuantitativos, así como cualitativos. Una medición preferente en particular de la "cantidad de un miR específico" es el valor de punto de corte (Cp) obtenido mediante qRT-PCR como se describe a continuación y en los ejemplos, pero la "cantidad" también se puede cuantificar mediante PCR digital, o diversos

procedimientos de secuenciación de nueva generación. En determinadas situaciones, por ejemplo, cuando se usan proporciones de niveles de expresión de miR para calcular una puntuación pronóstica, son suficientes los niveles de expresión determinados absolutos de los miR. Sin embargo, como alternativa a realizar determinaciones en base al nivel de expresión absoluto de los miR, las determinaciones se pueden basar en los niveles de expresión normalizados de los miR.

Los niveles de expresión se normalizan corrigiendo el nivel de expresión absoluto de un miR comparando su expresión con la expresión de un gen que se expresa de forma constitutiva o casi constitutiva. Los genes adecuados usados a menudo para la normalización incluyen genes de mantenimiento, tales como el gen actina. En el presente estudio se usa miR-151a-5p para la normalización.

Como se usa en el presente documento, los términos: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p, miR-625-3p y miR-151a-5p se refieren a las secuencias de miR humanas encontradas en la publicación de la base de datos de registro de miARN, 12.0 o posterior, y facilitadas por el Sanger Institute, Reino Unido, así como sus equivalentes animales. Véase también la tabla 9.

El término "pronóstico" se define en el presente documento para que englobe los siguientes procedimientos de forma individual o bien acumulativa dependiendo del contexto clínico: predecir el desenlace probable de un estado de enfermedad o determinar la posibilidad de recuperación de una enfermedad como se indica mediante la naturaleza y síntomas de un caso.

En el presente contexto, "perfil de expresión de referencia" designa el nivel de expresión de dos o más miR seleccionados de un grupo de 13 miR que consiste en: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p en muestras de células obtenidas de pacientes de cáncer de próstata cuyo cáncer no progresó significativamente en un periodo de 5 años después del muestreo.

El término "solidez" en conexión con el clasificador de miR se usa en el presente documento para describir un clasificador que proporciona un resultado relativamente similar con respecto a la evaluación del pronóstico de un paciente de cáncer de próstata, aunque se usen metodologías de muestreo y cuantificación un tanto diferentes.

"UniRT" es un procedimiento de qRT-PCR comercializado por Exiqon A/S. El procedimiento y su funcionamiento se describen en el ejemplo 1 y 7 y en la solicitud de patente danesa PA 2009 00156, el documento EP2391736 y Mestdagh *et al.* Nat Methods. Agosto de 2014; 11 (8):809-15.

Los modos de realización de la presente invención se describen a continuación, solo a modo de ejemplo.

#### **DIVULGACIÓN DETALLADA DE LA INVENCION:**

El problema técnico que subyace a la invención es la provisión de un procedimiento *in vitro* alternativo para evaluar el pronóstico de un paciente de cáncer de próstata. En particular, un procedimiento que puede complementar a, o quizás incluso sustituir a, otros indicadores pronóstico disponibles actualmente (principalmente la prueba de PSA, el número de biopsias positivas, el estadio TNM clínico y la puntuación de Gleason).

En el ejemplo 1, se describe la aplicación de dos enfoques estadísticos diferentes (1. una validación cruzada dejando uno fuera (LOOCV) y 2. combinaciones de microARN significativos en el análisis de regresión de Cox univariante) para llegar a cinco clasificadores de miR pronósticos, véase la tabla 1, 2 y 3. Conjuntamente, los dos enfoques proporcionaron clasificadores, cuando se combinaron, que consistían en 13 miR, a saber: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p (tabla 3).

De forma sorprendente, se ha descubierto que se pueden usar tantos como 2 miR extraídos de este grupo de 13 miR para diferenciar entre casos recidivantes frente a no recidivantes, y predecir el tiempo hasta la recidiva bioquímica después de la prostatectomía radical (RBQ; véanse el ejemplo 1+6, la figura 5 y la tabla 8). En consecuencia, un aspecto de la presente invención es un procedimiento *in vitro* para evaluar el pronóstico de un paciente de cáncer de próstata, que comprende medir el nivel de expresión de al menos dos miR seleccionados del grupo de miR que consiste en: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p en ARN de una muestra de tejido prostático obtenida de dicho paciente, en el que un nivel de expresión distinto de dichos al menos 2 miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.

Los 8 miARN con el impacto más alto en la precisión en el pronóstico del estudio de descubrimiento y validación que usa el procedimiento de clasificación de máxima verosimilitud con validación cruzada dejando uno fuera (LOOCV) (véase el ejemplo 1) parece interesante. Por tanto, un modo de realización de la presente invención es un procedimiento *in vitro* para evaluar el pronóstico de un paciente de cáncer de próstata (es decir, el riesgo de que un

5 sujeto tenga un cáncer de próstata de gran malignidad) que comprende medir el nivel de expresión de al menos dos miR seleccionados del grupo de miR que consiste en: miR-106a-5p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-374b-5p y miR-615-3p en una muestra de células obtenidas de dicho paciente, en el que un nivel de expresión distinto de dichos al menos 2 miR, en comparación con un perfil de expresión de referencia, indica un pronóstico menos favorable.

10 Para cumplir con los requisitos de la clínica cotidiana, un procedimiento de pronóstico debe proporcionar resultados relativamente similares con respecto a la diferenciación entre pacientes con un pronóstico bueno o uno malo, aunque se usen metodologías de muestreo y cuantificación un tanto diferentes. El procedimiento debe ser sólido. De forma alentadora, a pesar del hecho de que las muestras de tumores en la cohorte 3 fueron de diferente origen nacional (EE. UU.), se muestrearon de una manera diferente (ultracongeladas), se sometieron a diferentes procedimientos de extracción de ARN y se analizaron mediante una diferente plataforma de detección de la expresión de los microARN (micromatriz de miARN humanos de Agilent 2.0) y diferentes características de cohorte (la cohorte 3 fue, en general, de menor malignidad y tuvo menos acontecimientos de recidiva que la cohorte 1), el clasificador pronóstico de 8 microARN funcionó igualmente bien en esta cohorte, lo que subraya la solidez del clasificador pronóstico de 8 microARN, véanse el ejemplo 2 y la tabla 4.

20 Medir 13 o incluso 8 miARN puede parecer impracticable para las pruebas de pronóstico destinadas a la clínica cotidiana. En el ejemplo 1+3-6, tabla 5-8, se muestra una serie de clasificadores de miARN reducidos con valor pronóstico extraídos de esta lista de 13 miARN (tabla 3).

25 De forma sorprendente, estos clasificadores de miARN reducidos parecen incluso más sólidos que el clasificador pronóstico de 8 microARN. En consecuencia, otros aspectos de la presente invención son procedimientos *in vitro* para evaluar el pronóstico de un paciente de cáncer de próstata, que comprenden medir el nivel de expresión de: miR-10b-5p, miR-133a-3p, miR-23a-3p y miR-374b-5p; o el nivel de: miR-185-5p, miR-221-3p y miR-326; o el nivel de: miR-152-3p, miR-185-5p y miR-221-3p; o incluso el nivel de solo: miR-10b-5p y miR-374b-5p en una muestra de células obtenidas de dicho paciente, en los que un nivel de expresión distinto de al menos 2 miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.

30 La "cantidad de un miR específico" se puede expresar en mediciones absolutas, relativas o bien normalizadas y se refiere a los valores obtenidos mediante procedimientos cuantitativos, así como cualitativos.

35 Como alternativa a realizar determinaciones en base al nivel de expresión absoluto de los miR, las determinaciones se pueden basar en los niveles de expresión normalizados de los miR. Los niveles de expresión se normalizan corrigiendo el nivel de expresión absoluto de un miR comparando su expresión con la expresión de un gen que se expresa de forma constitutiva o casi constitutiva. Los genes de mantenimiento, tales como el gen actina, se usan a menudo para la normalización. Sin embargo, en el caso de los ARN cortos, tales como los miR, son preferentes determinados miARN invariables como normalizadores.

40 En consecuencia, en un modo de realización de la invención, los niveles de expresión son niveles de expresión normalizados. En el presente estudio se usan miR para la normalización. Usando el algoritmo NormFinder (10), se identificó a miR-151a-5p como el miARN expresado de forma más estable en el estudio de descubrimiento. La comparación de dos estrategias de normalización (media global y miR-151-5p) dio resultados muy similares en términos de los principales miARN expresados diferencialmente, así como en la clasificación general de los miARN. Por lo tanto, en un modo de realización de la invención, los niveles de expresión se normalizan al nivel de expresión de miR-151a-5p.

50 Cuando se aplica un ensayo pronóstico en la práctica, es ventajoso usar los valores del ensayo para calcular una puntuación pronóstica (P), lo que permite definir los valores de corte y evaluar el pronóstico de un paciente de cáncer de próstata en base a la puntuación pronóstica. Por tanto, un modo de realización de la presente invención es un procedimiento en el que la evaluación del pronóstico de los pacientes de cáncer implica detectar el nivel de dichos al menos dos miR en dicha muestra y calcular una puntuación pronóstica (P) en base a un conjunto de datos que comprende los datos del nivel de expresión de al menos dos miR.

55 El nivel de miR se puede cuantificar convenientemente mediante el procedimiento de reacción en cadena de la polimerasa mediada por transcriptasa inversa en tiempo real cuantitativa, qRT-PCR (17). Por tanto, en un modo de realización de la invención, el nivel de expresión de dichos miR se determina mediante el procedimiento de RT-QPCR, RT-qPCR o qRT-PCR, que son sinónimos.

60 Una medición preferente en particular de la "cantidad de un miR específico" es el valor de punto de corte (Cp) obtenido mediante qRT-PCR. Otra medición preferente de la "cantidad de un miR específico" es el valor "valor de ciclo umbral (Ct)" obtenido igualmente mediante qRT-PCR como se describe en los ejemplos. Las mediciones de Cp y Ct de la "cantidad de un miR específico" proporcionan mediciones aproximadamente similares, véase (12). La elección de Cp o Ct es en gran medida una cuestión de elección de la máquina a la que el ensayo se vinculó y sobre la que se realizó. Si la amplificación se realiza en un sistema de PCR en tiempo real LightCycler® 480 usando el programa informático Roche LC, la cantidad de un miR específico se expresa mediante el Cp. Si la amplificación se realiza en un

instrumento ABI Prism 7900HT de 384 pocillos de Applied Biosystems con programa informático proporcionado con él, la cantidad de un miR específico se expresa mediante el Ct. Lo siguiente se refiere al valor de Cp, pero también se aplica al valor de Ct y al valor de "ciclo de cuantificación" (Cq).

5 El valor de Cp se relaciona con el nivel, por ejemplo, de un miR específico, mediante la proporción:

$$\text{nivel de expresión de miRx (lineal)} \sim 2^{-Cp(\text{miRx normalizado})}$$

10 En la que  $Cp(miRx)$  designa la lectura de Cp de un instrumento qRT-PCR en tiempo real que detecta específicamente un miR específico llamado miRx. El ejemplo 1 describe en detalle dicho ensayo.

15 Los marcadores basados en proporción ofrecen un modelo de clasificador interesante debido a su independencia de la normalización de los datos. Se ha descubierto que un clasificador de proporciones extraído de los 13 miR (o los 11 miR de la "regresión de Cox") se podría usar como biomarcadores pronóstico independientes para la predicción del tiempo hasta la recidiva del cáncer de próstata, véase el ejemplo 5.

En consecuencia, cuando los valores de Cp se usan como cuantificadores de los niveles de miR, por ejemplo, la expresión:

$$20 \frac{(\text{nivel de miR23a-3p}) \times (\text{nivel de miR10b-5p})}{(\text{nivel de miR133a-3p}) \times (\text{nivel de miR374b-5p})}$$

es equivalente a:

$$25 Cp(miR23a-3p) + Cp(miR10b-5p) - Cp(miR133a-3p) - Cp(miR374b-5p)$$

En consecuencia, en un modo de realización de la invención, la puntuación pronóstica (P) de diagnóstico para el clasificador de 4 miR (miR10b-5p, miR-133a-3p, miR-23a-3p y miR-374b-5p) se calcula como:

$$30 P = XxC(miR23a-3p) + YxC(miR10b-5p) + ZxC(miR133a-3p) + WxC(miR374b-5p)$$

35 donde los coeficientes X, Y, Z y W se determinan mediante el análisis de regresión de acuerdo con la configuración particular, y C es el valor de ciclo umbral (Ct), o el valor de punto de corte (Cp), o el valor de "ciclo de cuantificación" (Cq), o cualquier valor similar a estos.

De forma sorprendente, se ha descubierto que un clasificador basado en proporción que consiste en solo dos de los cuatro miARN extraídos del grupo de 13 miR se podría usar como biomarcadores pronóstico independientes para la predicción de la recidiva del cáncer de próstata, véase el ejemplo 6.

40 En consecuencia, otro modo de realización de la presente invención es un procedimiento *in vitro* para evaluar el pronóstico de un paciente de cáncer de próstata, que comprende medir el nivel de expresión de los dos miR: miR-374b-5p y miR-10b-5p.

45 Cuando los miR se cuantifican mediante qRT-PCR y los valores de Cp se usan como cuantificadores de los niveles de miR, la expresión:

$$\frac{(\text{nivel de miR10b-5p})}{(\text{nivel de miR374b-5p})}$$

es equivalente a:

$$50 Cp(miR10b-5p) - Cp(miR374b-5p)$$

De forma similar, en un modo de realización de la invención, la puntuación pronóstica (P) de diagnóstico para el clasificador de 2 miR es

$$55 P = XxCp(miR10b-5p) + YxCp(miR374b-5p)$$

donde los coeficientes X e Y se determinan mediante el análisis de regresión, y C es el valor de ciclo umbral (Ct), o el valor de punto de corte (Cp), o el valor de "ciclo de cuantificación" (Cq), o cualquier valor similar a estos.

60 Se contemplan tanto tipos de regresión lineal como otros.

Por "aprendizaje de máquina" se hace referencia a un procedimiento que aprovecha los algoritmos informáticos que mejoran automáticamente a través de la experiencia, en la técnica, este procedimiento de mejora de los algoritmos se

denomina a menudo "entrenamiento". Se puede usar el aprendizaje de máquina para descubrir reglas generales en grandes conjuntos de datos, se puede usar el aprendizaje de máquina, por ejemplo, para extraer datos clínicos informativos de un conjunto de datos que comprende la expresión de miR en muestras cancerosas y no cancerosas de la próstata. Un tratado general del concepto de aprendizaje de máquina se puede encontrar en (Tom Mitchell, Machine Learning, McGraw Hill, 1997). En consecuencia, en un modo de realización de la invención, se alcanzó el algoritmo para calcular la puntuación pronóstica (P) de diagnóstico aplicando el aprendizaje de máquina.

Aunque el procedimiento de reacción en cadena de la polimerasa mediada por transcriptasa inversa en tiempo real cuantitativa, qRT-PCR, es el procedimiento de cuantificación preferente. Los ejemplos 2, 3, 4 y 6 muestran que la invención es compatible con la cuantificación basada en micromatrices y, en consecuencia, un modo de realización de la invención es un procedimiento en el que los niveles de expresión de dichos miR se miden mediante micromatrices.

Se contemplan otros procedimientos de cuantificación, tales como transferencias Northern, PCR digital cuantitativa y diversos procedimientos cuantitativos de secuenciación de nueva generación y, por tanto, modos de realización de la invención.

Otro aspecto de la invención es un kit para la evaluación *in vitro* del pronóstico de un paciente de cáncer de próstata, que comprende medir el nivel de expresión de al menos dos miR seleccionados del grupo de 13 miR que consiste en: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p en una muestra de ARN de células prostáticas obtenidas de dicho paciente, en el que un nivel de expresión distinto de dichos al menos 2 miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.

La invención se ilustra además en los siguientes ejemplos, tablas y figuras no limitantes.

## Leyendas

**Figura 1.** Análisis de supervivencia de Kaplan-Meier de la supervivencia sin recidiva (SSR) en base al clasificador pronóstico de 8 miARN (tabla 4) en muestras de prostatectomía radical (PR) de dos cohortes independientes. Los pacientes en la cohorte de entrenamiento 1 (A) se dividieron en grupos de bajo frente a alto riesgo después del análisis de ROC. Los pacientes en la cohorte de validación 3 (B) se dividieron en grupos de bajo/alto riesgo de acuerdo con el valor de corte (fracción) definido en la cohorte 1. Se dan valores de *p* significativos para la prueba del orden logarítmico bilateral. El estado de alto riesgo molecular, como se define mediante el clasificador pronóstico de 8 miARN, se asoció significativamente con la recidiva bioquímica precoz después de la PR.

**Figura 2.** Análisis de supervivencia de Kaplan-Meier de la supervivencia sin recidiva (SSR) en base al clasificador pronóstico de 8 miARN reducido (miR-185-5p+miR-221-3p+miR-152-3p, tabla 5) en muestras de prostatectomía radical (PR) de tres cohortes independientes. Los pacientes en la cohorte de entrenamiento 1 (A) se dividieron en grupos de bajo frente a alto riesgo después del análisis de ROC. Los pacientes en la cohorte de validación 2 (B) se dividieron en grupos de bajo/alto riesgo de acuerdo con el valor de corte (fracción) definido en la cohorte 1. Esto se hizo de la misma manera para la cohorte de validación externa 3 (Taylor *et al.* y Hieronymus *et al.*; C). Se dan valores de *p* significativos para la prueba del orden logarítmico bilateral. El estado de alto riesgo molecular, como se define mediante el clasificador pronóstico de 8 miARN reducido, se asoció significativamente con la recidiva bioquímica precoz después de la PR en tres cohortes independientes.

**Figura 3.** Análisis de supervivencia de Kaplan-Meier de la supervivencia sin recidiva (SSR) en base al clasificador pronóstico de 3 miARN (miR-185-5p+miR-221-3p+miR-326, tabla 6) en muestras de prostatectomía radical (PR) de tres cohortes independientes. Los pacientes en la cohorte de entrenamiento 1 (A) se dividieron en grupos de bajo frente a alto riesgo después del análisis de ROC. Los pacientes en la cohorte de validación 2 (B) se dividieron en grupos de bajo/alto riesgo de acuerdo con el valor de corte (fracción) definido en la cohorte 1. Esto se hizo de la misma manera para la cohorte de validación externa 3 (Taylor *et al.* y Hieronymus *et al.*; C). Se dan valores de *p* significativos para la prueba del orden logarítmico bilateral. El estado de alto riesgo molecular, como se define mediante el clasificador pronóstico de 3 miARN, se asoció significativamente con la recidiva bioquímica precoz después de la PR en tres cohortes independientes.

**Figura 4.** Análisis de supervivencia de Kaplan-Meier de la supervivencia sin recidiva (SSR) en base al modelo de proporción pronóstico de 4 miARN (proporción miR-10b-5p/miR-23a-3p/miR-133a-3p/miR-374b-5p, tabla 7) en muestras de prostatectomía radical (PR) de tres cohortes independientes. Los pacientes en la cohorte de entrenamiento 1 (A) se dividieron en grupos de bajo frente a alto riesgo después del análisis de ROC. Los pacientes en la cohorte de validación 2 (B) se dividieron en grupos de bajo/alto riesgo de acuerdo con el valor de corte absoluto definido en la cohorte 1. Esto se hizo de la misma manera para la cohorte de validación externa 3 (Taylor *et al.* y Hieronymus *et al.*; C). Se dan valores de *p* para la prueba del orden logarítmico bilateral. El estado de alto riesgo molecular, como se define mediante el modelo de proporción pronóstico de 4 miARN, se asoció significativamente con la recidiva bioquímica precoz después de la PR.

**Figura 5.** Análisis de supervivencia de Kaplan-Meier de la supervivencia sin recidiva (SSR) en base al modelo de

proporción pronóstico de 2 miARN (miR-10b-5p/miR-374b-5p, tabla 8) en muestras de prostatectomía radical (PR) de tres cohortes independientes. Los pacientes en la cohorte de entrenamiento 1 (A) se dividieron en grupos de bajo frente a alto riesgo después del análisis de ROC. Los pacientes en la cohorte de validación 2 (B) se dividieron en grupos de bajo/alto riesgo de acuerdo con el valor de corte absoluto definido en la cohorte 1. Esto se hizo de la misma manera para la cohorte de validación externa 3 (Taylor *et al.* y Hieronymus *et al.*; C). Se dan valores de  $p$  para la prueba del orden logarítmico bilateral. El estado de alto riesgo molecular, como se define mediante el modelo de proporción pronóstico de 2 miARN, se asoció significativamente con la recidiva bioquímica precoz después de la PR en tres cohortes independientes.

## 10 EJEMPLO 1

**Clasificadores pronóstico novedosos para la predicción de la recidiva del cáncer de próstata identificados mediante identificación de microARN del genoma completo.**

### 15 Objeto del estudio

Investigar el potencial como biomarcador pronóstico de la expresión de los microARN en el cáncer de próstata y construir un clasificador para la estratificación del riesgo.

### 20 Procedimientos

#### *Muestras*

Todas las muestras se recogieron en el Departamento de Urología y se obtuvieron del Institute of Pathology, Aarhus University Hospital, Dinamarca (de 1997 a 2005). La cohorte de entrenamiento (cohorte 1) consistió en 127 PR con intención curativa de cáncer de próstata clínicamente localizado y verificado histológicamente. Todas las muestras de tejido se evaluaron por un anatomopatólogo cualificado. Las regiones representativas con >90 % de tumor se marcaron en las secciones teñidas con hematoxilina y eosina (H y E), y se tomaron biopsias en sacabocados de los bloques FFPE correspondientes para la extracción de ARN como se describe en (Haldrup 2013). El ARN total se aisló de biopsias en sacabocados de 1,5 mm usando el miRNeasy FFPE Kit (Qiagen), de acuerdo con las instrucciones del fabricante. La proporción de absorbancia de 260/280 nm (proporción óptima: 2,0) se usó como garantía de calidad para las muestras de ARN. Las muestras de ARN con una proporción <1,75 se excluyeron de un análisis adicional. El ARN purificado se almacenó a -80 °C.

#### 35 *Perfil de expresión de los microARN*

Se realizó el perfil de expresión de los microARN (todos los reactivos de Exiqon) en Exiqon A/S, Vedbaek, Dinamarca, usando la plataforma miRCURY LNA™ Universal RT microRNA PCR. En resumen, se transcribieron de forma inversa 40 ng de ARN total en reacciones de 40 µl usando el miRCURY LNA™ microRNA PCR, Polyadenylation and cDNA synthesis kit II. El ADNc se diluyó 100x y se analizó en reacciones de PCR de 10 µl. Para la cohorte 1 (entrenamiento), se analizaron los niveles de expresión relativos de 752 miARN usando microRNA Ready-to-Use PCR, Human panel I+II, V3R, en placas de PCR de 384 pocillos. Los controles negativos (sin molde en la reacción de transcripción inversa) se efectuaron en paralelo. La amplificación se realizó usando el sistema de PCR en tiempo real LightCycler® 480 (Roche) y la mezcla maestra ExiLENT SYBR® Green. Las curvas de amplificación se analizaron usando el programa informático Roche LC para la determinación de los valores de ciclo de cuantificación (Cq) (mediante el procedimiento de la 2.<sup>a</sup> derivada (Livak, 2001) y para los análisis de la curva de fusión). Los microARN detectados con <3 Cq menos que el control negativo o para los que todos los valores de Cq excedieron de 37 en todas las muestras se excluyeron de los análisis adicionales (cohorte 1: 97 miARN, cohorte 2: sin miARN).

Para la cohorte 1 (entrenamiento), los datos se normalizaron a la media global (es decir, la media para todos los miARN detectados en todas las muestras; aquí n=61, véase la tabla 10), que se informó previamente como el mejor procedimiento de normalización para los datos de qRT-PCR que implican numerosos ensayos (15). Usando el algoritmo NormFinder (10), se descubrió que miR-151a-5p era un gen de normalización individual óptimamente estable en la cohorte 1. Las dos estrategias de normalización (normalización con miR-151a-5p y media global) dieron resultados muy similares en términos de los principales miARN expresados diferencialmente, así como en la clasificación general de los microARN en la cohorte 1, lo que respalda, por tanto, la validez de miR-151a-5p como un gen de normalización en un estudio de validación posterior con menos miARN sometidos a prueba. La normalización se realizó de acuerdo con la fórmula  $\Delta Cq = Cq_{\text{Factor de normalización}} - Cq_{\text{miARN}}$ . Las diferencias en los niveles de expresión se calculan como  $\Delta\Delta Cq = \Delta Cq_{\text{grupo 1}} - \Delta Cq_{\text{grupo 2}}$ . Para convertir esto en un valor de cambio, se usa la fórmula  $2^{-\Delta\Delta Cq}$ . El número recíproco, multiplicado por un factor de -1, se usó para los miARN regulados por disminución.

#### *Análisis estadísticos*

A menos que se establezca de otro modo, los análisis estadísticos se llevaron a cabo en STATA, versión 11 (StataCorp, Texas, EE. UU.). Para todos los análisis, los valores de  $p < 0,05$  se consideraron estadísticamente significativos.

Análisis de la expresión: Se usó la prueba de Shapiro Wilk para evaluar si los datos estaban distribuidos normalmente. Como varios de los microARN no lo estaban, se usó una prueba del orden con signo de Wilcoxon estadística no paramétrica para las comparaciones entre pares de la expresión de los microARN entre los diferentes grupos. Los valores de  $p$  se corrigieron para las múltiples pruebas usando el procedimiento de Benjamini-Hochberg (11).

#### Construcción del clasificador:

Los procedimientos de clasificación de máxima verosimilitud con validación cruzada dejando uno fuera (LOOCV) se han entrenado y sometido a prueba en el programa informático 'R' Bioconductor, versión 3.0.0 (Bell Laboratories, Lucent Technologies, <http://www.cran.r-project.org/>) como se describe anteriormente (13). Se usaron valores de expresión de los microARN normalizados de 127 (70 sin y 57 con recidiva bioquímica) muestras de PR para generar un clasificador de múltiples microARN pronóstico construido para clasificar la recidiva del cáncer de próstata. Solo se incluyeron los microARN expresados en al menos un 70 % de las muestras ( $n=236$  miARN) en el análisis. El funcionamiento de los clasificadores se sometió a prueba usando una prueba de  $\chi^2$  y cálculos de sensibilidad y especificidad (tabla 1). Para generar un modelo unificado que comprenda los 8 miARN en el clasificador pronóstico de 8 microARN (descrito más adelante) para los análisis de SSR, se ponderó cada miARN mediante los coeficientes de regresión estimados en el modelo de riesgos proporcionales multivariante, y se calculó una suma ponderada combinada para el clasificador de miARN.

Además, se usaron dos enfoques diferentes para entrenar a los clasificadores de miARN pronóstico de los 11 miARN significativos en el análisis de regresión de Cox univariante en la cohorte de entrenamiento ( $n=126$ ; uno de los 127 pacientes con PR se excluyó en los análisis de SSR debido al tratamiento endocrino posoperatorio; tabla 2). En un enfoque, se ponderó cada miARN incluido en dichos clasificadores mediante los coeficientes de regresión estimados en el modelo de riesgos proporcionales de Cox multivariante, y se calculó una suma ponderada combinada para el clasificador de miARN. En el segundo enfoque, se generaron clasificadores de miARN basados en proporción en base a los valores de Cq sin procesar para cada miARN incluido en el modelo.

Valor pronóstico: Para los análisis de supervivencia sin recidiva (SSR), se usó la recidiva bioquímica (RBQ; valor de corte de PSA  $\geq 0,2$  ng/ml, en base a la práctica clínica local) como criterio de valoración. Los pacientes que no habían experimentado una RBQ fueron objeto de censura estadística en su última medida de PSA normal. El valor pronóstico de la expresión de los microARN y la suma ponderada de los clasificadores basados en microARN se evaluaron mediante el análisis de Kaplan-Meier y la prueba del orden logarítmico bilateral, y mediante análisis de regresión de Cox univariantes y multivariantes como variables continuas, así como dicotómicas. Para el análisis de la expresión de los microARN y los clasificadores basados en microARN como variables dicotómicas, los pacientes de la cohorte 1 se dividieron en grupos de expresión alta y baja usando un valor de corte determinado después del análisis de ROC del estado sin recidiva/con recidiva. Todos los parámetros clinicopatológicos significativos en el análisis univariante se incluyeron en los análisis multivariantes. Las variables que no respondieron en el análisis multivariante se excluyeron del modelo multivariante final a través de selección regresiva gradual. El supuesto de riesgos proporcionales se verificó mediante la función de distribución de supervivencia log-negativa-log para todas las variables. La precisión en el pronóstico de los microARN y clasificadores se evalúa usando el índice de concordancia de Harrell (índice C); definido como la proporción de todos los pares de pacientes en los que las predicciones y los desenlaces son concordantes, como se establece mediante un análisis de regresión de Cox univariante y multivariante.

## **Resultados**

Para investigar el potencial pronóstico de la expresión de los microARN en el cáncer de próstata, se realizó inicialmente un análisis de supervivencia sin recidiva (SSR) bioquímica para 45 miARN candidatos que se descubrió que estaban significativamente desregulados en la comparación entre pares de pT2 frente a pT3-4, puntuación de Gleason baja frente a alta y/o tumores recidivantes frente a no recidivantes en la cohorte 1 (datos no mostrados). Mediante un análisis de regresión de Cox univariante, 11 de estos miARN se asociaron significativamente con el tiempo de SSR en la cohorte 1 ( $n=126$ ) (tabla 2). Más específicamente, la expresión alta de miR-10b-5p, miR-23a-3p, miR-185-5p, miR-615-3p y miR-625-3p, y la expresión baja de miR-30d-3p, miR-133a-3p, miR-193a-5p, miR-221-3p, miR-326 y miR-374b-5p, se asoció con la recidiva bioquímica (RBQ) precoz en esta cohorte. Sin embargo, ninguno de estos 11 miARN siguió siendo significativo en un modelo multivariante que incluía factores clinicopatológicos habituales (PSA, estadio pT, puntuación de Gleason y estado de los márgenes; datos no mostrados) y su valor pronóstico en un análisis univariante en la cohorte 1, en general, no se pudo confirmar en la cohorte 2 y la cohorte 3 ( $n = 99$ ) independientes (tabla 2; las cohortes se describen en los ejemplos 2 y 3).

En su lugar, se investigó si una combinación de varios microARN en clasificadores pronóstico podría mejorar la predicción del tiempo de SSR frente a microARN individuales. Los datos se analizaron en paralelo mediante procedimientos estadísticos: 1) validación cruzada dejando uno fuera (LOOCV; tabla 1) y 2) combinación (modelos ponderados y modelos basados en proporción) de los 11 microARN significativos en el análisis de regresión de Cox univariante en la cohorte 1 (tabla 2) (ambos procedimientos como se describe en la sección de procedimientos). Usando estos dos procedimientos, se descubrió una lista de 13 microARN candidatos con potencial pronóstico (tabla 3).

**Construcción de un clasificador conforme a LOOCV:**

**Procedimiento 1):** Se ha investigado si las combinaciones de varios microARN podrían incrementar la precisión de la predicción del riesgo de RBQ después de la PR. En consecuencia, se han usado los datos de expresión normalizados para 236 microARN detectados en más de un 70 % de las muestras en la cohorte 1. Se ha construido un clasificador de microARN pronóstico de máxima verosimilitud para distinguir a los 70 pacientes sin recidiva de los 57 pacientes con recidiva. Un clasificador que comprendía 8 microARN fue el más preciso para la clasificación de las muestras de PR en grupos sin recidiva/con recidiva (tabla 1). Por tanto, miR-615-3p, miR-185-5p, miR-23a-3p, miR-374b-5p, miR-193a-5p, miR-221-3p, miR-106a-5p y miR-152-3p se usaron en al menos un 70 % de los bucles de validación cruzada y constituyeron el clasificador pronóstico de 8 microARN final:

***Clasificador 1. Clasificador de 8 microARN (basado en la validación cruzada dejando uno fuera)***

El clasificador pronóstico de 8 microARN identificado mediante el procedimiento 1) miR-615-3p, miR-185-5p, miR-23a-3p, miR-374b-5p, miR-193a-5p, miR-221-3p, miR-106a-5p y miR-152-3p (tabla 1), clasificó correctamente un 74,6 % de las muestras recidivantes frente a no recidivantes en la cohorte 1 ( $p < 0,0001$ , prueba de la  $\chi^2$ ).

Para evaluar el funcionamiento del clasificador pronóstico de 8 microARN para la predicción del tiempo hasta la RBQ (en lugar del estado de RBQ), se calculó una suma ponderada de la expresión de todos los 8 microARN en el clasificador pronóstico. Luego, para evaluar el valor pronóstico del clasificador pronóstico de 8 miARN, se realizaron análisis de SSR basados en PSA. En la cohorte 1, el clasificador pronóstico (analizado como una variable continua de la suma ponderada) se asoció significativamente con un corto tiempo de SSR en el análisis de regresión de Cox univariante ( $p < 0,001$ ; tabla 4, arriba). El clasificador pronóstico de 8 miARN siguió siendo significativo en un modelo multivariante conjuntamente con la puntuación de Gleason, el estado de los márgenes quirúrgicos y el PSA preoperatorio ( $p < 0,001$ ; tabla 4, arriba). Cabe destacar que, cuando el clasificador pronóstico de 8 miARN se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,73 a 0,79 en la cohorte 1, lo que sugiere un funcionamiento moderadamente mejorado del modelo pronóstico combinado. Finalmente, los análisis de Kaplan-Meier demostraron una asociación significativa entre el clasificador pronóstico de 8 miARN y la corta SSR en la cohorte 1 (figura 1A;  $p < 0,0001$ , prueba del orden logarítmico).

En conclusión, se ha construido un clasificador pronóstico de 8 miARN, que fue un factor pronóstico independiente significativo de la RBQ después de la PR en la cohorte 1.

***Construcción de clasificadores reducidos******Clasificador 2: clasificador de 3 microARN (clasificador pronóstico de 8 miARN reducido generado a partir del clasificador de 8 microARN anterior)***

Usando una exclusión gradual de los 8 microARN en la tabla 1, se identificó una combinación de 3 microARN de miR-152-3p, miR-185-5p y miR-221-3p. La suma ponderada combinada de estos 3 microARN fue, como el clasificador pronóstico de 8 microARN, significativa tanto en los análisis de regresión de Cox uni como multivariantes (tabla 5, arriba). El clasificador pronóstico de 8 miARN reducido siguió siendo significativo en un modelo multivariante conjuntamente con la puntuación de Gleason y el estado de los márgenes quirúrgicos ( $p < 0,013$ ; tabla 5, arriba). Además, cuando el clasificador pronóstico de 8 miARN reducido se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,69 a 0,72 en la cohorte 1, lo que sugiere un funcionamiento moderadamente mejorado del modelo pronóstico combinado. Finalmente, los análisis de Kaplan-Meier demostraron una asociación significativa entre el clasificador pronóstico de 8 miARN reducido y la corta SSR en la cohorte 1 (figura 2A;  $p = 0,0001$ , prueba del orden logarítmico). En conclusión, se ha construido un clasificador pronóstico de 8 miARN reducido, que fue un factor pronóstico independiente significativo de la RBQ después de la PR en la cohorte 1.

**Construcción de un clasificador mediante análisis de regresión de Cox univariante:**

**Procedimiento 2):** Mediante un análisis de regresión de Cox univariante de la expresión de los microARN (analizada como variables continuas) en la cohorte 1 (cohorte de entrenamiento,  $n=126$ ), el corto tiempo de SSR se asoció de forma estadísticamente significativa (valores de  $p$  que variaban de  $p= 0,049$  a  $p < 0,001$ ) con la expresión de 11 de los 45 microARN analizados (tabla 2). Por tanto, la expresión alta de miR-10b-5p, miR-23a-3p, miR-185-5p, miR-615-3p y miR-625-3p, así como la expresión baja de miR-30d-3p, miR-133a-3p, miR-193a-5p, miR-221-3p, miR-326 y miR-374b-5p se asociaron significativamente con la RBQ precoz. Además, los factores pronóstico clinicopatológicos habituales establecidos: PSA alto, puntuación de Gleason alta, estadio T avanzado y estado de los márgenes quirúrgicos positivos se asociaron significativamente con la SSR, lo que indica que esta cohorte es representativa (tabla 2). A continuación, se investigó si una combinación de los 11 microARN podría mejorar la predicción del tiempo hasta la SSR frente a miARN individuales. Se usaron dos enfoques diferentes para entrenar a los clasificadores de miARN pronóstico de los 11 miARN significativos en el análisis de regresión de Cox univariante en la cohorte de entrenamiento. En un enfoque, se ponderó cada miARN incluido en el clasificador mediante los coeficientes de

regresión estimados en el modelo de riesgos proporcionales de Cox multivariante, y se calculó una suma ponderada combinada para el clasificador de miARN. En el segundo enfoque, se generaron clasificadores de miARN basados en proporción en base a los valores de Cq sin procesar para cada miARN incluido en el modelo.

5 **Clasificador 3:** *Clasificador de 3 microARN (clasificador basado en la suma ponderada de microARN significativos en el análisis de regresión de Cox univariante)*

10 Usando una exclusión gradual, se identificó una combinación de 3 microARN; miR-185-5p miR-221-3p y miR-326 (analizada como una suma ponderada combinada de estos 3 microARN) asociada significativamente con el corto tiempo de SSR en el análisis de regresión de Cox univariante en la cohorte 1 ( $p < 0,001$ ) y también siguió siendo significativa ( $p = 0,031$ ) en un modelo multivariante ajustado para el estadio pT, la puntuación de Gleason, el estado de los márgenes quirúrgicos y el PSA preoperatorio (tabla 6, arriba). Cabe destacar que la adición del clasificador pronóstico de 3 microARN a un modelo multivariante que incluía solo factores clinicopatológicos, incrementó la precisión predictiva (estimada mediante el índice C de Harrell) de 0,72 a 0,74 en la cohorte 1 (tabla 3, arriba), lo que sugiere un funcionamiento moderadamente mejorado. Además, los análisis de Kaplan-Meier mostraron una asociación significativa entre el clasificador pronóstico de 3 microARN y la SSR en la cohorte 1 ( $p = 0,0005$ , prueba del orden logarítmico, figura 3A). En resumen, se ha entrenado y sometido a prueba con éxito un clasificador pronóstico de 3 microARN (miR-185-5p+miR-221-3p+miR-326) novedoso que ha predicho el tiempo hasta la RBQ después de la PR independientemente de los parámetros clinicopatológicos habituales en la cohorte 1.

15 **Clasificador 4:** *clasificador de 4 microARN (basado en proporciones de cuatro microARN de los 11 microARN significativos en el análisis de regresión de Cox univariante)*

20 Los biomarcadores candidatos basados en proporciones de los valores de Cq sin procesar de los microARN serían favorables en una prueba clínica, puesto que evita la dependencia de los factores de normalización. Se usaron los 11 microARN significativos en el análisis univariante en la cohorte 1 (tabla 2) para construir biomarcadores de microARN basados en proporción candidatos. El objeto fue construir un simple modelo basado en una proporción 2:2 de miARN a partir de un subconjunto de los 11 miARN candidatos que fueron significativos en el análisis de SSR univariante en la cohorte 1 (tabla 2). Para garantizar la solidez del modelo de proporción, se priorizaron los miARN que se expresaban en al menos 125 de las 126 muestras de PR en la cohorte 1 (entrenamiento) y que tenían los valores de Cq más bajos. Además, para incrementar el rango dinámico, se separaron los miARN regulados por incremento y por disminución en el denominador y en el numerador.

25 Usando estos criterios, se desarrolló un modelo de proporción pronóstico de 4 miARN (miR-10b-5p/miR-23a-3p)/(miR-133a-3p/miR-374b-5p) que se asoció significativamente con la RBQ precoz en la cohorte 1 mediante Kaplan-Meier ( $p < 0,0001$ ; figura 4A), así como mediante análisis de regresión de Cox uni ( $p < 0,001$ ) y multivariante ( $p = 0,001$ ) (tabla 7, arriba). Además, cuando el modelo de proporción pronóstico de 4 microARN se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,73 a 0,76 en la cohorte 1, lo que sugiere un funcionamiento moderadamente mejorado del modelo pronóstico combinado.

30 En conclusión, se ha construido un modelo de proporción pronóstico de 4 microARN, que fue un factor pronóstico independiente significativo de la RBQ después de la PR en la cohorte 1.

35 **Construcción del modelo de proporción reducido**

**Clasificador 5:** *clasificador de 2 microARN (basado en proporciones de solo dos microARN reducido del clasificador de 4 microARN anterior)*

40 El clasificador basado en proporción de 4 microARN anterior se redujo además para simplificar este modelo para un potencial uso clínico en el futuro, se redujo a un modelo de proporción de 2 miARN (miR-10b-5p/miR-374b-5p). En la cohorte 1, este modelo de proporción pronóstico de 2 miARN fue un factor pronóstico significativo del tiempo hasta la RBQ en el análisis de regresión de Cox univariante ( $p = 0,001$ ), así como multivariante ( $p < 0,001$ ), que incluía variables clinicopatológicas habituales (tabla 8, arriba). Cabe destacar que la adición del modelo de proporción pronóstico de 2 miARN a un modelo multivariante que incluía solo factores clinicopatológicos, incrementó la precisión predictiva (estimada mediante el índice C de Harrell) de 0,73 a 0,77 en la cohorte 1, lo que sugiere un funcionamiento mejorado. Finalmente, los análisis de Kaplan-Meier también demostraron una asociación significativa entre el modelo de proporción pronóstico de 2 miARN y el tiempo de SSR en la cohorte 1 ( $p = 0,0259$ ; prueba del orden logarítmico; figura 5A).

45 En conclusión, se ha construido un modelo de proporción pronóstico de 2 miARN, que fue un factor pronóstico independiente significativo del tiempo hasta la RBQ después de la PR en la cohorte 1.

50 **Conclusión**

55 Se han creado clasificadores pronóstico que comprenden 2-8 microARN que predijeron el tiempo hasta la RBQ

después de la PR en una gran cohorte de cáncer de próstata independientemente de las variables clinicopatológicas habituales.

## EJEMPLO 2

### Validación del clasificador pronóstico de 8 microARN para la predicción del riesgo de recidiva en una cohorte de validación externa

#### Objeto del estudio

Validar el clasificador de 8 microARN (descrito en el ejemplo 1) en una cohorte de pacientes con PR independiente.

#### Procedimientos

##### *Muestras y datos de expresión de los miARN*

Se ha investigado el funcionamiento del potencial pronóstico del clasificador de 8 microARN pronóstico en el *conjunto de datos de expresión de los microARN GSE21036* externo de Taylor *et al* disponible en la página web de GEO ((Taylor *et al.* y Hieronymus *et al.*). Este conjunto de datos incluyó 99 muestras de tejido de PR ultracongeladas (cohorte 3). Se midió la expresión de 368 microARN usando Human microRNA Microarray 2.0 de Agilent ((Taylor *et al.* y Hieronymus *et al.*).

##### *Análisis estadísticos*

Los análisis estadísticos se llevaron a cabo en STATA, versión 11 (StataCorp, Texas, EE. UU.). Para todos los análisis, los valores de  $p < 0,05$  se consideraron estadísticamente significativos.

Precisión en el pronóstico: Se usó la recidiva bioquímica (RBQ; valor de corte de PSA  $\geq 0,2$  ng/ml, en base a la práctica clínica local) como criterio de valoración. Los pacientes que no habían experimentado una RBQ fueron objeto de censura estadística en su última medida de PSA normal. El valor pronóstico de la expresión de los microARN y del clasificador de 8-microARN se evaluó mediante el análisis de Kaplan-Meier y la prueba del orden logarítmico bilateral, y mediante análisis de regresión de Cox univariantes y multivariantes como variables continuas, así como dicotómicas. Para la validación del clasificador de 8 microARN, los pacientes en la cohorte 3 se dividieron en dos grupos usando el valor de corte (fracción) definido en la cohorte 1 (valor de corte determinado después del análisis de ROC del estado sin recidiva/con recidiva). El supuesto de riesgos proporcionales se verificó mediante la función de distribución de supervivencia log-negativa-log para todas las variables. Para las pruebas multivariantes, se incluyeron todos los parámetros clinicopatológicos significativos en el análisis univariante. Las variables que no respondieron en el análisis multivariante se excluyeron del modelo multivariante final a través de selección regresiva gradual. El estadio T patológico se dicotomizó en enfermedad localizada (T2) y localmente avanzada (T3-4). La puntuación de Gleason se agrupó en dos categorías que contenían las puntuaciones 5-6 (baja) y 7-10 (alta), respectivamente. El estado de los márgenes quirúrgicos se dicotomizó en márgenes negativos y positivos. Los niveles de PSA preoperatorio se analizaron como una variable continua. Se estimó la precisión en el pronóstico usando el índice de concordancia de Harrell.

#### Resultado

A pesar del hecho de que las muestras de tumores en la cohorte 3 fueron de diferente origen nacional (EE. UU.), se muestrearon de una manera diferente (ultracongeladas), se sometieron a diferentes procedimientos de extracción de ARN, se analizaron mediante una diferente plataforma de detección de la expresión de los microARN y diferentes características de cohorte (la cohorte 3 fue, en general, de menor malignidad y tuvo menos acontecimientos de recidiva que la cohorte 1), el clasificador pronóstico de 8 microARN funcionó igualmente bien en la cohorte externa, lo que subraya la solidez de este clasificador pronóstico de 8 microARN.

En la cohorte 3, el estado de alto riesgo molecular, como se define mediante el clasificador pronóstico de 8 miARN, se asoció significativamente con la RBQ precoz después de la PR tanto en el análisis de regresión de Cox univariante ( $p=0,003$ ) como multivariante ( $p=0,006$ ) (tabla 4, abajo). En esta cohorte, el estadio T patológico no tuvo valor pronóstico independiente significativo y la puntuación de Gleason se omitió del análisis debido a que no hubo ningún acontecimiento en uno de los grupos.

Además, se estimó la precisión predictiva del modelo multivariante mediante el índice C de Harrell. Cabe destacar que, cuando el clasificador pronóstico de 8 microARN se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,73 a 0,80 en la cohorte 3 (tabla 4, abajo), lo que respalda un funcionamiento moderadamente mejorado del modelo pronóstico combinado. Finalmente, los análisis de Kaplan-Meier demostraron una asociación significativa entre el clasificador pronóstico de 8 microARN y la corta SSR en la cohorte 3 (figura 1, abajo;  $p < 0,019$ , prueba del orden logarítmico).

## Conclusión

En conclusión, se ha construido un clasificador pronóstico de 8 microARN, que fue un factor pronóstico independiente significativo de la RBQ después de la PR en dos cohortes de pacientes independientes (ejemplo 1 y 2). Se validó el funcionamiento del clasificador pronóstico de 8 microARN; un factor pronóstico independiente significativo de la RBQ después de la PR en la cohorte 1 en una cohorte 3 de pacientes completamente diferente, lo que demuestra la capacidad del clasificador como una herramienta de pronóstico para el cáncer de próstata.

## EJEMPLO 3

### Validación de un clasificador pronóstico de 8 miARN reducido para la predicción del riesgo de recidiva en dos cohortes de validación independientes

#### Objeto del estudio

Validar el clasificador de 3 microARN (miR-152-3p/miR-185-5p/miR-221-3p); que es el clasificador reducido del clasificador de 8 microARN descrito en el ejemplo 1 y 2), en dos cohortes de pacientes con PR independientes: cohorte 2 y cohorte 3.

#### Procedimientos

##### Muestras

Se usaron dos cohortes para la validación. Cohorte 2: Para la identificación de miARN mediante RT-qPCR, se usaron muestras de tejido prostático incluido en parafina fijado con formol (FFPE). Todas las muestras se recogieron en el Departamento de Urología y se obtuvieron del Institute of Pathology, Aarhus University Hospital, Dinamarca (de 1997 a 2005). Para las muestras cancerosas, las regiones representativas con >90 % de tumor se marcaron en las secciones teñidas con hematoxilina y eosina (H y E), y se tomaron biopsias en sacabocados de los bloques FFPE correspondientes para la extracción de ARN como se describe en (Haldrup 2013). El ARN total se aisló de biopsias en sacabocados de 1,5 mm usando el miRNeasy FFPE Kit (Qiagen), de acuerdo con las instrucciones del fabricante. La proporción de absorbancia de 260/280 nm (proporción óptima: 2,0) se usó como garantía de calidad para las muestras de ARN. Las muestras de ARN con una proporción <1,75 se excluyeron de un análisis adicional. El ARN purificado se almacenó a -80 °C. Esta cohorte de validación (cohorte 2) consistió en 112 PR con intención curativa de cáncer de próstata clínicamente localizado y verificado histológicamente (para los análisis de SSR n=110; dos de los 112 pacientes con PR se excluyeron en los análisis de SSR debido al tratamiento endocrino posoperatorio). Cohorte 3: Consiste en un conjunto de datos de expresión de los microARN GSE21036 externo de Taylor *et al.* disponible en la página web de GEO ((Taylor *et al.* y Hieronymus *et al.*). Este conjunto de datos incluyó 99 muestras de tejido de PR ultracongeladas. Se midió la expresión de 368 microARN usando Human microRNA Microarray 2.0 de Agilent.

##### Perfil de expresión de los microARN

Se realizó el perfil de expresión de los microARN (todos los reactivos de Exiqon) en Exiqon A/S, Vedbaek, Dinamarca, usando la plataforma miRCURY LNA™ Universal RT microRNA PCR. En resumen, se transcribieron de forma inversa 40 ng de ARN total en reacciones de 40 µl usando el miRCURY LNA™ microRNA PCR, Polyadenylation and cDNA synthesis kit II. El ADNc se diluyó 100x y se analizó en reacciones de PCR de 10 µl. Para la cohorte 2 (validación), se analizaron 94 miARN seleccionados (incluido el gen de normalización miR-151a-5p) usando un panel de miRCURY LNA™ Universal RT Pick-&-Mix microRNA PCR (4x96 en 384 pocillos, listo para usar). Los controles negativos (sin molde en la reacción de transcripción inversa) se efectuaron en paralelo. La amplificación se realizó usando el sistema de PCR en tiempo real LightCycler® 480 (Roche) y la mezcla maestra ExiLENT SYBR® Green. Las curvas de amplificación se analizaron usando el programa informático Roche LC para la determinación de los valores de ciclo de cuantificación (Cq) (mediante el procedimiento de la segunda derivada (Livak, 2001) y para los análisis de la curva de fusión). Los microARN detectados con <3 Cq menos que el control negativo o para los que todos los valores de Cq excedieron de 37 en todas las muestras se excluyeron de los análisis adicionales (cohorte 2: sin miARN).

Se realizó la normalización usando el gen de normalización miR-151a-5p, establecido como un gen de normalización individual óptimamente estable en el ejemplo 1, identificado mediante el algoritmo NormFinder (10) como un gen de normalización individual óptimamente estable en ambas cohortes 1 y 2. La normalización se realizó de acuerdo con la fórmula  $\Delta Cq = Cq_{\text{Factor de normalización}} - Cq_{\text{miARN}}$ . Las diferencias en los niveles de expresión se calculan como  $\Delta\Delta Cq = \Delta Cq_{\text{grupo 1}} - \Delta Cq_{\text{grupo 2}}$ . Para convertir esto en un valor de cambio, se usa la fórmula  $2^{-\Delta\Delta Cq}$ . El número recíproco, multiplicado por un factor de -1, se usó para los miARN regulados por disminución.

##### Análisis estadístico:

Los análisis estadísticos se llevaron a cabo en STATA, versión 11 (StataCorp, Texas, EE. UU.). Para todos los análisis, los valores de  $p < 0,05$  se consideraron estadísticamente significativos.

Precisión en el pronóstico: Se usó la recidiva bioquímica (RBQ; valor de corte de PSA  $\geq 0,2$  ng/ml, en base a la práctica clínica local) como criterio de valoración. Los pacientes que no habían experimentado una RBQ fueron objeto de

censura estadística en su última medida de PSA normal. El valor pronóstico del clasificador se evaluó mediante el análisis de Kaplan-Meier y la prueba del orden logarítmico bilateral, y mediante análisis de regresión de Cox univariantes y multivariantes como variable continua, así como dicotómica. Para la validación del clasificador, los pacientes en la cohorte 2 y 3 se dividieron en dos grupos usando el valor de corte (fracción) definido en la cohorte 1 (valor de corte determinado después del análisis de ROC del estado sin recidiva/con recidiva). El supuesto de riesgos proporcionales se verificó mediante la función de distribución de supervivencia log-negativa-log para todas las variables. Para las pruebas multivariantes, se incluyeron todos los parámetros clinicopatológicos significativos en el análisis univariante. Las variables que no respondieron en el análisis multivariante se excluyeron del modelo multivariante final a través de selección regresiva gradual. El estadio T patológico se dicotomizó en enfermedad localizada (T2) y localmente avanzada (T3-4). La puntuación de Gleason se agrupó en dos categorías que contenían las puntuaciones 5-6 (baja) y 7-10 (alta), respectivamente. El estado de los márgenes quirúrgicos se dicotomizó en márgenes negativos y positivos. Los niveles de PSA preoperatorio se analizaron como una variable continua. Se estimó la precisión en el pronóstico usando el índice de concordancia de Harrell.

## 15 Resultados

Como se observa en la tabla 5, el estado de alto riesgo molecular, como se define mediante el clasificador pronóstico de 8 miARN reducido, se asoció significativamente con la RBQ precoz después de la PR tanto en el análisis de regresión de Cox univariante ( $p=0,001$ ) como multivariante ( $p=0,003$ ) en la cohorte 3 (tabla 5, abajo). En esta cohorte, el nivel de PSA preoperatorio no tuvo valor pronóstico independiente significativo y la puntuación de Gleason se omitió del análisis debido a que no hubo ningún acontecimiento en uno de los grupos. Además, el clasificador pronóstico de 8 miARN reducido se asoció significativamente con la RBQ precoz después de la PR en el análisis de regresión de Cox univariante ( $p=0,047$ ) y estuvo en el límite de la significación estadística en el multivariante ( $p=0,091$ ) en la cohorte 2 (tabla 5, centro). En esta cohorte, el estadio T patológico y la puntuación de Gleason tampoco tuvieron valor pronóstico independiente significativo.

Además, se estimó la precisión predictiva de cada uno de los modelos multivariantes mediante el índice C de Harrell. De forma interesante, cuando el clasificador pronóstico de 8 miARN reducido se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,74 a 0,78 en la cohorte 3 (tabla 5, abajo) y de 0,71 a 0,73 en la cohorte 2 (tabla 5, abajo), lo que respalda un funcionamiento moderadamente mejorado del modelo pronóstico combinado. Finalmente, los análisis de Kaplan-Meier demostraron una asociación significativa entre el clasificador pronóstico de 8 miARN reducido y la corta SSR en la cohorte 2 (figura 2B;  $p=0,014$ , prueba del orden logarítmico) y esta asociación también se validó con éxito en la cohorte 3 (figura 2C);  $p=0,019$ , prueba del orden logarítmico).

A pesar del hecho de que las muestras de tumores en la cohorte 3 fueron de diferente origen nacional (EE. UU.), se muestrearon de una manera diferente (ultracongeladas), se sometieron a diferentes procedimientos de extracción de ARN, se analizaron mediante una diferente plataforma de detección de la expresión de los microARN y diferentes características de cohorte (la cohorte 3 fue, en general, de menor malignidad y tuvo menos acontecimientos de recidiva que las cohortes 1 y 2), el clasificador pronóstico de 8 miARN reducido funcionó igualmente bien en la cohorte externa, lo que subraya la solidez de este clasificador pronóstico de 8 miARN reducido.

## Conclusión

Se ha construido un clasificador pronóstico de 8 miARN reducido, que fue un factor pronóstico significativo del tiempo hasta la RBQ después de la PR en tres cohortes de pacientes independientes (ejemplo 1 y 3). Se validó el funcionamiento del clasificador pronóstico de 8 miARN reducido; un factor pronóstico independiente significativo de RBQ después de la PR en la cohorte 1 en una cohorte 3 de pacientes completamente diferente, lo que demuestra la capacidad del clasificador como una herramienta de pronóstico para el cáncer de próstata.

## EJEMPLO 4

### Validación de un clasificador de 3 microARN pronóstico para la predicción del riesgo de recidiva en dos cohortes de validación independientes

#### Objeto del estudio

Validar el clasificador de 3 microARN (miR-185-5p+miR-221-3p+miR-326 en base a la suma ponderada de los miARN significativos en el análisis de regresión de Cox univariante, descrito en el ejemplo 1) en dos cohortes de pacientes independientes: cohorte 2 y cohorte 3.

#### Procedimientos

Descritos en el ejemplo 3

## Resultados

Como se observa en la tabla 6, el valor pronóstico del clasificador de 3 microARN se validó con éxito mediante análisis de regresión de Cox univariantes, así como multivariantes, en la cohorte 2 y la cohorte 3. El estado de alto riesgo molecular, como se define mediante el clasificador pronóstico de 3 miARN, se asoció significativamente con la RBQ precoz después de la PR tanto en el análisis de regresión de Cox univariante ( $p=0,006$  y  $p<0,001$ ) como multivariante ( $p=0,048$  y  $p=0,012$ ) en la cohorte 2 y la cohorte 3 respectivamente (tabla 6, centro y abajo). En la cohorte 2, el estado de los márgenes quirúrgicos y la puntuación de Gleason no tuvieron valor pronóstico independiente significativo, y en la cohorte 3, el estadio T patológico no fue significativa en el modelo multivariante, y la puntuación de Gleason se omitió del análisis debido a que no hubo ningún acontecimiento en uno de los grupos.

Además, se estimó la precisión predictiva de cada uno de los modelos multivariantes mediante el índice C de Harrell. Cabe destacar que, cuando el clasificador pronóstico de 3 microARN se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,73 a 0,75 en la cohorte 3 (tabla 6, centro) y de 0,74 a 0,80 en la cohorte 3 (tabla 5, abajo), lo que respalda un funcionamiento mejorado del modelo pronóstico combinado. Finalmente, los análisis de Kaplan-Meier demostraron una asociación significativa entre el clasificador pronóstico de 3 microARN y la corta SSR en la cohorte 2 (figura 3B;  $p=0,035$ , prueba del orden logarítmico) y esta asociación también se validó con éxito en la cohorte 3 (figura 3C);  $p=0,008$ , prueba del orden logarítmico).

A pesar del hecho de que las muestras de tumores en la cohorte 3 fueron de diferente origen nacional (EE. UU.), se muestrearon de una manera diferente (ultracongeladas), se sometieron a diferentes procedimientos de extracción de ARN, se analizaron mediante una diferente plataforma de detección de la expresión de los microARN y diferentes características de cohorte (la cohorte 3 fue, en general, de menor malignidad y tuvo menos acontecimientos de recidiva que las cohortes 1 y 2), el clasificador pronóstico de 3 microARN funcionó igualmente bien en la cohorte externa, lo que subraya la solidez de este clasificador pronóstico de 3 microARN.

## Conclusión

Se ha construido un clasificador pronóstico de 3 microARN con valor de predicción pronóstico independiente significativo para predecir el tiempo hasta la RBQ después de la PR en tres cohortes de pacientes independientes (ejemplo 1 y 4). Se validó el funcionamiento del clasificador pronóstico de 3 microARN; un factor pronóstico independiente significativo del tiempo hasta la RBQ después de la PR en la cohorte 1 en una cohorte 3 de pacientes completamente diferente, lo que demuestra la capacidad del clasificador como una herramienta de pronóstico para el cáncer de próstata.

## EJEMPLO 5

### Validación de un clasificador (basado en proporción) de 4 microARN pronóstico para la predicción del riesgo de recidiva en una cohorte de validación independiente

#### Objeto del estudio

Validar el clasificador de 4 microARN basado en proporción (miR-10b-5p/miR-23a-3p/miR-133a-3p/miR-374b-5p, descrito en el ejemplo 1) en dos cohortes de pacientes independientes: cohorte 2 y cohorte 3.

#### Procedimientos

Descritos en el ejemplo 3

#### Resultados

Como se observa en la tabla 7, el estado de alto riesgo molecular, como se define mediante el modelo de proporción pronóstico de 4 microARN, se asoció significativamente con la RBQ precoz después de la PR tanto en el análisis de regresión de Cox univariante ( $p=0,021$ ) como multivariante ( $p=0,024$ ) en la cohorte 2 (tabla 7, centro). En esta cohorte, el estadio T patológico y la puntuación de Gleason no tuvieron valor pronóstico independiente significativo. Además, el modelo de proporción pronóstico de 4 microARN se asoció significativamente con la RBQ precoz después de la PR en el análisis de regresión de Cox univariante ( $p=0,014$ ) y estuvo en el límite de la significación estadística en el multivariante ( $p=0,115$ ) en la cohorte 3 (tabla 7, abajo). En la cohorte 3, el estadio T patológico tampoco tuvo valor pronóstico independiente significativo y la puntuación de Gleason se omitió del análisis debido a que no hubo ningún acontecimiento en uno de los grupos.

Además, se estimó la precisión predictiva de cada uno de los modelos multivariantes mediante el índice C de Harrell. Cuando el modelo de proporción pronóstico de 4 microARN se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,73 a 0,75 en la cohorte 2 (tabla 7, centro) y de 0,73 a 0,78 en la cohorte 3 (tabla 7, abajo), lo que respalda un funcionamiento moderadamente mejorado del modelo pronóstico combinado. Finalmente, los análisis de Kaplan-Meier demostraron una asociación significativa

entre el modelo de proporción pronóstico de 4 microARN y la corta SSR en la cohorte 2 (figura 4B;  $p=0,0355$ , prueba del orden logarítmico) y esta asociación estuvo en el límite de la significación estadística en la cohorte 3 (figura 4C;  $p=0,1097$ , prueba del orden logarítmico).

## 5 Conclusión

En conclusión, se ha construido un modelo de proporción pronóstico de 4 microARN, que fue un factor pronóstico significativo de la RBQ después de la PR en tres cohortes de pacientes independientes y con valor pronóstico independientemente de los parámetros clinicopatológicos habituales en dos cohortes de pacientes de cáncer de próstata independientes (ejemplo 1 y 5).

## EJEMPLO 6

### 15 Validación de un clasificador (basado en proporción) de 2 microARN pronóstico para la predicción del riesgo de recidiva en dos cohortes de validación independientes

#### Objeto del estudio

20 Validar el clasificador de 2 microARN basado en proporción (miR10b-5p/ miR-374b-5p, descrito en el ejemplo 1) en dos cohortes de pacientes independientes: cohorte 2 y cohorte 3.

#### Procedimientos

25 Descritos en el ejemplo 3

#### Resultados

30 Como se observa en la tabla 8, el valor pronóstico del modelo de proporción pronóstico de 2 microARN se validó con éxito mediante análisis de regresión de Cox univariantes, así como multivariantes, en la cohorte 2 y la cohorte 3. El estado de alto riesgo molecular, como se define mediante el modelo de proporción pronóstico de 2 microARN, se asoció significativamente con la RBQ precoz después de la PR tanto en el análisis de regresión de Cox univariante ( $p=0,047$  y  $p=0,023$ ) como multivariante ( $p=0,047$  y  $p=0,015$ ) en la cohorte 2 y la cohorte 3 respectivamente (tabla 8, centro y abajo). En la cohorte 2, el estadio T patológico quirúrgico y la puntuación de Gleason no tuvieron valor pronóstico independiente significativo, y en la cohorte 3, el estadio T patológico no fue significativa en el modelo multivariante, y la puntuación de Gleason se omitió del análisis debido a que no hubo ningún acontecimiento en uno de los grupos.

35 Además, se estimó la precisión predictiva de cada uno de los modelos multivariantes mediante el índice C de Harrell. Cabe destacar que cuando el modelo de proporción pronóstico de 2 microARN se añadió a los modelos multivariantes que incluían solo factores clinicopatológicos, las precisiones predictivas se incrementaron de 0,73 a 0,74 en la cohorte 2 (tabla 8, centro) y de 0,73 a 0,79 en la cohorte 3 (tabla 8, abajo), lo que respalda un funcionamiento moderadamente mejorado del modelo pronóstico combinado. Finalmente, los análisis de Kaplan-Meier demostraron una asociación significativa entre el modelo de proporción pronóstico de 2 microARN y la corta SSR en la cohorte 2 (figura 5B;  $p=0,0487$ , prueba del orden logarítmico) y esta asociación también se validó con éxito en la cohorte 3 (figura 5C);  $p=0,0005$ , prueba del orden logarítmico).

45 A pesar del hecho de que las muestras de tumores en la cohorte 3 fueron de diferente origen nacional (EE. UU.), se muestrearon de una manera diferente (ultracongeladas), se sometieron a diferentes procedimientos de extracción de ARN, se analizaron mediante una diferente plataforma de detección de la expresión de los microARN y diferentes características de cohorte (la cohorte 3 fue, en general, de menor malignidad y tuvo menos acontecimientos de recidiva que las cohortes 1 y 2), el modelo de proporción pronóstico de 2 microARN funcionó igualmente bien en la cohorte externa, lo que subraya la solidez de este modelo de proporción pronóstico de 2 microARN.

#### Conclusión

55 Se ha construido un clasificador pronóstico de 2 microARN basado en proporción con valor independiente significativo para predecir el tiempo hasta la RBQ después de la PR en tres cohortes de pacientes independientes (ejemplo 1 y 6). Se validó el funcionamiento del modelo de proporción pronóstico de 2 microARN; un factor pronóstico independiente significativo del tiempo hasta la RBQ después de la PR en la cohorte 1 en una cohorte 3 de pacientes completamente diferente, lo que demuestra la capacidad y solidez del clasificador como una herramienta de pronóstico para el cáncer de próstata.

60

Tablas

**Tabla 1.** Los 8 miARN usados en al menos un 70 % de los bucles de validación cruzada al construir el clasificador de miARN pronóstico. Se muestran los resultados de la comparación entre pares de pacientes con y sin recidiva (ejemplo 1).

Regulado por incremento/disminución en la recidiva	Cohorte 1 (n=70 frente a n=57)				Cohorte 2 (n=62 frente a n=50)				Cohorte 3 (n=74 frente a n=25)			
	VC	Valor de p	Valor de p corrigido mediante BH	ABC (IC de un 95 %)	VC	Valor de p	Valor de p corrigido mediante BH	ABC (IC de un 95 %)	VC	Valor de p	Valor de p corrigido mediante BH	ABC (IC de un 95 %)
miR-615-3p	2,86	<b>0,001</b>	<b>0,002</b>	0,68 (0,59-0,78)	1,64	0,124	0,495	0,58 (0,48-0,69)	-1,14	0,646	0,652	0,53 (0,38-0,68)
miR-185-5p	1,54	<b>0,002</b>	<b>0,002</b>	0,68 (0,59-0,78)	1,12	0,631	0,804	0,53 (0,42-0,64)	1,16	<b>0,019</b>	0,115	0,66 (0,53-0,78)
miR-374b-5p	-1,49	<b>0,003</b>	<b>0,004</b>	0,66 (0,57-0,76)	-1,05	0,753	0,861	0,52 (0,40-0,62)	-1,16	0,074	0,198	0,62 (0,49-0,75)
miR-193a-5p	-1,36	<b>0,010</b>	<b>0,013</b>	0,64 (0,54-0,74)	1,14	0,266	0,532	0,56 (0,45-0,67)	1,07	0,263	0,526	0,58 (0,45-0,70)
miR-23a-3p	1,19	<b>0,008</b>	<b>0,013</b>	0,64 (0,54-0,74)	-1,16	0,215	0,532	0,57 (0,46-0,68)	1,08	0,652	0,652	0,53 (0,39-0,67)
miR-221-3p	-1,26	<b>0,025</b>	<b>0,030</b>	0,62 (0,52-0,72)	-1,53	<b>0,049</b>	0,388	0,61 (0,50-0,71)	-1,42	<b>0,029</b>	0,115	0,65 (0,52-0,77)
miR-106a-5p	-1,26	<b>0,028</b>	<b>0,031</b>	0,61 (0,51-0,72)	-1,05	0,704	0,804	0,52 (0,41-0,63)	-1,18	0,535	0,652	0,54 (0,40-0,68)
miR-152-3p	-1,39	0,069	0,069	0,58 (0,48-0,68)	-1,09	0,648	0,804	0,53 (0,42-0,63)	1,06	0,421	0,652	0,55 (0,41-0,69)

ABC, área bajo la curva de análisis de ROC; IC, intervalo de confianza; VC: valor de cambio calculado en base a la diferencia media entre los dos grupos. El número recíproco, multiplicado por -1, se usó para los miARN regulados por disminución. Los valores de p se calcularon con la prueba del orden con signo de Wilcoxon y se ajustaron para múltiples pruebas mediante el procedimiento de Benjamini- Hochberg (BH).  $p < 0,05$  se consideró significativo y aparece resaltado en negrita.

**Tabla 2.** Análisis de regresión de Cox univariantes del tiempo de supervivencia sin recidiva bioquímica (ejemplo 1). Se sometieron a prueba miARN significativos de la cohorte 1 en las cohortes 2 y 3.

Variable	Características	Cohorte 1, n=126, 56 con recidiva			Cohorte 2, n=110, 49 con recidiva			Cohorte 3, n=99, 25 con recidiva					
		HR (IC de un 95 %)	Valor de p	Valor de p corregido mediante BH	índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Valor de p corregido mediante BH	índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Valor de p corregido mediante BH	índice C <sup>a</sup>
Edad en el momento del diagnóstico	Continua	1,00 (0,94-1,05)	0,858	0,815	0,53	0,97 (0,93-1,03)	0,319	0,425	0,53	1,03 (0,98-1,10)	0,278	0,379	0,56
Estadio tumoral	pT2a-c frente a pT3a-c	3,12 (1,81-5,36)	<0,001	0,001	0,64	3,00 (1,69-5,30)	<0,001	0,001	0,64	4,05 (1,80-9,12)	0,001	0,005	0,68
Puntuación de Gleason	5-6 frente a 7-10	2,72 (1,51-4,93)	0,001	0,010	0,61	2,42 (1,23-4,73)	0,010	0,032	0,59	-	-	-	-
Estado de los márgenes quirúrgicos	Negativos frente a positivos	2,73 (1,59-4,70)	<0,001	0,001	0,63	3,37 (1,89-6,00)	<0,001	0,001	0,64	3,81 (1,70-8,54)	0,001	0,005	0,63
PSA preoperatorio	Continuo	1,05 (1,02-1,08)	<0,001	0,001	0,62	1,05 (1,03-1,07)	<0,001	0,001	0,72	1,09 (1,06-1,13)	<0,001	0,002	0,66
miR-374b-5p	Continuo	0,72 (0,60-0,86)	<0,001	0,005	0,63	1,02 (0,83-1,25)	0,831	0,878	0,55	0,57 (0,31-1,03)	0,062	0,133	0,62
miR-23a-3p	Continuo	2,57 (1,43-4,63)	0,002	0,017	0,63	0,80 (0,56-1,13)	0,205	0,298	0,54	1,33 (0,60-2,94)	0,476	0,549	0,50
miR-625-3p	Continuo	1,22 (1,06-1,40)	0,006	0,043	0,57	0,92 (0,82-1,04)	0,172	0,278	0,56	0,71 (0,41-1,24)	0,227	0,378	0,53
miR-615-3p	Continuo	1,16 (1,04-1,28)	0,007	0,044	0,63	1,04 (0,94-1,16)	0,439	0,540	0,53	1,00 (0,78-1,27)	0,973	0,973	0,50
miR-185-5p	Continuo	1,47 (1,10-1,97)	0,009	0,050	0,64	1,06 (0,85-1,33)	0,610	0,697	0,52	4,59 (1,52-13,90)	0,007	0,021	0,65
miR-133a-3p	Continuo	0,80 (0,67-0,96)	0,017	0,085	0,57	0,78 (0,66-0,92)	0,003	0,012	0,62	0,74 (0,48-1,14)	0,174	0,326	0,57
miR-193a-5p	Continuo	0,78 (0,62-0,97)	0,025	0,114	0,60	0,98 (0,78-1,24)	0,878	0,878	0,52	1,43 (0,65-3,16)	0,374	0,468	0,57
miR-221-3p	Continuo	0,68 (0,49-0,96)	0,029	0,121	0,55	0,83 (0,69-0,98)	0,033	0,075	0,56	0,59 (0,41-0,86)	0,005	0,019	0,67

Variable	Características	Cohorte 1, n=126, 56 con recidiva			Cohorte 2, n=110, 49 con recidiva			Cohorte 3, n=99, 25 con recidiva					
		HR (IC de un 95 %)	Valor de p	Valor de p corregido mediante BH	Índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Valor de p corregido mediante BH	Índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Valor de p corregido mediante BH	Índice C <sup>a</sup>
miR-326	Continuo	0,90 (0,82-0,99)	<b>0,032</b>	<b>0,123</b>	0,57	0,91 (0,84-0,99)	<b>0,023</b>	<b>0,061</b>	0,61	1,46 (0,76-2,83)	0,256	0,379	0,54
miR-10b-5p	Continuo	1,28 (1,00-1,63)	<b>0,046</b>	<b>0,163</b>	0,57	1,26 (0,93-1,71)	0,135	0,270	0,53	1,12 (0,65-1,90)	0,690	0,739	0,50
miR-30d-3p	Continuo	0,89 (0,79-1,00)	<b>0,049</b>	<b>0,163</b>	0,57	0,91 (0,81-1,03)	0,174	0,278	0,55	1,79 (1,09-2,92)	<b>0,020</b>	<b>0,050</b>	0,68

<sup>a</sup> Precisión predictiva, estimada mediante el índice de concordancia de Harrell (índice C). IC: intervalo de confianza, HR: proporción de riesgos instantáneos, PSA: antígeno prostático específico. Los valores de p significativos (p < 0,05) aparecen resaltados en negrita. Los valores de p de miARN individuales se ajustaron para múltiples pruebas mediante el procedimiento de Benjamini-Hochberg (BH), FDR <0,2 se consideró significativa (resaltada en negrita).

**Tabla 3.** Lista combinada de microARN con valor pronóstico potencial a partir de los que se construyen los clasificadores.

Descubrimiento de microARN combinados	Procedimiento de descubrimiento 1 (validación cruzada dejando uno fuera)		Procedimiento de descubrimiento 2 (regresión de Cox)			
	Clasificador de 8 microARN (1) <sup>a</sup>	Clasificador pronóstico de 8 miARN reducido (2) <sup>a</sup>	11 microARN con potencial pronóstico	Clasificador de 3 microARN (3) <sup>a</sup>	Clasificador de 4 microARN (basado en proporción) (4) <sup>a</sup>	Clasificador de 2 microARN (basado en proporción) (5) <sup>a</sup>
<b>miR-374b-5p</b>	miR-374b-5p		miR-374b-5p		miR-374b-5p	miR-374b-5p
<b>miR-23a-3p</b>	miR-23a-3p		miR-23a-3p		miR-23a-3p	
<b>miR-625-3p</b>			miR-625-3p			
<b>miR-615-3p</b>	miR-615-3p		miR-615-3p			
<b>miR-185-5p</b>	miR-185-5p	miR-185-5p	miR-185-5p	miR-185-5p		
<b>miR-133a-3p</b>			miR-133a-3p		miR-133a-3p	
<b>miR-193a-5p</b>	miR-193a-5p		miR-193a-5p			
<b>miR-221-3p</b>	miR-221-3p	miR-221-3p	miR-221-3p	miR-221-3p		
<b>miR-326</b>			miR-326	miR-326		
<b>miR-10b-5p</b>			miR-10b-5p		miR-10b-5p	miR-10b-5p
<b>miR-30d-3p</b>			miR-30d-3p			
<b>miR-152-3p</b>	miR-152-3p	miR-152-3p				
<b>miR-106a-5p</b>	miR-106a-5p					

<sup>a</sup> Corresponde al listado en el ejemplo 1.

5

**Tabla 4. Clasificador pronóstico de 8 miARN (ejemplo 1+2).** Potencial pronóstico del clasificador pronóstico de 8 miARN evaluado mediante análisis de regresión de Cox uni y multivariantes del tiempo de supervivencia sin recidiva bioquímica en tres cohortes de PR.

Variable	Características	Univariante			Multivariante - Final <sup>c</sup>		
		HR (IC de un 95 %)	Valor de p	Índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Índice C <sup>b</sup>
<b>Cohorte 1, n=126, 56 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	1,00 (0,94-1,05)	0,858	0,53	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,12 (1,81-5,36)	<b>&lt;0,001</b>	0,64	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,73 (1,51-4,93)	<b>0,001</b>	0,61	1,94 (1,03-3,69)	<b>0,042</b>	0,73
Estado de los márgenes quirúrgicos	Negativos frente a positivos	2,73 (1,59-4,70)	<b>&lt;0,001</b>	0,63	2,91 (1,65-5,12)	<b>&lt;0,001</b>	
PSA preoperatorio	Continuo	1,05 (1,02-1,08)	<b>&lt;0,001</b>	0,62	1,03 (1,01-1,07)	<b>0,005</b>	0,79
Clasificador pronóstico de 8 miARN <sup>d</sup>	Continuo	2,24 (1,97-3,75)	<b>&lt;0,001</b>	0,74	2,02 (1,52-2,68)	<b>&lt;0,001</b>	
<b>Cohorte 2, n=110, 49 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	0,97 (0,93-1,03)	0,319	0,53	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,00 (1,69-5,30)	<b>&lt;0,001</b>	0,63	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,42 (1,23-4,73)	<b>0,010</b>	0,59	-	-	0,71
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,37 (1,89-6,00)	<b>&lt;0,001</b>	0,64	3,24 (1,81-5,82)	<b>&lt;0,001</b>	
PSA preoperatorio	Continuo	1,05 (1,03-1,07)	<b>&lt;0,001</b>	0,72	1,05 (1,03-1,07)	<b>&lt;0,001</b>	
Clasificador pronóstico de 8 miARN <sup>d</sup>	Continuo	0,97 (0,73-1,28)	0,817	0,52	-	-	
<b>Cohorte 3, n=99, 25 con recidiva *</b>							
Edad en el momento del diagnóstico	Continua	1,03 (0,98-1,09)	0,278	0,56	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	4,05 (1,80-9,12)	<b>0,001</b>	0,68	-	-	
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,81 (1,70-8,54)	<b>0,001</b>	0,63	2,94 (1,29-6,72)	<b>0,011</b>	0,73
PSA preoperatorio	Continuo	1,09 (1,06-1,13)	<b>&lt;0,001</b>	0,66	1,09 (1,04-1,13)	<b>&lt;0,001</b>	
Clasificador pronóstico de 8 miARN <sup>d</sup>	Continuo	1,86 (1,23-2,81)	<b>0,003</b>	0,66	1,88 (1,20-2,96)	<b>0,006</b>	0,80

Abreviaturas: IC, intervalo de confianza; HR, proporción de riesgos instantáneos; PSA, antígeno prostático específico; pT, estadio tumoral patológico; PR, prostatectomía radical.

5 <sup>a</sup> Precisión predictiva estimada mediante el índice de concordancia de Harrell (índice C).

<sup>b</sup>Columna de la izquierda, índice C basado solo en variables clinicopatológicas (es decir, excluyendo la expresión del clasificador de miARN); columna de la derecha, índice C basado en todas las variables incluidas en el modelo.

10 <sup>c</sup> El clasificador pronóstico de 8 miARN se analizó en un análisis multivariante que incluía el estadio tumoral, la puntuación de Gleason, los márgenes quirúrgicos y el PSA preoperatorio. En el modelo multivariante final, las variables que no respondieron en el análisis multivariante global se excluyeron mediante selección regresiva gradual.

15 <sup>d</sup> Para la generación de este clasificador pronóstico de 8 miARN, se calculó una suma ponderada. El nivel de expresión de cada miARN se ponderó mediante los coeficientes de regresión estimados en un modelo de riesgos proporcionales multivariante (entrenado en la cohorte 1 y sometido a prueba en las cohortes 2 y 3).

\* La puntuación de Gleason se excluyó del análisis en la cohorte 3, puesto que en el grupo de puntuaciones de Gleason bajas (5-6) no hubo ningún acontecimiento, lo que provocó un error en el análisis.

20 Los valores de p significativos (p <0,05) aparecen resaltados en negra.

25 **Tabla 5. Clasificador de 8 miR reducido (ejemplo 1+3):** Potencial pronóstico del clasificador pronóstico de 8 miARN reducido (miR-152-3p+miR-185-5p+miR-221-3p) evaluado mediante análisis de regresión de Cox uni y multivariantes del tiempo de supervivencia sin recidiva bioquímica en tres cohortes de PR.

Variable	Características	Univariante			Multivariante - Final <sup>c</sup>		
		HR (IC de un 95 %)	Valor de p	Índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Índice C <sup>b</sup>
<b>Cohorte 1, n=126, 56 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	1,00 (0,94-1,05)	0,858	0,53	-	-	
PSA preoperatorio	<10 frente a <10 ng/ml	1,81 (0,91-3,58)	0,090	0,54	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,12 (1,81-5,36)	<b>&lt;0,001</b>	0,64	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,73 (1,51-4,93)	<b>0,001</b>	0,61	2,24 (1,20-3,45)	<b>0,011</b>	0,69
Estado de los márgenes quirúrgicos	Negativos frente a positivos	2,73 (1,59-4,70)	<b>&lt;0,001</b>	0,63	2,76 (1,59-4,78)	<b>&lt;0,001</b>	0,72
Clasificador pronóstico de 8 miARN reducido <sup>d</sup>	Continuo	1,67 (1,26-2,21)	<b>&lt;0,001</b>	0,64	2,00 (1,16-3,45)	<b>0,013</b>	
<b>Cohorte 2, n=110, 49 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	0,97 (0,93-1,03)	0,319	0,53	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,00 (1,69-5,30)	<b>&lt;0,001</b>	0,63	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,42 (1,23-4,73)	<b>0,010</b>	0,59	-	-	
PSA preoperatorio	≤10 frente a <10 ng/ml	3,54 (1,72-7,31)	<b>0,001</b>	0,64	3,52 (1,70-7,31)	<b>0,001</b>	0,71
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,37 (1,89-6,00)	<b>&lt;0,001</b>	0,64	3,63 (2,01-6,53)	<b>&lt;0,001</b>	0,73
Clasificador pronóstico de 8 miARN reducido <sup>d</sup>	Continuo	1,31 (1,00-1,71)	<b>0,047</b>	0,56	1,27 (0,96-1,67)	0,091	
<b>Cohorte 3, n=99, 25 con recidiva *</b>							
Edad en el momento del diagnóstico	Continua	1,03 (0,98-1,09)	0,278	0,56	-	-	
PSA preoperatorio	≤10 frente a <10 ng/ml	3,67 (1,64-8,20)	<b>0,002</b>	0,64	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	4,05 (1,80-9,12)	<b>0,001</b>	0,68	2,91 (1,26-6,72)	<b>0,012</b>	0,74
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,81 (1,70-8,54)	<b>0,001</b>	0,63	3,38 (1,47-7,76)	<b>0,004</b>	0,78
Clasificador pronóstico de 8 miARN reducido <sup>d</sup>	Continuo	1,88 (1,31-2,70)	<b>0,001</b>	0,70	1,82 (1,23-2,69)	<b>0,003</b>	

Abreviaturas: IC, intervalo de confianza; HR, proporción de riesgos instantáneos; PSA, antígeno prostático específico; pT, estadio tumoral patológico; PR, prostatectomía radical.

5 <sup>a</sup> Precisión predictiva estimada mediante el índice de concordancia de Harrell (índice C).

<sup>b</sup>Columna de la izquierda, índice C basado solo en variables clinicopatológicas (es decir, excluyendo la expresión del clasificador de miARN); columna de la derecha, índice C basado en todas las variables incluidas en el modelo.

10 <sup>c</sup> El clasificador pronóstico de 8 miARN reducido se analizó en un análisis multivariante que incluía el estadio tumoral, la puntuación de Gleason, los márgenes quirúrgicos y el PSA preoperatorio. En el modelo multivariante final, las variables que no respondieron en el análisis multivariante global se excluyeron mediante selección regresiva gradual.

15 <sup>d</sup> Para la generación de este clasificador pronóstico de 8 miARN reducido, se calculó una suma ponderada. El nivel de expresión de cada miARN se ponderó mediante los coeficientes de regresión estimados en un modelo de riesgos proporcionales multivariante (entrenado en la cohorte 1 y sometido a prueba en las cohortes 2 y 3).

20 \* La puntuación de Gleason se excluyó del análisis en la cohorte 3, puesto que en el grupo de puntuaciones de Gleason bajas (5-6) no hubo ningún acontecimiento, lo que provocó un error en el análisis.

Los valores de p significativos (p <0,05) aparecen resaltados en negrita.

25 **Tabla 6. Combinación de candidatos para la regresión de Cox de microARN individuales (ejemplo 1+4).** Potencial pronóstico del clasificador pronóstico de 3 miARN (miR-185-5p+miR-221-3p+miR-326) evaluado mediante análisis de regresión de Cox uni y multivariantes del tiempo de supervivencia sin recidiva bioquímica en tres cohortes de PR.

Variable	Características	Univariante			Multivariante - Final <sup>c</sup>		
		HR (IC de un 95 %)	Valor de <i>p</i>	Índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de <i>p</i>	Índice C <sup>b</sup>
<b>Cohorte 1, n=126, 56 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	1,00 (0,94-1,05)	0,858	0,53	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,12 (1,81-5,36)	<b>&lt;0,001</b>	0,64	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,73 (1,51-4,93)	<b>0,001</b>	0,61	2,68 (1,46-4,93)	<b>0,001</b>	0,72
Estado de los márgenes quirúrgicos	Negativos frente a positivos	2,73 (1,59-4,70)	<b>&lt;0,001</b>	0,63	2,46 (1,39-4,34)	<b>0,002</b>	
PSA preoperatorio	Continuo	1,05 (1,02-1,08)	<b>&lt;0,001</b>	0,62	1,04 (1,01-1,07)	<b>0,005</b>	0,74
Clasificador de 3 miARN <sup>d</sup>	Continuo	1,71 (1,31-2,24)	<b>&lt;0,001</b>	0,66	1,36 (1,03-1,79)	<b>0,031</b>	
<b>Cohorte 2, n=110, 49 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	0,97 (0,93-1,03)	0,319	0,53	-	-	
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,37 (1,89-6,00)	<b>&lt;0,001</b>	0,64	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,42 (1,23-4,73)	<b>0,010</b>	0,59	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,00 (1,69-5,30)	<b>&lt;0,001</b>	0,63	3,21 (1,76-5,84)	<b>&lt;0,001</b>	0,73
PSA preoperatorio	Continuo	1,05 (1,03-1,07)	<b>&lt;0,001</b>	0,72	1,05 (1,02-1,07)	<b>&lt;0,001</b>	
Clasificador de 3 miARN <sup>d</sup>	Continuo	1,44 (1,11-1,88)	<b>0,006</b>	0,58	1,28 (1,00-1,64)	<b>0,048</b>	0,75
<b>Cohorte 3, n=99, 25 con recidiva *</b>							
Edad en el momento del diagnóstico	Continua	1,03 (0,98-1,09)	0,278	0,56	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	4,05 (1,80-9,12)	<b>0,001</b>	0,68	-	-	
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,81 (1,70-8,54)	<b>0,001</b>	0,63	2,40 (0,94-6,12)	<b>0,007</b>	0,74
PSA preoperatorio	Continuo	1,09 (1,06-1,13)	<b>&lt;0,001</b>	0,66	1,06 (1,02-1,11)	<b>0,008</b>	
Clasificador de 3 miARN <sup>d</sup>	Continua	2,10 (1,42-3,10)	<b>&lt;0,001</b>	0,70	1,91 (1,26-2,91)	<b>0,012</b>	0,80

Abreviaturas: IC, intervalo de confianza; HR, proporción de riesgos instantáneos; PSA, antígeno prostático específico; pT, estadio tumoral patológico; PR, prostatectomía radical.

5 <sup>a</sup> Precisión predictiva estimada mediante el índice de concordancia de Harrell (índice C).

<sup>b</sup> Columna de la izquierda, índice C basado solo en variables clinicopatológicas (es decir, excluyendo la expresión del clasificador de miARN); columna de la derecha, índice C basado en todas las variables incluidas en el modelo.

10 <sup>c</sup> El clasificador pronóstico de 3 miARN se analizó en un análisis multivariante que incluía el estadio tumoral, la puntuación de Gleason, los márgenes quirúrgicos y el PSA preoperatorio. En el modelo multivariante final, las variables que no respondieron en el análisis multivariante global se excluyeron mediante selección regresiva gradual.

15 <sup>d</sup> Para la generación de este clasificador pronóstico de 3 miARN, se calculó una suma ponderada. El nivel de expresión de cada miARN se ponderó mediante los coeficientes de regresión estimados en un modelo de riesgos proporcionales multivariante (entrenado en la cohorte 1 y sometido a prueba en las cohortes 2 y 3).

\* La puntuación de Gleason se excluyó del análisis en la cohorte 3, puesto que en el grupo de puntuaciones de Gleason bajas (5-6) no hubo ningún acontecimiento, lo que provocó un error en el análisis.

20 Los valores de *p* significativos (*p* < 0,05) aparecen resaltados en negrita.

25 **Tabla 7. Modelo de proporción pronóstico de 4 miARN (ejemplo 1+5).** Potencial pronóstico del modelo de proporción de 4 miARN [miR-10b-5p/miR-23a-3p]/(miR-133a-3p/miR-374b-5p) evaluado mediante análisis de regresión de Cox uni y multivariantes de la supervivencia sin recidiva bioquímica en tres cohortes de PR.

Variable	Características	Univariante			Multivariante - Final <sup>c</sup>		
		HR (IC de un 95 %)	Valor de p	Índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Índice C <sup>b</sup>
<b>Cohorte 1, n=126, 56 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	1,00 (0,94-1,05)	0,858	0,53	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,12 (1,81-5,36)	<b>&lt;0,001</b>	0,64	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,73 (1,51-4,93)	<b>0,001</b>	0,61	1,98 (1,04-3,75)	<b>0,037</b>	0,73
Estado de los márgenes quirúrgicos	Negativos frente a positivos	2,73 (1,59-4,70)	<b>&lt;0,001</b>	0,63	3,04 (1,71-5,42)	<b>&lt;0,001</b>	
PSA preoperatorio	Continuo	1,05 (1,02-1,08)	<b>&lt;0,001</b>	0,62	1,04 (1,02-1,07)	<b>0,002</b>	0,76
Modelo de proporción de 4 miARN	Continuo	0,77 (0,68-0,87)	<b>&lt;0,001</b>	0,69	0,78 (0,67-0,90)	<b>0,001</b>	
<b>Cohorte 2, n=110, 49 con recidiva</b>							
Edad en el momento del diagnóstico	Continua	0,97 (0,93-1,03)	0,319	0,53	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	3,00 (1,69-5,30)	<b>&lt;0,001</b>	0,63	-	-	
Puntuación de Gleason	5-6 frente a 7-10	2,42 (1,23-4,73)	<b>0,010</b>	0,59	-	-	
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,37 (1,89-6,00)	<b>&lt;0,001</b>	0,64	3,11 (1,71-5,68)	<b>&lt;0,001</b>	0,73
PSA preoperatorio	Continuo	1,05 (1,03-1,07)	<b>&lt;0,001</b>	0,72	1,05 (1,03-1,07)	<b>&lt;0,001</b>	
Modelo de proporción de 4 miARN	Continuo	0,81 (0,67-0,97)	<b>0,021</b>	0,57	0,81 (0,66-0,97)	<b>0,024</b>	0,75
<b>Cohorte 3, n=99, 25 con recidiva *</b>							
Edad en el momento del diagnóstico	Continua	1,03 (0,98-1,09)	0,278	0,56	-	-	
Estadio tumoral	pT2a-c frente a pT3a-c	4,05 (1,80-9,12)	<b>0,001</b>	0,68	-	-	
Estado de los márgenes quirúrgicos	Negativos frente a positivos	3,81 (1,70-8,54)	<b>0,001</b>	0,63	2,52 (1,06-6,00)	<b>0,037</b>	0,73
PSA preoperatorio	Continuo	1,09 (1,06-1,13)	<b>&lt;0,001</b>	0,66	1,07 (1,03-1,12)	<b>&lt;0,001</b>	
Modelo de proporción de 4 miARN	Continuo	0,62 (0,42-0,91)	<b>0,014</b>	0,59	0,75 (0,53-1,07)	0,115	0,78

Abreviaturas: IC, intervalo de confianza; HR, proporción de riesgos instantáneos; PSA, antígeno prostático específico; pT, estadio tumoral patológico; PR, prostatectomía radical.

5 <sup>a</sup> Precisión predictiva, estimada mediante el índice de concordancia de Harrell (índice C).

<sup>b</sup>Columna de la izquierda, índice C basado solo en variables clinicopatológicas (es decir, excluyendo el modelo de proporción pronóstico de 4 miARN); columna de la derecha, índice C basado en todas las variables incluidas en el modelo.

10 <sup>c</sup>El modelo de proporción pronóstico de 4 miARN se analizó en un análisis multivariante que incluía el estadio tumoral, la puntuación de Gleason, el estado de los márgenes quirúrgicos y el PSA preoperatorio. En el modelo multivariante final, las variables que no respondieron en el análisis multivariante global se excluyeron mediante selección regresiva gradual.

15 \*La puntuación de Gleason se excluyó del análisis en la cohorte 3, puesto que en el grupo de puntuaciones de Gleason bajas (5-6) no hubo ningún acontecimiento, lo que provocó un error en el análisis. Los valores de p significativos (p < 0,05) aparecen resaltados en negrita.

20 **Tabla 8. Modelo de proporción pronóstico de 2 miARN (ejemplo 1+6).** Potencial pronóstico del modelo de proporción de 2 miARN [miR-10b-5p/miR-374b-5p] evaluado mediante análisis de regresión de Cox uni y multivariantes de la supervivencia sin recidiva bioquímica en tres cohortes de PR.

Variable	Características	Univariante			Multivariante - Final <sup>c</sup>		
		HR (IC de un 95 %)	Valor de p	Índice C <sup>a</sup>	HR (IC de un 95 %)	Valor de p	Índice C <sup>b</sup>
<b>Cohorte 1, n=126, 56 con recidiva</b>							
Edad en el momento del diagnóstico	<i>Continua</i>	1,00 (0,94-1,05)	0,858	0,53	-	-	
Estadio tumoral	<i>pT2a-c frente a pT3a-c</i>	3,12 (1,81-5,36)	<b>&lt;0,001</b>	0,64	-	-	
Puntuación de Gleason	<i>5-6 frente a 7-10</i>	2,73 (1,51-4,93)	<b>0,001</b>	0,61	2,60 (1,41-4,79)	<b>0,002</b>	0,73 0,77
Estado de los márgenes quirúrgicos	<i>Negativos frente a positivos</i>	2,73 (1,59-4,70)	<b>&lt;0,001</b>	0,63	3,20 (1,81-5,65)	<b>&lt;0,001</b>	
PSA preoperatorio	<i>Continuo</i>	1,05 (1,02-1,08)	<b>&lt;0,001</b>	0,62	1,04 (1,01-1,07)	<b>0,003</b>	
Modelo de proporción de 2 miARN	<i>Continuo</i>	0,73 (0,61-0,88)	<b>0,001</b>	0,65	0,73 (0,62-0,87)	<b>&lt;0,001</b>	
<b>Cohorte 2, n=110, 49 con recidiva</b>							
Edad en el momento del diagnóstico	<i>Continuo</i>	0,97 (0,93-1,03)	0,319	0,53	-	-	
Estadio tumoral	<i>pT2a-c frente a pT3a-c</i>	3,00 (1,69-5,30)	<b>&lt;0,001</b>	0,63	-	-	
Puntuación de Gleason	<i>5-6 frente a 7-10</i>	2,42 (1,23-4,73)	<b>0,010</b>	0,59	-	-	
Estado de los márgenes quirúrgicos	<i>Negativos frente a positivos</i>	3,37 (1,89-6,00)	<b>&lt;0,001</b>	0,64	3,20 (1,76-5,84)	<b>&lt;0,001</b>	0,73 0,74
PSA preoperatorio	<i>Continuo</i>	1,05 (1,03-1,07)	<b>&lt;0,001</b>	0,72	1,05 (1,03-1,07)	<b>&lt;0,001</b>	
Modelo de proporción de 2 miARN	<i>Continuo</i>	0,80 (0,64-1,00)	<b>0,047</b>	0,54	0,79 (0,62-1,00)	<b>0,047</b>	
<b>Cohorte 3, n=99, 25 con recidiva *</b>							
Edad en el momento del diagnóstico	<i>Continua</i>	1,03 (0,98-1,09)	0,278	0,56	-	-	
Estadio tumoral	<i>pT2a-c frente a pT3a-c</i>	4,05 (1,80-9,12)	<b>0,001</b>	0,68	-	-	
Estado de los márgenes quirúrgicos	<i>Negativos frente a positivos</i>	3,81 (1,70-8,54)	<b>0,001</b>	0,63	3,46 (1,43-8,35)	<b>0,006</b>	0,73 0,79
PSA preoperatorio	<i>Continuo</i>	1,09 (1,06-1,13)	<b>&lt;0,001</b>	0,66	1,07 (1,03-1,11)	<b>&lt;0,001</b>	
Modelo de proporción de 2 miARN	<i>Continuo</i>	0,44 (0,22-0,89)	<b>0,023</b>	0,58	0,38 (0,18-0,83)	<b>0,015</b>	

Abreviaturas: IC, intervalo de confianza; HR, proporción de riesgos instantáneos; PSA, antígeno prostático específico; pT, estadio tumoral patológico; PR, prostatectomía radical.

5 <sup>a</sup> Precisión predictiva estimada mediante el índice de concordancia de Harrell (índice C).

<sup>b</sup>Columna de la izquierda, índice C basado solo en variables clinicopatológicas (es decir, excluyendo el modelo de proporción de miARN); columna de la derecha, índice C basado en todas las variables incluidas en el modelo.

10 <sup>c</sup>El modelo de proporción pronóstico de 2 miARN se analizó en un análisis multivariante que incluía el estadio tumoral, la puntuación de Gleason, los márgenes quirúrgicos y el PSA preoperatorio. En el modelo multivariante final, las variables que no respondieron en el análisis multivariante global se excluyeron mediante selección regresiva gradual.

15 \* La puntuación de Gleason se excluyó del análisis en la cohorte 3, puesto que en el grupo de puntuaciones de Gleason bajas (5-6) no hubo ningún acontecimiento, lo que provocó un error en el análisis.

Los valores de p significativos (p <0,05) aparecen resaltados en negrita.

**Tabla 9.**

Todos los miR en la nomenclatura de miRBase, versión 21	Seq ID No	Secuencia
miR-106a-5p	1	AAAAGUGCUUACAGUGCAGGUAG
miR-10b-5p	2	UACCCUGUAGAACCGAAUUUGUG
miR-133a-3p	3	UUUGGUCCCCUUAACCAGCUG
miR-152-3p	4	UCAGUGCAUGACAGAACUUGG
miR-185-5p	5	UGGAGAGAAAGGCAGUUCCUGA
miR-193a-5p	6	UGGGUCUUUGCGGGCGAGAUGA
miR-221-3p	7	AGCUACAUUGUCUGCUGGGUUUC
miR-23a-3p	8	AUCACAUUGCCAGGGAUUUCC
miR-30d-3p	9	CUUUCAGUCAGAUGUUUGCUGC
miR-326	10	CCUCUGGGCCCUUCCUCCAG
miR-374b-5p	11	AUAUAAUACAACCUGCUAAGUG
miR-615-3p	12	UCCGAGCCUGGGUCUCCCUCUU
miR-625-3p	13	GACUAUAGAACUUUCCCCUCA
miR-151a-5p	14	UCGAGGAGCUCACAGUCUAGU

**Tabla 10.** 61 miARN detectados en todas las muestras.

hsa-let-7a-5p	hsa-miR-15a-5p	hsa-miR-29c-3p
hsa-let-7b-5p	hsa-miR-16-5p	hsa-miR-30b-5p
hsa-let-7c	hsa-miR-186-5p	hsa-miR-30c-5p
hsa-let-7f-5p	hsa-miR-191-5p	hsa-miR-30e-5p
hsa-let-7g-5p	hsa-miR-193b-3p	hsa-miR-320a
hsa-let-7i-5p	hsa-miR-1972	hsa-miR-320b
hsa-miR-103a-3p	hsa-miR-197-3p	hsa-miR-328
hsa-miR-106a-5p	hsa-miR-199a-3p	hsa-miR-342-3p
hsa-miR-107	hsa-miR-19b-3p	hsa-miR-34a-5p
hsa-miR-10b-5p	hsa-miR-21-5p	hsa-miR-361-5p
hsa-miR-125a-5p	hsa-miR-221-3p	hsa-miR-365a-3p
hsa-miR-125b-5p	hsa-miR-22-3p	hsa-miR-423-3p
hsa-miR-1260a	hsa-miR-23a-3p	hsa-miR-423-5p
hsa-miR-126-3p	hsa-miR-23b-3p	hsa-miR-490-3p
hsa-miR-128	hsa-miR-24-3p	hsa-miR-574-3p
hsa-miR-141-3p	hsa-miR-25-3p	hsa-miR-663a
hsa-miR-143-3p	hsa-miR-26a-5p	hsa-miR-664a-3p
hsa-miR-145-5p	hsa-miR-27b-3p	hsa-miR-92a-3p
hsa-miR-150-5p	hsa-miR-29a-3p	hsa-miR-93-5p
hsa-miR-151a-5p	hsa-miR-29b-3p	hsa-miR-99a-5p
		hsa-miR-99b-5p

**REFERENCIAS**

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer Journal international du cancer*. 2014.

2. Carthew RW, Sontheimer EJ. Origins and Mechanisms of microRNAs and siRNAs. *Cell*. 2009; 136:642-55.

3. Zeng Y, Cullen BR. Sequence requirements for micro RNA processing and function in human cells. *RNA*. 2003;9:112-23.

4. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell*. 2005;123:631-40.

5. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2014;42:D68-73.

6. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome research*. 2009;19:92-105.

7. Di Leva G, Garofalo M, Croce CM. MicroRNAs in cancer. *Annual review of pathology*. 2014;9:287-314.

8. Bartels CL, Tsongalis GJ. MicroRNAs: novel biomarkers for human cancer. *Clin Chem*. 2009;55:623-31.

9. Tong AW, Fulgham P, Jay C, Chen P, Khalil I, Liu S, et al. MicroRNA profile analysis of human prostate cancers. *Cancer Gene Ther*. 2009;16:206-16

10. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res*. 2004;64:5245-50

11. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Wiley: *Journal of the Royal Statistical Society. Series B (Methodological)*, 1995.

12. Draminski M, Rada-Iglesias, Enroth S, Wadelius C, Koronacki, et al. Monte Carlo feature selection for supervised classification (2008) *Bioinformatics* (2008) 24 (1): 110-117.

13. Dyrskjot L, Thykjaer T, Kruhoffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S, et al. Identifying distinct classes of bladder carcinoma using microarrays. *Nature genetics*. 2003;33:90-6.

14. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. *Cancer cell*. 2010;18:11-22.

15. Mestdagh P, Van Vlierberghe P, De Weer A, Muth D, Westermann F, Speleman F, et al. A novel and universal method for microRNA RT-qPCR data normalization. *Genome biology*. 2009;10:R64.

16. Mestdagh et al. *Nat Methods*. 2014 Aug; 11(8):809-15.

17. Bustin, S.A. (ed.) *A-Z of quantitative PCR*, IUL Biotechnology Series 5 (2004) 882 pages.

18. Harrell (2001) *Regression modelling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer-Verlag.

19. Tom Mitchell, Tom. *Machine Learning*, McGraw Hill, 1997.

20. Hieronymus H, Schultz N, Gopalan A, Carver BS, Chang MT, Xiao Y, et al. Copy number alteration burden predicts prostate cancer relapse. *Proc Natl Acad Sci U S A*. 2014; 111: 11139-44.

21. Haldrup C, Mundbjerg K, Vestergaard EM, Lamy P, Wild P, Schulz WA, et al. DNA methylation signatures for prediction of biochemical recurrence after radical prostatectomy of clinically localized prostate cancer. *J Clin Oncol*. 2013;31:3250-8.

22. Livak, K.J. and T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>( $\Delta\Delta C(T)$ ) Method. *Methods*, 2001.25(4): p. 402-8.

**PUNTOS DE LA DIVULGACIÓN**

- 5 1) Un procedimiento *in vitro* para evaluar el pronóstico de un paciente de cáncer de próstata, que comprende medir el nivel de expresión de al menos dos miR seleccionados del grupo de 13 miR que consiste en: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p en una muestra de células prostáticas obtenidas de dicho paciente, en el que un nivel de expresión distinto de dichos al menos 2 miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.
- 10 2) El procedimiento de acuerdo con el punto 1, en el que los al menos dos miR se seleccionan del grupo de miR que consiste en: miR-106a-5p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-374b-5p y miR-615-3p.
- 15 3) El procedimiento de acuerdo con el punto 1, en el que los al menos dos miR se seleccionan del grupo de miR que consiste en: miR10b-5p, miR-133a-3p, miR-23a-3p y miR-374b-5p.
- 4) El procedimiento de acuerdo con el punto 1, en el que los al menos dos miR se seleccionan del grupo de miR que consiste en: miR-185-5p, miR-221-3p y miR-326.
- 20 5) El procedimiento de acuerdo con el punto 1 o 2, en el que los al menos dos miR se seleccionan del grupo de miR que consiste en: miR-152-3p, miR-185-5p y miR-221-3p.
- 6) El procedimiento de acuerdo con el punto 1 o 3, en el que los al menos dos miR se seleccionan del grupo de miR que consiste en: miR-10b-5p y miR-374b-5p.
- 25 7) El procedimiento de acuerdo con cualquiera de los puntos precedentes, en el que los niveles de expresión son niveles de expresión normalizados.
- 8) El procedimiento de acuerdo con el punto 7, en el que los niveles de expresión se normalizan al nivel de expresión de miR-151a-5p.
- 30 9) El procedimiento de acuerdo con cualquiera de los puntos precedentes, en el que la evaluación de la probabilidad incrementada de que el cáncer de próstata progrese implica detectar el nivel de dichos al menos dos miR en dicha muestra y calcular una puntuación pronóstica (P) en base a un conjunto de datos que comprende datos del nivel de expresión de dichos al menos dos miR.
- 35 10) Procedimiento de acuerdo con el punto 9, en el que la puntuación pronóstica (P) se calcula como la proporción de los niveles de expresión:

$$P = \frac{(nivel\ de\ miR23a - 3p) \times (nivel\ de\ miR10b - 5p)}{(nivel\ de\ miR133a - 3p) \times (nivel\ de\ miR374b - 5p)}$$

- 40 11) Procedimiento de acuerdo con el punto 9, en el que la puntuación pronóstica (P) se calcula como la proporción de los niveles de expresión:

$$P = \frac{(nivel\ de\ miR10b - 5p)}{(nivel\ de\ miR374b - 5p)}$$

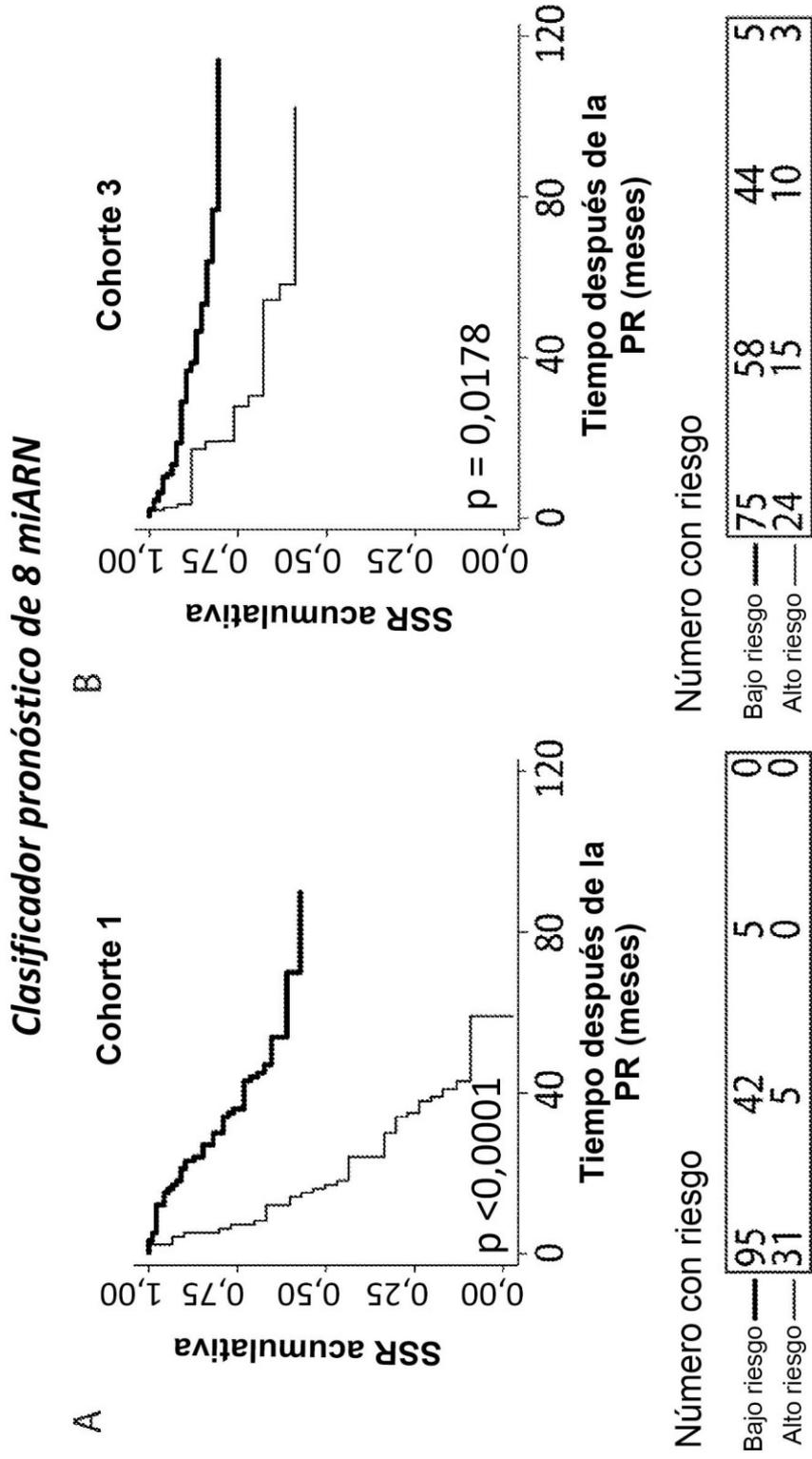
- 45 12) El procedimiento de cualquiera de los puntos precedentes, en el que el nivel de expresión de dichos miR se determina mediante el procedimiento de qRT-PCR.
- 50 13) El procedimiento de cualquiera de los puntos precedentes 1 - 11, en el que los niveles de expresión de dichos miR se miden mediante micromatrices.
- 14) El procedimiento de cualquiera de los puntos precedentes, en el que el ARN usado para medir el nivel de expresión de dichos al menos dos miR se extrae de biopsias en sacabocados de regiones de secciones de bloques FFPE o tejido ultracongelado evaluadas para comprender >90 % de tejido tumoral.
- 55 15) El procedimiento de acuerdo con cualquiera de los puntos precedentes, en el que la proporción de absorbancia de 260/280 nm del ARN usada para medir el nivel de expresión de dichos al menos dos miR está por encima de 1,75.
- 60 16) Un kit para la evaluación *in vitro* del pronóstico de un paciente de cáncer de próstata, que comprende medir el nivel de expresión de al menos dos miR seleccionados de un grupo de 13 miR que consiste en: miR-106a-5p, miR-10b-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326,

miR-374b-5p, miR-615-3p y miR-625-3p en una muestra de células prostáticas obtenidas de dicho paciente, en el que un nivel de expresión distinto de dichos al menos 2 miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.

- 5 17) El kit de acuerdo con el punto 16, en el que los al menos dos miR seleccionados son: miR-106a-5p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-374b-5p y miR-615-3p.
- 18) El kit de acuerdo con el punto 16, en el que los al menos dos miR seleccionados son: miR-10b-5p, miR-133a-3p, miR-23a-3p y miR-374b-5p.
- 10 19) El kit de acuerdo con el punto 16, en el que los al menos dos miR seleccionados son: miR-185-5p, miR-221-3p y miR-326.
- 20) El kit de acuerdo con el punto 16, en el que los al menos dos miR seleccionados son: miR-10b-5p y miR-374b-5p.
- 15 21) Un procedimiento de tratamiento de un paciente que necesita tratamiento para el cáncer de próstata, comprendiendo el procedimiento: realizar una prueba de acuerdo con cualquiera de los puntos precedentes para caracterizar el pronóstico del paciente y seleccionar un tratamiento apropiado para el paciente en base a esta información.
- 20 22) Un procedimiento de tratamiento de un paciente que necesita tratamiento para el cáncer de próstata, comprendiendo el procedimiento: realizar una prueba de acuerdo con cualquiera de los puntos precedentes para determinar si el paciente pertenece a un grupo de bajo o alto riesgo con respecto al pronóstico, y seleccionar un tratamiento apropiado para el paciente en base a esta información.
- 25

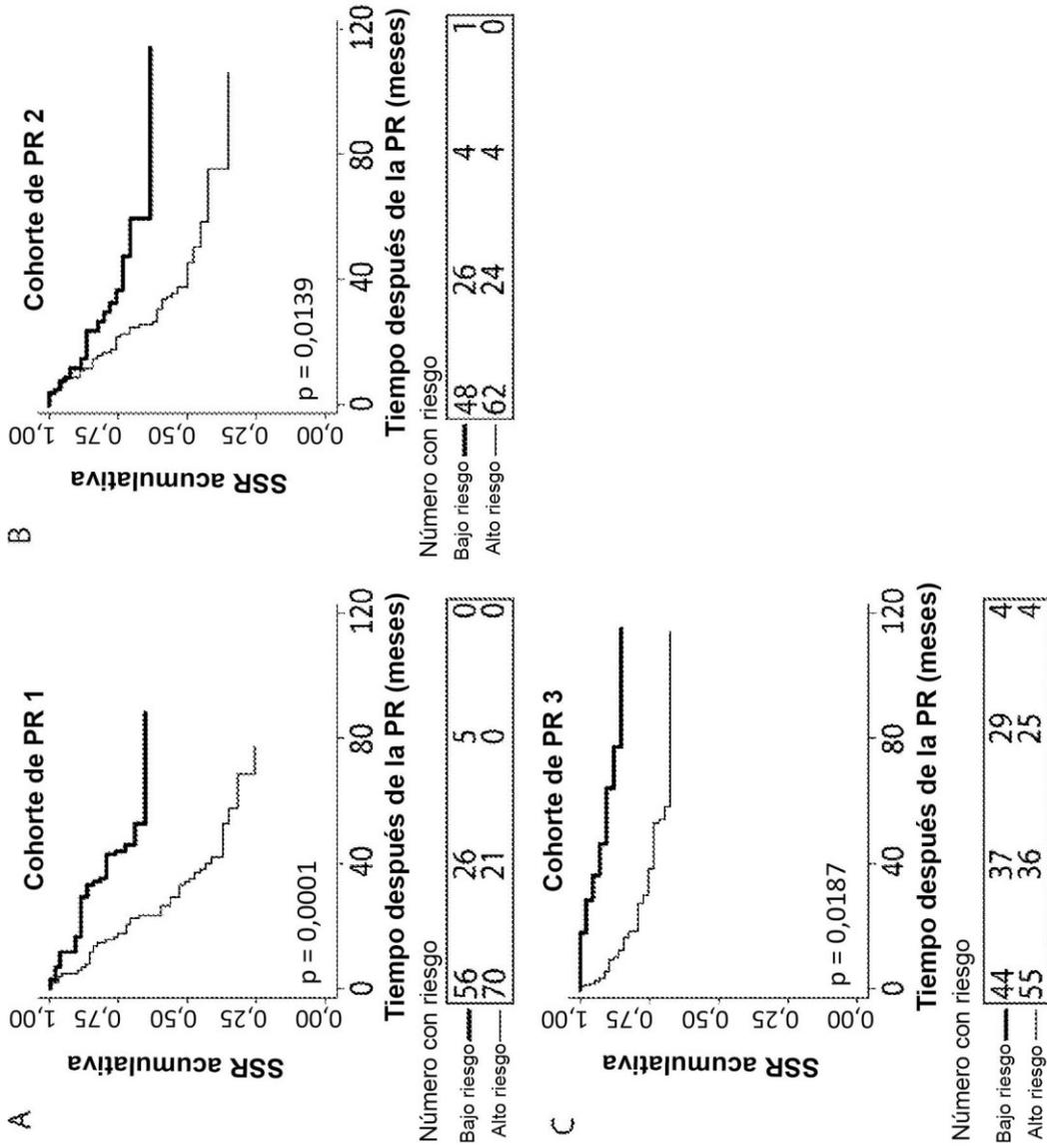
**REIVINDICACIONES**

- 5 **1.** Un procedimiento *in vitro* para evaluar el pronóstico de un paciente de cáncer de próstata, que comprende medir el nivel de expresión de: miR-10b-5p y miR-374b-5p en una muestra que comprende células prostáticas obtenidas de dicho paciente, en el que un nivel de expresión distinto de dichos miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.
- 10 **2.** El procedimiento de acuerdo con la reivindicación 1, que comprende además medir el nivel de expresión de un miR seleccionado del grupo que consiste en miR-133a-3p miR-106a-5p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p.
- 3.** El procedimiento de acuerdo con la reivindicación 2, en el que los miR son: miR10b-5p, miR-133a-3p, miR-23a-3p y miR-374b-5p.
- 15 **4.** El procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que los niveles de expresión son niveles de expresión normalizados.
- 5.** El procedimiento de acuerdo con la reivindicación 4, en el que los niveles de expresión se normalizan al nivel de expresión de miR-151a-5p.
- 20 **6.** El procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que la evaluación de la probabilidad incrementada de que el cáncer de próstata progrese implica detectar el nivel de al menos dos miR en dicha muestra y calcular una puntuación pronóstica (P) en base a un conjunto de datos que comprende datos del nivel de expresión de dichos al menos dos miR.
- 25 **7.** Procedimiento de acuerdo con la reivindicación 6, en el que la puntuación pronóstica (P) se calcula como la proporción de los niveles de expresión:
- $$P = \frac{(\text{level of miR23a} - 3p) \times (\text{level of miR10b} - 5p)}{(\text{level of miR133a} - 3p) \times (\text{level of miR374b} - 5p)}$$
- 30 **8.** Procedimiento de acuerdo con la reivindicación 6, en el que la puntuación pronóstica (P) se calcula como la proporción de los niveles de expresión:
- $$P = \frac{(\text{level of miR10b} - 5p)}{(\text{level of miR374b} - 5p)}$$
- 35 **9.** El procedimiento de cualquiera de las reivindicaciones precedentes, en el que el nivel de expresión de dichos miR se determina mediante el procedimiento de qRT-PCR.
- 10.** El procedimiento de cualquiera de las reivindicaciones precedentes 1 - 8, en el que los niveles de expresión de dichos miR se miden mediante micromatrices.
- 40 **11.** El procedimiento de cualquiera de las reivindicaciones precedentes, en el que el ARN usado para medir el nivel de expresión de dichos al menos dos miR se extrae de biopsias en sacabocados de regiones de secciones de bloques FFPE o tejido ultracongelado evaluadas para comprender >90 % de tejido tumoral.
- 45 **12.** El procedimiento de acuerdo con cualquiera de las reivindicaciones precedentes, en el que la proporción de absorbancia de 260/280 nm del ARN usada para medir el nivel de expresión de dichos al menos dos miR está por encima de 1,75.
- 50 **13.** Un kit para la evaluación *in vitro* del pronóstico de un paciente de cáncer de próstata, que comprende medir el nivel de expresión de miR-10b-5p y miR-374b-5p en una muestra que comprende células prostáticas obtenidas de dicho paciente, en el que un nivel de expresión distinto de dichos miR, en comparación con un perfil de expresión de referencia, es indicativo del pronóstico de dicho paciente.
- 55 **14.** El kit de acuerdo con la reivindicación 13, que comprende además medir el nivel de expresión de un miR seleccionado del grupo que consiste en los miR: miR-106a-5p, miR-133a-3p, miR-152-3p, miR-185-5p, miR-193a-5p, miR-221-3p, miR-23a-3p, miR-30d-3p, miR-326, miR-374b-5p, miR-615-3p y miR-625-3p.
- 60 **15.** El kit de acuerdo con la reivindicación 14, en el que los miR son: miR-10b-5p, miR-133a-3p, miR-23a-3p y miR-374b-5p.

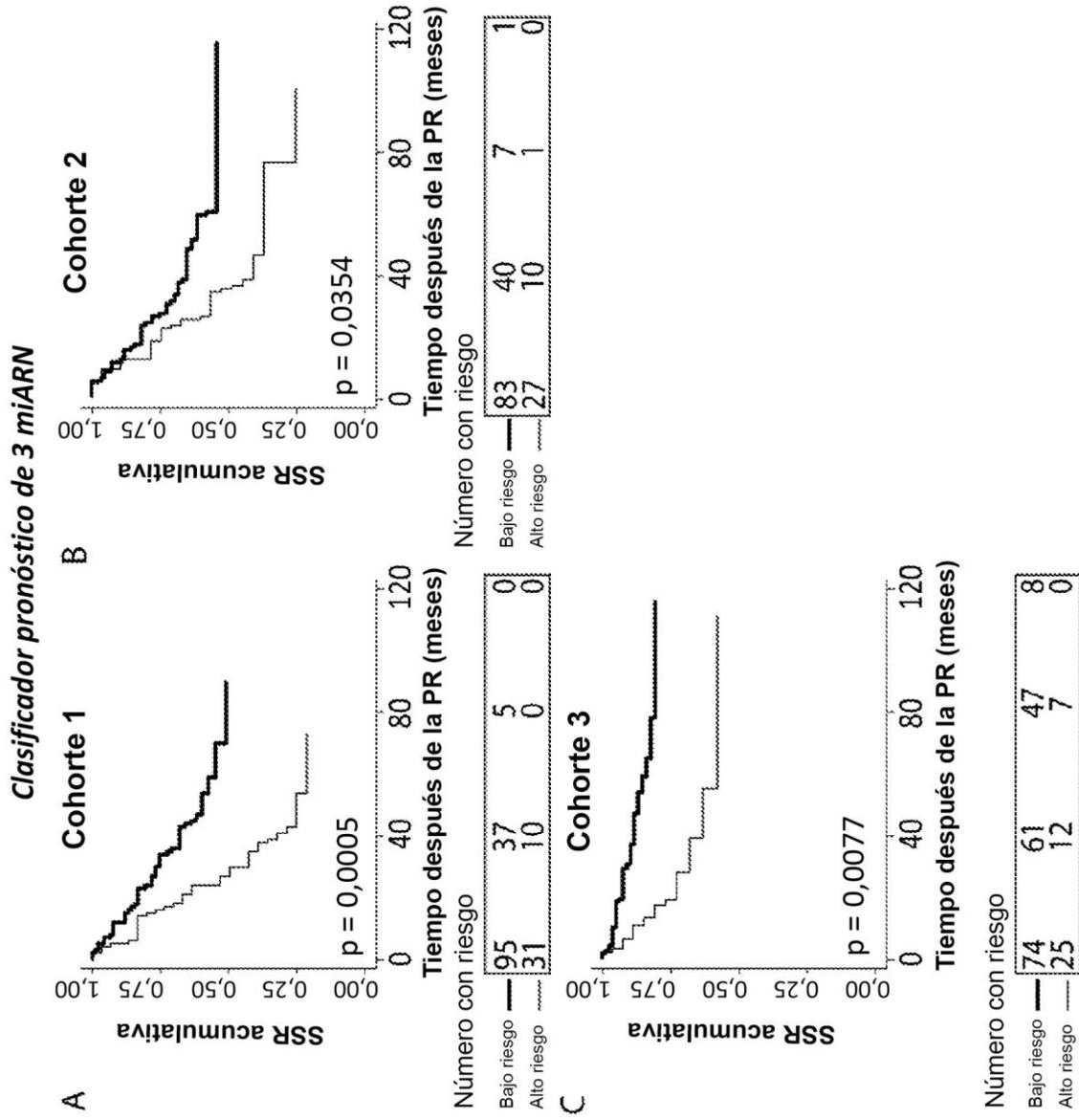


**Fig. 1**

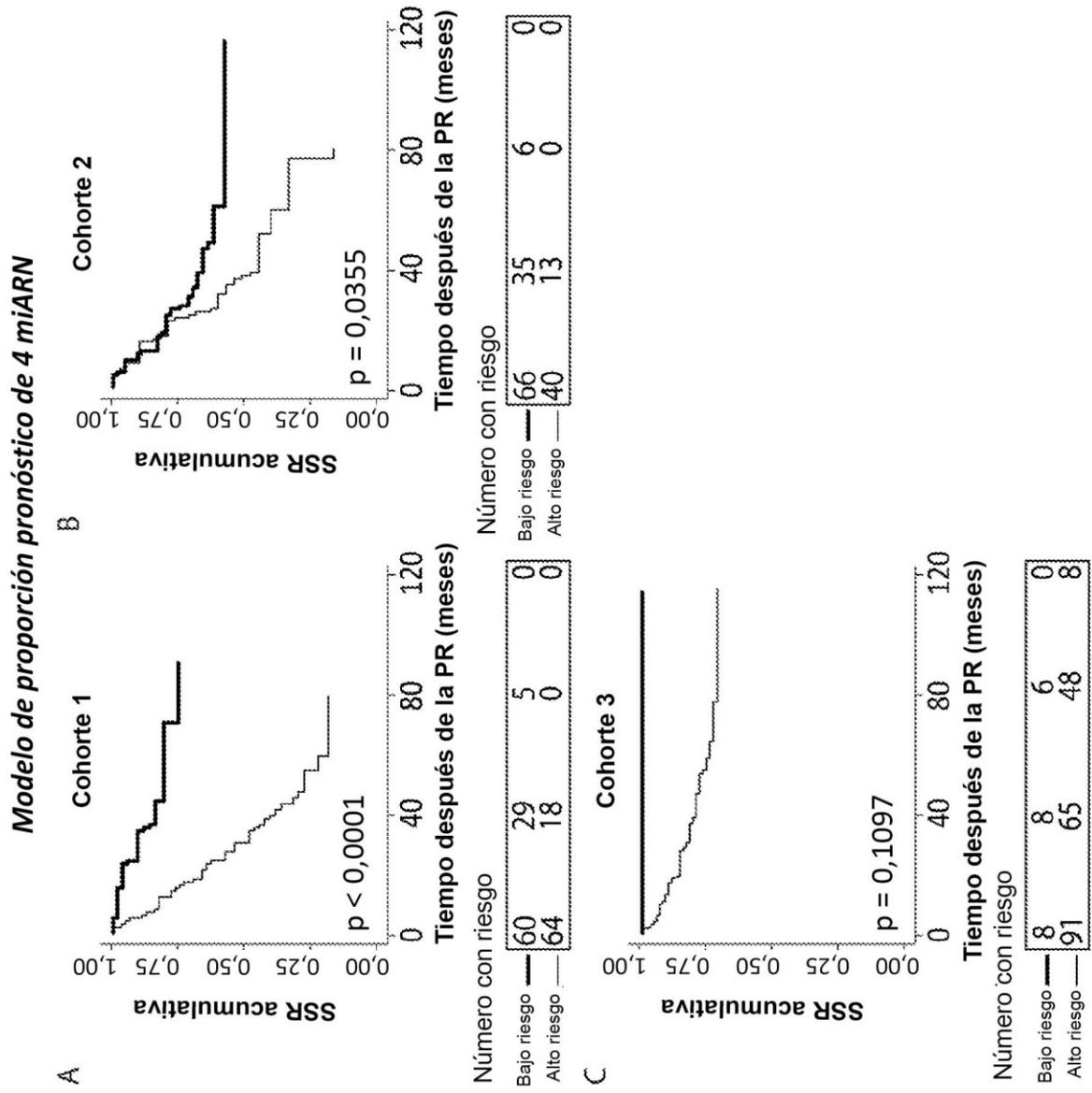
**Clasificador pronóstico de 3 miARN (clasificador pronóstico de 8 miARN reducido)**



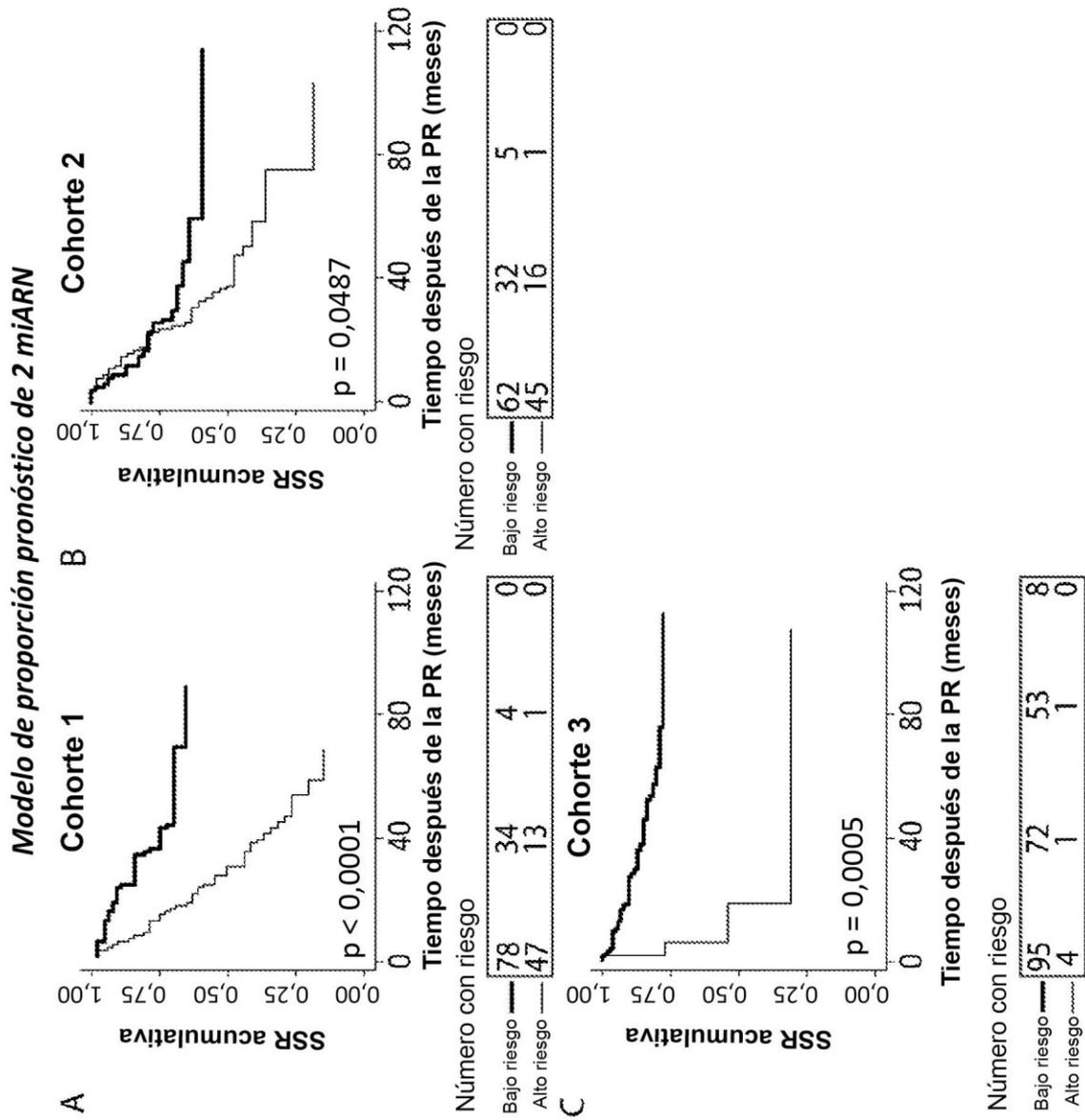
**Fig. 2**



**Fig. 3**



**Fig. 4**



**Fig. 5**