

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 646 826**

21 Número de solicitud: 201630628

51 Int. Cl.:

**C12Q 1/68** (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

**13.05.2016**

43 Fecha de publicación de la solicitud:

**18.12.2017**

56 Se remite a la solicitud internacional:

**PCT/ES2017/070307**

71 Solicitantes:

**SERVICIO ANDALUZ DE SALUD (75.0%)  
Avda. de la Constitucion, 18  
41071 Sevilla ES y  
UNIVERSIDAD DE CÓRDOBA (25.0%)**

72 Inventor/es:

**RODRÍGUEZ ARIZA, Antonio;  
LÓPEZ PEDRERA, Rosario;  
BARBARROJA PUERTO, Nuria y  
PÉREZ SÁNCHEZ, Carlos**

74 Agente/Representante:

**FÚSTER OLAGUIBEL, Gustavo Nicolás**

54 Título: **MicroARNs como biomarcadores para el diagnóstico del cáncer de pulmón**

57 Resumen:

MicroARN como biomarcadores para el diagnóstico del cáncer de pulmón.

Uso de los microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miARN-17 para el diagnóstico, clasificación y/o seguimiento del cáncer de pulmón, método de obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón, kit o dispositivo, microarray y usos.

**ES 2 646 826 A1**

**DESCRIPCIÓN****MicroARNs como biomarcadores para el diagnóstico del cáncer de pulmón****CAMPO DE LA INVENCION**

La presente invención se encuentra dentro del campo de la Biología Molecular y la Medicina  
5 Clínica, y específicamente se refiere a un método de obtención de datos útiles para el  
diagnóstico, clasificación y/o seguimiento del cáncer de pulmón.

**ANTECEDENTES DE LA INVENCION**

El cáncer de pulmón es una de las principales causas de muerte en los países  
industrializados, ocupando el segundo lugar solamente por detrás de las enfermedades  
10 cardiovasculares. El tabaco es el principal factor de riesgo, siendo el causante del 85-90%  
de todos los casos. A pesar de su alta incidencia y mortalidad, presenta una baja  
prevalencia debido a que la mayoría de los casos son detectados en estadios tardíos, por lo  
que sería muy importante disponer de pruebas de diagnóstico capaces de detectar la  
enfermedad lo antes posible, así como el tipo de tumor y la invasividad del mismo para  
15 orientar el tratamiento.

El análisis de condensación del aire exhalado en la respiración es una prueba no invasiva en  
la que el aire exhalado entra en un sistema de refrigeración para condensarse. Este aire  
condensado contiene proteínas y microARNs, que pueden constituir marcadores de  
patologías respiratorias, incluyendo el cáncer de pulmón (Conrad *et al.*, 2008. *Gen. Intern.*  
20 *Med.* 23 Suppl. 1, 78-84. doi: 10.1007/s11606-007-0411-1).

Hay muy pocos estudios proteómicos en el condensado del aire exhalado, y en los  
existentes se han conseguido identificar un número relativamente bajo de proteínas  
(Gianazza *et al.*, 2004. *Am. J. Med.* 117(1), 51-54; Fumagalli *et al.*, 2008. *Int. J. Mol. Sci.*  
13(11), 13894-13910; Kurova *et al.*, 2009. *Clin. Chem. Lab. Med.* 47(6), 706-712; Bloemen  
25 *et al.*, 2009. *Proteomics Clin. Appl.* 3(4), 498-504; Chang *et al.*, 2010. *Eur. Respir. J.* 35(5),  
1182-1215; Bloemen *et al.*, 2010. *Biomarkers* 15(7), 583-93). Solamente en dos estudios  
más recientes (Fumagalli *et al.*, 2012. *Int. J. Mol. Sci.* 13(11), 13894-13910.; Bredberg *et al.*,  
2012. *Clin. Chem.* 58(2), 431-440) se han conseguido identificar 44 y 124 proteínas  
diferentes, respectivamente. Sin embargo, ambos estudios analizaron “pools” de muestras, y  
30 no pacientes individuales. La gran mayoría de los estudios realizados hasta la fecha están  
relacionados con asma o con otras enfermedades pulmonares, como la enfermedad  
pulmonar crónica obstructiva (EPOC), y únicamente el estudio de Chang y cols., analizó

muestras de pacientes de cáncer de pulmón consiguiendo identificar 20 proteínas alteradas en todas las muestras analizadas (Chang *et al.*, 2010. *Eur. Respir. J.* 35(5), 1182-1215).

Los microARNs (miARN ó miR) son moléculas pequeñas de ARN, 20-25 nucleótidos, no codificantes para proteínas que regulan el procesamiento postranscripcional del ARN, mediante el apareamiento de bases con su ARN mensajero (ARNm) complementario, conduciendo a la represión de la transcripción o a la degradación del ARNm (Filipowicz *et al.*, 2008. *Nature Reviews Genetics*, 9(2),102-114). En la especie humana se han identificado aproximadamente 2500 microARNs, habiéndose demostrado su participación clave en la regulación de muchos procesos celulares y, su alteración está asociada con el desarrollo de diferentes patologías, especialmente cáncer.

Recientemente, se ha detectado la presencia de determinados microARNs en el esputo de pacientes con cáncer de pulmón y ha sido relacionada con el tipo de cáncer, destacando su papel en la diagnóstico de la enfermedad (Xie *et al.*, 2010. *Lung Cancer*, 67 (2), 170-176; Yu *et al.*, 2010. *International Journal of Cancer*, 127 (12), 2870-2878). Los microARNs también se han aislado en condensado de aire exhalado (Puneet y Faoud, 2013. *Journal of Analytical Sciences, Methods and Instrumentation*, 3, 17-29). En concreto elevados niveles de microARN-21 y microARN-486 se han encontrado diferencialmente expresados en condensado de aire exhalado de pacientes con cáncer de pulmón microcitico (Mozonni *et al.*, 2013. *Biomarkers.*, 18 (8), 679-686).

20

## BREVE DESCRIPCIÓN DE LA INVENCIÓN

Un **primer aspecto** de la invención se refiere al uso de los biomarcadores de microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17 para el diagnóstico, clasificación y/o seguimiento del cáncer de pulmón.

25 En

Un **segundo aspecto** de la invención se refiere a un método de obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón, de ahora en adelante primer método de la invención, que comprende:

30 a) cuantificar el producto de expresión de los biomarcadores miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y/o miR-17 en una muestra biológica aislada de dicho individuo. Preferiblemente la muestra biológica es condensado de aire exhalado.

En una realización preferida de este aspecto de la invención, el primer método de la invención además comprende:

b) calcular el índice A según la ecuación:

$$A = -0,083 + 0,172*(\text{miR-122}/\text{miR-29b}) + 0,009*(\text{miR-29b}/\text{miR-146}) - 0,038*(\text{miR-214}/\text{miR-146}) - 0,041*(\text{miR-193}/\text{miR-146}) + 0,004*(\text{miR-30}/\text{miR-146}) + 0,056*(\text{miR-146}/\text{miR-372}),$$

c) calcular el índice B según la ecuación:

$$B = 3,793 - 0,501*(\text{miR-17}/\text{miR-148}) + 0,167*(\text{miR-29b}/\text{miR-Let7c}) + 0,278*(\text{miR-214}/\text{miR-Let7c}) - 0,045*(\text{miR-148}/\text{miR-let7c}),$$

d) calcular el índice C según la ecuación:

$$C = 1,417 - 0,139*(\text{miR-17}/\text{miR-214}) + 0,05*(\text{miR-122}/\text{miR-148}) + 0,106*(\text{miR-214}/\text{miR-148}),$$

e) calcular el índice D según la ecuación:

$$D = 4,372 - 0,026*(\text{miR-17}/\text{miR-214}) - 0,049*(\text{miR-122}/\text{miR-214}) - 0,129*(\text{miR-29b}/\text{miR-214}) + 0,022*(\text{miR-214}/\text{miR-30})$$

y/o

f) calcular el índice E según la ecuación:

$$E = 16,12 - 0,630*(\text{miR-29b}/\text{miR-214}) + 0,157*(\text{miR-214}/\text{miR-30}),$$

asignando en cada ecuación los valores de los biomarcadores correspondientes obtenidos en el paso (a).

Un **tercer aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante segundo método de la invención, que comprende los pasos (a) y (b) según el primer método de la invención, y además comprende:

g) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón cuando el valor del índice A del paso (b) sea preferiblemente menor que 0,88, más preferiblemente menor que 0,75, y aún más preferiblemente menor que 0,579.

Un **cuarto aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante tercer método de la invención, que comprende los pasos (a) y (c) según el primer método de la invención, y además comprende:

- 5 h) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón de células no pequeñas de tipo adenocarcinoma y no epidermoide cuando el valor del índice B del paso (c) sea preferiblemente menor que 0,91, más preferiblemente menor que 0,72, y aún más preferiblemente menor que 0,50.

Un **quinto aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante cuarto método de la invención, que comprende los pasos (a) y (d) según el primer método de la invención, y además comprende:

- 15 i) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón de tumor invasivo cuando el valor del índice C del paso (d) sea preferiblemente mayor que 0,30, más preferiblemente mayor que 0,39, y aún más preferiblemente mayor que 0,65.

Un **sexto aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante quinto método de la invención, que comprende los pasos (a) y (e) según el primer método de la invención, y además comprende:

- 20 j) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón con metástasis cuando el valor del índice D del paso (e) sea preferiblemente mayor que 0,28, más preferiblemente mayor que 0,34, y aún más preferiblemente mayor que 0,45.

Un **séptimo aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante sexto método de la invención, que comprende los pasos (a) y (f) según el primer método de la invención, y además comprende:

- k) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón en estadio 4 cuando el valor del índice E del paso (f) sea preferiblemente mayor que 0,18, más preferiblemente mayor que 0,39, y aún más preferiblemente mayor que 0,64.

30 En una realización preferida de cualquiera de los aspectos de la invención, la muestra biológica aislada del paso (a) es una muestra de condensado de aire exhalado.

Un **octavo aspecto** de la invención se refiere a un kit o dispositivo, de ahora en adelante kit o dispositivo de la invención, que comprende los elementos necesarios para detectar los niveles de expresión de los microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17, tal y como se han definido estos en el primer aspecto de la invención.

En una realización preferida de este aspecto de la invención, el kit o dispositivo de la invención comprende cebadores, sondas y/o anticuerpos capaces de cuantificar el producto de expresión los microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17, y donde:

10 - los cebadores o *primers* son secuencias de polinucleótidos de entre 10 y 30 pares de bases, más preferiblemente de entre 15 y 25 pares de bases, aún más preferiblemente de entre 18 y 22 pares de bases, y aún mucho más preferiblemente de alrededor de 20 pares de bases, que presentan una identidad de al menos un 80%, más preferiblemente de al menos un 90%, aún más preferiblemente de al menos un 95%, aún mucho más preferiblemente de al menos un 98%, y particularmente de un 100%, con un fragmento de las secuencias complementarias a la SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

15 - las sondas son secuencias de polinucleótidos de entre 80 y 1100 pares de bases, más preferiblemente de entre 100 y 1000 pares de bases, y aún más preferiblemente de entre 200 y 500 pares de bases, que presentan una identidad de al menos un 80%, más preferiblemente de al menos un 90%, aún más preferiblemente de al menos un 95%, aún mucho más preferiblemente de al menos un 98%, y particularmente de un 100%, con un fragmento de las secuencias complementarias a la SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

20 - los anticuerpos son capaces de unirse a una región formada por cualquiera de las secuencias nucleotídicas SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID

Nº: 17, SEQ ID Nº: 18, SEQ ID Nº: 19, SEQ ID Nº: 20, SEQ ID Nº: 21, SEQ ID Nº: 22 y/o SEQ ID Nº: 23.

5 Un **noveno aspecto** de la invención se refiere a un microarray, de ahora en adelante microarray de la invención, que comprende oligonucleótidos o microarreglos de canal único diseñados a partir de las secuencias nucleotídicas SEQ ID Nº: 1, SEQ ID Nº: 2, SEQ ID Nº: 3, SEQ ID Nº: 4, SEQ ID Nº: 5, SEQ ID Nº: 6, SEQ ID Nº: 7, SEQ ID Nº: 8, SEQ ID Nº: 9, SEQ ID Nº: 10, SEQ ID Nº: 11, SEQ ID Nº: 12, SEQ ID Nº: 13, SEQ ID Nº: 14, SEQ ID Nº: 15, SEQ ID Nº: 16, SEQ ID Nº: 17, SEQ ID Nº: 18, SEQ ID Nº: 19, SEQ ID Nº: 20, SEQ ID Nº: 21, SEQ ID Nº: 22 y/o SEQ ID Nº: 23.

10 Así, por ejemplo, las secuencias de oligonucleótidos pueden ser construidas en la superficie de un chip mediante la elongación secuencial de una cadena en crecimiento con un sólo nucleótido utilizando fotolitografía. Así, los oligonucleótidos son anclados por el extremo 3' mediante un método de activación selectiva de nucleótidos, protegidos por un reactivo fotolábil, mediante la incidencia selectiva de luz a través de una fotomáscara. La  
15 fotomáscara puede ser física o virtual.

Así, las sondas de oligonucleótidos pueden ser de entre 10 y 100 nucleótidos, más preferiblemente, de entre 20 y 70 nucleótidos, y aún más preferiblemente, de entre 24 y 30 nucleótidos.

20 La síntesis *in situ* sobre un soporte sólido (por ejemplo, vidrio), podría hacerse mediante tecnología chorro de tinta (ink-jet), lo que requiere sondas más largas. Los soportes podrían ser, pero sin limitarse a, filtros o membranas de NC o nylon (cargadas), silicio, o Portas de vidrio para microscopios cubiertos con aminosilanos, polilisina, aldehídos o epoxy. La sonda es cada una de las muestras del chip. El target es la muestra a analizar: miARN, ARN mensajero, ARN total, un fragmento de PCR, etc.

25 En una realización preferida de este aspecto de la invención, el microarray de la invención presenta oligonucleótidos modificados, como se han descrito en el anterior aspecto de la invención.

30 Un **décimo aspecto** de la invención se refiere al uso del kit o dispositivo de la invención o del microarray de la invención, para la obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón.

**DESCRIPCIÓN DE LAS FIGURAS**

**Figura 1.** Diagrama de caja (Tukey) de la media de expresión ( $\text{Ln}2^{-\text{Ct}}$ ) de cada miARN diferencialmente expresados entre control y cáncer.

**Figura 2.** Diagrama de caja (Tukey) de los ratios de miARNs diferencialmente expresados entre control y cáncer.

**Figura 3.** Modelo de combinación de los ratios de microARNs que discrimina cáncer de control y curva ROC.

Ecuación resultante=  $-0,083+0,172*(\text{miR-122}/\text{miR-29b})+0,009*(\text{miR-29b}/\text{miR-146})-0,038*(\text{miR-214}/\text{miR-146})-0,041*(\text{miR-193}/\text{miR-146})+0,004*(\text{miR-30}/\text{miR-146})+0,056*(\text{miR-146}/\text{miR-372})$

**Figura 4.** Modelo de combinación de los ratios de microARNs que discriminan entre los tipos de tumor de células no pequeñas, Adenocarcinoma y Epidermoide. Ecuación resultante=  $3,793-0,501*(\text{miR-17}/\text{miR-148})+0,167*(\text{miR-29b}/\text{miR-Let7c})+0,278*(\text{miR-214}/\text{miR-Let7c})-0,045*(\text{miR-148}/\text{miR-Let7c})$

**Figura 5.** Modelo de combinación de los ratios de microARNs que discriminan entre invasividad o no del tumor.

Ecuación resultante=  $1,417-0,139*(\text{miR-17}/\text{miR-214})+0,05*(\text{miR-122}/\text{miR-148})+0,106*(\text{miR-214}/\text{miR-148})$

**Figura 6.** Modelo de combinación de los ratios de microARNs que discriminan entre la presencia o ausencia de metástasis.

Ecuación resultante=  $4,372-0,026*(\text{miR-17}/\text{miR-214})-0,049*(\text{miR-122}/\text{miR-214})-0,129*(\text{miR-29b}/\text{miR-214})+0,022*(\text{miR-214}/\text{miR-30})$

**Figura 7.** Modelo de combinación de los ratios de microARNs que discriminan entre el estadio 3 y 4 del tumor.

Ecuación resultante=  $16,12-0,630*(\text{miR-29b}/\text{miR-214})+0,157*(\text{miR-214}/\text{miR-372})$

**DESCRIPCIÓN DETALLADA DE LA INVENCION**

Los autores de la presente invención han analizado los niveles de microARNs en muestras de aire exhalado de pacientes con cáncer de pulmón en comparación con otros pacientes sanos y con diferentes factores de riesgo. Los resultados indican que los microARNs

presentes en el aire exhalado pueden ser usados como marcadores de un sistema diagnóstico, clasificación y/o seguimiento del cáncer de pulmón.

BIOMARCADORES DE LA INVENCION

Por tanto, un **primer aspecto** de la invención se refiere al uso de los biomarcadores de microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17, o cualquiera de sus combinaciones, para el diagnóstico, clasificación y/o seguimiento del cáncer de pulmón, de ahora en adelante biomarcadores de la invención. En una realización preferida, el uso de los microARNs es simultáneo.

Los biomarcadores de la invención pueden usarse simultáneamente, por separado o en cualquier combinación de los mismos para el diagnóstico, clasificación y/o seguimiento del cáncer de pulmón.

En una realización preferida de este aspecto de la invención, se utilizan conjuntos de ratios entre los biomarcadores de la invención para el diagnóstico, clasificación y/o seguimiento del cáncer de pulmón. Preferiblemente, los conjuntos de ratios formados entre biomarcadores de la invención para el diagnóstico, clasificación y/o seguimiento del cáncer de pulmón usados son:

- miR-122/miR-29b, miR-29b/miR-146, miR-193/miR-146, miR-146/miR-372, miR-214/miR-146 y/o miR-30/miR-146,
- miR-17/miR-148, miR-29b/miR-let7c, miR-214/miR-let7c y/o miR-148/ miR-let7c,
- miR-17/miR-214, miR-122/miR-148 y/o miR-214/miR-148,
- miR-17/miR-214, miR-122/miR-214, miR-29b/miR-214 y/o miR-214/miR-30,
- miR-29b/miR-214 y/o miR-214/miR-372

El término "microARN" se refiere a ARN monocatenario de aproximadamente 22 nucleótidos de longitud, no codificantes para proteínas y generados de transcritos endógenos que pueden formar estructuras en forma de horquilla (Lee *et al.*, 2004. *Embo. Journal*, 23 (20), 4051-4060). Los microARNs conforman una gran familia de genes reguladores postranscripcionales que controlan muchos procesos celulares y del desarrollo en eucariotas, cumpliendo una gran cantidad de funciones. Se estima que el 30% de todos los genes humanos son regulados por mecanismos dependientes de microARNs (Rajewsky N. 2006. *Nature Genetics*, 38, Suppl: S8-13) y que un solo microARN puede regular alrededor de 200 diferentes transcritos que pueden funcionar en diferentes vías en la célula (Krek *et al.*, 2005, *Nature Genetics*, 37 (5), 495-500), así como que un mismo ARNm puede ser

regulado por múltiples microARNs (Cai *et al.*, 2009. *Genomics Proteomics Bioinformatics*, 7 (4), 147-154).

El término "cáncer" pretende incluir cualquier miembro de una clase de enfermedades caracterizada por el crecimiento descontrolado de células aberrantes. El término incluye todos los cánceres y condiciones neoplásicas conocidos, caracterizados como malignos, benignos, tejido blandos o sólidos, y cánceres de todos los estadios y grados, incluyendo cánceres premetastásicos y postmetastásicos. Entre los ejemplos de diferentes tipos de cáncer se incluye, el cáncer de pulmón. Tal como se utiliza en la presente memoria, un "tumor" comprende una o más células cancerosas.

Los principales tipos de cáncer de pulmón son el cáncer de pulmón de células no pequeñas (de inglés *Non-small cell lung cancer* (NSCLC)) (≈ 85% de todos los cánceres de pulmón) y el cáncer de pulmón de células pequeñas (del inglés *Small cell lung cancer* (SCLC)) (≈ 15%). El cáncer de pulmón de células no pequeñas se puede dividir en tres grandes subtipos histológicos:

- Carcinoma de células escamosas, que supone del 25 al 30% de NSCLC. Este tipo es generalmente encontrado cerca de los bronquios, hacia el centro de la cavidad torácica (del pecho). Es también conocido como carcinoma epidermoide y está usualmente asociado a la exposición al humo de tabaco.
- Adenocarcinoma: conforma el 40% de todos los NSCLC. Este tipo de cáncer se encuentra generalmente en las regiones más externas del pulmón. También existe una forma rara de adenocarcinoma, llamado carcinoma broncoalveolar (*Bronchioalveolar Carcinoma* (BAC)) que se está viendo con mayor frecuencia a nivel mundial. BAC se disemina a lo largo de todo el pulmón a diferencia de otros tipos de cáncer de pulmón más comunes que forman tumores únicos. Se desconoce la causa del BAC. A pesar de presentarse en persona que fuman, generalmente se da en aquellas que nunca han fumado.
- Carcinoma de Células Grandes: conforma del 10% al 15% de NSCLC. Es de crecimiento rápido y puede aparecer en cualquier parte del pulmón., el cáncer de pulmón.

El tratamiento para cáncer de pulmón se basa en el tipo y estadio del cáncer. Los tratamientos, aunque sin limitarnos, pueden incluir: cirugía para remover el tumor, quimioterapia (medicamentos que matan o reducen el tamaño del tumor) o radiación (rayos X que destruyen o dañan las células cancerosas).

MÉTODOS DE LA INVENCION

Un **segundo aspecto** de la invención se refiere a un método de obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón, de ahora en adelante primer método de la invención, que comprende:

a) obtener una muestra biológica aislada de un individuo,

a') cuantificar el producto de expresión del miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y/o miR-17,

o alternativamente,

a) cuantificar el producto de expresión de los biomarcadores miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y/o miR-17 en una muestra biológica aislada de dicho individuo. Preferiblemente la muestra biológica es condensado de aire exhalado.

Aunque preferiblemente se cuantifica el producto de expresión de los 11 microARNs simultáneamente, para la obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón, la detección de cualquiera de ellos y, preferiblemente de las ratios: miR-122/miR-29b, miR-29b/miR-146, miR-193/miR-146, miR-146/miR-372, miR-214/miR-146, miR-30/miR-146, miR-17/miR-148, miR-29b/miR-let7c, miR-214/miR-let7c, miR-148/ miR-let7c, miR-17/miR-214, miR-122/miR-148, miR-214/miR-148, miR-17/miR-214, miR-122/miR-214, miR-29b/miR-214, miR-214/miR-30, miR-29b/miR-214 y/o miR-214/miR-372 resulta informativo.

El término "producto de expresión", también denominado "producto génico" se refiere al material bioquímico, ya sea ARN (por ejemplo microARN) o proteína, resultante de la expresión de un gen. Algunas veces se usa una medida de la cantidad de producto génico para inferir qué tan activo es un gen.

La cuantificación del producto de expresión de los microARNs, o la detección de la expresión de sus correspondientes genes, puede realizarse por cualquiera de los métodos conocidos en el estado de la técnica. Preferiblemente, la cuantificación se realiza mediante PCR a tiempo real (Q-RT-PCR).

La medida de los niveles de expresión de un gen, preferiblemente de manera cuantitativa, está basada en una señal que se obtiene directamente de los transcritos de dichos genes

(ARNm, microRNA, etc.), y que está correlacionada directamente con el número de moléculas de ARN producidas por los genes. Dicha señal – a la que también podemos referirnos como señal de intensidad – puede obtenerse, por ejemplo, midiendo un valor de intensidad de una propiedad química o física de dichos productos. Por otro lado, se podría

5 realizar mediante medida indirecta que se refiere a la medida obtenida de un componente secundario o un sistema de medida biológica (por ejemplo la medida de respuestas celulares, ligandos, “etiquetas” o productos de reacción enzimática).

La técnica preferiblemente utilizada es la llamada reacción en cadena de la polimerasa (PCR) en tiempo real. Según esta técnica, utilizando oligonucleótidos, enzimas, cebadores y

10 solución tampón, se amplifica de forma exponencial un fragmento de ADN original. Esta técnica es llevada a cabo en un aparato llamado termociclador que mantiene la temperatura necesaria en cada una de las etapas que conforman un ciclo. La detección en tiempo real se basa en la utilización de marcadores fluorescentes (sondas) que se unen a todas las

15 secuencias de doble cadena de ADN formadas en los ciclos de la reacción de PCR. Una vez que el marcador se une al ácido nucleico de doble cadena, éste emite una señal fluorescente que se procesa en tiempo real (Walter *et al.*, 2002, *Science*, 296, 557-559). El ciclo al cual la emisión de la intensidad del marcador fluorescente supera un umbral determinado es llamado Ct. Los niveles de expresión de un microARN son calculados como  $2^{-Ct}$ .

20 En una realización preferida de este aspecto de la invención, el primer método de la invención además comprende:

b) calcular el índice A según la ecuación:

$$A = -0,083 + 0,172*(\text{miR-122}/\text{miR-29b}) + 0,009*(\text{miR-29b}/\text{miR-146}) - 0,038*(\text{miR-214}/\text{miR-146}) - 0,041*(\text{miR-193}/\text{miR-146}) + 0,004*(\text{miR-30}/\text{miR-146}) + 0,056*(\text{miR-146}/\text{miR-372})$$

25 c) calcular el índice B según la ecuación:

$$B = 3,793 - 0,501*(\text{miR-17}/\text{miR-148}) + 0,167*(\text{miR-29b}/\text{miR-Let7c}) + 0,278*(\text{miR-214}/\text{miR-Let7c}) - 0,045*(\text{miR-148}/\text{miR-let7c})$$

d) calcular el índice C según la ecuación:

$$C = 1,417 - 0,139*(\text{miR-17}/\text{miR-214}) + 0,05*(\text{miR-122}/\text{miR-148}) + 0,106*(\text{miR-214}/\text{miR-148})$$

30 e) calcular el índice D según la ecuación:

$$D = 4,372 - 0,026*(\text{miR-17}/\text{miR-214}) - 0,049*(\text{miR-122}/\text{miR-214}) - 0,129*(\text{miR-29b}/\text{miR-214}) + 0,022*(\text{miR-214}/\text{miR-30})$$

y/o

f) calcular el índice E según la ecuación:

$$5 \quad E = 16,12 - 0,630*(\text{miR-29b}/\text{miR-214}) + 0,157*(\text{miR-214}/\text{miR-30})$$

asignando en cada ecuación los valores de los biomarcadores correspondientes obtenidos en el paso (a).

De esta manera, para el cálculo de los índices de la invención A-E hay que considerar que el valor de los cocientes entre los biomarcadores de microARNs (miRX/miRY) se obtiene mediante el  $\log_2(2^{-\text{CtX}}/2^{-\text{CtY}})$ , donde miRX o X se refiere al microARN indicado en el numerador y miRY o Y indicado microARN en el denominador.

Un **tercer aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante segundo método de la invención, que comprende los pasos (a) y (b) según el primer método de la invención, y además comprende:

g) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón cuando el valor del índice A del paso (b) sea preferiblemente menor que 0,88, más preferiblemente menor que 0,75, y aún más preferiblemente menor que 0,579.

Un **cuarto aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante tercer método de la invención, que comprende los pasos (a) y (c) según el primer método de la invención, y además comprende:

h) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón de células no pequeñas de tipo adenocarcinoma y no epidermoide cuando el valor del índice B del paso (c) sea preferiblemente menor que 0,91, más preferiblemente menor que 0,72, y aún más preferiblemente menor que 0,50.

Un **quinto aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante cuarto método de la invención, que comprende los pasos (a) y (d) según el primer método de la invención, y además comprende:

i) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón de tumor invasivo cuando el valor del índice C del paso (d) sea preferiblemente mayor que 0,30, más preferiblemente mayor que 0,39, y aún más preferiblemente mayor que 0,65.

5 En esta memoria se entiende por “cáncer o tumor invasivo” o “infiltrante” al cáncer que se disemina más allá de la capa de tejido en la que se inició y crece o progresa en tejidos sanos que la rodean. También se podría definir como la migración y la penetración directa de las células del cáncer en los tejidos vecinos.

10 Un **sexto aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante quinto método de la invención, que comprende los pasos (a) y (e) según el primer método de la invención, y además comprende:

j) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón con metástasis cuando el valor del índice D del paso (e) sea preferiblemente mayor que 0,28, más preferiblemente mayor que 0,34, y aún más preferiblemente mayor que 0,45.

15 El término “metástasis” se refiere a la capacidad de las células del cáncer de penetrar en los vasos sanguíneos y linfáticos, circular a través de de los mismos, y establecerse o crecer en tejidos normales de otra parte del cuerpo.

20 Un **séptimo aspecto** de la invención se refiere a un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón, de ahora en adelante sexto método de la invención, que comprende los pasos (a) y (f) según el primer método de la invención, y además comprende:

k) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón en estadio 4 cuando el valor del índice E del paso (f) sea preferiblemente mayor que 0,18, más preferiblemente mayor que 0,39, y aún más preferiblemente mayor que 0,64.

25 La “etapa o estadio” de un cáncer describe el grado de propagación de la enfermedad. Así, en el estadio 1 o T1: el tumor no mide más de 3 centímetros (cm), no ha alcanzado las membranas que rodean los pulmones (pleura visceral), y no afecta las ramas principales de los bronquios. En el estadio 2 o T2, el tumor presenta una o más de las siguientes características: a) Mide más de 3 cm, pero no mide más de 7 cm, b) Involucra a un bronquio principal, pero está a más de 2 cm de la carina (el punto donde la tráquea se divide en los  
30 bronquios principales izquierdo y derecho), c) ha crecido hacia el interior de las membranas que rodean a los pulmones (pleura visceral), d) el tumor obstruye parcialmente las vías

respiratorias, pero esto no ha causado el colapso de todo el pulmón ni la aparición de neumonía. En el estadio 3 o T3 el tumor presenta una o más de las siguientes características: a) su tamaño es mayor de 7 cm, b) ha crecido hacia el espesor de la pared del tórax, el músculo que separa el tórax del abdomen (diafragma), las membranas que rodean el espacio entre los dos pulmones (pleura mediastinal), o a las membranas del saco que rodea el corazón (pericardio parietal), c) el cáncer crece en un bronquio principal y está a menos de 2 cm de la carina, pero no afecta la carina en sí, d) ha crecido hacia el interior de las vías respiratorias lo suficiente como para causar el colapso total de un pulmón o neumonía en la totalidad del pulmón o e) dos o más nódulos tumorales separados se encuentran presentes en el mismo lóbulo de un pulmón. Un tumor de pulmón en estadio 4 o T4 presenta una o mas de las siguientes características: a) ha crecido hacia el espacio que existe entre los pulmones (mediastino), el corazón, los vasos sanguíneos grandes cercanos al corazón (tal como la aorta), la tráquea, el tubo que conecta la garganta con el estómago (esófago), la columna vertebral o la carina y/o b) dos o más nódulos tumorales separados se encuentran en lóbulos diferentes del mismo pulmón

En una realización preferida de cualquiera de los aspectos de la invención, la muestra biológica aislada del paso (a) es una muestra de condensado de aire exhalado.

Una "muestra biológica aislada" incluye, pero sin limitarnos a, células, tejidos y/o fluidos biológicos de un organismo, obtenidos mediante cualquier método conocido por un experto en la materia. Preferiblemente, la muestra biológica aislada del paso (a) es condensado de aire exhalado.

Los métodos de la invención pueden desarrollarse independientemente o conjuntamente, en cualquier combinación de los mismos.

Los pasos (a), (b), (c), (d), (e), (f), (g), (h), (i), (j) y/o (k) de los métodos descritos anteriormente pueden ser total o parcialmente automatizados, por ejemplo, por medio de un equipo robótico sensor para la cuantificación del paso (a) o el cálculo computarizado de cualquiera de los índices de los pasos (b), (c), (d), (e) y/o (f) o la clasificación computarizada en los pasos (g), (h), (i), (j) y/o (k).

#### *KIT O DISPOSITIVO DE LA INVENCION, MICROARRAY DE LA INVENCION Y USOS*

Un **octavo aspecto** de la invención se refiere a un kit o dispositivo, de ahora en adelante kit o dispositivo de la invención, que comprende los elementos necesarios para detectar los niveles de expresión de los microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372,

miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17, tal y como se han definido estos en el primer aspecto de la invención.

En una realización preferida de este aspecto de la invención, el kit o dispositivo de la invención comprende cebadores, sondas y/o anticuerpos capaces de cuantificar el producto de expresión los microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17, y donde:

- los cebadores o *primers* son secuencias de polinucleótidos de entre 10 y 30 pares de bases, más preferiblemente de entre 15 y 25 pares de bases, aún más preferiblemente de entre 18 y 22 pares de bases, y aún mucho más preferiblemente de alrededor de 20 pares de bases, que presentan una identidad de al menos un 80%, más preferiblemente de al menos un 90%, aún más preferiblemente de al menos un 95%, aún mucho más preferiblemente de al menos un 98%, y particularmente de un 100%, con un fragmento de las secuencias complementarias a la SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

- las sondas son secuencias de polinucleótidos de entre 80 y 1100 pares de bases, más preferiblemente de entre 100 y 1000 pares de bases, y aún más preferiblemente de entre 200 y 500 pares de bases, que presentan una identidad de al menos un 80%, más preferiblemente de al menos un 90%, aún más preferiblemente de al menos un 95%, aún mucho más preferiblemente de al menos un 98%, y particularmente de un 100%, con un fragmento de las secuencias complementarias a la SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

- los anticuerpos son capaces de unirse a una región formada por cualquiera de las secuencias nucleotídicas SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

En otra realización preferida de este aspecto de la invención, los anticuerpos están modificados. En otra realización preferida de este aspecto de la invención, el anticuerpo es humano, humanizado o sintético. En otra realización preferida, el anticuerpo es monoclonal y/o se encuentra marcado con un fluorocromo. Preferiblemente, el fluorocromo se selecciona de la lista que comprende Fluoresceína (FITC), Tetrametilrodamina y derivados, Ficoeritrina (PE), PerCP, Cy5, Texas, alofocianina, o cualquiera de sus combinaciones.

**Tabla 1.** microARNs de la invención.

Nombre del gen	Gen ID	Nombre del miARN	Secuencia del gen	Secuencia del miARN
microARN 146 ( <i>MIR146A</i> )	406938	hsa-miR-146a-5p	SEQ ID NO: 1	SEQ ID NO: 13
microARN 122 ( <i>MIR122</i> )	406906	hsa-miR122-5p	SEQ ID NO: 2	SEQ ID NO: 14
microARN 148a ( <i>MIR148A</i> )	406940	hsa-miR148a-3p	SEQ ID NO: 3	SEQ ID NO: 15
microARN 214 ( <i>MIR214</i> )	406996	hsa-miR214-3p	SEQ ID NO: 4	SEQ ID NO: 16
microRNA372 ( <i>MIR372</i> )	442917	hsa-miR372	SEQ ID NO: 5	SEQ ID NO: 17
microRNA let-7c ( <i>MIRLET7C</i> )	406885	hsa-miRlet7c	SEQ ID NO: 6	SEQ ID NO: 18
microRNA 30c ( <i>MIR30C</i> )	407031 407032	hsa-miR30c-5p	SEQ ID NO: 7 SEQ ID NO: 8	SEQ ID NO: 19
microRNA 19A ( <i>MIR19A</i> )	406979	hsa-miR19a-3p	SEQ ID NO: 9	SEQ ID NO: 20

microRNA 193b ( <i>MIR193B</i> )	574455	hsa-miR193b-3p	SEQ ID NO: 10	SEQ ID NO: 21
microRNA 29b-1 ( <i>MIR29B1</i> )	407024	hsa-miR29b-3p	SEQ ID NO: 11	SEQ ID NO: 22
microRNA17 ( <i>MIR17</i> )	406952	hsa-miR17-5p	SEQ ID NO: 12	SEQ ID NO: 23

En el contexto de la presente invención, los genes *MIR146A*, *MIR122*, *MIR148A*, *MIR214*, *MIR372*, *MIRLET7C*, *MIR30C*, *MIR19A*, *MIR193B*, *MIR29B1* y *MIR17* se definen también por una secuencia de nucleótidos o polinucleótido, que constituye la secuencia codificante de los microARNs recogidas respectivamente en las SEQ ID recogidas en la tabla 1, y que comprendería a diversas variantes procedentes de:

- a) moléculas de ácido nucleico que codifican un microARN que comprende la secuencia nucleotídica de la SEQ ID recogidas en la tabla 1,
- b) moléculas de ácido nucleico cuya cadena complementaria híbrida con la secuencia polinucleotídica de a),
- c) moléculas de ácido nucleico cuya secuencia difiere de a) y/o b) debido a la degeneración del código genético,
- d) moléculas de ácido nucleico que codifican un microARN que comprende la secuencia nucleotídica con una identidad de al menos un 80%, un 90%, un 95%, un 98% o un 99% con las SEQ ID recogidas en la tabla 1, respectivamente, y en las que el microARN codificado por dichos ácidos nucleicos posee la actividad y las características estructurales de los microARNs *MIR146A*, *MIR122*, *MIR148A*, *MIR214*, *MIR372*, *MIRLET7C*, *MIR30C*, *MIR19A*, *MIR193B*, *MIR29B1* y *MIR17*. Entre dichas moléculas de ácido nucleico se encuentra las recogidas en las SEQ ID indicadas en la tabla 1.

Secuencias de la invención:

SEQ ID NO: 1:

CCGATGTGTATCCTCAGCTTTGAGAACTGAATTCCATGGGTTGTGTCAGTGTCAGACCT  
CTGAAATTCAGTTCTTCAGCTGGGATATCTCTGTATCGT

SEQ ID NO: 2:

5 CCTTAGCAGAGCTGTGGAGTGTGACAATGGTGTGGTGTCTAAACTATCAAACGCCATT  
ATCACACTAAATAGCTACTGCTAGGC

SEQ ID NO: 3:

GAGGCAAAGTTCTGAGACACTCCGACTCTGAGTATGATAGAAGTCAGTGCACTACAGAA  
CTTTGTCTC

SEQ ID NO: 4:

10 GGCCTGGCTGGACAGAGTTGTCATGTGTCTGCCTGTCTACACTTGCTGTGCAGAACATC  
CGCTCACCTGTACAGCAGGCACAGACAGGCAGTCACATGACAACCCAGCCT

SEQ ID NO: 5:

GTGGGCCTCAAATGTGGAGCACTATTCTGATGTCCAAGTGAAAGTGCTGCGACATTTG  
AGCGTCAC

15

SEQ ID NO: 6:

GCATCCGGGTTGAGGTAGTAGGTTGTATGGTTTAGAGTTACACCCTGGGAGTTAACTGT  
ACAACCTTCTAGCTTTCCTTGGAGC

SEQ ID NO: 7:

20 ACCATGCTGTAGTGTGTGTAACATCCTACACTCTCAGCTGTGAGCTCAAGGTGGCTGG  
GAGAGGGTTGT TTAACCTTCTGCCATGGA

SEQ ID NO: 8:

AGATACTGTAAACATCCTACACTCTCAGCTGTGGAAAGTAAGAAAGCTGGGAGAAGGCT  
GTTTACTCTTTCT

25

SEQ ID NO: 9:

GCAGTCCTCTGTTAGTTTTGCATAGTTGCACTACAAGAAGAATGTAGTTGTGCAAATCTA  
TGCAAAACTGATGGTGGCCTGC

SEQ ID NO: 10:

GTGGTCTCAGAATCGGGGTTTTGAGGGCGAGATGAGTTTATGTTTTATCCAACCTGGCCC  
TCAAAGTCCCG CTTTTGGGGTCAT

SEQ ID NO: 11:

5 CTTCAGGAAGCTGGTTTCATATGGTGGTTTAGATTTAAATAGTGATTGTCTAGCACCATT  
TGAAATCAGTGTTCTTGGGGG

SEQ ID NO: 12:

GTCAGAATAATGTCAAAGTGCTTACAGTGCAGGTAGTGATATGTGCATCTACTGCAGTG  
AAGGCACTTGTAGCATTATGGTGAC

10 SEQ ID NO: 13:

UGAGAACUGAAUCCAUGGGUU

SEQ ID NO: 14:

UGGAGUGUGACAAUGGUGUUUG

SEQ ID NO: 15:

15 UCAGUGCACUACAGAACUUUGU

SEQ ID NO: 16:

ACAGCAGGCACAGACAGGCAGU

SEQ ID NO: 17:

CCUCAAAUGUGGAGCACUAUUCU

20 SEQ ID NO: 18:

UGAGGUAGUAGGUUGUAUGGUU

SEQ ID NO: 19:

UGUAAACAUCCUACACUCUCAGC

SEQ ID NO: 20:

25 AGUUUUGCAUAGUUGCACUACA

SEQ ID NO: 21:

CGGGGUUUUGAGGGCGAGAUGA

SEQ ID NO: 22:

UAGCACCAUUUGAAAUCAGUGUU

5 SEQ ID NO: 23:

CAAAGUGCUUACAGUGCAGGUAG

El kit además puede contener, sin ningún tipo de limitación, tampones, soluciones de extracción de proteínas, agentes para prevenir la contaminación, inhibidores de la degradación de las proteínas, etc.

10 Por otro lado, el kit o dispositivo de la invención puede incluir todos los soportes y recipientes necesarios para su puesta en marcha y optimización. Preferiblemente, el kit comprende además las instrucciones para llevar a cabo cualquiera de los métodos de la invención.

15 El kit de la invención puede incluir controles positivos y/o negativos. El kit además puede contener, sin ningún tipo de limitación, tampones, soluciones de extracción de proteínas, agentes para prevenir la contaminación, inhibidores de la degradación de las proteínas, etc.

20 En otra realización preferida de este aspecto de la invención, el kit o dispositivo de la invención es un kit de partes, que comprende un componente A, formado por un dispositivo para la recogida de la muestra del paso a), y un componente B, formado por los elementos necesarios para llevar a cabo el análisis cuantitativo en la muestra del paso a) o cualquiera de los métodos de la invención.

25 En una realización preferida el kit o dispositivo de la invención comprende oligonucleótidos. En otra realización preferida, los oligonucleótidos presentan modificaciones en alguno de sus nucleótidos, como por ejemplo, pero sin limitarnos a, nucleótidos que tengan alguno de sus átomos con un isótopo radiactivo, normalmente  $^{32}\text{P}$  o tritio, nucleótidos marcados inmunológicamente, como por ejemplo con una molécula de digoxigenina, y/o inmovilizadas en una membrana. Varias posibilidades son conocidas en el estado de la técnica.

Un **noveno aspecto** de la invención se refiere a un microarray, de ahora en adelante microarray de la invención, que comprende oligonucleótidos o microarreglos de canal único diseñados a partir de las secuencias nucleotídicas SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, 5 SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

Así, por ejemplo, las secuencias de oligonucleótidos pueden ser construidas en la superficie un chip mediante el elongamiento secuencial de una cadena en crecimiento con un sólo 10 nucleótido utilizando fotolitografía. Así, los oligonucleótidos son anclados por el extremo 3' mediante un método de activación selectiva de nucleótidos, protegidos por un reactivo fotolábil, mediante la incidencia selectiva de luz a través de una fotomáscara. La fotomáscara puede ser física o virtual.

Así, las sondas de oligonucleótidos pueden ser de entre 10 y 100 nucleótidos, más 15 preferiblemente, de entre 20 y 70 nucleótidos, y aún más preferiblemente, de entre 24 y 30 nucleótidos.

La síntesis *in situ* sobre un soporte sólido (por ejemplo, vidrio), podría hacerse mediante tecnología chorro de tinta (ink-jet), lo que requiere sondas más largas. Los soportes podrían ser, pero sin limitarse a, filtros o membranas de NC o nylon (cargadas), silicio, o Portas de 20 vidrio para microscopios cubiertos con aminosilanos, polilisina, aldehídos o epoxy. La sonda es cada una de las muestras del chip. El target es la muestra a analizar: miARN, ARN mensajero, ARN total, un fragmento de PCR, etc.

En una realización preferida de este aspecto de la invención, el microarray de la invención 25 presenta oligonucleótidos modificados, como se han descrito en el anterior aspecto de la invención.

Un **décimo aspecto** de la invención se refiere al uso del kit o dispositivo de la invención o del microarray de la invención, para la obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón.

30 La invención se extiende también a programas de ordenador adaptados para que cualquier medio de procesamiento pueda llevar a la práctica los métodos de la invención. Tales programas pueden tener la forma de código fuente, código objeto, una fuente intermedia de código y código objeto, por ejemplo, como en forma parcialmente compilada, o en cualquier

otra forma adecuada para uso en la puesta en práctica de los procesos según la invención. Los programas de ordenador también abarcan aplicaciones en la nube basadas en dicho procedimiento.

5 Un **décimo primer aspecto** de la invención se refiere a un programa de ordenador que comprende instrucciones de programa para hacer que un ordenador lleve a la práctica el procedimiento de acuerdo con cualquiera de los métodos de la invención.

10 En particular, la invención abarca programas de ordenador dispuestos sobre o dentro de una portadora. La portadora puede ser cualquier entidad o dispositivo capaz de soportar el programa. Cuando el programa va incorporado en una señal que puede ser transportada directamente por un cable u otro dispositivo o medio, la portadora puede estar constituida por dicho cable u otro dispositivo o medio. Como variante, la portadora podría ser un circuito integrado en el que va incluido el programa, estando el circuito integrado adaptado para ejecutar, o para ser utilizado en la ejecución de, los procesos correspondientes.

15 Por ejemplo, los programas podrían estar incorporados en un medio de almacenamiento, como una memoria ROM, una memoria CD ROM o una memoria ROM de semiconductor, una memoria USB, o un soporte de grabación magnética, por ejemplo, un disco flexible o un disco duro. Alternativamente, los programas podrían estar soportados en una señal portadora transmisible. Por ejemplo, podría tratarse de una señal eléctrica u óptica que podría transportarse a través de cable eléctrico u óptico, por radio o por cualesquiera otros  
20 medios.

Un **décimo segundo aspecto** de la invención se refiere a un medio de almacenamiento legible por un ordenador que comprende instrucciones de programa capaces de hacer que un ordenador lleve a cabo los pasos de cualquiera de los métodos de la invención.

25 Un **décimo tercer aspecto** de la invención se refiere a una señal transmisible que comprende instrucciones de programa capaces de hacer que un ordenador lleve a cabo los pasos de cualquiera de los métodos de la invención.

30 Una "secuencia de ácido nucleico o polinucleótido" puede comprender las cinco bases que aparecen biológicamente (adenina, guanina, timina, citosina y uracilo) y/o bases distintas de las cinco que aparecen biológicamente. Estas bases pueden servir para distintos propósitos, por ejemplo, para estabilizar o desestabilizar la hibridación; para estimular o inhibir la degradación de la sonda; o como puntos de unión para restos detectables o restos de apantallamiento. Por ejemplo, un polinucleótido de la invención puede contener uno o más restos de base modificados, no estándar, derivatizados, incluyendo, pero sin limitarse a, N<sup>6</sup>-

metil-adenina, N<sup>6</sup>-terc-butil-bencil-adenina, imidazol, imidazoles sustituidos, 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroximetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-metiladenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxycarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopenteniladenina, ácido uracil-5-oxiacético, wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo (es decir, timina), éster metílico del ácido uracil-5-oxiacético, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)<sub>w</sub>, 2,6-diaminopurina, y 5-propinil pirimidina. Otros ejemplos de restos de bases modificados, no estándar, o derivatizados pueden encontrarse en las Patentes de EEUU Nos. 6.001.611; 5.955.589; 5.844.106; 5.789.562; 5.750.343; 5.728.525; y 5.679.785. Además, una secuencia de ácido nucleico o polinucleótido puede comprender uno o más restos de azúcares modificados incluyendo, pero sin limitarse a, arabinosa, 2-fluoroarabinosa, xilulosa, y una hexosa.

Los términos "polinucleótido" y "ácido nucleico" se usan aquí de manera intercambiable, refiriéndose a formas poliméricas de nucleótidos de cualquier longitud, tanto ribonucleótidos (ARN ó RNA) como desoxirribonucleótidos (ADN ó DNA).

Los términos "secuencia aminoacídica", "péptido", "oligopéptido", "polipéptido" y "proteína" se usan aquí de manera intercambiable, y se refieren a una forma polimérica de aminoácidos de cualquier longitud, que pueden ser codificantes o no codificantes, química o bioquímicamente modificados.

En la presente invención se entiende por variante o fragmento biológicamente activo, aquellas variantes o fragmentos de los péptidos indicados que tienen un efecto fisiológico, metabólico o inmunológico igual, o presentan la misma utilidad que los descritos. Esto es, son funcionalmente equivalentes. Dichos efectos se pueden determinar mediante métodos convencionales.

El término "identidad", tal y como se utiliza en esta memoria, hace referencia a la proporción de nucleótidos o aminoácidos idénticos entre dos secuencias nucleotídicas o aminoacídicas que se comparan. Los métodos de comparación de secuencias son conocidos en el estado de la técnica, e incluyen, aunque sin limitarse a ellos, el programa GAG, incluyendo GAP (Devereux *et al.*, *Nucleic Acids Research* 12: 287 (1984) *Genetics Computer Group*

*University of Wisconsin, Madison, (WI); BLAST, BLASTP o BLASTN, y FASTA (Altschul et al., 1999. J. Mol. Biol. 215: 403-410.*

El término "individuo" o "sujeto", tal y como se utiliza en la descripción, se refiere a un animal, preferiblemente a un mamífero, y más preferiblemente a un ser humano. El término "individuo" en esta memoria, es sinónimo de "paciente", y no pretende ser limitativo en ningún aspecto, pudiendo ser éste de cualquier edad, sexo y condición física. El individuo puede ser cualquiera, un individuo predispuesto a una enfermedad (por ejemplo, cáncer de pulmón) o un individuo que padece dicha enfermedad.

A lo largo de la descripción y las reivindicaciones la palabra "comprende" y sus variantes no pretenden excluir otras características técnicas, aditivos, componentes o pasos. Para los expertos en la materia, otros objetos, ventajas y características de la invención se desprenderán en parte de la descripción y en parte de la práctica de la invención. Los siguientes ejemplos y dibujos se proporcionan a modo de ilustración, y no se pretende que sean limitativos de la presente invención.

#### 15 **EJEMPLO DE LA INVENCION**

El objeto de esta patente es la utilización de los niveles de microARNs circulantes en condensado de aire exhalado como biomarcadores para la detección de cáncer de pulmón.

Para ello se aisló ARN mediante miRNAeasy serum/plasma kit (qiagen) de 200 µl de aire exhalado de vías inferiores de 127 sujetos, divididos en 4 grupos:

- 20 - Grupo I: 32 Pacientes diagnosticados de cáncer de pulmón con distintos estadios de la enfermedad, procedentes del servicio de Oncología del Hospital Reina Sofía, Córdoba
- Grupo II: 31 Pacientes diagnosticados con EPOC procedentes del Servicio de Neumología del Hospital Reina Sofía, Córdoba.
- 25 - Grupo III: 32 Controles sanos con Factor de riesgo, siendo fumadores o ex fumadores procedentes del servicio de Medicina Interna del Hospital Reina Sofía, Córdoba.
- Grupo IV: 32 Sujetos sanos, procedentes del Servicio de Medicina Interna del Hospital Reina Sofía, Córdoba.

30 El ARN total de las muestras incluyendo la fracción de los microARNs se aisló a partir de 200 µl de aire exhalado condensado utilizando el kit QIAzol miRNeasy (Qiagen, Valencia, CA, USA). Los 200 µl de aire exhalado se incubaron con 1 ml de QIAzol Lysis Reagent

(Qiagen) para inhibir la función de las enzimas ARNasas. El protocolo se realizó siguiendo exhaustivamente las recomendaciones del fabricante. Finalmente, el ARN aislado se diluyó en 14 µl de agua libre de ARNasas y se guardó a -80°C.

5 Se realizó un PCR array de microARNs (miScript miRNA PCR array cancer, Quiagen; siguiendo las instrucciones del fabricante) descritos en distintos tipos de tumores: cerebral, colon, gástrico, cabeza, cuello, hígado, riñón, linfoma, pancreático, próstata, piel, ovario, pecho y útero. Con esta técnica se evaluó la expresión de 84 microARNs en aire exhalado de vías inferiores de un pool de 17 controles sanos y 17 pacientes con cáncer (Grupo I y grupo IV).

10 Una vez realizado el array y obtenidas las Cts, los resultados de la PCR fueron normalizados con dos *housekeepings* diferentes:

- Tomando como referencia el valor del microARN sintético Cel-miRNA-39.
- Tomando de referencia el valor del gen SNORD68 que se expresaba de manera constitutiva en nuestras muestras.

15 Los niveles de expresión relativa se calcularon según el método de  $2^{-Ct}$ . Utilizando un punto de corte de 2 veces de cambio (2-Fold change), observamos que 40 microARNs estaban incrementados y 9 reducidos en las muestras de cáncer comparadas con los sujetos sanos. Los resultados mostraron que 11 microARNs se expresaban solamente en los pacientes con tumor o se hallaban incrementados en cáncer; en todos ellos las veces de cambio de  
20 expresión respecto al control eran mayores de 10. Algunos de estos microARN se han descrito en orina y plasma de pacientes con cáncer de pulmón, otros se han descrito en otros tipos de tumores, pero ninguno en aire exhalado.

La expresión de estos 11 microARNs se validó mediante RT-PCR de forma independiente en cada sujeto de los cuatro grupos: casos, controles sin factor de riesgo, controles con  
25 factor de riesgo y EPOC. Los 11 microARNs analizados fueron: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17.

Para realizar las PCRs, la metodología seguida es la que comprende TaqMan MicroRNA Assay kit (Applied Biosystem), estandarizada en nuestro grupo y la que se encuentra en la mayoría de las publicaciones relacionadas con microARNs. El protocolo consiste en una  
30 primera retrotranscripción (RT) del 11 ng de ARN (cuantificado en el nanodrop) con *stem-loop primers* específicos de microARNs, seguida de una *real-time* PCR con sondas TaqMan específicas para cada microARN que se pretenda amplificar.

Cada reacción de retrotranscripción se realizó en 15 µl conteniendo 1,5 µl de 10 X RT buffer 1, 0,2 µl de inhibidor de ARNasas (20 U/ml), 0,15 µl de 100 mmol/L de “deoxynucleoside triphosphates”, 1 µl de 50 U/µl MultiScribe transcriptasa inversa, 1,5 µl de cada RT primers (4 miARN RT primers/reacción de RT), 11 ng de ARN y agua libre de ARNasas. La reacción se llevó a cabo en un termociclador GeneAmp PCR System 9700 (Applied BioSystems) a 16°C durante 30 minutos, 42°C durante 30°C y 85°C 5 minutos. Cada preamplificación se hizo con una mezcla de 4 sondas de 4 microARNs. Así realizamos 3 preamplificaciones para realizar posteriormente las PCRs de los 11 microARNs seleccionados.

Después se procedió con la reacción de preamplificación. Para ello, 2,5 µl de producto de RT diluido (2.5 µl en 10 µl de agua libre de ARNasas) se combinó con 5 µl de PCR Master Mix Taqman 2x, No AmpErase UNG y 2,5 µl de un pool de 4 primers Taqman miRNA assay (0.2x). La reacción se realizó en un termociclador GeneAmp PCR System 9700 (Applied BioSystems) a 95°C durante 10 minutos, más 20 ciclos a 95°C durante 15 segundos y 60°C durante 4 minutos.

Posteriormente el producto de preamplificación se diluyó en 60 µl de agua libre de ARNasas para proceder a la reacción de PCR. 4 µl de este producto de preamplificación y se combinó con 5 µl de Taqman 2x Universal PCR Master Mix, No AmpErase UNG, con 0,5 µl del primer específico 20X Taqman miRNA Assay y 0,5 µl de agua destilada. La PCR se realizó en el termociclador de Roche LightCycler 480 a 95°C durante 10 min, seguido de 40 ciclos a 95°C durante 15 s y 60°C durante 1 min. Los niveles de expresión de los miARNs se calcularon utilizando el método  $2^{-Ct}$

Así se realizó las PCRs de estos 11 microARNs en las muestras de aire exhalado de vías inferiores de 127 sujetos, se procedió al análisis de estos resultados.

Teniendo en cuenta la ausencia de un método establecido de normalización de datos de expresión de microARNs circulantes, el análisis se llevó a cabo teniendo en cuenta los distintos modelos publicados hasta el momento (Mestdagh *et al.*, 2009. *Genome Biology*, Vol.10, Nº 6, PP R6; Boeri *et al.*, 2011. *Proceedings of the National Academy of Sciences*, Vol. 108, Nº 9, PP 3713-3718).

Utilizando el método de normalización de la expresión utilizando como *housekeeping* la media de los miARNs expresados en todas las muestras (Mestdagh *et al.*, 2009. *Genome Biology*, Vol.10, Nº 6, PP R6) hallamos que los niveles del miR-29b y miR-146a estaban diferencialmente expresados de forma significativa en los pacientes con cáncer respecto al grupo control (Figura 1).

Utilizando un segundo análisis (Boeri *et al.*, 2011. *Proceedings of the National Academy of Sciences*, Vol. 108, Nº 9, PP 3713-3718), el nivel de expresión de cada miARN se calculó mediante el método de  $2^{-Ct}$ . Después se calcularon los distintos ratios recíprocos, considerando un solo ratio para cada par de miARNs y expresado como log2. De esta forma, hallamos que una firma de 7 microARNs compuesta por los microARNs 122, 29b, 146, 193, 372, 214, 30 generaban una serie de ratios que estaban diferencialmente expresados entre Control y Cáncer (Tabla 2; Figura 2).

**Tabla 2.** Medias y SEM de los ratios de miARNs diferencialmente expresados entre Control y Cáncer

miR Ratios (Log2)	Control (Media ± SEM)	Cáncer (Media ± SEM)	P
miR-122/miR-29b	-1,1 ± 0,4	-4,3 ± 1,2	0,01
miR-29b/miR-146	7,9 ± 1,1	12,8 ± 1,3	0,00
miR-193/miR-146	-8,6 ± 1,5	-3,6 ± 1,4	0,02
miR-146/miR-372	12,5 ± 1,2	8,2 ± 1,7	0,04
miR-214/miR-146	-12,2 ± 1,5	-8,1 ± 1,4	0,03
miR-30/miR-146	2 ± 1,3	5,4 ± 1,2	0,05

10

Asimismo, generamos un modelo que integraba la combinación de los ratios diferencialmente expresados (regresión logística binaria), el cual era capaz de discriminar pacientes con cáncer respecto a controles con un 80% de Especificidad, un 70% de Sensibilidad y un área bajo la curva de 0.76. Este modelo se ajusta a los datos de microARNs en nuestra cohorte de individuos. Sería necesaria su validación en otra cohorte de pacientes independiente (Figura 3).

15

Por otro lado encontramos que dentro del grupo de cáncer, existían distintas firmas de microARNs cuyos ratios se asociaron a determinadas características de la clínica de los pacientes con cáncer de pulmón como tipo de tumor, invasión, metástasis y estadio.

#### 20 Tipos de Tumor

Una firma de 5 microARNs compuesta por el miR-17, -29b, -214, -148 y -let7c generaban ratios que estaban significativamente diferencialmente expresados entre los subtipos de cáncer adenocarcinoma y epidermoide (Tabla 3):

**Tabla 3.** Media y SEM de los ratios de miARNs diferencialmente expresados entre el subtipo de tumor Adenocarcinoma y Epidermoide

<b>miR Ratios (Log2)</b>	<b>Adenocarcinoma (Media ± SEM)</b>	<b>Epidermoide (Media ± SEM)</b>	<b>P</b>
miR-17/miR-148	16,4 ± 1,6	9,7 ± 3,05	0,05
miR-29b/miR-let7c	15,9 ± 2,1	22,7 ± 1,6	0,02
miR-214/miR-let7c	-7,1 ± 1,9	1,5 ± 2,8	0,01
miR-148/ miR-let7c	-6,4 ± 2,2	2,6 ± 2,6	0,01

El modelo de combinación de estos ratios permitía discriminar el tipo de tumor Adenocarcinoma del Epidermoide con una Especificidad del 91%, una Sensibilidad del 87% y un Área bajo la curva de 0.92. Este modelo se ajustó exclusivamente a los datos de miARNs y tipo de tumor de nuestra cohorte de 32 pacientes con cáncer (Figura 4).

#### Invasividad

Una firma de 4 microARNs compuesta por el miR-17, -214, -122 y -148 generaban ratios que estaban diferencialmente expresados de forma significativa entre la invasividad o no del tumor (Tabla 4):

**Tabla 4.** Media y SEM de los ratios microARN diferencialmente expresados entre la invasividad o no del tumor

<b>miR Ratios (Log2)</b>	<b>No Invasión (Media ± SEM)</b>	<b>Invasión (Media ± SEM)</b>	<b>P</b>
miR-17/miR-214	17,1 ± 2	12,1 ± 1,1	0,04
miR-122/miR-148	14,5 ± 2,8	19,4 ± 1,7	0,05
miR-214/miR-148	-3,1 ± 0,9	2,5 ± 2,1	0,02

El modelo de combinación de estos ratios permitía discriminar la invasividad o no del tumor con una Especificidad del 91 %, una Sensibilidad del 64 % y un Área bajo la curva de 0.85. Este modelo se ajustó exclusivamente a los datos de microARN e invasividad del tumor de nuestra cohorte de 32 pacientes con cáncer (Figura 5).

Metástasis

Una firma de 5 microARNs compuesta por el miR-17, -214, -122, -29b y -30 generaban ratios que estaban diferencialmente expresados de forma significativa entre la presencia o no de metástasis (Tabla 5):

5 **Tabla 5.** Medias y SEM de los ratios de microARNs diferencialmente expresados entre la presencia o ausencia de metástasis

<b>miR Ratios (Log2)</b>	<b>No Metástasis (Media ± SEM)</b>	<b>Metástasis (Media ± SEM)</b>	<b>P</b>
miR-17/miR-214	16,1 ± 1,6	10,3 ± 2,2	0,04
miR-122/miR-214	18,8 ± 1,6	13 ± 2,2	0,05
miR-29b/miR-214	23,6 ± 1,1	17,1 ± 2	0,02
miR-214/miR-30	-16 ± 1,6	-10,3 ± 2,3	0,02

10 El modelo de combinación de estos ratios permitía discriminar la presencia o no de metástasis con una Especificidad del 75 %, una Sensibilidad del 71 % y un Área bajo la curva de 0.78. Este modelo se ajustó exclusivamente a los datos de microARN y metástasis de nuestra cohorte de 32 pacientes con cáncer.

Estadio

15 Una firma de 3 microARNs compuesta por el miR-29b, -214, y -372 generaban ratios que estaban diferencialmente expresados de forma significativa entre la presencia o no de metástasis (Tabla 6):

**Tabla 6.** Medias y SEM de los ratios de microARNs diferencialmente expresados entre el estadio 3 y 4 del tumor

<b>miR Ratios (Log2)</b>	<b>Estadio 3 (Media ± SEM)</b>	<b>Estadio 4 (Media ± SEM)</b>	<b>P</b>
miR-29b/miR-214	25,4 ± 0,89	19,2 ± 1,73	0,00
miR-214/miR-372	-3,9 ± 1,8	2,07 ± 2,3	0,03

20 El modelo de combinación de estos ratios permitía discriminar entre el estadio 3 y 4 del tumor con una Especificidad del 72 %, una Sensibilidad del 78 % y un Área bajo la curva de

0.84. Este modelo se ajustó exclusivamente a los datos de microARN y estadio del tumor de nuestra cohorte de 32 pacientes con cáncer.

**REIVINDICACIONES**

1.- Uso simultáneo de los biomarcadores de microARNs: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17 para el diagnóstico, clasificación y/o seguimiento del cáncer de pulmón.

5 2.- Un método de obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón, que comprende:

a) cuantificar el producto de expresión de los biomarcadores miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b  
10 y/o miR-17 en una muestra biológica aislada de dicho individuo.

3.- El método según la reivindicación anterior, que además comprende:

b) calcular el índice A según la ecuación:

$$A = -0,083 + 0,172*(miR-122/miR-29b) + 0,009*(miR-29b/miR-146) - 0,038*(miR-214/miR-146) - 0,041*(miR-193/miR-146) + 0,004*(miR-30/miR-146) + 0,056*(miR-146/miR-372),$$

15

c) calcular el índice B según la ecuación:

$$B = 3,793 - 0,501*(miR-17/miR-148) + 0,167*(miR-29b/miR-Letc7c) + 0,278*(miR-214/miR-Let7c) - 0,045*(miR-148/miR-let7c),$$

d) calcular el índice C según la ecuación:

$$C = 1,417 - 0,139*(miR-17/miR-214) + 0,05*(miR-122/miR-148) + 0,106*(miR-214/miR-148),$$

20

e) calcular el índice D según la ecuación:

$$D = 4,372 - 0,026*(miR-17/miR-214) - 0,049*(miR-122/miR-214) - 0,129*(miR-29b/miR-214) + 0,022*(miR-214/miR-30)$$

25 y/o

f) calcular el índice E según la ecuación:

$$E = 16,12 - 0,630*(miR-29b/miR-214) + 0,157*(miR-214/miR-30),$$

asignando en cada ecuación los valores de los biomarcadores correspondientes obtenidos en el paso (a).

4.- Un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón que comprende los pasos (a) y (b) según cualquiera de las reivindicaciones 2-3, y además comprende:

g) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón cuando el valor del índice A del paso (b) sea preferiblemente menor que 0,88, más preferiblemente menor que 0,75, y aún más preferiblemente menor que 0,579.

5.- Un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón que comprende los pasos (a) y (c) según cualquiera de las reivindicaciones 2-3, y además comprende:

h) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón de células no pequeñas de tipo adenocarcinoma y no epidermoide cuando el valor del índice B del paso (c) sea preferiblemente menor que 0,91, más preferiblemente menor que 0,72, y aún más preferiblemente menor que 0,50.

6.- Un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón que comprende los pasos (a) y (d) según cualquiera de las reivindicaciones 2-3, y además comprende:

i) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón de tumor invasivo cuando el valor del índice C del paso (d) sea preferiblemente mayor que 0,30, más preferiblemente mayor que 0,39, y aún más preferiblemente mayor que 0,65.

7.- Un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón que comprende los pasos (a) y (e) según cualquiera de las reivindicaciones 2-3, y además comprende:

j) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón con metástasis cuando el valor del índice D del paso (e) sea preferiblemente mayor que 0,28, más preferiblemente mayor que 0,34, y aún más preferiblemente mayor que 0,45.

8.- Un método para la clasificación, diagnóstico y seguimiento del cáncer de pulmón que comprende los pasos (a) y (f) según cualquiera de las reivindicaciones 2-3, y además comprende:

k) clasificar al individuo del paso (a) en el grupo de individuos con cáncer de pulmón en estadio 4 cuando el valor del índice E del paso (f) sea preferiblemente mayor que 0,18, más preferiblemente mayor que 0,39, y aún más preferiblemente mayor que 0,64.

5 9.- El método según cualquiera de las reivindicaciones 2-8, donde la muestra biológica aislada del paso (a) es una muestra de condensado de aire exhalado.

10.- Un kit o dispositivo según la reivindicación anterior, que comprende cebadores, sondas y/o anticuerpos capaces de cuantificar el producto de expresión los microARN: miR-146a, miR-122, miR-148a, miR-214, miR-372, miR-let7c, miR-30c, miR-19a, miR-193b, miR-29b y miR-17, y donde:

- los cebadores o primers son secuencias de polinucleótidos de entre 10 y 30 pares de bases, más preferiblemente de entre 15 y 25 pares de bases, aún más preferiblemente de entre 18 y 22 pares de bases, y aún mucho más preferiblemente de alrededor de 20 pares de bases, que presentan una identidad de al menos un 80%, más preferiblemente de al menos un 90%, aún más preferiblemente de al menos un 95%, aún mucho más preferiblemente de al menos un 98%, y particularmente de un 100%, con un fragmento de las secuencias complementarias a la SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

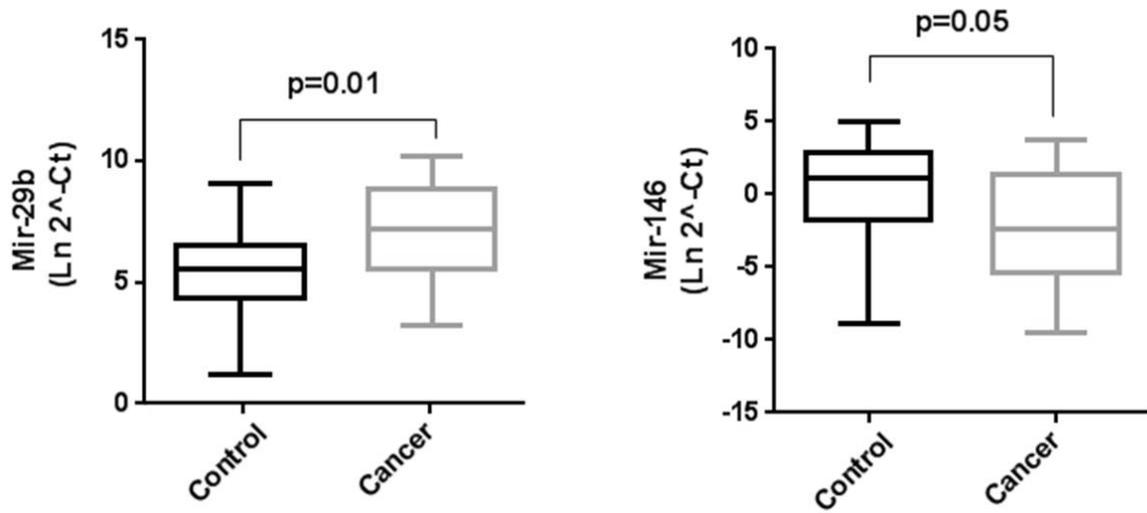
- las sondas son secuencias de polinucleótidos de entre 80 y 1100 pares de bases, más preferiblemente de entre 100 y 1000 pares de bases, y aún más preferiblemente de entre 200 y 500 pares de bases, que presentan una identidad de al menos un 80%, más preferiblemente de al menos un 90%, aún más preferiblemente de al menos un 95%, aún mucho más preferiblemente de al menos un 98%, y particularmente de un 100%, con un fragmento de las secuencias complementarias a la SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

- los anticuerpos son capaces de unirse a una región formada por cualquiera de las secuencias nucleotídicas SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°:

10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

5 11.- Un microarray que comprende oligonucleótidos o microarreglos de canal único diseñados a partir de las secuencias nucleotídicas SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22 y/o SEQ ID N°: 23.

10 12.- Uso del kit o dispositivo tal y como se define en la reivindicación 10 o de un microarray tal y como se define en la reivindicación 11, para la obtención de datos útiles para el diagnóstico, clasificación y/o seguimiento de un individuo o sujeto que potencialmente sufra cáncer de pulmón.



**Mir 29b (Media $\pm$ SEM): Control= 5.3  $\pm$  0.3; Cancer= 6.6  $\pm$  0.4; p=0.01**  
**MiR 146 (Media $\pm$ SEM): Control= -0.2 $\pm$ 0.6 ; Cancer= -2.07 $\pm$ 0.7; p=0.05**

**Fig.1**

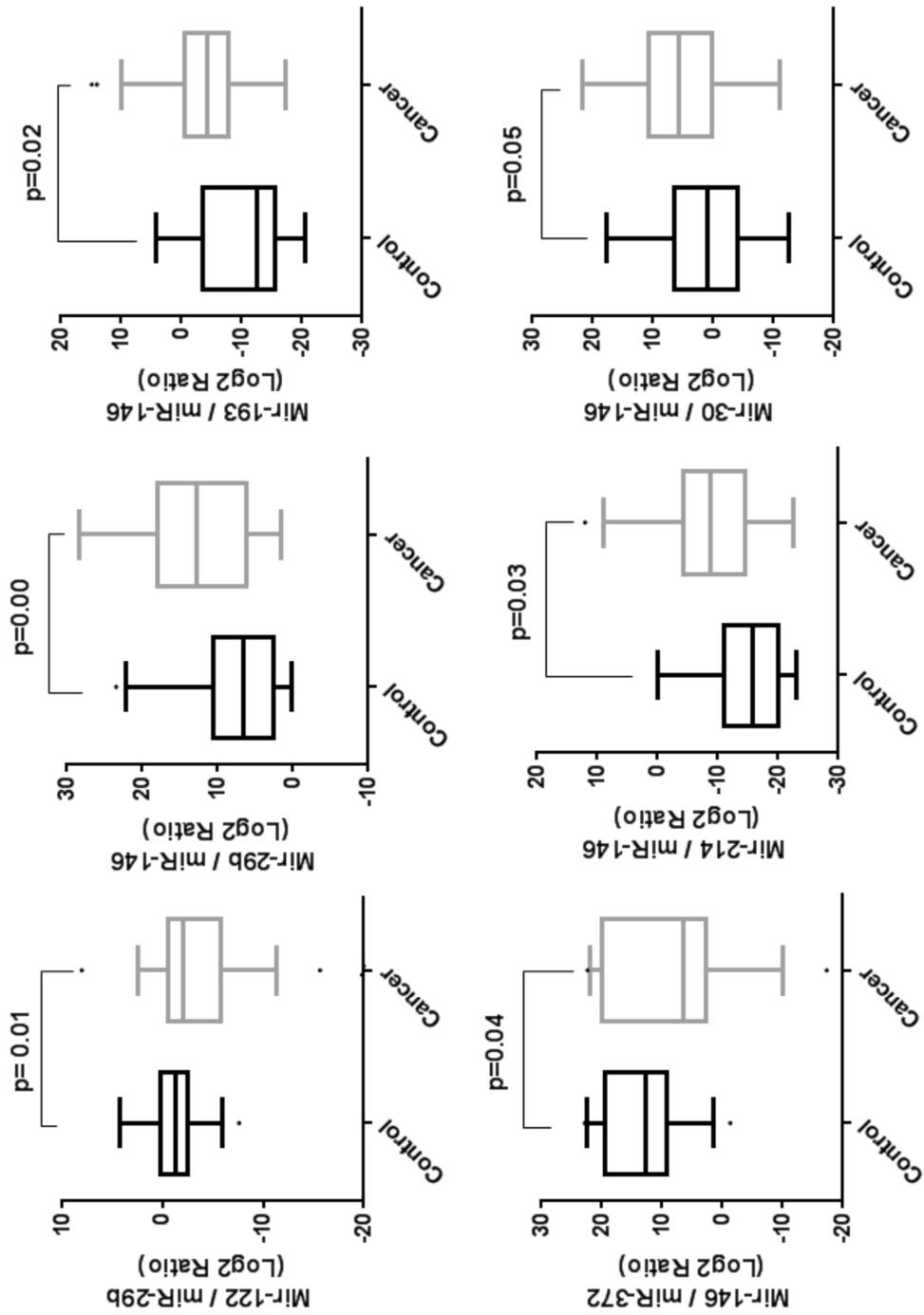
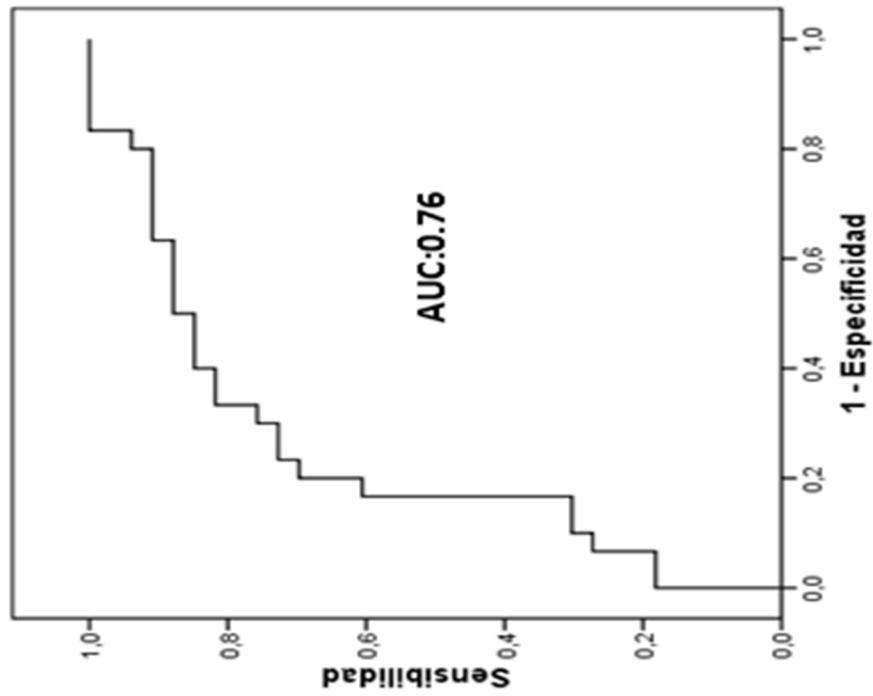
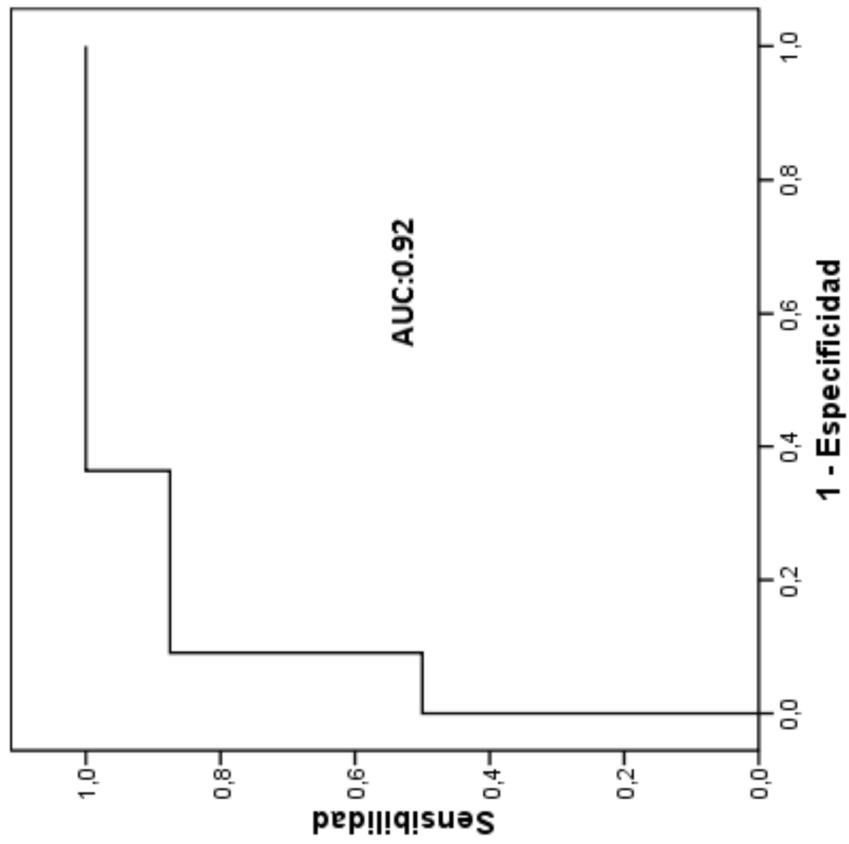


Fig.2



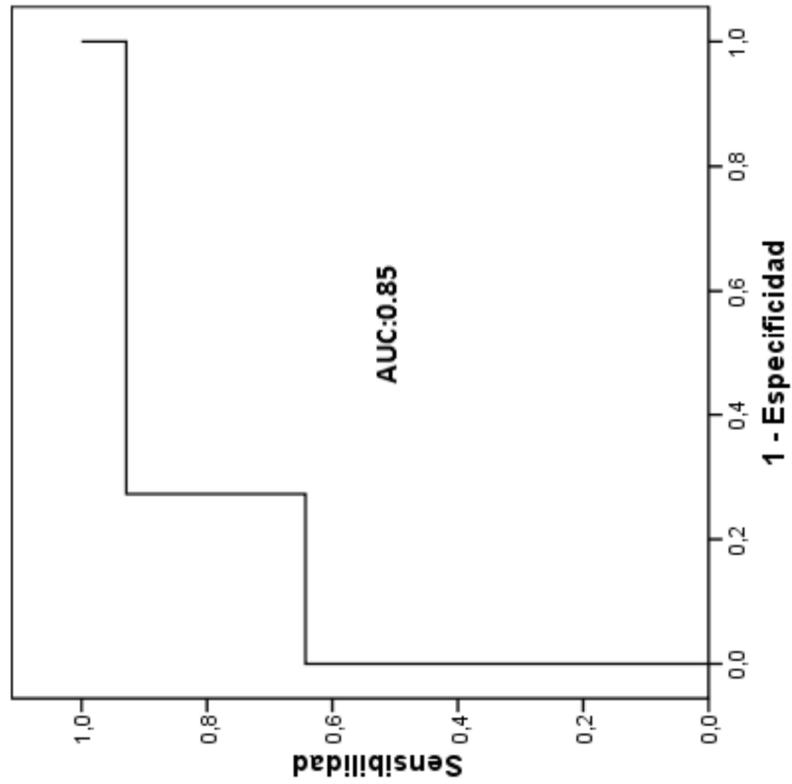
ROC Análisis	
Área bajo la curva (AUC)	0.76
Sensibilidad (%)	70
Especificidad (%)	80
Punto de corte	0.579

Fig.3



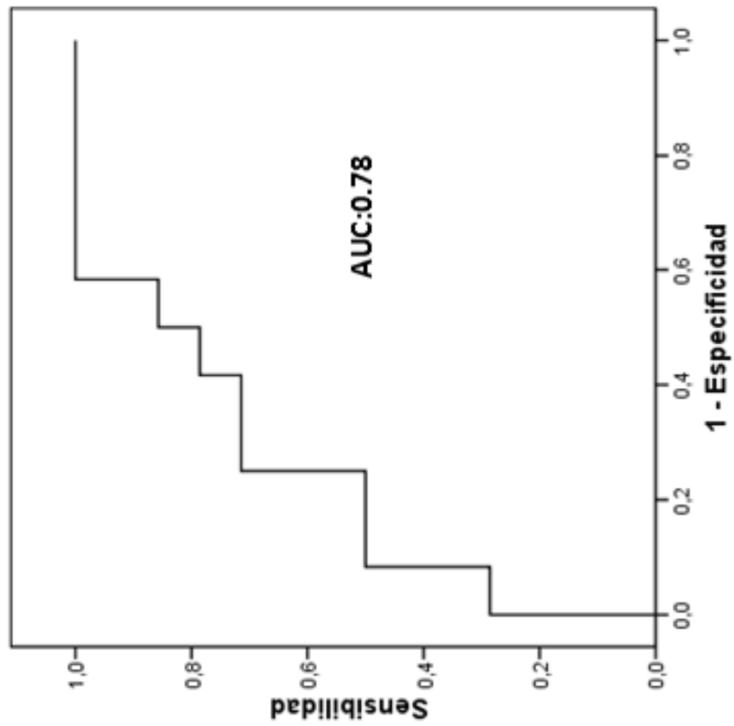
ROC Análisis	
Área bajo la curva (AUC)	0.92
Sensibilidad (%)	87
Especificidad (%)	91
Punto de corte	0.50

Fig.4



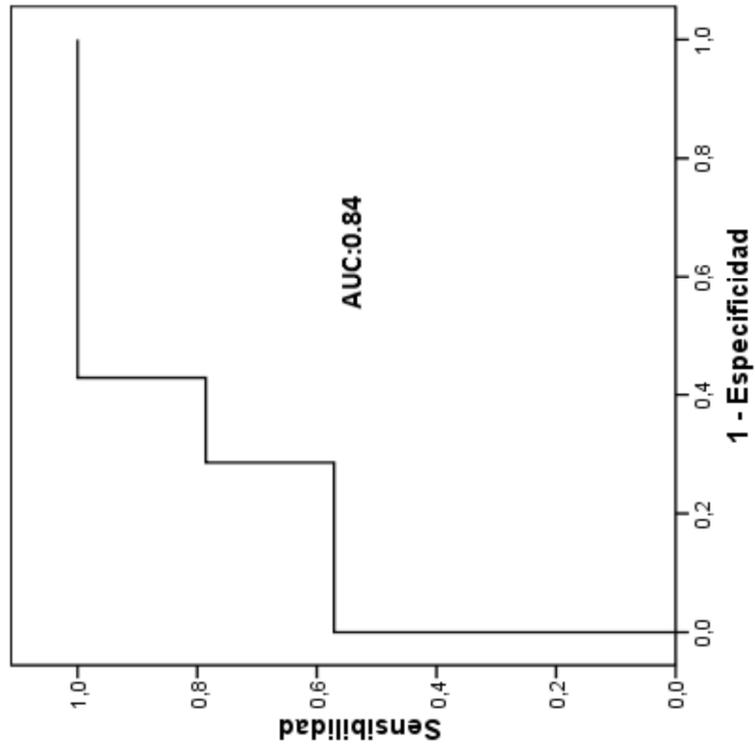
ROC Análisis	
Área bajo la curva (AUC)	0.85
Sensibilidad (%)	64
Especificidad (%)	91
Punto de corte	0.65

Fig.5



ROC Análisis	
Área bajo la curva (AUC)	0.78
Sensibilidad (%)	71
Especificidad (%)	75
Punto de corte	0.45

Fig.6



ROC Análisis	
Área bajo la curva (AUC)	0.84
Sensibilidad (%)	78
Especificidad (%)	72
Punto de corte	0.64

Fig.7

# ES 2 646 826 A1

## Listado de Secuencias

<110> SERVICIO ANDALUZ DE SALUD  
UNIVERSIDAD DE CÓRDOBA

<120> MicroARNs como biomarcadores para el diagnóstico del cáncer de pulmón

<130> FIBICO-15011

<160> 23

<170> PatentIn version 3.5

<210> 1

<211> 99

<212> DNA

<213> Homo sapiens

<400> 1

ccgatgtgta tcctcagctt tgagaactga attccatggg ttgtgtcagt gtcagacctc  
60

tgaaattcag ttcttcagct gggatatctc tgatcatcgt  
99

<210> 2

<211> 85

<212> DNA

<213> Homo sapiens

<400> 2

ccttagcaga gctgtggagt gtgacaatgg tgtttgtgtc taaactatca aacgccatta  
60

tcacactaaa tagctactgc taggc  
85

<210> 3

<211> 68

<212> DNA

<213> Homo sapiens

<400> 3

gaggcaaagt tctgagacac tccgactctg agtatgatag aagtcagtg c actacagaac  
60

tttgtctc

68

<210> 4

<211> 110

<212> DNA

<213> Homo sapiens

# ES 2 646 826 A1

<400> 4  
ggcctggctg gacagagttg tcatgtgtct gcctgtctac acttgctgtg cagaacatcc  
60

gctcacctgt acagcaggca cagacaggca gtcacatgac aaccagcct  
110

<210> 5  
<211> 67  
<212> DNA  
<213> Homo sapiens

<400> 5  
gtgggcctca aatgtggagc actattctga tgtccaagtg gaaagtgcctg cgacatttga  
60

gcgtcac  
67

<210> 6  
<211> 84  
<212> DNA  
<213> Homo sapiens

<400> 6  
gcacccgggt tgaggtagta gtttgtatgg tttagagtta caccctggga gtaactgta  
60

caaccttcta gctttccttg gagc  
84

<210> 7  
<211> 89  
<212> DNA  
<213> Homo sapiens

<400> 7  
accatgctgt agtgtgtgta aacatcctac actctcagct gtgagctcaa ggtggctggg  
60

agagggttgt ttactccttc tgccatgga  
89

<210> 8  
<211> 72  
<212> DNA  
<213> Homo sapiens

<400> 8  
agatactgta aacatcctac actctcagct gtggaaagta agaaagctgg gagaaggctg  
60

tttactcttt ct  
72

# ES 2 646 826 A1

<210> 9  
<211> 82  
<212> DNA  
<213> Homo sapiens

<400> 9  
gcagtcctct gttagttttg catagttgca ctacaagaag aatgtagttg tgcaaatcta  
60

tgcaaaactg atggtggcct gc  
82

<210> 10  
<211> 83  
<212> DNA  
<213> Homo sapiens

<400> 10  
gtggtctcag aatcgggggt ttgagggcga gatgagtta tgttttatcc aactggcct  
60

caaagtcccg cttttggggg cat  
83

<210> 11  
<211> 81  
<212> DNA  
<213> Homo sapiens

<400> 11  
cttcaggaag ctggtttcat atggtgggtt agatttaa atgtgattgtc tagcaccatt  
60

tgaaatcagt gttcttgggg g  
81

<210> 12  
<211> 84  
<212> DNA  
<213> Homo sapiens

<400> 12  
gtcagaataa tgtcaaagtg cttacagtgc aggtagtgat atgtgcatct actgcagtga  
60

aggcacttgt agcattatgg tgac  
84

<210> 13  
<211> 22  
<212> RNA  
<213> Homo sapiens

# ES 2 646 826 A1

<400> 13  
ugagaacuga auuccauggg uu  
22

<210> 14  
<211> 22  
<212> RNA  
<213> Homo sapiens

<400> 14  
uggaguguga caaugguguu ug  
22

<210> 15  
<211> 22  
<212> RNA  
<213> Homo sapiens

<400> 15  
ucagugcacu acagaacuuu gu  
22

<210> 16  
<211> 22  
<212> RNA  
<213> Homo sapiens

<400> 16  
acagcaggca cagacaggca gu  
22

<210> 17  
<211> 23  
<212> RNA  
<213> Homo sapiens

<400> 17  
ccucaaauuu ggagcacuau ucu  
23

<210> 18  
<211> 22  
<212> RNA  
<213> Homo sapiens

<400> 18  
ugagguagua gguuguaugg uu  
22

<210> 19  
<211> 23

# ES 2 646 826 A1

<212> RNA  
<213> Homo sapiens

<400> 19  
uguaaacauc cuacacucuc agc  
23

<210> 20  
<211> 22  
<212> RNA  
<213> Homo sapiens

<400> 20  
aguuuugcau aguugcacua ca  
22

<210> 21  
<211> 22  
<212> RNA  
<213> Homo sapiens

<400> 21  
cgggguuuug agggcgagau ga  
22

<210> 22  
<211> 23  
<212> RNA  
<213> Homo sapiens

<400> 22  
uagcaccuu ugaaaucagu guu  
23

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

<400> 23  
caaagugcuu acagugcagg uag  
23