

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 785 682**

51 Int. Cl.:

**C12Q 1/6886** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.07.2017** E 17181646 (5)

97 Fecha y número de publicación de la concesión europea: **12.02.2020** EP 3431609

54 Título: **Método de diagnóstico de cáncer colorrectal**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**07.10.2020**

73 Titular/es:

**MASARYKOVA UNIVERZITA (50.0%)**

**Zerotinovo namesti 617/9**

**60177 Brno, CZ y**

**MASARYKUV ONKOLOGICKY USTAV (50.0%)**

72 Inventor/es:

**SLABY, ONDREJ;**

**VYCHYTILOVA, PETRA;**

**SVOBODA, MAREK y**

**SACHLOVA, MILANA**

74 Agente/Representante:

**ZUAZO ARALUZE, Alexander**

**ES 2 785 682 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

Método de diagnóstico de cáncer colorrectal

5 **Campo de la técnica**

La presente invención se refiere a un método de diagnóstico de cáncer colorrectal a partir de muestras de líquido corporal, usando piARN y miARN como biomarcadores.

10 **Antecedentes de la técnica**

El cáncer colorrectal (CCR) representa el 9,7% de todos los cánceres, es el tercer cáncer más común en todo el mundo y la cuarta causa más común de muertes relacionadas con cáncer, representando más del 13% de todas las muertes por cáncer y representando la segunda causa principal de mortalidad por tumores malignos en países industrializados. La tasa de supervivencia es significativamente mejor para pacientes diagnosticados de manera temprana. La posibilidad de sobrevivir al CCR está estrechamente relacionada con el estadio de la enfermedad en el diagnóstico; cuanto más temprano sea el diagnóstico, mayor será la probabilidad de supervivencia. Por ejemplo, hay menos de un 20% de posibilidades de supervivencia a los 5 años cuando se diagnostica tarde en el periodo de la enfermedad (estadio IV), mientras que hay más de un 90% de posibilidades de supervivencia a los 5 años cuando se diagnostica de manera temprana (estadio I). El diagnóstico temprano conduce normalmente a la posibilidad de intervención quirúrgica curativa que sólo es factible para pacientes con enfermedad limitada a tumor primario y ganglios linfáticos regionales.

La mayoría de los cánceres colorrectales se desarrollan lentamente, empezando como pequeños adenomas colorrectales benignos que progresan a lo largo de varias décadas a lesiones mayores y más displásicas que finalmente se vuelven malignas. Esta progresión gradual proporciona múltiples oportunidades para la prevención y la intervención. Dado que el diagnóstico temprano ofrece la mejor oportunidad para un tratamiento curativo, los pacientes con CCR se beneficiarían en gran medida de la detección temprana debido a la eficacia del tratamiento quirúrgico temprano.

Desafortunadamente, la mayoría de los casos de CCR todavía se diagnostican en estadios ya avanzados cuando no es posible un tratamiento quirúrgico curativo, y la quimioterapia sigue siendo la única opción a pesar de los altos costes y los efectos secundarios no deseados. Debido a la ausencia de métodos de diagnóstico de CCR eficaces, los pacientes con neoplasia colorrectal se beneficiarían más de pruebas de CCR de diagnóstico temprano eficaces que puedan revelar estadios tempranos de CCR, permitiendo de ese modo intervenciones preventivas.

El riesgo de CCR empieza a aumentar después de los 50 años de edad; después de eso, el riesgo continúa aumentando, duplicándose aproximadamente con cada década siguiente. El riesgo aumentado es menor en mujeres y, antes de los 75 años de edad, las mujeres tienen una incidencia menor de CCR que los hombres.

Los procedimientos más comunes y de detección más temprana disponibles en la actualidad para CCR son (i) la prueba de sangre oculta en las heces (FOBT, por sus siglas en inglés), que se basa en la suposición de que los cánceres sangrarán y, por tanto, puede detectarse en las heces usando ensayos químicos o inmunológicos; normalmente debe existir un tamaño de tumor significativo antes de que se detecte sangre en las heces; (ii) métodos de obtención de imágenes tales como colonoscopia virtual; y (iii) métodos invasivos que identifican anomalías graves tales como sigmoidoscopia o colonoscopia.

La FOBT es la prueba clínicamente útil más extendida usada para CCR, e implica una prueba rudimentaria para determinar la actividad similar a peroxidasa de sangre en hemoglobina (prueba guayaco) o un anticuerpo específico contra la hemoglobina (prueba inmunoquímica fecal, FIT, por sus siglas en inglés). Sin embargo, esta técnica de cribado presenta un determinado número de inconvenientes: el principal inconveniente es su sensibilidad mediocre de aproximadamente el 50%, con una sensibilidad del 20% para adenomas (que, si son de tamaño grande, darán como resultado el desarrollo de cáncer en 1 de cada 10 casos), debido al hecho de que no todos los adenomas y CCR sangran. La prueba tampoco es muy específica ya que la aparición de sangre en las heces puede estar relacionada con un estado no tumoral (por ejemplo, colitis ulcerosa, hemorroides, fístulas, etc.). Por tanto, también debe llevarse a cabo una colonoscopia, con los inconvenientes descritos a continuación.

La colonografía de tomografía computarizada (CTC), o colonoscopia virtual, es una técnica no invasiva reciente para la obtención de imágenes del colon, con informes que varían drásticamente en cuanto a las características del rendimiento del ensayo (que oscilan entre el 39% y el 94% de especificidad), debido principalmente a diferencias tecnológicas en la preparación del paciente y el hardware y software usado para el análisis. Otras limitaciones de la CTC incluyen las altas lecturas de falsos positivos, la incapacidad de detectar adenomas planos, la no capacidad de extirpar pólipos, las dosis de radiación repetitivas y acumulativas y el coste.

Los métodos invasivos para identificar anomalías graves incluyen sigmoidoscopia y colonoscopia. La colonoscopia es normalmente el método preferido para cribar individuos con riesgo aumentado y promedio en torno a los 50 años

de edad que tienen antecedentes de CCR o pólipos adenomatosos anteriores, u otras enfermedades predisponentes tales como enfermedad inflamatoria intestinal. Aunque la colonoscopia es todavía la prueba convencional para detectar la presencia o ausencia de pólipos y CCR, puede pasar por alto hasta el 15% de las lesiones de diámetro >1 cm. Las complicaciones con la colonoscopia pueden incluir perforación, hemorragia, insuficiencia respiratoria, arritmias e infección. Aproximadamente uno de cada 1.000 pacientes padece perforaciones y tres de cada 1.000 experimentan hemorragias. Como resultado del procedimiento, se producen entre una y tres muertes de cada 10.000 pruebas. Otras desventajas tales como la falta de personal cualificado, la incomodidad del paciente, y el alto coste y el bajo nivel de cumplimiento (alrededor del 2%) para la colonoscopia evitarán probablemente que la colonoscopia se convierta en un método de cribado de CCR rutinario para la población general. Se cree que la mayoría de CCR esporádicos se desarrollan a partir de adenomas benignos, de los cuales sólo un pequeño número evolucionará a maligno. Dado que el periodo de tiempo para el desarrollo del tumor maligno a partir de un adenoma benigno es de cinco a diez años, la detección de adenomas en toda la población general mediante colonoscopia requeriría un sobretratamiento excesivo de pacientes, siendo tanto costoso como potencialmente perjudicial. La colonoscopia es la siguiente prueba tras una FOBT positiva, y con una tasa de falsos positivos del 60%, impone riesgos innecesarios que requieren una mejora.

Por tanto, son altamente necesarios biomarcadores que permitan un diagnóstico temprano, preferiblemente también un pronóstico, y más adelante una monitorización precisa de la progresión y respuesta al tratamiento de los pacientes. Se prefieren biomarcadores presentes en muestras que pueden obtenerse de los pacientes mediante métodos poco invasivos o no invasivos (tales como muestras de sangre), ya que pueden usarse en el cribado de la población.

Actualmente, no existe una prueba de cribado serológica o una prueba de diagnóstico específica que detecte CCR temprano. Warren *et al.* (BMC Med 9 (2011) 133) han descrito recientemente la metilación de septina 9 como una prueba de sangre sensible y específica para detectar cáncer colorrectal. Aunque hay una prueba comercial, no se usa de manera rutinaria para el cribado, dado que esta prueba requiere mucho trabajo y es costosa.

En los últimos años se ha informado una gran cantidad de los denominados genes específicos de CCR. La inmensa mayoría de las solicitudes de patente o los artículos de investigación correspondientes se basan en datos obtenidos mediante el análisis de patrones de expresión de ARN en tejido de cáncer de colon frente a un tejido diferente, un tejido normal adyacente o tejido de colon sano. Tales enfoques pueden resumirse como técnicas de visualización de ARNm diferenciales. Sin embargo, las diferencias en el nivel de ARNm no se ven reflejadas necesariamente por el nivel de las proteínas correspondientes. Una proteína codificada por un ARNm raro puede encontrarse en cantidades muy altas y, no obstante, una proteína codificada por un ARNm abundante puede ser difícil de detectar y encontrar. Esta falta de correlación entre el nivel de ARNm y el nivel de proteína se debe a motivos como la estabilidad del ARNm, la eficacia de traducción, la estabilidad de la proteína, etc.

También hay enfoques recientes que investigan las diferencias en los patrones de proteínas entre tejidos diferentes o entre tejido sano y enfermo con el fin de identificar moléculas marcadoras candidatas que puedan usarse en el diagnóstico de CCR. Bruenagel *et al.* (Cancer Research 62 (2002) 2437-2442) han identificado siete proteínas de la matriz nuclear que parecen ser más abundantes en tejido de CCR en comparación con tejido normal adyacente. Además, el documento WO 2010/061996 A1 da a conocer varios biomarcadores de proteínas, algunos de los cuales pueden detectarse en suero de pacientes con CCR.

Otros ejemplos representativos conocidos como marcadores moleculares de cáncer colorrectal son los siguientes: el documento WO 2005/015224 da a conocer un método para diagnosticar cáncer colorrectal usando un anticuerpo contra la proteína RLA-0 (proteína ribosómica ácida 60S P0). El documento WO 2004/079368 da a conocer que HSP90 se expresa altamente en cánceres colorrectales. El documento WO 2004/071267 describe un método para diagnosticar cáncer colorrectal en un estadio temprano midiendo la NNMT (nicotinamida N-metiltransferasa) en una muestra de heces, y el documento WO 2005/015234 da a conocer que la proteína SAHH (S-adenosilhomocisteína hidrolasa) puede utilizarse en el diagnóstico de cáncer colorrectal. Además, el documento de patente estadounidense n.º 7.501.243 da a conocer la TTK (tirosina treonina cinasa) como un marcador de colon cáncer.

Sin embargo, a pesar de que se han propuesto marcadores de proteína candidatos, hasta la fecha no se ha demostrado su utilidad clínica/de diagnóstico. Idealmente, un nuevo marcador de diagnóstico como marcador único también debe conducir a un progreso en la sensibilidad y/o especificidad de diagnóstico o bien si se usa solo o bien en combinación con uno o más de otros marcadores, en comparación con los métodos de diagnóstico usados actualmente. En la actualidad, por ejemplo, las pruebas de sangre de diagnóstico basadas en la detección de antígeno carcinoembrionario (CEA) o CA19-9 están disponibles para ayudar al diagnóstico en el campo de CCR. El desarrollo de los primeros ensayos de antígeno carcinoembrionario (CEA) sérico en 1965 aumentaron de manera considerable la esperanza sobre la posibilidad de una prueba de cribado en sangre. Sin embargo, el trabajo posterior demostró que la sensibilidad de CEA era menor del 35% en individuos con cáncer invasivo; no pudo detectar formas tempranas y su especificidad era insuficiente, dado que sus niveles séricos aumentan en varias patologías. Las limitaciones del antígeno CA19-9, introducido más recientemente, son comparables a las de CEA. El uso actual de estas dos moléculas se limita a la vigilancia postoperatoria. Además, este biomarcador se expresa por otros tipos de cáncer y en patologías benignas. A pesar de todo, es posible aumentar la sensibilidad sin perder

especificidad combinando, con CEA, otro marcador tumoral.

Recientemente, se notificaron paneles de firma de miARN circulantes que permitían una detección temprana no invasiva y una predicción de pronóstico de cáncer de colon a partir de un líquido corporal (en particular, suero sanguíneo) de los pacientes en Vychytilova-Faltejskova *et al.*, *Carcinogenesis*, 2016, 941-950. El panel de diagnóstico comprendía miR-23a-3p, miR-27a-3p, miR-142-5p, miR-276c-3p y mostró una sensibilidad del 89% y una especificidad del 81% en la validación independiente. El panel de pronóstico comprendía miR-23a-3p, miR-276c-3p. Sin embargo, algunos miembros de este panel de diagnóstico y pronóstico se describieron como que estaban implicados con otros cánceres sólidos además del cáncer colorrectal y, por tanto, su especificidad en el cribado de población podría verse comprometida.

Tomados en conjunto, los enfoques de diagnóstico mencionados anteriormente presentan baja sensibilidad en la detección de estadios tempranos (I y II) de cáncer colorrectal y, hasta la fecha, no se han identificado biomarcadores definitivos que puedan usarse de manera fiable para detectar CCR en muestras de sangre. Por tanto, existe la necesidad clínica de identificar biomarcadores específicos para la detección temprana de CCR que puedan someterse a prueba de una manera no invasiva.

La presente invención tiene como objetivo identificar biomarcadores de líquidos corporales adecuados para diagnosticar y pronosticar CCR que puedan usarse en pruebas de cribado y que sean relevantes también en estadios tempranos de la enfermedad, para aumentar las tasas de supervivencia. Las pruebas deseables comprenderían la mayor cantidad posible de las siguientes características: riesgo promedio, individuos asintomáticos, altamente sensible, no invasiva, bajo riesgo, rentable y fácil de implementar en toda una población grande.

## Divulgación de la invención

La presente invención proporciona piARN recién identificados que son útiles como biomarcadores de cáncer colorrectal. Los piARN (o ARN asociados a PIWI) son pequeñas moléculas de ARN no codificadas expresadas en células eucariotas. Son distintos de los miARN en tamaño (26-31 nucleótidos en lugar de 21-24 nucleótidos), carecen de conservación de secuencia y complejidad aumentada. Parecen ser marcadores más específicos que los miARN. Ozawa *et al.*, *Gastroenterology*, 2017:152(5) supl. 1, página S-152 describe que pueden usarse los niveles en plasma de piARN como biomarcadores para el pronóstico y la predicción de la recaída tumoral en pacientes con cáncer colorrectal.

Los piARN recién identificados son piRNA-hsa-5937 (o piR-5937, alias piR-43771), que tiene la secuencia TCCCTGGTGGTCTAGTGGTTAGGATTCGGCA (SEQ ID NO. 1), y piR-28876, que tiene la secuencia GTTCCGTTAGTGTAGTGGTCATCACGTTCCG (SEQ ID NO. 15). Ambos piARN son útiles como biomarcadores de cáncer colorrectal que se regulan por disminución en pacientes con cáncer.

Además, dichos piARN pueden usarse en una combinación de biomarcadores con otros piARN y miARN. Se usan los niveles de expresión de piARN y miARN circulantes, y los métodos de diagnóstico y/o pronóstico se llevan a cabo usando muestras de líquidos corporales de los pacientes. Los líquidos corporales que pueden usarse incluyen, en particular, sangre, suero sanguíneo, plasma sanguíneo, orina y saliva.

Por tanto, en un primer aspecto de la presente invención, se proporciona un método para diagnosticar cáncer colorrectal, en el que se determina el nivel de expresión de piRNA-hsa-5937 en una muestra de líquido corporal, luego se compara el nivel de expresión de piRNA-hsa-5937 en la muestra con un nivel de expresión de piRNA-hsa-5937 en el líquido corporal de un ser humano sano, mientras que cuando el nivel de expresión de piRNA-hsa-5937 en la muestra es menor que el nivel de expresión de piRNA-hsa-5937 en el líquido corporal de un ser humano sano, se diagnostica cáncer colorrectal en la muestra. El piRNA-hsa-5937 se usa en esta invención junto con miR-23a-3p (tal como se describe a continuación en el presente documento) y, opcionalmente, también junto con miARN adicionales.

El nivel de expresión en el líquido corporal de un ser humano sano se obtiene estableciendo el nivel de expresión en muestras de líquido corporal obtenidas de voluntarios sanos que no padecen cáncer colorrectal. Si se desea, el nivel de expresión promedio puede obtenerse mediante métodos estadísticos conocidos.

Con el fin de comparar los niveles de expresión, es ventajoso cuando el contenido de piRNA-hsa-5937 en la muestra se referencia con respecto a una sustancia que tiene un contenido constante en el líquido corporal (es decir, normalizado). Tal sustancia puede ser miR-93-5p, que tiene la secuencia CAAAGUGCUGUUCGUGCAGGUAG (SEQ ID NO. 2).

El establecimiento del resultado del método puede basarse en lo siguiente: si el nivel normalizado es menor que un valor de punto de corte predeterminado, entonces la muestra es positiva (se diagnostica cáncer colorrectal). Si el nivel normalizado es mayor que un valor de punto de corte predeterminado, entonces la muestra es negativa (no se diagnostica cáncer colorrectal). El valor de punto de corte predeterminado puede obtenerse y determinarse usando

metodologías conocidas, por ejemplo, un análisis estadístico de los niveles de líquido corporal normalizados en grupos de voluntarios sanos (controles) y pacientes con cáncer colorrectal.

5 En una realización, los niveles de suero sanguíneo normalizados de piRNA-hsa-5937 en una muestra son  $X = 2^{-(Ct.piR - Ct.miR93)}$ . Si el nivel normalizado es menor que un valor de punto de corte predeterminado, entonces la muestra es positiva (se diagnostica cáncer colorrectal). Si el nivel normalizado es mayor que un valor de punto de corte predeterminado, entonces la muestra es negativa (no se diagnostica cáncer colorrectal).

10 El método de diagnóstico basado en la determinación del nivel de expresión de piRNA-hsa-5937 en una muestra de líquido corporal puede lograr una sensibilidad del 92,5% y una especificidad del 82,5%.

15 En un segundo aspecto de la invención, se proporciona un método de pronóstico de supervivencia global de paciente con cáncer colorrectal, en el que se determina el nivel de expresión de piRNA-hsa-5937 en una muestra de líquido corporal, luego se compara el nivel de expresión de piRNA-hsa-5937 con un nivel de expresión de referencia, mientras que el nivel de expresión de piRNA-hsa-5937 en la muestra que es mayor que el nivel de expresión de referencia indica un pronóstico de supervivencia global favorable, mientras que el nivel de expresión de piRNA-hsa-5937 en la muestra que es menor que el nivel de expresión de referencia indica un pronóstico de supervivencia global malo.

20 El nivel de expresión de referencia se obtiene a partir de un análisis estadístico de muestras de un grupo de pacientes con cáncer colorrectal con un tiempo de supervivencia global favorable (por ejemplo, al menos más de 3 años desde el diagnóstico) y un grupo de pacientes con cáncer colorrectal con un tiempo global de supervivencia malo (por ejemplo, menor de 3 años desde el diagnóstico). El experto en la técnica conoce los métodos de obtención y evaluación de los datos y de realización del análisis estadístico. Por ejemplo, el método de análisis de características operativas del receptor (ROC, por sus siglas en inglés) puede usarse para identificar el nivel de expresión de referencia.

25 El pronóstico de supervivencia global favorable significa un pronóstico de supervivencia global de al menos un número predeterminado de años o meses, por ejemplo, un pronóstico de supervivencia global de al menos 20 ó 24 ó 26 meses o 3 años desde el diagnóstico.

30 El pronóstico de supervivencia global malo significa un pronóstico de supervivencia global de menos de un número predeterminado de años o meses, por ejemplo, un pronóstico de supervivencia global de menos de 20 ó 24 ó 26 meses o 3 años desde el diagnóstico.

35 Con el fin de comparar los niveles de expresión, es ventajoso cuando el contenido de piRNA-hsa-5937 en la muestra se refiere a una sustancia que tiene un contenido constante en el líquido corporal, tal como miR-93-5p.

40 El método para diagnosticar cáncer colorrectal usando piRNA-hsa-5937 como marcador incluye además determinar los niveles de expresión en la muestra de líquido corporal de

miR-23a-3p, que tiene la secuencia AUCACAUUGCCAGGGAUUUUCC (SEQ ID NO. 3),

45 y, opcionalmente, también de al menos un miARN seleccionado de:

miR-23a-3p, que tiene la secuencia AUCACAUUGCCAGGGAUUUUCC (SEQ ID NO. 3),

miR-27a-3p, que tiene la secuencia UUCACAGUGGCUAAGUUCGCG (SEQ ID NO. 4),

50 miR-142-5p, que tiene la secuencia CAUAAAGUAGAAAGCACUACU (SEQ ID NO. 5),

compararlos con los niveles de expresión de los mismos miARN en el líquido corporal de un ser humano sano, mientras que cuando el nivel de expresión de dichos miARN en la muestra es mayor que el nivel de expresión de dichos miARN en el líquido corporal de un ser humano sano, se diagnostica cáncer colorrectal en la muestra.

55 Con el fin de comparar los niveles de expresión, es ventajoso cuando el contenido de los miARN definidos anteriormente en la muestra se referencia con respecto a una sustancia que tiene un contenido constante en el líquido corporal (es decir, un marcador de normalización). Preferiblemente, el marcador de normalización es miR-93-5p.

60 En una realización, la puntuación de diagnóstico (DxScore) puede calcularse a partir de los niveles de expresión, o preferiblemente a partir de los niveles de expresión normalizados de los biomarcadores. Entonces, el valor de la puntuación de diagnóstico puede usarse para clasificar a los pacientes en pacientes que padecen cáncer colorrectal o pacientes que no padecen cáncer colorrectal.

65 Si la DxScore es mayor que un valor de punto de corte, la muestra es positiva (se diagnostica cáncer colorrectal), y

si la DxScore es menor que un valor de punto de corte, la muestra es negativa (no se diagnostica cáncer colorrectal). El valor de punto de corte puede obtenerse usando metodologías conocidas, por ejemplo, un análisis estadístico de los niveles de líquido corporal normalizados en grupos de voluntarios sanos (controles) y pacientes con cáncer colorrectal.

5 Usando la DxScore, el método permite lograr una sensibilidad del 96,25% y una especificidad del 97,50%.

10 En una realización, el nivel en suero normalizado de hsa-miR-23a en la muestra se calcula como  $X(\text{miR-23a}) = 2^{-(\text{Ct.miR23a} - \text{Ct.miR93})}$ , el nivel en suero normalizado de hsa-miR-27a en la muestra se calcula como  $X(\text{miR-27a}) = 2^{-(\text{Ct.miR27a} - \text{Ct.miR93})}$ , el nivel en suero normalizado de hsa-miR-142-5p en la muestra se calcula como  $X(\text{miR-142-5p}) = 2^{-(\text{Ct.miR142-5p} - \text{Ct.miR93})}$ , el nivel en suero normalizado de piRNA-hsa-5937 en la muestra se calcula como  $X(\text{piR-5937}) = 2^{-(\text{Ct.piR5937} - \text{Ct.miR93})}$ .

15 Entonces, puede usarse la siguiente fórmula para calcular la puntuación de diagnóstico (DxScore):

$$\text{DxScore} = -0,8895 + 62,92648 * X(\text{miR-23a}) - 8,08928 * X(\text{miR-27a}) + 73,16745 * X(\text{miR-142-5p}) - 1,6757 * X(\text{piR-5937}).$$

20 Pueden usarse los niveles normalizados de los biomarcadores y la puntuación de diagnóstico (DxScore). Preferiblemente, el marcador de normalización es miR-93-5p. Si los niveles normalizados o la DxScore son menores que un valor de punto de corte predeterminado, entonces la muestra es positiva (se diagnostica cáncer colorrectal). Si los niveles normalizados o la DxScore son mayores que un valor de punto de corte predeterminado, entonces la muestra es negativa (no se diagnostica cáncer colorrectal). El valor de punto de corte predeterminado y/o la DxScore pueden obtenerse y determinarse usando metodologías conocidas, por ejemplo, un análisis estadístico de los niveles de líquido corporal normalizados en grupos de voluntarios sanos (controles) y pacientes con cáncer colorrectal.

25 Aún más, el método para diagnosticar cáncer colorrectal puede incluir además la determinación de niveles de expresión en la muestra de líquido corporal de al menos uno de los siguientes piARN, compararlos con los niveles de expresión del mismo piARN, respectivamente, en el líquido corporal de un ser humano sano, mientras que cuando el nivel de expresión de dicho piARN muestra el cambio que determina cáncer colorrectal con respecto al nivel de expresión del mismo piARN en el líquido corporal de un ser humano sano, se diagnostica cáncer colorrectal en la muestra:

piARN	Secuencia	Cambio en el nivel de expresión que determina cáncer colorrectal
piR-hsa-32161	CCCCCACTGCTAAATTTGACTGG (SEQ ID NO. 6)	disminución
piR-hsa-32238	CCCCCACTGCTAAATTTGACTGGTT (SEQ ID NO. 7)	disminución
piR-hsa-1242	AGCCCGGCTAGCTCAGTCGGTAGAGCATGAG (SEQ ID NO. 8)	disminución
piR-hsa-32187	CCCCCACTGCTAAATTTGACTG (SEQ ID NO. 9)	disminución
piR-hsa-27620	GCCCGGCTAGCTCAGTCGGTAGAGCATGAG (SEQ ID NO. 10)	disminución
piR-hsa-32162	CCCCCACTGCTAAATTTGACTGGC (SEQ ID NO. 11)	disminución
piR-hsa-32195	CCCCCACTGCTAAATTTGACTGGT (SEQ ID NO. 12)	disminución
piR-hsa-32158	CCCCCACTGCTAAATTTGACTGGT (SEQ ID NO. 13)	disminución
piR-hsa-29218	TAAAGTGCTGACAGTGCAGATAGTGGTCCTC (SEQ ID NO. 14)	disminución

ES 2 785 682 T3

piR-hsa-28876	GTTTCCGTAGTGTAGTGGTCATCACGTTCCGC (SEQ ID NO. 15)	disminución
piR-hsa-24672	TTCCCTGGTGGTCTAGTGGTTAGGATTCGGC (SEQ ID NO. 16)	disminución
piR-hsa-26872	GAGGAATGATGACAAGAAAAGGCCGAA (SEQ ID NO. 17)	aumento
piR-hsa-8226	TCTGCTGCCTCAGCCTCCCGAGTAGCTGA (SEQ ID NO. 18)	aumento
piR-hsa-30715	TACTCAGGAGGCTGAGGCAGGAGAATGGC (SEQ ID NO. 19)	aumento
piR-hsa-6746	TCCTGGGTTCAAGTGATTCTCCTGCCTCAGT (SEQ ID NO. 20)	aumento
piR-hsa-30937	TAGTCCCAGCTACTTGGGAGGCTGAGGCA (SEQ ID NO. 21)	aumento
piR-hsa-26543	GACCAGCCTGGCCAACATGGTGAAACCCCGCC (SEQ ID NO. 22)	aumento
piR-hsa-15278	TGCCTCCCGGATTCAAGTGATTCTCCTGCCT (SEQ ID NO. 23)	aumento
piR-hsa-5505	TCCCAGCTACCTAGGAGGCTGAGGCAGGAG (SEQ ID NO. 24)	aumento
piR-hsa-30714	TACTCAGGAGGCTGAGACAGGAGAATTGC (SEQ ID NO. 25)	aumento
piR-hsa-1359	ATCGAGGCTAGAGTCACGCTTGGGTATCGGCT (SEQ ID NO. 26)	disminución
piR-hsa-27139	GCCTGGGTAGCTCAGTCGGTAGAGCATCAGA (SEQ ID NO. 27)	disminución
piR-hsa-28488	GGTCGCTGGTTCGAATCCGGCTCGAAGGACC (SEQ ID NO. 28)	disminución
piR-hsa-27616	GCCCCGATGATCCTCAGTGGTCTGGGGTGC (SEQ ID NO. 29)	disminución
piR-hsa-29716	TAAGGTGCATCTAGTGCAGATAGTGAAGTA (SEQ ID NO. 30)	disminución
piR-hsa-28019	GGAGGTGATGAACTGTCTGAGCCTGACC (SEQ ID NO. 31)	disminución
piR-hsa-23231	CCCCTGGTGGTCTAGTGGTTAGGATTCGGC (SEQ ID NO. 32)	disminución
piR-hsa-28846	GTTCACTGATGAGAGCATTGTTCTGAGCCA (SEQ ID NO. 33)	disminución
piR-hsa-24360	TTCACTGATGAGAGCATTGTTCTGAGC (SEQ ID NO. 34)	disminución
piR-hsa-6046	TCCGATCCTCGTTGTTTTGGCTATGGCCAGA (SEQ ID NO. 35)	disminución

piR-hsa-23940	CTGACCTCAAGTGATCCACCTGCCTCAGCC (SEQ ID NO. 36)	aumento
piR-hsa-32182	CCCCCACTGCTAAATTTGACTGGCTA (SEQ ID NO. 37)	disminución
piR-hsa-32167	CCCCCACTGCTAAATTTGACTGGCT (SEC ID NO. 38)	disminución
piR-hsa-28190	GGCCGTGATCGTATAGTGGTTAGTACTCTG (SEQ ID NO. 39)	disminución
piR-hsa-23210	CCCCCACTGCTAAATTTGACTGGTT (SEQ ID NO. 40)	disminución
piR-hsa-11362	TGAGGTAGTAGGTTGTATGGTTTAGAGTTACA (SEQ ID NO. 41)	disminución
piR-hsa-23209	CCCCCACTGCTAAATTTGACTGGCTA (SEQ ID NO. 42)	disminución
piR-hsa-32159	CCCCCACTGCTAAATTTGACTGGCT (SEQ ID NO. 43)	disminución

Todos los miARN y piARN usados en la presente invención pueden referirse a la misma sustancia que tiene un contenido constante en el líquido corporal, tal como miR-93-5p.

- 5 En una realización, la puntuación de diagnóstico (DxScore) puede calcularse a partir de los niveles de expresión, o preferiblemente a partir de los niveles de expresión normalizados de los biomarcadores. Entonces, el valor de la puntuación de diagnóstico puede usarse para clasificar a los pacientes en pacientes que padecen cáncer colorrectal o pacientes que no padecen cáncer colorrectal.
- 10 El valor de punto de corte predeterminado y/o la DxScore pueden obtenerse y determinarse usando metodologías conocidas, por ejemplo, un análisis estadístico de los niveles de líquido corporal normalizados en grupos de voluntarios sanos (controles) y pacientes con cáncer colorrectal.

15 El presente método de diagnóstico permite diagnosticar cáncer colorrectal incluso en un estadio muy temprano (estadio clínico I y II), que aumenta considerablemente las oportunidades de tratamiento con éxito para los pacientes. Puede llevarse a cabo con sólo un piARN, y su sensibilidad y especificidad pueden mejorarse adicionalmente mediante la adición de marcadores adicionales. Las sensibilidades y especificidades logradas para muestras de estadio temprano indican que el presente método de diagnóstico es muy adecuado para su uso en el cribado de la población. A diferencia de los métodos de diagnóstico usados actualmente en la práctica clínica, el método de la presente invención es no invasivo, aumenta la comodidad del paciente y no requiere equipos complejos para extraer la muestra del paciente. Estos beneficios darán como resultado que más personas estén dispuestas a participar en el cribado, y en la detección más temprana del cáncer colorrectal que mejora en gran medida las oportunidades de recuperación y curación completa de los pacientes.

25 Los biomarcadores específicos de CCR de la invención son piARN y miARN circulantes, es decir, sus niveles en líquidos corporales son indicativos de diagnóstico y pronóstico de CCR. Una prueba basada en estos biomarcadores puede aceptarse ampliamente por la población general ya que es mínimamente invasiva y puede usarse para monitorizar la susceptibilidad del individuo a una enfermedad antes de recurrir, o en combinación con, métodos de cribado convencionales. Esto será extremadamente beneficioso en la gestión del riesgo, la prevención y el tratamiento de CCR.

30 A partir de esta perspectiva, la detección de pequeños ARN no codificantes (microARN y ARN asociados a PIWI) presenta un enfoque de diagnóstico muy prometedor.

35 Los microARN (miARN) son reguladores importantes de la expresión génica que comprenden una clase abundante de pequeños ARN no codificantes endógenos (18-25 nucleótidos de longitud). Son capaces de o bien promover la degradación de ARNm o bien atenuar la traducción de proteínas. Estudios bioinformáticos han estimado que los miARN pueden regular más del 50% de todos los genes humanos y cada miARN puede controlar cientos de dianas génicas. Algunos miARN se expresan de una manera específica de células, específica de tejidos y/o específica de estadios de desarrollo, mientras que otros se expresan de manera ubicua. El número de miARN verificados todavía sigue creciendo; la última versión de la base de datos web miRBase ha anotado alrededor de 1800 precursores y

40

2342 secuencias maduras en el genoma humano. Basándose en las anotaciones para la posición genómica de los miARN que indican que una inmensa mayoría de miARN se ubican en regiones intergénicas (>1 kb de distancia de los genes anotados o predichos), se ha postulado que la mayoría de los genes de miARN se transcriben como unidades de transcripción autónomas.

5 Los miARN pueden servir como reguladores maestros de muchos procesos biológicos fundamentales, tales como embriogénesis, desarrollo de órganos, diferenciación celular, proliferación, apoptosis, etc., que afectan a tales sistemas biológicos principales como severidad e inmunidad.

10 Cuando se compara con los ARNm, los microARN representan biomarcadores mucho más adecuados debido a su larga semivida, alta estabilidad en condiciones que incluyen exposición a la RNasa, extremos de pH, almacenamiento a largo plazo y múltiples ciclos de congelación-descongelación.

15 Los ARN asociados a Piwi (piARN) es la clase más grande de pequeñas moléculas de ARN no codificantes expresadas en células animales. Los piARN forman complejos ARN-proteína a través de interacciones con proteínas piwi. Son distintos del microARN (miARN) en tamaño (26-31 nt en lugar de 21-24 nt), carecen de conservación de secuencia y complejidad aumentada. Estos pequeños ARN regulan la expresión génica a nivel transcripcional y posttranscripcional, sin embargo, están implicados preferentemente en el silenciamiento de elementos transponibles LINE y SINE y, por tanto, contribuyen a la estabilidad genómica. Además, los piARN también participan en otros procesos biológicos importantes, tales como gametogénesis, segregación de cromosomas o autorrenovación de células madre. Aunque su expresión se ha descrito por primera vez en células de la línea germinal, se ha demostrado que están presentes en todos los tipos de tejidos y su expresión es altamente específica de tejido. Además, se ha descrito la expresión desregulada de los piARN en tejidos tumorales. No obstante, no se conoce la función exacta de estas moléculas en cancerogénesis. Recientemente, se ha notificado que los piARN de circulación libres están presentes de manera estable en líquidos corporales.

Detección de biomarcadores:

30 La detección de biomarcadores piARN y miARN de la presente invención puede llevarse a cabo mediante métodos conocidos, tales como un ensayo de PCR en tiempo real, un ensayo de micromatriz, un ensayo de histoquímica, un ensayo inmunológico o un ensayo de secuenciación.

35 Tanto los miARN como los piARN pueden detectarse mediante tecnologías convencionales de PCR con RT (transcripción inversa) en tiempo real disponibles en los departamentos de medicina de laboratorio a nivel de hospital. Las principales ventajas de RT-PCR en tiempo real en comparación con pruebas serológicas es su alta sensibilidad, fiabilidad y especificidad. Actualmente, los costes de ensayos de PCR con miARN/piARN individuales es lo suficientemente bajo para permitir su aplicación en toda la población en los programas de cribado. Se encuentran disponibles ensayos comerciales para detectar miARN/piARN mediante RT-PCR en tiempo real (por ejemplo, Thermo Fischer Scientific).

40 En los métodos de RT-PCT, se utiliza en primer lugar transcripción inversa (RT) para convertir el ARN diana en su ADNc, que a continuación se amplifica y cuantifica posteriormente por medio de uno de los varios métodos de PCR convencionales. Sin embargo, la traducción simple de estos métodos a miARN/piARN se complica por el corto tamaño de la diana, ya que la longitud de los cebadores normalmente usados en la etapa de PCR es tan larga como los propios miARN/piARN maduros. Los cebadores más cortos normalmente no son útiles, ya que su baja temperatura de fusión de dúplex con los miARN/piARN puede introducir un sesgo de señal. Para evitar estos desafíos, los investigadores han desarrollado varios de los enfoques creativos basados en la modificación enzimática de cebadores convencionales o cebadores completamente nuevos para miARN/piARN maduros. Actualmente, existen dos métodos principales para PCR cuantitativa de miARN/piARN (método basado en RT de tallo-bucle y métodos basados en poliadenilación).

Métodos estadísticos:

55 La curva característica operativa del receptor (ROC) es la herramienta analítica convencional para evaluar pruebas de diagnóstico. La precisión de diagnóstico de un biomarcador se mide de manera más común calculando su sensibilidad y especificidad. La sensibilidad es la proporción de pacientes que se categorizan de manera correcta como que tienen enfermedad entre aquellos que realmente tienen la enfermedad. De manera similar, la especificidad es la proporción de pacientes que se categorizan de manera correcta como que no tienen la enfermedad entre todos los pacientes que realmente no tienen la enfermedad. Dado que la mayoría de los biomarcadores de diagnóstico proporcionan resultados a escala continua, la sensibilidad y especificidad de la prueba depende del umbral específico seleccionado. El análisis de ROC es un método estadístico para identificar estos umbrales específicos. El área bajo la curva ROC (AUC) es la sensibilidad promedio del biomarcador a lo largo del intervalo de especificidades. A menudo se usa como una estadística resumen que representa el rendimiento global del biomarcador. Un biomarcador sin valor predictivo tendría un AUC de 0,5 (representado también por la línea diagonal de "oportunidad" anterior), mientras que un biomarcador con una capacidad perfecta para predecir una enfermedad tendría un AUC de 1.

**Breve descripción de los dibujos**

5 Figura 1. Validación de piR-5937 (A,B) y piR-28876 (C,D) en suero sanguíneo como biomarcadores de diagnóstico en cáncer colorrectal. A,C - prueba de Mann-Whitney, B,D - análisis de ROC.

Figura 2. Los niveles de piR-5937 en suero sanguíneo se correlacionan fuertemente con la supervivencia global de pacientes con cáncer colorrectal (prueba del rango logarítmico,  $P < 0,0001$ ).

10 Figura 3. Niveles de miARN y piARN en pacientes con cáncer colorrectal (todos los estadios clínicos, TNM I-IV) y controles sanos (\*\*\*,  $P < 0,0001$ ).

Figura 4. Niveles de miARN y piARN en pacientes con cáncer colorrectal (estadios tempranos - TNM I a II) y controles sanos.

15 Figura 5. Validación de DxScore1 basándose en niveles de miR-23a-3p, miR-27a-3p, miR-142-5p y piR-5937 en suero en pacientes con cáncer colorrectal de todos los estadios (A,C) y pacientes con cáncer colorrectal de estadios tempranos (B,D).

20 Figura 6. Validación de DxScore2 basándose en niveles de miR-23a-3p y piR-5937 en suero en pacientes con cáncer colorrectal de todos los estadios (A,C) y pacientes con cáncer colorrectal de estadios tempranos (B,D).

**Ejemplos****25 Pacientes y métodos***Pacientes*

30 Se extrajeron muestras de suero sanguíneo de 144 casos de cáncer colorrectal (36 pacientes en TNM estadio I, 36 pacientes en TNM estadio II, 36 pacientes en TNM estadio III, 36 pacientes en TNM estadio IV; 82 hombres, 62 mujeres; edad media de 65 años) y 96 controles (48 hombres, 48 mujeres; edad media de 62 años) incluidos en el estudio de obtención de perfiles de piARN. Para la validación independiente de piARN candidatos, se incluyeron otros 80 casos (19 pacientes en TNM estadio I, 21 pacientes en TNM estadio II, 21 pacientes en TNM estadio III, 19 pacientes en TNM estadio IV; 44 hombres, 36 mujeres; edad media de 65 años) y 80 controles (47 hombres, 33 mujeres; edad media de 60 años). Además, se usaron otros 90 casos de cáncer colorrectal con estadio temprano de la enfermedad (40 pacientes en TNM estadio 1 y 50 pacientes en TNM estadio 2, 50 hombres, 40 mujeres; edad media de 61 años) y 100 controles (52 hombres, 48 mujeres; edad media de 59 años) para la evaluación de biomarcadores y modelos de miARN/piARN identificados. Todos los sujetos inscritos en el estudio fueron de la misma raza (ascendencia europea), y los pacientes con cáncer colorrectal no recibieron ningún tratamiento con neoadyuvante. Se excluyeron del estudio muestras de suero hemolizado, ya que la hemólisis puede influir en la expresión de algunos miARN. Se obtuvo el consentimiento informado por escrito de todos los participantes (pacientes, donantes sanos), y la Junta de Ética local del Masaryk Memorial Cancer Institute aprobó el estudio.

*45 Extracción de ARN*

Antes de la extracción de ARN, se comprobó la hemólisis en todas las muestras usando el método espectrofotométrico de Harboe, que utiliza un método de corrección de Allen de 3 puntos (380 nm, 415 nm y 450 nm) para cuantificar la concentración de hemoglobina. Sólo se usaron posteriormente en este estudio las muestras con una concentración de hemoglobina menor de  $5 \text{ mg} \cdot \text{dl}^{-1}$ . El ARN total enriquecido para ARN pequeños se aisló del suero sanguíneo usando el kit de suero/plasma miRNeasy de Qiagen (Qiagen, GmbH, Hilden, Alemania) según el protocolo modificado por los fabricantes. En resumen, se descongelaron  $250 \mu\text{l}$  de suero sobre hielo y se centrifugaron a  $14000 \times g$  a  $4^\circ\text{C}$  durante 5 minutos para retirar el residuo celular. Posteriormente, se lisaron  $200 \mu\text{l}$  de sobrenadante con 1 ml de reactivo de lisis QIAzol. Para cada muestra, se añadieron  $1,25 \mu\text{l}$  de portador de ARN MS2  $0,8 \mu\text{g} \cdot \mu\text{l}^{-1}$  (n.º. de cat.10165948001, Roche, Basilea, Suiza) a 1 ml de disolución de QIAzol. Se eluyó el ARN extraído mediante  $2 \times 20 \mu\text{l}$  de disolución de elución calentada previamente. Con fines de preparación de la biblioteca, se aisló siempre el ARN por separado de 12 muestras ( $12 \times 250 \mu\text{l}$ ; en el caso de pacientes con cáncer de colon, los 12 pacientes presentaban el mismo estadio. Después de la separación de fases, se combinó la fase acuosa superior de las 12 muestras, se mezcló con 1,5 volúmenes de etanol al 100% y se pipeteó en una columna de centrifugación RNeasy MinElute. La elución se realizó usando  $14 \mu\text{l}$  de disolución de elución calentada previamente. La concentración y pureza del ARN se determinaron de manera espectrofotométrica midiendo su densidad óptica ( $A_{260/280} > 2,0$ ;  $A_{260/230} > 1,8$ ) usando un espectrofotómetro NanoDrop ND-1000 (Thermo Fisher Scientific, Wilmington, DE, EE.UU.). Además, las concentraciones y calidad del ARN de muestras agrupadas para NGS también se midieron usando un fluorímetro Qubit 2.0 (Thermo Fisher Scientific) y un bioanalizador Agilent 2100 (Agilent Technologies, Santa Clara, CA, EE.UU.). Las muestras o bien se almacenaron a  $-80^\circ\text{C}$  o bien se procesaron posteriormente.

*Secuenciación y construcción de biblioteca de ARN pequeños*

Se prepararon todas las bibliotecas usando el protocolo de ARN pequeños Illumina TruSeq (Illumina, San Diego, CA, EE.UU.) siguiendo las instrucciones del fabricante. En resumen, se usó al menos 1 µg de ARN total enriquecido para ARN pequeño, y se ligó un par de adaptadores a los extremos 3' y 5' de los piARN. Posteriormente, se realizaron 13 ciclos de amplificación por PCR con cebadores de amplificación marcados con códigos de barras únicos, y se llevó a cabo la selección por tamaño en gel de poliacrilamida natural al 6%. A continuación, se escindieron del gel fragmentos de ADNc de entre 145 y 160 pb correspondientes a las poblaciones de piARN, se eluyeron y precipitaron usando precipitación con etanol. El sedimento de ADNc final se secó al aire y volvió a suspenderse en 8 µl de agua libre de nucleasa. La concentración de las bibliotecas preparadas se midió usando un chip de ADN de alta sensibilidad y un bioanalizador Agilent 2100 (Agilent Technologies, Santa Clara, CA, EE.UU.). Se agruparon cantidades equimolares de cada biblioteca a una concentración final de ADNc 2 nM, y las muestras se secuenciaron en una celda de flujo con lecturas de 50 pb de un solo extremo usando un secuenciador MiSeq (Illumina, San Diego, CA, EE.UU.).

*Procesamiento de datos de secuenciación y análisis diferenciales de piARN*

Los datos de expresión de miARN basados en recuento se generaron mediante la herramienta Chimera de archivos fastq. Todas las secuencias se recortaron en el adaptador y se mapearon contra piRBase permitiendo hasta dos apareamientos erróneos por secuencia. Se realizaron análisis adicionales usando paquetes R/Bioconductor. Se eliminaron los piARN que tenían menos de 1 lectura por millón en más de 17 muestras agrupadas. Los recuentos de lectura se normalizaron previamente añadiendo factores de normalización dentro del paquete edgeR y más entre muestras normalizadas mediante la función voom en el paquete LIMMA. Tras determinar los niveles de expresión normalizados, se cribaron los piARN expresados de manera diferencial entre pacientes con cáncer colorrectal y controles sanos aplicando un ajuste de modelo lineal y un enfoque de Bayes. Los valores de P obtenidos se ajustaron para múltiples pruebas usando el método de Benjamini-Hochberg.

*Transcripción inversa y PCR cuantitativa en tiempo real*

Se sintetizó ADN complementario a partir del ARN total usando cebadores específicos de genes según el protocolo de ensayo de microARN TaqMan (Applied Biosystems, Foster City, CA, EE.UU.). Para las reacciones de transcriptasa inversa, se usaron 10 ng de muestra de ARN, 50 nM de cebador de RT de tallo-bucle (de miR-23a-3p, miR-27a-3p, miR-142-5p, piR-5937, Applied Biosystems, Thermo Fisher, EE.UU.), 1 x tampón de RT, 0,25 mM cada uno de los dNTP, transcriptasa inversa MultiScribe 3,33 U.µl<sup>-1</sup> e inhibidor de RNasa 0,25 U.µl<sup>-1</sup> (todos del kit de transcripción inversa de microARN TaqMan, Applied Biosystems, Foster City, CA, EE.UU.). Se incubaron las mezclas de reacción (10 µl) durante 30 min a 16°C, 30 min a 42°C, 5 min a 85°C y luego se mantuvieron a 4°C (termociclador T100TM; Bio-Rad, Hercules, CA, EE.UU.). Se realizó PCR en tiempo real usando el sistema de PCR en tiempo real QuantStudio 12K Flex (Applied Biosystems, Foster City, CA, EE.UU.). La mezcla de reacción de PCR de 20 µl incluía 1,33 µl de producto de RT, 1 x mezcla maestra de PCR universal TaqMan (NoUmpErase UNG) y 1 µl de cebador (de ensayos de miR-23a-3p, miR-27a-3p, miR-142-5p, piR-5937) y mezcla de sonda del kit de ensayo de microARN TaqMan (Applied Biosystems, Foster City, CA, EE.UU.). Las reacciones se incubaron en una placa óptica de 96 pocillos a 95°C durante 10 min, seguido por 40 ciclos a 95°C durante 15 s y 60°C durante 1 min.

*Normalización de datos y análisis estadísticos*

Los datos de ciclo umbrales se calcularon mediante el software QuantStudio 12K Flex (Applied Biosystems, Foster City, CA, EE.UU.). Todas las reacciones de PCR en tiempo real se ejecutaron por triplicado. Los niveles de expresión promedio de todos los miARN medidos se normalizaron usando miR-93-5p (ensayo n°. 001090; Applied Biosystems, Foster City, CA, EE.UU.) y posteriormente se analizaron mediante el método 2<sup>-ΔCt</sup>. Se seleccionó miR-93-5p como control endógeno a través de una combinación de los algoritmos convencionales geNorm y NormFinder (véase Vychytilova-Faltejskova *et al.*, 2016). Se evaluaron las diferencias estadísticas entre los niveles de miARN analizados en muestras de suero de pacientes con cáncer de colon y muestras de suero de donantes sanos mediante la prueba no paramétrica de Mann-Whitney bilateral. Se analizaron muestras emparejadas antes y después de la cirugía usando la prueba no paramétrica de Wilcoxon bilateral para muestras emparejadas. Se realizaron el análisis de curvas de supervivencia de Kaplan-Meier y la prueba del rango logarítmico para evaluar el potencial de pronóstico de piR-5937. Todos los cálculos se realizaron usando GraphPad Prism versión 5.00 (GraphPad Software, La Jolla, CA, EE.UU.) y R environment (R Development Core Team). Los valores de P de menos de 0,05 se consideraron estadísticamente significativos.

Ejemplo 1

ARN asociados a PIWI con niveles diferenciales en líquido corporal de pacientes con cáncer colorrectal y donantes sanos

5 La secuenciación de ARN pequeños de ARN purificado de muestras de suero sanguíneo de 144 pacientes con  
 10 cáncer colorrectal y 96 controles sanos se llevó a cabo usando un secuenciador MiSeq (Illumina). En total, se  
 prepararon y secuenciaron 20 bibliotecas de ARN pequeños (12 bibliotecas por pacientes con cáncer de colon, 8  
 bibliotecas por controles sanos, muestras de ARN de 12 pacientes/donantes sanos en cada biblioteca). En  
 promedio, más del 96% de las lecturas tuvo una puntuación Q mayor de 30, por tanto, se consideró que los datos  
 obtenidos eran de alta calidad. Las muestras secuenciadas contenían en promedio  $8.925.000 \pm 2.139.597$  lecturas y  
 $8.219.664 \pm 1.954.177$  lecturas que pasaron el filtro. En total, pudieron detectarse 470 y 453 piARN (más de 50 copias  
 por 1 millón de lecturas) en pacientes con cáncer colorrectal y donantes sanos, respectivamente. De estos, se  
 encontró que 39 miARN tenían niveles significativamente diferentes (10 regulados por incremento, 29 regulados por  
 disminución; tabla 1) en muestras de suero de pacientes con cáncer colorrectal en comparación con donantes sanos  
 (P ajustado < 0,005).

Tabla 1

Resumen de 39 piARN con niveles significativamente diferentes en suero sanguíneo de pacientes con cáncer colorrectal y donantes sanos (P ajustado < 0,005)				
piARN	Secuencia	Cambio con CCR	log FC	Valor de P
piR-hsa-32161	CCCCCACTGCTAAATTTGACTGG (SEQ ID NO. 6)	DISMINUCIÓN	-1,208	$6,232 \cdot 10^{-4}$
piR-hsa-32238	CCCCCACTGCTAAATTTGACTGGTT (SEQ ID NO. 7)	DISMINUCIÓN	-1,114	$3,002 \cdot 10^{-3}$
piR-hsa-1242	AGCCCGGCTAGCTCAGTCGGTAGAGCATGAG (SEQ ID NO. 8)	DISMINUCIÓN	-1,203	$4,31 \cdot 10^{-4}$
piR-hsa-32187	CCCCCACTGCTAAATTTGACTG (SEQ ID NO. 9)	DISMINUCIÓN	-1,173	$2,292 \cdot 10^{-3}$
piR-hsa-27620	GCCCGGCTAGCTCAGTCGGTAGAGCATGAG (SEQ ID NO. 10)	DISMINUCIÓN	-1,081	$1,678 \cdot 10^{-3}$
piR-hsa-32162	CCCCCACTGCTAAATTTGACTGGC (SEQ ID NO. 11)	DISMINUCIÓN	-1,567	$3,98 \cdot 10^{-6}$
piR-hsa-32195	CCCCCACTGCTAAATTTGACTGGT (SEQ ID NO. 12)	DISMINUCIÓN	-1,048	$1,0366 \cdot 10^{-3}$
piR-hsa-32158	CCCCCACTGCTAAATTTGACTGGT (SEQ ID NO. 13)	DISMINUCIÓN	-2,021	$2,1181 \cdot 10^{-3}$
piR-hsa-29218	TAAAGTGCTGACAGTGCAGATAGTGGTCCTC (SEQ ID NO. 14)	DISMINUCIÓN	-1,093	$1,1209 \cdot 10^{-3}$
piR-hsa-28876	GTTTCCGTAGTGTAGTGGTCATCACGTTCCG (SEQ ID NO. 15)	DISMINUCIÓN	-2,367	$8,51 \cdot 10^{-9}$
piR-hsa-5937	TCCCTGGTGGTCTAGTGGTTAGGATTCGGCA (SEQ ID NO. 1)	DISMINUCIÓN	-2,999	$8,46 \cdot 10^{-12}$
piR-hsa-24672	TTCCCTGGTGGTCTAGTGGTTAGGATTCGGC (SEQ ID NO. 16)	DISMINUCIÓN	-3,130	$4,03 \cdot 10^{-11}$
piR-hsa-26872	GAGGAATGATGACAAGAAAAGCCGAA (SEQ ID NO. 17)	INCREMENTO	1,149	$2,6204 \cdot 10^{-3}$
piR-hsa-8226	TCTGCTGCCTCAGCCTCCCGAGTAGCTGA (SEQ ID NO. 18)	INCREMENTO	1,910	$7,435 \cdot 10^{-4}$
piR-hsa-30715	TACTCAGGAGGCTGAGGCAGGAGAATGGC (SEQ ID NO. 19)	INCREMENTO	1,320	$3,304 \cdot 10^{-3}$
piR-hsa-6746	TCCTGGGTTGAGGTGATTCTCCTGCCTCAGT (SEQ ID NO. 20)	INCREMENTO	1,693	$2,4257 \cdot 10^{-3}$

ES 2 785 682 T3

piR-hsa-30937	TAGTCCCAGCTACTTGGGAGGCTGAGGCA (SEQ ID NO. 21)	INCREMENTO	1,917	0,0001423
piR-hsa-26543	GACCAGCCTGGCCAACATGGTGAAACCCCGCC (SEQ ID NO. 22)	INCREMENTO	1,597	2,5176 · 10 <sup>-3</sup>
piR-hsa-15278	TGCCTCCCGGATTCAAGTGATTCTCCTGCCT (SEQ ID NO. 23)	INCREMENTO	1,849	9,70 · 10 <sup>-5</sup>
piR-hsa-5505	TCCCAGCTACCTAGGAGGCTGAGGCAGGAG (SEQ ID NO. 24)	INCREMENTO	1,903	1,464 · 10 <sup>-4</sup>
piR-hsa-30714	TACTCAGGAGGCTGAGACAGGAGAATTGC (SEQ ID NO. 25)	INCREMENTO	1,574	9,617 · 10 <sup>-4</sup>
piR-hsa-1359	ATCGAGGCTAGAGTCACGCTTGGGTATCGGCT (SEQ ID NO. 26)	DISMINUCIÓN	-1,359	2,6683 · 10 <sup>-3</sup>
piR-hsa-27139	GCCTGGGTAGCTCAGTCGGTAGAGCATCAGA (SEQ ID NO. 27)	DISMINUCIÓN	-1,341	4,7537 · 10 <sup>-3</sup>
piR-hsa-28488	GGTCGCTGGTTCGAATCCGGCTCGAAGGACC (SEQ ID NO. 28)	DISMINUCIÓN	-1,443	2,1103 · 10 <sup>-3</sup>
piR-hsa-27616	GCCCGGATGATCCTCAGTGGTCTGGGGTGC (SEQ ID NO. 29)	DISMINUCIÓN	-1,277	4,8853 · 10 <sup>-3</sup>
piR-hsa-29716	TAAGGTGCATCTAGTGCAGATAGTGAAGTA (SEQ ID NO. 30)	DISMINUCIÓN	-1,323	3,3901 · 10 <sup>-3</sup>
piR-hsa-28019	GGAGGTGATGAACTGTCTGAGCCTGACC (SEQ ID NO. 31)	DISMINUCIÓN	-1,945	3,2583 · 10 <sup>-3</sup>
piR-hsa-23231	CCCCTGGTGGTCTAGTGGTTAGGATTCGGC (SEQ ID NO. 32)	DISMINUCIÓN	-1,855	9,07 · 10 <sup>-5</sup>
piR-hsa-28846	GTTCACTGATGAGAGCATTGTTCTGAGCCA (SEQ ID NO. 33)	DISMINUCIÓN	-2,050	2,142 · 10 <sup>-4</sup>
piR-hsa-24360	TTCACTGATGAGAGCATTGTTCTGAGC (SEQ ID NO. 34)	DISMINUCIÓN	-1,430	4,8228 · 10 <sup>-3</sup>
piR-hsa-6046	TCCGATCCTCGTTGTTTTGGCTATGGCCAGA (SEQ ID NO. 35)	DISMINUCIÓN	-1,396	1,4122 · 10 <sup>-3</sup>
piR-hsa-23940	CTGACCTCAAGTGATCCACCTGCCTCAGCC (SEQ ID NO. 36)	INCREMENTO	1,217	3,495 · 10 <sup>-3</sup>
piR-hsa-32182	CCCCACTGCTAAATTTGACTGGCTA (SEQ ID NO. 37)	DISMINUCIÓN	-1,044	2,0513 · 10 <sup>-3</sup>
piR-hsa-32167	CCCCACTGCTAAATTTGACTGGCT (SEC ID NO. 38)	DISMINUCIÓN	-1,108	1,2705 · 10 <sup>-3</sup>
piR-hsa-28190	GGCCGTGATCGTATAGTGGTTAGTACTCTG (SEQ ID NO. 39)	DISMINUCIÓN	-1,277	7,75 · 10 <sup>-5</sup>
piR-hsa-23210	CCCCCACTGCTAAATTTGACTGGTT (SEQ ID NO. 40)	DISMINUCIÓN	-1,418	3,86 · 10 <sup>-5</sup>
piR-hsa-11362	TGAGGTAGTAGGTTGTATGGTTTAGAGTTACA (SEQ ID NO. 41)	DISMINUCIÓN	-0,890	3,728 · 10 <sup>-3</sup>

piR-hsa-23209	CCCCCCTGCTAAATTTGACTGGCTA (SEQ ID NO. 42)	DISMINUCIÓN	-1,642	3,79·10 <sup>-5</sup>
piR-hsa-32159	CCCCCCTGCTAAATTTGACTGGCT (SEQ ID NO. 43)	DISMINUCIÓN	-1,723	3,02·10 <sup>-5</sup>

Validación de piARN candidatos en cohortes independientes de pacientes y controles

5 En total, se incluyeron muestras de suero de 80 pacientes con cáncer colorrectal y 80 donantes sanos en la fase de validación del estudio. De 39 piARN identificados mediante secuenciación de ARN pequeños que tenían niveles significativamente menores en suero sanguíneo de pacientes con cáncer colorrectal en comparación con controles sanos, se seleccionaron piR-5937 y piR-28876 como los más prometedores para validación independiente. Se demostró que ambos piARN disminuían significativamente en suero sanguíneo de pacientes con cáncer colorrectal en comparación con donantes sanos: piR-5937 (P<0,0001, AUC=0,916, sensibilidad = 92,5%, especificidad = 82,5%), piR-28876 (P<0,0001, AUC=0,896, sensibilidad = 92,5%, especificidad = 72,5%) (véase la figura 1). Además, los mayores niveles de piR-5937 en suero sanguíneo se asociaron significativamente con una supervivencia más favorable que aquellos con niveles menores (P<0,0001; HR 0,09, CI 0,02-0,036, figura 2), lo que muestra su efecto de pronóstico.

15 Ejemplo 2

Desarrollo y validación de modelos de diagnóstico de cáncer colorrectal (DxScores) basados en la combinación de niveles de piARN y miARN en suero sanguíneo

20 El desarrollo de modelos de diagnóstico de cáncer colorrectal (DxScores) basados en la combinación de niveles de piARN y miARN en suero sanguíneo se realizó usando regresión logística (para diagnóstico). Para generar la firma relacionada con el diagnóstico, se introdujeron piR-5937, piR-28876 y los 4 mejores miARN identificados en el estudio de Vychytilova-Faltejskova *et al.* (2016, citado en el capítulo de antecedentes de la técnica) (miR-23a-3p, miR-27a-3p, miR-142-5p, miR-376c-3p) en un modelo de regresión logística bidireccional por pasos en el conjunto de muestras de entrenamiento. El modelo final se tomó como el que maximiza el criterio de información Akaike. Se generó una curva ROC (característica operativa del receptor) para el análisis posterior de las DxScores. Se usó el índice máximo de Youden para obtener un valor de punto de corte óptimo para la discriminación de donantes sanos y pacientes con cáncer de colon. Posteriormente, se aplicó un análisis de curvas ROC para cada conjunto de muestras para calcular la sensibilidad y especificidad de DxScore. La validación del modelo obtenido y el valor de punto de corte se realizó en el conjunto de muestras de validación independiente con estadios tempranos de la enfermedad. Este enfoque condujo al establecimiento de los dos modelos de diagnóstico eficaces: (i) basado en la expresión de 4 biomarcadores (miR-23a-3p, miR-27a-3p, miR-142-5p y piR-5937) (DxScore1) y (ii) basado en la expresión de 2 biomarcadores (miR-23a-3p y piR-5937) (DxScore2). Ambas se caracterizaron por una muy alta sensibilidad y especificidad en el diagnóstico de cáncer colorrectal a partir de una muestra de suero sanguíneo. Todos los miARN/piARN incluidos en las DxScores indicaron niveles significativamente diferentes en suero sanguíneo de pacientes con cáncer colorrectal y donantes sanos. Los niveles de miARN fueron mayores y los niveles de piARN fueron menores en suero sanguíneo de pacientes en ambas cohortes de validación (1ª cohorte - TNM todos los estadios, 2ª cohorte - TNM estadios tempranos) (figuras 3 y 4).

40 El mejor rendimiento analítico indicó un modelo basado en 4 biomarcadores (miR-23a-3p, miR-27a-3p, miR-142-5p y piR-5937) (DxScore1). La puntuación de diagnóstico se calculó según la siguiente fórmula:

$$\text{DxScore1} = -0,8895 + 62,92648 \cdot \text{miR-23a} - 8,08928 \cdot \text{miR-27a} + 73,16745 \cdot \text{miR-142-5p} - 1,6757 \cdot \text{piR-5937}$$

45 DxScore1 permitió identificar cáncer colorrectal con una sensibilidad del 99% y una especificidad del 90% en la cohorte de pacientes con TNM todos los estadios (figura 5A,C). Si la misma DxScore1 y el valor de punto de corte se aplicaban en cohortes de pacientes de validación independiente con TNM estadios tempranos de la enfermedad, la sensibilidad era del 84% y la especificidad del 83% (figura 5B,D), lo que supera significativamente los enfoques de diagnóstico actuales usados en cáncer colorrectal (resumido en la tabla 2).

50 Tabla 2. Resumen del rendimiento analítico de DxScore1 en el diagnóstico de cáncer colorrectal.

	Fase de validación (TNM estadio I-IV)	Fase de validación I + II (TNM estadio I + II)
AUC	0,9908	0,9192
sensibilidad	96,25%	84,27%

especificidad	97,50%	82,65%
precisión	96,86%	83,46%
PPV	90,70%	81,72%
NPV	97,30%	85,26%
punto de corte	- 0,4799	- 0,4799

El rendimiento analítico ligeramente menor en comparación con DxScore1 indicó un modelo basado en 2 biomarcadores (miR-23a-3p y piR-5937) (DxScore2). La puntuación de diagnóstico se calculó según la siguiente fórmula:

5

$$\text{DxScore2} = -0,64243 + 39,96723 * \text{miR-23a-3p} - 1,62861 * \text{piR-5937}$$

DxScore2 permitió identificar cáncer colorrectal con una sensibilidad del 94% y una especificidad del 90% en la cohorte de pacientes con TNM todos los estadios (figura 6A,C). Si la misma DxScore2 y el valor de punto de corte se aplicaban en cohortes de pacientes de validación independiente con TNM estadios tempranos de la enfermedad, la sensibilidad era del 82% y la especificidad del 75% (figura 6B,D), que es un rendimiento todavía satisfactorio cuando se compara con los enfoques de diagnóstico actuales usados en cáncer colorrectal (resumido en la tabla 3).

10

Tabla 3. Resumen del rendimiento analítico de DxScore2 en el diagnóstico de cáncer colorrectal.

15

	Fase de validación (TNM estadio I-IV)	Fase de validación (TNM estadio I + II)
AUC	0,9793	0,8754
sensibilidad	93,67%	82,01%
especificidad	90,00%	75,51%
precisión	91,82%	78,69%
PPV	90,24%	74,00%
NPV	92,31%	81,82%
valor de punto de corte	- 0,09549	- 0,09549

**Lista de secuencias**

20

<110> Masarykova univerzita

<120> Método de diagnóstico de cáncer colorrectal

<130> P

25

<160> 43

<170> PatentIn versión 3.5

30

<210> 1

<211> 31

<212> ADN

<213> *Homo sapiens*

35

<400> 1

tccctggtgg tctagtgggt aggattcggc a 31

<210> 2

<211> 23

<212> ARN

40

<213> *Homo sapiens*

<400> 2

caaagugcug uucgucagg uag 23

45

<210> 3

<211> 21

<212> ADN

<213> *Homo sapiens*

ES 2 785 682 T3

	<400> 3 aucacauugc cagggauuuc c	21
5	<210> 4 <211> 21 <212> ADN <213> <i>Homo sapiens</i>	
10	<400> 4 uucacagugg cuaaguuccg c	21
15	<210> 5 <211> 21 <212> ADN <213> <i>Homo sapiens</i>	
20	<400> 5 cauaaaguag aaagcacuac u	21
25	<210> 6 <211> 24 <212> ADN <213> <i>Homo sapiens</i>	
30	<400> 6 ccccccactg ctaaattga ctgg	24
35	<210> 7 <211> 25 <212> ADN <213> <i>Homo sapiens</i>	
40	<400> 7 ccccccactgc taaattgac tggtt	25
45	<210> 8 <211> 31 <212> ADN <213> <i>Homo sapiens</i>	
50	<400> 8 agcccggcta gctcagtcgg tagagcatga g	31
55	<210> 9 <211> 23 <212> ADN <213> <i>Homo sapiens</i>	
60	<400> 9 ccccccactg ctaaattga ctg	23
65	<210> 10 <211> 30 <212> ADN <213> <i>Homo sapiens</i>	
70	<400> 10 gcccggctag ctcagtcggt agagcatgag	30
75	<210> 11 <211> 25 <212> ADN <213> <i>Homo sapiens</i>	
80	<400> 11 ccccccactg ctaaattga ctggc	25

ES 2 785 682 T3

<210> 12  
 <211> 25  
 <212> ADN  
 5 <213> *Homo sapiens*  
  
 <400> 12  
 cccccactg ctaaattga ctggt 25  
  
 10 <210> 13  
 <211> 25  
 <212> ADN  
 <213> *Homo sapiens*  
  
 15 <400> 13  
 cccccactg ctaaattga ctggt 25  
  
 <210> 14  
 <211> 31  
 20 <212> ADN  
 <213> *Homo sapiens*  
  
 <400> 14  
 25 taaagtctg acagtcaga tagtggtcct c 31  
  
 <210> 15  
 <211> 31  
 <212> ADN  
 <213> *Homo sapiens*  
 30  
 <400> 15  
 gtttccgtag tgtagtggc atcacgttcg c 31  
  
 <210> 16  
 <211> 31  
 <212> ADN  
 <213> *Homo sapiens*  
 35  
 <400> 16  
 40 ttccctggtg gtctagtgg taggattcgg c 31  
  
 <210> 17  
 <211> 27  
 <212> ADN  
 45 <213> *Homo sapiens*  
  
 <400> 17  
 gaggaatgat gacaagaaaa ggccgaa 27  
  
 50 <210> 18  
 <211> 29  
 <212> ADN  
 <213> *Homo sapiens*  
  
 55 <400> 18  
 tctgtgcct cagcctccg agtagctga 29  
  
 <210> 19  
 <211> 29  
 60 <212> ADN  
 <213> *Homo sapiens*  
  
 <400> 19  
 65 tactcaggag gctgaggcag gagaatggc 29  
  
 <210> 20

	<211> 31		
	<212> ADN		
	<213> <i>Homo sapiens</i>		
5	<400> 20		
	tcctgggttc aggtgattct cctgcctcag t	31	
	<210> 21		
	<211> 29		
10	<212> ADN		
	<213> <i>Homo sapiens</i>		
	<400> 21		
15	tagtcccagc tacttgggag gctgaggca	29	
	<210> 22		
	<211> 32		
	<212> ADN		
	<213> <i>Homo sapiens</i>		
20	<400> 22		
	gaccagcctg gccaacatgg tgaaaccccg cc	32	
	<210> 23		
25	<211> 31		
	<212> ADN		
	<213> <i>Homo sapiens</i>		
	<400> 23		
30	tgccctccgg attcaagtga ttctcctgcc t	31	
	<210> 24		
	<211> 30		
	<212> ADN		
35	<213> <i>Homo sapiens</i>		
	<400> 24		
	tcccagctac ctaggaggct gaggcaggag	30	
40	<210> 25		
	<211> 29		
	<212> ADN		
	<213> <i>Homo sapiens</i>		
45	<400> 25		
	tactcaggag gctgagacag gagaattgc	29	
	<210> 26		
	<211> 32		
50	<212> ADN		
	<213> <i>Homo sapiens</i>		
	<400> 26		
55	atcgaggcta gagtcacgct tgggtatcgg ct	32	
	<210> 27		
	<211> 31		
	<212> ADN		
	<213> <i>Homo sapiens</i>		
60	<400> 27		
	gcttgggtag ctcagtcggt agagcatcag a	31	
	<210> 28		
65	<211> 31		
	<212> ADN		

	<213> <i>Homo sapiens</i>	
	<400> 28	
5	ggtcgcctggt tcgaatccgg ctcaaggac c	31
	<210> 29	
	<211> 30	
	<212> ADN	
10	<213> <i>Homo sapiens</i>	
	<400> 29	
	gcccgatga tctcagtgg tctggggtgc	30
	<210> 30	
15	<211> 30	
	<212> ADN	
	<213> <i>Homo sapiens</i>	
	<400> 30	
20	taaggtgcat ctagtgcaga tagtgaagta	30
	<210> 31	
	<211> 28	
	<212> ADN	
25	<213> <i>Homo sapiens</i>	
	<400> 31	
	ggaggtgatg aactgtctga gcctgacc	28
30	<210> 32	
	<211> 30	
	<212> ADN	
	<213> <i>Homo sapiens</i>	
35	<400> 32	
	cccctggtgg tctagtgggt aggattcggc	30
	<210> 33	
40	<211> 30	
	<212> ADN	
	<213> <i>Homo sapiens</i>	
	<400> 33	
45	gtcactgat gagagcattg ttctgagcca	30
	<210> 34	
	<211> 27	
	<212> ADN	
50	<213> <i>Homo sapiens</i>	
	<400> 34	
	ttcactgatg agagcattgt tctgagc	27
	<210> 35	
55	<211> 31	
	<212> ADN	
	<213> <i>Homo sapiens</i>	
	<400> 35	
60	tccgatcctc gttgttttgg ctatggccag a	31
	<210> 36	
	<211> 30	
	<212> ADN	
65	<213> <i>Homo sapiens</i>	

ES 2 785 682 T3

	<400> 36 ctgacctcaa gtgatccacc tgcctcagcc	30
5	<210> 37 <211> 26 <212> ADN <213> <i>Homo sapiens</i>	
10	<400> 37 ccccactgc taaattgac tggcta	26
15	<210> 38 <211> 25 <212> ADN <213> <i>Homo sapiens</i>	
20	<400> 38 ccccactgc taaattgac tggct	25
25	<210> 39 <211> 30 <212> ADN <213> <i>Homo sapiens</i>	
30	<400> 39 ggccgtgatc gtatagtgg tagtactctg	30
35	<210> 40 <211> 26 <212> ADN <213> <i>Homo sapiens</i>	
40	<400> 40 ccccccactg ctaaattga ctggtt	26
45	<210> 41 <211> 32 <212> ADN <213> <i>Homo sapiens</i>	
50	<400> 41 tgaggtagta ggtgtatgg tttagagtta ca	32
55	<210> 42 <211> 27 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 42 ccccccactg ctaaattga ctggcta	27
	<210> 43 <211> 26 <212> ADN <213> <i>Homo sapiens</i>	
	<400> 43 ccccccactg ctaaattga ctggct	26

**REIVINDICACIONES**

1. Método para diagnosticar cáncer colorrectal, caracterizado porque se determinan un nivel de expresión de piRNA-hsa-5937, que tiene la secuencia TCCCTGGTGGTCTAGTGGTTAGGATTCCGGCA (SEQ ID NO. 1), y un nivel de expresión de miR-23a-3p, que tiene la secuencia AUCACAUUGCCAGGGAUUUCC (SEQ ID NO. 3), en una muestra de líquido corporal, luego se comparan los niveles de expresión de piRNA-hsa-5937 y miR-23a-3p en la muestra con los niveles de expresión de piRNA-hsa-5937 y miR-23a-3p en el líquido corporal de un ser humano sano, mientras que cuando el nivel de expresión de piRNA-hsa-5937 en la muestra es menor que el nivel de expresión de piRNA-hsa-5937 en el líquido corporal de un ser humano sano y el nivel de expresión de miR-23a-3p en la muestra es mayor que el nivel de expresión de miR-23a-3p en el líquido corporal de un ser humano sano, se diagnostica cáncer colorrectal en la muestra.
2. Método según la reivindicación 1, que incluye además determinar los niveles de expresión en la muestra de líquido corporal de al menos un miARN seleccionado de:
  - miR-27a-3p, que tiene la secuencia UUCACAGUGGCUAAGUCCGC (SEQ ID NO. 4),
  - miR-142-5p, que tiene la secuencia CAUAAAGUAGAAAGCACUACU (SEQ ID NO. 5),
 compararlos con los niveles de expresión de los mismos miARN en el líquido corporal de un ser humano sano, mientras que cuando el nivel de expresión de dichos miARN en la muestra es mayor que el nivel de expresión de dichos miARN en el líquido corporal de un ser humano sano, se diagnostica cáncer colorrectal en la muestra.
3. Método según una cualquiera de las reivindicaciones anteriores, que incluye además determinar los niveles de expresión en la muestra de líquido corporal de al menos uno de los siguientes piARN, compararlos con los niveles de expresión del mismo piARN, respectivamente, en el líquido corporal de un ser humano sano, mientras que cuando el nivel de expresión de dicho piARN muestra el cambio que determina cáncer colorrectal con respecto al nivel de expresión del mismo piARN en el líquido corporal de un ser humano sano, se diagnostica cáncer colorrectal en la muestra:

piARN	Secuencia	Cambio en el nivel de expresión que determina cáncer colorrectal
piR-hsa-32161	CCCCCACTGCTAAATTTGACTGG (SEQ ID NO. 6)	disminución
piR-hsa-32238	CCCCCACTGCTAAATTTGACTGGTT (SEQ ID NO. 7)	disminución
piR-hsa-1242	AGCCCGGCTAGCTCAGTCGGTAGAGCATGAG (SEQ ID NO. 8)	disminución
piR-hsa-32187	CCCCCACTGCTAAATTTGACTG (SEQ ID NO. 9)	disminución
piR-hsa-27620	GCCCGGCTAGCTCAGTCGGTAGAGCATGAG (SEQ ID NO. 10)	disminución
piR-hsa-32162	CCCCCACTGCTAAATTTGACTGGC (SEQ ID NO. 11)	disminución
piR-hsa-32195	CCCCCACTGCTAAATTTGACTGGT (SEQ ID NO. 12)	disminución
piR-hsa-32158	CCCCCACTGCTAAATTTGACTGGT (SEQ ID NO. 13)	disminución
piR-hsa-29218	TAAAGTGCTGACAGTGCAGATAGTGGTCCTC (SEQ ID NO. 14)	disminución
piR-hsa-28876	GTTTCCGTAGTGTAGTGGTCATCACGTTCCGC (SEQ ID NO. 15)	disminución
piR-hsa-24672	TTCCCTGGTGGTCTAGTGGTTAGGATTCCGGC (SEQ ID NO. 16)	disminución
piR-hsa-26872	GAGGAATGATGACAAGAAAAGGCCGAA (SEQ ID NO. 17)	aumento
piR-hsa-8226	TCTGCTGCCTCAGCCTCCCGAGTAGCTGA (SEQ ID NO. 18)	aumento
piR-hsa-30715	TACTCAGGAGGCTGAGGCAGGAGAATGGC (SEQ ID NO. 19)	aumento
piR-hsa-6746	TCCTGGGTTCCAGGTGATTCTCCTGCCTCAGT (SEQ ID NO. 20)	aumento

piR-hsa-30937	TAGTCCCAGCTACTTGGGAGGCTGAGGCA (SEQ ID NO. 21)	aumento
piR-hsa-26543	GACCAGCCTGGCCAACA TGGTGAAACCCCGCC (SEQ ID NO. 22)	aumento
piR-hsa-15278	TGCCTCCCGGATTCAAGTGATTCTCCTGCCT (SEQ ID NO. 23)	aumento
piR-hsa-5505	TCCCAGCTACCTAGGAGGCTGAGGCAGGAG (SEQ ID NO. 24)	aumento
piR-hsa-30714	TACTCAGGAGGCTGAGACAGGAGAATTGC (SEQ ID NO. 25)	aumento
piR-hsa-1359	ATCGAGGCTAGAGTCACGCTTGGGTATCGGCT (SEQ ID NO. 26)	disminución
piR-hsa-27139	GCCTGGGTAGCTCAGTCGGTAGAGCATCAGA (SEQ ID NO. 27)	disminución
piR-hsa-28488	GGTCGCTGGTTCGAATCCGGCTCGAAGGACC (SEQ ID NO. 28)	disminución
piR-hsa-27616	GCCCGGATGATCCTCAGTGGTCTGGGGTGC (SEQ ID NO. 29)	disminución
piR-hsa-29716	TAAGGTGCATCTAGTGCAGATAGTGAAGTA (SEQ ID NO. 30)	disminución
piR-hsa-28019	GGAGGTGATGAACTGTCTGAGCCTGACC (SEQ ID NO. 31)	disminución
piR-hsa-23231	CCCCTGGTGGTCTAGTGGTTAGGATTCGGC (SEQ ID NO. 32)	disminución
piR-hsa-28846	GTTCACTGATGAGAGCATTGTTCTGAGCCA (SEQ ID NO. 33)	disminución
piR-hsa-24360	TTCACTGATGAGAGCATTGTTCTGAGC (SEQ ID NO. 34)	disminución
piR-hsa-6046	TCCGATCCTCGTTGTTTTGGCTATGGCCAGA (SEQ ID NO. 35)	disminución
piR-hsa-23940	CTGACCTCAAGTGATCCACCTGCCTCAGCC (SEQ ID NO. 36)	aumento
piR-hsa-32182	CCCCACTGCTAAATTTGACTGGCTA (SEQ ID NO. 37)	disminución
piR-hsa-32167	CCCCACTGCTAAATTTGACTGGCT (SEC ID NO. 38)	disminución
piR-hsa-28190	GGCCGTGATCGTATAGTGGTTAGTACTCTG (SEQ ID NO. 39)	disminución
piR-hsa-23210	CCCCCACTGCTAAATTTGACTGGTT (SEQ ID NO. 40)	disminución
piR-hsa-11362	TGAGGTAGTAGGTTGTATGGTTTAGAGTTACA (SEQ ID NO. 41)	disminución
piR-hsa-23209	CCCCCACTGCTAAATTTGACTGGCTA (SEQ ID NO. 42)	disminución
piR-hsa-32159	CCCCCACTGCTAAATTTGACTGGCT (SEQ ID NO. 43)	disminución

4. Método según una cualquiera de las reivindicaciones anteriores, en el que se normaliza el nivel de expresión de al menos un piARN o miARN como marcador de diagnóstico en la muestra con respecto a una sustancia que tiene un contenido constante en el líquido corporal, preferiblemente con respecto a miR-93-5p, que tiene la secuencia CAAAGUGCUGUUCGUGCAGGUAG (SEQ ID NO. 2).
5. Método según una cualquiera de las reivindicaciones anteriores, en el que se compara el nivel de expresión de al menos un piARN o miARN como marcador de diagnóstico en la muestra con un valor de punto de corte predeterminado, obtenido preferiblemente usando un análisis estadístico de los niveles de expresión, opcionalmente niveles de expresión normalizados, en grupos de controles sanos y pacientes con cáncer colorrectal.
6. Método según una cualquiera de las reivindicaciones 2-5, en el que se usan los niveles de expresión de al menos dos piARN o miARN como marcador de diagnóstico en la muestra para calcular una puntuación de diagnóstico, que luego se compara con un valor de punto de corte predeterminado, obtenido preferiblemente usando un análisis estadístico de los niveles de expresión, opcionalmente niveles de expresión normalizados, en grupos de controles sanos y pacientes con cáncer colorrectal.
7. Método de pronóstico de supervivencia global de un paciente con cáncer colorrectal, caracterizado porque

se determina un nivel de expresión de piRNA-hsa-5937 en una muestra de líquido corporal, luego se compara el nivel de expresión de piRNA-hsa-5937 con un nivel de expresión de referencia, mientras que el nivel de expresión de piRNA-hsa-5937 en la muestra que es mayor que el nivel de expresión de referencia indica un pronóstico de supervivencia global favorable, mientras que el nivel de expresión de piRNA-hsa-5937 en la muestra que es menor que el nivel de expresión de referencia indica un pronóstico de supervivencia global malo, en el que el nivel de expresión de referencia se obtiene a partir de un análisis estadístico de muestras de un grupo de pacientes con cáncer colorrectal con un tiempo global de supervivencia favorable y un grupo de pacientes con cáncer colorrectal con tiempo global de supervivencia malo.

- 5
- 10
- 15
8. Método según la reivindicación 7, en el que se hace referencia al nivel de expresión de piRNA-hsa-5937 en la muestra con respecto a una sustancia que tiene un nivel de expresión constante en el líquido corporal, tal como miR-93-5p.
  9. Métodos según una cualquiera de las reivindicaciones anteriores, en los que el líquido corporal se selecciona de sangre, suero sanguíneo, plasma sanguíneo, orina y saliva.

Figura 1

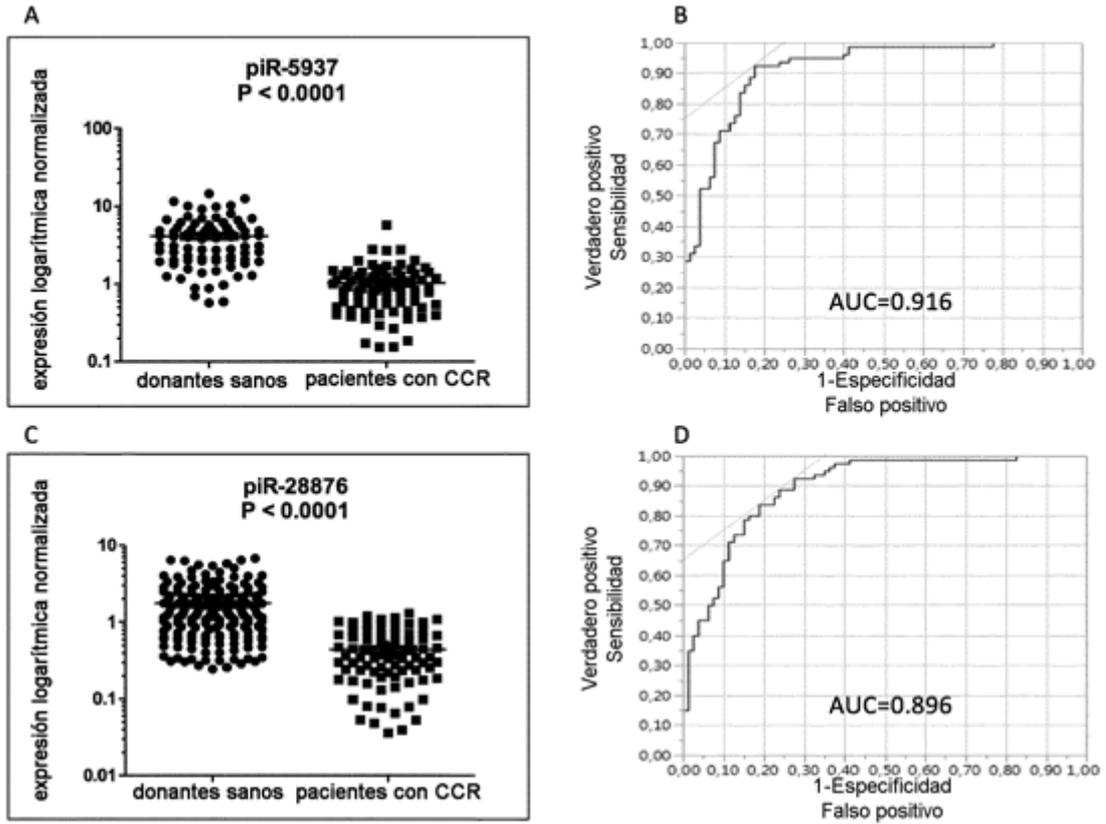


Figura 2

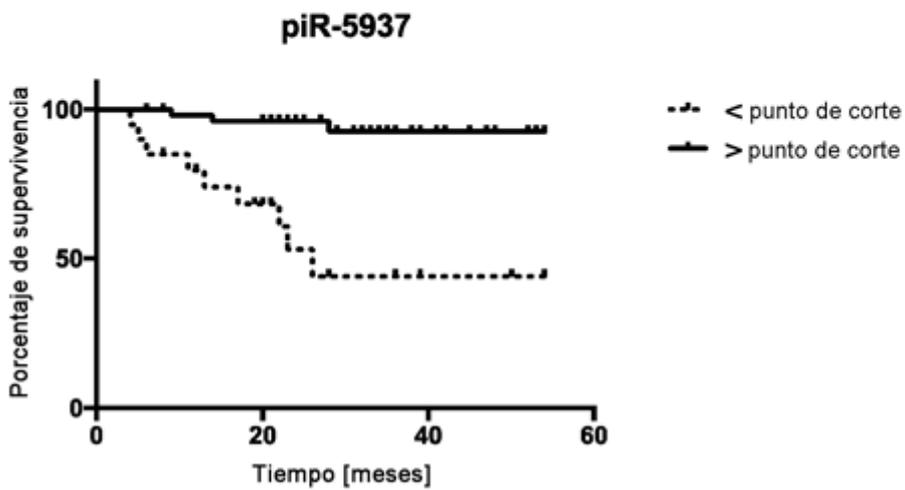


Figura 3

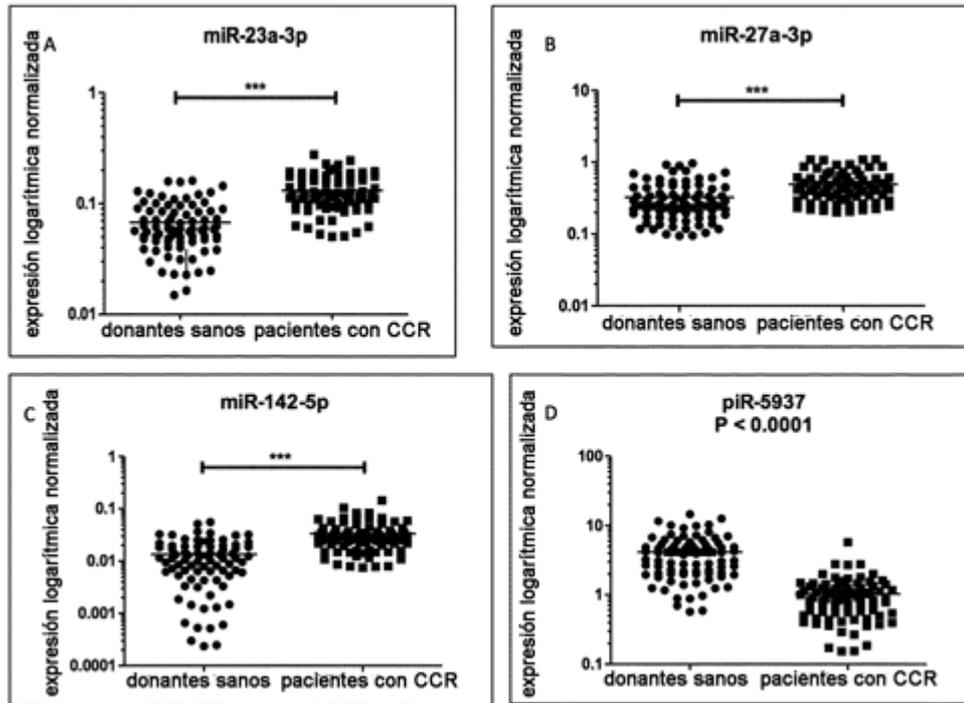


Figura 4

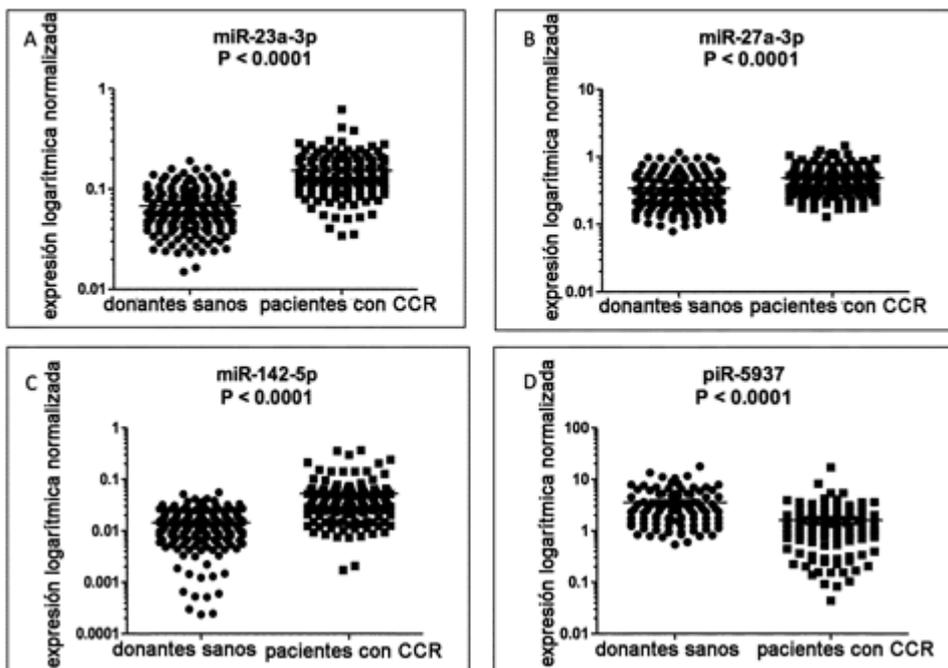


Figura 5

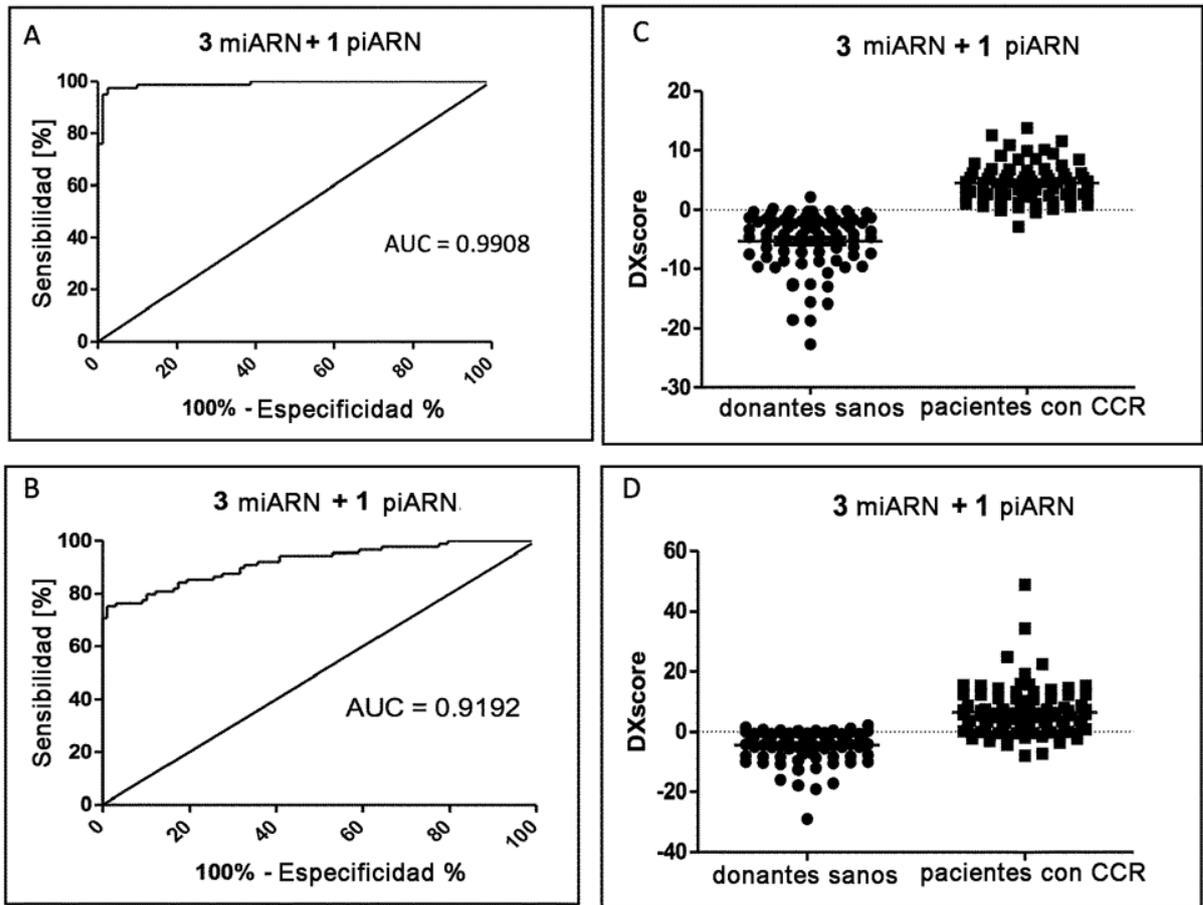


Figura 6

