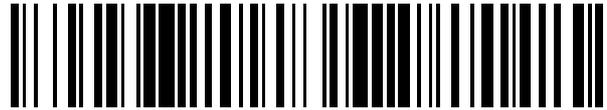


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 608 675**

21 Número de solicitud: 201531301

51 Int. Cl.:

G01N 33/50 (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

11.09.2015

43 Fecha de publicación de la solicitud:

12.04.2017

56 Se remite a la solicitud internacional:

PCT/ES2016/070640

71 Solicitantes:

SERVICIO ANDALUZ DE SALUD (50.0%)

Avda. de la Constitucion, 18

41071 Sevilla ES y

UNIVERSIDAD DE CÓRDOBA (50.0%)

72 Inventor/es:

JURADO GÁMEZ, Bernabé;

LUQUE DE CASTRO, María Dolores ;

PRIEGO CAPOTE , Feliciano;

CALDERÓN SANTIAGO , Mónica y

PERALBO MOLINA , Ángela

74 Agente/Representante:

FÚSTER OLAGUIBEL, Gustavo Nicolás

54 Título: **Biomarcadores en aire exhalado para el diagnóstico, clasificación y seguimiento del cáncer de pulmón**

57 Resumen:

Biomarcadores en aire exhalado para el diagnóstico, clasificación y seguimiento del cáncer de pulmón.

Método de obtención de datos útiles para el diagnóstico, clasificación y seguimiento del cáncer de pulmón, kit o dispositivo y uso.

ES 2 608 675 A1

DESCRIPCIÓN

Biomarcadores en aire exhalado para el diagnóstico, clasificación y seguimiento del cáncer de pulmón

CAMPO DE LA INVENCION

5 La presente invención se encuentra dentro del campo de la biomedicina y la biotecnología, y específicamente se refiere a un método de obtención de datos útiles para el diagnóstico, clasificación y seguimiento del cáncer, preferiblemente cáncer de pulmón, mediante el análisis de los biomarcadores: *p*-cresol, alcohol cumílico, eicosenamida, hexadecilindano, 2,4-bis-dimetilbencil-6-*t*-butilfenol, monostearina, espiro-2,4-heptano-1,5-6-metileno, 13-
10 heptadecin-1-ol, metil estearato, glicerol-1-palmitato, alcohol bencílico, 2,4-difenil-4-metil-2-E-pentano en una muestra biológica aislada, preferiblemente condensado de aire exhalado.

ANTECEDENTES DE LA INVENCION

La enfermedad respiratoria crónica es un término genérico que abarca condiciones patológicas a largo plazo que afectan a las vías respiratorias y otras estructuras pulmonares
15 implicadas en la respiración. Entre estas enfermedades crónicas se pueden destacar dos patologías: enfermedad pulmonar obstructiva crónica (EPOC) y el cáncer de pulmón, este último perteneciente a la lista de las principales causas de muerte relacionada con la enfermedad en los países más desarrollados [1,2]. La alta tasa de mortalidad a causa de las enfermedades respiratorias más graves está asociada a la baja eficiencia de los métodos de
20 cribado para su detección en etapas iniciales [3]. La búsqueda de pruebas de diagnóstico para la detección precoz de las enfermedades respiratorias crónicas se ha intensificado en las últimas décadas. La tomografía computarizada (TC) a dosis baja es en la actualidad la prueba más utilizada para la detección del cáncer de pulmón, ya que reduce la mortalidad de la población de alto riesgo alrededor del 20% en comparación con la radiografía de tórax,
25 según el Estudio Nacional de Exámenes de pulmón [4]. La TC también ha demostrado ser útil para la evaluación de la EPOC [5].

Además de las pruebas basadas en técnicas de imagen, diferentes estudios basados en las disciplinas "ómicas" se han centrado en el desarrollo de herramientas de evaluación para diagnosticar enfermedades respiratorias, buscando principalmente potenciales
30 biomarcadores ya sea en tejidos o en biofluidos. Los métodos que utilizan muestras de biofluidos obtenidos de manera no invasiva están ganando popularidad frente a los muestreos invasivos. Ese es el caso de la orina o el esputo y, más recientemente, del aliento exhalado o el sudor [6]. Un estudio reciente sobre el sudor que discrimina patrones de

metabolitos para el cribado del cáncer de pulmón [7], ha dado lugar a un panel óptimo de 5 compuestos que proporcionan el 80% de especificidad y 79% de sensibilidad y conducen a tasas de falsas positivas y negativas en torno al 20% [8]. En cuanto a la respiración exhalada, los niveles de nitrito en el líquido epitelial han demostrado su capacidad como biomarcador ya que un aumento de este anión está directamente asociado con el cáncer [9]. Se han investigado citoquinas individuales también en el aliento exhalado de pacientes con cáncer de pulmón utilizando inmunoensayos enzimáticos (EIA), pero faltan pruebas disponibles comercialmente con la sensibilidad exigida para la detección de estos compuestos [10–13]. Otros biomarcadores (por ejemplo, peróxido de hidrógeno [14], 8-isoprostano [14] o el valor del pH [15]) permite la discriminación entre los pacientes con EPOC y controles sanos; sin embargo, no son capaces de discriminar entre estos dos grupos de individuos y pacientes con cáncer de pulmón [16]. Phillips *et al.*, centrándose en vapor de aire exhalado, han desarrollado un modelo matemático constituido por 22 compuestos (principalmente alcanos y sus derivados y los derivados del benceno) como biomarcadores de cáncer de pulmón primario. El modelo reportó valores de sensibilidad y selectividad de 71,7% y 66,7%, respectivamente, incluso para las etapas avanzadas de la enfermedad [17]. Este estudio se mejoró posteriormente mediante la inclusión en el modelo de un conjunto de productos del estrés oxidativo excretados en la respiración [18], y se logró una sensibilidad y especificidad del 85,1% y 80,5%, respectivamente. Más adelante, otro estudio concluyó que una prueba de aliento de dos minutos puede predecir el cáncer pulmonar independientemente de la histología, estadio de la enfermedad o el consumo de tabaco, con el 84,6% de sensibilidad y 80% de especificidad [19]. Un estudio reciente ha proporcionado un panel de 23 compuestos orgánicos volátiles en el aliento exhalado con capacidad para distinguir entre pacientes y controles de cáncer de pulmón con 96,5% de sensibilidad y 97,5% de especificidad [20]. Sin embargo, los criterios para la clasificación de los compuestos endógenos no fue restrictiva y algunos de los compuestos incluidos en el panel también se detectaron en el aire de la sala de toma de muestras.

Además del análisis directo del aire exhalado utilizando la interfaz adecuada, la muestra a analizar puede ser el aire exhalado condensado (EBC); es decir el gas exhalado que condensa como una solución líquida por enfriamiento, cuyo análisis puede permitir conocer de forma aproximada la composición del líquido extracelular y de gases exhalados solubles. A pesar de que el componente principal del EBC es vapor de agua, posee cientos de componentes diferentes, se pueden encontrar desde pequeños iones inorgánicos pasando por grandes moléculas orgánicas hasta péptidos, proteínas, tensioactivos y macromoléculas [23] en concentraciones traza [24–27].

En cuanto el protocolo de muestreo para el EBC, existen dispositivos comerciales que permiten recolección de dos fracciones del EBC separando el aire procedente de las vías respiratorias superiores y centrales (aliento superior, UA) y de la vía aérea distal (aliento inferior, DA). A pesar de que se cree que el UA contiene compuestos sin relevancia clínica [23], no se han establecido claramente diferencias entre ambos tipos de muestras. El fraccionamiento del EBC permite la comparación de los perfiles de composición de la UA y DA para encontrar diferencias metabólicas en pacientes con cáncer de pulmón en comparación con un grupo de factor de riesgo compuesto por pacientes con EPOC y con los fumadores activos. Se utilizó un tercer grupo como grupo de control, que estaba formado por individuos sanos no fumadores. Los perfiles de composición de UA y DA para los tres grupos se obtuvieron usando cromatografía de gases acoplada a espectrometría de masas en modo de alta resolución (GC-TOF/MS).

BREVE DESCRIPCIÓN DE LA INVENCION

Un **primer aspecto** de la invención se refiere al uso de los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida, el hexadecilindano, el 2,4-dis-dimetilbencil-6-t-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol, el metil estearato, el glicerol-1-palmitato, el alcohol bencílico y el 2,4-difenil-4-metil-2-E-pentano para el diagnóstico, clasificación y seguimiento del cáncer.

En una realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

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

a) cuantificar los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida y el hexadecilindano en una muestra biológica aislada de dicho individuo.

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

b) comparar las cantidades obtenidas en el paso (a) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos sanos.

En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

En otra realización preferida de este aspecto de la invención, la muestra biológica aislada del paso (a) es aire exhalado, preferiblemente condensado de aire exhalado.

5 En otra realización preferida de este aspecto de la invención, la cuantificación de los metabolitos del paso (a) se realiza por análisis GC–TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

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

10 c) incluir al individuo del paso a) en el grupo de individuos que presentan cáncer, cuando se detectan niveles de al menos 3 metabolitos, preferiblemente de los 4 metabolitos del paso (a), en cantidades significativamente distintas a la cantidad de referencia.

En una realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

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

a) cuantificar los metabolitos 2,4-bis-dimetilbencil-6-t-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol y el metil estearato en una muestra biológica aislada de dicho individuo.

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

b) comparar las cantidades obtenidas en el paso (a) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos que presentan al menos un factor de riesgo de padecer cáncer, preferiblemente tabaquismo o con EPOC.

25 En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

En otra realización preferida de este aspecto de la invención, la muestra biológica aislada del paso (a) es aire exhalado, preferiblemente condensado de aire exhalado.

En otra realización preferida de este aspecto de la invención, la cuantificación de los metabolitos del paso (a) se realiza por análisis GC–TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

5 Un **quinto aspecto** de la invención se refiere a un método de diagnóstico, clasificación y seguimiento del cáncer, de ahora en adelante cuarto método de la invención, que comprende los pasos (a)–(b) del tercer método de la invención, y además comprende:

10 c) incluir al individuo del paso a) en el grupo de individuos que presentan cáncer, cuando se detectan niveles de al menos 3 metabolitos, preferiblemente de 4 metabolitos, y más preferiblemente de 5 metabolitos del paso (a), en cantidades significativamente distintas cantidad de referencia.

En una realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

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

15 a) cuantificar los metabolitos glicerol-1-palmitato, alcohol bencílico, la monostearina, el 2,4-difenil-4-metil-2-E-pentano y el *p*-cresol en una muestra biológica aislada de dicho individuo.

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

20 b) comparar las cantidades obtenidas en el paso (a) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos sanos.

En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

25 En otra realización preferida de este aspecto de la invención, la muestra biológica aislada del paso (a) es aire exhalado, preferiblemente condensado de aire exhalado.

En otra realización preferida de este aspecto de la invención, la cuantificación de los metabolitos del paso (a) se realiza por análisis GC–TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

Un **séptimo aspecto** de la invención se refiere a un método de diagnóstico, clasificación y seguimiento del cáncer, de ahora en adelante sexto método de la invención, que comprende los pasos (a)–(b) del quinto método de la invención, y además comprende:

5 c) incluir al individuo del paso a) en el grupo de individuos de alto riesgo de padecer cáncer, cuando se detectan niveles de al menos 2 metabolitos, preferiblemente de 3 metabolitos, más preferiblemente de 4 metabolitos, y aún más preferiblemente de 5 metabolitos del paso (a), en cantidades significativamente distintas cantidad de referencia.

Un **octavo aspecto** de la invención se refiere a un kit o dispositivo que comprende los elementos necesarios para cuantificar los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida, el hexadecilindano, el 2,4-bis-dimetilbenzil-6-*t*-butilfenol, la monostearina, el 10 espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol, el metil estearato, el glicerol-1-palmitato, el benzil alcohol y/o el 2,4-difenil-4-metil-2-*E*-pentano, según se describe en cualquiera de los métodos de la invención.

Un **noveno aspecto** de la invención se refiere al uso del kit o dispositivo según la 15 reivindicación anterior, para el diagnóstico, clasificación y seguimiento del cáncer.

En una realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

BREVE DESCRIPCIÓN DE LAS FIGURAS

Figura 1. Diagrama de barras que compara la composición de muestras de EBC recogidas 20 de las vías aéreas superiores y centrales (UA) y de la vía aérea distal (DA).

Figura 2. Análisis discriminante por mínimos cuadrados parciales (PLS-DA) realizado a partir del conjunto de datos obtenidos tras el análisis de extractos de aire expirado condensado de vías respiratorias superiores y centrales comparando pacientes (A) con 25 cáncer de pulmón (LC) y las personas con factor de riesgo (FR), (B) pacientes con cáncer de pulmón e individuos control (control), y (C) los individuos con factor de riesgo e individuos control (control).

Figura 3. Análisis discriminante por mínimos cuadrados parciales (PLS-DA) realizado a partir del conjunto de datos obtenidos tras el análisis de extractos de aire expirado condensado de la vía aérea distal comparando pacientes (A) con cáncer de pulmón (LC) y 30 las personas con factor de riesgo (FR), (B) pacientes con cáncer de pulmón e individuos control (control), y (C) los individuos con factor de riesgo e individuos control (control).

Figura 4. Diagramas de caja y bigotes que presentan la variabilidad de (A) cuatro metabolitos —monopalmitin, n-hexadecilindano, monoestearina y escualeno— en pacientes con cáncer de pulmón (LC) en comparación con los individuos con factor de riesgo (FR) y la cohorte de individuos control, y (B) cuatro metabolitos —11-eicosenamida, *p*-cresol, indol, ácido benzoico 4-etoxi etil éster en los individuos de control en comparación con los pacientes de cáncer de pulmón y los individuos con factor de riesgo.

Figura 5. Curva de calibración del índice de retención proporcionada por la mezcla estándar de alcanos aplicando el método cromatográfico utilizado.

Figura 6. Análisis de Componentes Principales (PCA) construido a partir de los datos obtenidos con los extractos de aire expirado condensado de las vías aéreas superiores y centrales y de la vía aérea distal de los pacientes con cáncer de pulmón (LC), personas con factor de riesgo (FR) y los individuos control.

Figura 7. Análisis de Componentes Principales (PCA) construido a partir del conjunto de datos obtenidos mediante el análisis de extractos de aire expirado condensado de las vías aéreas superiores y centrales comparando pacientes (A) de cáncer de pulmón (LC) y las personas con factor de riesgo (FR), (B) de los pacientes con cáncer de pulmón e individuos de control (Control), e individuos de control (C) y los individuos con factor de riesgo.

Figura 8. Análisis de Componentes Principales (PCA) construido a partir del conjunto de datos obtenidos tras el análisis de extractos de aire expirado condensado de la vía aérea distal comparando pacientes (A) de cáncer de pulmón (LC) y las personas con factor de riesgo (FR), (B) de los pacientes con cáncer de pulmón y los individuos de control (control), y de control (C) y los individuos con factor de riesgo particulares.

DESCRIPCIÓN DETALLADA DE LA INVENCION

Los autores de la presente invención han identificado y calculado la concentración relativa de los distintos metabolitos de condensado de aire exhalado en individuos que padecen cáncer de pulmón, en individuos que presentan algún factor de riesgo de padecerlo, y en individuos sanos. Han encontrado una serie de marcadores para el diagnóstico de los individuos con cáncer de pulmón, diferenciando los sujetos con cáncer de pulmón de aquéllos que no lo padecen. Así pues, la presente invención proporciona un método de obtención de datos útiles para el diagnóstico, clasificación y seguimiento de individuos con cáncer de pulmón.

Por tanto, un **primer aspecto** de la invención se refiere al uso de los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida, el hexadecilindano, el 2,4-bis-dimetilbencil-6-*t*-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol, el metil estearato, el glicerol-1-palmitato, el alcohol bencílico y el 2,4-difenil-4-metil-2-*E*-pentano para el diagnóstico, clasificación y seguimiento del cáncer.

En una realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

Aunque se puede usar cualquier metabolito, o cualquiera de sus combinaciones, más preferiblemente los metabolitos se usan agrupados de la siguiente forma

Grupo I) *p*-cresol, alcohol cumílico, eicosenamida y/o hexadecilindano,

Grupo II) 2,4-bis-dimetilbencil-6-*t*-butilfenol, monostearina, espiro-2,4-heptano-1,5-6-metileno, 13-heptadecin-1-ol y/o metil estearato, y

Grupo III) glicerol-1-palmitato, alcohol bencílico, monostearina, 2,4-difenil-4-metil-2-*E*-pentano y/o *p*-cresol,

usando al menos un grupo.

En una realización preferida de este aspecto de la invención el cáncer es el cáncer de pulmón.

A continuación se describen las características de los metabolitos que son objeto del estudio:

p-Cresol: Compuesto fenólico de bajo peso molecular conocido, entre otras cosas, por ser un producto de degradación del tolueno y que, por tanto, puede relacionarse directamente con el consumo de tabaco. Además, es un metabolito del aminoácido tirosina y en cierta medida de la fenilalanina, por la conversión en ácido 4-hidroxifenilacético provocada por las bacterias intestinales antes de ser descarboxilado a *p*-cresol (putrefacción).

Alcohol cumílico: Monoterpeno propuesto como marcador para la detección de cáncer de pulmón en otros estudios que utilizan de aire exhalado (vapor) como biofluido.

Eicosenamida: Derivado ácido graso conocido por su potencial como agente antimicrobiano.

Hexadecilindano: Derivado de los indanos propuesto como marcador para la detección de cáncer de pulmón en otros estudios que utilizan aire exhalado (vapor) como biofluido.

2,4-Bis-dimetilbencil-6-t-butilfenol: Compuesto derivado del benceno relacionado con el consumo de tabaco.

5 Monostearina: Pertenece a la familia de monoacilgliceroles, glicéridos que consisten en una cadena de ácido graso unido covalentemente a una molécula de glicerol a través de un enlace éster.

Espiro-2,4-heptano-1,5-6-metileno: No existe información sobre su existencia en aire exhalado o en cualquier otro biofluido.

13-Heptadecin-1-ol: No existe información sobre su existencia en aire exhalado o en cualquier otro biofluido.

10 Estearato de metilo: Derivado de ácido graso previamente identificado en heces.

Glicerol-1-palmitato: Pertenece a la familia de monoacilgliceroles, glicéridos que consisten en una cadena de ácido graso unido covalentemente a una molécula de glicerol a través de un enlace éster.

15 Alcohol bencílico: Derivado del benceno conocido por ser producto de la degradación del tolueno y que, por tanto, puede relacionarse directamente con el consumo de tabaco.

2,4-Difenil-4-metil-2-E-pentano: Compuesto derivado del benceno relacionado con el consumo de tabaco.

20 Un **segundo aspecto** de la invención se refiere a un método de obtención de datos útiles para el diagnóstico, clasificación y seguimiento del cáncer, de ahora en adelante primer método de la invención, que comprende:

a) obtener una muestra biológica aislada de un individuo,

b) cuantificar los metabolitos p-cresol, el alcohol cumílico, la eicosenamida y el hexadecilindano en la muestra del paso a).

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

c) comparar las cantidades obtenidas en el paso (b) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos sanos.

En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

En otra realización preferida de este aspecto de la invención, la muestra biológica aislada del paso (a) es aire exhalado, preferiblemente condensado de aire exhalado.

- 5 En otra realización preferida de este aspecto de la invención, la cuantificación de los metabolitos del paso (b) se realiza por análisis mediante GC–TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

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

d) incluir al individuo del paso a) en el grupo de individuos que presentan cáncer, cuando se detectan niveles de al menos 3 metabolitos, preferiblemente de los 4 metabolitos del paso (b), en cantidades significativamente distintas a la cantidad de referencia.

En una realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

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

a) obtener una muestra biológica aislada de un individuo,

20 b) cuantificar los metabolitos 2,4-bis-dimetilbenzil-6-t-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol y el estearato de metilo en la muestra del paso a).

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

25 c) comparar las cantidades obtenidas en el paso (a) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos que presentan al menos un factor de riesgo de padecer cáncer, preferiblemente tabaquismo o con EPOC.

En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

En otra realización preferida de este aspecto de la invención, la muestra biológica aislada del paso (a) es aire exhalado, preferiblemente condensado de aire exhalado.

En otra realización preferida de este aspecto de la invención, la cuantificación de los metabolitos del paso (b) se realiza por análisis mediante GC–TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

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

d) incluir al individuo del paso a) en el grupo de individuos que presentan cáncer, cuando se detectan niveles de al menos 3 metabolitos, preferiblemente de 4 metabolitos, y más preferiblemente de 5 metabolitos del paso (b), en cantidades significativamente distintas a la cantidad de referencia.

En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

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

a) obtener una muestra biológica aislada un individuo,

b) cuantificar los metabolitos glicerol-1-palmitato, alcohol bencílico, la monostearina, el 2,4-difenil-4-metil-2-E-pentano y el *p*-cresol en la muestra del paso a).

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

c) comparar las cantidades obtenidas en el paso (b) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos sanos.

En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

En otra realización preferida de este aspecto de la invención, la muestra biológica aislada del paso (a) es aire exhalado, preferiblemente condensado de aire exhalado.

En otra realización preferida de este aspecto de la invención, la cuantificación de los metabolitos del paso (b) se realiza por análisis mediante GC-TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

5 Un **séptimo aspecto** de la invención se refiere a un método de diagnóstico, clasificación y seguimiento del cáncer, de ahora en adelante sexto método de la invención, que comprende los pasos (a)-(c) del quinto método de la invención, y además comprende:

10 c) incluir al individuo del paso a) en el grupo de individuos de alto riesgo de padecer cáncer, cuando se detectan niveles de al menos 2 metabolitos, preferiblemente de 3 metabolitos, más preferiblemente de 4 metabolitos, y aún más preferiblemente de 5 metabolitos del paso (b), en cantidades significativamente distintas cantidad de referencia.

En otra realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

Una "muestra biológica", como se define aquí, es una pequeña parte de un sujeto, representativa del conjunto y puede estar constituido por una biopsia o una muestra de fluido corporal. Las biopsias son pequeñas piezas de tejido y pueden ser frescas, congeladas o fijadas, como fijada con formalina y embebidas en parafina (*formalin-fixed and paraffin embedded* FFPE). Muestras de fluidos corporales puede ser sangre, plasma, suero, orina, esputo, líquido cefalorraquídeo, leche o muestras de fluido ductal y pueden asimismo ser frescos, congelados o fijadas. Las muestras se pueden extirpar quirúrgicamente, mediante extracción es decir, por agujas hipodérmicas o de otro tipo, por microdissección o captura láser. La muestra debe contener cualquier material biológico adecuado para detectar el biomarcador o biomarcadores deseado/s, por lo tanto, dicha muestra podría comprender material de las células del sujeto. La/s muestra/s usada/s para desarrollar los métodos de la invención son una muestra gaseosa o de lavado broncoalveolar. Preferiblemente la muestra es aire exhalado, y más preferiblemente condensado de aire exhalado.

Una "muestra de referencia", como se usa aquí, significa una muestra obtenida de los individuos, preferiblemente dos o más individuos, de los que se sabe que están libres de la enfermedad (cáncer, preferiblemente de cáncer de pulmón) o, alternativamente, de la población general. Los niveles adecuados de metabolitos se pueden determinar mediante la medición de los niveles de dichos metabolitos en varios individuos adecuados, y tales niveles de referencia se pueden ajustar para poblaciones de individuos o sujetos específicos. En una realización preferida, la muestra de referencia se obtiene de un grupo

de individuos o sujetos sanos o de sujetos sin historia previa de padecer cáncer de pulmón. La cantidad y/o concentración de los metabolitos en la muestra de referencia puede, preferiblemente, generarse a partir de una población de dos o más individuos; por ejemplo, la población puede comprender 3, 4, 5, 10, 15, 20, 30, 40, 50 o más individuos o sujetos. En
5 otra realización preferida la muestra de referencia se obtiene de un grupo de individuos de riesgo, preferiblemente fumadores y/o pacientes de EPOC.

Un "individuo" o "sujeto", como se usa aquí, se refiere a un mamífero, humano o no humano, en observación, y más preferiblemente un ser humano. El individuo puede ser cualquiera, un
10 individuo predispuesto a una enfermedad (por ejemplo, cáncer de pulmón) o un individuo que padece dicha enfermedad.

El término "cantidad", tal y como se utiliza en la descripción, se refiere pero no se limita, a la cantidad absoluta o relativa de los metabolitos, de su concentración en el exhalado de aire, preferiblemente en el condensado de aire exhalado, así como a cualquier otro valor o
15 parámetro relacionado con los mismos o que pueda derivarse de éstos. La cantidad de los metabolitos puede medirse de manera directa o indirecta.

El término "comparación", tal y como se utiliza en la descripción, se refiere pero no se limita, a la comparación de la cantidad y/o concentración de los metabolitos de la muestra biológica a analizar, también llamada muestra biológica problema, con una cantidad y/o concentración de los metabolitos de una o varias muestras de referencia deseable. La muestra de
20 referencia puede ser analizada, por ejemplo, simultánea o consecutivamente, junto con la muestra biológica problema. La comparación descrita en el apartado (c) del método de la presente invención puede ser realizada manualmente o asistida por ordenador.

Las cantidades de referencia adecuadas pueden ser determinadas por el método de la presente invención a partir de una muestra de referencia que puede ser analizada, por
25 ejemplo, simultánea o consecutivamente, junto con la muestra biológica problema. Así, por ejemplo pero sin limitarnos, la muestra de referencia pueden ser los controles negativos, esto es, las cantidades detectadas por los métodos de la invención en muestras de individuos que no padecen la enfermedad o en individuos con factor de riesgo de padecer la enfermedad (tabaquismo y EPOC).

30 Los pasos (b) y/o (c) de los métodos descritos anteriormente pueden ser total o parcialmente automatizados. Los metabolitos que se determinan en el paso (b) pueden determinarse individualmente, o se puede determinar cualquiera de sus combinaciones.

La determinación de los metabolitos se puede hacer por cualquier medio conocido por el experto en la materia. La cuantificación de los metabolitos del paso (b) de los métodos de la invención se realiza preferiblemente por análisis mediante GC–TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

- 5 En esta memoria el equipo analítico utilizado y designado por GC–TOF/MS, del inglés *Gas Chromatography-Time of Flight Mass Spectrometry*, ha permitido desarrollar un método adecuado para la detección de moléculas orgánicas con un pesos moleculares de hasta aproximadamente 1000 Da. El GC–Q/TOF ofrece alta sensibilidad y selectividad con el valor añadido de proporcionar información precisa y de alta resolución para la confirmación
- 10 estructural de los metabolitos.

La determinación de la cantidad y/o concentración metabolitos, se puede hacer, por ejemplo, aunque sin limitarnos, mediante un sistema indicador preparado sobre un soporte sólido (papel o sorbente sólido) en el que se han inmovilizado reactivos selectivos o específicos para los compuestos identificados como marcadores.

- 15 Un **octavo aspecto** de la invención se refiere a un kit o dispositivo que comprende los elementos necesarios para cuantificar los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida, el hexadecilindano, el 2,4-bis-dimetilbencil-6-*t*-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol, el estearato de metilo, el glicerol-1-palmitato, el alcohol bencílico y/o el 2,4-difenil-4-metil-2-*E*-pentano, según se describe en
- 20 cualquiera de los métodos de la invención.

En otra realización preferida de este aspecto de la invención, el kit o dispositivo de la invención además comprende todos aquellos elementos necesarios para llevar a cabo un análisis mediante GC–TOF/MS.

- El kit además puede incluir, sin ningún tipo de limitación, tampones, disoluciones para la
- 25 extracción de proteínas, agentes para prevenir la contaminación, inhibidores de la degradación de las proteínas, reactivos derivatizantes, etc.

- 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
- 30 necesarios para llevar a cabo el análisis cualitativo, semicuantitativo o cuantitativo en la muestra del paso a) o cualquiera de los métodos de la invención

Por otro lado, el kit 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 los métodos de la invención.

5 Un **noveno aspecto** de la invención se refiere al uso del kit o dispositivo según la reivindicación anterior, para el diagnóstico, clasificación y seguimiento del cáncer.

En una realización preferida de este aspecto de la invención, el cáncer es cáncer de pulmón.

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 desprenden 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.

EJEMPLO DE LA INVENCION

15 **Materiales y métodos**

Reactivos

Se usó hexano *TraceSELECT*® de Sigma-Aldrich (St. Louis, EE.UU.) como disolvente orgánico para la preparación de muestras, una mezcla estándar de alcanos (de C10 a C40) también de Sigma-Aldrich para las pruebas de separación de GC y para establecer la calibración del índice de retención (RI). También se utilizó agua desionizada (18 mΩ • cm) de un sistema de purificación de agua *Millipore Milli-Q*.

Instrumentos y dispositivos

Para la recogida de EBC se utilizó un dispositivo de ECOScreen2 (FILT Thorax-und LungenDiagnostik GmbH, Berlín, Alemania). La homogeneización de los extractos se llevó a cabo por un MS2 Minishaker Vortex (IKA, Alemania).

Se utilizó un sistema Agilent 7890A Series GC acoplado a un Agilent 7200 UHD Accurate-Mass QTOF espectrómetro de masas híbrido equipado con una fuente de impacto electrónico (EI) (Santa Clara, CA, EE.UU.). La muestra analítica se monitorizó en el modo de alta resolución.

30 *Cohorte seleccionada para el estudio*

Las muestras de EBC se obtuvieron de 239 individuos en ayunas, y se almacenaron a -80°C en el bio-repositorio del Hospital Reina Sofía (Córdoba, España). Todos los individuos se reclutaron en el Departamento de Medicina Respiratoria. La cohorte incluyó a 48 pacientes con diagnóstico de cáncer de pulmón entre noviembre de 2012 y mayo de 2014. Todos los
5 pacientes se diagnosticaron con cáncer de pulmón y se confirmaron citohistológicamente después de las pruebas clínicas basadas en la broncoscopia, biopsia con aguja fina o videotoracoscopia. Estos pacientes tenían una media de 63 ± 7 años y el 94% de ellos eran varones. El diagnóstico más frecuente fue el carcinoma de células escamosas (15
10 pacientes, 31,25%), seguido de adenocarcinoma (13 pacientes, 27%), carcinoma de células pequeñas (7 pacientes, 14,5%), carcinoma de células grandes (6 pacientes, 12,5%). Siete personas (14%) fueron diagnosticadas de cáncer de pulmón de células no pequeñas y sin clasificación histológica.

El grupo de factor de riesgo estaba formado por 130 personas, 83 de ellos fumadores activos (>20 paquetes/año) y 47 individuos diagnosticados con EPOC mediante espirometría
15 (FVC/FEV1 ratio $<0,7$). Dos razones apoyaron la inclusión de los fumadores y pacientes con EPOC en el grupo de factores de riesgo: en primer lugar, el tabaquismo es considerado el factor de riesgo más importante para desarrollar cáncer de pulmón y la EPOC; en segundo lugar, el aumento del estrés oxidativo o la presencia de células inflamatorias infiltradas en la EPOC y en el cáncer de pulmón son formas comunes de explicación teórica de daño
20 pulmonar (especialmente esta última). El grupo con factor de riesgo, control clínico al menos por un año, se caracterizó por una edad promedio de 61 ± 8 años, con el 82,4% de los individuos del sexo masculino. La existencia de cáncer de pulmón fue descartada en este grupo a través de la TC y la broncoscopia. El grupo de control formado por 61 individuos sanos con una media de 60 ± 9 años, 87% de ellos los individuos masculinos. Todos ellos
25 eran no fumadores activos ni pasivos, sin síntomas clínicos y con un perfil normal establecido por la radiografía de tórax.

Los criterios para la exclusión de pacientes fueron: a) la coexistencia de patología tumoral extrapulmonar o un tratamiento con fármacos citostáticos para una neoplasia diferente; b) diagnóstico de neoplasia en los últimos cinco años; c) la pérdida de peso injustificada (≥ 7
30 kg) durante el año pasado; d) grave trastorno de cualquier órgano con influencia negativa en el pronóstico o que impidió aplicar el protocolo (en los casos se trataba de incluir el grado IV de insuficiencia cardíaca según la Asociación del Corazón de Nueva York, la cirrosis hepática avanzada, la insuficiencia renal en fase V con el tratamiento sustitutivo con hemodiálisis o diálisis peritoneal, y la enfermedad pulmonar de diagnóstico no relacionado
35 con fumar, incluyendo neumopatía intersticial, neumonía, tuberculosis, etc. [8]).

Todos los experimentos se desarrollaron de acuerdo con los principios éticos de la investigación médica en humanos (Asociación Médica Mundial, Declaración de Helsinki). El comité de ética de la Reina del Hospital Sofía (Córdoba, España) aprobó y supervisó el estudio clínico.

5 *Procedimiento de recolección del EBC*

El dispositivo ECOScreen2 utilizado para el muestreo permite recoger directamente y condensar la EBC en bolsas de polietileno desechables. Se opera a $-20\text{ }^{\circ}\text{C}$, se realiza la recolección de EBC controlado en dos bolsas separadas para la separación física entre el aire contenido dentro de UA de la de DA [23]. La principal modificación fue la inserción de un filtro de protección comercial de Scharlab (Barcelona, España) sobre la válvula de entrada de aire para evitar la entrada de compuestos orgánicos exógenos y las partículas de la atmósfera ambiente. Este filtro se cambió periódicamente para evitar la saturación.

El muestreo de la respiración (con un clip de nariz) se realizó durante 15 minutos, tiempo necesario para recoger un volumen medio de EBC de 1,5 ml de la DA y 1 ml de la UA. Sólo 101 personas fueron capaces de proporcionar un volumen suficiente de las dos fracciones para el análisis. Las muestras se dividieron en alícuotas de 100 μl y los viales se almacenaron a $-80\text{ }^{\circ}\text{C}$ hasta su análisis. Se analizaron todas las muestras dentro de los 3 meses después de la recolección.

Preparación de la muestra

La preparación de la muestra consistió en extracción líquido-líquido utilizando hexano como extractante. En todos los casos, 100 μl de alícuotas de EBC y el mismo volumen de hexano se agitaron mediante vórtex en un inserto de vidrio a temperatura ambiente durante 1 min. Después, la fase orgánica se aisló y se puso en un nuevo inserto de vidrio para su análisis. Para eliminar las interferencias exógenas, los blancos se prepararon mediante el uso de agua tratada como las muestras.

Análisis GC-TOF/MS

Los análisis mediante GC-TOF/MS se realizaron con ionización por impacto de electrones (EI) a 70 eV y controlado por *MassHunter Acquisition B.06*. La separación cromatográfica se llevó a cabo en una columna capilar de sílice fundida DB-5MS-UI 30 m \times 0,25 mm i.d., de 0,25 μm de grosor. El programa de temperatura para la separación en el GC comenzó a 60 $^{\circ}\text{C}$ (1 min), seguido por un ascenso de temperatura de 10 $^{\circ}\text{C}/\text{min}$ hasta 300 $^{\circ}\text{C}$ finales (2 min). Se programó un aumento de temperatura de la columna hasta 310 $^{\circ}\text{C}$ durante 4 min

para asegurar la elución completa de todos los componentes de la muestra inyectada. Se llevaron a cabo inyecciones de 1 µl de muestra a 250 °C sin división del flujo, y se utilizó helio de grado ultrapuro como gas portador a un caudal de 1,0 ml/min. Las temperaturas de interfaz, fuente de iones y cuadrupolo se fijaron en 280, 300 y 200 °C, respectivamente. El filamento de la fuente de iones se apagó hasta el minuto 5.5 para evitar daños por la salida del frente del disolvente. El detector TOF realizó 5 espectros/s en el intervalo de m/z de 50 a 550 y la resolución fue 8500 (anchura a la mitad de la altura máxima del pico, FWHM) a m/z 501,9706. El PFTBA (perfluorotri-n-butilamina) de alta pureza para espectrometría de masas se utilizó para la calibración diaria del eje de masas. La identificación de los metabolitos se realizó mediante la búsqueda en los espectros de MS en la base de datos NIST 11 teniendo en cuenta los valores de RI.

Procesamientos de los datos y análisis estadísticos

Para procesar todos los datos obtenidos por GC–TOF/MS en el modo de análisis completo se utilizó el *Unknown Analysis* software (versión 7.0, Agilent Technologies, Santa Clara, CA, EE.UU.). El tratamiento de archivos de datos brutos se inició por deconvolución de las entidades moleculares (MFs) potenciales con el algoritmo adecuado incluido en el software. Para este propósito, el algoritmo de deconvolución para el parámetro altura absoluta consideró todos los iones superiores a 1.500 cuentas. Además, se establecieron el error de precisión y el tamaño de la ventana en de 50 ppm y 150, respectivamente. Después de la extracción de las MFs, se crearon para cada muestra los archivos de datos en formato de intercambio de archivos de compuestos (.cef) y se exportaron al software *Mass Profiler Professional* (MPP) (versión 12.1, Agilent Technologies, Santa Clara, CA, EE.UU.) para su posterior procesamiento.

En el siguiente paso, los datos fueron procesados por la alineación de las MFs potenciales de acuerdo con su tiempo de retención y el valor m/z utilizando una ventana de tolerancia de 0,3 min y un error en la precisión de 15 ppm. Las MFs a partir del análisis presentes en los blancos tratados como muestras se eliminaron del conjunto de datos de MFs definitivo para las muestras de EBC. El algoritmo de extracción confirmó la eficacia de esta etapa de filtrado. Esta corrección se aplicó a todos los tratamientos de las muestras analizadas. La reducción gradual del número MFs se basa en la frecuencia de ocurrencia mediante la comparación de repeticiones de un mismo grupo de personas. Se fijó un filtro por frecuencia de 100%, asegurando así la detección de cada MF en todas las repeticiones de cada grupo de muestras inyectadas (cáncer de pulmón, grupo con factor de riesgo y el control de los individuos sanos).

En el último paso, los MFs resultantes se exportaron (archivo .cef) para un análisis recursivo. Con este propósito, se utilizó el software *Quantitative Analysis* (versión 7.0, Agilent Technologies, Santa Clara, CA, EE.UU.) para reintegrar a todos los compuestos potenciales que se encuentran en las muestras analizadas. La tabla resultante se exporta en formato de valores separados por comas (archivo .csv) y se vuelve a procesar con el paquete de software *Mass Profiler Professional* (MPP). Se aplicó un filtro para eliminar las muestras con una réplica dentro de la variabilidad por encima de 10% para asegurar la eficacia del análisis recursivo. Por último, el conjunto de datos se normalizó por transformación logarítmica de la relación entre el área del pico de cada entidad molecular y la suma total de las áreas de las MFs presentes en el 100% de las muestras (MSTUS).

A continuación, los datos resultantes establecidos de cada fracción EBC se sometieron a análisis supervisado y no supervisado mediante análisis de componentes principales (PCA) y análisis discriminante por mínimos cuadrados parciales (PLS-DA). Como método de validación se seleccionó la validación cruzada mediante el uso de un modelo de N veces. Con este modelo, las clases en los datos de entrada se dividen aleatoriamente en N partes iguales; N-1 partes se utilizan para el entrenamiento y la parte restante se utiliza para la prueba. El proceso se repitió N veces, con un subconjunto diferente que se utiliza para realizar pruebas en un proceso iterativo. Por lo tanto, cada fila se utiliza al menos una vez en la formación y una vez en la prueba, y se genera una matriz de confusión. El proceso completo se puede repetir tantas veces como se especifique por el número de repeticiones. La validación en esta investigación consistió en diez repeticiones y un número de veces de tres.

Por último, se aplicó un análisis *t*-test pareado para comparar las dos fracciones de la EBC del mismo paciente, y se utilizó un *t*-test no pareado para comparar los grupos en estudio utilizando un test de corrección Bonferroni-Holm.

Identificación de las MFs potenciales detectadas por GC-TOF/MS

La identificación se llevó a cabo en primer lugar mediante la búsqueda de espectros MS en la base de datos NIST11. Se consideraron como válidas sólo las identificaciones con un factor de coincidencia y un factor de coincidencia inversa mayor que 700. Los valores RI incluidos en la base de datos NIST también se tuvieron en cuenta para apoyar las identificaciones. Se construyó un modelo de calibración RI mediante la comparación de los valores de RI de una mezcla estándar de alcanos (compuesto por alcano C10 y C40 entre con un número par de átomos de carbono) usando el método cromatográfico propuesto en esta investigación y los valores de RI proporcionados por la base de datos NIST. La figura 1

suplementaria muestra la línea de calibración RI obtenida por este enfoque. El requisito para aceptar las identificaciones del NIST fue que la diferencia entre el RI teórico y el experimental, obtenida por extrapolación de la curva de calibración, debía estar dentro de ± 100 unidades.

- 5 La base de datos NIST no contiene información de alta resolución MS conforme a lo proporciona el detector TOF. Por esta razón, se incluyó un tercer paso para validar la identificación de cada MF mediante el uso de espectrometría de masas de alta resolución. Por lo tanto, la fórmula molecular para el ion precursor tentativo ($[M]^+$) y los iones productos más intensos obtenidos para cada MF debe encajar la identificación NIST fijando un valor de
- 10 corte de precisión de masa de 10 ppm. La Tabla 1 contiene los compuestos identificados clasificados por familias químicas.

Tabla 1. Compuestos en los extractos de EBC identificados por GC–TOF/MS

NOMBRE DEL COMPUESTO	TIEMPO DE RETENCIÓN	FORMULA	CAS ID	Fragmentos	Familia
Eucaliptol	5,995	C ₁₀ H ₁₈ O	470-82-6	154.1361 – [C ₁₀ H ₁₈ O] ⁺ 139.1119 – [C ₉ H ₁₅ O] ⁺ 93.0695 – [C ₇ H ₉] ⁺	Compuestos heteropolicíclicos alifáticos (oxanos)
Indol	9,904	C ₈ H ₇ N	120-72-9	117.0558 – [C ₈ H ₇ N] ⁺ 90.0448 [C ₇ H ₆] ⁺ 74.0145 – [C ₆ H ₂] ⁺	Compuestos heterocíclicos aromáticos (indoles)
Ácido benzoico 4-etoxy-etil ester	12,923	C ₁₁ H ₁₄ O ₃	23676-09-7	194.0425 – [C ₁₁ H ₁₄ O ₃] ⁺ 149.0581 – [C ₉ H ₉ O ₂] ⁺ 121.0269 – [C ₇ H ₅ O ₂] ⁺	Compuestos aromáticos homomonocíclicos (benceno y derivados sustituidos)
Éster metílico del ácido benzoico	6,951	C ₈ H ₈ O ₂	93-58-3	136.0514 – [C ₈ H ₈ O ₂] ⁺ 105.0332 – [C ₇ H ₅ O] ⁺ 77.0378 – [C ₆ H ₅] ⁺	Compuestos aromáticos homomonocíclicos (ácido benzoico y derivados)
Ácido 3,5-di-t-butil-4-hidroxicinámico	20,155	C ₁₇ H ₂₄ O ₃	22014-01-3	276.1712 – [C ₁₇ H ₂₄ O ₃] ⁺ 261.1479 – [C ₁₆ H ₂₁ O ₃] ⁺ 177.0896 – [C ₁₁ H ₁₃ O ₂] ⁺	Compuestos aromáticos homomonocíclicos (ácido cinámico y derivados)
Octinoxato	21,005	C ₁₈ H ₂₆ O ₃	5466-77-3	290.1867 – [C ₁₈ H ₂₆ O ₃] ⁺ 178.0611 – [C ₁₀ H ₁₀ O ₃] ⁺ 161.0578 – [C ₁₀ H ₉ O ₂] ⁺	Compuestos aromáticos homomonocíclicos (ácido cinámico y derivados)
p-Cresol	6,536	C ₇ H ₈ O	106-44-5	136.0514 – [C ₈ H ₈ O] ⁺ 105.0332 – [C ₇ H ₅ O] ⁺ 77.0378 – [C ₆ H ₅] ⁺	Compuestos aromáticos homomonocíclicos (fenoles y derivados-cresol)

ES 2 608 675 A1

Alcohol bencílico	5,963	C ₇ H ₈ O	100-51-6	108.0565 – [C ₇ H ₈ O] ⁺ 91.0535 – [C ₇ H ₇] ⁺ 79.0533 – [C ₆ H ₇] ⁺	Compuestos aromáticos homomonocíclicos (alcoholes primarios)
Laurato de isopropilo	14,97	C ₁₅ H ₃₀ O ₂	10233-13-3	201.1835 – [C ₁₂ H ₂₅ O ₂] ⁺ 157.1203 – [C ₉ H ₁₇ O ₂] ⁺ 102.0656 – [C ₅ H ₁₀ O ₂] ⁺	Lípidos (ésteres de ácidos grasos)
Éster metílico del ácido palmítico	17,342	C ₁₇ H ₃₄ O ₂	112-39-0	270.2545 – [C ₁₇ H ₃₄ O ₂] ⁺ 227.1998 – [C ₁₄ H ₂₇ O ₂] ⁺ 143.1048 – [C ₈ H ₁₅ O ₂] ⁺	Lípidos (ésteres de ácidos grasos)
Éster metílico del ácido oléico	19,041	C ₁₉ H ₃₆ O ₂	112-62-9	296.2702 – [C ₁₉ H ₃₆ O ₂] ⁺ 264.2442 – [C ₁₈ H ₃₄ O] ⁺ 81.0685 – [C ₆ H ₉] ⁺	Lípidos (ésteres de ácidos grasos)
Éster metílico del ácido esteárico	19,28	C ₁₉ H ₃₈ O ₂	112-61-8	298.2862 – [C ₁₉ H ₃₈ O ₂] ⁺ 255.2315 – [C ₁₆ H ₃₁ O ₂] ⁺ 87.0436 – [C ₄ H ₇ O ₂] ⁺	Lípidos (ésteres de ácidos grasos)
Ácido palmítoleico	17,452	C ₁₆ H ₃₀ O ₂	373-49-9	236.2122 – [C ₁₆ H ₂₈ O] ⁺ 98.0710 – [C ₆ H ₁₀ O] ⁺ 69.0689 – [C ₅ H ₉] ⁺	Lípidos (ésteres de ácidos grasos)
Ácido palmítico	17,697	C ₁₆ H ₃₂ O ₂	57-10-3	227.1997 – [C ₁₄ H ₂₇ O ₂] ⁺ 129.0891 – [C ₇ H ₁₃ O ₂] ⁺ 73.0279 – [C ₃ H ₅ O ₂] ⁺	Lípidos (ácidos grasos y conjugados)
Ácido esteárico	19,584	C ₁₈ H ₃₆ O ₂	57-11-4	284.2706 – [C ₁₈ H ₃₆ O ₂] ⁺ 129.0908 – [C ₇ H ₁₃ O ₂] ⁺ 73.0281 – [C ₃ H ₅ O ₂] ⁺	Lípidos (ácidos grasos y conjugados)
Glicidol estearato	22,476	C ₂₁ H ₄₀ O ₃	7460-84-6	297.2436 – [C ₁₈ H ₃₃ O ₃] ⁺ 98.0719 – [C ₆ H ₁₀ O] ⁺ 71.0848 – [C ₅ H ₁₁] ⁺	Lípidos (ácidos grasos y conjugados)
Undecanol	1,125,125	C ₁₁ H ₂₄ O	112-42-5	111.1157 – [C ₈ H ₁₅] ⁺ 83.0844 – [C ₆ H ₁₁] ⁺ 69.0691 – [C ₅ H ₉] ⁺	Lípidos (alcoholes grasos)
1-Hexadecanol 2-metilo	15,994	C ₁₇ H ₃₆ O	2490-48-4	111.1160 – [C ₈ H ₁₅] ⁺ 97.1006 – [C ₇ H ₁₃] ⁺ 69.0691 – [C ₅ H ₉] ⁺	Lípidos (alcoholes grasos)
Oleamida	21,328	C ₁₈ H ₃₅ NO	301-02-0	281.2679 – [C ₁₈ H ₃₅ NO] ⁺ 126.0914 – [C ₇ H ₁₂ NO] ⁺ 72.0438 – [C ₃ H ₆ NO] ⁺	Lípidos (amidas grasas)
11-Eicosenamida	23,01	C ₂₀ H ₃₉ NO	10436-08-5	126.0915 – [C ₇ H ₁₂ NO] ⁺ 72.0439 – [C ₃ H ₆ NO] ⁺ 309.2973 – [C ₂₀ H ₃₉ NO] ⁺	Lípidos (amidas grasas)
Erucamida	24,572	C ₂₂ H ₄₃ NO	112-84-5	337.3338 – [C ₂₂ H ₄₃ NO] ⁺	Lípidos (amidas)

ES 2 608 675 A1

				126.0916 – [C7H12NO] ⁺ 72.0440 – [C3H6NO] ⁺	grasas)
Monopalmitin	22,57	C ₁₉ H ₃₈ O ₄	542-44-9	299.2577 – [C18H35O3] ⁺ 257.2462 – [C16H33O2] ⁺ 239.2366 – [C16H31O] ⁺	Lípidos (glicerolípidos)
Monostearin	24,14	C ₂₁ H ₄₂ O ₄	123-94-4	327.2897 – [C20H39O3] ⁺ 267.2677 – [C18H35O] ⁺ 98.0723 – [C6H10O] ⁺	Lípidos (glicerolípidos)
Hediona	14,408	C ₁₃ H ₂₂ O ₃	24851-98-7	83.0479 – [C5H7O] ⁺ 97.0623 – [C6H9O] ⁺ 226.1566 – [C13H22O3] ⁺	Lípidos (ácido lineólico y derivados del ácido jasmónico)
Esclareol	18,772	C ₂₀ H ₃₆ O ₂	515-03-7	272.2494 – [C20H32] ⁺ 121.0988 – [C9H13] ⁺ 95.0837 – [C7H11] ⁺	Lípidos (prenol lípidos-diterpenos)
D-Limoneno	5,938	C ₁₀ H ₁₆	5989-27-5	136.1227 – [C10H16] ⁺ 121.0992 – [C9H13] ⁺ 79.0524 – [C6H7] ⁺	Lípidos (prenol lípidos- monoterpenos)
Alcohol cumílico	6,805	C ₁₀ H ₁₄ O	617-94-7	121.0644 – [C8H9O] ⁺ 103.0530 – [C8H7] ⁺ 91.0537 – [C7H7] ⁺	Lípidos (prenol lípidos- monoterpenos)
Linalol	6,999	C ₁₀ H ₁₈ O	78-70-6	136.1218 – [C10H16] ⁺ 93.0679 – [C7H9] ⁺ 71.0844 – [C5H4] ⁺	Lípidos (prenol lípidos- monoterpenos)
Alcanfor	7,835	C ₁₀ H ₁₆ O	464-48-2	152.1192 – [C10H16] ⁺ 137.0963 – [C9H13O] ⁺ 95.0852 – [C7H11]	Lípidos (prenol lípidos- monoterpenos)
Camphol	8,160,999	C ₁₀ H ₁₈ O	507-70-0	121.1000 – [C9H13] ⁺ 95.0853 – [C7H11] ⁺ 77.0381 – [C6H5] ⁺	Lípidos (prenol lípidos- monoterpenos)
Levomentol	8,251	C ₁₀ H ₂₀ O	2216-51-5	138.1379 – [C10H18] ⁺ 95.0837 – [C7H11] ⁺ 81.0683 – [C6H9] ⁺	Lípidos (prenol lípidos- monoterpenos)
Terpineol	8,516	C ₁₀ H ₁₈ O	98-55-5	136.1244 – [C10H16] ⁺ 121.1007 – [C9H13] ⁺ 93.0695 – [C7H9] ⁺	Lípidos (prenol lípidos- monoterpenos)
Escualeno	24,835	C ₃₀ H ₅₀	111-02-4	410.3907 – [C30H50] ⁺ 121.0994 – [C9H13] ⁺ 81.0686 – [C6H9] ⁺	Lípidos (prenol lípidos-triterpenos)
Colestadieno	25,554	C ₂₇ H ₄₄	747-90-0	368.3437 – [C27H44] ⁺ 247.2412 – [C18H31] ⁺ 147.1141 – [C11H15] ⁺	Lípidos (esteroides y derivados de esteroides)

ES 2 608 675 A1

Citrato de trietilo	14,404	C ₁₂ H ₂₀ O ₇	77-93-0	203.0913 – [C ₉ H ₁₅ O ₅] ⁺ 157.0496 – [C ₇ H ₉ O ₄] ⁺ 83.0486 – [C ₅ H ₇ O] ⁺	Ácidos orgánicos y derivados (ácidos carboxílicos y derivados)
Espiro [2,4] heptano-1,5-dimetil-6-metileno	5,953	C ₁₀ H ₁₆	62238-24-8	136.1225 – [C ₁₀ H ₁₆] ⁺ 121.0990 – [C ₉ H ₃] ⁺ 79.0524 – [C ₆ H ₇] ⁺	Otros compuestos orgánicos
2-propanol 1-(2-butoxi-1-metil-etoxi)	9,21	C ₁₀ H ₂₂ O ₃	29911-28-2	59.0485 – [C ₃ H ₇ O] ⁺ 86.0715 – [C ₅ H ₁₀ O] ⁺ 103.0728 – [C ₅ H ₁₁ O ₂] ⁺	Otros compuestos orgánicos
2,4,6-Triisopropilfenol	12,685	C ₁₇ H ₂₆ O	08/07/1934	220.1822 – [C ₁₅ H ₂₄ O] ⁺ 205.1584 – [C ₁₄ H ₂₁ O] ⁺ 77.0369 – [C ₆ H ₅] ⁺	Otros compuestos orgánicos
1,3-Heptadecyn-1-ol	15,11	C ₁₇ H ₃₂ O	56554-77-9	225.1826 – [C ₁₄ H ₂₅ O ₂] ⁺ 81.0681 – [C ₆ H ₉] ⁺ 67.0529 – [C ₅ H ₇] ⁺	Otros compuestos orgánicos
2,4-difenil-4-metil-2 (E) penteno	16,45	C ₁₈ H ₂₀	22768-22-5	236.1567 – [C ₁₈ H ₂₀] ⁺ 143.0809 – [C ₁₁ H ₁₁] ⁺ 91.0513 – [C ₇ H ₇] ⁺	Otros compuestos orgánicos
7,9-Di-t-butil-oxaespíro (4,5) deca-6,9-dieno-2,8-diona	17,182	C ₁₇ H ₂₆ O ₂	82304-66-3	175.1104 – [C ₁₂ H ₁₅ O] ⁺ 133.0638 – [C ₉ H ₉ O] ⁺ 77.0369 – [C ₆ H ₅] ⁺	Otros compuestos orgánicos
10,18-Bisnorabieta-8,11,13-trieno	17,999,014	C ₁₈ H ₂₆	32624-67-2	242.2007 – [C ₁₈ H ₂₆] ⁺ 227.1790 – [C ₁₇ H ₂₃] ⁺ 143.0864 – [C ₁₁ H ₁₁ O] ⁺	Otros compuestos orgánicos
2,6-Di-t-butil-4-(2-fenilpropan-2-il) fenol	18,709	C ₂₃ H ₃₂ O	34624-81-2	324.2438 – [C ₂₃ H ₃₂ O] ⁺ 309.2212 – [C ₂₂ H ₂₉ O] ⁺ 119.0836 – [C ₉ H ₁₁] ⁺	Otros compuestos orgánicos
Fenol 2,2'-metileno-bis [6-(1,1-dimetil)-4-metilo	21,768	C ₂₃ H ₃₂ O ₂	119-47-1	330.1984 – [C ₂₄ H ₂₆ O] ⁺ 315.1748 – [C ₂₃ H ₂₃ O] ⁺ 237.1263 – [C ₁₇ H ₁₇ O] ⁺	Otros compuestos orgánicos
Fenol 2,4-bis (1-metil-1-feniletíl) -	22,522	C ₂₄ H ₂₆ O	2772-45-4	330.1984 – [C ₂₄ H ₂₆ O] ⁺ 315.1748 – [C ₂₃ H ₂₃ O] ⁺ 237.1263 – [C ₁₇ H ₁₇ O] ⁺	Otros compuestos orgánicos
2,4-Bis (dimetilbencil) -6-t-butilfenol	22,57	C ₂₈ H ₃₄ O	244080-16-8	386.2617 – [C ₂₈ H ₃₄ O] ⁺ 371.2370 – [C ₂₇ H ₃₁ O] ⁺ 293.1897 – [C ₂₁ H ₂₅ O] ⁺	Otros compuestos orgánicos
N-hexadecilindano	23,928	C ₂₅ H ₄₂	55334-29-7	117.0351 – [C ₉ H ₉] ⁺ 130.0427 – [C ₁₀ H ₁₀] ⁺ 154.1345 – [C ₁₁ H ₂₂] ⁺	Otros compuestos orgánicos

Resultados

Comparación entre las muestras UA y DA

El dispositivo de muestreo contiene válvulas que permiten la separación según la vía de expiración, además del fraccionamiento del volumen exhalado de acuerdo con un umbral de volumen en las dos cavidades. Esta configuración hace la contaminación por saliva altamente improbable [28]. A pesar de que se supone el fraccionamiento de EBC para separar los componentes exógenos y endógenos, no hay estudios previos que hayan evaluado la composición de las dos fracciones de EBC. Todos los compuestos identificados estuvieron presentes en ambas muestras UA y DA, pero algunos de ellos mostraron diferencias entre las fracciones en cuanto a su concentración relativa. La Figura 1 muestra el diagrama de barras que compara la media de la concentración relativa para los 47 compuestos detectados en las muestras. Como se puede ver, la mayoría de los compuestos identificados se encontraron en DA en concentraciones más altas que en la UA. Un test-t pareado reveló que 18 de estos compuestos tenían concentraciones significativamente diferentes en las fracciones en las muestras recogidas en los dos compartimentos (valor de $p < 0,05$), lo que significa que casi el 38% de los componentes de EBC identificados presentaron diferentes concentraciones en UA y DA (Figura 1). Nueve compuestos (entre ellos los ésteres de ácidos grasos y escualeno) estaban más concentrados en DA, mientras que otros como el esclareol, el limoneno, ácido benzoico o 11-eicosenamida estaban más concentrados en las muestras UA. Como las dos fracciones de EBC resultaron en composición cuantitativamente diferentes, se llevó a cabo un análisis estadístico multivariante de cada fracción independientemente.

Análisis multivariante de cada fracción de vía respiratoria

Se aplicó un análisis no supervisado para encontrar diferencias en EBC recogidas en los tres grupos de muestras para las fracciones UA y DA. El análisis de componentes principales se llevó a cabo con el conjunto de datos que incluyó 44 de los 47 compuestos identificados. Se excluyeron tres compuestos ya que su origen sólo podría ser explicado por fuentes externas tales como cosméticos (esclareol y octinoxato) o las bolsas de plástico empleadas para la recogida de EBC (7,9-di-*t*-butil-1-oxaespíro (4,5) deca -6,9-dieno-2,8-diona). Debido a la alta variabilidad asociada a los individuos, los tres grupos estudiados (cáncer de pulmón, grupo con factor de riesgo y grupo de individuos sanos) aparecieron superpuestos en las tres dimensiones del gráfico *PCA score* para UA y DA (Fig. suplementaria 2). Por esta razón, el análisis no supervisado se dividió en tres estudios mediante la inclusión de sólo dos grupos: el cáncer de pulmón en comparación con grupo

con factor de riesgo, el cáncer de pulmón en comparación con individuos sanos de control, y el grupo con factor de riesgo en comparación con individuos sanos de control. Las figuras suplementarias 3 y 4 ilustran los gráficos 3D de los *scores* para los tres casos utilizando la UA y DA, respectivamente. En todos los casos se observaron tendencias de discriminación entre los dos grupos. Ambas fracciones EBC mostraron una tendencia de separación similar; sin embargo, las fuentes de variabilidad intra-individuales e inter-individuales no permiten la completa separación de los grupos evaluados. En las diferentes pruebas de PCA, la combinación de PC1/PC2 /PC3 no explica la variabilidad por encima de 50% de la variabilidad total contenida en el estudio.

Se aplicó el análisis supervisado para encontrar patrones de discriminación asociados al diagnóstico de cáncer de pulmón. Como en el estudio anterior, el análisis se dividió en tres estudios: cáncer de pulmón en comparación con grupo de factor de riesgo, el cáncer de pulmón en comparación con individuos sanos de control, y el grupo con factor de riesgo en comparación con los individuos de control. Las figuras 2 y 3 ilustran los *scores* en gráficos 3D de los tres estudios independientes para cada fracción de EBC. Como se puede ver, el porcentaje de variabilidad explicada no superó el 50%, pero en este caso las tendencias de discriminación se observaron claramente en todos los casos. El porcentaje de muestras correctamente clasificados para la validación y formación conjuntos del resultado PLS-DA para UA y DA se recoge en la Tabla 2. Los pacientes con cáncer de pulmón se clasifican con precisión en el modelo combinado con individuos sanos, especialmente en la fracción de la AU que reportó una capacidad de reconocimiento del 97,0% en el conjunto de entrenamiento y del 88,2% en la etapa de validación. La capacidad de reconocimiento disminuyó cuando el grupo con factor de riesgo se incluyó en el análisis de clasificación, lo que es bastante lógico ya que es el grupo intermedio. Así, en la UA la capacidad de reconocimiento para la separación de este grupo tanto de los pacientes con cáncer de pulmón como de individuos sanos fue similar: 76,5% para el conjunto de entrenamiento y 60,5% para el conjunto de validación. Estos valores se incrementaron cuando la fracción DA fue la muestra objetivo. La capacidad de reconocimiento también fue similar en la fracción DA, con 83 y 84% en el conjunto de entrenamiento utilizado para la separación del grupo con factor de riesgo de pacientes con cáncer de pulmón y de los individuos sanos, respectivamente, y el 73 y el 77,5% para los conjuntos de validación. Desde un punto de vista clínico, el modelo más interesante para ayudar en el diagnóstico de cáncer de pulmón es la comparación entre pacientes con cáncer de pulmón y el grupo con factor de riesgo. Atendiendo al parámetro con capacidad de reconocimiento, la fracción DA se caracterizó por una capacidad de clasificación más alta que la fracción UA. El modelo proporcionado por el

análisis DA está bien equilibrado en términos de sensibilidad y especificidad, siendo estos parámetros 89,8 y 79,5%, respectivamente. Por esta razón, se seleccionó la fracción DA para otros estudios que se ocupan de identificación de compuestos importantes que contribuyen a explicar los patrones observados.

- 5 **Tabla 2.** Frecuencias de muestreo clasificadas para la validación, formación y capacidad de reconocimiento para el desarrollo de los modelos PLS-DA proporcionados por las vías respiratorias superiores/centro (UA) y la vía aérea distal (DA) para la comparación de los tres grupos entre sí (controles —Control—, controles con factor de riesgo —RF— y pacientes con cáncer de pulmón —LC—)

Vía respiratoria	UA			DA		
	Modelos	RF vs LC	Control vs LC	Control vs RF	CRF vs LC	Control vs LC
Modelo de entrenamiento						
Sensibilidad (%)	92,857	92,857	67,532	89,796	89,796	81,818
Especificidad (%)	67,532	97,057	97,059	79,542	94,000	90,000
Capacidad de reconocimiento (%)	76,471	94,737	76,577	83,212	91,919	84,783
Modelo de validación						
Sensibilidad (%)	73,810	76,190	55,844	83,673	77,551	75,000
Especificidad (%)	53,810	88,235	73,529	67,045	82,000	82,000
Capacidad de reconocimiento (%)	60,504	71,579	61,261	72,993	79,798	77,536

10

Identificación de compuestos significativos que contribuyen a explicar la variabilidad clínica

Como se ha mencionado antes, los compuestos identificados en el EBC de los tres grupos estudiados y utilizados para construir los modelos de predicción se muestra en la Tabla 1. Se aplicó un t-test no pareado para identificar los compuestos más importantes que contribuyeron a explicar las diferencias observadas entre los tres grupos (tabla 3). Cuatro compuestos resultaron significativos en la comparación de los pacientes con cáncer de pulmón en comparación con el grupo de factor de riesgo y que forman por individuos sanos.

15

Entre estos compuestos, es de destacar la presencia de dos monoacilgliceroles saturados (monopalmitina y monoestearina), y un triterpenoide acíclico (escualeno), que es el precursor de los esteroides, incluyendo el colesterol, y ácidos biliares [29]. La presencia de escualeno en el aliento exhalado se ha descrito ya ampliamente [30]. Este compuesto es estructuralmente similar al isopreno, que se considera uno de los compuestos más concentrados en la respiración humana [22]. De hecho, algunos autores han propuesto que poliisoprenos tales como escualeno se consideran fuentes potenciales de isopreno por peroxidación, uno de los mecanismos del estrés oxidativo [30]. Sin embargo, el isopreno no se detectó en EBC en esta investigación, lo que podría explicarse por su alta volatilidad. Las diferencias en la concentración relativa de los dos monoacilgliceroles y escualeno en los tres grupos estudiados pueden visualizarse en la Figura 4. La monopalmitina y monoestearina se caracterizaron por diferentes comportamientos: la monopalmitina estuvo más concentrada en el grupo de factor de riesgo que en los pacientes con cáncer de pulmón, que también presentó mayor concentración de este monoacilglicérol que los individuos sanos. Por otro lado, la monoestearina presentó el perfil opuesto: como el grupo de factor de riesgo dio una concentración relativa más baja que los pacientes con cáncer de pulmón. El escualeno dio un perfil de concentración similar al de la monopalmitina. El cuarto compuesto (hexadecilindano, un derivado del indano) no se ha relacionado con alguna fuente endógena y, por lo tanto, podría atribuirse a un origen exógeno. Sin embargo, Phillips *et al.* [31] también han detectado derivados de indano en la respiración, e incluso seleccionaron uno de ellos como biomarcador de cáncer de pulmón [32].

Tabla 3. Análisis de t-test no pareado usando un test de corrección múltiple de Bonferroni-Holm para evaluar la importancia de los metabolitos identificados en condensado de aire exhalado para discriminar entre los pacientes con cáncer de pulmón, los individuos con factores de riesgo y los individuos control.

CÁNCER DE PULMÓN VS CONTROL			
COMPUESTO	P	REGULACIÓN	FC
p-Cresol	9,00E-04	BAJA	-3,10466
Indol	3,51E-04	BAJA	-2,77727
Ácido benzoico 4-etoxy-etil ester	0,003489	BAJA	-1,58336
Citrato de trietilo	0,009728	ALTA	2,233156
Monopalmitin	0,015986	ALTA	2,205814
11-Eicosenamida	3,13E-05	ALTA	2,16227
n- Hexadecilindano	0,002707	ALTA	1,952551
Monostearin	0,003911	ALTA	2,608156
LUNG CANCER VS CONTROL WITH RISK FACTOR			
COMPUESTO	P	REGULATION	FC
Hedione	0,041098	ALTA	1,234381

13-Heptadecin-1-ol	0,002017	ALTA	1,4398
Monopalmitina	0,003899	ALTA	2,630078
n- Hexadecilindano	4,81E-04	ALTA	1,815379
Monostearina	1,07E-04	ALTA	3,018592
Escualeno	2,12E-04	ALTA	2,431075
CONTROL VS CONTROL WITH RISK FACTOR			
COMPOUND	P	REGULATION	FC
<i>p</i>-Cresol	0,009083	BAJA	-2,27088
Indol	0,004347	BAJA	-2,0566
Undecanol	0,043269	BAJA	-1,41217
Ácido benzoico 4-etoxy-etil éster	0,004672	BAJA	-1,35418
Citrato de trietilo	0,027726	ALTA	1,854397
Hediona	0,01682	ALTA	1,223202
Monostearin	1,36E-07	ALTA	5,801463
c11-Eicosenamida	8,68E-05	ALTA	1,799589
n- Hexadecilindano	3,75E-12	ALTA	3,544619
Monostearin	1,74E-12	ALTA	7,87296
13-Docosenamida	5,19E-04	ALTA	1,485012

Tomando el grupo de individuos sanos como referencia, seis compuestos resultaron significativos en la comparación frente a los pacientes con cáncer de pulmón y el grupo con factor de riesgo. Entre ellos el citrato de trietilo, que se ha encontrado en plásticos y en filtros de cigarrillos, detectado en una concentración más alta en pacientes con cáncer de pulmón y en los individuos con factor de riesgo que en individuos sanos, [33]; por lo tanto, su presencia en el aliento exhalado podría vincularse sólo a fuentes exógenas. Un compuesto fenólico (*p*-cresol) y un derivado del fenol (éster del ácido 4-etoxietilo benzoico) también fueron encontrados a diferentes concentraciones en los tres grupos. El *p*-cresol se detectó a una concentración más baja en los pacientes con cáncer de pulmón que en los individuos con factor de riesgo, que también reportaron niveles más bajos que los individuos sanos. El derivado del éster de ácido benzoico se encontró en concentraciones más bajas en el grupo de cáncer de pulmón y los individuos con factor de riesgo que en el grupo sano. Philipps *et al.* han identificado este compuesto en la respiración humana como un marcador potencial candidato para discriminar a los pacientes de cáncer de pulmón [32].

Otro compuesto interesante fue el indol, que está implicado en el metabolismo del triptófano, particularmente en bacterias [34]. Además, se ha encontrado indol en el humo del cigarrillo y, en este contexto, podría estar asociado con el consumo de tabaco [35]. El perfil de concentración relativa de este compuesto se caracteriza por una concentración más alta en el grupo con factor de riesgo seguido de pacientes con cáncer de pulmón y, a continuación, a un nivel inferior en individuos sanos. Por último, otros compuestos que contribuyen a

diferenciar individuos sanos fueron eicosenamida y erucamida, presente a niveles inferiores en este grupo en comparación con los individuos del grupo factor de riesgo y de cáncer de pulmón. Estas amidas grasas no se habían descrito de forma endógena, pero algún análogo tal como oleamida, que también se halló en EBC, está estructuralmente relacionado con el cannabinoide endógeno anandamida, implicado en muchas funciones biológicas. Las amidas grasas también se utilizan como aditivos de plástico y, por esta razón, una fuente de contaminación también podría explicar la presencia de eicosenamida y erucamida en el EBC.

Otros tres compuestos que fueron significativas en la comparación del grupo de factor de riesgo y los otros dos grupos fueron dos alcoholes grasos (13-heptadecyn-1-ol y undecanol) y un derivado del ácido jasmónico (hediona). Respecto al 13 heptadecyn-1-ol, no hay información anterior que lo describa en el aliento o sobre su implicación biológica en los procesos humanos. El otro ácido graso (undecanol), junto con otros 200 compuestos, se han detectado previamente en heces por Garner *et al.*, quienes estudiaron el potencial del undecanol para el diagnóstico de enfermedades gastrointestinales [36]. El último (hediona), que ha sido detectada previamente en la saliva [37], podría ser un producto de la oxidación de ácido linoleico.

En resumen, y como se ha expuesto anteriormente, el EBC no ha sido ampliamente explotado en el campo clínico a pesar de las ventajas asociadas a su muestreo. Un método de análisis metabolómico de EBC basado en perfiles GC–TOF/MS en modo de alta resolución ha sido desarrollado usando la extracción líquido–líquido para la preparación de la muestra. Los compuestos identificados se han utilizado para discriminar entre tres grupos diferentes: los individuos diagnosticados con cáncer de pulmón, los individuos con factor de riesgo (incluyendo fumadores y pacientes con EPOC) y el control formado por individuos sanos. Entre estos compuestos cabe mencionar la presencia de derivados del monoacilglicerol de dos de los cuatro ácidos grasos saturados principales y el escualeno, que podría ser considerado como un producto intermedio en la vía para la formación *in vivo* por la peroxidación de isopreno en aliento humano [39], y también involucrado en la síntesis de colesterol [40]. Estos resultados apoyan el potencial de EBC como biofluido para discriminar entre los pacientes con cáncer de pulmón y el grupo con factor de riesgo, lo que podría ayudar en el diagnóstico de esta enfermedad en la búsqueda de un método de cribado para reducir el uso de una prueba de confirmación para el caso de respuesta positiva de los primeros.

Los resultados obtenidos se podrían resumir en la siguiente tabla:

Tabla 4. Resumen de los marcadores útiles en el diagnóstico, clasificación y seguimiento del cáncer de pulmón.

PACIENTES CON CÁNCER DE PULMÓN VS. SANOS*	
METABOLITO	THRESHOLDS
<i>p</i> -Cresol	<0.0143707
Alcohol cumílico	>0.0172175
Eicosenamida	>0.00450892
Hexadecilindano	>0.0268941
PACIENTES CON CÁNCER DE PULMÓN VS. FACTOR DE RIESGO*	
METABOLITO	THRESHOLDS
2,4-Bis-dimetilbencil-6-t-butilfenol	>0.0773647
Monostearina	<0.0666031
Spiro-2,4-heptane-1,5-6-methylene	>210.452
13-Heptadecyn-1-ol	<0.00438287
Estearato de metilo	>0.00245645
PACIENTES FACTOR DE RIESGO VS. SANOS**	
METABOLITO	THRESHOLDS
Glycerol-1-palmitato	>0.0818449
Alcohol bencílico	>0.109225
Monostearina	>0.053136
2,4-Diphenyl-4-metil-2-E-pentano	<0.00557449
<i>p</i> -Cresol	<0.00680814

*(positivo (cáncer) cuando al menos 3 metabolitos cumplen la condición)

** (positivo (riesgo) cuando al menos 2 metabolitos cumplen la condición)

- 5 Los valores proporcionados en esta tabla están expresados como el tanto por uno respecto a todos los metabolitos detectados en la muestra (44 metabolitos). Es decir, el área de cada uno de los compuestos se dividió por la suma del área de los 44 metabolitos, obteniendo así la contribución de cada compuesto al total de metabolitos detectados. Sobre estos valores

finales se han obtenido los paneles, de forma que los resultados puedan compararse con los obtenidos usando otra instrumentación.

5 En cuanto al funcionamiento de los paneles en sí, están configurados de forma que, aunque estén compuestos por 4-5 metabolitos, sólo es necesario que 2 o 3 de ellos cumplan el punto de corte establecido para que el panel se considere positivo. Por ejemplo, en el primer panel, un individuo cuyo EBC presente niveles, expresados en tanto por uno, de *p*-cresol menores de 0.0143707, de alcohol cumílico mayores de 0.172175 y de eicosenamida mayores de 0.00450892, se clasificará como paciente con cáncer, sin importar los niveles de dehexadecilindano que presente.

BIBLIOGRAFÍA

- [1] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA. Cancer J. Clin. 61 (2011) 69–90.
- [2] P. Burney, A. Jithoo, B. Kato, C. Janson, D. Mannino, E. Nizankowska-Mogilnicka, et al.,
 5 Chronic obstructive pulmonary disease mortality and prevalence: the associations with smoking and poverty: a bold analysis, Thorax. 69 (2014) 465–473.
- [3] A. Jemal, M.M. Center, C. DeSantis, E.M. Ward, Global patterns of cancer incidence and mortality rates and trends, Cancer Epidemiol. Biomarkers Prev. 19 (2010) 1893–1907.
- [4] P.M. Boiselle, Computed tomography screening for lung cancer, JAMA. 309 (2013) 1163–
 10 1170.
- [5] O.M. Mets, M. Schmidt, C.F. Buckens, M.J. Gondrie, I. Isgum, M. Oudkerk, et al., Diagnosis of chronic obstructive pulmonary disease in lung cancer screening computed tomography scans: independent contribution of emphysema, air trapping and bronchial wall thickening, Respir. Res. 14 (2013) 59.
- [6] R.D. Beger, A review of applications of metabolomics in cancer., Metabolites. 3 (2013)
 15 552–74.
- [7] M. Calderón-Santiago, F. Priego-Capote, B. Jurado-Gámez, M.D. Luque de Castro, Optimization study for metabolomics analysis of human sweat by liquid chromatography-tandem mass spectrometry in high resolution mode, J. Chromatogr. A. 1333 (2014) 70–78.
- [8] M. Calderón-Santiago, F. Priego-Capote, N. Turck, X. Robin, B. Jurado-Gámez, J.C. Sanchez, et al., Human sweat metabolomics for lung cancer screening., Anal. Bioanal. Chem. (2015).
 20
- [9] C.Y. Liu, C.H. Wang, T.C. Chen, H.C. Lin, C.T. Yu, H.P. Kuo, Increased level of exhaled nitric oxide and up-regulation of inducible nitric oxide synthase in patients with primary lung cancer., Br. J. Cancer. 78 (1998) 534–41.
 25
- [10] I. Horváth, Z. Lázár, N. Gyulai, M. Kollai, G. Losonczy, Exhaled biomarkers in lung cancer, Eur. Respir. J. 34 (2009) 261–275.
- [11] G.E. Carpagnano, O. Resta, M.P. Foschino-Barbaro, E. Gramiccioni, F. Carpagnano, Interleukin-6 is increased in breath condensate of patients with non-small cell lung cancer., Int. J. Biol. Markers. 17 141–5.
 30

- [12] D.L. Bayley, H. Abusriwil, A. Ahmad, R.A. Stockley, Validation of assays for inflammatory mediators in exhaled breath condensate., *Eur. Respir. J.* 31 (2008) 943–8.
- [13] E. Dalaveris, T. Kerenidi, A. Katsabeki-Katsafli, T. Kiropoulos, K. Tanou, K.I. Gourgoulianis, et al., VEGF, TNF-alpha and 8-isoprostane levels in exhaled breath condensate and serum of patients with lung cancer, *Lung Cancer.* 64 (2009) 219–225.
- [14] K. Kostikas, G. Papatheodorou, K. Psathakis, P. Panagou, S. Loukides, Oxidative stress in expired breath condensate of patients with COPD., *Chest.* 124 (2003) 1373–80.
- [15] C. Gessner, S. Hammerschmidt, H. Kuhn, H.-J. Seyfarth, U. Sack, L. Engelmann, et al., Exhaled breath condensate acidification in acute lung injury., *Respir. Med.* 97 (2003) 1188–94.
- [16] B. Antus, I. Barta, Exhaled breath condensate pH in patients with lung cancer., *Lung Cancer.* 75 (2012) 178–80.
- [17] M. Phillips, K. Gleeson, J.M. Hughes, J. Greenberg, R.N. Cataneo, L. Baker, et al., Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study., *Lancet.* 353 (1999) 1930–3.
- [18] M. Phillips, R.N. Cataneo, A.R.C. Cummin, A.J. Gagliardi, K. Gleeson, J. Greenberg, et al., Detection of lung cancer with volatile markers in the breath, *Chest.* 123 (2003) 2115–2123.
- [19] M. Phillips, N. Altorki, J.H.M. Austin, R.B. Cameron, R.N. Cataneo, J. Greenberg, et al., Prediction of lung cancer using volatile biomarkers in breath, *Cancer Biomark.* 3 (2007) 95–109.
- [20] Y. Wang, Y. Hu, D. Wang, K. Yu, L. Wang, Y. Zou, et al., The analysis of volatile organic compounds biomarkers for lung cancer in exhaled breath, tissues and cell lines., *Cancer Biomark.* 11 (2012) 129–37.
- [21] G. Hedlin, J. Konradsen, A. Bush, An update on paediatric asthma, *Eur. Respir. Rev.* 21 (2012) 175–185.
- [22] G. Hillas, S. Loukides, K. Kostikas, P. Bakakos, Biomarkers obtained by non-invasive methods in patients with COPD: where do we stand, what do we expect?, *Curr. Med. Chem.* 16 (2009) 2824–2838.

- [23] P. Kubáň, F. Foret, Exhaled breath condensate: determination of non-volatile compounds and their potential for clinical diagnosis and monitoring. A review, *Anal. Chim. Acta.* 805 (2013) 1–18.
- [24] C. Costa, C. Bucca, M. Bergallo, P. Solidoro, G. Rolla, R. Cavallo, Unsuitability of
5 exhaled breath condensate for the detection of herpesviruses DNA in the respiratory tract, *J. Virol. Methods.* 173 (2011) 384–386.
- [25] S. Dragonieri, P. Brinkman, E. Mouw, A.H. Zwinderman, P. Carratú, O. Resta, et al., An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls, *Respir. Med.* 107 (2013) 1073–1078.
- 10 [26] J. Hunt, Exhaled breath condensate: An evolving tool for noninvasive evaluation of lung disease, *J. Allergy Clin. Immunol.* 110 (2002) 28–34.
- [27] A. Mazzatenta, C. Di Giulio, M. Pokorski, Pathologies currently identified by exhaled biomarkers, *Respir. Physiol. Neurobiol.* 187 (2013) 128–134.
- [28] F. Hoffmeyer, M. Raulf-Heimsoth, V. Harth, J. Bünger, T. Brüning, Comparative analysis
15 of selected exhaled breath biomarkers obtained with two different temperature-controlled devices, *BMC Pulm. Med.* 9 (2009) 48.
- [29] G. Salvioli, R. Lugli, J.M. Pradelli, Relationships between squalene and cholesterol in bile: Effect of ursodeoxycholic acid administration in patients with radiolucent gallstones, *Metabolism.* 33 (1984) 641–645.
- 20 [30] M. Phillips, J. Greenberg, Method for the collection and analysis of volatile compounds in the breath, *J. Chromatogr. B Biomed. Sci. Appl.* 564 (1991) 242–249.
- [31] M. Phillips, Variation in volatile organic compounds in the breath of normal humans, *J. Chromatogr. B.* 729 (1999) 75–88.
- [32] M. Phillips, N. Altorki, J.H.M. Austin, R.B. Cameron, R.N. Cataneo, R. Kloss, et al.,
25 Detection of lung cancer using weighted digital analysis of breath biomarkers., *Clin. Chim. Acta.* 393 (2008) 76–84.
- [33] Smoke filters, (1968).
- [34] A. Nowak, Z. Libudzisz, Influence of phenol, p-cresol and indole on growth and survival of intestinal lactic acid bacteria., *Anaerobe.* 12 (2006) 80–4.

- [35] M.E. Snook, R.F. Arrendale, H.C. Higman, O.T. Chortyk, Isolation of indoles and carbazoles from cigarette smoke condensate, *Anal. Chem.* 50 (1978) 88–90.
- [36] C.E. Garner, S. Smith, B. de Lacy Costello, P. White, R. Spencer, C.S.J. Probert, et al., Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease., *FASEB J.* 21 (2007) 1675–88.
5
- [37] B. de Lacy Costello, A. Amann, H. Al-Kateb, C. Flynn, W. Filipiak, T. Khalid, et al., A review of the volatiles from the healthy human body., *J. Breath Res.* 8 (2014) 014001.
- [38] J.J. Kabara, D.M. Swieczkowski, A.J. Conley, J.P. Truant, Fatty acids and derivatives as antimicrobial agents, *Antimicrob. Agents Chemother.* 2 (1972) 23–28.
- 10 [39] R.A. Stein, J.F. Mead, Small hydrocarbons formed by the peroxidation of squalene, *Chem. Phys. Lipids.* 46 (1988) 117–120.
- [40] R.B. Woodward, K. Bloch, The cyclization of squalene in cholesterol synthesis, *J. Am. Chem. Soc.* 75 (1953) 2023–2024.

REIVINDICACIONES

- 1.- Uso de los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida, el hexadecilindano, el 2,4-bis-dimetilbencil-6-*t*-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol, el estearato de metilo, el glicerol-1-palmitato, el alcohol bencílico y el
- 5 2,4-difenil-4-metil-2-*E*-pentano para el diagnóstico, clasificación y seguimiento del cáncer.
- 2.- El uso de los metabolitos según la reivindicación anterior, donde el cáncer es cáncer de pulmón.
- 3.- Un método de obtención de datos útiles para el diagnóstico, clasificación y seguimiento de un individuo o sujeto que potencialmente sufra cáncer que comprende:
- 10 a) cuantificar los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida y el hexadecilindano en una muestra biológica aislada de dicho individuo.
- 4.- El método según la reivindicación anterior, que además comprende:
- b) comparar las cantidades obtenidas en el paso (a) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos
- 15 metabolitos en individuos sanos.
- 5.- El método según cualquiera de las reivindicaciones 3-4, donde el cáncer es cáncer de pulmón.
- 6.- El método según cualquiera de las reivindicaciones 3-5, donde la muestra biológica aislada del paso (a) es aire exhalado.
- 20 7.- El método según cualquiera de las reivindicaciones 3-6, donde la muestra biológica aislada del paso a) es condensado de aire exhalado.
- 8.- El método según cualquiera de las reivindicaciones 3-7, donde la cuantificación de los metabolitos del paso (a) se realiza por análisis mediante GC-TOF/MS, preferiblemente con ionización por impacto electrónico (IE).
- 25 9.- Un método de diagnóstico, clasificación y seguimiento del cáncer, que comprende los pasos (a)-(b) según cualquiera de las reivindicaciones 3-8, que además comprende:
- c) incluir al individuo del paso a) en el grupo de individuos que presentan cáncer, cuando se detectan niveles de al menos 3 metabolitos, preferiblemente de los 4 metabolitos del paso (a), en cantidades significativamente distintas a la cantidad de referencia.

10.- El método según la reivindicación anterior, donde el cáncer es cáncer de pulmón.

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

5 a) cuantificar los metabolitos 2,4-bis-dimetilbencil-6-t-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol y el estearato de metilo en una muestra biológica aislada de dicho individuo.

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

10 b) comparar las cantidades obtenidas en el paso (a) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos que presentan al menos un factor de riesgo de padecer cáncer, preferiblemente tabaquismo o con EPOC.

13.- El método según cualquiera de las reivindicaciones 11-12, donde el cáncer es cáncer de pulmón.

15 14.- El método según cualquiera de las reivindicaciones 11-13, donde la muestra biológica aislada del paso (a) es aire exhalado condensado.

15- El método según cualquiera de las reivindicaciones 11-14, donde la muestra biológica aislada del paso a) es condensado de aire exhalado.

20 16.- El método según cualquiera de las reivindicaciones 11-15, donde la cuantificación de los metabolitos del paso (a) se realiza por análisis mediante GC-TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

17.- Un método de diagnóstico, clasificación y seguimiento del cáncer, que comprende los pasos (a)-(b) según cualquiera de las reivindicaciones 11-16, que además comprende:

25 c) incluir al individuo del paso a) en el grupo de individuos que presentan cáncer, cuando se detectan niveles de al menos 3 metabolitos, preferiblemente de 4 metabolitos, y más preferiblemente de 5 metabolitos del paso (a), en cantidades significativamente distintas cantidad de referencia.

18.- El método según la reivindicación anterior, donde el cáncer es cáncer de pulmón.

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

a) cuantificar los metabolitos glicerol-1-palmitato, el alcohol bencílico, la monostearina, el 2,4-difenil-4-metil-2-E-pentano y el *p*-cresol en una muestra biológica aislada de dicho individuo.

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

5 b) comparar las cantidades obtenidas en el paso (a) con una cantidad de referencia, donde la cantidad de referencia para cada metabolito son los niveles medios de dichos metabolitos en individuos sanos.

21.- El método según cualquiera de las reivindicaciones 19–20, donde el cáncer es cáncer de pulmón.

10 22.- El método según cualquiera de las reivindicaciones 19–21, donde la muestra biológica aislada del paso (a) es aire exhalado.

23.- El método según cualquiera de las reivindicaciones 19–22, donde la muestra biológica aislada del paso a) es condensado de aire exhalado.

15 24.- El método según cualquiera de las reivindicaciones 19–23, donde la cuantificación de los metabolitos del paso (a) se realiza por análisis mediante GC–TOF/MS, preferiblemente con ionización por impacto electrónico (IE).

25.- Un método de diagnóstico, clasificación y seguimiento del cáncer, que comprende los pasos (a)–(b) según cualquiera de las reivindicaciones 19–23, que además comprende:

20 c) incluir al individuo del paso a) en el grupo de individuos de alto riesgo de padecer cáncer, cuando se detectan niveles de al menos 2 metabolitos, preferiblemente de 3 metabolitos, más preferiblemente de 4 metabolitos, y aún más preferiblemente de 5 metabolitos del paso (a), en cantidades significativamente distintas cantidad de referencia.

26.- El método según la reivindicación anterior, donde el cáncer es cáncer de pulmón.

25 27.- Un kit o dispositivo que comprende los elementos necesarios para cuantificar los metabolitos *p*-cresol, el alcohol cumílico, la eicosenamida, el hexadecilindano, el 2,4-bis-dimetilbenzil-6-*t*-butilfenol, la monostearina, el espiro-2,4-heptano-1,5-6-metileno, el 13-heptadecin-1-ol, el estearato de metilo, el glicerol-1-palmitato, el alcohol bencílico y/o el 2,4-difenil-4-metil-2-E-pentano, según se describe en cualquiera de las reivindicaciones 3–27.

30 28.- Uso del kit o dispositivo según la reivindicación anterior, para el diagnóstico, clasificación y seguimiento del cáncer.

29.- El uso del kit o dispositivo según la reivindicación anterior, donde el cáncer es cáncer de pulmón.

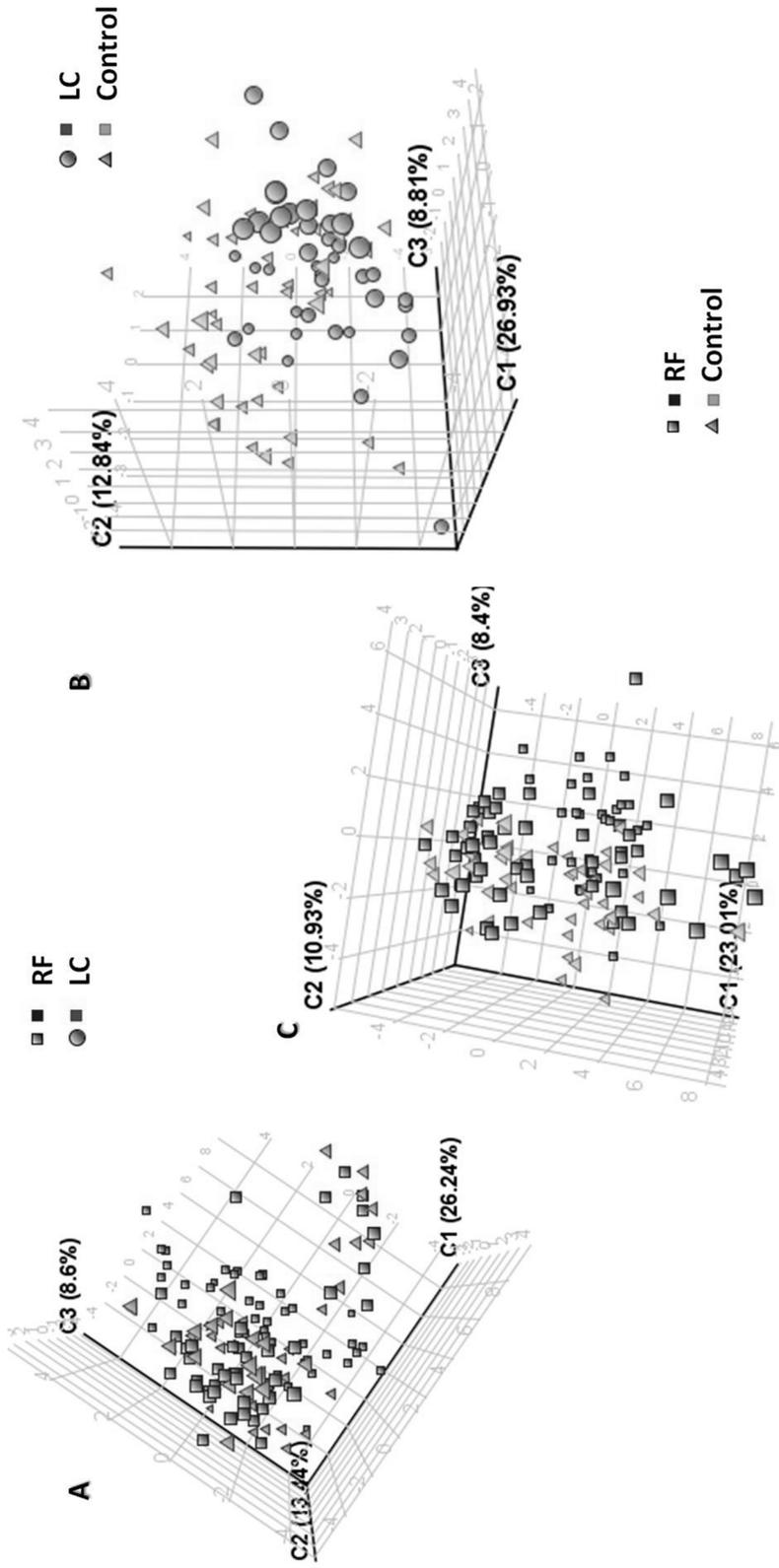


Fig.2

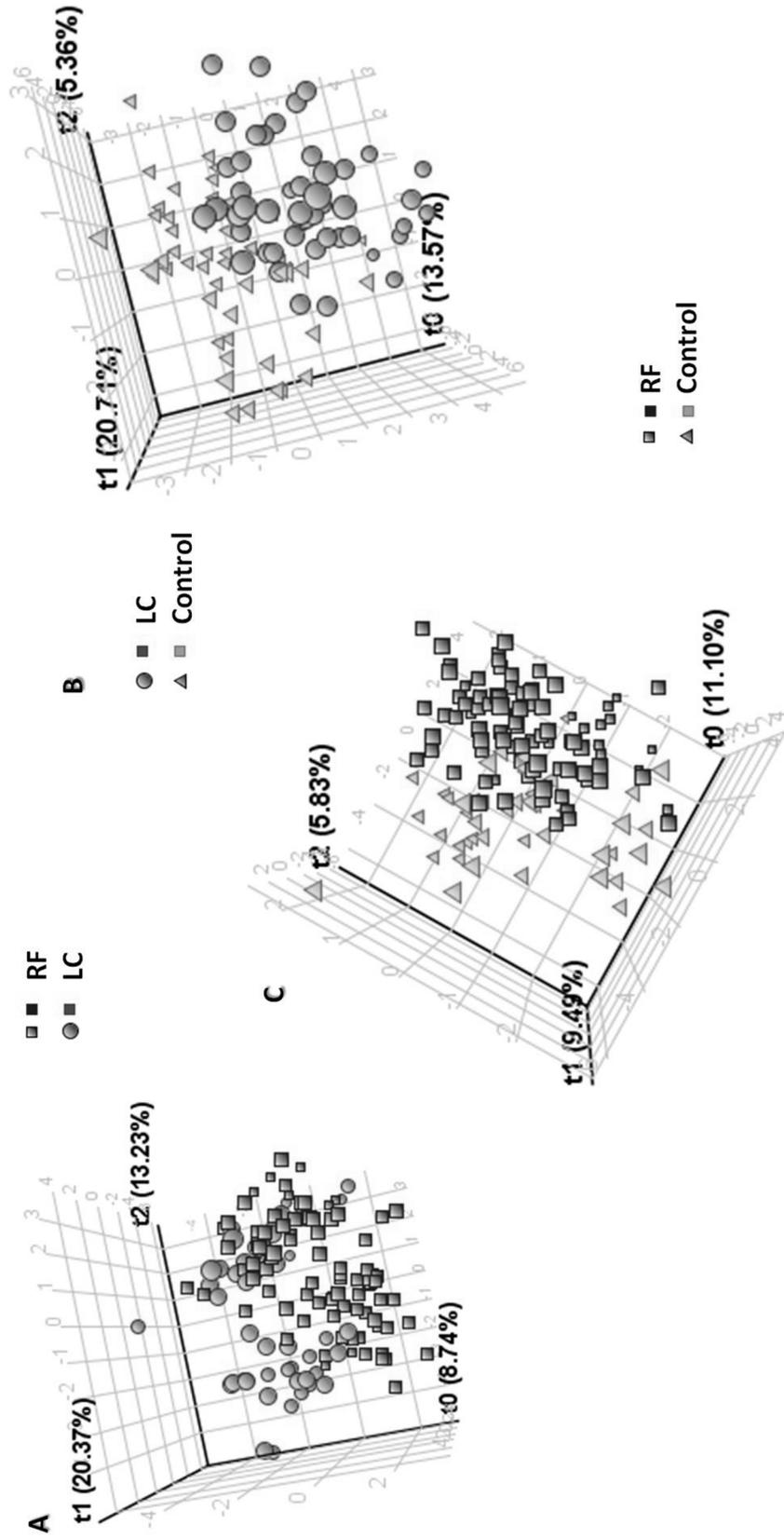


Fig.3

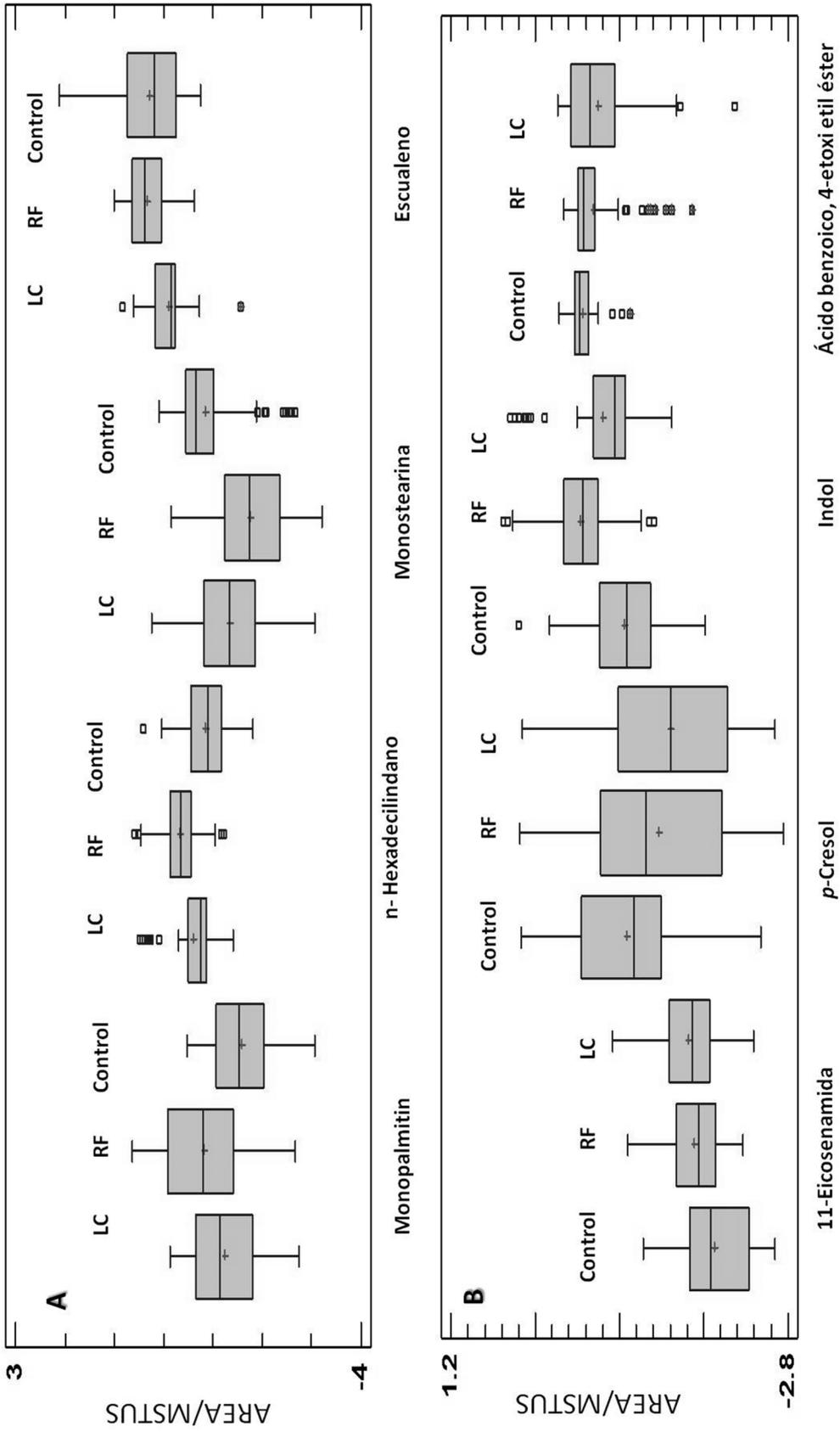


Fig.4

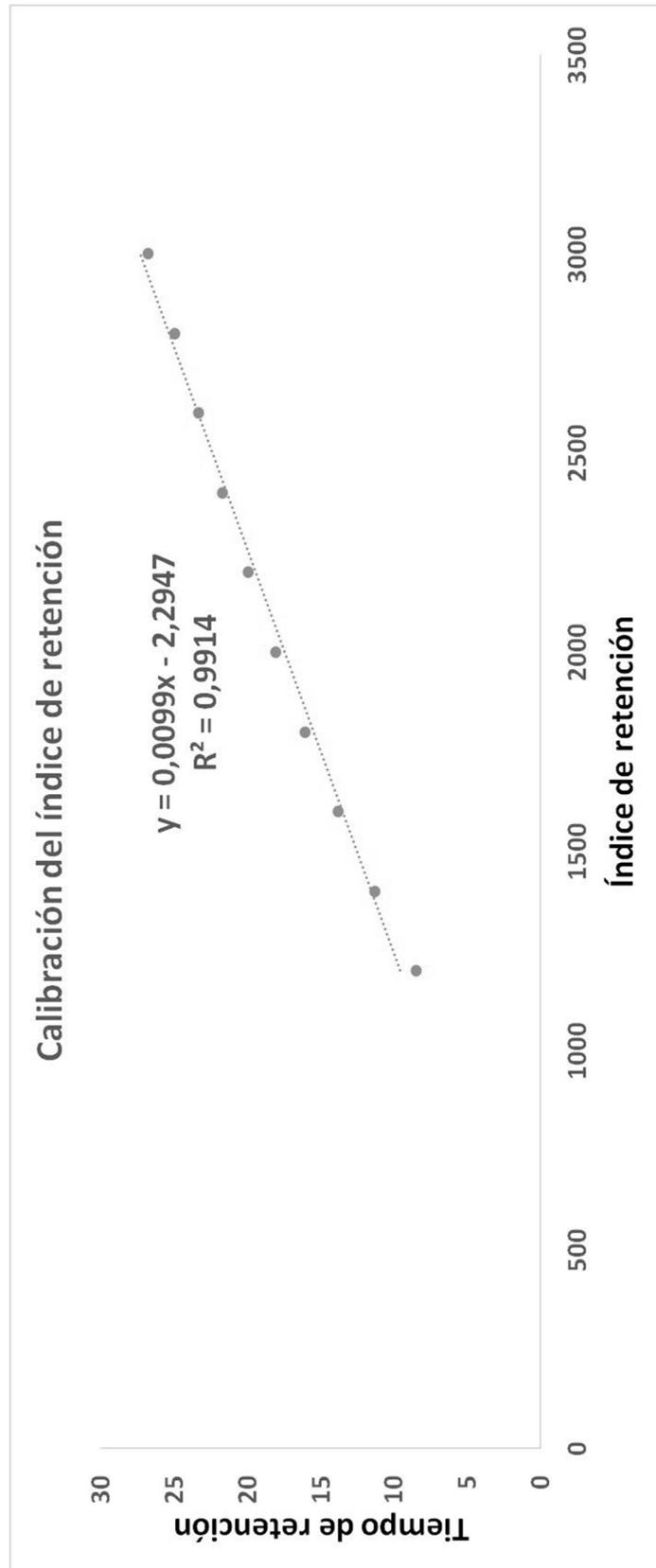


Fig.5.

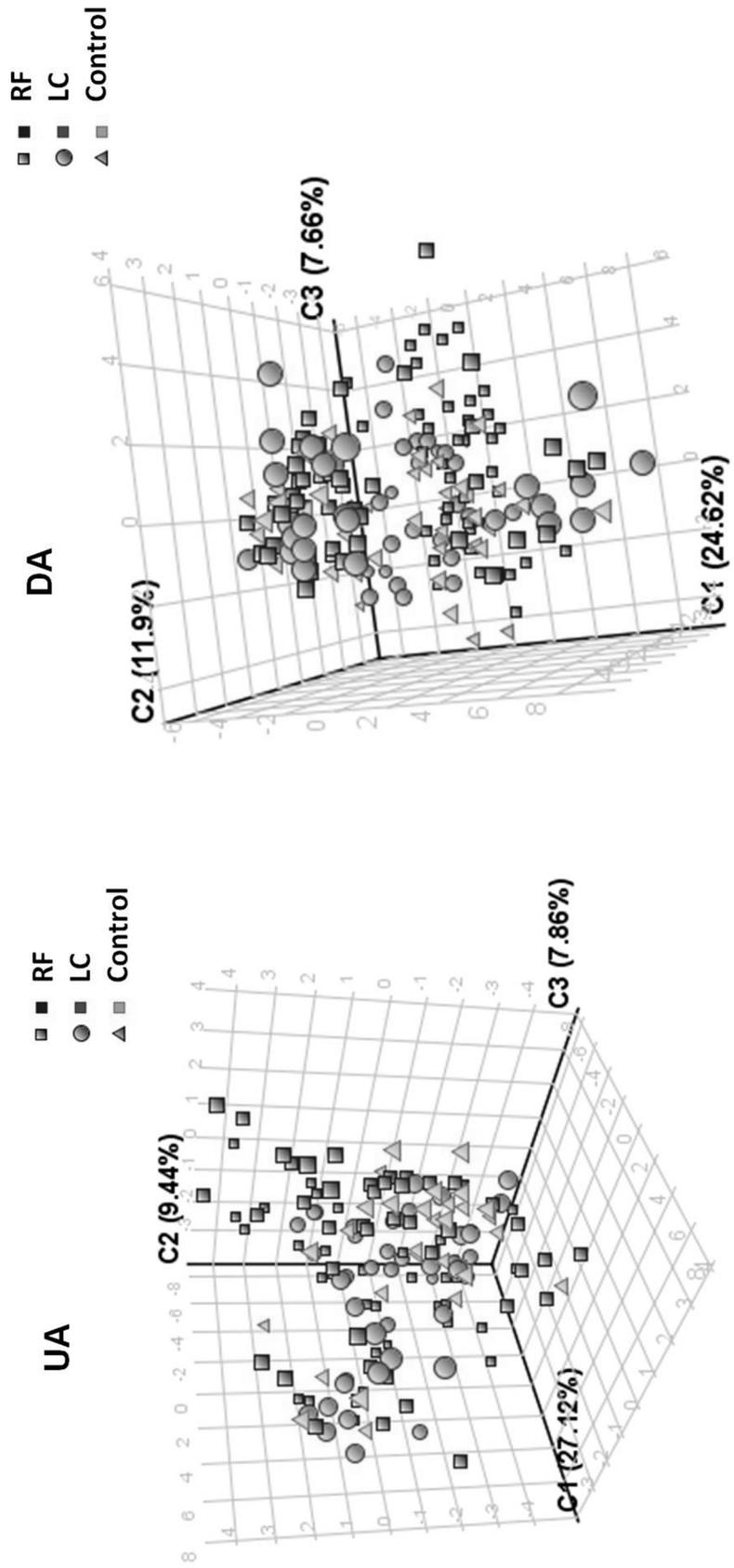


Fig._6.

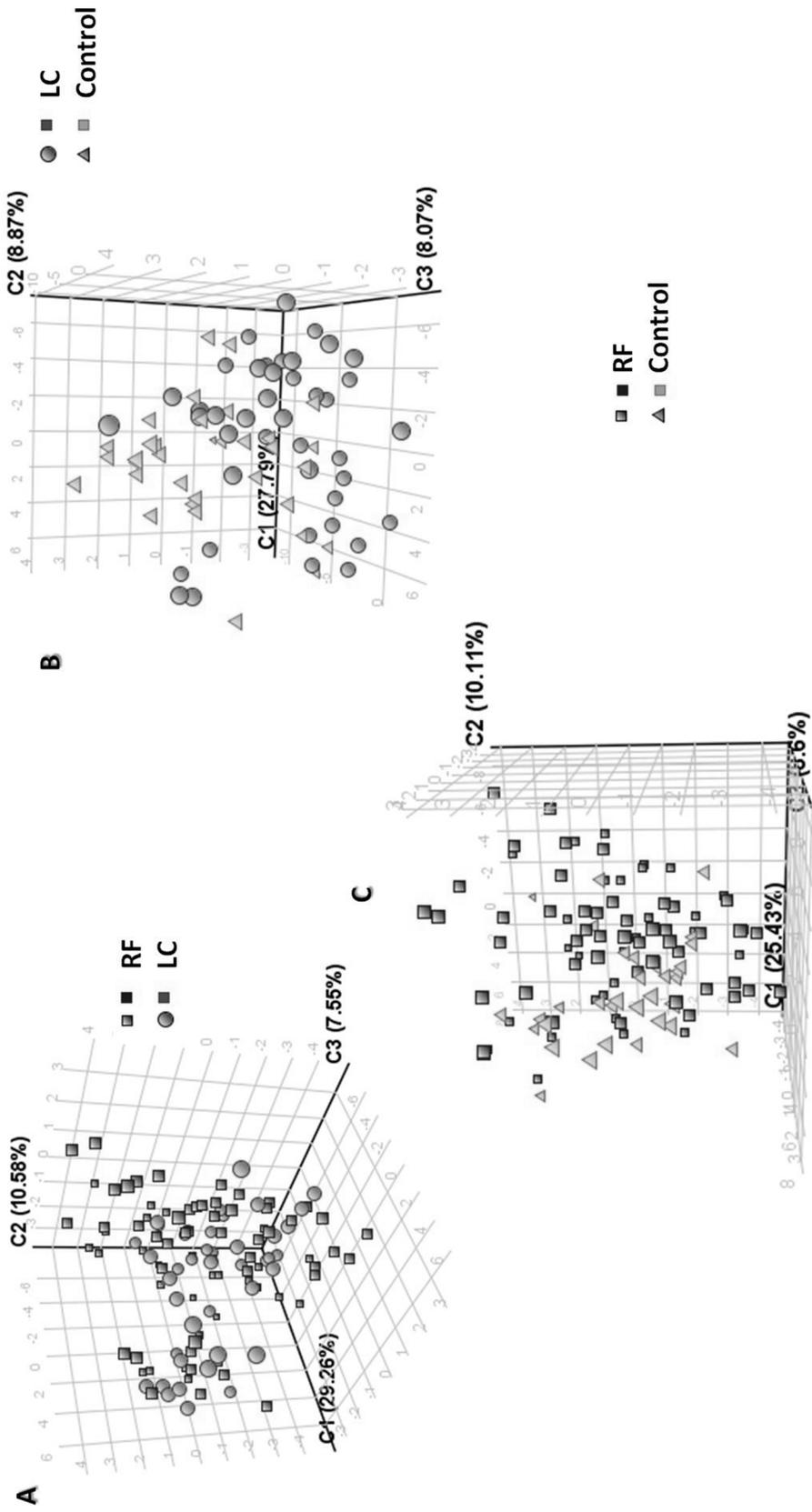


Fig.7.

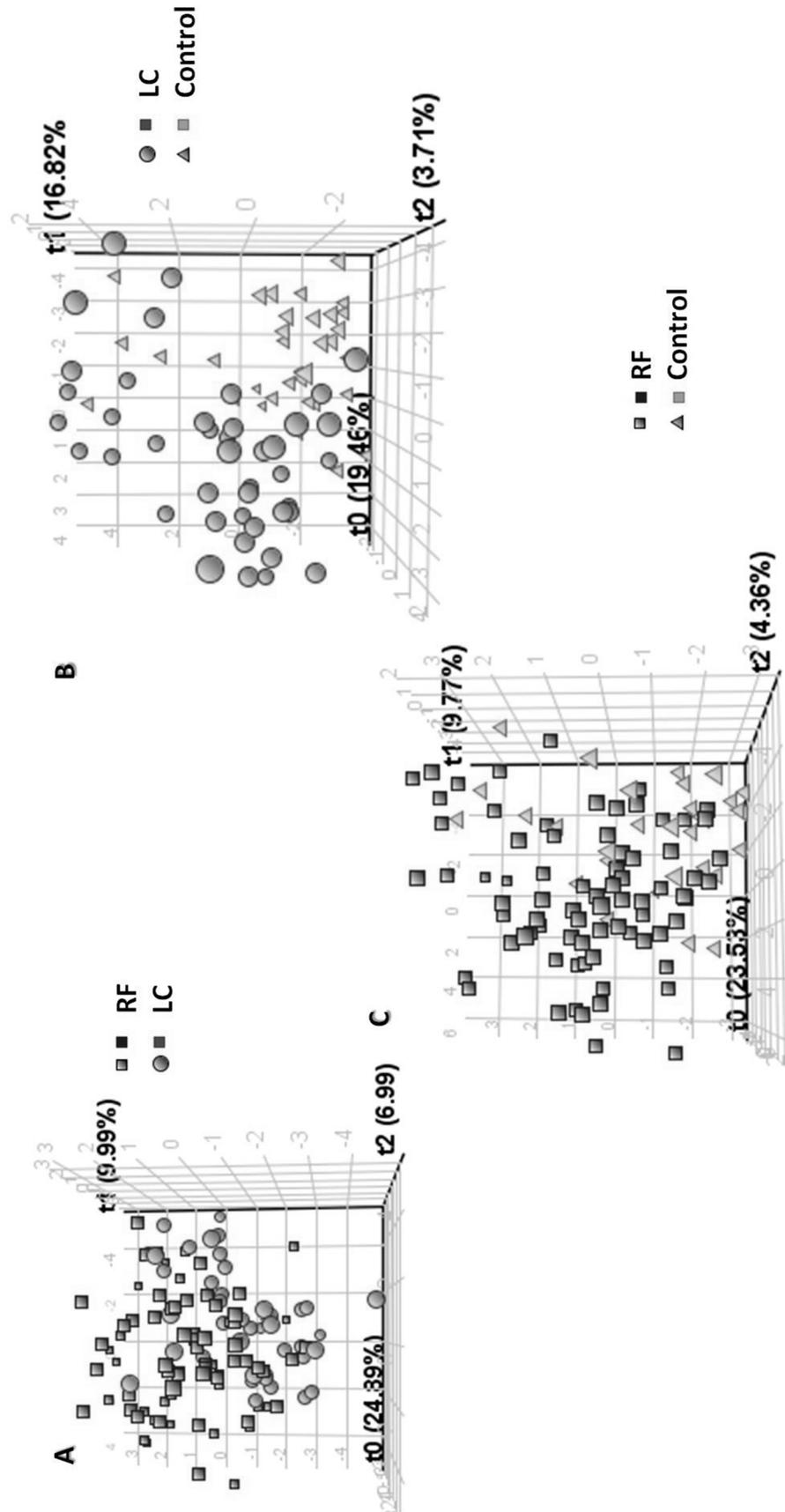


Fig.8.