Universidad de Córdoba
FACULTAD DE CIENCIAS
Departamento de Química Analítica

Avances en metabolómica basados en la preparación de muestra y en plataformas analíticas de cromatografía y espectrometría de masas

Ángela Peralbo Molina

Tesis Doctoral

Córdoba 2016

TITULO: *AVANCES EN METABOLÓMICA BASADOS EN LA PREPARACIÓN DE MUESTRA Y EN PLATAFORMAS ANALÍTICAS DE CROMATOGRAFÍA Y ESPECTROMETRÍA DE MASAS*

AUTOR: *Ángela Peralbo Molina*

© Edita: UCOPress. 2016
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

www.uco.es/publicaciones
publicaciones@uco.es

UNIVERSIDAD DE CÓRDOBA

FACULTAD DE CIENCIAS
DEPARTAMENTO DE QUÍMICA ANALÍTICA

AVANCES EN METABOLÓMICA BASADOS EN LA PREPARACIÓN DE MUESTRA Y EN PLATAFORMAS ANALÍTICAS DE CROMATOGRAFÍA Y ESPECTROMETRÍA DE MASAS

Ángela Peralbo Molina

Córdoba, 2016

AVANCES EN METABOLÓMICA BASADOS EN LA PREPARACIÓN DE MUESTRA Y EN PLATAFORMAS ANALÍTICAS DE CROMATOGRAFÍA Y ESPECTROMETRÍA DE MASAS

Los Directores,

Fdo. María Dolores Luque de Castro
Catedrática
Departamento de Química Analítica
Universidad de Córdoba

Fdo. Feliciano Priego Capote
Profesor Contratado Doctor
Departamento de Química Analítica
Universidad de Córdoba

Trabajo presentado para optar al grado de
Doctora en Ciencias, Sección Químicas

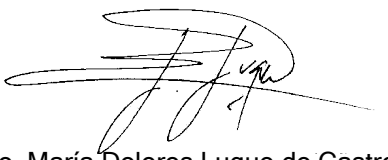
Fdo. Ángela Peralbo Molina
Licenciada en Química

María Dolores Luque de Castro, Catedrática, y **Feliciano Priego Capote**, Profesor Contratado Doctor, ambos del Departamento de Química Analítica, Facultad de Ciencias, Universidad de Córdoba, en calidad de Directores de la Tesis Doctoral presentada por la Licenciada en Química Ángela Peralbo Molina, con el título "Avances en metabolómica basados en la preparación de muestra y en plataformas analíticas de cromatografía y espectrometría de masas",

CERTIFICAN:

Que la citada Tesis Doctoral se ha realizado en los laboratorios del Departamento de Química Analítica, Facultad de Ciencias, Universidad de Córdoba y que, a su juicio, reúne los requisitos necesarios exigidos en este tipo de trabajos.

Y para que conste y surta los efectos pertinentes, expiden el presente certificado en Córdoba, 10 de Junio de 2016.



Fdo. María Dolores Luque de Castro

Fdo. Feliciano Priego Capote

TÍTULO DE LA TESIS: Avances en metabolómica basados en la preparación de muestra y en plataformas analíticas de cromatografía y espectrometría de masas

DOCTORANDA: Ángela Peralbo Molina

INFORME RAZONADO DE LOS DIRECTORES DE LA TESIS

(se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

La doctoranda Ángela Peralbo Molina ha desarrollado a lo largo de 4 años una investigación cuyo interés queda de manifiesto en el número de artículos a los que ha dado lugar; investigación que, con el denominador común de "Metabolómica" ha abarcado las áreas de productos naturales y clínica. La profundización en la documentación existente sobre sus temas de investigación le ha proporcionado también una sólida base en este aspecto. La investigación experimental realizada puso de manifiesto por una parte el enorme interés que productos considerados de desecho, e incluso indeseables, pueden tener un enorme valor añadido; por otra parte, la investigación en el área clínica ha demostrado la capacidad de una muestra poco común como es el aliento condensado para la búsqueda de biomarcadores de cáncer de pulmón.
El interés de la doctoranda en otras áreas la ha llevado a realizar colaboraciones que han dado lugar a los 4 artículos que se incluyen en la Memoria como anexos y a la publicación de una revisión bibliográfica sobre el uso de los ultrasonidos en química analítica.

El número de artículos que conforman el cuerpo de la Tesis son 9 en total, que han dado lugar a otros tantos capítulos de la Memoria:

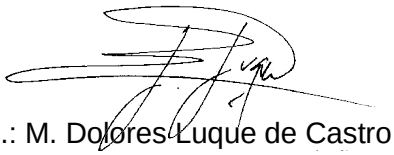
Una revisión bibliográfica sobre el potencial de los residuos del área agroalimentaria de la cuenca mediterránea: Olivo/aceite y vid/vino.
Un capítulo de libro sobre la composición fenólica de las uvas y sus potenciales efectos en biomedicina.
Dos artículos en los que se ha profundizado en el potencial de los desechos de las alcoholeras como materias primas para la obtención de productos de alto valor añadido.
Un artículo sobre la identificación de la gran variedad de compuestos fenólicos que forman parte del alperujo; lo que ha puesto de manifiesto la riqueza de este desecho, tan abundante en nuestra región y tan mal explotado.
Cuatro artículos sobre el potencial del aliento condensado para la búsqueda de enfermedades respiratorias, en los que,
-Se ha desarrollado un método para el análisis de aliento basado en GC-QTOF MS/MS con un estudio exhaustivo de la preparación de la muestra.
-Se han identificado 51 compuestos del aliento condensado, mayoritariamente endógenos, pero también algún compuesto exógeno de interés para conocer las características del individuo.
-Se han desarrollado paneles de compuestos para discriminar entre individuos sanos y pacientes de cáncer de pulmón.
-Se han desarrollado paneles de compuestos para discriminar entre individuos con cáncer de pulmón, aquéllos con riesgo de padecerlo e individuos sanos.

Por ello consideramos que la investigación desarrollada y recogida en la Memoria reúne los requisitos de originalidad, innovación y calidad y autorizamos la presentación de la Tesis Doctoral de Ángela Peralbo Molina.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 8 de junio de 2016

Firma de los directores



Fdo.: M. Dolores Luque de Castro

Fdo.: Feliciano Priego Capote

Agradecimientos

Una vez finalizada la tesis doctoral, me enfrento al capítulo más complicado de este trabajo, resumir en unas pocas líneas mi agradecimiento a todas las personas que han contribuido a la elaboración de este proyecto.

En primer lugar, y de forma muy especial, quiero transmitir mi más sincero agradecimiento a mis directores de tesis por su dedicación, su apoyo y su esfuerzo. A la Catedrática María Dolores Luque, por invitarme a realizar esta tesis doctoral bajo su dirección confiando siempre en mi trabajo y capacidad, nunca podré corresponder como merece por tantos años de conocimiento y experiencia empleados en mi formación. Al Doctor Feliciano Priego, por ser tan brillante y tan agradable, por su excelente disposición y paciencia, por sus siempre buenos consejos, por sus increíbles ideas. Ambos con vuestra experiencia y dedicación, además de contribuir a mi formación, me habéis transmitido esa ilusión por el mundo de la investigación. Gracias, de corazón, por ser unos verdaderos maestros.

En segundo lugar, quiero agradecer a mis compañeros de grupo, no solo su ayuda profesional sino también su inestimable apoyo y amistad. Este camino tan largo y duro se anda mejor con una compañía como la vuestra. A Chema, por sus bromas y las innumerables comidas con charlas reconfortantes. A Mara, Sole, Bea y Miguel por los buenos ratos que hemos pasado. A José, nuestro "veci" por amenizarnos el ratito del almuerzo. A Carlos, por nuestros desayunos en agosto. A Auxi, por sus increíbles masajes. A Pili, por nuestras charlas enriquecedoras y por seguir estando ahí animándome en la recta final. A Vero, por su apoyo profesional y personal, por escucharme tantas noches de invierno y ayudarme a sacar la parte buena de lo malo. A Antonio, por ser grande por fuera y más por dentro y por presentarnos a Mayte, una maravillosa persona que formará parte de mi familia (guiño, guiño). A María por tan buenos ratos que me ha hecho pasar. A Chuck, por dejarnos ser su familia en su estancia en España, ojalá vuelvas pronto. A Maria del Mar, por ser un ejemplo de superación y una gran persona. A Asun, por nuestros ratos con "el gases" y sus consejos de moda. A Carolina, por ser tan

adorablemente friki y por aportarme tanto en tan poco tiempo. A Fabi, por ser una más durante su estancia en España. Por último, quiero expresar mi agradecimiento especial a la Dra. Mónica Calderón ("Moni" para los amigos), quien ha influido en mi formación desde el principio siendo "mi primera jefa" y supervisando además la última etapa de este largo camino. No hay otra compañera tan dispuesta a compartir conocimientos y ayudar como ella. Además de gran profesional es una gran persona. Gracias a todos, de corazón.

También quiero dar las gracias a otros miembros fugaces y del departamento con quienes he compartido innumerables ratitos: Óscar, Joana, Zaca, Mabel, Ilker, Mivan, Fredi, Sara, Antonia, M. José, Isa, María… No puede faltar mi agradecimiento a Diego, Juani y Rafa, que han estado siempre disponibles cuando he necesitado algo.

A mi familia, especialmente mis padres, dos pilares fundamentales en mi vida, por su apoyo incondicional, porque soy lo que soy gracias a ellos, por vuestro sacrificio, por haberme enseñado a seguir adelante y a no rendirme nunca. Os lo debo todo a vosotros. GRACIAS. A mi hermana Laura y Marcos (casi mi hermano), por estar siempre dispuestos a ayudarme, por haber cuidado tanto de mí. A Natalia y Ainhoa, mis tesoros, por ser mi alegría y la mejor ayuda para evadirme del mundo. A mi familia política, por sus muestras de apoyo y cariño.

A Kiko, por creer en mí, por su apoyo y comprensión, por poner música y sol a mis días nublados.

"*And should I fall behind, wait for me, I'll wait for you*".

ÍNDICE

OBJETIVOS

El **objetivo global** de una tesis, que es investigar aspectos de un área de conocimiento que aporten información para un mejor soporte, o para un nuevo enfoque en determinada materia, se ha dividido en la presente Tesis en dos **objetivos genéricos** como consecuencia de que se han abarcado las dos áreas que constituyen las líneas de investigación del grupo en el que se integra la doctoranda, cuyo nombre en el PAIDI es "Plataformas analíticas en metabolómica: áreas clínica y agroalimentaria". Por tanto, y como resultado del interés de la doctoranda en ambas áreas y de su capacidad para abarcarlas, se ha realizado investigación sobre la caracterización de productos naturales cuyo conocimiento lleva a su mejor aprovechamiento, y sobre el uso de una muestra escasamente utilizada en el área clínica (el aliento condensado) para la búsqueda de compuestos con potencial marcador en el diagnóstico de cáncer de pulmón; lo que ha dado lugar a las dos partes, A y B, en las que se ha dividido la Memoria que se presenta. El nexo de unión entre ambas líneas es la metabolómica, como disciplina clave en los dos casos y, concretamente, el proceso analítico utilizado en esta disciplina, ya que los avances planteados se han centrado en dos etapas claves para asegurar la calidad de los resultados generados en un estudio metabolómico: la preparación de la muestra y la adquisición y tratamiento de datos.

La investigación recogida en la Parte A ha sido el resultado de los siguientes **objetivos concretos**:

-Conocer la situación actual descrita en la bibliografía de la explotación, real y potencial, de los residuos más destacables de la agricultura y la industria agroalimentaria mediterráneas (vid/vino y olivo/aceite); así como la consideración de la producción, composición fenólica y aplicaciones biomédicas de la uva, que han dado lugar a una revisión bibliográfica y a un capítulo de libro, que constituyen los Capítulos I y II, respectivamente, de esta Tesis Doctoral.

-Aportar nuevos conocimientos sobre la forma de obtención de compuestos de interés existentes en residuos de las industrias agroalimentarias más importantes de la comunidad andaluza (el vino y el aceite de oliva) con el fin último de un potencial aprovechamiento de estos materiales de escaso o nulo valor. En este caso las materias objeto de estudio han sido un residuo de residuo (el resultante de la obtención de etanol de los desechos de la industria vinícola) y el alperujo (desecho de la producción de aceite de oliva por el sistema de dos fases). En todos los casos se han puesto a punto métodos de extracción de los compuestos de interés (mayoritariamente compuestos fenólicos). Se han utilizado los hollejos y pepitas desechados tras la obtención de etanol para comprobar su capacidad antioxidante y caracterizar su composición; lo que ha dado lugar a los Capítulos III y IV, respectivamente, de la Memoria. En el caso del alperujo la investigación estuvo orientada a los componentes minoritarios del AOV, que confieren a este aceite propiedades nutricionales, identificándolos, estableciendo sus patrones de fragmentación mediante espectrometría de masas en tándem y dando lugar a los resultados que se recogen en el Capítulo V.

-En la investigación que constituye esta Parte A se utilizaron varias técnicas de extracción sólido-líquido asistidas por energías auxiliares, mientras que en relación a la determinación la herramienta analítica principal y necesaria para la consecución de los resultados pretendidos fue un equipo constituido por un cromatógrafo de líquidos y un detector de masas de alta resolución.

La investigación que ha dado lugar a la Parte B de esta Memoria ha sido el resultado de los siguientes **objetivos concretos**:

-Proporcionar un método para la determinación de los componentes del aliento condensado, optimizando tanto la preparación de la muestra con mínimo consumo de este escaso biofluido, la identificación de sus componentes, así como la evaluación de diferentes alternativas de normalización de los datos obtenidos para minimizar la influencia de fuentes de variabilidad externas, tales

como el volumen de muestra proporcionado por paciente. Todo ello ha dado lugar al Capítulo VI de esta Memoria de Tesis.

-Discriminar en función de la composición de aliento exhalado entre pacientes con cáncer de pulmón, individuos con factor de riesgo e individuos sanos. Para el desarrollo de este estudio se utilizó de forma independiente el aliento exhalado procedente de las vías respiratorias alta y baja. Este estudio junto a la búsqueda de una interpretación biológica de los principales cambios detectados en la composición de aliento exhalado de los diferentes grupos estudiados se recoge como Capítulo VII de la Memoria.

-Configurar paneles de compuestos potencialmente marcadores para el diagnóstico de cáncer de pulmón basados en la discriminación de pacientes afectados por esta enfermedad e individuos con factor de riesgo, esencialmente fumadores activos. La evaluación estadística de estos paneles permitirá seleccionar los paneles de mayor potencial en función de si el objetivo de los mismos es priorizar la sensibilidad, la especificidad o la combinación de ambos parámetros. Esta investigación constituye el Capítulo VIII de la Memoria.

-Identificar diferencias cualitativas y cuantitativas entre individuos fumadores, exfumadores y no fumadores en aliento condensado con el fin estudiar el origen de esos cambios: compuestos endógenos, exógenos o compuestos generados por el microbioma existente en las vías respiratorias. La discriminación entre los tres grupos de invididuos y la interpretación biológica de los resultados generados se describe en el Capítulo VIII de la Memoria.

-La herramientas analíticas clave para la obtención de los resultados de esta Parte B de la Tesis han sido un equipo especialmente adquirido para estudiar de la forma más adecuada en la actualidad el tipo de muestra en cuestión: un cromatógrafo de gases conectado a un espectrómetro de masas de tiempo de vuelo (GC-TOF/MS) teniendo en cuenta el carácter volátil o parcialmente volátil de los compuestos presentes en el aliento exhalado y la combinación de software y herramientas bioinformáticas para el tratamiento

multivariante de los datos que permitan identificar compuestos potencialmente marcadores de un proceso biológico.

La formación de la futura doctora, que constituye el *objetivo último* de toda tesis doctoral, ha incluido la realización del máster en "Química Fina" con el número de créditos correspondientes, y en paralelo con éste y con la investigación recogida en la parte principal de la Memoria, se ha pretendido una formación más amplia de la doctoranda mediante la realización de otras actividades que se recogen como anexos, tales como:

- Investigación simultánea con la tesis, también relacionada en cierto modo con la tesis, realizada en colaboración con el Departamento de Bioquímica y Biología Molecular, que ha dado lugar a la publicación de un artículo; mientras que la colaboración con otros miembros del Grupo ha resultado en la publicación de tres artículos además de la redacción de uno más que se ha enviado para su publicación a la revista Analytical Methods. Estos trabajos constituyen el Anexo I.
- Una revisión bibliográfica sobre el papel de los ultrasonidos en la aceleración de las reacciones de derivatización constituye el Anexo II.
- La patente de tres paneles de marcadores para el diagnóstico precoz de cáncer de pulmón mediante el análisis de aliento humano constituye el Anexo III.
- La asistencia a conferencias nacionales e internacionales, con presentaciones en cartel, con un total de 9 comunicaciones, que se recogen como Anexo IV.

INTRODUCCIÓN

La planificación de una Tesis Doctoral implica dos aspectos claves: (i) el desarrollo de nueva investigación de interés en un ámbito dado o de aspectos de ella que no han sido abordados previamente; (ii) la formación del/la doctorando/a, siempre en concordancia con sus aptitudes e inquietudes.

La investigación a desarrollar se encuadra, generalmente, en la línea de investigación del grupo en el que se integra el/la doctorando/a. Cuando el grupo abarca dos líneas de investigación de muy diferente naturaleza puede aparecer un conflicto si las dos son igualmente interesantes, aunque en diferentes ámbitos, si las dos conducen a una formación para el/la doctorando/a que lo/la facultará para su futuro laboral, si las aptitudes e inquietudes del/la doctorando/a son similarmente adecuadas para desarrollar investigación en cualquiera de las dos líneas y si el/la doctorando/a valora su formación por encima del tiempo que tendrá que dedicar para realizar investigación en ambas líneas.

El conflicto surgido en el caso de la presente Tesis se resolvió considerando que la doctoranda tenía capacidad para desarrollar investigación en las dos líneas básicas del grupo, cuyo ámbito queda de manifiesto en el nombre con el que aparece en el Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI), con el código FQM-227: "Plataformas analíticas en metabolómica: áreas clínica y agroalimentaria". Por ello, la Tesis que aquí se presenta abarca investigación que en el plano temporal se dirigió en primer lugar al área de productos naturales y concretamente –siguiendo la pauta de la línea de aprovechamiento de residuos agroalimentarios– a los residuos más abundantes de la región andaluza que resultan del cultivo y las industrias más características de esta región: la vid/vino y el olivo/aceite. La formación adquirida en el desarrollo de esta investigación y la integración del grupo en el Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC) dio lugar a la investigación en el área clínica: búsqueda de compuestos con potencial marcador de cáncer de pulmón utilizando como muestra el aliento condensado.

El nexo de unión de la investigación que constituye esta Tesis Doctoral lo establecen la disciplina ómica en la que se sustenta toda ella (la

metabolómica), la capacidad de la doctoranda para abarcar áreas diferentes, el interés de los directores en conseguir una formación lo más completa posible para sus doctorandos y la disponibilidad de instrumentación analítica adecuada para desarrollar los diferentes aspectos en estudio.

Tras esta necesaria aclaración, la introducción de esta Tesis Doctoral comienza con la descripción del estado actual de la metabolómica para considerar después las dos áreas de investigación abarcadas y poner de manifiesto, de forma separada, los aspectos básicos y el estado actual de cada una de ellas.

1. Metabolómica: definición y características

Del latín "oma", que significa masa en el sentido de masivo, mucho o muchos, el término ómica es un sufijo de nuevo cuño que se utiliza para referirse al estudio de un sistema biológico, entendiendo por sistema un todo o parte funcional de un organismo. Los estudios pueden estar relacionados con el ADN, ARN mensajero (ARNm), proteínas y metabolitos, que dan lugar a ómicas primarias (genómica, transcriptómica, proteómica y metabolómica, respectivamente), o pueden ser familias específicas, tales como lípidos, glicanos, compuestos exógenos o xenobióticos, incluso productos de su interacción, dando lugar a las ómicas secundarias (lipidómica, glicómica, xenometabolómica e interactómica, respectivamente), todas ellas derivadas de las ómicas primarias. La complementariedad de varias ómicas primarias da lugar a la "Biología de Sistemas", término acuñado por Nicholson y Wilson [1] para expresar la integración de la información procedente de estudios realizados mediante dos o más ómicas para la mejor comprensión del funcionamiento de un determinado sistema biológico. Un aspecto fundamental que diferencia la biología de sistemas del dogma central de la biología molecular es que este último está regido por la unidireccionalidad en la expresión de la información desde los genes de una célula u organismo, es decir, el ADN, a su transcripción al ARNm y su traducción

a las proteínas que catalizan las reacciones metabólicas, tal como se muestra en la Figura 1A. La biología de sistemas rompe esta visión tradicional al poner de manifiesto la interacción de los diferentes niveles, el flujo de información que tiene lugar no sólo corriente abajo (down-stream), sino también corriente arriba (up-stream) y que explica que el fenotipo de un individuo sea el reflejo directo de su metabolismo, pero que este último pueda proporcionar información de niveles superiores que justifican la situación de los inferiores, tal como se esquematiza en la Figura 1B.

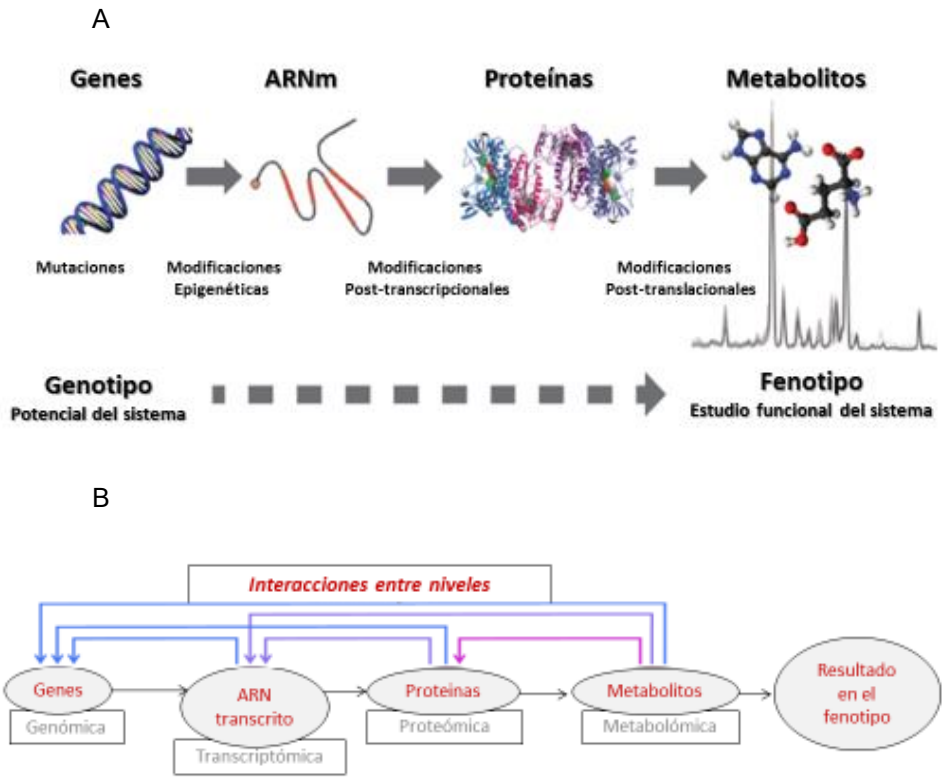


Figura 1. A) Dogma central de la biología molecular. B) Esquema general que justifica la biología de sistemas.

La metabolómica fue definida por Fiehn en 2002 como "un análisis global, exhaustivo en el que se identifican y cuantifican todos los metabolitos de un sistema biológico" [2]. Esta primera definición, bastante pobre al identificarla

con un análisis, ha evolucionado, de forma que Beyoğlu e Idle han propuesto una definición más amplia de la metabolómica como "una disciplina para el estudio global y no sesgado del conjunto de moléculas pequeñas (<1 kDa) en un biofluido, tejido, órgano u organismo [3]".

Al contrario de lo que ocurre con las proteínas o los genes, que para ser activos experimentan modificaciones post-traduccionales o epigenéticas, respectivamente, los metabolitos, que resultan de la actividad celular, sirven como indicadores directos de la actividad bioquímica de la célula en un momento determinado. Así pues, no es de extrañar que el estudio del metaboloma se esté utilizando cada vez más en el diagnóstico clínico ya que el perfil metabólico ofrece una situación única para interrogar el funcionamiento celular y desentrañar los mecanismos bioquímicos implicados para relacionarlos con el fenotipo observado [4].

El metaboloma, por analogía con el genoma, transcriptoma o proteoma, se define como el complemento cuantitativo de los metabolitos presentes en un fluido biológico, célula u organismo en unas condiciones fisiológicas dadas [5]. Estas condiciones pueden abarcar diferentes perturbaciones, tales como variaciones genéticas, estados patológicos o respuestas a estímulos externos. El metaboloma está formado por diferentes partes o fracciones: el metaboloma endógeno, proveniente del metabolismo de las células y tejidos del individuo; el "*food metabolome"* procedente de los metabolitos derivados de la digestión de los compuestos nutricionales esenciales y no esenciales, el metabolismo derivado de la interacción del individuo (o huésped) y su microbiota; el "*drug metabolome"* procedente de compuestos xenobióticos derivados de fármacos, etc., y el metabolismo derivado del ambiente, contaminación, etc. "*pollutant metabolome"* [6]. Por tanto, el metaboloma, y en consecuencia el estado metabólico de un ser vivo, puede verse afectado por diversos factores intrínsecos (edad, estado de salud, estado reproductivo, etc.) y extrínsecos (nutrientes, compuestos artificiales como pesticidas o fertilizantes, en caso de plantas, y fármacos en caso de animales y seres humanos, contaminantes, etc.).

Por ello, el estudio del metaboloma es de gran importancia para conocer el estado del organismo y cómo le afecta su entorno.

Además, los cambios fisiológicos que se producen como consecuencia de la expresión de genes o proteínas se amplifican a través del transcriptoma y del proteoma por lo que la detección a nivel de metaboloma es más sensible. Este hecho explica que las rutas metabólicas se reflejen de forma más precisa en la concentración de metabolitos que en la actividad de las enzimas implicadas o del ARNm) que las codifican [7]. Por último, la tecnología que se requiere en metabolómica es más genérica puesto que un determinado metabolito, a diferencia de un gen o una proteína, es el mismo en todo organismo que lo contenga.

El metabolismo, en definitiva, es una red extensa de reacciones metabólicas donde los productos de una reacción son los reactantes de otra subsiguiente. Por ello, se define metabolito como "cualquier intermedio o producto final del metabolismo, usualmente restringido a pequeñas moléculas que no están genéticamente codificadas". En función tanto de las rutas metabólicas en las que están implicados los metabolitos como de su función se pueden clasificar en metabolitos primarios (aquellos que están directamente implicados en el crecimiento, desarrollo y reproducción del organismo) o secundarios (los que no están directamente implicados en los procesos de crecimiento, desarrollo y reproducción, pero generalmente tienen una función biológica importante como, por ejemplo, la defensa contra predadores, parásitos o enfermedades, competencia entre especies, para facilitar los procesos de reproducción o en mecanismos de señalización celular).

La metabolómica es a día de hoy una disciplina suficientemente madura que posee un estatus en cierto modo similar al de otras grandes ómicas debido a la relevancia del metaboloma en el contexto biológico. La Figura 2 muestra, de forma comparativa, la evolución de las publicaciones sobre las grandes ómicas hasta final de 2015, así como algunas fechas claves en su desarrollo.

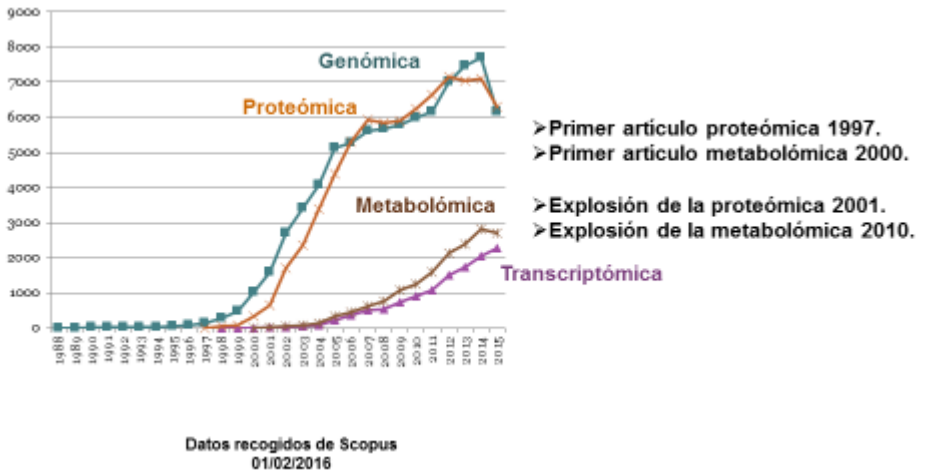


Figura 2. Evolución del número de publicaciones de las diferentes ómicas desde su aparición hasta final de 2015, y algunas fechas claves en la evolución.

La metabolómica presenta importantes ventajas para el estudio de un sistema biológico con respecto al resto de las ómicas. Para llevar a cabo este tipo de estudios se utilizan plataformas analíticas que se caracterizan por tener una excelente reproducibilidad analítica y biológica con un bajo coste del análisis por muestra, en contraposición a los costes de aplicación de cualquier otra de las grandes ómicas.

Desde el punto de vista del número de componentes que abarca cada una de las ómicas, la metabolómica vuelve a presentar ventajas. Según un estudio reciente [8], el número predecible de metabolitos en humanos es de sólo 40.000 (29.000 de ellos endógenos) [9], mientras que el de genes, transcritos y proteínas puede llegar a 25.000, 300.000 y 1.000.000, respectivamente [8]. El menor número de potenciales analitos reduce notablemente la complejidad del análisis. En el caso de la genómica, la inclusión de las modificaciones de genes que estudia la epigenómica aumenta enormemente el número de componentes [10]. Por tanto los estudios en metabolómica conllevan una reducción en la complejidad de los datos, por lo que se simplifica el análisis y aumenta la

probabilidad de detectar cambios significativos que reflejen el problema biológico a estudiar [11].

Sin embargo, la metabolómica presenta ciertas características que la hacen más compleja que otras ómicas en algunos aspectos. Los metabolitos son moléculas muy diversas químicamente, no se encuentran confinados, como en el caso del ADN, y están presentes en el organismo en un amplio rango de concentraciones que puede abarcar hasta 9 órdenes de magnitud (de picomoles a milimoles). La gran variedad de analitos y su amplio rango de concentraciones hacen necesaria la aplicación de técnicas muy diversas tanto para la preparación de la muestra como para el análisis de los metabolitos en cuestión, jugando un papel clave la experiencia del investigador para seleccionar las más adecuadas en función de la naturaleza de los metabolitos a determinar y de la matriz de la muestra.

El estado actual de la metabolómica permite distinguir en esta disciplina características positivas, que pueden incluso mejorarse, y negativas, que deben superarse. Las **características positivas** de la metabolómica que promueven su uso y establecen una diferencia con otras ómicas son las siguientes:

- La existencia de plataformas analíticas robustas y estables para el análisis de un número importante de metabolitos, como resultado de la investigación en esta área durante las dos últimas décadas del siglo XX, previa a la aparición de las ómicas como tales.

- La excelente precisión analítica y biológica de las plataformas existentes.

- El coste de análisis por muestra y analito, que es, en general, bajo en comparación con los análisis en genómica y en transcriptómica. Este hecho facilita la aplicación a un número de muestras grande, permitiendo llevar a cabo estudios extensivos en cohortes con una variabilidad biológica significativa; lo que proporciona una información amplia y bien soportada.

- El tiempo de análisis, mucho más corto que el requerido en genómica, transcriptómica y proteómica, especialmente para el perfil metabolómico, que se obtiene en unos pocos minutos.

- La capacidad para integrar datos obtenidos con diferentes técnicas (metabolómica integrada) aplicada a muestras de diferente procedencia.

Aspectos negativos o no resueltos de la metabolómica son los siguientes:

- La necesidad de plataformas de análisis múltiples, como resultado de la enorme variedad de metabolitos, que oscila desde lípidos no polares a compuestos iónicos; desde moléculas de alto peso molecular tales como colesterol a muy pequeñas tales como iones amonio.

- Las concentraciones a las que aparecen los metabolitos en los organismos abarcan más de nueve órdenes de magnitud (de picomoles a milimoles), y hacen obligatorio el uso tanto de etapas de fraccionamiento, dilución y preconcentración, como de equipos analíticos con diferente rango dinámico.

- La necesidad de métodos estandarizados todavía no establecidos, como consecuencia de ser la última de las ómicas primarias.

- El número de metabolitos que constituyen el metaboloma, que en un sistema vegetal puede alcanzar 200.000, y la diversidad dentro de una misma familia.

2. Tipos de estrategias en metabolómica

En metabolómica se emplean en términos generales cuatro estrategias en función del conocimiento previo del problema biológico y del objetivo del estudio:

• *Análisis dirigido u orientado ("targeted analysis"):* es el que se realiza cuando existe un conocimiento previo de los metabolitos implicados en el proceso biológico en estudio. Este enfoque permite escoger la técnica analítica más adecuada y optimizar la etapa de preparación de la muestra en función de la naturaleza y abundancia de los metabolitos a analizar. Estos estudios se centran en la detección y cuantificación de un número reducido de compuestos relacionados directamente con el problema biológico y proporcionan una medida o valor absoluto de la concentración de los metabolitos con una elevada selectividad y precisión [4,12].

• *Análisis no dirigido o perfil metabólico ("untargeted analysis" o "global metabolomics profiling"):* se emplea cuando no se parte de una hipótesis sobre los mecanismos implicados o metabolitos alterados en un determinado sistema biológico. Con esta aproximación, en contraposición al enfoque anterior, el análisis no se restringe a un grupo seleccionado de compuestos, sino que se pretende analizar simultáneamente tantos metabolitos como sea posible para obtener una visión global lo más completa posible del sistema en cuestión. La detección de esta amplia gama de metabolitos puede hacerse mediante una sola plataforma analítica o una combinación de plataformas complementarias –principalmente basadas en resonancia magnética nuclear (NMR) o espectrometría de masas (MS), junto con cromatografía de gases (GC), cromatografía de líquidos (LC) o electroforesis capilar (CE). En estos estudios generalmente se calculan las concentraciones relativas de los metabolitos analizados y se estudian sus variaciones entre dos o más situaciones del sistema en estudio [4,12].

• *Huella dactilar metabolómica ("metabolomics fingerprinting")* [13]: se emplea como metodología rápida y de alto rendimiento para el análisis de muestras biológicas con un perfil metabolómico concreto para la clasificación y el cribado de muestras.

• *Huella pedicular metabolómica ("metabolomics footprinting")* [14]: tiene por objetivo el estudio de los metabolitos en el fluido extracelular, también

conocido como exometaboloma o secretoma. Al igual que en la estrategia anterior, se pretende conseguir información para la clasificación de grupos de muestras o el cribado, principalmente.

Obviamente, la complejidad de la preparación de la muestra depende de la estrategia [14]. Si el objetivo es un análisis orientado, la preparación de la muestra se focalizará en los metabolitos de interés y se utilizarán etapas de limpieza, aislamiento y preconcentración de los analitos; por el contrario, los métodos para análisis global son no selectivos y con ellos se pretende abarcar el mayor número posible de metabolitos, por lo que en la etapa de preparación de la muestra se persigue que el contenido de la muestra analítica sea representativo de la muestra original. El estado físico de la muestra también influye en el número de etapas requeridas para la preparación de la muestra, que es más laboriosa en general para muestras sólidas. Cada una de estas estrategias tiene sus propias ventajas e inconvenientes, pero pueden ser altamente complementarias cuando se combinan.

Las técnicas de detección empleadas dependen también de la estrategia elegida. Por ejemplo, para obtener la huella dactilar metabolómica se requieren técnicas de detección que permitan el análisis directo y rápido de la muestra. Para estos análisis se usan principalmente la NMR, la MS (dependiendo de la complejidad de la muestra) y, en menor proporción, las espectroscopías infrarroja y Raman.

En el caso de análisis orientados y análisis globales metabolómicos se usan más ampliamente los métodos que implican separación previa a la detección, ya que permiten la detección individual de los metabolitos para su cuantificación o identificación, siendo las técnicas cromatográficas (de gases o de líquidos) las más usadas en este ámbito. El detector más utilizado en metabolómica, tanto para análisis orientado como para el global es el MS, que pueden aportar una gran resolución espectral y, por tanto, una gran exactitud en la medida de la relación masa/carga, y una excelente sensibilidad.

3. Herramientas analíticas en metabolómica

La selección de las herramientas analíticas más adecuadas en cada caso depende del objetivo del estudio, que a su vez delimitará la estrategia analítica a utilizar y, por tanto, el método analítico que se requiere poner a punto:

- Para obtener el perfil metabolómico, la situación más común es un análisis directo de la muestra o bien la implementación de etapas muy simples tales como la dilución o la eliminación de proteínas.

- Cuando se trata de análisis orientado, los protocolos son más selectivos, ya que es necesario obtener la fracción deseada eliminando o minimizando la presencia de especies interferentes, lo que en definitiva requiere un mayor número de etapas en la preparación de la muestra.

- Si se trata del análisis metabolómico global se requieren protocolos genéricos no selectivos con el fin de identificar el mayor número posible de metabolitos.

En la Figura 3 se muestra el diagrama de flujo de un proceso analítico, que comprende la toma de muestra y su preparación para la detección, que proporcionará datos cuyo tratamiento permitirá expresar el resultado en la forma requerida. Como las herramientas para el tratamiento de datos se describen en el siguiente punto de esta Memoria, esta sección se va a dedicar principalmente a la descripción de las herramientas analíticas empleadas en la preparación de la muestra y en la separación de los analitos, que será más o menos exhaustiva dependiendo del tipo de detección.

Figura 3. Esquema general de un proceso analítico.

3.1. El muestreo

Una vez planificado el estudio a realizar, la primera etapa del proceso analítico en general, y de la metabolómica en particular, es la selección del tipo de muestra. El muestreo es de gran importancia, ya que influirá decisivamente en el resultado final. La representatividad es la característica fundamental de un buen muestreo, que implica la selección del material biológico más adecuado para el objetivo del análisis y que debe perseguir la obtención de muestras representativas cuya conservación hasta la siguiente etapa ha de asegurarse.

Realizado el muestreo, una primera etapa de pretratamiento en el caso de muestras líquidas es la eliminación de partículas en suspensión o la precipitación de proteínas, si se trata de muestras clínicas. En el caso de muestras de origen vegetal es frecuente realizar un secado y pulverización o bien su trituración con el fin de obtener un tamaño de partícula homogéneo. Estas operaciones facilitarán la posterior extracción de los metabolitos con el extractante adecuado.

La etapa de interrupción del metabolismo es clave para el estudio de muestras con gran actividad enzimática, ya que su objetivo es evitar las alteraciones producidas tras el muestreo hasta el momento del análisis. Esta etapa debe ser muy rápida, prácticamente instantánea, compatible con las siguientes etapas del proceso analítico, y no debe producir cambios físicos o químicos en los metabolitos. Estos requerimientos hacen que las alternativas más empleadas en esta etapa sean la disminución de la temperatura hasta un intervalo entre -20 y -80 ºC o el uso de valores extremos de pH, tanto ácidos como básicos con el fin de bloquear la actividad enzimática que se desarrolla en un intervalo de pH intermedio.

Las muestras más comunes utilizadas en metabolómica son los biofluidos, las células y tejidos animales y las plantas, que requieren protocolos de muestreo y conservación muy diferentes. Entre los biofluidos, los dos más

comunes que pueden obtenerse por medios poco o no invasivos son la sangre y la orina. Son muestras idóneas para la prognosis/diagnosis de enfermedades y para el seguimiento de dietas y terapias con fármacos. Otros biofluidos utilizados en menor proporción son la saliva, el sudor, el aliento, la leche, la bilis y los líquidos amniótico, cerebroespinal, seminal, digestivo, sinovial, etc. Cuando se trabaja con células o tejidos, el objetivo en el primer caso puede ser el estudio del endometaboloma, el exometaboloma o ambos, mientras que en los tejidos el objetivo es comúnmente conocer los metabolitos formados en un comportamiento anormal del tejido en cuestión.

Las plantas ofrecen muy variados tipos de muestreo según se trate de hojas, raíces, frutos, ramas, flores o productos derivados (ej. aceites, resinas, vino). Existen numerosas revisiones bibliográficas recientes, en las que se discuten resultados contradictorios obtenidos por diversos autores, el tiempo máximo que debe transcurrir desde el muestreo hasta el almacenamiento (generalmente ultracongelación), así como la temperatura de este último y el tiempo de almacenamiento máximo (periodo en el cual la muestra se mantiene inalterada) [15,16].

3.2. La preparación de la muestra

En esta etapa del proceso analítico se pueden utilizar diferentes metodologías teniendo en cuenta que puede requerir varias operaciones, como la lixiviación (si la muestra es sólida), la extracción líquido-líquido (LLE) o líquido-sólido (si la muestra es líquida, o tras el proceso de lixiviación para el tratamiento del lixiviado) de los metabolitos, la preconcentración, la limpieza y, en algunos casos, la derivatización. La extracción de los metabolitos es una etapa clave en el proceso analítico en metabolómica y su eficacia afecta a la calidad de los resultados. La selectividad requerida en la etapa de extracción dependerá del objetivo del estudio. Por lo tanto, mientras que en el análisis orientado se buscan estrategias de separación selectivas que proporcionen muestras analíticas

limpias y concentradas, en el análisis no dirigido la extracción es principalmente no selectiva, considerando sólo las sales y las macromoléculas como potenciales interferentes que deben eliminarse en esta etapa. El protocolo de extracción está condicionado principalmente por la muestra biológica; así, en el caso de muestras sólidas, la lixiviación se realiza mediante técnicas clásicas (agitación del sistema sólido-líquido), o técnicas que permitan la aceleración y/o la automatización de esta etapa (uso de energías auxiliares como microondas, ultrasonidos o temperatura y presión altas) [17,18]. Por otra parte, los metabolitos de las muestras líquidas se separan de la matriz de la muestra principalmente mediante LLE, extracción en fase sólida (SPE) o microextracción en fase sólida (SPME). La necesidad de operaciones adicionales de preparación de la muestra la establece la plataforma utilizada para la etapa de análisis. A modo de ejemplo, la baja volatilidad de muchos metabolitos hace necesaria una etapa de derivatización –generalmente sililación– previa a la introducción de la muestra analítica en un equipo GC-MS [19].

La introducción de una etapa cromatográfica previa a la detección es de gran importancia cuando se estudian metabolitos concretos (análisis orientado) o el perfil global. La cromatografía es una técnica de separación donde los componentes de la muestra se distribuyen entre una fase móvil (el disolvente en LC y un gas inerte en GC) y una fase estacionaria (el relleno de la columna). La separación se produce en función de las distintas afinidades de los constituyentes de la muestra hacia cada fase.

La GC permite la separación de compuestos volátiles, como, por ejemplo, compuestos aromáticos. Para la separación de compuestos menos volátiles es necesaria una reacción de derivatización de los metabolitos con la que se aumenta su volatilidad.

En LC el componente más determinante en la separación es la columna, cuyas características definen el tipo y modo de la cromatografía. Las columnas de fase reversa, principalmente las empaquetadas con sílice (C18 o C8), se caracterizan por una fuerte interacción con compuestos de baja polaridad,

mientras que las columnas de fase normal, como las HILIC, son más eficaces para compuestos polares.

El diámetro interno de la columna y los caudales utilizados dan lugar a la clasificación que se muestra en la Tabla 1. Generalmente se considera HPLC (del inglés, high performance liquid chromatography) cuando la separación cromatográfica se lleva a cabo en columnas de diámetro interno entre 1.5 y 4.5 mm, para columnas con diámetros alrededor de 800 µm se define como microLC, cuando se usan columnas de diámetro entre 100 y 500 µm se conoce como LC capilar y en caso de columnas de entre 10 y 100 µm se denomina nanoLC.

Tabla 1: Principales diferencias entre los distintos tipos de LC

	UPLC/RRLC	HPLC	MicroLC	LC capilar	NanoLC
Diámetro interno de la columna (mm)	1.5-4.5	1.5-4.5	0.8	0.18-0.32	0.075-0.1
Longitud de la columna (cm)	3-15	3-30	5-25	5-25	5-15
Tamaño de partícula (µm)	<2	3-40	3-5	3-5	3-5
Flujo de fase móvil	0.2-5 mL/min	0.2-2.5 mL/min	10-100 µL/min	1-10 µL/min	0.1-1 µL/min

La nanoLC ha ganado mucho protagonismo en los últimos años debido a las ventajas que presenta frente a la HPLC convencional ya que, al utilizar columnas capilares de diámetro interno muy pequeño, disminuye considerablemente el consumo de fases móviles (los flujos oscilan entre 100 y 500 nL/min). Se consigue así disminuir los volúmenes de inyección de muestra, con la consiguiente reducción de costes y los beneficios medioambientales derivados. Además, desde un punto de vista teórico, se ha demostrado que una reducción en el diámetro interno de la columna aumenta la sensibilidad debido a

que se produce una menor dilución del analito [20] y se consigue una mayor eficacia [21].

También se puede distinguir otra categoría en LC en función del tamaño de partícula del material empaquetado en la columna, de forma que se denomina comúnmente cromatografía líquida de ultra alta resolución (UPLC) cuando se emplean columnas con tamaños de partícula inferiores a 2 µm. En estos casos, normalmente la longitud de la columna es menor y se usan flujos relativamente elevados. Este tipo de LC permite realizar análisis más rápidos manteniendo la eficacia, o bien conseguir eficacias mayores para el mismo tipo de análisis. La combinación del uso de pequeños diámetros de partícula y presiones de hasta 1250 bares proporciona una eficacia de hasta 100000 platos teóricos frente a los 15000-25000 platos teóricos de la HPLC convencional [22].

3.3. Herramientas analíticas empleadas en la etapa de detección

Debido al gran número de moléculas que constituyen el metaboloma, sus muy diferentes abundancias y la gran variedad de estructuras químicas, es imposible cubrir todo el metaboloma utilizando una sola plataforma analítica. La selección de la técnica de detección más adecuada para cada situación depende del propósito final del estudio. Las espectrometrías infrarroja y Raman tienen un pequeño coste por análisis y tiempos de toma y tratamiento de datos cortos, pero la información estructural y la sensibilidad que proporcionan son limitadas. Las técnicas más recomendadas y utilizadas en metabolómica son la MS y la NMR.

La NMR (basada en las diferencias de contraste que producen los espacios vecinos a los núcleos atómicos ante campos magnéticos muy intensos) es una técnica no destructiva, por tanto, muy útil para muestras valiosas, con una preparación de muestra simple y rápida (normalmente sólo requiere dilución). La MS (basada en la ionización, posterior fragmentación y detección de los iones formados por los analitos de interés) presenta en análisis sin separación previa una ventaja respecto a otras técnicas: un espectrómetro de

masas se comporta, en cierto modo, como un equipo de separación ya que permite separar las señales en función de la relación masa/carga (*m/z*) de los compuestos.

Las principales limitaciones de la NMR comparada con la MS son la menor sensibilidad y menor resolución cuando se analizan muestras con una matriz compleja como es el caso de biofluidos con una alta concentración proteica. Las ventajas de la MS en comparación con la RMN se reflejan en el número de publicaciones anuales, donde los estudios en los que se emplea la MS superan a los basados en RMN. Ambas técnicas son adecuadas para el análisis metabolómico, cada una con sus fortalezas y debilidades, por lo que pueden considerarse complementarias (Tabla 2).

La MS por lo general va combinada con técnicas de separación analítica como la GC o la LC (que más recientemente también se han incorporado a la NMR). La combinación de ambas proporciona un mayor nivel de información puesto que hay metabolitos que sólo pueden detectarse mediante LC-MS o GC-MS debido a que son volátiles, no se ionizan adecuadamente a presión atmosférica o son termolábiles [23].

3.4. La espectrometría de masas en metabolómica

La MS es una técnica de detección basada en el desplazamiento diferencial de moléculas ionizadas (o átomos ionizados) por aplicación de un campo eléctrico en condiciones de vacío. De forma simplista, un espectrómetro de masas consiste en una fuente de iones, un analizador de masas, un detector y un sistema de recogida de datos. Las moléculas de la muestra analítica se insertan en la fuente de iones, donde se ionizan. Los iones así formados se encuentran en fase gaseosa y se separan entre sí en función de su relación masa/carga *(m/z)* en el analizador de masas, donde se detectan.

Tabla 2: Ventajas y limitaciones de la RMN y la MS para análisis metabolómico.

	RMN	EM
Sensibilidad	Baja sensibilidad, del orden de micromolar (10^{-6})	Elevada sensibilidad, del orden de picomolar (10^{-12}) con técnicas estándar
Reproducibilidad	Muy alta	Moderada
Preparación de muestras	Mínima: añadir tampón, D_2O y compuesto de referencia si se requiere	Necesaria la extracción o separación de los metabolitos mediante cromatografía
Volumen muestra	Normalmente 300-500 µL. Pueden analizarse 50 µL empleando capilares	Inferior a 100 µL
Recuperación de muestra	No se destruye la mestra, puede ser recuperada, almacenada y reanalizada posteriormente	Técnica destructiva, no es posible recuperar la muestra
Análisis de tejido	Si. Tejido intacto utilizando una sonda HR-MAS	No. Se requiere extracción previa de los metabolitos
Tiempo adquisición experimentos	5 min. Experimentos estándar 1D (^{1}H-RMN)	10 min. UPLC-MS
Tipos de metabolitos detectados	Todos los metabolitos que posean un hidrógeno (^{1}H) en su estructura, asumiendo concentración del orden 10^6	Mayor cantidad y variedad de metabolitos. Se requiere extracción y separación cromatográfica específica en función del tipo de metabolito a identificar
Identificación de metabolitos	Se combina la información de los experimentos 1D y 2D y las bases de datos. La identificación es más directa que con EM	Una vez conocida la masa y el tiempo de retención, se requiere un análisis adicional para conocer el patrón de fragmentación del compuesto

Existen diferentes técnicas de ionización para generar los iones dando lugar a diferentes fuentes de ionización. La ionización por impacto electrónico (EI) e ionización química (CI) son las que se utilizan generalmente en el acoplamiento GC-MS, mientras que la ionización por electroespray (ESI) y la ionización química a presión atmosférica (APCI) son las alternativas más comunes en LC-MS [24]. Para el objetivo de la investigación que se recoge en esta Memoria, los analizadores de ionización por electroespray se han utilizado en el acoplamiento de LC con MS, y la ionización por impacto electrónico en el acoplamiento de GC con MS. Ambas formas de ionización son las más utilizadas en los dos acoplamientos y acaparan la mayor parte de las aplicaciones basadas en detección por espectrometría de masas con separación previa.

La ESI es particularmente útil en el acoplamiento de MS con LC, ya que la ionización se produce a presión atmosférica, por lo que resulta especialmente adecuada para compuestos polares e iónicos. En ESI la muestra se pulveriza en un metal o en un capilar de sílice fundida. El electroespray se consigue aumentado el potencial en el capilar a 4 kV, tanto para el modo de ionización positivo como para el negativo. El espray resultante, formado por gotitas cargadas, se dirige al contra-electrodo a un potencial más bajo; en él las gotitas pierden el disolvente por evaporación quedando como especies iónicas en fase gaseosa. El contra-electrodo contiene un orificio a través del cual los iones se transportan a la cámara de vacío del espectrómetro en un proceso facilitado por la alta concentración de carga del mismo signo, positiva o negativa, atravesando regiones bombeadas diferencialmente mediante lentes selectoras (“skimmers”) [25]. Los iones generados se transportan al analizador de masas mediante un campo eléctrico o magnético.

Por otro lado, la ionización por EI se produce por la interacción de los electrones energéticos emitidos por un filamento caliente de tungsteno o renio y son acelerados por un potencial de aproximadamente 70 V que se aplica entre el filamento y el ánodo. Las trayectorias de los electrones y las moléculas forman ángulo recto y se cruzan en el centro de la fuente, donde colisionan y tiene lugar

la ionización. El producto primario son iones de una única carga positiva que se forman cuando los electrones de elevada energía se acercan suficientemente a las moléculas como para causarles la pérdida de electrones por repulsiones electrostáticas; por tanto, el producto o ión molecular es un ion radical que tiene el mismo peso molecular que la molécula [26].

Una vez que en la interfase se ha llevado a cabo la transferencia de los iones, éstos son dirigidos hacia el analizador de masas. Los analizadores de masas permiten la separación, detección y cuantificación de los analitos en estudio con un grado de sensibilidad y selectividad muy elevado, proporcionando información sobre su masa molecular.

Existen diversos tipos de analizadores: cuadrupolos, trampas de iones, analizadores de tiempo de vuelo, etc. Puesto que la gama de analizadores de masas comercialmente asequibles es amplia, sólo se describirán de forma breve los más utilizados en metabolómica.

Un detector de masas de cuadrupolo simple consiste en cuatro barras paralelas de sección hiperbólica en la cara interna, generalmente de unos 15-20 cm de largo y 0.5 cm de radio, separadas entre sí unos 2 cm, a las que se aplica un potencial combinado de corriente continua y de radiofrecuencia que crean en su interior un campo denominado cuadrupolar (Figura 4). Los iones generados en la fuente deben atravesar longitudinalmente el recinto limitado por estas barras para incidir en el detector. Estos iones, que entran en el analizador con una energía de unos pocos electronvoltios, son sometidos al efecto del campo cuadrupolar que los hace oscilar y los desvía en función de su valor *m/z* de forma que para una combinación de potenciales sólo los iones en un estrecho rango de valores *m/z* llega a incidir en el detector. Los analizadores de cuadrupolo actúan por tanto como filtros de iones y los espectros de masas en estos sistemas se obtienen mediante un barrido de potencial aplicado a las barras. En consecuencia, en cada instante sólo una pequeña fracción del total de iones es monitorizada mientras que el resto se desecha.

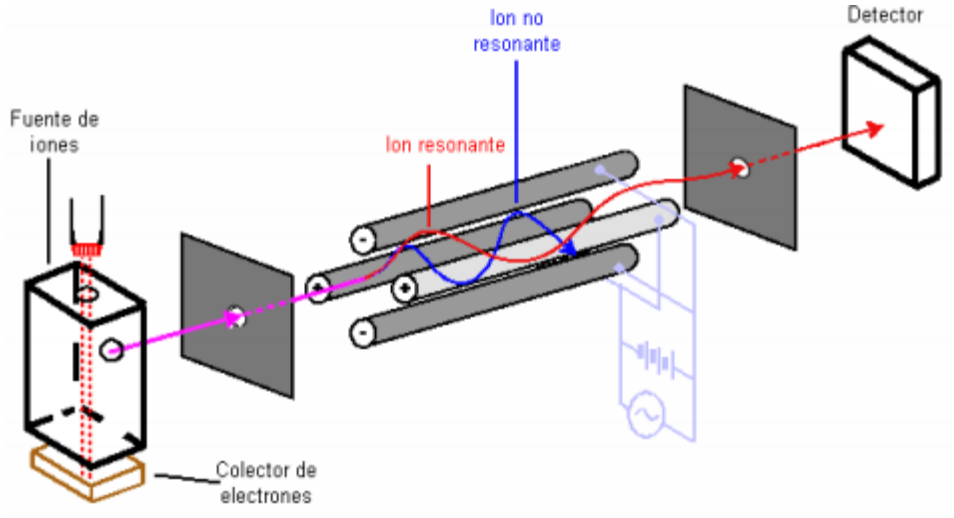


Figura 4: Esquema de un analizador de masas de cuadrupolo simple.

Como los analizadores de cuadrupolo simple trabajan sólo con campos eléctricos, los barridos pueden ser extraordinariamente rápidos (0.01 s); por otra parte, no precisan rendijas para el enfoque del haz con lo que la sensibilidad del equipo aumenta significativamente. Otra ventaja es que la escala de masas es lineal con respecto a los potenciales utilizados, de forma que en los espectros obtenidos por este procedimiento pueden realizarse interpolaciones de masas con facilidad.

El principal inconveniente que presenta este tipo de analizador es su poder de resolución, relativamente pequeño –del orden de 500 a 1000– y su limitado rango de utilización, ya que sólo permite la separación de iones con relaciones *m/z* menores de 1000.

El analizador de trampa de iones permite el confinamiento de iones dentro de una cámara de pequeño tamaño utilizando campos eléctricos. Este tipo de analizadores permite almacenar, seleccionar y analizar los iones formados en la misma trampa o en fuentes de ionización externas. Los iones pueden mantenerse en el interior de la trampa durante tiempos largos con objeto de favorecer la observación de descomposiciones metaestables o de fragmentos producidos por colisión de moléculas de gas. Los iones fragmentos generados

en estos procesos pueden ser a su vez seleccionados de nuevo en la misma trampa de forma que el sistema equivale a un sistema de espectrometría de masas en tándem múltiple.

Uno de los analizadores más usuales es el denominado IT. Este tipo de analizador es una modificación del analizador de cuadrupolo simple, y está basado en la utilización de una zona de confinamiento electromagnética generada por dos señales de radiofrecuencia. El analizador de IT (Figura 5) está formado por tres electrodos de superficie hiperbólica; de ellos, el electrodo central es anular, mientras que los electrodos superior e inferior forman el cierre de los extremos del anillo. Los tres electrodos forman una cavidad en la que se produce la ionización, la fragmentación y el análisis de masas.

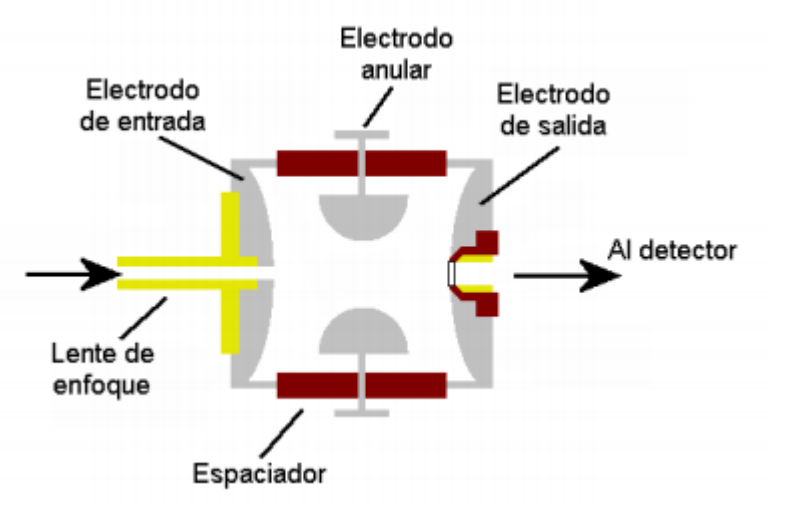


Figura 5: Esquema de un analizador de trampa de iones.

Durante el proceso de análisis de masas, se aplica entre los electrodos superior e inferior un potencial de radiofrecuencia de 525 kHz (voltaje de modulación axial); al mismo tiempo, sobre el electrodo central se aplica otro voltaje de radiofrecuencia de 1.1 MHz y de amplitud variable entre 0 y 7500 V. Estos dos voltajes de radiofrecuencia dan lugar a un campo electromagnético cuadrupolar tridimensional en el que quedan confinados los iones con una

trayectoria oscilante estable, dependiendo el movimiento exacto de cada ion de los voltajes aplicados y de su relación *m/z*. Para detectar los iones, se altera la señal de radiofrecuencia del electrodo anular, lo que da lugar a la desestabilización de la trayectoria de un ion concreto que es eyectado de la trampa. Un cambio gradual en la amplitud del campo de radiofrecuencia dará lugar a que los iones sean eyectados de la trampa en orden creciente de su relación *m/z*, lo que producirá un espectro de masas.

El analizador de IT presenta la ventaja de una sensibilidad mayor que las conseguidas utilizando otros analizadores. Al mismo tiempo, su velocidad de barrido es también muy alta, por lo que sus prestaciones como detector cromatográfico son francamente buenas. Por otro lado, debe tenerse en cuenta que, al existir una concentración de iones bastante elevada en el interior de la trampa, es relativamente frecuente la existencia de reacciones bimoleculares entre los iones, lo que puede dar lugar a espectros con esquemas de fragmentación diferentes a los que se obtienen con otros tipos de analizadores.

Un detector de masas de QqQ consta de las unidades que se muestran en la Figura 6 [25]: Una fuente de iones (generalmente una fuente de ESI) seguida de un conjunto de lentes para la transferencia de iones al primer cuadrupolo, formado por cuatro barras paralelas a las que se aplican voltajes específicos de corriente continua y de radiofrecuencia que hacen que todos los iones excepto los de uno o varios valores de *m/z* se filtren a su través eliminándose. El voltaje aplicado es variable, por lo que secuencialmente unos iones pasan entre ellos y otros siguen circulando hasta alcanzar la celda de colisión en la que se fragmentan. Esta celda, que generalmente recibe el nombre de segundo cuadrupolo, es en realidad un hexapolo relleno de un gas inerte, nitrógeno o argón, en el que se fragmentan los iones, que se envían al tercer cuadrupolo para una segunda etapa de filtrado que permite aislar y examinar las múltiples transiciones desde el precursor al ión producto. Éste es el modo llamado monitorización de reacciones seleccionadas (selected reaction monitoring o SRM). Puesto que los fragmentos iónicos formados son partes de la molécula precursora, representan porciones de su estructura global. Estos

detectores se utilizan preferentemente para análisis orientado, ya que permiten cuantificar con excelente sensibilidad y selectividad en el modo SRM.

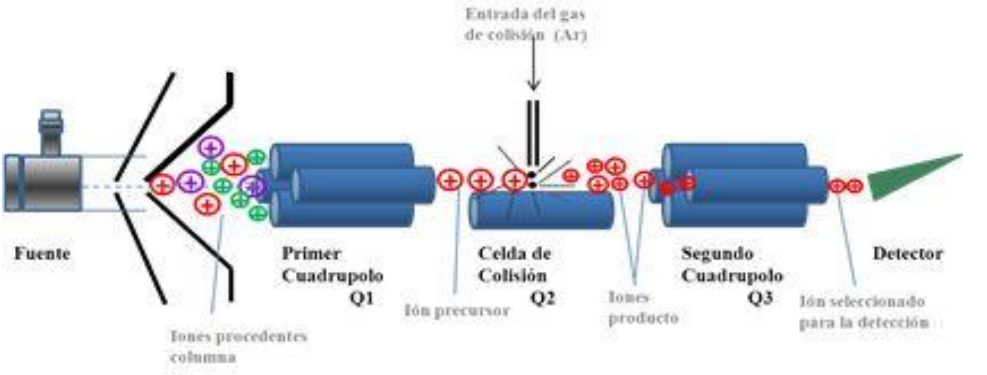


Figura 6: Esquema de un analizador de triple cuadrupolo.

El detector de masas del tipo QTOF ha sido el espectrómetro utilizado en la investigación recogida en esta Memoria. En la Figura 7 se puede ver un esquema de este detector [27] que consta de: una fuente de iones (ESI para el acoplamiento con LC y EI para el acoplamiento con GC) seguida de un conjunto de lentes para la transferencia de iones al primer cuadrupolo, formado por cuatro barras paralelas a las que se aplican voltajes específicos de corriente continua y de radiofrecuencia para el filtrado de iones de *m/z* concreta con el fin de fragmentarlos en la célula de colisión. El voltaje aplicado puede ser fijo para aislar un valor de *m/z* característico o variable para aislar de forma secuencial un rango de iones determinado con diferente valor de *m/z*. Los iones aislados se dirigen a la celda de colisión en la que se fragmentan mediante la aplicación de una energía de colisión. Esta celda, que generalmente recibe el nombre de segundo cuadrupolo, generalmente se trata de un hexapolo en el que se inyecta un gas inerte, nitrógeno o argón, como gas de colisión para fragmentar los iones que se envían al tubo de tiempo de vuelo (TOF, del inglés "time of flight"). La unidad TOF se encuentra en configuración ortogonal con los dos cuadrupolos. A los iones que llegan al TOF se les aplica un pulso de energía constante de forma que recorren el tubo a una velocidad inversamente proporcional a su relación *m/z* que produce la separación entre ellos llegando al detector y generando el

espectro. La longitud del tubo de vuelo establece la resolución espectral ya que, a mayor longitud, mayor resolución espectral podrá conseguirse.

El analizador de masas QTOF ofrece una exactitud de masas mayor que la del Orbitrap, además de un coste de adquisición mucho menor. Sin embargo, el Orbitrap permite realizar masas en tándem (MS^n) y, además, ofrece una mayor resolución que el QTOF.

Comparando con el analizador de QqQ, el QTOF ofrece una mayor selectividad que el de QqQ, aunque la sensibilidad es considerablemente menor. Por tanto, el QTOF está orientado al análisis cualitativo o semicuantitativo, mientras que el QqQ encuentra su principal área de aplicación en el análisis cuantitativo y confirmatorio. El QTOF puede funcionar en modo MS con el TOF como herramienta de análisis, aprovechando la alta precisión en masa, o en modo MS/MS para la elucidación estructural. Por otra parte, y por su exactitud en la medida de masas (error menor de 2 ppm), se puede llevar a cabo una muy buena identificación lo que lo convierte en una herramienta adecuada para análisis global (obtención de perfiles metabolómicos).

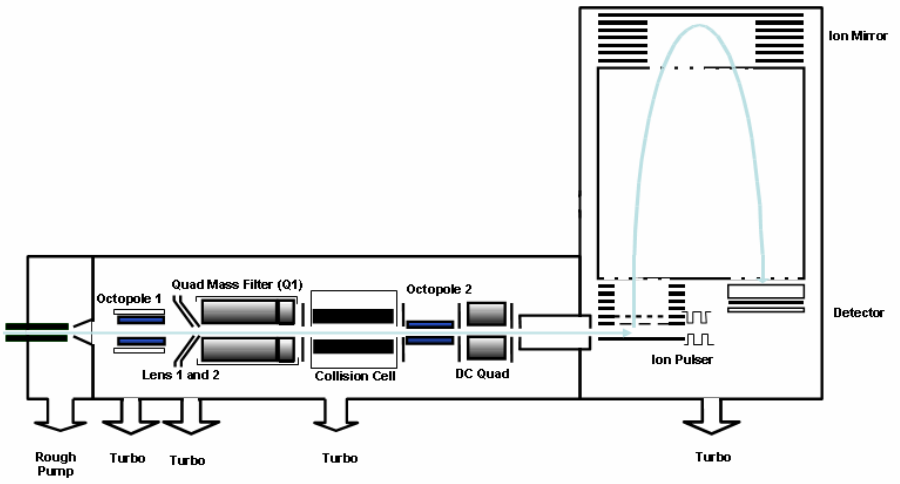


Figura 7. Esquema general de un espectrómetro de masas de tiempo de vuelo de cuadrupolo.

El Orbitrap (Figura 8) es un analizador que consiste en un electrodo exterior con forma de barril y un electrodo interior coaxial con forma de huso para atrapar iones en el campo electrostático formado entre ellos. Los iones introducidos de forma perpendicular en este sistema adquieren un movimiento radial alrededor del huso combinado con un movimiento axial periódico cuya frecuencia es función de su valor *m/z*. El electrodo externo está dividido en su centro y perpendicularmente en dos mitades, de forma que el movimiento periódico de los paquetes de iones de un lado a otro de la cavidad induce entre las dos mitades una señal imagen cuya frecuencia puede medirse para determinar el valor *m/z* correspondiente. En la práctica, la señal registrada es una señal compleja o interferograma formada por la combinación de las señales senoidales de distintas frecuencias (*m/z*) e intensidades (abundancias) de cada uno de los iones. La aplicación a esta señal de la transformada de Fourier, método matemático para separar los componentes de una señal periódica compleja, permite determinar la frecuencia derivada de cada uno de los iones y obtener el espectro de masas. A pesar de ser uno de los espectrómetros de masas más reciente, su alta resolución (pudiendo alcanzar resoluciones del orden de 100.000 FWHM) y la exactitud de masa que ofrece (pudiendo alcanzar los <2.0 ppm) han permitido que se introduzca rápidamente en el área de la proteómica en una configuración trampa de iones lineal-Orbitrap. Recientemente se ha propuesto una configuración Q-Orbitrap con una alta resolución para su implementación en metabolómica. Esta versión, mucho más económica que la configuración utilizada en proteómica, se ha convertido rápidamente en un competidor directo al QTOF en metodologías de análisis global.

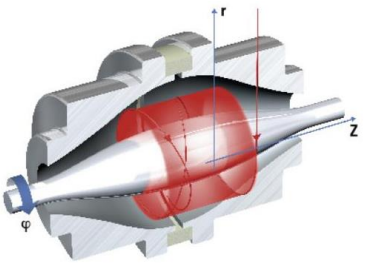


Figura 8: Esquema de un Orbitrap.

4. Aplicaciones de la metabolómica I: explotación de residuos agroalimentarios

Tradicionalmente, los cultivos se han utilizado mayoritariamente para la obtención de un único producto de alto valor, como es el caso del aceite de oliva, la sidra o el vino, en los que una parte importante de la planta o la fruta constituye un residuo de nulo o escaso valor; la mayoría de las veces muy contaminante y, por tanto, indeseable. Éste es el caso del alperujo y del alpechín, residuos de la industria del aceite de oliva. En la actualidad, el notable crecimiento de la generación de residuos es una consecuencia directa del aumento de la producción. Los desechos agroindustriales se generan a partir de la transformación de productos primarios, resultando materiales, sólidos o líquidos, que no son útiles para el proceso en el que se generaron, pero que podrían utilizarse o transformarse en productos con otra aplicación.

Con frecuencia los términos subproductos, residuos y desechos se utilizan indistintamente, a pesar de que existe una diferencia conceptual entre ellos. De hecho, un "subproducto" es un producto secundario que resulta de un proceso industrial y que es útil en algún grado, comercializable y con un cierto valor añadido. El término "residuo" se aplica a materiales que pueden o no tener valor comercial, porque son raros o generados a niveles bajos. Sin embargo, algunos componentes de los residuos, incluso en baja proporción, les pueden conferir interés. Desde este punto de vista, los términos "subproductos" y "residuos" se pueden utilizar como sinónimos, pero no el término "desechos", lo que se conoce como materiales sin valor comercial, o ningún interés para su uso; por lo tanto, los "desechos" se consideran basura.

En general, las características de los residuos agroindustriales son variadas y dependen de la materia prima y del proceso por el que se generan. Sin embargo, todos los desechos agrícolas comparten una característica: el contenido de materia orgánica, compuesta por diferentes porcentajes de celulosa, lignina, hemicelulosa y pectina, por lo que reciben el nombre de

"residuos orgánicos". En esta categoría se incluyen otros residuos como los lodos de plantas de tratamiento, aguas residuales, así como los residuos domésticos y residuos sólidos urbanos. Como resultado del aumento de la generación de residuos, miles de toneladas de ellos se queman, se depositan en vertederos, se arrojan a ríos, etc., con el daño correspondiente al ecosistema. Esta situación ha propiciado la implantación de políticas nacionales e internacionales para sancionar prácticas ilegales en este campo.

Las políticas para la explotación de los residuos procedentes de la industria agroalimentaria de la cuenca mediterránea afectan principalmente a los dos cultivos más abundantes de esta región: la vid y el olivo. En ambos casos, tanto los residuos de la planta como de la industria correspondiente contienen compuestos de interés para industrias como la farmacológica, cosmética, alimentaria y nutracéutica.

Los residuos en estos casos contienen cantidades significativas de fenoles, que los dotan de un alto valor añadido, propiciando, por tanto, su explotación, que debería abarcar:

(1) Los residuos ricos en compuestos de alto valor añadido que han sido reconocidos como tales por la Autoridad Europea de Seguridad Alimentaria (EFSA), como es el caso de hidroxitirosol y los derivados de oleuropeína, presentes en las hojas de olivo, alpechín o alperujo, además de en el fruto y el aceite. Estos compuestos disminuyen el riesgo de padecer aterosclerosis [28] y ayudan a mantener los niveles normales de colesterol HDL en sangre [29].

(2) Los compuestos contenidos en los residuos del apartado (1) cuyos efectos beneficiosos no están suficientemente estudiados para ser aceptados por la EFSA, pero son ampliamente utilizados en alimentación y cosmética.

(3) Residuos, cuyo potencial no ha sido reconocido o está mal explotado.

El objetivo de este apartado es mostrar la importancia de los residuos de los cultivos y productos de las dos principales agroindustrias de la cuenca del

Mediterráneo (vino y aceite de oliva). La explotación de estos residuos puede proporcionar beneficios a la industria agroalimentaria gracias a (entre otros) el alto contenido en compuestos con potencial antioxidante con propiedades nutricionales testadas y su utilidad también en el campo de la cosmética, además del beneficio que supone la reducción de la contaminación por la acumulación de residuos.

El Capítulo I de esta Memoria ofrece una amplia visión de la situación actual de la explotación de residuos de las industrias del vino y del aceite de oliva, además de tratar de forma separada los estudios realizados sobre el aprovechamiento de los residuos de la vid y el olivo (sarmientos y hojas de vid y hojas de olivo) de los llevados a cabo con los procedentes de la producción del vino (semillas, hollejos y lías de vinificación) y del aceite (alpechín/orujo y alperujo).

El Capítulo II presenta una amplia visión de la familia de compuestos más importante de los presentes en estos residuos: los compuestos fenólicos. Se trata de un capítulo de libro desarrollado desde la perspectiva de los fenoles en la uva, en el que se estudian también sus efectos positivos para la salud.

Finalmente, los Capítulos III-V de esta Memoria ponen de manifiesto el valor de algunos de los residuos descritos en el Capítulo I. Cabe destacar que en la investigación recogida en los Capítulos III y IV se utilizaron residuos procedentes de otros residuos, asignatura pendiente en el campo de la explotación de estos materiales. Los hollejos y las pepitas de uvas tinta y blanca se utilizaron tras soportar el proceso de destilación de alcohol. El Capítulo V recoge la investigación sobre la extracción y la caracterización de otro residuo hasta ahora desaprovechado desde el punto de vista de extracción de compuestos de interés, el alperujo.

5. Aplicaciones de la metabolómica II: medicina personalizada. Búsqueda de biomarcadores

5.1. Medicina personalizada

A pesar de que la metabolómica es la más reciente de las ómicas primarias, su papel en la medicina personalizada merece ser discutido debido a su significativa contribución al campo clínico. Las ómicas secundarias en las que se han desarrollado las contribuciones más destacadas son la lipidómica y la nutrimetabolómica. En el mundo de las ómicas, los lípidos han pasado desapercibidos mientras que los ácidos nucleicos y las proteínas eran el centro de atención. Recientemente, la investigación sobre los lípidos ha aumentado, gracias a las mejoras en las plataformas analíticas empleadas para el análisis metabolómico, demostrándose su implicación en enfermedades como la obesidad, la aterosclerosis, la diabetes o el Alzheimer [3,30–35]. La nutrimetabolómica también ha demostrado ser uno de los apoyos más firmes de la medicina personalizada debido a la estrecha relación entre la nutrición y la salud [36–49].

Las áreas médicas en las que es más común el empleo de la metabolómica para el diagnóstico de enfermedades son la cardiología [50], la reproducción humana [51], la diabetes [31,52], la oncología [11,53,54] y las enfermedades del sistema nervioso central [32,55,56]. Los estudios sobre el diagnóstico de enfermedades se orientan al descubrimiento y la caracterización de nuevos biomarcadores. La búsqueda y el uso de biomarcadores en metabolómica se exponen en el siguiente apartado.

También existen estudios clínicos en metabolómica cuyo objetivo es entender o encontrar qué vías metabólicas se perturban o modifican por la presencia de una enfermedad determinada. Este tipo de investigación ha permitido conocer, por ejemplo, cómo el metabolismo de lípidos se modifica por

la enfermedad de Alzheimer [57] o cómo el aumento de la obesidad a nivel mundial está relacionado con la diabetes tipo 2 (DM2) y los trastornos cardiovasculares. Una cadena de aminoácidos ramificada se ha relacionado con la resistencia a la insulina en humanos obesos [58].

El estudio de los trastornos respiratorios a través de la metabolómica se ha incrementado también en los últimos años a través del análisis de los compuestos volátiles contenidos en el aliento. A modo de ejemplo, Fowler y col. analizaron el aliento de pacientes intubados para el estudio de los patógenos de las vías respiratorias [59]. Las infecciones relacionadas con la intubación de pacientes son difíciles de diagnosticar y tratar, por lo que el objetivo de este estudio fue demostrar la presencia de patógenos clínicamente relevantes en el aliento que proviene del tracto respiratorio inferior. Se recogieron muestras de 46 pacientes y se tuvieron en cuenta para el análisis estadístico tanto el perfil individual como la duración de la intubación. El análisis de la fracción volátil del aliento permitió distinguir entre los pacientes con y sin presencia significativa de patógenos en el tracto respiratorio inferior. Si este estudio se validara a gran escala, estos hallazgos podrían conducir al desarrollo de sistemas y medios de diagnóstico para la infección del tracto respiratorio inferior.

El asma es otra enfermedad común de las vías respiratorias. Para comprender mejor los fenotipos metabólicos del asma, Yu y col. han identificado dieciséis metabolitos como potenciales biomarcadores de la enfermedad. En este estudio se ha presentado un mapa completo de las rutas metabólicas perturbadas en el plasma de ratones con asma inducido. Se han propuesto como marcadores dieciséis metabolitos debido a su implicación en la alteración de 6 rutas metabólicas. Estos resultados demuestran el potencial del estudio metabolómico del plasma para contribuir a una comprensión en profundidad de la regulación metabólica del asma a nivel molecular [60].

La medicina personalizada ofrece nuevas oportunidades para la medicina preventiva, proporcionando nuevas estrategias de pronóstico de enfermedades. Además, la importancia de la metabolómica queda demostrada

por su presencia en todas las definiciones actuales de medicina personalizada [61] como, por ejemplo: “La medicina personalizada busca mejorar las medidas preventivas y terapéuticas mediante la información biológica y la búsqueda de biomarcadores a nivel molecular para cualquier enfermedad haciendo uso de la genética, la proteómica, y la metabolómica; así como disminuir el tiempo del análisis”.

5.2. Búsqueda de biomarcadores

En 2001, el Instituto Nacional de la Salud definió el término biomarcador como "una característica que se mide y se evalúa objetivamente como un indicador de procesos biológicos normales, procesos de patogénesis, o respuestas farmacológicas a una intervención terapéutica" [62]. A pesar de que los biomarcadores no son nuevos en el campo de la medicina [63] han ganado gran interés científico y clínico en los últimos años gracias, en muchos aspectos, a las disciplinas ómicas.

Los biomarcadores son útiles para monitorizar el desarrollo de una enfermedad de modo que pueden utilizarse en el contexto de la prevención primaria (para prevenir la enfermedad en sí), pueden facilitar la prevención secundaria (para la detección temprana de la enfermedad) e incluso pueden contribuir en la prevención terciaria (para ayudar a guiar el tratamiento) [64].

Las características de un marcador ideal son:

1) Ser seguro y fácil de medir, de manera que la medición se pueda repetir, si es necesario, sin peligro para el paciente y que pueda realizarse por personal no muy especializado.

2) Bajo costo de la medida. No sólo el análisis, sino también el costo de las pruebas de seguimiento debe ser bajos para monitorizar los niveles de los marcadores y la eficacia del tratamiento.

3) La presencia de la enfermedad en una proporción moderada de la población debe ser explicada por el biomarcador.

4) Alta precisión predictiva para detectar la presencia de la enfermedad (sensibilidad) o la ausencia de la misma (especificidad). Ambos parámetros, sensibilidad y especificidad, del biomarcador deben ser relativamente altos en función de la finalidad de la prueba, que puede estar orientada a reducir la tasa de falsos positivos, la tasa de falsos negativos o ambas.

5) Consistente a través de sexos y grupos étnicos.

Tanto la plataforma utilizada para obtener los datos analíticos primarios como las herramientas quimiométricas usadas para el tratamiento de datos, junto con las correspondientes bases de datos, son claves para el éxito en la búsqueda de biomarcadores [65].

5.1. El proceso para el desarrollo de biomarcadores

El proceso de desarrollo de biomarcadores se puede dividir en tres etapas principales [66]: el descubrimiento, la verificación y la aprobación. La primera etapa, el descubrimiento, puede dividirse en dos sub-etapas: identificación del biomarcador candidato y la validación analítica. La primera generalmente implica la aplicación de una estrategia de análisis no orientado o de huella dactilar con el fin de identificar moléculas con potencial marcador que, básicamente, corresponden a moléculas presentes en una concentración significativamente diferente en grupos de individuos. Una vez identificados los compuestos que presentan la capacidad de discriminar entre grupos de individuos debe evaluarse su potencial marcador utilizando algoritmos adecuados tales como el análisis de curvas ROC. Los biomarcadores pueden ser independientes o combinarse en paneles con el fin de mejorar la capacidad discriminante respecto a las moléculas de forma independiente. Una vez identificadas las moléculas candidatas a marcadores, la siguiente etapa tiene

como objetivo realizar una validación analítica para evaluar su comportamiento predictivo mediante el uso de métodos más sensibles. Un ejemplo puede ser el análisis de proteínas y metabolitos mediante LC-MS/MS utilizando como detector un analizador de triple cuadrupolo, especialmente ideal para el análisis confirmatorio y cuantitativo. Un ejemplo representativo de una plataforma analítica para el descubrimiento/identificación de biomarcadores mediante análisis proteómico o metabolómico puede verse en la Figura 9. La plataforma está basada en el acoplamiento LC-MS/MS e integra la utilización de dos espectrómetros de masas de diferentes prestaciones: un QTOF, ideal para la primera fase de identificación, y un QqQ que es adecuado para la etapa de validación.

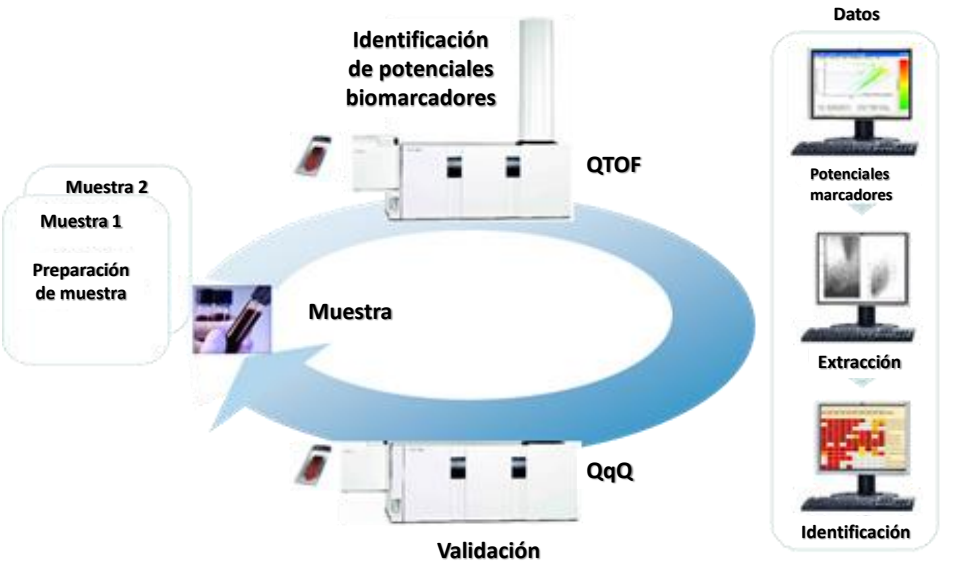


Figura 9. Plataforma analítica para la identificación y validación de biomarcadores.

La etapa de verificación enlaza el marcador molecular con la alteración del proceso bioquímico en que se fundamenta la enfermedad o con el efecto de un medicamento. La verificación también puede combinarse con la validación analítica comprobando el efecto del marcador a otros niveles. Por ejemplo, si el

compuesto marcador es un metabolito, podría desarrollarse el proceso de verificación a nivel proteico, de ARNm o de genes.

El proceso de aprobación es el puente que conecta los resultados de las mediciones de marcadores moleculares, los efectos de los medicamentos sintomáticos y los resultados de la enfermedad. Se define como "un proceso cuya finalidad es evidenciar el nexo entre un biomarcador y la biología con criterios de valoración clínicos" [67]. Se diferencia de la validación en que se centra en las características de fiabilidad y rendimiento del ensayo analítico utilizado para medir los marcadores moleculares [66,68]. El paso de la etapa de descubrimiento al desarrollo clínico depende en gran medida de la disponibilidad de ensayos robustos, precisos y sensibles para la medición de un número mayor de muestras [69]. Durante la fase de descubrimiento, una validación parcial es suficiente, pero esta etapa debe ser más rigurosa si las estrategias de desarrollo de fármacos y las decisiones clínicas van a depender de tal marcador molecular. La cuantificación absoluta de los marcadores y la validación completa es crítica durante el desarrollo clínico, especialmente cuando se pretende desarrollar una herramienta de diagnóstico. Un marcador molecular adecuadamente validado y que soporte los resultados primarios puede ayudar a comprender y controlar los mecanismos de toxicidad, las interacciones de medicamentos, las interacciones de la enfermedad con las drogas y los efectos de los genotipos, el sexo y la edad. Además, pueden utilizarse para estratificar poblaciones de pacientes [70].

5.2. Biomarcadores metabolómicos

Se ha seleccionado el estudio del cáncer de pulmón como ejemplo para comentar con mayor profundidad la importancia de los biomarcadores metabolómicos porque ha sido ésta el área en la que se ha investigado en la parte clínica de esta Memoria.

El cáncer de pulmón es la principal causa de muerte por cáncer en los países desarrollados [71]. Esta situación es consecuencia de que el 60% de los

pacientes son diagnosticados en etapas avanzadas de la enfermedad [72]. La tasa de mortalidad anual por cáncer de pulmón es superior a la combinación de las tasas anuales de cáncer de mama, próstata y colon, para todos los cuales existen herramientas clínicas que permiten la detección de la enfermedad en estadios más tempranos [73]. Por lo tanto, la búsqueda de estrategias de diagnóstico para la detección precoz del cáncer de pulmón se ha intensificado en la última década. La detección temprana implica el estudio de poblaciones de individuos con alto riesgo, tanto antes como después de que se detecte la enfermedad (Figura 10).

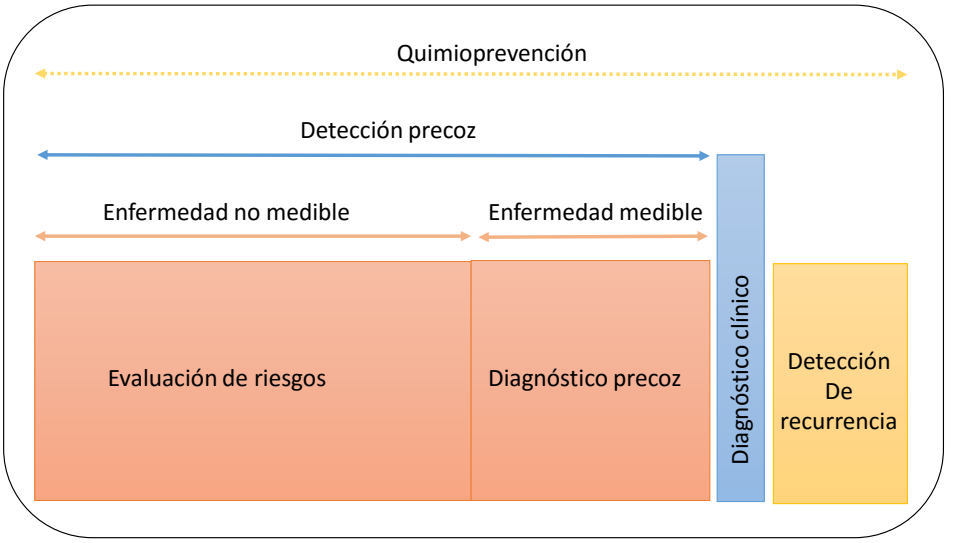


Figura 10. Contextos clínicos para el desarrollo de biomarcadores en la detección precoz del cáncer de pulmón.

La búsqueda de biomarcadores para la detección temprana de cáncer de pulmón ha incluido las siguientes muestras-ómicas:

(i) Los tejidos y la genómica (la hipermetilación de un número de genes supresores del tumor [74-76]), transcriptómica (regiones de amplificación cromosómica [77]), la variación de la expresión del ARNm [78,79], la expresión

diferencial de varios microRNAs [80]), la proteómica [81,82]) y, por supuesto, la metabolómica, que se considera a continuación con más detalle.

El estudio del perfil metabolómico de tejido de pulmón obtenido por CE-TOF/MS proporcionó la identificación de 114 metabolitos, que permitió distinguir tejido afectado de tejidos normales. Además, fue posible diferenciar el carcinoma de células escamosas de otros tipos de tumores de pulmón. Las concentraciones de la mayoría de los aminoácidos, especialmente los de cadena ramificada, fueron significativamente más altos en los tejidos tumorales [83].

(ii) Los biofluidos, incluyendo la sangre periférica y sus componentes (células circulantes en plasma y en suero), el aliento condensado, la orina, el esputo y el sudor, permiten un muestreo mínimamente invasivo o no invasivo, así como un amplio abanico de opciones de análisis. Las alteraciones debidas a la presencia de la enfermedad pueden conducir a la generación de especies moleculares específicas tales como ADN alterado o metilado, ARNm sobreexpresado, microRNAs o proteínas que potencialmente pueden ser liberadas en el microambiente extracelular. Por lo tanto, el análisis molecular de fluidos biológicos relacionados con el cáncer de pulmón en etapa temprana representa una opción atractiva para el descubrimiento y validación de biomarcadores de diagnóstico [84,85].

El análisis de suero procedente de 29 voluntarios sanos y 33 pacientes con cáncer de pulmón (adenocarcinoma, n=12; carcinoma de células escamosas, n=11; o carcinoma de células pequeñas, n=10 que van desde el estadio I de la enfermedad al IV), y el tejido pulmonar de 7 pacientes con cáncer de pulmón (incluyendo el tejido tumoral y el normal circundante) permitió la detección de 58 metabolitos en el suero, y 71 metabolitos en el tejido pulmonar. Los niveles de 23 de los 58 metabolitos en suero cambiaron significativamente en todos los pacientes con cáncer de pulmón en comparación con voluntarios sanos, y los niveles de 48 de los 71 metabolitos cambiaron de manera significativa en el tejido tumoral en comparación con el no tumoral. Los datos de

la muestra de suero se sometieron a análisis estadístico que mostraron discriminación por subtipo histológico y estadio [86].

Uno de los estudios metabolómicos más recientes en la búsqueda de biomarcadores de cáncer de pulmón se basó en el empleo de sudor como muestra analítica. Un protocolo de análisis basado en LC-QTOF MS/MS permitió la identificación de un número grande de componentes (en total 43 metabolitos) pertenecientes a una amplia variedad de familias. Entre ellos se identificaron 19 aminoácidos, incluyendo todos los esenciales, excepto la lisina. También se identificaron compuestos exógenos, como la cafeína y la teofilina, así como ácidos dicarboxílicos como el sebácico, el subérico, el azeálico o el butanodioico. Este estudio de identificación ha contribuido a conocer este biofluido tan poco estudiado y ha sugerido que, puesto que algunos de los metabolitos identificados habían sido anteriormente evaluados en otros biofluidos como potenciales biomarcadores, el sudor puede también proponerse como biofluido para la búsqueda de biomarcadores [87]. Con estas premisas, el potencial del sudor como biofluido para su implantación en el diagnóstico del cáncer se demostró de la siguiente forma: se construyó un modelo de predicción basado en un panel de metabolitos que incluyó aminoácidos, azúcares y algunos lípidos y que permitió discriminar entre pacientes con cáncer y un grupo control. Su alto valor predictivo permite prever su potencial uso para reducir el número de posibles casos que tienen que someterse a pruebas confirmatorias [88]. Se desarrollaron dos paneles de marcadores que incluían maltotriosa y ácido nonanedioico en combinación con γ-GluLeu y MG(22:2). Ambos paneles mejoraban de forma significativa la capacidad de discriminación de los metabolitos independientes. El primer panel se caracterizó por el 100% de especificidad y el 63.6% de sensibilidad; por tanto, la presencia de falsos negativos fue del 0%. El segundo panel proporcionó valores de especificidad y sensibilidad del 82%. Estos resultados preliminares hacen muy conveniente un estudio a gran escala para validar los paneles propuestos con un doble objetivo: reducir el número de individuos que tengan que someterse a un test confirmatorio y detectar el cáncer de pulmón en una etapa lo más temprana posible [88].

Mención especial merece el uso reciente del aire exhalado condensado (EBC), comúnmente conocido como aliento, como muestra para la búsqueda de biomarcadores de cáncer de pulmón, como se discute en los Capítulos VII-VIII de esta Memoria [89,90]. El aliento está cargado básicamente de vapor de agua y de un conjunto de moléculas que expresan la situación funcional del pulmón y de otros tejidos. Estas moléculas (volátiles o no) pueden analizarse mediante la condensación del aliento haciéndolo pasar por un circuito frío. El muestreo del aliento condensado es sencillo, no invasivo, requiere poca colaboración del paciente (de especial interés para su uso en pediatría y geriatría) y es reproducible en el tiempo, con períodos de recogida cortos, que no necesita de instalaciones especiales ni de personal particularmente entrenado, además de realizarse con equipos portátiles y con un bajo coste [91]. Los compuestos que hasta ahora han sido identificados en aliento condensado incluyen, además de las proteínas y del ADN [92], los productos de peroxidación de lípidos, los del metabolismo del óxido de nitrógeno, los iones hidrógeno, el peróxido de hidrógeno y las citoquinas. El EBC presenta potencial interés en la investigación metabolómica para el diagnóstico de cáncer de pulmón [93]. Varios estudios recientes han utilizado el análisis mediante GC-MS de los compuestos orgánicos volátiles [94-96]. Otros grupos han utilizado GC-MS y nanosensores de diseño personalizado en los que cambios en la resistencia eléctrica producidos por compuestos orgánicos existentes en el aliento de los pacientes se han utilizado para su detección. Por ejemplo, en un estudio realizado por Peng y col., un grupo de compuestos orgánicos volátiles permitió distinguir entre pacientes con cáncer de pulmón, colorrectal, de mama e individuos sanos [97-99]. Otros estudios orientados a la identificación de proteínas y péptidos volátiles presentes en EBC como marcadores potenciales para la detección precoz de cáncer de pulmón [98,99] proporcionaron pruebas de la viabilidad de esta estrategia para aislar e identificar proteínas útiles para este fin. A pesar de la existencia de numerosos estudios que muestran que la investigación en este campo ha aumentado exponencialmente, sigue siendo necesaria la estandarización del proceso de recogida, así como un protocolo de normalización de los datos de manera que pueda demostrarse de forma más concluyente la utilidad de este

tipo de muestra en la práctica clínica. El Capítulo VI de esta Memoria presenta alternativas a estos dos puntos débiles que permiten el desarrollo de estudios posteriores en los que se consideran las diferencias entre una población afectada con cáncer de pulmón y dos grupos de controles (con y sin factor de riesgo) (Capítulo VII) así como la propuesta de paneles de marcadores que permiten un primer cribado para el diagnóstico de esta enfermedad (Capítulo VIII). En el Capítulo IX se recoge el estudio sobre las diferencias entre una población formada por fumadores, exfumadores e individuos que nunca han fumado.

Referencias

[1] J.K. Nicholson, I.D. Wilson, Opinion: understanding "global" systems biology: metabonomics and the continuum of metabolism, Nat. Rev. Drug Discov. 2 (2003) 668–76.

[2] O. Fiehn, Metabolomics--the link between genotypes and phenotypes, Plant Mol. Biol. 48 (2002) 155–71.

[3] D. Beyoğlu, J.R. Idle, The metabolomic window into hepatobiliary disease., J. Hepatol. 59 (2013) 842–58.

[4] G.J. Patti, O. Yanes, G. Siuzdak, Innovation: Metabolomics: the apogee of the omics trilogy, Nat. Rev. Mol. Cell Biol. 13 (2012) 263–9.

[5] J.C. Lindon, J.K. Nicholson, Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery, Trends Anal. Chem. 27 (2008) 194–204.

[6] A. Scalbert, L. Brennan, C. Manach, C. Andres-Lacueva, L.O. Dragsted, J. Draper, et al., The food metabolome: a window over dietary exposure., Am. J. Clin. Nutr. 99 (2014) 1286–308.

[7] R. Goodacre, Metabolomics of a superorganism., J. Nutr. 137 (2007) 259S–266S.

[8] D.S. Wishart, T. Jewison, A.C. Guo, M. Wilson, C. Knox, Y. Liu, et al., HMDB 3.0--The Human Metabolome Database in 2013, Nucleic Acids Res. 41 (2013) D801–7.

[9] D. Wishart, R. Mandal, A. Stanislaus, M. Ramirez-Gaona, Cancer Metabolomics and the Human Metabolome Database, Metabolites. 6 (2016) 10.

[10] C.E. Mason, S.G. Porter, T.M. Smith, Characterizing multi-omic data in systems biology, Adv. Exp. Med. Biol. 799 (2014) 15–38.

[11] V.W. Davis, O.F. Bathe, D.E. Schiller, C.M. Slupsky, M.B. Sawyer, Metabolomics and surgical oncology: Potential role for small molecule biomarkers, J. Surg. Oncol. 103 (2011) 451–9.

[12] W.B. Dunn, D.I. Broadhurst, H.J. Atherton, R. Goodacre, J.L. Griffin, Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy, Chem. Soc. Rev. 40 (2011) 387–426.

[13] E. Allard, D. Bäckström, R. Danielsson, P.J.R. Sjöberg, J. Bergquist, Comparing capillary electrophoresis-mass spectrometry fingerprints of urine samples obtained after intake of coffee, tea, or water, Anal. Chem. 80 (2008) 8946–55.

[14] D.B. Kell, M. Brown, H.M. Davey, W.B. Dunn, I. Spasic, S.G. Oliver, Metabolic footprinting and systems biology: the medium is the message, Nat. Rev. Microbiol. 3 (2005) 557–65.

[15] B. Álvarez-Sánchez, F. Priego-Capote, M.D. Luque de Castro, Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation, TrAC Trends Anal. Chem. 29 (2010) 111–119.

[16] M.A. Fernández-Peralbo, M.D. Luque de Castro, Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis, Trends Anal. Chem. 41 (2012) 75–85.

[17] Á. Peralbo-Molina, F. Priego-Capote, M. Dolores Luque de Castro, Comparison of extraction methods for exploitation of grape skin residues from ethanol distillation, Talanta. 101 (2012) 292–298.

[18] Á. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro, Characterization of grape seed residues from the ethanol-distillation industry, Anal. Methods. 5 (2013) 1922.

[19] W.B. Dunn, D.I. Ellis, Metabolomics: Current analytical platforms and methodologies, TrAC Trends Anal. Chem. 24 (2005) 285–294.

[20] J.P.C. Vissers, Recent developments in microcolumn liquid chromatography, J. Chromatogr. A. 856 (1999) 117–143.

[21] S. Hsieh, J.W. Jorgenson, Preparation and evaluation of slurry-packed liquid chromatography microcolumns with inner diameters from 12 to 33 microns, Anal. Chem. 68 (1996) 1212–7.

[22] G. Desmet, D. Cabooter, P. Gzil, H. Verelst, D. Mangelings, Y. Vander Heyden, et al., Future of high pressure liquid chromatography: do we need porosity or do we need pressure? J. Chromatogr. A. 1130 (2006) 158–66.

[23] I.R. Lanza, S. Zhang, L.E. Ward, H. Karakelides, D. Raftery, K.S. Nair, Quantitative metabolomics by H-NMR and LC-MS/MS confirms altered metabolic pathways in diabetes, PLoS One. 5 (2010) e10538.

[24] P. Viñas, N. Campillo, M. Hernández-Pérez, M. Hernández-Córdoba, A comparison of solid-phase microextraction and stir bar sorptive extraction coupled to liquid chromatography for the rapid analysis of resveratrol isomers in wines, musts and fruit juices, Anal. Chim. Acta. 611 (2008) 119–25.

[25] W.J. Griffiths, Metabolomics, Metabonomics and Metabolite Profiling, 2008.

[26] H.-J. Hübschmann, Handbook of GC-MS: Fundamentals and Applications, 2015.

[27] H.L. Newmark, Squalene, olive oil, and cancer risk: a review and hypothesis., Cancer Epidemiol. Biomarkers Prev. 6 (1997) 1101–3.

[28] E.F.S. Authority, Scientific and technical guidance for the preparation and presentation of an application for authorisation of a health claim (revision 1), (n.d.).

[29] E.F.S. Authority, Scientific Opinion on the safety and efficacy of aliphatic and alicyclic ethers (chemical group 16) when used as flavourings for all animal species, (n.d.).

[30] A. D'Alessandro, B. Giardina, F. Gevi, A.M. Timperio, L. Zolla, Clinical metabolomics: the next stage of clinical biochemistry, Blood Transfus. 10 Suppl 2 (2012) s19–24.

[31] N. Friedrich, Metabolomics in diabetes research, J. Endocrinol. 215 (2012) 29–42.

[32] A. Zhang, H. Sun, X. Wang, Recent advances in metabolomics in neurological disease, and future perspectives, Anal. Bioanal. Chem. 405 (2013) 8143–50.

[33] R. Priori, R. Scrivo, J. Brandt, M. Valerio, L. Casadei, G. Valesini, et al., Metabolomics in rheumatic diseases: the potential of an emerging methodology for improved patient diagnosis, prognosis, and treatment efficacy, Autoimmun. Rev. 12 (2013) 1022–30.

[34] V. Fanos, C. Fanni, G. Ottonello, A. Noto, A. Dessì, M. Mussap, Metabolomics in adult and pediatric nephrology, Molecules. 18 (2013) 4844–57.

[35] Y.-Y. Zhao, Metabolomics in chronic kidney disease, Clin. Chim. Acta. 422 (2013) 59–69.

[36] E. Ibáñez, Applications of advanced omics technologies: from genes to metabolites, Elsevier, 2014.

[37] S.E. Richards, E. Holmes, Metabolomics as a tool in nutrition research, Elsevier, 2015.

[38] L. V Hooper, T. Midtvedt, J.I. Gordon, How host-microbial interactions shape the nutrient environment of the mammalian intestine, Annu. Rev. Nutr. 22 (2002) 283–307.

[39] F.-P.J. Martin, N. Sprenger, I. Montoliu, S. Rezzi, S. Kochhar, J.K. Nicholson, Dietary modulation of gut functional ecology studied by fecal metabonomics, J. Proteome Res. 9 (2010) 5284–95.

[40] F.-P.J. Martin, Y. Wang, I.K.S. Yap, N. Sprenger, J.C. Lindon, S. Rezzi, et al., Topographical variation in murine intestinal metabolic profiles in relation to microbiome speciation and functional ecological activity, J. Proteome Res. 8 (2009) 3464–74.

[41] Y. Wang, E. Holmes, E.M. Comelli, G. Fotopoulos, G. Dorta, H. Tang, et al., Topographical variation in metabolic signatures of human

gastrointestinal biopsies revealed by high-resolution magic-angle spinning 1H NMR spectroscopy, J. Proteome Res. 6 (2007) 3944–51.

[42] Y. Wang, H. Tang, E. Holmes, J.C. Lindon, M.E. Turini, N. Sprenger, et al., Biochemical characterization of rat intestine development using high-resolution magic-angle-spinning 1H NMR spectroscopy and multivariate data analysis, J. Proteome Res. 4 1324–9.

[43] F.-P.J. Martin, Y. Wang, N. Sprenger, E. Holmes, J.C. Lindon, S. Kochhar, et al., Effects of probiotic Lactobacillus paracasei treatment on the host gut tissue metabolic profiles probed via magic-angle-spinning NMR spectroscopy, J. Proteome Res. 6 (2007) 1471–81.

[44] M.D. Collins, G.R. Gibson, Probiotics, prebiotics, and synbiotics: approaches for modulating the microbial ecology of the gut, Am. J. Clin. Nutr. 69 (1999) 1052S–1057S.

[45] J.A.J. Madden, J.O. Hunter, A review of the role of the gut microflora in irritable bowel syndrome and the effects of probiotics, Br. J. Nutr. 88 Suppl 1 (2002) S67–72.

[46] S.J. Bickston, L.W. Comerford, F. Cominelli, Future therapies for inflammatory bowel disease, Curr. Gastroenterol. Rep. 5 (2003) 518–523.

[47] F.-P.J. Martin, S. Collino, S. Rezzi, S. Kochhar, Metabolomic applications to decipher gut microbial metabolic influence in health and disease, Front. Physiol. 3 (2012) 113.

[48] J.T. Bjerrum, O.H. Nielsen, F. Hao, H. Tang, J.K. Nicholson, Y. Wang, et al., Metabonomics in ulcerative colitis: diagnostics, biomarker identification, and insight into the pathophysiology, J. Proteome Res. 9 (2010) 954–62.

[49] J. van Duynhoven, E.E. Vaughan, F. van Dorsten, V. Gomez-Roldan, R. de Vos, J. Vervoort, et al., Interactions of black tea polyphenols with human gut microbiota: implications for gut and cardiovascular health, Am. J. Clin. Nutr. 98 (2013) 1631S–1641S.

[50] G.D. Lewis, A. Asnani, R.E. Gerszten, Application of Metabolomics to Cardiovascular Biomarker and Pathway Discovery, J. Am. Coll. Cardiol. 52 (2008) 117–123.

[51] C.G. Vergouw, L.L. Botros, P. Roos, J.W. Lens, R. Schats, P.G.A. Hompes, et al., Metabolomic profiling by near-infrared spectroscopy as a

tool to assess embryo viability: a novel, non-invasive method for embryo selection, Hum. Reprod. 23 (2008) 1499–504.

[52] A. Floegel, N. Stefan, Z. Yu, K. Mühlenbruch, D. Drogan, H.-G. Joost, et al., Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach, Diabetes. 62 (2013) 639–48.

[53] J.L. Spratlin, N.J. Serkova, S.G. Eckhardt, Clinical applications of metabolomics in oncology: a review, Clin. Cancer Res. 15 (2009) 431–40.

[54] T.M. O'Connell, Recent advances in metabolomics in oncology, Bioanalysis. 4 (2012) 431–51.

[55] M.A. Akanmu, S.O. Adeosun, O.R. Ilesanmi, Neuropharmacological effects of oleamide in male and female mice, Behav. Brain Res. 182 (2007) 88–94.

[56] R. Kaddurah-Daouk, K.R.R. Krishnan, Metabolomics: a global biochemical approach to the study of central nervous system diseases, Neuropsychopharmacology. 34 (2009) 173–86.

[57] E. Trushina, M.M. Mielke, Recent advances in the application of metabolomics to Alzheimer's Disease, Biochim. Biophys. Acta. 1842 (2014) 1232–9.

[58] C.B. Newgard, J. An, J.R. Bain, M.J. Muehlbauer, R.D. Stevens, L.F. Lien, et al., A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance, Cell Metab. 9 (2009) 311–26.

[59] S.J. Fowler, M. Basanta-Sanchez, Y. Xu, R. Goodacre, P.M. Dark, Surveillance for lower airway pathogens in mechanically ventilated patients by metabolomic analysis of exhaled breath: a case-control study, Thorax. 70 (2015) 320–5.

[60] M. Yu, F.-X. Cui, H.-M. Jia, C. Zhou, Y. Yang, H.-W. Zhang, et al., Aberrant purine metabolism in allergic asthma revealed by plasma metabolomics, J. Pharm. Biomed. Anal. 120 (2016) 181–9.

[61] S. Schleidgen, C. Klingler, T. Bertram, W.H. Rogowski, G. Marckmann, What is personalized medicine: sharpening a vague term based on a systematic literature review, BMC Med. Ethics. 14 (2013) 55.

[62] Biomarkers and surrogate endpoints: preferred definitions and conceptual framework, Clin. Pharmacol. Ther. 69 (2001) 89-95.

[63] U. Pinder, Epiphanie medicorum, Speculum videndi urinas hominum, Clavis aperiendi portas pulsuum, Berillus discernendi causas et differentias febrium, 1506.

[64] R.S. Vasan, Biomarkers of cardiovascular disease: molecular basis and practical considerations, Circulation. 113 (2006) 2335-62.

[65] R. Rousseau, B. Govaerts, Comparison of some chemometric tools for metabonomics biomarker identification, Intell. Lab. 91 (2008) 54-66.

[66] P.Y. Muller, F. Dieterle, Tissue-specific, non-invasive toxicity biomarkers: translation from preclinical safety assessment to clinical safety monitoring, Expert Opin. Drug Metab. Toxicol. 5 (2009) 1023-38.

[67] J.A. Wagner, S.A. Williams, C.J. Webster, Biomarkers and surrogate end points for fit-for-purpose development and regulatory evaluation of new drugs, Clin. Pharmacol. Ther. 81 (2007) 104-7.

[68] G.J. Burckart, S. Amur, F.M. Goodsaid, L.J. Lesko, F.W. Frueh, S.-M. Huang, et al., Qualification of biomarkers for drug development in organ transplantation, Am. J. Transplant. 8 (2008) 267-70.

[69] J.W. Lee, R.S. Weiner, J.M. Sailstad, R.R. Bowsher, D.W. Knuth, P.J. O'Brien, et al., Method validation and measurement of biomarkers in nonclinical and clinical samples in drug development: a conference report, Pharm. Res. 22 (2005) 499-511.

[70] CFR - Code of Federal Regulations Title 21, (n.d.).

[71] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA. Cancer J. Clin. 61 (2011) 69-90.

[72] 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.

[73] D.E. Brenner, D.P. Normolle, Biomarkers for cancer risk, early detection, and prognosis: the validation conundrum, Cancer Epidemiol. Biomarkers Prev. 16 (2007) 1918-1920.

[74] M. Castro, L. Grau, P. Puerta, L. Gimenez, J. Venditti, S. Quadrelli, et al., Multiplexed methylation profiles of tumor suppressor genes and clinical outcome in lung cancer, J. Transl. Med. 8 (2010) 86.

[75] K.L. Richards, B. Zhang, M. Sun, W. Dong, J. Churchill, L.L. Bachinski, et al., Methylation of the candidate biomarker TCF21 is very frequent across a spectrum of early-stage nonsmall cell lung cancers, Cancer. 117 (2011) 606–617.

[76] C. Kneip, B. Schmidt, A. Seegebarth, S. Weickmann, M. Fleischhacker, V. Liebenberg, et al., SHOX2 DNA methylation is a biomarker for the diagnosis of lung cancer in plasma, J. Thorac. Oncol. 6 (2011) 1632–1638.

[77] P.P. Massion, Y. Zou, H. Uner, P. Kiatsimkul, H.J. Wolf, A.E. Baron, et al., Recurrent genomic gains in preinvasive lesions as a biomarker of risk for lung cancer, PLoS One. 4 (2009) e5611.

[78] T. Blomquist, E.L. Crawford, D. Mullins, Y. Yoon, D.-A. Hernandez, S. Khuder, et al., Pattern of antioxidant and DNA repair gene expression in normal airway epithelium associated with lung cancer diagnosis, Cancer Res. 69 (2009) 8629–8635.

[79] M.D. Wilkerson, X. Yin, K.A. Hoadley, Y. Liu, M.C. Hayward, C.R. Cabanski, et al., Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types, Clin. Cancer Res. 16 (2010) 4864–4875.

[80] N. Yanaihara, N. Caplen, E. Bowman, M. Seike, K. Kumamoto, M. Yi, et al., Unique microRNA molecular profiles in lung cancer diagnosis and prognosis, Cancer Cell. 9 (2006) 189–198.

[81] K. Yanagisawa, S. Tomida, Y. Shimada, Y. Yatabe, T. Mitsudomi, T. Takahashi, A 25-signal proteomic signature and outcome for patients with resected non-small-cell lung cancer, J. Natl. Cancer Inst. 99 (2007) 858–867.

[82] S.M.J. Rahman, A.L. Gonzalez, M. Li, E.H. Seeley, L.J. Zimmerman, X.J. Zhang, et al., Lung cancer diagnosis from proteomic analysis of preinvasive lesions, Cancer Res. 71 (2011) 3009–3017.

[83] K. Kami, T. Fujimori, H. Sato, M. Sato, H. Yamamoto, Y. Ohashi, et al., Metabolomic profiling of lung and prostate tumor tissues by capillary electrophoresis time-of-flight mass spectrometry, Metabolomics. 9 (2013) 444–453.

[84] S. Hanash, Harnessing immunity for cancer marker discovery, Nat. Biotechnol. 21 (2003) 37–38.

[85] S.M. Hanash, S.J. Pitteri, V.M. Faca, Mining the plasma proteome for cancer biomarkers, Nature. 452 (2008) 571–579.

[86] S. Hori, S. Nishiumi, K. Kobayashi, M. Shinohara, Y. Hatakeyama, Y. Kotani, et al., A metabolomic approach to lung cancer, Lung Cancer. 74 (2011) 284–92.

[87] 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.

[88] 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).

[89] A. Peralbo-Molina, M. Calderón-Santiago, F. Priego-Capote, B. Jurado-Gámez, M.D. Luque de Castro, Development of a method for metabolomic analysis of human exhaled breath condensate by gas chromatography-mass spectrometry in high resolution mode, Anal. Chim. Acta. 887 (2015) 118–126.

[90] A. Peralbo-Molina, M. Calderón-Santiago, F. Priego-Capote, B. Jurado-Gámez, M.D. Luque de Castro, Metabolomics analysis of exhaled breath condensate for discrimination between lung cancer patients and risk factor individuals, J. Breath Res. 10 (2016) 016011.

[91] J. Hunt, Exhaled breath condensate: An evolving tool for noninvasive evaluation of lung disease, J. Allergy Clin. Immunol. 110 (2002) 28–34.

[92] M.Y.L. and P.S. Thomas, Biomarkers in exhaled breath condensate and serum of chronic obstructive pulmonary disease and non-small-cell lung cancer, Int. J. Chronic Dis. 2013 (2013) 15.

[93] P.J. Mazzone, Analysis of volatile organic compounds in the exhaled breath for the diagnosis of lung cancer, (2008) 774–780.

[94] P. Fuchs, C. Loeseken, J.K. Schubert, W. Miekisch, Breath gas aldehydes as biomarkers of lung cancer, Int. J. Cancer. 126 (2010) 2663–70.

[95] A. Bajtarevic, C. Ager, M. Pienz, M. Klieber, K. Schwarz, M. Ligor, et al., Noninvasive detection of lung cancer by analysis of exhaled breath, BMC Cancer. 9 (2009) 348.

[96] M. Ligor, T. Ligor, A. Bajtarevic, C. Ager, M. Pienz, M. Klieber, et al., Determination of volatile organic compounds in exhaled breath of patients with lung cancer using solid phase microextraction and gas chromatography mass spectrometry, Clin. Chem. Lab. Med. 47 (2009) 550–60.

[97] G. Peng, M. Hakim, Y.Y. Broza, S. Billan, R. Abdah-Bortnyak, A. Kuten, et al., Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors, Br. J. Cancer. 103 (2010) 542–51.

[98] V.S. Kurova, A.S. Kononikhin, D.A. Sakharov, I.A. Popov, I.M. Larina, A.G. Tonevitskiĭ, et al., [Exogenic proteins in the human exhaled breath condensate], Bioorg. Khim. 37 55–60.

[99] C. Gessner, B. Rechner, S. Hammerschmidt, H. Kuhn, G. Hoheisel, U. Sack, et al., Angiogenic markers in breath condensate identify non-small cell lung cancer, Lung Cancer. 68 (2010) 177–84.

HERRAMIENTAS Y EQUIPOS ANALÍTICOS

En este apartado de la Memoria se describen someramente los diferentes tipos de muestras que han sido objeto de investigación, así como los instrumentos y aparatos usados durante el desarrollo experimental de la Tesis. En los diferentes capítulos se incluye una explicación más detallada de los que se han utilizado en la investigación recogida en cada uno de ellos.

1. Muestras

Para el desarrollo de la investigación que se recoge en la Parte A se han utilizado dos tipos de muestras; desechos de desechos, resultantes de la producción de etanol a partir de los residuos de la industria vinícola, como hollejos y pepitas, y desechos de la industria aceitera, como el alperujo.

En los estudios de comparación de métodos de extracción y caracterización de fenoles (Capítulos III y IV, respectivamente) se utilizaron hollejos y pepitas de uvas blancas y tintas provenientes de la planta industrial ‘‘Alcoholeras Reunidas, S.A.’’ (Ciudad Real). En el último capítulo de esta sección se usó alperujo proporcionado por “Núñez de Prado, C.B.” (de Baena, Córdoba), producido en la cosecha 2009/2010.

Para la sección B se ha utilizado un único tipo de muestra: aire exhalado condensado procedente de individuos sanos, fumadores o no, y pacientes de cáncer de pulmón.

2. Sistemas no automáticos para la preparación de muestra

Los sistemas de tipo discontinuo utilizados para la preparación de muestra en la sección A fueron los siguientes:

Dispositivo de microondas. Se utilizó un digestor comercial Microdigest 301 fabricado por Prolabo, basado en microondas focalizadas, para la extracción de los compuestos fenólicos de los hollejos y las pepitas (Capítulos III y IV).

Dispositivo de ultrasonidos. La sonda de ultrasonidos utilizada fue una Branson 450 digital, que permite la selección de la potencia de los ultrasonidos, así como el modo de aplicación, continuo o discontinuo (el llamado ciclo útil). Esta sonda se utilizó para acelerar la extracción de los compuestos fenólicos de los hollejos y de las pepitas. Para ello la sonda se introdujo en el recipiente que contenía la muestra sólida y el extractante, tal como se detalla en los Capítulos III y IV.

Equipo de extracción mediante líquidos sobrecalentados. La extracción con mezclas etanol-agua a temperaturas altas y presión suficiente para mantener el extractante en estado líquido se llevó a cabo con un extractor diseñado y construido en el laboratorio con los siguientes elementos: un horno, una bomba de alta presión (Hitachi L-6200A) para impulsar el extractante a través del sistema, una cámara de extracción de acero inoxidable, una válvula de alta presión para permitir el paso de extractante y llenar el sistema y para purgarlo una vez terminada la extracción, un restrictor (para mantener la presión en el sistema) y tubo de acero inoxidable para construir las zonas de transporte y de calentamiento previo del extractante. Una descripción más detallada del extractor se encuentra en los Capítulos III, IV y V. El extractor se utilizó para la obtención de extractos polares y de polaridad media a partir de hollejos y pepitas procedentes de la industria vitivinícola, y de alperujo, procedente de la industria oleícola (Capítulos III a V).

Molino de bolas. Se utilizó un molino Restch MM301 para triturar los hollejos y las pepitas de procedentes de la industria vitivinícola hasta conseguir un tamaño homogéneo de partícula de 40 mesh (0.42 mm de diámetro) (Capítulos III y IV).

Agitador eléctrico. Para favorecer la extracción sólido–líquido en el análisis de extractos de hollejos y pepitas de la industria vitivinícola, así como en

la extracción de compuestos del aliento condensado, se utilizó un agitador eléctrico MS2 minishaker (Capítulos III, IV, VI, VII, VIII, IX).

Los sistemas sólido-líquido tratados con los dispositivos anteriores se centrifugaron mediante una Centrífuga Selecta Mixtasel-BL para separar el residuo sólido del extracto.

En el desarrollo de la investigación que recoge la Parte B, y dada la naturaleza no selectiva, semi- o no cuantitativa de las plataformas analíticas empleadas para análisis metabólico global, se hizo especial hincapié en el desarrollo de métodos con mínima o nula preparación de muestra, evitándose, por tanto, los sistemas continuos para este tratamiento. En general, la preparación de la muestra se basó en extracción líquido-líquido o extracción en fase sólida (SPE) usando μ-SpinColumn (un sistema de SPE diseñado para pequeños volúmenes de muestra y en el que la fuerza centrífuga posibilita el paso del líquido a través del cartucho).

3. Sistemas de separación y de detección utilizados

Los métodos desarrollados en la parte experimental de esta Tesis Doctoral se han basado en una separación cromatográfica (mediante LC o GC) previa a la detección mediante un espectrómetro de masas.

En las plataformas dedicadas a la obtención del perfil metabolómico (metabolomics profiling) de muestras de hollejos y pepitas (Capítulos III y IV) y alperujo (Capítulo V), se utilizó un equipo HPLC Agilent 1200 Series acoplado a un detector de masas de tiempo de vuelo de alta resolución, Agilent 6540. Las columnas para la separación cromatográfica fueron C18 (fase reversa) En todos los casos se usó el software MassHunter para la adquisición de espectros y el análisis cualitativo.

En los Capítulos VI-IX, dedicados a estudios del perfil metabolómico (metabolomics profiling) del aliento condensado, se empleó un cromatógrafo de

gases (GC) Agilent 7890A, equipado con un automuestreador, un inyector con/sin división del flujo (split/splitless) y un detector de masas de tiempo de vuelo de alta resolución, Agilent 7200. La columna analítica fue de tipo DB-5MS-UI (30 m × 0.25 mm i.d., 0.25 µm). En todos los casos se usó el software MassHunter para la adquisición de espectros y el análisis cualitativo.

4. Técnicas quimiométricas

De acuerdo con la importancia que ha adquirido la quimiometría en metabolómica, en la investigación que se recoge en esta Memoria se han utilizado extensamente herramientas quimiométricas, tanto para el desarrollo y optimización de métodos analíticos como para el tratamiento de datos multivariantes.

El tratamiento de datos se realizó con distintos programas informáticos según el objetivo:

a) Alineamiento de entidades moleculares en análisis no orientado utilizando un cromatógrafo de líquidos acoplado a un detector de tiempo de vuelo (profiling):

   - Por un lado se usó la combinación de dos paquetes del lenguaje de programación R de uso libre: XCMS y CAMERA. El primero permite extraer y alinear las entidades moleculares potenciales de un análisis no orientado, y el segundo reconoce aductos e isótopos y los agrupa para eliminar falsos positivos y crear la lista definitiva de entidades.

   - Por otro lado se empleó la combinación de dos softwares de Agilent: Qualitative Workstation y MassProfiler Professional (Agilent). El primero permite extraer las entidades teniendo en cuenta aductos e isótopos, mientras que el segundo posibilita el alineamiento de las entidades potenciales.

b) Alineamiento de entidades moleculares en análisis no orientado utilizando un cromatógrafo de gases acoplado a un detector de tiempo de vuelo (profiling):

   - Por un lado se usó la combinación de tres softwares de Agilent: Unknown Analysis, MassProfiler Professional y Quantitative Workstation. El primero permite extraer las entidades moleculares potenciales de un análisis no orientado, y el segundo permite alinearlas, restar la señal del blanco y aplicar diferentes tipos de filtro para crear la lista definitiva de entidades. Una vez creada esta lista, el tercer software se utiliza para re-extraer las entidades seleccionadas en todas las muestras (análisis recursivo).

c) Análisis estadístico:

   - Statgraphics: software que permite realizar distintos análisis estadísticos univariantes y multivariantes, así como el diseño y evaluación de modelos de cribado (screening) y superficies de respuesta.

   - MassProfiler Professional: permite la aplicación de diferentes algoritmos de análisis estadístico especialmente adecuados para el análisis metabolómico.

d) Evaluación de la capacidad de predicción de metabolitos o paneles con potencial marcador: esta evaluación se ha realizado a través de las curvas ROC, que representan especificidad frente a sensibilidad. Para su obtención se han usado las herramientas gratuitas de acceso online ROCCET (http://www.roccet.ca/ROCCET/), y METABOANALYST (http://www.metaboanalyst.ca/).

e) Diseño de modelos de predicción y creación de paneles de marcadores: en este caso se han empleado tanto el MassProfiler Professional, que permite la creación de modelos por PLS-DA, como las herramientas en línea ROCCET y METABOANALYST o el software PanelomiX, que

permiten crear paneles de marcadores y evaluar su capacidad de predicción/discriminación.

5. Bases de datos

Aunque queda un largo camino por recorrer, existen varias bases de datos de metabolitos a disposición del usuario que contienen información para la identificación y caracterización de muchos de los compuestos presentes en diversas matrices biológicas. Entre ellas, las bases de datos 'Metabolites and Tandem MS Database' (METLIN), 'Human Metabolome Database' (HMDB) y 'MassBank' se han empleado para la identificación de compuestos presentes en biofluidos (aliento condensado) y en los extractos obtenidos a partir de residuos de la industria agroalimentaria (hollejos, pepitas y alperujo) a partir de los datos espectrales obtenidos mediante espectrometría de masas en los Capítulos III, IV, V y VI de esta Memoria. Además, la información biológica disponible en la HMDB y las rutas biosintéticas de la base de datos 'Kyoto Encyclopedia of Genes and Genomes' (KEGG) se han utilizado para la interpretación de resultados obtenidos en los Capítulos VI, VII, VIII y IX.

PARTE EXPERIMENTAL

PARTE A

Aportaciones al aprovechamiento de residuos: extracción de compuestos de alto valor añadido e identificación de compuestos de desechos

Esta parte A de la Tesis abarca, en primer lugar, la investigación sobre los estudios previos publicados como punto de apoyo de la realización de investigación experimental para aportar nuevos conocimientos sobre las materias objeto de esta parte A, el aprovechamiento de residuos, a lo que se contribuye con: (i) nuevos métodos para la extracción de compuestos de interés existentes en ellos y la adecuación de cada método a cada uno de los compuestos o una familia de ellos; (ii) la identificación de los compuestos existentes en los extractos como aspecto clave para su mejor explotación.

La investigación sobre los estudios previos dio lugar a una revisión bibliográfica en la que se abarcó el potencial de los residuos más destacables de la agricultura y la industria agroalimentaria mediterráneas (vid/vino y olivo/aceite), sobre la cual se soportaron los trabajos experimentales en esta área. Esta investigación también permitió escribir un capítulo de libro sobre uvas, considerando su producción, su composición fenólica y sus aplicaciones biomédicas. Estos estudios previos, que sitúan el estado actual en esta área, constituyen los Capítulos I y II, respectivamente, de esta Tesis.

Los dos primeros estudios experimentales se dedicaron a establecer las posibilidades de aprovechamiento de un residuo de residuo: el que resulta de la producción de etanol a partir de los residuos de la industria vinícola. Se utilizaron separadamente los hollejos y las pepitas de uvas blancas y tintas para estudiar las posibilidades de estas materias, muy diferentes. Utilizando los hollejos se compararon métodos de extracción de compuestos polares —principalmente fenoles— asistidos por diferentes energías para favorecer el proceso: ultrasonidos, microondas y presión+temperatura altas. En los extractos se estudiaron parámetros genéricos como el contenido total de fenoles y la capacidad antioxidante, además de llevar a cabo su caracterización tentativa mediante cromatografía de líquidos-espectrometría de masas de alta resolución con detector de tiempo de vuelo (LC-TOF/MS). Los componentes polares de las pepitas se extrajeron con los mismos métodos del caso anterior y los extractos fueron igualmente caracterizados en cuanto a contenido total de fenoles, capacidad antioxidante y composición fenólica mediante LC-TOF/MS. Los resultados,

recogidos en los Capítulos III y IV, respectivamente, muestran que ambas materias primas son fuentes de compuestos de interés para las industrias alimentaria y nutracéutica.

La investigación que recoge el último capítulo de esta parte A (Capítulo V) se dedicó al alperujo, el residuo resultante de la producción de aceite de oliva mediante el sistema de dos fases. Se estudió en este residuo, hasta ahora indeseable, la presencia de los componentes minoritarios del aceite de oliva; aquéllos que confieren a este aceite sus especiales características (hidroxitirosol y derivados, secoiridoides, sus precursores y sus derivados, flavonoides, lignanos y ácidos fenólicos), identificándolos, estableciendo sus patrones de fragmentación mediante espectrometría de masas en tándem y discutiendo sus propiedades beneficiosas para la salud.

Capítulo I:

Potential of residues from the Mediterranean agriculture and agrifood industry

Trends in Food Science and Technology

Volume 32 (1), July 2013

16-24

Potential of residues from the Mediterranean agriculture and agrifood industry

Ángela Peralbo Molina[1-3], María Dolores Luque de Castro[1-3]

[1]Department of Analytical Chemistry, Annex C-3, Campus of Rabanales, 14014 Córdoba, Spain

[2]University of Córdoba Agroalimentary Excellence Campus, ceiA3, Campus of Rabanales, 14014 Córdoba, Spain

[3]Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital,

University of Córdoba, E-14014, Córdoba, Spain

All the authors belong to the 3 institutions.

Potential of residues from the Mediterranean agriculture and agrifood industry

Ángela Peralbo Molina, María Dolores Luque de Castro

Abstract

The most abundant agrifood residues in the Mediterranean basin derive from the two traditional crops: olive trees and vineyards. They are vine shoots and leaves (both fresh and senescent leaves) lees, grape skins, seeds and stalk, on the one hand; olive tree leaves and olive pomace, on the other. The potential of these residues as a powerful source of well-known phenols with antioxidant properties has been proved; nevertheless, there are a number of other unknown phenols together with other families of very valuable compounds that should be studied. This article constitutes a call of attention on this unexploited wealth.

Keywords: phenolic compounds, olive oil residues, wine residues, extraction, exploitation of residues

1. Introduction

Traditionally, most crops have been focused on obtaining a single high-value product (e.g. oil, wine, cider, horchata) discarding the remaining parts of the plant or fruit, or exploiting them to obtain low value products (sometimes highly polluting and, therefore, undesirable products as is the case with residues from the oil industry). A rational use of unwanted, useless or low valuable residues would create a source of additional wealth, which in many cases could have a value even higher than that traditionally given to the only reason of the target crop, thus presenting interesting perspectives for creation of new industries.

Implementation of policies for exploitation of residues from the agrifood industry in the Mediterranean basin should by the two traditional cultures of this region: vineyards and olive trees. In both cases residues from the plant and the corresponding industry contain compounds of interest to the pharmacological, nutraceutical, cosmetics, and food industries.

Vineyards produce vast amounts of agricultural residues (particularly vine shoots and leaves). According to Jiménez *et al.* (2006), 1.4–2.0 tons of shoots can be obtained per hectare of vine crop. Since the vineyard cultivated area in the world is around 8 million ha, an estimated total of 11.2–16 million tons of vine shoots are produced each year. On the other hand, the most abundant residue of the winemaking process is the solid waste remaining after grape pressing: the grape pomace. Since 20%–25% of the weight of processed grape remains as pomace (Kammerer, Gajdos, Carle & Schieber, 2005) and about 80% of the world grape output is used to make wine, around 10.5–13.1 million tons of grape pomace are produced in the world each year.

The olive oil industry gives also place to vast amounts of residues from the production line. More than 8 million ha of olive trees are cultivated worldwide, especially in the Mediterranean basin. It has been estimated that an average of three tons of pruning biomass is obtained each year from one olive tree hectare,

making these residues a huge, cheap, and unexploited source of energy or chemicals (Conde *et al.*, 2009). Depending on the particular technology used for oil production, mill waste or olive pomace is constituted by alpechín (an aqueous liquor coming from the plant water and soft tissues of the olive fruit plus water added during processing) plus orujo (a solid waste consisting of olive pulp and stones) or alperujo (an emulsion of alpechín and orujo) as a result of applying the three- or the two-phase process, respectively. An indirect residue of olive oil production consists of leaves accompanying the olive fruits to the mill. Both olive leaves and olive pomace contain substantial amounts of phenols, which endow them with a high added value.

Although some of these residues start to be poorly managed, higher benefits should (and must) be obtained from these materials by adequate information, planning, and development; with the involvement of both professionals and the appropriate institutional organisms in each case. Actually, exploitation of residues could encompass three gradational states, namely:

(1) Residues from which compounds of high added value have been recognized as such by the European Food Safety Authority (EFSA), as is the case with hydroxytyrosol, extracted from either olive leaves or oil pomace. This compound is claimed to reduce an emerging risk factor of atherosclerosis (oxidized LDL cholesterol) (Authority, EFSA, 2011) and maintain normal blood HDL-cholesterol levels (Authority, EFSA, 2012).

(2) Products from the residues in (1), the beneficial effects of which are not enough supported at present for EFSA acceptance, but widely used in the food and cosmetics fields, as is the case of oleuropein.

(3) Residues, the potential of which has not been (or has been poorly) recognized and exploited.

There is a growing demand for exploitation of agricultural residues. One of them is the present pressing need to obtain natural dyes able to substitute existing artificial dyes, particularly in the food industry. This time pressure makes

desirable to obtain the demanded products from raw materials traditionally used as foods, thus eliminating the long research and experimentation period for regulation of materials that have not constituted a part of the human diet. Cases of long tradition as food are olives and grapes; therefore, the colorants obtained from their residues should not be subjected to legal delays in their use. The need for more natural colorants began with a study of the University of Southampton, in late 2009, which reported an experiment with 6 artificial colorants that caused hyperactivity and allergy on children subjected to the test (Stevenson, 2009). Subsequently, the EFSA conducted a toxicological study with the 6 colorants, after which the organism ruled that the target colorants did not show hyperactive, allergens, or other toxic effects (Authority, EFSA, 2010). Nevertheless, the DG SANCO of the European Parliament decided, as a precautionary measure, to include in the labelling of foods containing the target colorants a sentence about the possibility that they "can produce actions and negative effects on children's activity" (www.cen-online.org, June 9 (2008) 32, January 21 (2008) 33). This situation has pushed the search for, and application of, natural colorants to replace synthetic colorants (www.cen-online.org, March 24 (2008) 28, December 15 (2008) 18).

The aim of this review was to show the importance of residues from both the crops and products of the two major industries in the Mediterranean basin (wine and olive oil), also characteristic of a number of other regions and countries. The exploitation of these residues can provide lucrative benefits to the agrifood industry thanks to (among others) the high content in antioxidants with proved healthy, cosmetics and nutritional properties, in addition to the reduction of pollution from accumulation of unexploited residues.

2. Potential exploitation of residues from olive trees and olive oil production

Phenolic compounds are among the most conflictive fractions of residues from olive oil production (Priego-Capote, Ruiz-Jiménez & Luque de Castro, 2004) as they are the major contributors to the antibacterial and phytotoxic activity of black olive residues by limiting their microbial degradability and potential for biogas production (Montero, Miranda, Arranz & Rojas, 2011). Therefore, removal of phenols from the residue of oil production involves two positive aspects: obtainment of high-added value products and a secondary residue which can be either used for biogas production or easily degraded to compost.

A number of olive phenols, including oleuropein (the most abundant phenol in olive leaves), hydroxytyrosol (the most bioactive olive phenol), tyrosol, verbascoside, apigenin-7-glucoside and luteolin-7-glucoside are present in variable amounts in the different parts of olive trees, in olive oil and oil production residues (Fig.1). As a general rule, the o-dihydroxy structure characteristic of hydroxytyrosol and other secoiridoid derivatives seems to be responsible for the high antioxidant activity of these phenols, followed by 4-O-monohydroxy compounds (ligstroside and tyrosol) and 3-O-hydroxy-substituted catechols (Torres de Pinedo, Peñalver & Morales, 2007). A number of published studies has reported the beneficial effects of olive phenols on human health, the most salient being antioxidative (Cioffi *et al.*, 2010), antimicrobial (Obied, Bedgood, Prenzler & Robards, 2007), antiviral, anti-ischemic, anti-inflammatory and hypolipidemic effects (Petty & Scully, 2009). Some of them have also shown to be cardioprotective (Covas, 2007), neuroprotective (Omar, 2010), and preventtive of tumoral diseases –in addition to be endowed with scavenger and antioxidant properties higher than typical antioxidants such as vitamins C and E or 2,6-di-tert-butyl-4-methylphenol (BHT). Besides, it is known that flavones possess healthy effects, as is the case with apigenin-7-glucoside, used to fight against Alzheimer (Ansari *et al.* 2009) and liver diseases (Oh, Kim, Cho & Kim, 2004); and luteolin-7-glucoside that avoids the abnormal proliferation of aortic vascular smooth

muscle cells, a frequent cause of pathogenesis such as atherosclerosis and restenosis (Kim, Kim, Jin & Yun, 2006). These proved effects of olive phenols may turn olive residues into a cheap source of natural antioxidants to be used as pharmacological principles, nutraceuticals or food supplements.

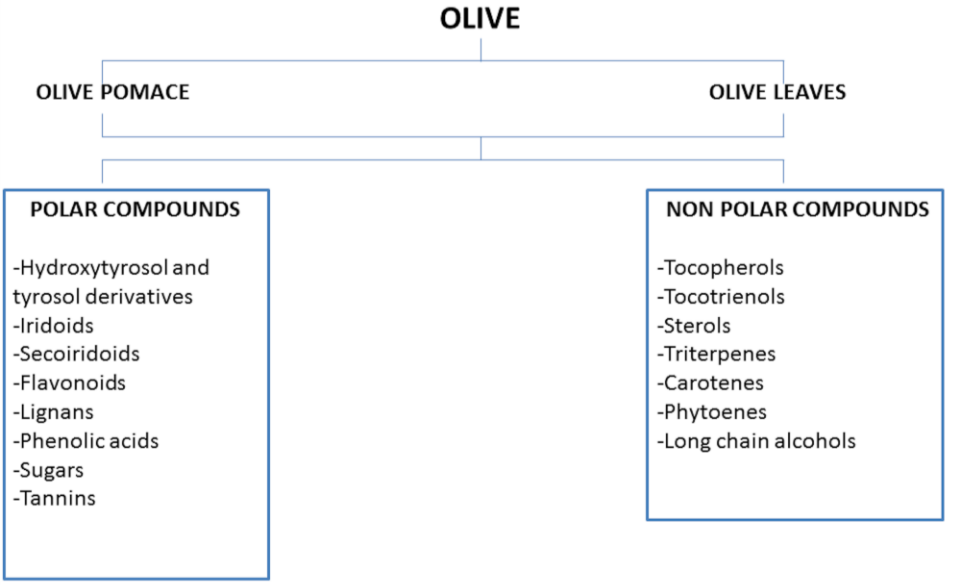


Figure 1. Valuable compounds obtained from olive residues.

2.1. Olive pomace as a source of high-added value compounds

After virgin olive oil (VOO) decantation in the two-phase system, vast amounts of olive pomace remain as a very undesirable residue. As no (or little amount of) water is added during this process, the quality of the resulting residue is higher than that from the three-phase system since the latter produces two different residues the volume of their sum being increased by the amount of added water (Authority, EFSA, part II of III, 2012). Once VOO is decanted, additional centrifugation steps or a second extraction process (beating and decanting) on the remaining olive pomace leads to second order olive oils characterized by high wax content. The isolation of these low-quality oils is economically viable; therefore, this strategy is implemented in most olive oil

factories (Lozano-Sánchez *et al.*, 2011). This second extraction step has decreased the interest for olive pomace oil isolated with the aid of organic solvents and heat (refined olive oil). Currently, olive pomace from this second extraction is used either as fuel in cogeneration plants to produce energy or as organic fertilizer after a composting process enlarged by the presence of antioxidant compounds (Sampedro *et al.*, 2004). However, the profitability of cogeneration plants is doubtful, and the fertilization industry only absorbs a very small proportion of the produced olive pomace. For these reasons, a frequent option is dumping olive pomace in controlled areas owing to its high polluting organic load. Therefore, it is clear that there is a demand for alternatives capable of obtaining benefit from olive pomace (either before or after extraction of the low-quality oil), which in addition to phenols is characterized by a high content of other organic substances such as sugars, tannins, polyalcohols and lipids (Lesage-Messen *et al.*, 2001).

Research for valorization of olive pomace has so far been mainly focused on panels or target known compounds considered of interest such as hydroxytyrosol (Manna *et al.*, 1999), tyrosol (Mazzoti *et al.*, 2012) or oleuropein (Luque de Castro & Priego-Capote, 2010), but global characterization had not been made. Therefore, a global profiling of compounds in pomace would be of interest both to better understand their commercial potential and know the biochemical pathways in which they are involved for proper modification if required. Similar studies on other olive mill residues such as wastewater have been reported (Herrero *et al.*, 2011). A global metabolomics profile of phenols obtained from olive pomace by subjecting the extract to liquid chromatography separation and tandem mass spectrometry detection using a QqTOF hybrid analyzer to take benefit from high mass accuracy has allowed identification of 53 compounds belonging to different families (Peralbo-Molina, Priego-Capote & Luque de Castro, 2012). Hydroxytyrosol and tyrosol derivatives, iridoids precursors, secoiridoids and derivatives, flavonoids, lignans, phenolic acids and tocopherols are some of the compounds present in the extracts from olive oil production residues.

In addition to phenols, interesting non-polar compounds in olive pomace are tocopherols, tocotrienols, sterols, triterpenes, carotenes, phytoenes or long chain alcohols (Phillips *et al.*, 2002); nevertheless, a comprehensive characterization of this fraction has not been performed yet. Long chain alcohols (Fernández-Arche *et al.*, 2009) and pentacyclic triterpenes (Rodríguez-Rodríguez & Ruíz-Gutiérrez, 2010) have so far been the only studied families. The reason for the practically absence of research on this fraction of olive pomace could be the foreseeable low content of these compounds in pomace taking into account that their partition coefficient is more favorable to their permanence in oil. A bioactive compound appreciated in food, pharmaceutical and cosmetics industries that could be obtained from this raw material is squalene. The interest in establishing novel sources to obtain squalene is increasing as the conventional ones are rather limited and the demand for this compound is growing (Ki & Karadeniz, 2012).

Finally, an unexploited area is that of dyes present in olive pomace. The growing tendency towards the search for low cost natural dyes (Bechtold, Mahmud-Ali & Mussak, 2007, Ali & El-Mohamedy, 2011) could find an excellent source in this raw material. Despite olive pomace looks to be a valuable potential source of abundant natural dyes, available in large quantities and low cost, few researchers have considered olive pomace an easy matter for exploitation.

A recent study focuses on valorization of olive pomace extracts as possible dye baths for dyeing wool (Meksi, Haddar, Hammami & Mhenni, 2012). More research on this subject would be of great interest because, in addition to obtaining natural dyes, this exploitation would contribute to reduce the environmental problem caused by olive pomace.

2.2. Olive leaves and their rich composition

Olive leaves are an agricultural residue resulting from the pruning of olive trees, and also an industrial residue as they find their way to the olive mill in amounts around 10% of the total weight of olives arriving to mills (Herrero *et al.*,

2011). The scientifically proven healthy properties of olive leaf extracts are leading almost on a nearly daily basis to the presentation of new patents for functional foods or cosmetics based on them (Meksi, Haddar, Hammami & Mhenni, 2012, Shtukatur, 2003, Stueckler, 1998, Amari, 1998).

Depending on the compounds to be extracted from olive leaves, methods based on conventional stirring or Soxhlet extraction with different solvents (hexane, water, ethanol and methanol) (Luque de Castro, Priego-Capote, 2010) have been reported. Nevertheless, the most interesting extraction methods, which should be taken into account for industrial applications, are those based on present technologies that accelerate, automate and improve extraction efficiency. Thus, olive leaves have been subjected to supercritical extraction with CO2 (with or without assistance of a co-extractant in the case of polar compounds) for the removal of tocopherols (Xynos *et al.*, 2012), waxes (Tabera, *et al.*, 2004), hydrocarbons, squalene, β-carotene, triglycerides, α-tocopherol, β-sitosterol and alcohols (Tabera *et al.*, 2004), carboxylic acids and phenols (Le Floch, Tena, Ríos & Valcárcel, 1998). Less costly are methods based on the use of superheated liquid extraction (SHLE) such as those reported to obtain oleuropein and related phenols using either static, dynamic or static-dynamic approaches (Japón-Luján & Luque de Castro, 2006, 2007); or those based on the use of auxiliary energies such as ultrasound (Japón-Luján, Luque-Rodríguez, Luque de Castro, 2006) or microwaves (Li *et al.*, 2011) for which ethanol-water mixtures have been used as extractant with a view in the use of the extracts to supplement food or as nutraceuticals. Separation of extracts components and purification by the use of macroporous resins has been proposed (Li *et al.*, 2011). The removal of chlorophylls and lipids is considered to be necessary to achieve pure extracts of the target compounds. Some authors have used toxic extractants like dichloromethane that complicates obtainment of a final product for human use (Mylonaki, Kiassos, Makris &Kefalas, 2008).

A comprehensive characterization of the phenol composition of olive leaf extracts obtained by SHLE has been carried out using LC-ESI-MS (Herrero *et al.*,

2011). Around 25 phenols have been tentatively identified in water and ethanol extracts, including phenolic acids, secoiridoids, hydroxycinnamic acid derivatives, flavonols and flavones. Among them, the most important phenols identified in olive trees (viz. hydroxytyrosol, oleuropein or luteolin-glucoside) were found, thus reinforcing the evidences of the high potential of this raw material that should be studied more in depth.

A global characterization of the nonpolar fraction of olive leaves has not been made yet; therefore, despite olive leaves is a raw material better studied than olive pomace, more research on both is mandatory to promote full exploitation.

3. Potential exploitation of residues from vineyard and wine production

3.1. Potential of vineyard residues

Spain is the country with the largest area in the world dedicated to vineyards, with approximately 1.1 million hectares, being the third wine-producer country, followed by France and Italy. There are two residues from vineyards which have traditionally been poorly exploited: vine shoots and leaves (Fig. 2).

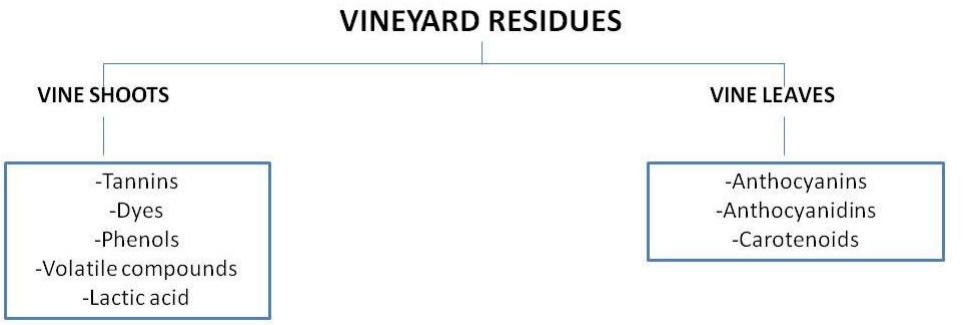


Figure 2. Valuable compounds obtained from vineyard residues.

Vine shoots and their unexploited richness. Despite vine shoots constitute an abundant agricultural residue, their present economic value is very small as they are mostly used as fuel or fertilizer (Mediavilla, Fernández & Esteban, 2009). Most of the scant research on vine shoots has been focused on the production of paper pulp and ethanol, the former requiring in-depth studies to improve production as vine shoots provide pulp of lower quality than other agricultural residues such as wheat straw (Jiménez *et al.*, 2008). Some other ways of vine shoot exploitation, such as tanning and dyeing of leather (Moldes, Torrado, Barral & Domínguez, 2007), production/extraction of phenols (Moldes, Torrado, Converti & Domínguez, 2006), volatile compounds (Luque-Rodríguez, Pérez-Juan & Luque de Castro, 2006), activated carbon for wine treatment (Max, Salgado, Cortes & Domínguez, 2010), lactic acid (Luque-Rodríguez, Luque de Castro, Japón-Luján, 2008), biosurfactants (Pan, Bolton &Leary, 1998) and ferulic and coumaric acids (Morales, Benítez & Troncoso, 2004); and smoke flavors (Bautista-Orín *et al.*, 2008) have been investigated without emphasis.

The composition of vine shoots is characterized by three main fractions: cellulose, hemicellulose, and lignin; the content of total cellulose being around 68% and that of lignin around 20% (dry weight). Lignin, a well-known component of secondary cell walls, is a high molecular mass cross-linked polymer, which is built up by random oxidative coupling of three major C6−C3 (phenylpropanoid) units (monolignols) due to the lack of enzymatic control. These units [namely, trans-p-coumaryl (4- hydroxycinnamyl), coniferyl (4-hydroxy-3-methoxycinnamyl, forming guayacyl units), and sinapyl(3,5-dimethoxy-4-hydroxycinnamyl, forming syringyl units)] are characterized by a phenolic structure. Lignin can be hydrolyzed to release aromatic phenolic compounds such as low molecular mass alcohols, aldehydes, ketones, or acids; therefore, vine shoots can be a suitable source of phenols. The abundance and richness of vine shoots make their exploitation highly interesting in economic terms as this raw material can be very useful to obtain products of a high-added value in the nutraceuticals, cosmetics, pharmacological, oenological, and food additive industries. Phenol extracts from vine shoots have proved their effectiveness in animals by reducing proliferation of

leukemic cells (Cerdá, Rodríguez & Ancn, 2002), against epilepsy (Arapitsas, Antonopoulos, Stefanou & Dourtoglou, 2004), and for prevention of aging and diseases such as atherosclerosis, diabetes, and inflammatory processes (Fernández de Simón, Cadahía, del Álamo & Nevares, 2010) as a result of the antioxidant properties of phenols and their ability to act as efficient free radical scavengers.

Research on this subject has involved three different extraction protocols (based on microwave-assisted extraction –MAE–, ultrasound-assisted extraction –USAE– and superheated liquid extraction –SHLE) for isolation of interesting compounds from vine shoots (Delgado-Torre, Ferreiro-Vera, Priego-Capote & Luque de Castro, 2012). After selection of SHLE (the comparison was estimated by the Folin-Ciocalteu test) as the most suitable approach for isolation of phenols, it was applied to 18 vine shoot cultivars. The purpose was to compare the antioxidant potential of the different cultivars (using the ORAC assay) and the content of interesting compounds from the enological and nutraceutical points of view. Compounds like ferulic acid, pyrocatechol, protocatechuic acid, pyrogallol, guaiacol, and vanillic acid, among others, were identified by LC-TOF/MS. One of the most characteristic phenols found in the extracts from all studied cultivars was gallic acid, a final product from the hydrolysis of elagitanning which contributes to the astringency of wines. From an oenological point of view, the presence of elagitanning and other representative phenols in extracts from vine shoots makes foreseeable their use to improve wine quality in a way similar to wine aging either in contact with oak chips or in oak barrels. Following the above study, the extracts from the 18 vine-shoot cultivars were compared with those from five varieties of oak chips and similar patterns for many of the compounds with oenological interest were found. Statistical analyses enabled identification of the vine shoot varieties providing extracts with a composition more similar to that of extracts from oak chips. Therefore, these vine shoots varieties could be proposed as an alternative to the use of oak barrels or oak chips in the ageing process of wine and spirits (Delgado-Torre, Ferreiro-Vera, Priego-Capote &

Luque de Castro, 2012, Delgado de la Torre, Priego-Capote & Luque de Castro, 2012).

In relation to potential health benefits of phenolic compounds from vine shoots, the nutraceutical interest of these extracts should be evaluated.

Vine leaves. Vine leaves have been unequal used depending on their maturation state. Some high-priced cosmetics included in their composition extracts of outbreaks of vine leaves, with the consequent damage to the plant. Green leaves are used as food in some countries such as Greece and Turkey, both fresh and pickled (Xynos *et al.*, 2012). In addition, fresh leaves have been used to treat hypertension, diarrhea, bleeding, varicose veins, inflammation and diabetes, and they seem also to exert a protective effect on the liver (Rodrigo, Miranda & Vergara, 2011).

There is a completely ignored area in this field: senescent leaves. It is well known that in this state, vine leaves have a color ranging from yellow to red and brown, depending on the variety, which reveals the presence of flavonoids, especially flavonols, anthocyanins/anthocyanidins and carotenoids with very different characteristics. Senescent vine leaves are, therefore, a potential raw material for the production of natural colorants that do not require an experimentation period prior to regulating their use.

3.2. Potential of winemaking residues

Grape production worldwide is about 70 million tonnes, of which 80% is dedicated to winemaking. The unequal exploitation of residues from the winemaking process has given place in few cases to high-added value products, despite the variety and features that make valuable some of their components. Distinction is here made between skin, seeds and lees (Fig. 3).

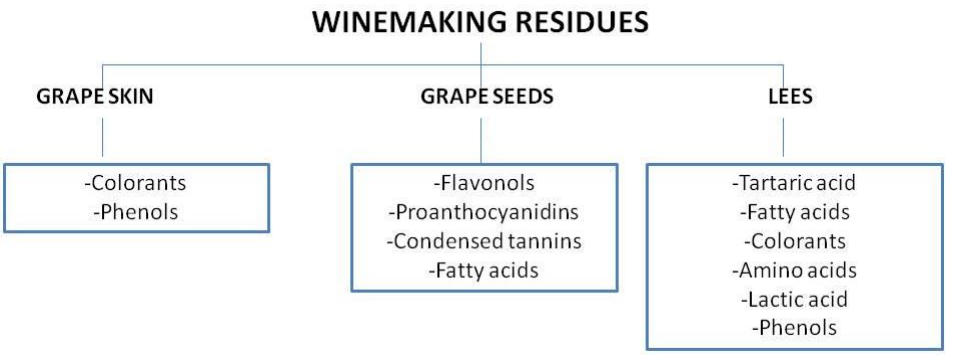


Figure 3. Valuable compounds obtained from winemaking residues.

Grape skin. The characteristics of grape skin differ depending on grape variety, and especially on color, being skin from red grapes the most appreciated. Most of the research carried out so far has been devoted to red grape skin directly from grapes, representing a significant cost of raw material (Corrales, García, Butz & Tauscher, 2009). Several studies have demonstrated that the grape skin from winemaking of red wine has a number and concentration of colorants that make this residue profitable as a raw material for colorants production (Corrales, García, Butz & Tauscher, 2009).

Most aromas and phenols are colorless, the latter being endowed with a high antioxidant capacity and existing at higher concentrations in skin and seeds than in pulp. Nevertheless, interesting products from this raw material can be obtained prior to colorants –as is the case with aromas– and either simultaneous or subsequent to obtainment of colorants –as is the case for non-colored phenol compounds.

The need for removal of the high moisture content of this raw material –an expensive step– prior to its use, could be drastically reduced by using the SHLE method reported by Luque-Rodríguez *et al.*, 2006.

Grape seeds. Grape seeds are the most studied and exploited residues from winemaking due to the large variety of compounds in this raw material.

Between 38 and 52% of grape is made by seeds, which give an overall idea of the amount of this residues generated in each harvest. The high content in phenols of seeds (mainly flavonols and proanthocyanidins or condensed tannins) endows them with a high antioxidant capacity (Passos *et al.*, 2010), which has been the subject of many studies. Proanthocyanidins –compounds that provide wine with flavor and, especially, astringency– are recognized by their beneficial health effects, such as anti-inflammatory, cardioprotective, radioprotective, anti-hyperglycemic, anti-tumor and antigenotoxic, prevention and healing of cataracts (Ferreira, Marais, Coleman & Slade, 2010). These compounds are the only that have been commercialized as nutritional supplements in US, Australia, Japan, and Korea, while they are used in Brazil as antioxidants to prevent atherosclerosis and other cardiovascular diseases as well as for treatment of dislipidemy (Ferreira, Marais, Coleman & Slade, 2010). Recently, phenol extracts from grape seeds have been used to improve the antioxidant quality of bread (Peng *et al.*, 2010), and more recent studies have demonstrated the powerful action of these compounds as free radical scavengers (Ferreira, Marais, Coleman & Slade, 2010). In the light of the research developed so far and the healthy properties of proanthocyanidins from grape seeds, we should wonder why not an extent use of grape seeds has been promoted yet.

In addition to the interest on the above components, the lipid fraction of grape seeds has a high commercial value, which has been only partially exploited. Grape-seed oil is made by 90% of poly- and mono-unsaturated fatty acids, particularly linoleic acid (58–78%, 18:2n-6), followed by oleic acid (3-15%, 18:1n-9), which are responsible for its nutritive value as edible oil. This oil only contains low amounts of saturated fatty acids (10%). The high content of unsaturated fatty acids makes grape-seed oil a high-quality nutritional oil, which exhibits properties for prevention of thrombosis, inhibition of cardiovascular diseases, reduction of cholesterol in serum, dilation of blood vessels and regulation of autonomic nerves (Ferreira, Marais, Coleman & Slade, 2010). Methods based on superheated liquids, with reduced extraction costs, have also been developed (Luque-Rodríguez, Luque de Castro & Pérez-Juan, 2005, 2007).

The use of this oil in cosmetics is widely supported on the large number of commercial products which include it in their composition.

A complete characterization of the composition of grape seed would be of interest to know if other compounds, alone or in mixture, could also be obtained from this raw material.

Vinification lees. Vinification lees, defined by the EU as "the residue formed at the bottom of wine containers after fermentation, during the storage or after authorized treatments, and the residue from filtering or centrifugating this product" have a key role in winemaking (e.g. removal of undesirable compounds, interaction with the volatile fraction, biogenic amines and phenolic compounds and the release of mannoproteins, which work as natural stabilizer of colloidal charge in wine when they join to other proteins, tartrates, etc.) (Pérez-Serradilla & Luque de Castro, 2008).

Lees constitute an abundant and subexploited residue from the wine industry. Their present use is reduced to the recovery of tartaric acid, which is employed to correct the pH of must prior to fermentation (Pérez-Serradilla & Luque de Castro, 2008), some attempts to use lees as animal feed with low nutritional value, and as a complement of food and feed for its high content in polyunsaturated fatty acids. Other proposals without success have been as biosorbent for the removal of undesirable substances, for compost production, and as a nutrient to obtain lactic acid from Lactobacillus strains (Pérez-Serradilla & Luque de Castro, 2008).

The studies developed so far have allowed obtaining high-priced oils to be used in cosmetics and food; phenols with high antioxidant capacity, and amino acids of great interest for food enrichment. However, the two most promising uses of lees are as raw material for extraction of mannoproteins and colorants, both irreplaceable materials for wine clarification. The most valuable finding of a recent study was that the adsorptive capacity of lees makes the concentration of colorants in them up to ten times higher than in skin of red grapes (Pérez-

Serradilla & Luque de Castro, 2011). In addition, the high variety of vine cultivars provides grapes with a wide range of red tonalities by combination of a wide range of different colorants. A detailed study of the colorants characteristic of each cultivar will provide the necessary information for a rational exploitation of lees. It is worth noting that the easy extraction of colorants allows the matrix to remain unaltered, so it can be used for subsequent extraction of other valuable products.

Residues from residues. Finally, an exhaustive exploitation of residues could lead to the use of "residues from residues". An example of this potential exploitation would be the residue remaining after ethanol distillation from the wine pomace (Peralbo-Molina, Priego-Capote & Luque de Castro, 2012). Red- and white-grape skins and seeds from a distillery have been subjected to extraction with ethanol and water in equal proportion as extractant to obtain the largest possible number of compounds from these residues, which had so far been used only as a heat source. Four different extraction techniques were used (conventional maceration extraction –CME–, UAE, MAE, and SHLE and the last was selected as the best after the extracts were overall and individually studied by the Folin–Ciocalteu plus Ferric Reducing Antioxidant Power methods and LC–DAD plus LC–TOF/MS, respectively. The composition of the extracts under optimal conditions, different for red- and white-grape skins and seeds, was studied by LC–TOF/MS, and similarities were compared by Venn diagrams. Identification showed that the variety of compounds extracted under superheated-extractant conditions encompassed flavonoids, mainly anthocyanins; flavonols 3-glucoside derivatives and simple flavonols; phenolic compounds such as acids, esters and methoxy and ethoxy derivatives and stilbenes. The absence of pyroanthocyanins (except for vitisin B) in all extracts suggests their degradation during the ethanol-distillation process, supported by the characteristic lability of these compounds.

Despite the degradation caused by the drastic conditions of the distillation process, many interesting compounds exist in the extracts from the

waste of this process, which make it a useful matter for more than as a heating source. The variety of identified compounds in the extracts makes them exploitable as additives in the food industry (either as colorants, as flavor modifiers or as antioxidants), and also in the cosmetics and nutraceuticals industries.

Acknowledgement

The Spanish Ministerio de Ciencia e Innovación (MICINN) and FEDER program are thanked for financial support through project CTQ2009-07430.

References

Ali, N. F., & El-Mohamedy, R. S. R. (2011). Eco-friendly and protective natural dye from red prickly pear (Opuntia Lasiacantha Pfeiffer) plant. Journal of Saudi Chemical Society, 15, 257-261.

Amari, G. (1998). No. EP 0937455-A1.

Ansari, M. A., Abdul, H. M., Joshi, G., Opii, W. O., & Butterfield, D. A. (2009). Protective effect of quercetin in primary neurons against Aβ(1–42): relevance to Alzheimer's disease. Journal of Nutritional Biochemistry, 20, 269-275.

Arapitsas, P., Antonopoulos, A., Stefanou, E., & Dourtoglou, V. G. (2004). Artificial aging of wines using oak chips. Food Chemistry, 86, 563-570.

Authority, EFSA, (2010). Scientific Opinion on the appropriateness of the food azo-colours Tartrazine (E 102), Sunset Yellow FCF (E 110), Carmoisine (E 122), Amaranth (E 123), Ponceau 4R (E 124), Allura Red AC (E 129), Brilliant Black BN (E 151), Brown FK (E 154), Brown HT (E 155) and Litholrubine BK (E 180) for inclusion in the list of food ingredients set up in Annex IIIa of Directive 2000/13/EC. EFSA Journal, 8, 11 pp.

Authority, EFSA, (2011) Scientific opinion on the substantiation of health claims related to polyphenols in olive and protection of LDL particles from oxidative damage (ID 1333, 1638, 1639, 1696, 2865), maintenance of normal blood HDL cholesterol concentrations (ID 1639), maintenance of normal blood pressure (ID 3781), "anti-inflammatory properties" (ID 1882), "contributes to the upper respiratory tract health" (ID 3468), "can help to maintain a normal function of gastrointestinal tract" (3779), and "contributes to body defences against external agents" (ID 3467) pursuant to Article 13(1) of Regulation (EC) No 1924/2006. EFSA Journal, 9, 25 pp.

Authority, EFSA, (2012). Scientific Opinion on the substantiation of a health claim related to polyphenols in olive and maintenance of normal blood HDL cholesterol concentrations (ID 1639, further assessment) pursuant to Article 13(1) of Regulation (EC) No 1924/2006. EFSA Journal, 10, 8pp.

Authority, EFSA, (2012). Scientific Opinion on the evaluation of the substances currently on the list in the annex to Commission Directive 96/3/EC as acceptable previous cargoes for edible fats and oils – Part II of III. EFSA Journal, 10, 150 pp.

Bautista-Ortín, A. B., Lencina, A. G., Cano-López, M., Pardo-Mínguez, F., López-Roca, J. M., & Gómez-Plaza, E. (2008). The use of oak chips during the ageing of a red wine in stainless steel tanks or used barrels: effect of the contact time and size of the oak chips on aroma compounds. Australian Journal of Grape and Wine Research, 14, 63-70.

Bechtold, T., Mahmud-Ali, A., & Mussak, R. (2007). Natural dyes for textile dyeing: A comparison of methods to assess the quality of Canadian golden rod plant material. Dyes and Pigments, 75, 287-293.

Cerdá, T. G., Rodríguez Mozaz, S., & Azpilicueta, C. A. (2002). Volatile composition of aged wine in used barrels of French oak and of American oak. Food Research International, 35, 603-610.

Cioffi, G., Pesca, M. S., De Caprariis, P., Braca, A., Severino, L., & De Tommasi, N. (2010). Phenolic compounds in olive oil and olive pomace from Cilento (Campania, Italy) and their antioxidant activity. Food Chemistry, 121, 105-111.

Conde, E., Cara, C., Moure, A., Ruiz, E., Castro, E., & Domínguez, H. (2009). Antioxidant activity of the phenolic compounds released by hydrothermal treatments of olive tree pruning. Food Chemistry, 114(3), 806–812.

Corrales, M., García, A. F., Butz, P., & Tauscher, B. (2009). Extraction of anthocyanins from grape skins assisted by high hydrostatic pressure. Journal of Food Engineering, 90, 415-421.

Covas, M. I. (2007). Olive oil and the cardiovascular system. Pharmacological. Research, 55, 175-186.

Delgado-Torre, M. P., Ferreiro-Vera, C., Priego-Capote, F., Pérez-Juan, P. M., & Luque de Castro, M. D. (2012). Comparison of accelerated methods for the extraction of phenolic compounds from different vine-shoot cultivars. Journal of Agricultural and Food Chemistry, 60, 3051-3060.

Delgado de la Torre, M. P., Priego-Capote, F., & Luque de Castro, M. D. (2012). Evaluation of the composition of vine shoots and oak chips for oenological purposes by superheated liquid extraction and high-resolution liquid chromatography-time-of-flight/mass spectrometry analysis. Journal of Agricultural and Food Chemistry, 60, 3409-3417.

Fernández-Arche, A., Márquez-Martín, A., Vázquez, R. P., Perona, J. S., Terencio, C., Pérez-Camino, C., & Ruíz-Gutiérrez, V. (2009). Long-chain fatty alcohols from pomace olive oil modulate the release of proinflammatory mediators. Journal of Nutritional Biochemistry, 20, 155-162.

Fernández de Simón, B., Cadahía, E., del Álamo, M., & Nevares, I. (2010). Effect of size, seasoning and toasting in the volatile compounds in

toasted oak wood and in a red wine treated with them. Analytica Chimica Acta, 660, 211-220.

Ferreira, D., Marais, J. P. J., Coleman, C. M., & Slade, D. (2010) Proanthocyanidins: Chemistry and Biology. In Lew, M. & Hung-Wen, L. (eds), Comprehensive Natural Products II (pp 605-661). Oxford: Ed. Elsevier.

Herrero, M., Temirzoda, T. N., Segura-Carretero, A., Quirantes, R., Plaza, M., & Ibañez, E. (2011). New possibilities for the valorization of olive oil by-products. Journal of Chromatography A, 1218, 7511-7520.

Japón-Luján, R., & Luque de Castro, M. D. (2006). Superheated liquid extraction of oleuropein and related biophenols from olive leaves. Journal of Chromatography A, 1136, 185-191.

Japón-Luján, R., & Luque de Castro, M. D. (2007). Static−dynamic superheated liquid extraction of hydroxytyrosol and other biophenols from alperujo (a semisolid residue of the olive oil industry). Journal of Agricultural of Food Chemistry, 55, 3629-3634.

Japón-Luján, R., Luque-Rodríguez, J. M., & Luque de Castro, M. D. (2006). Dynamic ultrasound-assisted extraction of oleuropein and related biophenols from olive leaves. Journal of Chromatography A, 1108, 76-82.

Japón-Luján, R., Luque-Rodríguez, J., & Luque de Castro, M. D. (2006). Multivariate optimisation of the microwave-assisted extraction of oleuropein and related biophenols from olive leaves. Analytical and Bioanalytical Chemistry, 385, 753-759.

Jiménez, L., Angulo, V., Ramos, E., De la Torre, M.J., & Ferrer, J.L. (2006). Comparison of various pulping processes for producing pulp from vine shoots. Industrial Crops and Products, 23(2), 122–130.

Jiménez, L., Rodríguez, A., Pérez, A., Moral, A., & Serrano, L. (2008). Alternative raw materials and pulping process using clean technologies. Industrial Crops and Products, 28, 11-16.

Kammerer, D., Gajdos Kljusuric, J., Carle, R., & Schieber, A. (2005). Recovery of anthocyanins from grape pomace extracts (Vitis vinifera L. cv. Cabernet mitos) using a polymeric adsorber resin. European Food and Research Technology, 220(3-4), 431–437.

Kim, S.K., & Karadeniz, F. (2012). Biological Importance and Applications of Squalene and Squalane. In Se-Kwon K. (Ed), Advances in Food and Nutrition Research, (pp 223-233). San Diego: Academic Press.

Kim, T. J., Kim, J. H., Jin, Y. R., & Yun, Y. P. (2006). The inhibitory effect and mechanism of luteolin 7-glucoside on rat aortic vascular smooth muscle cell proliferation. Archives of Pharmacological Research, 29, 67-72.

Le Floch, F., Tena, M. T., Ríos, A., & Valcárcel, M. (1998). Supercritical fluid extraction of phenol compounds from olive leaves. Talanta, 46, 1123-1130.

Lesage-Meessen, L., Navarro, D., Maunier, S., Sigoillot, J. C., Lorquin, J., Delattre, M., Simon, J. L., Asther, M., Labat, M. (2001). Simple phenolic content in olive oil residues as a function of extraction systems. Food Chemistry, 75, 501-507.

Li, C., Zheng, Y., Wang, X., Feng, S., & Di, D. (2011). Simultaneous separation and purification of flavonoids and oleuropein from Olea europaea L. (olive) leaves using macroporous resin. Journal of Science and Food Agriculture, 91, 2826-2834.

Lozano-Sánchez, J., Giambanelli, E., Quirantes-Pine, R., Cerretani, L., Bendini, A., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2011). Wastes generated during the storage of extra virgin olive oil as a natural source of

phenolic compounds. Journal of Agricultural and Food Chemistry, 59, 11491-11500.

Luque de Castro, M. D., & Priego Capote, F. (2010). Extraction of Oleuropein and Related Phenols from Olive Leaves and Branches. In Víctor, R. P. & Ross W. R. (Eds), Olives and Olive Oil in Health and Disease Prevention (pp 259-273). San Diego: Academic Press.

Luque-Rodríguez, J. M., Luque de Castro, M. D., & Japón Luján, R. (2008). Exploitation of Residues from Vineyards, Olive Groves, and Wine and Oil Production to Obtain Phenolic Compounds of High-Added Value. In Hurst, W. J. (Ed), Methods of Analysis for Functional Foods and Nutraceuticals, second edition, CRC Press.

Luque-Rodríguez, J. M., Luque de Castro, M. D., & Pérez-Juan, P. (2005). Extraction of fatty acids from grape seed by superheated hexane. Talanta, 68, 126-130.

Luque-Rodríguez, J. M., Luque de Castro, M. D., & Pérez-Juan, P. (2007). Dynamic superheated liquid extraction of anthocyanins and other phenolics from red grape skins of winemaking residues. Bioresource Technology, 98, 2705-2713.

Luque-Rodríguez, J. M., Pérez-Juan, P., & Luque de Castro, M. D. (2006). Extraction of polyphenols from vine shoots of *vitis vinifera* by superheated ethanol–water mixtures. Journal of Agricultural and Food Chemistry, 54, 8775-8781.

Manna, C., Galletti, P., Cucciolla, V., Montedoro, G., & Zappia, V. (1999). Olive oil hydroxytyrosol protects human erythrocytes against oxidative damages. The Journal of Nutritional Biochemistry, 10, 159-165.

Max, B., Salgado, J. M., Cortes, S., & Domínguez, J. M. (2010). Extraction of phenolic acids by alkaline hydrolysis from the solid residue obtained

after prehydrolysis of trimming vine shoots. Journal of Agricultural and Food Chemistry, 58, 1909-1917.

Mazzotti, F., Benabdelkamel, H., Donna, L. D., Maiuolo, L., Napoli, A., & Sindona, G. (2012). Assay of tyrosol and hydroxytyrosol in olive oil by tandem mass spectrometry and isotope dilution method. Food Chemistry, 135(3), 1006-1010.

Mediavilla, I., Fernández, M. J., & Esteban, L. S. (2009). Optimization of pelletisation and combustion in a boiler of 17.5 kWth for vine shoots and industrial cork residue. Fuel Processing Technology, 90, 621-628.

Meksi, N., Haddar, W., Hammami, S., Mhenni, M. F. (2012). Olive mill wastewater: A potential source of natural dyes for textile dyeing. Industrial Crops and Products, 40, 103-109.

Moldes, A. B., Torrado, A. M., Barral, M. T., & Domínguez, J. M. (2007). Evaluation of biosurfactant production from various agricultural residues by Lactobacillus pentosus. Journal of Agricultural and Food Chemistry, 55, 4481-4486.

Moldes, A. B., Torrado, A., Converti, A., & Domínguez, J. M. (2006). Complete bioconversion of hemicellulosic sugars from agricultural residues into lactic acid by Lactobacillus pentosus. Applied Biochemistry and Biotechnology, 135, 219-228.

Montero, I., Miranda, T., Arranz, J. I., & Rojas, C. V. (2011). Thin layer drying kinetics of by-products from olive oil processing. International Journal of Molecular Science, 12, 7885-7897.

Morales, M. L., Benítez, B., & Troncoso, A. M. (2004). Accelerated aging of wine vinegars with oak chips: evaluation of wood flavour compounds. Food Chemistry, 88, 305-315.

Mylonaki, S., Kiassos, E., Makris, D., & Kefalas, P. (2008). Optimisation of the extraction of olive Olea europaea leaf phenolics using water/ethanol-based solvent systems and response surface methodology. Analytical and Bioanalytical Chemistry, 392, 977-985.

Obied, H. K., Bedgood Jr, D. R., Prenzler, P. D., & Robards, K. (2007). Bioscreening of Australian olive mill waste extracts: Biophenol content, antioxidant, antimicrobial and molluscicidal activities. Food and Chemical Toxicology, 45, 1238-1248.

Oh, H., Kim, D. H., Cho, J. H., & Kim, Y. C. (2004). Hepatoprotective and free radical scavenging activities of phenolic petrosins and flavonoids isolated from Equisetum arvense. Journal of Ethnopharmacology, 95, 421-424.

Omar, S. H. (2010). Cardioprotective and neuroprotective roles of oleuropein in olive. Saudi Pharmaceutical Journal, 18, 111-121.

Pan, G. X., Bolton, J. L., & Leary, G. J. (1998). Determination of ferulic and *p*-coumaric acids in wheat straw and the amounts released by mild acid and alkaline peroxide treatment. Journal of Agricultural and Food Chemistry, 46, 5283-5288.

Passos, C. P., Silva, R. M., Da Silva, F. A., Coimbra, M. A., & Silva, C. M. (2010). Supercritical fluid extraction of grape seed (*Vitis vinifera L.*) oil. Effect of the operating conditions upon oil composition and antioxidant capacity. Journal of Chemical Engineering, 160, 634-640.

Peng, X., Ma, J., Cheng, K.-W., Jiang, Y., Chen, F., & Wang, M. (2010). The effects of grape seed extract fortification on the antioxidant activity and quality attributes of bread. Food Chemistry, 119, 49-53.

Phillips, K. M., Ruggio, D. M., Toivo, J. I., Swank, M. A., & Simpkins, A. H. (2002). Free and esterified sterol composition of edible oils and fats. Journal of Food Composition and Analysis, 15, 123-142.

Pérez-Serradilla, J. A., & Luque de Castro, M. D. (2008). Role of lees in wine production: A review. Food Chemistry, 111, 447-456.

Pérez-Serradilla, J. A., & Luque de Castro, M. D. (2011). Microwave-assisted extraction of phenolic compounds from wine lees and spray-drying of the extract. Food Chemistry. 124, 1652-1659.

Peralbo-Molina, A.; Priego-Capote, F.; & Luque de Castro, M. D. (2012). Comparison of extraction methods for exploitation of grape skin residues from ethanol distillation. Talanta, 101, 292-298.

Peralbo Molina, A., Priego Capote, F., & Luque de Castro, M.D. (2012). Valorization of olive pomace by characterization of phenolic extracts using liquid chromatography-tandem mass spectrometry with a QQTOF mass analyzer. Journal of Agricultural and Food Chemistry, 60(46), 11542-11550.

Petti, S., & Scully, C. (2009). Polyphenols, oral health and disease: A review. The Journal of Clinical Dentistry, 37, 413-423.

Priego-Capote, F., Ruiz-Jiménez, J., & Luque de Castro, M. D. (2004). Fast separation and determination of phenolic compounds by capillary electrophoresis-diode array detection: Application to the characterisation of alperujo after ultrasound-assisted extraction. Journal of Chromatography A, 1045, 239-246.

Rodrigo, R., Miranda, A., & Vergara, L. (2011). Modulation of endogenous antioxidant system by wine polyphenols in human disease. Clinical Chimica Acta, 412, 410-424.

Rodríguez-Rodríguez, R., & Ruíz-Gutiérrez, V. (2010). Functional properties of pentacyclic triterpenes contained in pomace olive oil. In R.P. Víctor & W. Ronald Ross (Eds), Olives and olive oil in health and disease prevention (pp. 1431-1438). San Diego: Academic Press.

Sampedro, I., Aranda, E., Martín, J., García-Garrido, J. M., García-Romera, I., & Ocampo, J. A. (2004). Saprobic fungi decrease plant toxicity caused by olive mill residues. Applied Soil Ecology, 26, 149-156.

Shtukatur, I. (2003). No. WO 2004080203-A2.

Stevenson, J., (2009). Food additives and children's behaviour: evidence based policy at the margins of certainty. Journal of Children's Service, 4, 4-13.

Stueckler, F. (1998). No. WO 9948386-A1.

Tabera, J., Guinda, A., Ruíz-Rodríguez, A., Señorans, F. J., Ibáñez, E., Albi, T., & Reglero, G. (2004). Countercurrent supercritical fluid extraction and fractionation of high-added-value compounds from a hexane extract of olive leaves. Journal of Agricultural and Food Chemistry, 52, 4774-4779.

Torres de Pinedo, A., Peñalver, P., & Morales, J. C. (2007). Synthesis and evaluation of new phenolic-based antioxidants: Structure–activity relationship. Food Chemistry, 103, 55-61.

www.cen-online.org, June 9 (2008) 32, January 21 (2008) 33.

www.cen-online.org, March 24 (2008) 28, December 15 (2008) 18.

Xynos, N., Papaefstathiou, G., Psychis, M., Argyropoulou, A., Aligiannis, N., & Skaltsounis, A.L. (2012). Development of a green extraction procedure with super/subcritical fluids to produce extracts enriched in oleuropein from olive leaves. The Journal of Supercritical Fluids, 67, 89-93.

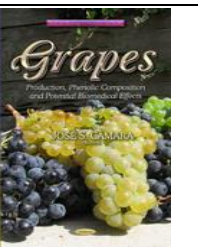
Capítulo II:

Grape phenols and their healthy properties

Nova Science Publisher

"Grapes: Production, phenolic composition and potential biomedical effects"

Chapter 7: Grape phenols and their healthy properties



Grape phenols and their healthy properties

A. Peralbo-Molina[1-3,] M.D. Luque de Castro[1-3*]

[1]*Department of Analytical Chemistry, Annex Marie Curie Building. Campus of Rabanales, University of Córdoba, Córdoba, Spain.*

[2]*University of Córdoba Agroalimentary Excellence Campus, ceiA3.*

[3]*Maimónides Institute of Biomedical Research (IMIBIC), Reina Sofía University Hospital, University of Córdoba, E-14071, Córdoba, Spain.*

Grape phenols and their healthy properties

Ángela Peralbo-Molina and Mª Dolores Luque de Castro

Abstract

Information and discussion on grapes in general and phenols in them in particular are the subject matter of this chapter. Therefore, after information on production (world-wide production and main producer countries, percentage of the different uses), description of families, classes and subclasses of phenols in grapes, their generic formulae and chemical properties are addressed. Discussion about the different parts of grapes and the absolute and relative concentrations of the types of phenols they contain precedes analytical information on these compounds, which is aimed at providing a critical view of: (i) the types of residues as raw materials from which phenols can be obtained; (ii) the extraction/leaching techniques in which the methods to obtain grape phenols are based: their efficiency, rapidity and performance; (iii) the global and/or individual analytical information about the target phenols the proposed methods provide; (iv) the present technologies to identify/quantify each type of target phenols.

The proved healthy properties of grape phenols, widely demonstrated but not assessed yet, in different clinical areas (e.g. improvement/prevention of neurodegenerative diseases, cardiovascular disorders and cancer) are discussed and potential research in the light of metabolomics is proposed.

Finally, the foreseeable trends in the grape-phenols field are outlined and the most important among them –integral exploitation of all parts of grape and residues from any type of grape-derivated industry– highlighted.

Keywords: Grape, phenols, exploitation of grape residues.

1. Introduction

Epidemiological studies provide convincing evidence that diets rich in fruits and vegetables are associated to prevention or delay of chronic degenerative diseases such as cancer, cardiovascular disease, atherosclerosis, and type 2 diabetes (Spormann *et al.* 2008), among others.

The Mediterranean diet is widely known thanks to its well documented health benefits (Kushi *et al.* 1995), which include the protective effects listed above (Biesalski *et al.* 2004; Choi, 2005; Dryden *et al.* 2006; Giacosa *et al.* 2013; Kok *et al.* 2004). One of the traditional cultures characteristic of this region is vineyard.

By comparing diets among different countries, researchers discovered that notwithstanding the French tendency to eat abundant animal fat, the incidence of heart disease remained lower than in other countries. This phenomenon was appointed as the "French Paradox", and it is thought to occur thanks to the protective benefits of regular consumption of red wine. Studies on the health effects of wine began with this discovery and polyphenols were ascribed to their antioxidant and free radical scavenging properties, being this fact more influential than wine consumption (Keys, 1970; Feinleib 1981).

Furthermore, there is at present a pressing need to obtain natural dyes to substitute artificial dyes. This need began with a study of the University of Southampton which reported, in late 2009, an experiment involving 6 artificial colorants that caused hyperactivity and allergy in children subjected to the test. Subsequently, the EFSA (European Food Safety Authority) conducted a toxicological study with 5 azoic colorants (E102, tartrazina yellow; E110, orange sunset yellow; E122, carmoisine; E124, Ponceau 4R; and E129, red Allura AC) and 1 quinolinic colorant (E104, quinoline yellow). After this study, the EFSA ruled that the target colorants did not produce hyperactive, allergenic, or other toxic effects. Nevertheless, the Directorate General for Health and Consumers (DG SANCO) of the European Parliament decided, as a precautionary measure,

to include in the labelling of foods containing the target colorants a sentence about the possibility that they "can produce actions and negative effects on children's activity", as set in Annex V of Regulation (EC) 1333/2008 (food additives) (www.cen-online.org January 21 (2008) 33; www.cen-online.org June 9 (2008) 32). This situation has increased the search for –and application of– natural colorants to replace synthetic colorants (www.cen-online.org December 15 (2008) 18; www.cen-online.org March 24 (2008) 28). Grapes have been subjected to numerous studies with the aim of combating this emerging need as the colorants from this fruit, or from any grape residue, should not be subjected to legal delays in their use (Wada *et al.* 2007).

A significant fraction of compounds in grapes is formed by phenols, which are a huge and wide family of compounds, many of which exist naturally in an array of food plants, especially in fruits, seeds and leaves. They contain more than 8000 known compounds, ranging from simple phenols such as phenol itself to complex phenols of variable composition such as tannins (Bravo, 1998; Harborne, 1995). In the wine jargon the term polyphenols is used collectively both to refer to simple and/or complex phenols involving mixed chemical structures which are often imprecise even to researchers (Tsao, 2010). In the authors' opinion, the use of the term phenols in this area is more appropriate as it involves both simple phenols and polyphenols. Chemically, phenols include a wide variety of biomolecules –such as flavonoids, stilbenes, lignans, etc.– which contain several hydroxyl groups (D'Archivio *et al.* 2007; Tsao, 2010). All them are considered key compounds that contribute to the antioxidant capacity of grapes, wine and other fruits, major responsible of the health benefits of these foods (Moini *et al.* 2000).

In this chapter description-discussion on grapes (general information, production, uses) in general, and on the phenols they contain (description of classes, subclasses, absolute and relative concentration in grapes) can be found. Also the chapter includes information about the extraction techniques used to

isolate grape phenols, their healthy properties and contribution to different clinical areas and, finally, the foreseeable trends in the grape-phenols field.

2. Grapes production

Grape production worldwide is about 70 million tons, 80% of which is devoted to winemaking, and practically the rest is the raw material of the grape juice industry (Moini *et al.* 2000). Only 2% is used as fresh or dried fruit. Most grapes come from Vitis vinifera, which is native from the Mediterranean basin and Central Asia. Minor contributors are American and Asia species such as Vitis labrusca, Vitis ripalia and Vitis rotund folia, among others. According to the Food and Agriculture Organization (FAO), the Mediterranean basin and Central Asia produced 16.7 million tons of wine in 2010, which was 64% of the total global wine production. This region covered 63% of the total vineyard area in the world with 4.5 million hectares dedicated to this cultivar (FAO, 2013). Mediterranean countries have the largest total area of vineyards, and also the three major wine producers (viz. France, Italy and Spain, responsible for 48% of global wine production). Over the past two decades wine production has decreased by 18% in this region due to the emergence of new prominent wine producers such as Australia, South Africa, and the United States (FAO, 2013).

Vineyards produce huge amounts of agricultural residues (particularly vine shoots and leaves). According to Jiménez *et al.*, an estimated total of 11.2-16.0 million tons of vine shoots are produced each year (Jiménez *et al.* 2006). On the other hand, grape pomace –the solid waste remaining after grape pressing– is the most abundant residue from the winemaking process –around 10.5-13.1 million tons of this waste produced in the world each year (Kammerer *et al.* 2005). At present, there is a growing demand for exploitation of agricultural residues, which have proved to be a key source of phenols (Peralbo-Molina and Luque de Castro, 2013).

Climate change is exerting an increased and profound influence on vine phenomenology and grape composition, affecting vinification, wine microbiology,

chemistry, and sensory aspects. According to Mira de Orduña (Mira de Orduña, 2010), there is evidence about association of climate change with grape quality, which will affect in the future to final wine quality, especially concerning the expression of varietal grape aromas, microbiological and chemical stability and sensory balance. Phenols quantity and quality in plant food can vary significantly according to different intrinsic and extrinsic factors such as plant genetics and cultivar, soil composition and growing conditions, maturity state and post-harvest conditions, among others (Faller and Fialho, 2010). These aspects should be studied in depth to provide basic scientific knowledge on how climate change may affect humans after wine consumption.

3. Grape phenols

Phenols contribute to the color and sensory properties such as bitterness and astringency of fruits, vegetables, nuts, seeds, flowers and also some herbs. These compounds are an integral part of the human diet that takes benefit from their antioxidant properties, mainly responsible for the growing interest in these compounds (Caillet *et al.* 2006).

3.1. Phenols classification

Phenols, one of the most abundant and widely distributed natural compounds in the plant kingdom, are highly diverse. The majority of phenols in plants exists as glycosides with different sugar units and acylated sugars at different positions of the phenol skeleton. Phenols are mainly classified as a function of their origin, their biological function or the chemical structure of the corresponding aglycon. The last classification is used in this chapter, in which a chart devoted to each class has been included with the corresponding derived subclasses and representative examples according to the substituents they are provided with.

Antioxidant compounds present in grape have been identified as phenolic acids (benzoic and hydroxynnamic acids), stilbene derivatives (resveratrol),

flavan-3-ols (catechin, epicatechin), flavonols (kaempferol, quercetin, myricetin) and anthocyanins; all them have been classified as non flavonoids and flavonoids compounds (Caillet *et al.* 2006).

(i) Non flavonoids

It is widely known that flavonoids are compounds which possess a C6-C3-C6 backbone, with two benzene rings connected via a heterocyclic pyrane or pyrane ring. Therefore, non-flavonoids compounds are all other compounds that do not possess this chemical structure, such as phenolic acids, stilbenes and lignans.

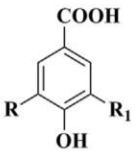
a. Phenolic acids

Phenolic acids in grapes, usually present in plants as esters, are also found in free form in grape pomace and leaf extracts as a result of partial hydrolysis of esters during extraction (Luque-Rodríguez *et al.* 2008). As can be seen in Figure 1, depending on the aromatic ring from which they come there are two classes of phenolic acids: benzoic and cinnamic acids and their derivatives (Manach *et al.* 2004).

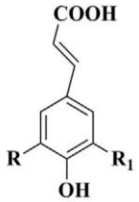
Hydroxybenzoic acids in plants are as such at very low concentrations; nevertheless, they are components of complex structures such as hydrolyzable tannins, of which gallic acid is the precursor (Garrido and Borges, 2013). Several types of hydroxybenzoic acids have been identified in both grapes and wine. The most abundant are *p*-hydroxybenzoic, protocatechuic, vanillic, gallic and syringic acids. Due to their scant presence in human edible plants, hydroxybenzoic acids have not been extensively studied and are not currently considered to be of great nutritional interest (Manach *et al.* 2004).

Hydroxycinnamic acids, which are also rarely found in free form, are located mainly in grape skin, leaves and stems. The most abundant of these acids are present either as tartaric esters or as a part of acylated anthocyanins

(Luque-Rodríguez *et al.* 2008). They are usually found at low concentrations in grapes (Morata *et al.* 2007); however, varieties with high contents of these acids also exist, and their use may stimulate phenolic pigments formation. Hydroxycinnamic acids may form a stack group with grape anthocyanins, thus increasing and stabilizing the color of anthocyanins and that of wine as a result – the copigmentation effect of caffeic acid on red wine color has been reported (Darias-Martín *et al.* 2001). Additionally, these compounds contribute to other sensory attributes such as bitterness, astringency and flavor (Andrade *et al.* 2001), being also recognized as responsible for several beneficial physiological effects on human health thanks to their antioxidant properties (Monagas *et al.*, 2005). The presence of hydroxycinnamic acids is a symptom of wine quality (Gambelli *et al.* 2004). Mulero *et al.* quantified hydroxycinnamic acids in grapes, finding that the most abundant hydroxycinnamic acid derivatives are *trans*-*p*-coumaroyltartaric and *trans*-caffeoyltartaric acids (Mulero *et al.* 2010).



Hydroxybenzoic acids	R	R1
Hydroxybenzoic acid	H	H
Protocatechuic acid	OH	H
Vanillic acid	OCH_3	H
Gallic acid	OH	OH
Syringic acid	OCH_3	OCH_3



Hydroxycinnamic acids	R	R1
p-Coumaric acid	H	H
Caffeic acid	OH	H
Ferulic acid	OCH_3	H
Sinapic acid	OCH_3	OCH_3

Figure 1: Chemical structure of hydroxybenzoic and hydroxycinnamic acids.

b. Stilbenes

Stilbenes are non-flavonoid phenols that occur in a number of plant families which constitute the main sources of them in diet. These phenols, endowed with two aromatic rings linked by an ethane bridge (see Figure 2), are constituents of the woody organs (roots, vine shoots, and stems), and

substances of induced production (in leaves and berries) acting as phytoalexins in the mechanisms of grape resistance against certain pathogens such as fungi (Luque-Rodríguez *et al.* 2008).

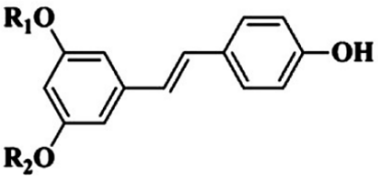
Stilbenes are also in grapes and, after processing, in wine (Zamora-Ros *et al.* 2008). The main grape stilbenes are *trans*- and *cis*-resveratrol, their glucosides (known as piceids), picetannol (Cantos *et al.* 2002), astringin, pterostilbene and oligomers (viniferins and hopeaphenol) (Guerrero *et al.* 2009); all them reported as responsible for numerous beneficial effects –antioxidant, antibacterial, antifungal, cardioprotective, neuroprotective, antiaging and anticancer effects, among others (Guerrero *et al.* 2009). Also stress resistance and lifespan in some organisms as caused by *trans*-resveratrol have been reported (Gruber *et al.* 2009)). Resveratrol, usually found in grape at higher concentrations than other stilbenes, is also found as glucosylated derivatives or oligomeric forms called viniferins (Luque-Rodríguez *et al.* 2008).

Sun *et al.* have studied the stilbenoid composition of grapes and wine (Sun *et al.* 2006). Their research yielded interesting results in which *cis*-piceid is the most abundant stilbene present in grape skin (150.64 mg kg^{-1}), followed by *trans*-piceid (72.19 mg kg^{-1}) and *trans*-resveratrol (66.67 mg kg^{-1}). Nevertheless, *cis*-resveratrol has been found in wine and non detected in grape skins (Roggero and García-Parrilla, 1995; Romero-Pérez *et al.* 2001). Astringin (either *trans* or *cis* forms) has not been detected in grape skins or in red wines, although several authors reported the presence of *trans*-astringin in red wines (Ribeiro de Lima *et al.* 1999).

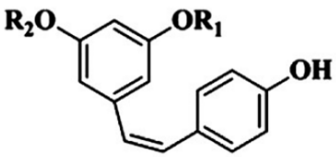
c. Lignin

Lignin, which has key biological functions in plants as to increase the mechanical strength of cell wall, accounts for about 17-28% of dry weight of grapevines –20% average in vine shoots (Luque-Rodríguez *et al.* 2008). Its three-dimensional structure is due to polymerization of phenolic alcohols units

such as coniferil or sinapyl alcohols (Garrido and Borges, 2013), as shown in Figure 3.



Stilbenes	R1	R2
trans-resveratrol	H	H
trans-piceid	H	Gluc
trans-pterostilbene	CH_3	CH_3



Stilbenes	R1	R2
cis-resveratrol	H	H
cis-piceid	H	Gluc

Figure 2. Chemical structure of stilbenes and derivatives.

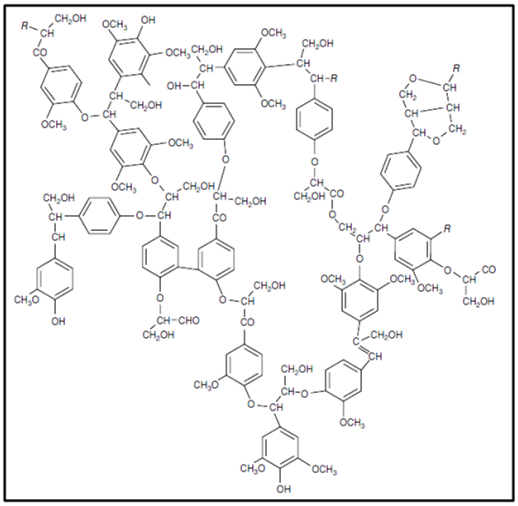


Figure 3: Lignin structure.

(ii) Flavonoids

Flavonoids are molecules with two aromatic rings linked by three carbons, usually arranged as an oxygenated heterocycle (C6-C3-C6). Flavonoids are classified, according to oxidation degree of the three carbons ring, into anthocyans, flavanols (also named as catechins), flavonols and flavanonols. Most flavonoids are present in nature as O-glycosides and other conjugates (catechins are an exception), which contribute to their complexity and the large number of individual molecules that have been identified (more than 5000) (Anhê *et al.* 2013).

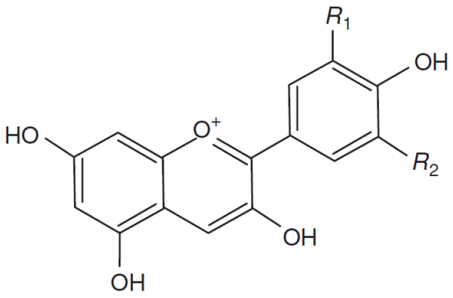
a. Anthocyanins

Anthocyanins account for most phenols remaining in skin from the red pomace, mainly 3-glycosides, 3-acetylglycosides and 3-*p*-coumaroylglycosides of malvidin, peonidin, delphinidin, petunidin and cyanidin (Luque-Rodríguez *et al.* 2007), but tartaric esters of hydroxycinnamic acids, monomeric and dimeric flavanols, flavonols and stilbenes are also found in skin and seeds (Rodríguez-Montealegre *et al.* 2006)). The main, and practically the only difference between red and white grape varieties is based on these compounds, virtually absent in the skin of white grapes (Luque-Rodríguez *et al.* 2008).

The basic skeleton structure of anthocyanins is shown in Figure 4. Structurally, anthocyanins are heterosides of an aglycone unit (anthocyanidin). The main differences among anthocyanins are the number of hydroxylated groups in the anthocyanidin skeletum, the nature and the number of bonded sugars in their structure, the aliphatic or aromatic carboxylate derivatives bonded to the sugar in the molecule, and the position of these bonds (Valls *et al.* 2009).

Potential health benefits –anticarcinogenic (Zhao *et al.* 1999), angioprotective (Vennat *et al.* 1988), antiinflammatory (Pietta *et al.* 2003), antibacterial activities (Fukai *et al.* 1991)– of these compounds have been widely reported. They are strong antioxidants and free radical scavengers (Woodman *et*

al. 2005) reducing the risk of coronary heart disease (Hung *et al.* 2000), some types of tumors (Williamson, 1990) and chronic diseases (Dryden *et al.* 2006).



Anthocyanins	R1	R2
Cyanidin	OH	H
Delphinidin	OH	OH
Peonidin	OCH_3	H
Petunidin	OCH_3	OH
Malvidin	OCH_3	OCH_3

Figure 4. Anthocyanins frequently present in grapes.

Anthocyanins are partially extracted into the must during the winemaking process, but the extraction is far for being complete. Therefore, the skins from the grape pomace constitute an abundant and inexpensive source of these phenols (Luque-Rodríguez *et al.* 2007) that are also present in other parts of the grapevine. Thus, leaves are rich in these compounds when they turn red in autumn (Luque-Rodríguez *et al.* 2008).

The anthocyanin content has been determined in different varieties of ripe grapes such as Gran Negro (1720 mg kg^{-1} of berry) or Mouraton (1000 mg kg^{-1} of berry), being malvidin 3-O-glucoside the most abundant anthocyanin (797 mg kg^{-1} of berry) followed by malvidin-3-O-(6-O-*p*-coumaryl)-glucoside (119 mg kg^{-1} of berry). Malvidin derivatives represent 52% of total anthocyanins, followed by peonidin compounds, which represent 38%. Minor contents –between 2 and 5%– correspond to petunidin, delphinidin and cyaniding derivatives (Figuereido-González *et al.* 2013).

b. Flavanols

Flavanols occur mainly in the outer coating of seeds but they are also present in some skin cells, stems, shoots and leaves (Luque-Rodríguez *et al.* 2008). The most abundant flavanols in grapes are flavan-3-ols, which are a subclass of flavonoids with variable number of hydroxyl groups at positions 5 and/or 7 on the A-ring and at positions 3′, 4′ and 5′ on the B-ring (Figure 5). Position 3 on the C-ring is commonly occupied by a hydroxyl group or it is esterified by gallic acid. Catechin, epicatechin, catechin gallate, epicatechin gallate, gallocatechin, epigallocatechin, gallocatechin gallate and epigallocatechin gallate are the most common flavan-3-ol monomers. Oligomers and polymers are mainly procyanidins, which are (+)-catechin, (-)-epicatechin and epicatechin-3-O-gallate in seeds and skins; partially galloylated units in shoots; and ungalloylated prodelphinidin derivatives in leaves (Luque-Rodríguez *et al.* 2008; Ali *et al.* 2010).

The content of flavan-3-ol monomers was determined by Figueiredo *et al.* (Figuereido-González *et al.* 2013) both in fresh grapes (24 mg kg^{-1}) and dried grapes (42 mg kg^{-1}). These results agree with those found in dried grapes by other authors (Márquez *et al.* 2012; Peinado *et al.* 2009; Serratosa *et al.* 2008a; 2008b; 2011). (+)-Catechin is the flavan-3-ol monomer found at the highest concentration, followed by (-)-epicatechin and epicatechin-3-O-gallate.

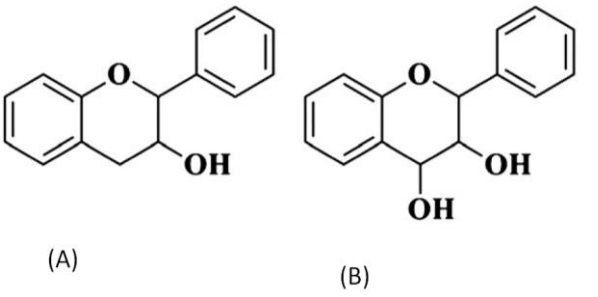
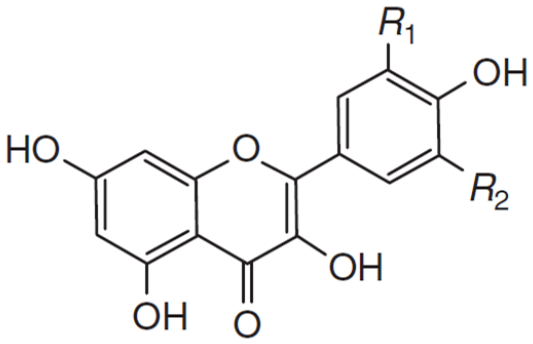


Figure 5. Chemical structure of flavan-3-ols (A) and flavan-3,4-diols (B).

c. Flavonols

Flavonols –flavonoids containing a carbonyl group at position 4 and a double bond between C2 and C3 (Figure 6)– are abundant in the skin of fresh grapes. These yellow pigments contribute directly to the color of white wines, but they are masked by anthocyanins in red wines. However, flavonols affect red wine color mainly by copigmentation (Schwarz *et al.* 2005).



Flavanols	R1	R2
Kaempferol	H	H
Quercetin	OH	H
Myricetin	OH	OH
Isorhamnetin	OCH_3	H

Figure 6. Flavonols frequently present in grapes.

Castillo-Muñoz *et al.* (Castillo-Muñoz *et al.* 2007)) have recently demonstrated that flavonols occurring in red grape cultivars belong to three glycosylated series (3-O-glucosyl, 3-O-galactosyl, and 3-O-glucuronyl) of the six possible flavonoid structures (kaempferol, quercetin, isorhamnetin, myricetin, laricitrin, and syringetin). With regards to white grape cultivars, the presence of isorhamnetin type flavonols in the skin of white grapes has been suggested on the basis of hydrolysis treatment and subsequent detection of the free aglycone

isorhamnetin (Mattivi *et al.* 2006), and also by the coincidence of the UV-vis spectra and chromatographic retention time with a true standard of isorhamnetin 3-O-glucoside (Rodríguez *et al.* 2006). However, the occurrence of isorhamnetin type flavonols and also myricetin type flavonols has been considered exclusive of red grape cultivars (Castillo-Muñoz *et al.* 2007).

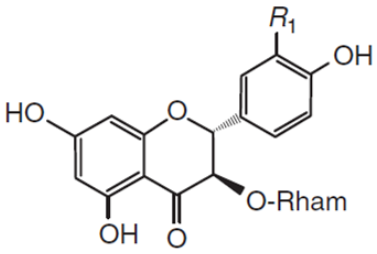
The most widespread roles of flavonols seem to be UV protection; nevertheless, these compounds have attracted much interest for their potential beneficial effects on human health linked principally to their antioxidant properties. Recent studies have highlighted their possible role in prevention of cardiovascular, neurological (López-Sánchez *et al.* 2007), and tumor diseases (Boots *et al.* 2008). In addition, either the flavonol profile or the ratio between quercetin and myricetin have been successfully applied to differentiate varietal wines (Castillo-Muñoz *et al.* 2007).

Figueiredo-González *et al.* have determined the total content of flavonols, resulting 44 mg kg^{-1} of berry. This result is similar to the previously reported by Castillo-Muñoz *et al.* (45 mg kg^{-1} of berry) for the same variety (Castillo-Muñoz *et al.* 2007). The most abundant flavonols in fresh grapes are quercetin derivatives (46%) followed by myricetin compounds (22%), being the glucoside derivatives present at higher concentrations than other derivatives (glucuronides, galactosides or rutinosides) (Figuereido-González *et al.* 2013).

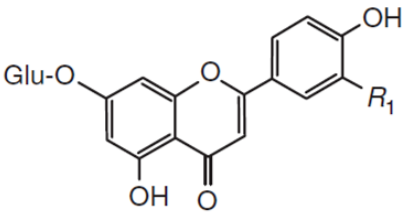
d. Flavanonols and flavones

Flavanonols are 2,3-dihydroflavonols, being astilbin and engeletin, identified in skin of white grapes (Trousdale and Singleton, 1983) and in stems (Souquet *et al.* 2000), the most known. Taxifolin, astilbin and dihydromyricetin 3-O-rhamnoside have often been reported in the literarure (Garrido and Borges, 2013). On the other hand, flavones are structurally similar to flavonols, but without a hydroxyl group at C3 (Figure 7). These compounds are known to have a significant pharmacological activity (Garrido and Borges, 2013). The most

important among them are apigenin-7-glucoside and luteolin-7-glucoside, detected in vine grape leaves (Hmamouchi *et al.* 1996).



Flavanonol	R1
Astilbin	OH
Engeletin	H



Flavone	R1
Luteolin-7-*O*-glucoside	OH
Apigenin-7-*O*-glucoside	H

Figure 7. Flavanonols and flavones found in grapes.

4. Analytical research on grape/phenols

Grape phenols have in the past been widely studied (Flamini *et al.* 2003; Jaitz *et al.* 2010; Modun *et al.* 2008; Serafini *et al.* 2000) because of their antioxidant and bactericide effects, beneficial for human health. The composition of grapes strongly affects the properties and sensory quality of wine (Bonghi *et al.* 2012). Effects from ageing in barrels have also been studied as tannins from wood can present an uptake route (Jaitz *et al.* 2010). Furthermore, the metabolism and gene expression of phenols in the skin of wine grape berries subjected to partial postharvest dehydration have been studied. This process seems to selectively affect specific pathways in skins, changing berries

composition, with consequences on the sensory characteristics of wines and affecting health promoting properties of grapes (Bonghi *et al.* 2012).

As mentioned before, grapes belong to the world's largest fruit crops (Moini *et al.* 2000), and they are appreciated for their rich content of phenols such as gallic acid, catechin and resveratrol, and a wide variety of procyanidins (Xu *et al.* 2010). Nutritional qualities of grapes are affected by environmental, cultural and post harvesting conditions, but genotype is the determinant factor leading to variation (Prior *et al.* 1998; Proteggente *et al.* 2002).

Processing of grapes for wine and juice globally leads to an annual amount of at least 10 million tons of press residues (Maier *et al.* 2009). Grape residue or wine orujo has become popular in recent years as nutritional supplement because grape skins and seeds accumulate a great proportion of phenolic compounds. The unequal exploitation of residues from winemaking process has given place in few cases to high added value products, despite the variety and features that make valuable some of their components. Distinction is here made among skin, seeds and lees.

The characteristics of grape skin depend on grape variety and color, being red grape skin the most appreciated after several studies have demonstrated the number and concentration of colorants that make this residue profitable as a raw material for colorant production, whether coming from the skin of fresh grapes (representing a significant cost of raw material) or after winemaking of red wine (a cheaper source) (Corrales *et al.* 2009).

In addition to colorants, aromas and colorless phenols are at higher concentrations in skin and seeds than in pulp, being possible to obtain them prior (aromas) and either simultaneous or subsequent (non colored phenol compounds) to colorants (Peralbo-Molina and Luque de Castro, 2013). Luque-Rodríguez *et al.* reported a superheated liquid extraction method to remove compounds of interest from this raw material at low costs (Luque-Rodríguez *et al.* 2007).

Between 38 and 52% of grape is constituted by seeds, which give an overall idea of the amount of this residue generated by each harvest (Maier *et al.* 2009). The composition of grape seed is basically 40% (w/w) fiber, 16% essential oil, 11% protein, 7% complex phenolic compounds like tannins, sugars, minerals, and other substances (de Campos *et al.* 2008). The lipid fraction of grape seeds has a high commercial value thanks to its high content of unsaturated fatty acids, such as linoleic acid (72–76% w/w), and tocopherols, which are also of interest because of their antioxidant properties.

Phenolic compounds are present at concentrations between 5 and 8%, depending on the cultivar; mainly as simple phenols such as gallic acid, flavonoids including monomeric flavan-3-ols catechin, epicatechin, gallocatechin, epigallocatechin, and epicatechin 3-O-gallate; also procyanidin dimmers, trimmers, and more polymerized procyanidins. Methods based on superheated liquids, with reduced extraction costs, have also been developed to obtain these compounds (Luque-Rodríguez *et al.* 2007).

Vinification lees, a residue formed at the bottom of wine containers after fermentation, during storage or after authorized treatments, has a key role in winemaking: for example, removal of undesirable compounds or interaction with the volatile fraction. Lees are a subexploited residue as their use is practically reduced to the recovery of tartaric acid (Pérez-Serradilla and Luque de Castro, 2011). The studies developed so far have allowed obtaining high priced oils to be used in cosmetic and food; phenols with high antioxidant capacity, and amino acids of great interest for food enrichment, in addition to mannoproteins and colorants (Pérez-Serradilla and Luque de Castro, 2008; 2011).

The residue remaining after ethanol distillation from the wine pomace (that is, a residue from a previous residue) has demonstrated to be a source of flavonoids, mainly anthocyanins, flavonols 3-glucoside derivatives and simple flavonols, phenolic compounds such as acids, esters, methoxy and ethoxy derivatives, and stilbenes (Peralbo-Molina *et al.* 2012; 2013). Despite the degradation caused by the high temperatures required in the distillation process,

the variety of compounds identified in the extracts makes them exploitable as additives in the food industry, and also in the cosmetics and nutraceutical industries.

4.1. Techniques and methods for extraction of valuable phenols from vineyard and wine residues

Green extraction processes (viz. extraction by superheated water, sub or supercritical carbon dioxide) are the most promising alternatives for implementation of industrial extraction, which require a prior development at the laboratory scale for proper and cheap optimization. In fact, methods based on the use of these extractants are an excellent way for isolation of anthocyanins and other phenols from residues of grape processing (Khanal *et al.* 2010).

(i) Supercritical fluid extraction (SFE)

SFE using CO_2 as extractant has, as main advantages its nontoxic, nonflammable, and noncorrosive character and the removal of the extractant after depressurization, thus providing a powder of the extracted compounds without the need for drying (Shi *et al.* 2005). However, the relatively high acquisition costs of the equipment and the price of the extractant with the appropriate purity degree constitute major constraints. This technique has been successfully used to extract phenols from grape pomace, generally using CO_2 modified with methanol (Casas *et al.* 2010; Mantell *et al.* 2001; 2003) or ethanol as cosolvent (Chafer *et al.* 2005) to increase extractant polarity.

(ii) Superheated liquid extraction (SHLE)

SHLE is more affordable than SFE. Also, it has two crucial advantages over conventional extraction, namely: (a) raising the temperature above the boiling point of the solvent increases the diffusion rate, solubility, and mass transfer of the compounds and decreases the viscosity and surface tension of the solvent. These changes result in improved contact of the compounds with the

solvent that facilitate extraction, which is thus faster and with less solvent requirement. (b) The absence of light and air significantly reduce degradation and oxidation of the target compounds during extraction (Santos-Buelga and Williamson, 2003).

SHLE has so far been used almost exclusively as a first step in methods for analyzing several phenols from fresh grapes. Thus, Piñero *et al.* determined *trans*-resveratrol in grapes by using a method involving a triple discrete SHLE step with methanol (Piñeiro *et al.* 2006). Also, Ju and Howard (Ju and Howard, 2003; 2005) examined the effect of various solvents and temperature conditions on the SHLE of anthocyanins from fresh grape skins obtained from a highly pigmented advanced breeding line of wine grape and concluded that either superheated water (110 ºC) or sulfured water provide extraction efficiencies similar to plein hot water or a 60:40 methanol–water mixture but with drastic shortening of the extraction time. SHLE methods applied to grape pomace have rarely been reported, particularly for red grape pomace as phenol skins from this material are extracted during the winemaking process; so, this last state pomace only contains strongly retained phenols. González-Rodríguez *et al.* (González-Rodríguez *et al.* 2003) and Luque-Rodríguez *et al.* (Luque-Rodríguez *et al.* 2007) optimized the extraction of anthocynins and other phenols using either superheated ethanol–water mixtures or a batch mode for grape pomace and a continuous mode for skins from grape pomace. The yields obtained with SHLE exceeded those provided by conventional solvent extraction; however, the total amounts obtained were, obviously, lower than those provided by fresh grapes.

Concerning the use of energy different from temperature+pressure to accelerate extraction from solid samples, the most important are microwaves (MW) and ultrasound (US), thus given place to microwave assisted extraction (MAE) and ultrasound assisted extraction (UAE).

(iii) Microwave assisted extraction

MAE is recognized as an outstanding way to promote efficient removal of target compounds from solid samples, cutting down extraction times and increasing yields and, frequently, quality of the extract. Also, MAE is considered faster, ecocompatible (less solvent required), repeatable and very effective (Cravotto *et al.* 2008). This technique has been applied to the development of methods for the extraction of organic compounds from matrices such as soils (Serrano and Gallego, 2006), sediments (Morales *et al.* 2005), seeds (Prado-Rosales *et al.* 2009), and foods (Robards, 2003; Wang and Sun, 2003). These studies show that microwaves expedite extraction from a number of different matrices as compared to conventional extraction methods. Thus, it has been reported that MAE is well suited for the extraction of phenolic compounds from grapes, which can be performed at relatively high temperatures (110–150 ºC) without alteration of the target compounds, a critical feature to handle antioxidants without degradation (Grigonis *et al.* 2005). Some other studies have also assessed the suitability of MAE for phenol removal without altering the antioxidant potential of the extracts and largely surpassing conventional extraction methods (Spigno *et al.* 2009). The key influence of the extractant on the expeditiousness of the process has been shown clearly by the extraction of phenols from grape seeds (Peng *et al.* 2001). It was found that the extractant was the most significant variable, even more than microwave power and extraction time; fact that can be ascribed to non interaction between MW and non polar extractants. Finally, a method for extraction of anthocyanins in grapes in which the simultaneous treatment of several samples multiplies sample throughput has recently been developed (Liazid et al 2011).

(iv) Ultrasound assisted extraction

UAE has demonstrated to pose a number of advantages over classical leaching methods, namely: shorter time, less drastic conditions –especially

temperature–, reduction of the use of hazardous reagents and less intervention of the analyst because of the possibility for automation.

One of the most important disadvantages of classical extraction methods such as Soxhlet extraction for analytical purposes or solvent extraction for analytical or industrial purposes is the requirement of long periods of solid sample–extractant contact. In a comparative study by Albu *et al.* (Albu *et al.* 2004) to extract antioxidants from rosemary by an ultrasonic bath, an ultrasonic probe and the stirring conventional method using different temperatures and extractants the use of US showed a drastic decrease of the extraction time. Similar behavior was reported by Cho *et al.* (Cho *et al.* 2006) in the extraction of resveratrol from grapes. The favorable effects of US on extraction, mainly due to the cavitation phenomenon –creation of high effective tiny temperatures and pressures– that efficiently acts at the interface between the solution subjected to US and a solid, can also result in denaturation of phenols (Kammerer et al. 2005).

(v) Enzyme assisted extraction

Assistance of extraction by enzymes is other approach that may be a useful alternative to conventional extraction of phenols, as demonstrated at laboratory and pilot plant scale. Preextracting the pomace with hot water, followed by treatment with cell wall degrading enzymes, provided extraction rates comparable to those obtained by extracting grape pomace with sulfite (Kammerer *et al.* 2005).

4.2. Analytical information from phenols extracts

There are different degrees of information to be mined from the extracts as a result of the complexity of the analytical equipment used, from overall determination to individual identification and/or quantitation.

(i) Overall analytical information from the extracts

Several types of overall information can be obtained from phenols extracts depending on the analytical method and the types of phenols.

a. Information about the total content of target families of compounds

This overall information can be obtained using general reagents that act on a given group in a molecule. In this case, a generic calibration curve is run using a representative compound of the target family; so semiquantitative results are thus obtained always relative to the standard used.

One of these methods, due to Folin-Ciocalteu (Singleton and Rossi, 1965), is used for overall determination of phenols. All compounds containing a phenol group in their molecule yield a colored compound with this reagent that can be monitored at 765 nm; thus allowing overall determination. Most authors use this method to determine the total phenolic content (Corrales *et al.* 2010; Negro *et al.* 2003), but there are other alternatives such as the enzymatic method described by Stevanato *et al.* (Stevanato *et al.* 2004) and also reported by Peinado *et al.* (Peinado *et al.* 2009) consisting on adding to the extract a 0.1 M potassium phosphate buffer, at pH 8, containing 3 mM horseradish peroxidase and monitored at 510 nm after 10 min rest. The results are expressed as equivalent of catechin.

Different authors have reported assays for given classes of phenols, as is the case for Di Stefano *et al.*, who proposed a method for total flavonoids and anthocyans based on measure of the absorbance at 280 and 540 nm, then expressing the results as equivalent of catechin and/or equivalent of malvidin (Di Stefano *et al.* 1989)). Also, the free anthocyans content can be determined by monitoring the absorbance at 540 nm after purification by solid phase extraction using C18 as sorbent (Di Stefano *et al.* 1989). Orak (Orak, 2007) reported other alternative to evaluate total anthocyans content using a method described by Wrolstad (Wrolstad, 1976). The grape extracts at both pH 1 and pH 4.5 are

monitored at 520 and 700 nm. The result, considered as the total anthocanins content, was calculated through a mathematical formulae and expressed as milligram of malvidin-3-O-glucoside per 1000 g of fruit.

Total flavonoids could be determined according to Makris *et al.* by diluting the grape extract with MeOH, mixing with distilled water and 5% $NaNO_2$ and allowing the mixture to react for 5 min. Then, 10% $AlCl_3$ was added and the mixture allowed standing for a further 5 min. Finally, 1 M Na_2CO_3 and distilled water were added to the reaction mixture and the absorbance at 510 nm was obtained against a blank that had been prepared in a similar manner, by replacing the extract by distilled water. The total flavonoid content was calculated from a calibration curve using catechin as standard, and expressed as mg of catechin per 100 g of fresh grape (Makris *et al.* 2007).

Total flavanols could be determined after derivatization with *p*-DMACA using an optimized methodology (Nagel and Glories, 1991). Grape extract suitably diluted with methanol was introduced into a microcentrifuge tube and HCl (0.24 N in methanol) and DMACA solution (0.2% in methanol) were added. The mixture was allowed to react for 5 min at room temperature, and the absorbance was monitored at 640 nm. A control sample was prepared by replacing sample with methanol. The results were expressed as mg of catechin per 100 g of fresh grape.

Procyanidins can be overally determined by the method from Habertson-Adams as modified by Skogerson *et al.* (Skogerson *et al.* 2007) consisting of precipitation of tannins with serum bovine albumin, washing of the precipitate, redissolution and determination of the phenols by FeCl3. Catechin was used as standard and the results expressed as equivalent of catechin.

b. Information about a chemical characteristic of the given extract

Free radical scavenging activity is a chemical characteristic of phenols extracts frequently demanded. Peralbo-Molina *et al.* reported this type of

information by applying the test for total antioxidant activity based on the ferric reducing antioxidant power (FRAP) (Peralbo-Molina *et al.* 2012; 2013). The assay consists of reduction of the ferric tripyridyltriazin [Fe(III)-TPTZ] complex to ferrous tri-pyridyl triazin [Fe(II)-TPTZ] at low pH by the action of antioxidants in the target solution. The ferrous complex [Fe(II)-TPTZ] has an intense blue color that can be monitored at 593 nm. The assay response was standardized against the antioxidant standard Trolox. The method has been automated by Benzie *et al.* (Benzie *et al.* 1996; 1999) by using a Cobas Fara centrifugal analyzer that performs the FRAP assay as follows: freshly prepared FRAP reagent is warmed to 37 ºC and a reagent blank reading is taken at 593 nm; the sample was then added with water. Absorbance readings were taken after 0.5 s and every 15 s thereafter during the monitoring period. The change in absorbance between the final reading selected and the blank reading was calculated for each sample and related to a calibration curve of a Fe(II) standard solution tested in parallel. Minor modifications of the method were developed by Bub *et al.* for miniaturization using a 96 well microtiter plate (Halvorsen *et al.* 2002).

Other alternatives to measure antioxidant activity are the 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox) equivalent antioxidant capacity (TEAC) assay of Miller *et al.* (Miller *et al.* 1993), and the oxygen radical absorbance capacity (ORAC) assay of DeLange and Glacer (DeLange and Glazer, 1989). Both assays are based on the ability of the antioxidant to react with or neutralize free radicals generated in the assay.

Among the three different assays, FRAP is the only that directly measures antioxidants or reductants in a sample because it uses antioxidant as reductants in a redox linked colorimetric reaction, while the other assays are indirect because they measure the inhibition of reactive species (free radicals) generated in the reaction mixture, and the results strongly depend on the type of reactive species used.

(ii) Individual quantitative information on phenols in the extracts

This type of information requires previous separation of the target phenols using the most appropriate technique as a function of their chemical characteristics: liquid or gas chromatography, or even capillary electrophoresis can be used with this aim, followed by a detector the features of which establish the less or higher selectivity of the determination and also the need for total or partial separation. Thus, molecular absorption detection (usually by a diode array detector –DAD) involves the use of the retention time and spectral features of the analytes together with their standards for proper quantitation. Similar requirements, but a more sensitive detection is achieved by a fluorescence detector, preferably with excitation by a laser source; nevertheless, this type of detection makes mandatory intrinsic emission of the analyte or a derivatization reaction to yield an emissive product. Also sensitive quantitation but with higher acquisition and maintenance costs is provided by mass detectors.

Some examples of the determination of given phenols families is in the research by Ling *et al.* who proposed a rapid and sensitive method for quantitation of four anthocyanins using a liquid chromatograph coupled to a mass spectrometer (LC-MS) (Ling *et al.* 2009).

Other alternative to quantify anthocyanins was reported by Oh *et al.* (Oh *et al.* 2008). The characterization and quantitation of anthocyanins in grape cultivars were carried out by partial purification through XAD-7 cartridges followed by C-18 LC-DAD, LC-MS, and LC-MSMS analyses.

(iii) Comprehensive characterization

Liquid chromatography coupled to hybrid quadrupole time-of-flight detection (LC-QqTOF) is one of the most sensitive and accurate approaches for comprehensive characterization. The LC-QqTOF assembly is at present an unquestionable powerful tool for separation and characterization of phenols. The detector provides accurate mass determination of components in complex

mixtures and, subsequently, the accurate composition can be suggested considering other chemical information, such as neutral loss and product ions. The system is very useful for characterization of most compounds in complex matrices, and also provides key information about unknown compounds and isomers, which is pivotal, most times, for characterization of natural products from plants.

Perestrelo *et al.* reported the first phytochemical research on phenolic compounds of Sercial and Tinta Negra Vitis vinifera L (Perestrelo *et al.* 2012). The phenol profile of grape skins (white and red varieties, respectively) was analyzed using an LC-DAD-ESI-MS^n, at different ripening stages. A total of 40 phenols were identified, including 3 hydroxybenzoic acids, 8 hydroxycinnamic acids, 4 flavanols, 5 flavanones, 8 flavonols, 4 stilbenes, and 8 anthocyanins; identification that was based on their MS^n fragmentation profile. These data represent valuable information that may be useful for oenological management and valorization of these varieties as a source of bioactive compounds.

A recent study describes for the first time a complete characterization of phenols in different parts of Albariño grape (Di Lecce *et al.* 2014). Complementary equipment (LC-QqTOF and LC coupled to triple quadrupole mass spectrometry –LC-QQQ– were used to identify and quantify the phenolic composition of the given varietal. Furthermore, a more complete phenolic profile was obtained by product ion and precursor ion scans, while a neutral loss scan at 152 mass units enabled a fast screening of procyanidin dimers, trimers and their galloylated derivatives. Confirmation by accurate mass measurements in TOF/MS and TOF-MS/MS modes at high resolution provided good fits for all investigated ions, with errors ranging from 0.2 to 4.5 mDa.

5. Potential biomedical effects of phenols

Phenols are secondary metabolites with key roles in plants as are pigmentation to attract pollinators and seed dispersers, and defence of the whole plant against both pathogenic microorganisms and predators (Petti *et al.* 2009).

Concerning human health, a number of studies support the protective effect of phenols against the oxidative damage associated to development of most major age related degenerative diseases, thus decreasing risk for cancer, cardiovascular diseases and neurodegenerative disorders (Manach *et al.* 2004; Petti *et al.* 2009).

Phenols effects depend on the amount and periodicity of intake. The available information on the quantities of phenols daily consumed over the world (Manach *et al.* 2004) indicates a mean daily intake by adults of 863 mg, of which phenolic acids represent 75% of total intake, followed by proanthocyanidins (14%), anthocyanidins and other flavonoids (10%). The main sources for intake are coffee, cereals, fruits and vegetables (Petti *et al.* 2009).

Furthermore, it is important to note that despite phenols are common in the human diet, they are not necessarily active in the organism, either because they are endowed with intrinsic activity not always high, or because they are poorly absorbed from the intestine, highly metabolized, or rapidly eliminated (Manach *et al.* 2004). Little is known about phenols absorption, distribution and metabolism, despite some authors propose a common pathway as probable (Petti *et al.* 2009). As a general rule, the metabolites are rapidly eliminated from plasma, which indicates that consumption of plant products is necessary on a daily basis to maintain the needed concentrations of these metabolites in blood (Graefe *et al.* 2001; Hollman *et al.* 1997; Manach *et al.* 1998).

Petti *et al.* (Petti *et al.* 2009) reviewed the studies published between 1993 and 2008 with the aim of finding a relationship between phenols and dental diseases. The results showed that the antioxidant capacity of phenols is

potentially valid to explain their preventive effect against diseases of the oral cavity (Halliwell *et al.* 2000). Besides, catechins, methoxylated flavonoids and proanthocyanidins proved potential preventive activity against oral squamous cell carcinoma (Ho *et al.* 2007; Hsu *et al.* 2001, Khafif *et al.* 1998). Also, diets rich in antioxidants seem to inhibit development and progression of periodontal diseases, particularly in subjects exposed to environmental and dietary sources of oxidative stress (Battino *et al.* 1999; Ritchie *et al.* 2003; Sculley *et al.* 2002).

Research on the effect of phenols on Alzheimer's disease concluded that these compounds may exert a variety of modulatory actions with positive influence on the pathogenesis of this disease (Jayasena *et al.* 2013). Antidiabetic effects seem also to be exerted by phenols, either directly (dependent on absorption) or indirectly (dependent on gut microbiota modulation and pancreatic lipase inhibition) (Anhê *et al.* 2013).

In short, phenols seem to exert benefits on human health thanks to (among others) their antioxidants properties.

6. Other phenols effects

Other interesting characteristics that make phenols useful are related to their use both in cosmetics and food industries.

Bittar *et al.* reported a study in which fresh grape berries were pressed and the residue from juice was extracted by a green extraction technique – microwave hydrodiffusion and gravity (Al Bittar *et al.* 2013)– to obtain phenols from it. Then, the extract was added to the total volume of natural juice to produce a juice enriched with phenols. As a result, the supplemented juice had a more attractive color and acceptable organoleptic characteristics. No studies on the effect of this juice as compared to conventional juice have been developed.

A study looking for the relationship between antimicrobial activity and phenols content was developed using extracts from grape seeds and bagasse

obtained using two different solvent extractions. The study concluded that the grape seed extracts had high total phenolics compared with those of bagasse, which did not inhibit any of the bacteria tested. Similarly, Shoko *et al.* (Shoko *et al.* 1999) confirmed that phenols were the most important compounds active against bacteria. They also identified gallic acid as the most active compound for inhibition of bacteria. Their results suggest that the use of grape seed extracts may be exploitable as antibacterial agents to prevent the deterioration of stored foods by bacteria, if any organoleptic effects are acceptable.

Osmotic treatment (OT) is widely used to modify the composition of solid foods by partially removing water and adding solutes. Recently, Rózek *et al.* reported that OT is a unit operation to introduce into a food matrix controlled quantities of solutes with antioxidant properties such as grape phenols. Using concentrated red grape must as a source of phenols, they found that the phenolic content in an osmotreated solid food was similar to that of some fruits rich in phenols (Rozek *et al.* 2007). Then, they tried to get other commercial sources of phenols, such as grape seed and skin extracts, and used these extracts to infuse a model food with grape phenols (Rozek *et al.* 2008; 2009); finally, they studied the influence of the source of grape phenols and the phenol concentration of the osmotic solution on the properties of the osmotreated food (Rozet *et al.* 2010a; 2010b).

The interest of winemakers in the phenol content of grapes is increasing, as it offers ways of influencing the color, bitterness, astringency, ‘‘mouth feel” and ‘‘age ability” of wines. Content in phenols is one of the main factors in the quality of grapes and wine. The phenolic composition of a wine depends primarily on the phenol content of the grapes and secondarily on the winemaking techniques employed. In a recent study, grapes of three varieties autochthonous from Andalusia (Jaén tinto, Palomino negro and Tintilla de Rota) and two others of the most commonly used in Spain, and all over the world (Cabernet Sauvignon and Tempranillo) were analyzed by LC–DAD–MS and compared with the aim to identify and quantify their phenol composition. The results of the characterization

can be used to select winemaking techniques aimed at improving the quality of the final wine (Guerrero *et al.* 2009). Schwarz *et al.* also studied the relationship between antioxidant activity and phenolic content and the correlation of these values with the wine age analyzing commercial and experimental sherry brandies. They found a significant relationship between them; however, it seems that some non phenol compounds could also contribute to the antioxidant activity. The results revealed certain differences between commercial and experimental sherry brandies; differences that could be due to the addition of caramel coloring. Besides, the antioxidant power of Brandy de Jerez was compared with other commercial aged distillates of various different geographical origins showing the former the highest antioxidant activity of all the products tested (Schwartz *et al.* 2009).

The volatile fraction of wine is composed by several hundreds of chemically different compounds. The presence or absence of some of the compounds in a particular wine depends on several factors such as environment (climate and soil), ripeness and grape variety, winemaking conditions, and aging, among the most significant. Some of these compounds are already present in the grapes, others are formed during wine fermentation and aging. Ruiz-Bejarano *et al.* studied the evolution, under different aging procedures, of the main volatile compounds of sweet Sherry wines obtained from Muscat grapes dried using drying chambers and fermented under different enological conditions (Ruiz-Bejarano *et al.* 2013). Therefore, two factors and their possible influence on the volatile profile were considered together: aging and vinification conditions. They observed a key influence of the three factors considered, fermentation type, aging type, and aging time, on volatile compounds and sensory characteristics of the wines being the aging procedure the highest influential factor on the volatile content of the wines (Ruiz-Bejarano *et al.* 2013).

The relationship between the content of phenols and furan compounds in commercial Sherry brandies and other aged distillates from different geographical origin has been studied to know the influence of the relative content of these

families to explain the highly specific character of Brandy de Jerez. The aging system was confirmed as having a great influence on the analytical profile of aged distillates (discrimination of up to 100%). Differences between commercial brandies and those aged experimentally with equivalent average age have also been confirmed by a clear differentiation between Solera Gran Reserva Brandies de Jerez and the rest of the distillates of different origin (discrimination of up to 80%), indicating their highly specific character (Rodríguez *et al.* 2010).

Conclusion

Despite the uses discussed in this chapter, an integral exploitation of both grape byproducts and residues generated by grape byproducts ("residues from residues") is not a present reality. Research about exploitation of these residues is scant, despite some of the results have demonstrated the usefulness of them as sources of compounds of great interest. As an example, the press residues of grape seed oil production as a phenols source has been studied by Maier *et al.* (Maier *et al.* 2009) classifying this residue as a phenol rich by product with antioxidant activity.

Peralbo-Molina *et al.* proposed an extraction method with interesting analytical characteristics for a favorable isolation of priced compounds from a raw material with very scant value: residues from distillery industries using the byproduct after alcohol distillation from red and white grape pomace (Peralbo-Molina *et al.* 2012; 2013). Besides, the identification of interesting compounds in the extracts showed that despite the degradation caused by the drastic conditions of the distillation process, the waste from this industry constitutes a useful raw matter for more than as a heating source. The variety of identified compounds in the extracts makes them exploitable as additives in the food industry (either as colorants, as flavor modifiers or as antioxidants), and also in the cosmetics and nutraceutical industries.

In addition, Farías-Campomanes *et al.* demonstrated the economic feasibility of large scale operations of supercritical fluid extraction for recovery of phenolic compounds using grape bagasse from Pisco residues (Farías Campomanes *et al.* 2013).

References

Al Bittar S, Perino-Issartier S, Dangles O, Chemat F. An innovative grape juice enriched in polyphenols by microwave-assisted extraction. Food Chem. 2013;141(3):3268-72.

Albu S, Joyce E, Paniwnyk L, Lorimer JP, Mason TJ. Potential for the use of ultrasound in the extraction of antioxidants from Rosmarinus officinalis for the food and pharmaceutical industry. Ultrason Sonochem. 2004;11(3-4):261–75.

Ali K, Maltese F, Choi YH, Verpoorte R. Metabolic constituents of grapevine and grape-derived products. Phytochem Rev. 2010;9(3):357–78.

Andrade PB, Oliveira BM, Seabra RM, Ferreira MA, Ferreres F, García-Viguera C. Analysis of phenolic compounds in Spanish Albariño and Portuguese Alvarinho and Loureiro wines by capillary zone electrophoresis and high-performance liquid chromatography. Electrophoresis. 2001;22(8):1568–72.

Anhê FF, Desjardins Y, Pilon G, Dudonné S, Genovese MI, Lajolo FM, *et al.* Polyphenols and type 2 diabetes: a prospective review. PharmaNutrition. (DOI: 10.1016/j.phanu.2013.07.004).

Battino M, Bullon P, Wilson M, Newman H. Oxidative injury and inflammatory periodontal diseases: the challenge of antioxidants to free radicals and reactive oxygen species. Crit Rev Oral Biol Med. 1999;10(4):458–76.

Benzie IFF, Strain JJ. The ferric reducing ability of plasma (FRAP) as a measure of “antioxidant power”: the FRAP assay. Anal Biochem. 1996;239(1):70-6.

Benzie IF, Chung W, Strain JJ. "Antioxidant" (reducing) efficiency of ascorbate in plasma is not affected by concentration. J Nutr Biochem. 1999;10(3):146-50.

Biesalski HK. Diabetes preventive components in the Mediterranean diet. Eur J Nutr. 2004;43(1):26-30.

Bonghi C, Rizzini FM, Gambuti A, Moio L, Chkaiban L, Tonutti P. Phenol compound metabolism and gene expression in the skin of wine grape (Vitis vinifera L.) berries subjected to partial postharvest dehydration. Postharvest Biol Technol. 2012;67(0):102-9.

Boots AW, Wilms LC, Swennen ELR, Kleinjans JCS, Bast A, Haenen GRMM. In vitro and ex vivo anti-inflammatory activity of quercetin in healthy volunteers. Nutrition. 2008;24(7-8):703-10.

Bravo L. Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance. Nutr Rev. 1998;56(11):317-33.

Caillet S, Salmiéri S, Lacroix M. Evaluation of free radical-scavenging properties of commercial grape phenol extracts by a fast colorimetric method. Food Chem. 2006;95(1):1-8.

Cantos E, Espín JC, Tomás-Barberán FA. Postharvest stilbene-enrichment of red and white table grape varieties using UV-C irradiation pulses. J Agric Food Chem. 2002;50(22):6322-9.

Casas L, Mantell C, Rodríguez M, Ossa EJMdl, Roldán A, Ory ID, *et al.* Extraction of resveratrol from the pomace of Palomino fino grapes by supercritical carbon dioxide. J Food Eng. 2010;96(2):304-8.

Castillo-Muñoz N, Gómez-Alonso S, García-Romero E, Hermosín-Gutiérrez I. Flavonol profiles of Vitis vinifera red grapes and their single-cultivar wines. J Agric Food Chem. 2007;55(3):992-1002.

Chafer A, Pascual-Marti MC, Salvador A, Berna A. Supercritical fluid extraction and HPLC determination of relevant polyphenolic compounds in grape skin. J Sep Sci. 2005;28(16):2050-6.

Cho Y-J, Hong J-Y, Chun HS, Lee SK, Min H-Y. Ultrasonication-assisted extraction of resveratrol from grapes. J Food Eng. 2006;77(3):725-30.

Choi HK. Dietary risk factors for rheumatic diseases. Curr Opin Rheumatol. 2005;17(2):141-6.

Corrales M, Fernández A, Vizoso Pinto MG, Butz P, Franz CM, Schuele E, *et al.* Characterization of phenolic content, in vitro biological activity, and pesticide loads of extracts from white grape skins from organic and conventional cultivars. Food Chem Toxicol. 2010;48(12):3471-6.

Corrales M, García AF, Butz P, Tauscher B. Extraction of anthocyanins from grape skins assisted by high hydrostatic pressure. J Food Eng. 2009;90(4):415-21.

Cravotto G, Boffa L, Mantegna S, Perego P, Avogadro M, Cintas P. Improved extraction of vegetable oils under high-intensity ultrasound and/or microwaves. Ultrason Sonochem. 2008;15(5):898-902.

D'Archivio M, Filesi C, Di Benedetto R, Gargiulo R, Giovannini C, Masella R. Polyphenols, dietary sources and bioavailability. Ann Ist Super Sanita. 2007;43(4):348-61.

Darias-Martín J, Carrillo M, Díaz E, Boulton RB. Enhancement of red wine colour by pre-fermentation addition of copigments. Food Chemistry. 2001;73(2):217-20.

de Campos LM, Leimann FV, Pedrosa RC, Ferreira SRS. Free radical scavenging of grape pomace extracts from Cabernet sauvingnon (Vitis vinifera). Bioresource Technol. 2008;99(17):8413-20.

DeLange RJ, Glazer AN. Phycoerythrin fluorescence-based assay for peroxy radicals: a screen for biologically relevant protective agents. Anal Biochem. 1989;177(2):300–6.

Di Lecce G, Arranz S, Jáuregui O, Tresserra-Rimbau A, Quifer-Rada P, Lamuela-Raventós RM. Phenolic profiling of the skin, pulp and seeds of Albariño grapes using hybrid quadrupole time-of-flight and triple-quadrupole mass spectrometry. Food Chem. 2014;145:874–82.

Di Stefano R, Cravero MC, Gentilizi M. Metodi per lo studio dei polifenoli dei vini. L'Enotecnico. 1989;May:83–9.

Dryden GW, Song M, McClain C. Polyphenols and gastrointestinal diseases. Curr Opin Gastroenterol. 2006;22(2):165–70.

Faller ALK, Fialho E. Polyphenol content and antioxidant capacity in organic and conventional plant foods. J Food Comp Anal. 2010;23(6):561–8.

Farías-Campomanes AM, Rostagno MA, Meireles MAA. Production of polyphenol extracts from grape bagasse using supercritical fluids: yield, extract composition and economic evaluation. J Supercrit Fluids. 2013;77(0):70–8.

Feinleib M. Seven countries: A multivariate analysis of death and coronary heart disease. JAMA. 1981;245(5):511–2.

Figueiredo-González M, Cancho-Grande B, Simal-Gándara J. Evolution of colour and phenolic compounds during Garnacha Tintorera grape raisining. Food Chem. 2013;141(3):3230–40.

Flamini R. Mass spectrometry in grape and wine chemistry. Part I: Polyphenols. Mass Spectrom Rev. 2003;22(4):218–50.

Food and Agriculture Organization of the United Nations. 2013.

Fukai K, Ishigami T, Kara Y. Antibacterial activity of tea polyphenols against phytopathogenic bacteria. Agric Biol Chem. 1991;55(7):1895–7.

Gambelli L, Santaroni GP. Polyphenols content in some Italian red wines of different geographical origins. J Food Comp Anal. 2004;17(5):613–8.

Garrido J, Borges F. Wine and grape polyphenols – A chemical perspective. Food Res Internat. (DOI: 10.1016/j.foodres.2013.08.002).

Giacosa A, Barale R, Bavaresco L, Gatenby P, Gerbi V, Janssens J, *et al.* Cancer prevention in Europe: the Mediterranean diet as a protective choice. Eur J Cancer Prev. 2013;22(1):90–5.

González-Rodríguez J, Pérez-Juan P, Luque de Castro MD. Superheated liquids for extraction of solid residues from winemaking processes. Anal Bioanal Chem. 2003;377(7-8):1190–5.

Graefe EU, Wittig J, Mueller S, Riethling AK, Uehleke B, Drewelow B, *et al.* Pharmacokinetics and bioavailability of quercetin glycosides in humans. J Clin Pharmacol. 2001;41(5):492–9.

Grigonis D, Venskutonis PR, Sivik B, Sandahl M, Eskilsson CS. Comparison of different extraction techniques for isolation of antioxidants from sweet grass (Hierochloë odorata). J Supercrit Fluids. 2005;33(3):223–33.

Gruber JAN, Tang SY, Halliwell B. Evidence for a trade-off between survival and fitness caused by resveratrol treatment of Caenorhabditis elegans. Ann New York Academy Sci. 2007;1100(1):530–42.

Guerrero RF, García-Parrilla MC, Puertas B, Cantos-Villar E. Wine, resveratrol and health: a review. Nat Prod Commun. 2009;4(5):635–58.

Guerrero RF, Liazid A, Palma M, Puertas B, González-Barrio R, Gil-Izquierdo A, *et al.* Phenolic characterisation of red grapes autochthonous to Andalusia. Food Chem. 2009;112(4):949–55.

Halliwell B, Zhao K, Whiteman M. The gastrointestinal tract: a major site of antioxidant action. Free Radic Res. 2000;33(6):819–30.

Halvorsen BL, Holte K, Myhrstad MCW, Barikmo I, Hvattum E, Remberg SF, *et al.* A systematic screening of total antioxidants in dietary plants. J Nutrition. 2002;132(3):461–71.

Harborne, J.B. The flavonoids: advances in research since 1986. J Chem Educ. 1995;72(3):A73.

Hmamouchi M, Es-Safi N, Lahrichi M, Fruchier A, Essassi EM. Flavones and flavonols in leaves of some Moroccan Vitis vinifera cultivars. Am J Enol Viticulture. 1996;47(2):186–92.

Ho YC, Yang SF, Peng CY, Chou MY, Chang YC. Epigallocatechin-3-gallate inhibits the invasion of human oral cancer cells and decreases the production of matrix metalloproteinases and urokinase-plasminogen activator. J Oral Pathol Med. 2007;36(10):588–93.

Hollman PC, van Trijp JM, Buysman MN, van der Gaag MS, Mengelers MJ, de Vries JH, *et al.* Relative bioavailability of the antioxidant flavonoid quercetin from various foods in man. FEBS Lett. 1997;418(1-2):152–6.

Hsu S, Lewis JB, Borke JL, Singh B, Dickinson DP, Caughman GB, *et al.* Chemopreventive effects of green tea polyphenols correlate with reversible induction of p57 expression. Anticancer Res. 2001;21(6A):3743–8.

Hung LM, Chen JK, Huang SS, Lee RS, Su MJ. Cardioprotective effect of resveratrol, a natural antioxidant derived from grapes. Cardiovasc Res. 2000;47(3):549–55.

Jaitz L, Siegl K, Eder R, Rak G, Abranko L, Koellensperger G, *et al.* LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. Food Chem. 2010;122(1):366–72.

Jayasena T, Poljak A, Smythe G, Braidy N, Munch G, Sachdev P. The role of polyphenols in the modulation of sirtuins and other pathways involved in Alzheimer's disease. Ageing Res Rev. 2013;12(4):867–83.

Jiménez L, Angulo V, Ramos E, de la Torre MJ, Ferrer JL. Comparison of various pulping processes for producing pulp from vine shoots. Industrial Crops Products. 2006;23(2):122–30.

Ju ZY, Howard LR. Effects of solvent and temperature on pressurized liquid extraction of anthocyanins and total phenolics from dried red grape skin. J Agric Food Chem. 2003;51(18):5207–13.

Kammerer D, Claus A, Schieber A, Carle R. A novel process for the recovery of polyphenols from grape (Vitis vinifera L.) pomace. J Food Sci. 2005;70(2):C157–C63.

Kammerer D, Gajdos Kljusuric J, Carle R, Schieber A. Recovery of anthocyanins from grape pomace extracts (Vitis vinifera L. cv. Cabernet Mitos) using a polymeric adsorber resin. Eur Food Res Technol. 2005;220(3-4):431–7.

Khafif A, Schantz SP, al-Rawi M, Edelstein D, Sacks PG. Green tea regulates cell cycle progression in oral leukoplakia. Head Neck. 1998;20(6):528–34.

Khanal RC, Howard LR, Prior RL. Effect of heating on the stability of grape and blueberry pomace procyanidins and total anthocyanins. Food Res Internat. 2010;43(5):1464–9.

Keys, A. Coronary heart disease in seven countries. Summary. Circulation. 1970;41(4 Suppl):186–95.

Kok FJ, Kromhout D. Atherosclerosis-epidemiological studies on the health effects of a Mediterranean diet. Eur J Nutr. 2004;43(1):2–5.

Kushi LH, Lenart EB, Willett WC. Health implications of Mediterranean diets in light of contemporary knowledge. 1. Plant foods and dairy products. Am J Clin Nutr. 1995;61(6 Suppl):1407S–15S.

Liazid A, Guerrero RF, Cantos E, Palma M, García-Barroso C. Microwave assisted extraction of anthocyanins from grape skins. Food Chem. 2011;124(3):1238–43.

Ling Y, Ren C, Mallery SR, Ugalde CM, Pei P, Saradhi UVRV, *et al.* A rapid and sensitive LC–MS/MS method for quantification of four anthocyanins and its application in a clinical pharmacology study of a bioadhesive black raspberry gel. J Chromatogr B. 2009;877(31):4027–34.

López-Sánchez C, Martín-Romero FJ, Sun F, Luis L, Samhan-Arias AK, García-Martí V, *et al.* Blood micromolar concentrations of kaempferol afford protection against ischemia/reperfusion-induced damage in rat brain. Brain Res. 2007;28:123–37.

Luque-Rodríguez JM, Pérez-Juan P, Luque de Castro MD. Dynamic superheated liquid extraction of anthocyanins and other phenolics from red grape skins of winemaking residues. Bioresource Technol. 2007;98(14):2705–13.

Luque-Rodríguez JL, Japón-Luján R, Luque de Castro MD. Exploitation of Residues from Vineyards, Olive Groves, and Wine and Oil Production to Obtain Phenolic Compounds of High-Added Value in Methods of Analysis for Functional Foods and Nutraceuticals, Second Edition: CRC Press; 2008.

Maier T, Schieber A, Kammerer DR, Carle R. Residues of grape (Vitis vinifera L.) seed oil production as a valuable source of phenolic antioxidants. Food Chem. 2009;112(3):551–9.

Makris DP, Boskou G, Andrikopoulos NK. Polyphenolic content and in vitro antioxidant characteristics of wine industry and other agri-food solid waste extracts. J Food Comp Anal. 2007;20(2):125–32.

Manach C, Morand C, Crespy V, Demigne C, Texier O, Regerat F, *et al.* Quercetin is recovered in human plasma as conjugated derivatives which retain antioxidant properties. FEBS Lett. 1998;426(3):331–6.

Manach C, Scalbert A, Morand C, Remesy C, Jiménez L. Polyphenols: food sources and bioavailability. Am J Clin Nutr. 2004;79(5):727–47.

Mantell C, de la Ossa EM, Rodríguez M. Supercritical Fluid Extraction of Anthocyanins from Grape Pomace. In: Hidalgo M, Valiente M, editors. Solvent Extraction for the 21st Century Proceedings of ISEC'99. Barcelona, Spain: Society of Chemical Industry; 2001.

Mantell C, Rodríguez M, de la Ossa EM. Measurement of the diffusion coefficient of a model food dye (malvidin 3,5-diglucoside) in a high pressure CO2+methanol system by the chromatographic peak-broadening technique. J Supercrit Fluids. 2003;25(1):57–68.

Márquez A, Serratosa MP, López-Toledano A, Mérida J. Colour and phenolic compounds in sweet red wines from Merlot and Tempranillo grapes chamber-dried under controlled conditions. Food Chem. 2012;130(1):111–20.

Mattivi F, Guzzon R, Vrhovsek U, Stefanini M, Velasco R. Metabolite profiling of grape: flavonols and anthocyanins. J Agric Food Chem. 2006;54(20):7692–702.

Miller NJ, Rice-Evans C, Davies MJ, Gopinathan V, Milner A. A novel method for measuring antioxidant capacity and its application to monitoring the antioxidant status in premature neonates. Clin Sci. 1993;84(4):407–12.

Mira de Orduña R. Climate change associated effects on grape and wine quality and production. Food Res Internat. 2010;43(7):1844–55.

Modun D, Music I, Vukovic J, Brizic I, Katalinic V, Obad A, *et al.* The increase in human plasma antioxidant capacity after red wine consumption is due to both plasma urate and wine polyphenols. Atherosclerosis. 2008;197(1):250–6.

Moini H, Guo Q, Packer L. Enzyme inhibition and protein-binding action of the procyanidin-rich French maritime pine bark extract, pycnogenol: effect on xanthine oxidase. J Agric Food Chem. 2000;48(11):5630-9.

Monagas M, Bartolomé B, Gómez-Cordovés C. Evolution of polyphenols in red wines from Vitis vinifera L. during aging in the bottle. Eur Food Res Technol. 2005;220(3-4):331–40.

Morales S, Canosa P, Rodríguez I, Rubi E, Cela R. Microwave assisted extraction followed by gas chromatography with tandem mass spectrometry for the determination of triclosan and two related chlorophenols in sludge and sediments. J Chromatogr A. 2005;5(2):128–35.

Morata A, González C, Suárez-Lepe JA. Formation of vinylphenolic pyranoanthocyanins by selected yeasts fermenting red grape musts supplemented with hydroxycinnamic acids. Internat J Food Microbiol. 2007;116(1):144–52.

Mulero J, Pardo F, Zafrilla P. Antioxidant activity and phenolic composition of organic and conventional grapes and wines. J Food Comp Anal. 2010;23(6):569–74.

Nagel CW, Glories Y. Use of a modified dimethylaminocinnamaldehyde reagent for analysis of flavanols. Am J Enol Viticulture. 1991;42(4):364–6.

Negro C, Tommasi L, Miceli A. Phenolic compounds and antioxidant activity from red grape marc extracts. Bioresource Technol. 2003;87(1):41–4.

Oh YS, Lee JH, Yoon SH, Oh CH, Choi DS, Choe E, *et al.* Characterization and quantification of anthocyanins in grape juices obtained from the grapes cultivated in Korea by HPLC/DAD, HPLC/MS, and HPLC/MS/MS. J Food Sci. 2008;73(5):C378–C89.

Orak HH. Total antioxidant activities, phenolics, anthocyanins, polyphenoloxidase activities of selected red grape cultivars and their correlations. Scientia Horticulturae. 2007;111(3):235–41.

Peinado J, López de Lerma N, Moreno J, Peinado RA. Antioxidant activity of different phenolics fractions isolated in must from Pedro Ximenez

grapes at different stages of the off-vine drying process. Food Chem. 2009;114(3):1050–5.

Peng Z, Hayasaka Y, Iland PG, Sefton M, Hoj P, Waters EJ. Quantitative analysis of polymeric procyanidins (Tannins) from grape (Vitis vinifera) seeds by reverse phase high-performance liquid chromatography. J Agric Food Chem. 2001;49(1):26–31.

Peralbo-Molina A, Priego-Capote F, Luque de Castro MD. Comparison of extraction methods for exploitation of grape skin residues from ethanol distillation. Talanta. 2012;101(0):292–8.

Peralbo-Molina A, Priego-Capote F, Luque de Castro MD. Characterization of grape seed residues from the ethanol-distillation industry. Anal Methods. 2013;5(8):1922–9.

Peralbo-Molina A, Luque de Castro MD. Potential of residues from the Mediterranean agriculture and agrifood industry. Trends Food Sci Technol. 2013;32(1):16–24.

Perestrelo R, Lu Y, Santos SAO, Silvestre AJD, Neto CP, Câmara JS, *et al.* Phenolic profile of Sercial and Tinta Negra Vitis vinifera L. grape skins by HPLC-DAD-ESI-MSn: novel phenolic compounds in Vitis vinifera L. grape. Food Chem. 2012;135(1):94–104.

Pérez-Serradilla JA, Luque de Castro MD. Role of lees in wine production: A review. Food Chem. 2008;111(2):447–56.

Pérez-Serradilla JA, Luque de Castro MD. Microwave-assisted extraction of phenolic compounds from wine lees and spray-drying of the extract. Food Chem. 2011;124(4):1652–9.

Petti S, Scully C. Polyphenols, oral health and disease: a review. J Dent. 2009;37(6):413–23.

Prados-Rosales RC, Luque García JL, Luque de Castro MD. Rapid analytical method for the determination of pesticide residues in sunflower seeds based on focused microwave-assisted Soxhlet extraction prior to gas chromatography–tandem mass spectrometry. J Chromatogr A. 2003;993(1-2):121-9.

Pietta P, Minoggio M, Bramati L. Plant Polyphenols: Structure, Occurrence and Bioactivity. In: Atta-ur R, editor. Studies in Natural Products Chemistry: Elsevier; 2003. p. 257-312.

Piñeiro Z, Palma M, García-Barroso C. Determination of *trans*-resveratrol in grapes by pressurised liquid extraction and fast high-performance liquid chromatography. J Chromatogr A. 2006;1110(1-2):61-5.

Prados-Rosales RC, Luque García JL, Luque de Castro MD. Rapid analytical method for the determination of pesticide residues in sunflower seeds based on focused microwave-assisted Soxhlet extraction prior to gas chromatography–tandem mass spectrometry. J Chromatogr A. 2003;993(1-2):121-9.

Prior RL, Cao G, Martin A, Sofic E, McEwen J, O'Brien C, *et al.* Antioxidant capacity as influenced by total phenolic and anthocyanin content, maturity, and variety of Vaccinium species. J Agric Food Chem. 1998;46(7):2686-93.

Proteggente AR, Pannala AS, Paganga G, Van Buren L, Wagner E, Wiseman S, *et al.* The antioxidant activity of regularly consumed fruit and vegetables reflects their phenolic and vitamin C composition. Free Radic Res. 2002;36(2):217-33.

Ribeiro De Lima MT, Waffo-Téguo P, Teissedre PL, Pujolas A, Vercauteren J, Cabanis JC, *et al.* Determination of stilbenes (*trans*-astringin, *cis*- and *trans*-piceid, and *cis*- and *trans*-resveratrol) in Portuguese wines. J Agric Food Chem. 1999;47(7):2666-70.

Ritchie CS, Kinane DF. Nutrition, inflammation, and periodontal disease. Nutrition. 2003;19(5):475-6.

Robards K. Strategies for the determination of bioactive phenols in plants, fruit and vegetables. J Chromatogr A. 2003;1000(1-2):657-91.

Rodríguez Dodero MC, Guillén Sánchez DA, Rodríguez MS, García-Barroso C. Phenolic compounds and furanic derivatives in the characterization and quality control of Brandy de Jerez. J Agric Food Chem. 2010;58(2):990-7.

Rodríguez Montealegre R, Romero Peces R, Chacón Vozmediano JL, Martínez Gascueña J, García Romero E. Phenolic compounds in skins and seeds of ten grape Vitis vinifera varieties grown in a warm climate. J Food Comp Anal. 2006;19(6-7):687-93.

Roggero JP, García-Parrilla C. Effects of ultraviolet irradiation on resveratrol and changes in resveratrol and various of its derivatives in the skins of ripening grapes. Sci Aliments. 1995;15(5):411-22.

Romero-Pérez AI, Lamuela-Raventós RM, Andrés-Lacueva C, de la Torre-Boronat MC. Method for the quantitative extraction of resveratrol and piceid isomers in grape berry skins. Effect of powdery mildew on the stilbene content. J Agric Food Chem. 2001;49(1):210-5.

Rozek A, Achaerandio I, Almajano MP, Guell C, López F, Ferrando M. Solid foodstuff supplemented with phenolics from grape: antioxidant properties and correlation with phenolic profiles. J Agric Food Chem. 2007;55(13):5147-55.

Rózek A, Achaerandio I, Güell C, López F, Ferrando M. Direct formulation of a solid foodstuff with phenolic-rich multicomponent solutions from grape seed: effects on composition and antioxidant properties. J Agric Food Chem. 2008;56(12):4564-76.

Rózek A, Achaerandio I, Güell C, López F, Ferrando M. Grape phenolic impregnation by osmotic treatment: influence of osmotic agent on mass transfer and product characteristics. J Food Eng. 2009;94(1):59–68.

Rózek A, Achaerandio I, Güell C, López F, Ferrando M. Use of commercial grape phenolic extracts to supplement solid foodstuff. LWT - Food Sci Technol. 2010;43(4):623–31.

Rózek A, García-Pérez JV, López F, Güell C, Ferrando M. Infusion of grape phenolics into fruits and vegetables by osmotic treatment: phenolic stability during air drying. J Food Eng. 2010;99(2):142–50.

Ruiz-Bejarano M, Castro-Mejías R, Rodríguez-Dodero MD, García-Barroso C. Study of the content in volatile compounds during the aging of sweet Sherry wines obtained from grapes cv. Muscat and fermented under different conditions. Eur Food Res Technol. 2013:1–18.

Santos-Buelga C, Williamson G, Chemistry RS. Methods in Polyphenol Analysis: Royal Society of Chemistry; 2003.

Schwarz M, Rodríguez M, Martínez C, Bosquet V, Guillén D, García-Barroso C. Antioxidant activity of Brandy de Jerez and other aged distillates, and correlation with their polyphenolic content. Food Chem. 2009;116(1):29–33.

Schwarz M, Picazo-Bacete JJ, Winterhalter P, Hermosin-Gutiérrez I. Effect of copigments and grape cultivar on the color of red wines fermented after the addition of copigments. J Agric Food Chem. 2005;53(21):8372–81.

Sculley DV, Langley-Evans SC. Salivary antioxidants and periodontal disease status. Proc Nutr Soc. 2002;61(1):137–43.

Serafini M, Laranjinha JAN, Almeida LM, Maiani G. Inhibition of human LDL lipid peroxidation by phenol-rich beverages and their impact on plasma total antioxidant capacity in humans. J Nutritional Biochem. 2000;11(11–12):585–90.

Serrano A, Gallego M. Continuous microwave-assisted extraction coupled on-line with liquid–liquid extraction: determination of aliphatic hydrocarbons in soil and sediments. J Chromatogr A. 2006;1104(1–2):323–30.

Serratosa MP, López-Toledano A, Mérida J, Medina M. Changes in color and phenolic compounds during the raisining of grape cv. Pedro Ximenez. J Agric Food Chem. 2008;56(8):2810–6.

Serratosa MP, Lopez-Toledano A, Medina M, Merida J. Drying of Pedro Ximenez grapes in chamber at controlled temperature and with dipping pretreatments. Changes in the color fraction. J Agric Food Chem. 2008;56(22):10739-46.

Serratosa MP, Márquez A, López-Toledano A, Medina M, Mérida J. Changes in hydrophilic and lipophilic antioxidant activity in relation to their phenolic composition during the chamber drying of red grapes at a controlled temperature. J Agric Food Chem. 2011;59(5):1882–92.

Shi J, Nawaz H, Pohorly J, Mittal G, Kakuda Y, Jiang Y. Extraction of polyphenolics from plant material for functional foods—engineering and technology. Food Rev Internat. 2005;21(1):139–66.

Singleton VL, Rossi JA. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. Am J Enol Viticulture. 1965;16(3):144–58.

Skogerson K, Downey M, Mazza M, Boulton R. Rapid determination of phenolic components in red wines from UV-visible spectra and the method of partial least squares. Am J Enol Viticulture. 2007;58(3):318–25.

Souquet JM, Labarbe B, Le Guerneve C, Cheynier V, Moutounet M. Phenolic composition of grape stems. J Agric Food Chem. 2000;48(4):1076–80.

Spigno G, De Faveri DM. Microwave-assisted extraction of tea phenols: a phenomenological study. J Food Eng. 2009;93(2):210-7.

Spormann TM, Albert FW, Rath T, Dietrich H, Will F, Stockis JP, *et al.* Anthocyanin/polyphenolic-rich fruit juice reduces oxidative cell damage in an intervention study with patients on hemodialysis. Cancer Epidemiol Biomarkers Prev. 2008;17(12):3372–80.

Stevanato R, Fabris S, Momo F. New enzymatic method for the determination of total phenolic content in tea and wine. J Agric Food Chem. 2004;52(20):6287–93.

Sun B, Ribes AM, Leandro MC, Belchior AP, Spranger MI. Stilbenes: quantitative extraction from grape skins, contribution of grape solids to wine and variation during wine maturation. Anal Chim Acta. 2006;563(1–2):382–90.

Trousdale Vernon L, Singleton EK. Astilbin engeletin in grapes and wine. Phytochem. 1983;22(2):619–20.

Tsao R. Chemistry and biochemistry of dietary polyphenols. Nutrients. 2010;2(12):1231–46.

Valls J, Millán S, Martí MP, Borràs E, Arola L. Advanced separation methods of food anthocyanins, isoflavones and flavanols. J Chromatogr A. 2009;1216(43):7143–72.

Vennat B, Pourrat A, Pourrat H, Gross D, Bastide P, Bastide J. Procyanidins from the roots of Fragaria vesca: characterization and pharmacological approach. Chem Pharm Bull. 1988;36(2):828–33.

Wada M, Kido H, Ohyama K, Ichibangase T, Kishikawa N, Ohba Y, *et al.* Chemiluminescent screening of quenching effects of natural colorants against reactive oxygen species: Evaluation of grape seed, monascus, gardenia and red radish extracts as multi-functional food additives. Food Chem. 2007;101(3):980–6.

Wang LJ, Sun DW. Recent developments in numerical modeling of heating and cooling processes in the food industry: a review. Trends Food Sci Technol. 2003;14:408-23.

Williamson EM. Selected bibliography. Phytotherapy Research. 1990;4(4):5-10.

Woodman OL, Meeker WF, Boujaoude M. Vasorelaxant and antioxidant activity of flavonols and flavones: structure-activity relationships. J Cardiovasc Pharmacol. 2005;46(3):302-9.

Wrolstad RE. Color and Pigment Analyses in Fruit Products: Oregon State University; 1976.

www.cen-online.org. June 9 (2008) 32.

www.cen-online.org. January 21 (2008) 33.

www.cen-online.org. March 24 (2008) 28.

www.cen-online.org. December 15 (2008) 18.

Xu C, Zhang Y, Cao L, Lu J. Phenolic compounds and antioxidant properties of different grape cultivars grown in China. Food Chem. 2010;119(4):1557-65.

Zamora-Ros R, Andrés-Lacueva C, Lamuela-Raventós RM, Berenguer T, Jakszyn P, Martínez C, *et al.* Concentrations of resveratrol and derivatives in foods and estimation of dietary intake in a Spanish population: European Prospective Investigation into Cancer and Nutrition (EPIC)-Spain cohort. Br J Nutr. 2008;100(1):188-96.

Zhao J, Wang J, Chen Y, Agarwal R. Anti-tumor-promoting activity of a polyphenolic fraction isolated from grape seeds in the mouse skin two-stage initiation-promotion protocol and identification of procyanidin B5-3′-gallate as the most effective antioxidant constituent. Carcinogenesis. 1999;20(9):1737-45.

Capítulo III:

Comparison of extraction methods for exploitation of grape skin residues from ethanol distillation

Talanta

Volume 101, 15 November 2012,

Pages 292-298

Comparison of extraction methods for exploitation of grape skin residues from ethanol distillation

Ángela Peralbo Molina, Feliciano Priego-Capote, María Dolores Luque de Castro*

[1]*Department of Analytical Chemistry, Annex C-3, Campus of Rabanales, 14071 Córdoba, Spain*

[2]*University of Córdoba Agroalimentary Excellence Campus, ceiA3, Campus of Rabanales, 14071 Córdoba, Spain*

[3]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital,*

University of Córdoba, E-14014, Córdoba, Spain

All the authors belong to the 3 institutions.

Comparison of extraction methods for exploitation of grape skin residues from ethanol distillation

Ángela Peralbo Molina, Feliciano Priego-Capote, María Dolores Luque de Castro*

Abstract

Four extraction techniques –namely, conventional maceration, ultrasound-assisted extraction, microwave-assisted extraction, and superheated liquid extraction (SHLE)– have been compared to evaluate their suitability to obtain valuable compounds from a raw material traditionally of scant interest: grape skin residues from ethanol-distillation. With this aim, red and white grape skins were separated from the rest of the pomace residue and subjected to extraction with 1:1 ethanol-acidic water as extractant in order to obtain the largest possible number of valuable compounds from this material, which has so far been used only as a heat source. The resulting extracts were characterized by the Folin-Ciocalteu and Ferric Reducing Antioxidant Power tests and by liquid chromatography-time-of-flight/mass spectrometry (LC-TOF/MS). The composition of the extracts under each of the optimal conditions was studied by LC-TOF/MS, and the information thus obtained compared by Venn diagrams. These diagrams, together with the extracted base peak chromatograms, were used to assess the optimal working conditions. Tentative identification of compounds was conducted using open-free databases. Grape skins from distillation industries are a source of compounds of interest for the food and nutraceutical fields.

Keywords: grape skin; superheated-liquid extraction; ultrasound-assisted extraction; microwave-assisted extraction; LC-TOF/MS analysis, statistical data-treatment

1. Introduction

Grape production worldwide is about 70 million tonnes, 80% of which is dedicated to winemaking and, practically the rest amount to the grape juice industry [1]. There is an unequal use of by-products obtained from vine cultivars or from the winemaking process such as vine-shoots, vinification lees or grape skins that, in any case, allow obtaining high-added value products [2]. Among these by-products, grape skins should be emphasized due to their significant content in monomeric and polymeric molecules such as anthocyanins, flavan-3-ols, flavonols, dihydroflavonols, hydroxycinnamoyl tartaric acids, hydroxybenzoic acids and hydroxystilbenes [3]. Attending to this composition, grape skins are considered a potential source for isolation of natural compounds [4].

The composition of grape skins is characterized by the vine cultivar, with special significance of the grape color. Overall, red grapes possess higher crude protein, fat and ash contents than white grapes [5,6]. They also have higher total extractable pectins, dietary fiber, neutral sugars, condensed tannins and resistant proteins. White grape skins have significantly higher soluble sugars, uronic acid and Klason lignin than red skins [7].

From a nutraceutical point of view, an interesting fraction of compounds in grapes is that formed by phenolic compounds, which are endowed in general terms with a high antioxidant capacity and are at higher concentrations in skin and seeds than in grape pulp [2,8]. Usually, the amount of total phenols in white grape varieties is lower than in red grapes, as the former ones do not synthesize anthocyanins [5]. These are phenolic compounds with flavonoid structure responsible for color of the red grapes [9,10], while the low-molecular weight flavan-3-ols (such as catechins) and procyanidin oligomers are responsible for bitterness [4]. The nutraceutical interest of phenolic compounds is out of doubt. Phenols are considered key compounds contributing to the antioxidant potential of grapes, wine and other fruits, which is responsible of the health benefits attributed to this fraction [11–14].

After the winemaking process, low-quality wines, grape pomace (skins and seeds) and vinification lees are mainly destined to distilleries to produce alcohol. In the case of pomace, this residue is subjected to an extraction process (cleaning) with water, yielding clean pomace and pomace extract, called pickaxe [15]. Currently, this extraction is performed by diffusion in industrial extractors type "broadcaster band" with a processing capacity of around 300.000 kg/day. After extraction, the clean pomace, completely depleted in sugar, is subjected to fermentation as a preparation step for alcohol distillation. Following alcohol isolation, the resulting residue is wasted although valorization studies are demanded. One example is the de-tartration process to isolate tartaric salts present in grape skins. The exploitation strategies have not considered the isolation of phenolic compounds, which could be a challenge taking into account the different exhaustive steps applied to the pomace residue. Phenolic extracts from this residue could increase its value.

The extraction of phenols remains as a challenge due to the variety of compounds and the chemical complexity of the sample. This step is focused on the removal of phenolic compounds from the vacuolar structures where they are found, either through rupturing plant tissue or through a diffusion process [16].

Maceration is the most commonly used procedure for leaching phenolic compounds from grapes and residues from them, where the solid sample is stirred with a suitable liquid for long times [17]. However, the different treatments to which grape pomace is subjected suggest the possibility of using auxiliary energies to enhance the leaching efficiency of strongly retained compounds. In this sense, the leaching step can be assisted by auxiliary energies such as microwaves [18,19] or ultrasound [20]. One other alternative is to use supercritical fluids, which has been reported for the leaching of relevant phenolic compounds from grape skin [21]; however, the hight cost of the common supercritical extractant –CO_2– and its non polar character make its application non-attractive for phenols extraction. In this context, the use of superheated solvents can be an interesting alternative for extracting phenolic compounds from

grapes and their residues. The dielectric constant of the extractant decreases by increasing the temperature above its boiling point and polar and relatively non-polar compounds can be extracted by ambient-temperature polar extractants [22].

The main aim of the present research was to compare the present and traditional extraction techniques to propose an optimum method for isolation of priced compounds from a raw material with very scant value: waste from distillery industries using the cake obtained after alcohol distillation from red and white grape pomace. Three present extraction techniques (*viz*. microwave-assisted extraction, MAE; ultrasound-assisted extraction, USAE; and superheated-liquid extraction, SHLE) have been selected and used in comparison with conventional maceration extraction, CME, using overall quantitation methods and individual characterization by LC-TOF/MS in order to know both the extraction efficiency of each at a whole, and the compounds extracted by the different working conditions provided by each extraction technique. A final aim was to demonstrate that not only the waste from the winemaking industry can be exploited to obtain valuable compounds, but that these compounds –and even others resulting from the drastic conditions to which grape pomace is subjected for ethanol distillation– can be obtained from the "waste from the waste" of this industry.

2. Experimental

2.1. Samples

Grape pomace generated in the winemaking process from red and white grapes was collected in an industrial plant ("Alcoholeras Reunidas, S.A.", Ciudad Real, Spain) after alcohol distillation. Skins from red and white grapes were separated from the corresponding pomace. After separation, both types of skins were dried for 72 h at 35 °C, then milled and sieved with a 40-mm mesh-particle size (less than 0.42 mm d). The unknown and heterogeneous origin of the massive raw materials subjected to distillation prevents from knowing the

degradation caused by the drastic conditions of ethanol-distillation through comparison of the materials before and after this step.

2.2. Reagents

Ethanol (96% v/v) PA from Panreac (Barcelona, Spain) and distilled water were used to prepare the different ethanol-water mixtures. LC-MS grade formic acid and acetonitrile (ACN) were purchased from Scharlab (Barcelona, Spain). Deionized water (18 MΩ·cm) was obtained from a Millipore (Bedfore, MA, USA) Milli-Q plus system, and n-hexane, for liquid-liquid extractions, was from LiChrosolv (Merck, Darmstadt, Germany). TPTZ (2,4,6-tris(2-pyridyl)-s-triazine) and Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid) were from Fluka (Buches, Switzerland). The Folin-Ciocalteu (F-C) reagent, sodium carbonate, gallic acid and AAPH (2,2'-azobis-2-methyl-propanimidamide dihydrochloride) were from Sigma (St. Louis, USA).

2.3. Apparatus and instruments

A grinder (Ball Mill Restch MM301) was used to mill the skins. Superheated liquid extractions were performed by a laboratory-made dynamic extractor [23], consisting of the following units: a) an extractant supply; b) a high-pressure pump (Shimadzu LD-AC10), which propels the extractant through the system; c) a switching valve placed next to the pump to develop static extractions; d) a stainless-steel cylindrical extraction chamber (550 mm x 10 mm i.d., 4.3 mL internal volume) where the sample is placed. This chamber was closed at both ends with screws whose caps contain cotton made filters to ensure the sample is not carried away by the extractant; e) a restriction valve to maintain the desired pressure in the system; f) a cooler made of a stainless steel tube (1 m length, 0.4 mm i.d.) and refrigerated with water; g) a gas chromatograph oven (Konix, Cromatix KNK-2000) where the extraction chamber was placed and heated. Both chamber and tubing system were of high quality stainless-steel to avoid or minimize corrosion under drastic working conditions.

Microwave-assisted extractions were carried out in a Microdigest 301 digestor of 200 W maximum power by Prolabo (Paris, France), furnished with a microprocessor programmer to control the microwave unit and used to accelerate solid–liquid extraction, and provided.

Ultrasound-assisted extractions were performed with a Branson 450 digital sonifier (20 kHz, 450W) equipped with a cylindrical titanium-alloy probe (12.70 mm diameter), which was immersed into a lab-made stainless-steel container with eight compartments to place test tubes, which were thus subjected to similar ultrasound intensity [20].

The absorbance of the extracts after reaction with the F–C reagent and the FRAP (Ferric Reducing Antioxidant Power) assay was monitored by a Thermo Spectronic Helios Gamma spectrometer (Waltham, MA, USA). Shaking and centrifugation of the extracts were carried out by means of an MS2 Minishaker Vortex (IKA, Germany) and a Mixtasel centrifuge (Selecta, Barcelona, Spain), respectively.

All samples were analyzed by a 1200 Series LC system (Agilent Technologies, Waldbronn, Germany) coupled to an Agilent 6530 TOF mass spectrometer with an electrospray ionization (ESI) source.

2.4. Superheated liquid extraction

One gram of milled grape skin was placed into the extraction cell installed into the gas chromatograph oven; then, a relative high flow rate (7 mL/min) was used for 1 min to fill the cell rapidly. To ensure the absence of air inside the extraction cell, the restrictor valve was kept open until the first drop of extractant appeared. At that moment, the restrictor valve was closed and when the desired pressure was reached, the switching valve was closed, the pump was turned off and the oven was switched on. During temperature rising, the switching valve had to be opened at short intervals to prevent the pressure from surpassing the working value. Once the selected temperature and pressure were reached, static

extraction was performed for a preset time. Finally, the oven was switched off, the chamber was cooled below the boiling point of ethanol and then, the switching valve and the restrictor valve were switched to enable new extractant to flow through the cell and flush out the extract.

2.5. Extraction protocols with auxiliary energies

Two grams grape skins were placed into the extraction vessel with 20 mL 50% (v/v) aqueous ethanol acidified with 0.8% (v/v) HCl. The suspension was subjected to microwave assistance for 10 min with a power irradiation of 140 W. After extraction, the liquid phase was isolated by centrifugation and stored at -20 ºC until analysis.

The same amount of material and extractant were located in a glass vessel that was introduced in a water bath at 25 ºC located in a sound-proof chamber. The ultrasonic probe was immersed in the extraction vessel through the upper part of the chamber. Ultrasound was applied for 10 min with a duty cycle of 0.5 s/s and 60% of the nominal power of the converter.

2.6. Conventional maceration extraction (CME)

Two grams of grape skins were extracted with 20 mL 50% (v/v) acidified with 0.8% (v/v) HCl aqueous ethanol by stirring at 40 °C for 24 h.

2.7. Determination of total phenols by the F-C method

The total amount of phenolic compounds was quantified by the F-C method using gallic acid as standard. With this purpose, a calibration curve was run using solutions of 1, 10, 25, 50, 75 and 100 mg/L of this acid (Absorbance = 0.0065 Concentration + 0.1286, R^2 = 0.9909). A 0.5-mL aliquot of extract, 10 mL of distilled water, 1 mL of F-C reagent and 3 mL of Na_2CO_3 (20%, w/v) were mixed, made to 25 mL with distilled water and heated at 50 °C for 5 min. After heating, the samples were kept at room temperature for 30 min and, finally, the absorbance was measured at 765 nm against a blank solution containing distilled

water instead of extract. The concentration of phenolic compounds thus obtained was multiplied by the dilution factor of the extract volume and divided by the amount of grape skins used. The results were expressed as equivalent to milligrams of gallic acid (mg GAE) per mL of grape skin extract (mg GAE/mL of extract) [24].

2.8. FRAP assay

The antioxidant (AOP) potential of grape skin extracts was determined using the FRAP assay, based on reduction of the ferric tripyridyltriazin [Fe(III)-TPTZ] complex to ferrous tri-pyridyl triazin [Fe(II)-TPTZ] at low pH by the action of antioxidants in the target solution. The ferrous complex [Fe(II)-TPTZ] has an intense blue color that can be monitored at 593 nm. The assay response was standardized against the antioxidant standard Trolox.

2.9. LC-TOF/MS analysis

An injection volume of 20 µL and a flow rate of 1 mL/min was used. A mobile phase A consisting of 0.2% (v/v) formic acid aqueous solution and a mobile phase B consisting of acetonitrile constituted the chromatographic phases to establish the following gradient method: from 96% to 50% A in 60 min, from 50% to 0% A in 5 min. MassHunter Workstation Data acquisition software (Agilent Technologies) was used to control the instrument. Data were processed using MassHunter Qualitative Analysis software (Agilent Technologies). Extraction of unknown molecular features from raw data was carried out by the Molecular Feature Extraction (MFE) algorithm in MassHunter Qualitative analysis software –taking as molecular feature any molecule, ion, etc. structural and isotopically identifiable as a separately distinct entity. The feature extraction algorithm took into account all ions exceeding 1000 counts and with a charge state equal to, or above, one and a feature had to be composed of two or more ions to be valid (e.g. two ions in the isotope cluster). The theoretical formula adjusted to the corresponding isotopic distribution of molecular features was generated with the Molecular Formula Generation software (Agilent

Technologies). Using background subtracted data, files in compound exchange format (.cef files) were created for each sample and exported into the Mass Profiler Professional (MPP) software package (version 2.0, Agilent Technologies, Santa Clara, CA, USA) for further processing. In the next step, alignment of RT and *m/z* values was carried out across the sample set using a tolerance window of 0.2 min and 5 ppm, respectively. Analyses were processed using MassHunter Qualitative software and tentative identification of compounds was performed using the METLIN Personal Metabolite Database and PlantCyc [25].

3. Results and discussion

3.1. Comparison of extraction techniques

Four extraction techniques were selected to compare their efficiency for isolation of compounds from grape skins as sample preparation strategy. These were CME as reference, USAE and MAE to check the influence of auxiliary energies for extraction assistance, and SHLE to take benefits from the superheated state of a suited extractant. Proper comparison required to use the same amount of material, volume and composition of the extractant. This was set at 50:50 (v/v) ethanol–water with 0.8% (v/v) HCl, according to preliminary experiments [23]. A hydroalcoholic extractant was selected to favor isolation of polar and mid-polar compounds. Ethanol was used as organic solvent avoiding other toxic solvents such as methanol, acetonitrile, chloroform or hexane, because of the potential use of the extracts for human consumption. Apart from that, an acid pH was required to enhance hydrolysis of polymeric structures and release monomeric metabolites, easily solubilized in this way. Other extraction conditions used were based on previous MAE [18,19,26], USAE [20] and SHLE [23] methods found in the literature for isolation of natural products from vegetal materials. Thus, in the case of MAE and USAE, the power was set at intermediate values, while the extraction time was planned for 10 min in both cases (0.5 s/s duty cycle for USAE). Concerning SHLE, the temperature was set at 180 ºC, and 60 min was adopted as extraction time.

Since the same ratio between sample weight and extractant volume was used in the four isolation protocols, the extracts obtained were analyzed by LC-TOF/MS in accurate mode with the same dilution pattern. Supplementary Figure 1 illustrates the TOF chromatograms (Base Peak Chromatograms, BPC) for extracts from red and white grape skins. As can be seen, the extracts were predominantly rich in polar compounds, mainly eluted within 28 min of the chromatographic gradient described under "LC-TOF/MS analysis". This elution time corresponds to a chromatographic mobile phase with 25% ACN. Visually, in the case of white grape skins extract, higher peaks were obtained with SHLE in the first five min of elution. Within this chromatographic time, the highest signals corresponded to the extracts obtained by CME. The opposite situation was found in the analysis of extracts from red grape skins. Thus, CME reported higher chromatographic peaks during the first 5 min, but the peaks were, in general, higher for SHLE in the rest of the chromatogram.

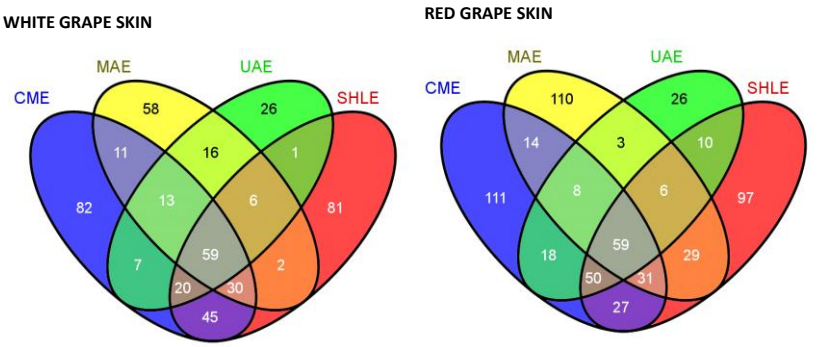


Figure 1: Venn diagrams obtained after processing the molecular features extracted in the analysis of extracts from white and red grape skins showing the number of potential compounds obtained for each extraction method used and those common to all extracts.

The first test for comparison of the efficiency of the extraction alternatives was to compare the molecular features obtained as described under "LC-TOF/MS analysis". For this purpose, data sets obtained by analysis of blanks (extraction solvents) were subtracted from the data set corresponding to the analytical samples in the step for extraction of molecular entities in each raw data file. Venn

diagrams are shown in Figure 1 for each type of grape skin. As can be seen, similar behavior showed the results obtained for white and red grape skins. In both cases, CME and SHLE reported the highest number of detected molecular features (267 and 318 in CME for white and red grape skins, respectively, versus 244 and 309 in SHLE for white and red grape skins, respectively). By contrast, MAE reported 195 and 260 molecular features for white and red grape skins, respectively, while USAE reported the lowest number of molecular features in both types of samples. It is important to mention that only 59 potential molecular features were constant for the four extraction methods with the maximum similarity for CME and SHLE with 154 and 167 common features for white and red grape skins, respectively. Both SHLE and CME were also characterized by the low number of features common to USAE and MAE. Maceration is the extraction alternative that operates under the softest conditions in terms of temperature or assisted energy. Attending to the profiles of potential molecular features obtained by the different techniques, MAE and USAE seem to alter the composition of the extracts by comparison to CME. On the other hand, SHLE provided an extract with coverage, in terms of molecular entities, quite similar to that obtained by CME. With these premises, MAE and USAE were discarded while SHLE was the preferred technique to optimize exhaustively the extraction conditions.

3.2. Influence of extractant pH on SHLE

The influence of extractant pH on SHLE was studied with the purpose of approaching to the efficiency attained with CME taking into account that less acidic pHs have also been used in SHLE methods for isolation of natural extracts [27]. Figure 2 plots the Venn diagrams for each type of grape skin used in this study. As can be seen, an improved effect was found by SHLE at pH 3 as compared with lower pH tests. The efficiency of SHLE was considerably enhanced at pH 3, providing more molecular features than maceration. This result could be justified by hydrolysis and degradation of labile compounds at extreme acid pHs.

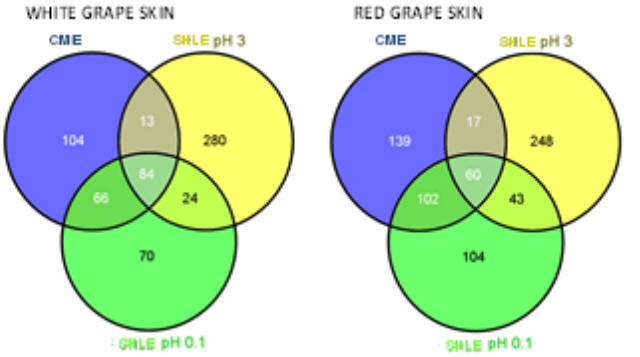


Figure 2. Venn diagrams obtained after processing the molecular features extracted in the analysis of extracts from white and red grape skins showing the number of potential compounds obtained for SHLE working at two pHs (0.1 and 3.0) and CME, showing common and uncommon compounds.

3.3. Influence of critical SHLE variables

The two most critical variables influencing SHLE (namely: the extraction temperature and processing time) were studied for isolation of high-added value compounds from grape skins. The results obtained after univariate study of both variables are discussed below.

Temperature. This variable was tested from 140 to 200 °C in the case of red-grape skins, while it was forced up to 220 °C for white-grape skins according to the chromatographic profiles obtained. Figure 3 shows the Venn diagrams representing the influence of the extraction temperature on the metabolites coverage for the two types of grape skins.

The highest number of molecular features (441 and 398 for white- and red- grape skins, respectively) was obtained at the maximum temperatures tested in each case, 220 and 200 °C for white and red grape skins, respectively. However, both temperature tests were characterized by a high proportion of molecular features exclusively detected in the extracts from these experiments, 35 and 46% for white and red grape skins, which could be indicative of

degradation of the target compounds in the extracts as compared to the other tested temperatures. Attending to the balance between number of molecular features and proportion of molecular features detected exclusively in one experiment, the optimum temperatures are 180 and 160 °C for extraction of white and red grape skins, respectively. Thus, 401 molecular features were detected in the extract of white grape skins at 180 °C; 25% of them were detected only in this extract. Similarly, 387 molecular features were detected in the extract from red grape skins at 160 °C, with a 27% detected only in it. Supplementary Figure 2 illustrates the BPC obtained for each type of grape skins at the tested temperatures, showing the differences between the chromatographic profiles of the extracts.

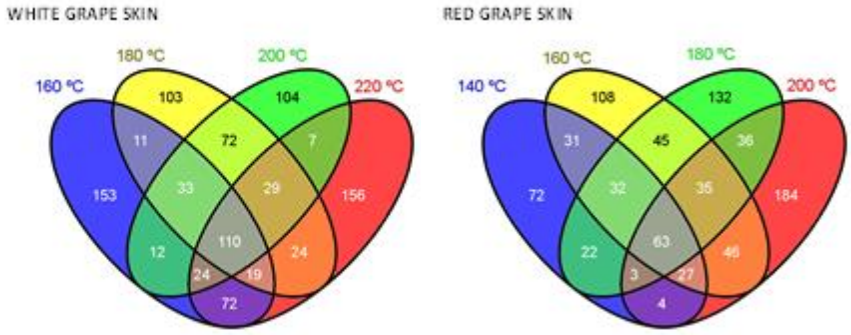


Figure 3. Venn diagrams obtained after processing the molecular features from the analysis of extracts from white and red grape skins showing both the number of potential compounds obtained at each extraction temperature used, and the common and uncommon compounds.

Time. A kinetics study was developed to set the optimum extraction time; thus, extraction times of 20, 40, 60 and 80 min were tested. The results, as Venn diagrams, are shown in Figure 4. These diagrams demonstrated that an extraction time of 60 min provided the highest number of molecular features for both matrices (401 and 364 for white- and red-grape skins, respectively). In the case of extracts from white grape skins, no statistical differences were observed, at 95% confidence level, in the number of molecular features between 40 and 60

min. However, there were significant differences for the extracts from red grape skins and, for this reason, the selected extraction time was 60 min.

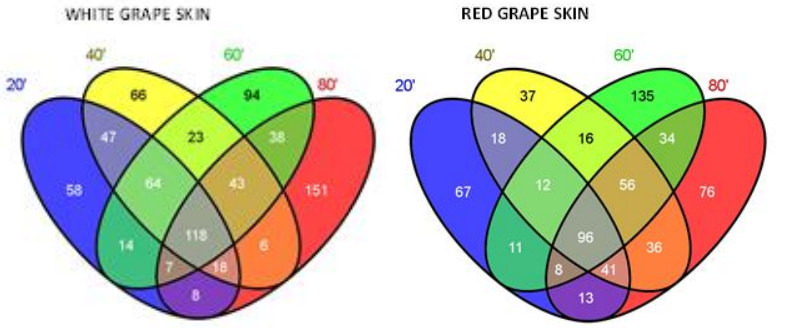


Figure 4. Venn diagrams obtained after processing the molecular features from the analysis of extracts from white and red grape skins showing both the number of potential compounds obtained for each extraction time and those common and uncommon to them.

Supplementary Figure 3 plots BPCs corresponding to the analysis of extracts obtained at different extraction times from the two types of samples.

3.4. Total phenolic content and evaluation of the antioxidant potential by the FRAP assay

Two different tests were applied to evaluate the potential of the two types of grape skins used in this research for isolation of interesting compounds such as phenolic compounds or, more generally, antioxidants. The F-C test revealed that the extracts from white-grape skins reported higher concentrations of total phenols (equivalent to 4.2 mg GAE/mL of extract) than those from red grape skins, with 1.5 mg GAE/mL. Concerning the antioxidant activity, the same trend as in the total phenols content was observed. Thus, the skin extracts from white grapes reported the highest antioxidant activity with 5.8 mg equivalents of Trolox/mL of extract, versus skin extracts from red grapes that gave 2.3 mg equivalents of Trolox/mL.

3.5. Identification of interesting compounds in skins from white- and red grapes

Once the extraction conditions were defined for isolation of the polar/mid-polar fraction from the two types of skins from residues obtained after grape pressing and alcohol distillation, characterization of the resulting extracts was demanded for valorization of the residues. Numerous compounds were tentatively identified in skins from white and red grapes. Identification was supported on mass accurate detection with 10 ppm maximum error in monoisotopic mass taking into account the isotopic distribution. Metabolites were identified by searching on PlantCyc database (www.plantcyc.org) and on a personnel database prepared after review of existing research in the literature about characterization of skin from grapes [27,28]. However, this section will be focused on two main families that should be emphasized because of their nutraceutical and oenological interest: phenols and furfural derivatives (the latter produced by sugars degradation). In all cases more polar than non-polar compounds were extracted from grape skins, which is a consequence of both the polar nature of the extractant used and the high content of polar compounds in the raw material.

The variety of phenolic compounds extracted under superheated-extractant conditions encompassed flavonoid and non-flavonoid compounds. The former can be classified as follows: (i) anthocyanins, only present in red grapes, were not detected in any extracts, except for the pyroanthocyanin vitisin B (malvidin-3-glucoside acetaldehyde); (ii) flavanols, that included catechin, epicatechin and gallate derivatives, were detected, but also a procyanidin B isomer was tentatively identified; (iii) flavonols, which were detected as free and 3-glucoside derivatives, were present at higher number in extracts from red-grape residues. Concerning non-flavonoid phenols, a great variety of phenolic acids was detected in the extracts from grape-skin residues. Thus, compounds such as caffeic acid, gallic acid or protocatechuic acid were detected in both types of extracts. An interesting fraction because of their biological activity was that

formed by stilbenes, indistinctly detected in the residue from both grape varieties. Other phenolic compounds with biological interest such as ethyl esters (ethyl ferulate and ethyl protocatechuate), lignin monomers and derivatives (pyrocatechol, pyrogallol and syringol), aldehydes (coniferaldehyde) and ketones (vanillin), among others, were also tentatively identified in the characterization analysis.

Concerning furanic derivatives, the removal of sugars in the previous industrial steps justified the scant number of furan derivatives identified (among them furfuryl alcohol). Despite the degradation caused by the drastic conditions of the distillation process, many interesting compounds are found in the extract from the waste of this process, which makes it a useful matter for more than as a heating source.

Table 1. Identified compounds in the extracts from grape skins.

FLAVONOIDS	THEORICAL *m/z*	EXPERIMENTAL *m/z*	FORMULA	GRAPE SKIN	
				RED	WHITE
ANTHOCYANINS					
MALVIDIN-3-GLUCOSIDE ACETALDEHYDE (VITISIN B)	528.1035	528.1028	$C_{23}H_{25}ClO_{12}$	v	χ
FLAVANOLS					
EPICATECHIN	290.079	290.0807	$C_{15}H_{14}O_6$	χ	v
CATECHIN	290.079	290.0807	$C_{15}H_{14}O_6$	v	v
EPICATECHIN-O-GALLATE	442.09	442.0902	$C_{22}H_{18}O_{10}$	v	χ
EPIGALLOCATECHIN	306.074	306.0725	$C_{15}H_{14}O_7$	v	v
GALLOCATECHIN-O-GALLATE	458.0849	458.0867	$C_{22}H_{18}O_{11}$	v	χ
PROCYANIDIN B2	578.1424	578.1436	$C_{30}H_{26}O_{12}$	v	χ
FLAVONOLS					
KAEMPFEROL	286.0477	286.0489	$C_{15}H_{10}O_6$	v	v
MYRICETIN	318.0376	318.0395	$C_{15}H_{10}O_8$	v	χ
QUERCETIN	302.0427	302.0427	$C_{15}H_{10}O_7$	v	v
ISORHAMNETIN	316.0583	316.0613	$C_{16}H_{12}O_7$	χ	v
MYRICETIN-3-O-GLUCOSIDE	480.0904	480.0925	$C_{21}H_{20}O_3$	v	χ
MYRICETIN-7-O-GLUCOSIDE	480.0904	480.0902	$C_{21}H_{20}O_3$	v	v
QUERCETIN-3-O-GLUCOSIDE	464.0955	464.0964	$C_{21}H_{20}O_{12}$	v	χ

NON FLAVONOIDS	THEORICAL *m/z*	EXPERIMENTAL *m/z*	FORMULA	GRAPE SKIN	
				RED	WHITE
PHENOLIC ACIDS					
CINNAMIC ACID	148.0524	148.0523	$C_9H_8O_2$	χ	v
CAFFEIC ACID	180.0423	180.0423	$C_9H_8O_4$	v	v
o-COUMARIC ACID	164.0473	164.0476	$C_9H_8O_3$	v	v
FERULIC ACID	194.0579	194.0574	$C_{10}H_{10}O_4$	χ	v
GALLIC ACID	170.0215	170.022	$C_7H_6O_5$	v	v
4-HYDROXYBENZOIC ACID	138.0317	138.0317	$C_7H_6O_3$	v	v
PROTOCATECHUIC ACID	154.0266	154.0268	$C_7H_6O_4$	v	v
SYRINGIC ACID	198.0528	198.0527	$C_9H_{10}O_5$	v	χ
STILBENES					
TRANS -RESVERATROL	228.0786	228.0786	$C_{14}H_{12}O_3$	v	v
α-VINIFERIN	678.189	678.188	$C_{42}H_{30}O_9$	v	v
ASTRINGININ	244.0736	244.0736	$C_{14}H_{12}O_4$	v	v
TRANS -ε-VINIFERIN	454.1416	454.1426	$C_{28}H_{22}O_6$	v	v
PICEID	390.1315	390.1331	$C_{20}H_{22}O_8$	v	v
TRANS -PTEROSTILBENE	256.1099	256.1097	$C_{16}H_{16}O_3$	χ	χ

OTHER PHENOLS	THEORICAL*m/z*	EXPERIMENTAL*m/z*	FORMULA	GRAPE SKIN	
				RED	WHITE
PHENOLIC ESTERS					
ETHYL FERULATE	222.0892	222.0902	$C_{12}H_{14}ClO_4$	v	v
ETHYL PROTOCATECHUATE	182.0579	182.0583	$C_9H_{10}O_4$	v	v
LIGNIN MONOMERS AND DERIVATIVES					
PYROCATECHOL	110.0368	110.0368	$C_6H_6O_2$	v	v
PYROGALLOL	126.0317	126.0311	$C_6H_6O_3$	v	v
SYRINGOL	154.063	154.225	$C_8H_{10}O_3$	v	χ
4-METHYLPYROCATECHOL	124.0524	124.0519	$C_7H_8O_2$	v	v
4-VINYLGUAIACOL	150.0681	150.0682	$C_9H_{10}O_2$	v	v
PHENOLIC ALDEHYDES					
2-PHENYLACETALDEHYDE	120.0575	120.0578	C_8H_8O	v	v
4-HYDROXYBENZALDEHYDE	122.0368	122.0365	$C_7H_6O_2$	v	v
CONIFERALDEHYDE	178.063	178.0627	$C_{10}H_{10}O_3$	v	v
PHENOLIC KETONES					
p-HYDROXYBENZALACETONE	162.0681	162.0683	$C_{10}H_{10}O_2$	v	v
VANILLIN	152.0473	152.0472	$C_8H_8O_3$	v	v
1-3-4-DIHIDROXYPHENYL -1-PROPANONE	166.063	166.065	$C_9H_{10}O_3$	χ	v
METHOXY AND ETHOXY PHENOLS					
2,6-DIMETHOXY -4-1-PROPENYL -PHENOL	194.0943	194.0941	$C_{11}H_{14}O_3$	χ	v
4-ETHOXYPHENOL	138.0681	138.0878	$C_8H_{10}O_2$	χ	χ
FURANIC					
5-ACETOXYMETHYLFURFURAL	168.0423	168.0423	$C_8H_8O_4$	v	v
2-FURANCARBOXYLIC ACID, ETHYL ESTHER	140.0473	140.0466	$C_7H_8O_3$	v	χ

Conclusions

The research, the results of which have been exposed and discussed, was aimed at demonstrating that: (i) very different profiles of compounds can be obtained from the extracts from the same raw material depending on the type of energy used to accelerated/improve the extraction step. Therefore, the different options of auxiliary energy should be tested in the light of the target compounds we are looking for before selecting one of them. (ii) Not only the waste from the winemaking industry can be exploited to obtain valuable compounds, but the same compounds –and even others resulting from the drastic conditions to which grape pomace is subjected for ethanol distillation– can be obtained from the "waste from the waste" of this industry.

The results obtained by subjecting the raw material from waste of red and white grapes to four extraction methods allow obtaining the following conclusions:

- In all cases, more polar than non-polar compounds were extracted, which is a consequence of both the polar extractant used and the high content of the former in the raw materials.

- The two extraction techniques based on the use of auxiliary energy (MAE, UAE) yielded lower efficiencies than CME and SHLE.

- The highest number of compounds was detected in the extracts obtained by SHLE and CME.

- The increase of pH also increased the efficiency of SHLE as compared with CME.

-The yield of extraction using ethanol-water mixtures under superheated conditions clearly surpassed that of the conventional method, which used the extractant at ambient temperature.

-The superheated conditions dramatically shortened the extraction time (60 min versus 24 h for the conventional method).

The final conclusion from this study is that despite the degradation caused by the drastic conditions of the distillation process, many interesting, valuable compounds are found in the extract from the waste of this process, which make it a useful matter for a better exploitation than as a heating source.

- The variety of tentatively identified compounds in the extracts makes them exploitable as additives in the food industry (either as colorants, as flavor modifiers or as antioxidants), and also in the cosmetics and nutraceuticals industries.

Acknowledgements

Spanish Ministerio de Ciencia e Innovación (MICINN) through Project CTQ2009-07430 and Complementary Action CTQ2009-08064-E. F.P.-C. is also sponsored by MICINN through a Ramón y Cajal Contract (RYC-2009-03921).

References

[1] H. Wang, E.J. Race, A.J. Shrikhande, J. Agric. Food Chem., 51 (2003) 1839-1844.

[2] K. Ali, F. Maltese, Y.H. Choi, R. Verpoorte, Phytochem. Rev., 9 (2010) 357-378.

[3] M. Palma, L.T. Taylor, J. Chromatogr. A, 849 (1999) 117-124.

[4] J.-M. Souquet, V. Cheynier, F. Brossaud, M. Moutounet, Phytochemistry, 43 (1996) 509-512.

[5] M. Serafini, G. Maiani, A. Ferro-Luzzi, J. Nutr., 128 (1998) 1003-1007.

[6] J.A. Vinson, B.A. Hontz, J. Agric. Food Chem., 43 (1995) 401-403.

[7] F. Saura-Calixto, I. Goñi, E. Mañas, R. Abia, Food Chem., 39 (1991) 299-309.

[8] E.Q. Xia, G.F. Deng, Y.J. Guo, H.B. Li, Int. J. Mol. Sci., 11 (2010) 622-646.

[9] S. Vidal, Y. Hayasaka, E. Meudec, V. Cheynier, G. Skouroumounis, J. Agric. Food Chem., 52 (2004) 713-719.

[10] S. Vidal, E. Meudec, V. Cheynier, G. Skouroumounis, Y. Hayasaka, J. Agric. Food Chem., 52 (2004) 7144-7151.

[11] S.B. Lotito, C.G. Fraga, Free Radic. Bio. Med., 24 (1998) 435-441.

[12] H. Moini, Q. Guo, L. Packe, Adv Exp Med Biol, 505 (2002) 141-149.

[13] B.C. Scott, J. Butler, B. Halliwell, O.I. Aruoma, Free Radic. Res. Commun., 19 (1993) 241-253.

[14] W. Bors, C. Michel, Free Radic. Biol. Med., 27 (1999) 1413-1426.

[15] Van der Merwe, A. Thesis doctoral. Department of Logistics, University of Stellenbosch, South Africa (2009).

[16] S.L. Nixdorf, I. Hermosín-Gutiérrez, Anal. Chim. Acta, 659 (2010) 208-215.

[17] S. Quideau, Angew Chem. Int. Edit, 43 (2004) 393-394.

[18] H. Li, B. Chen, L. Nie, S. Yao, Phytochem. Anal., 15 (2004) 306-312.

[19] X. Pan, G. Niu, H. Liu, Chem. Eng. Process, 42 (2003) 129-133.

[20] O.F. Ricárdez, J. Ruiz-Jiménez, L. Lagúnez-Rivera, M.D. de Castro, Phytochem. Anal., 22 (2011) 484-491.

[21] A. Chafer, M.C. Pascual-Marti, A. Salvador, A. Berna, J. Sep. Sci., 28 (2005) 2050-2056.

[22] M.D. Luque de Castro, J.L. Luque-García, Anal. Biol. Chem., 377 (2003) 1089-1090.

[23] J.M. Luque-Rodríguez, M.D. Luque de Castro, P. Perez-Juan, Bioresour. Technol., 98 (2007) 2705-2713.

[24] J.M. Luque-Rodríguez, P. Pérez-Juan, M.D. Luque de Castro, J. Agric. Food Chem., 54 (2006) 8775-8781.

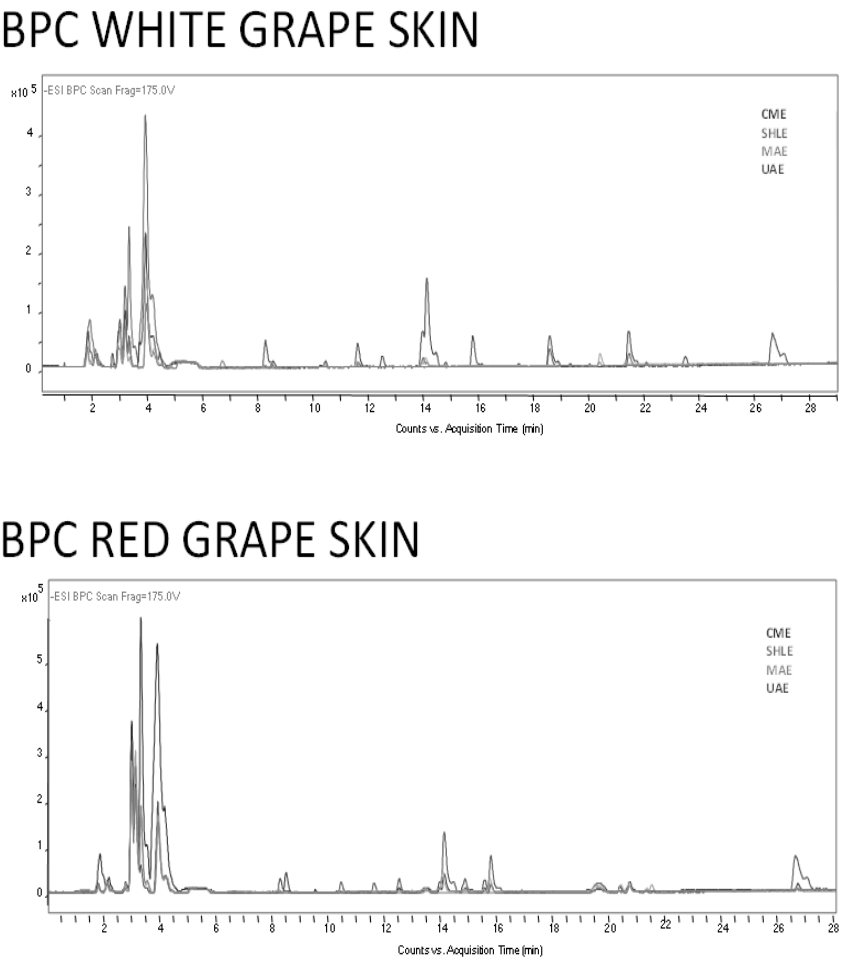
[25] http://www.plantcyc.org/

[26] J.A. Pérez-Serradilla, M.D. Luque de Castro, Food Chem., 124 (2011) 1652-1659.

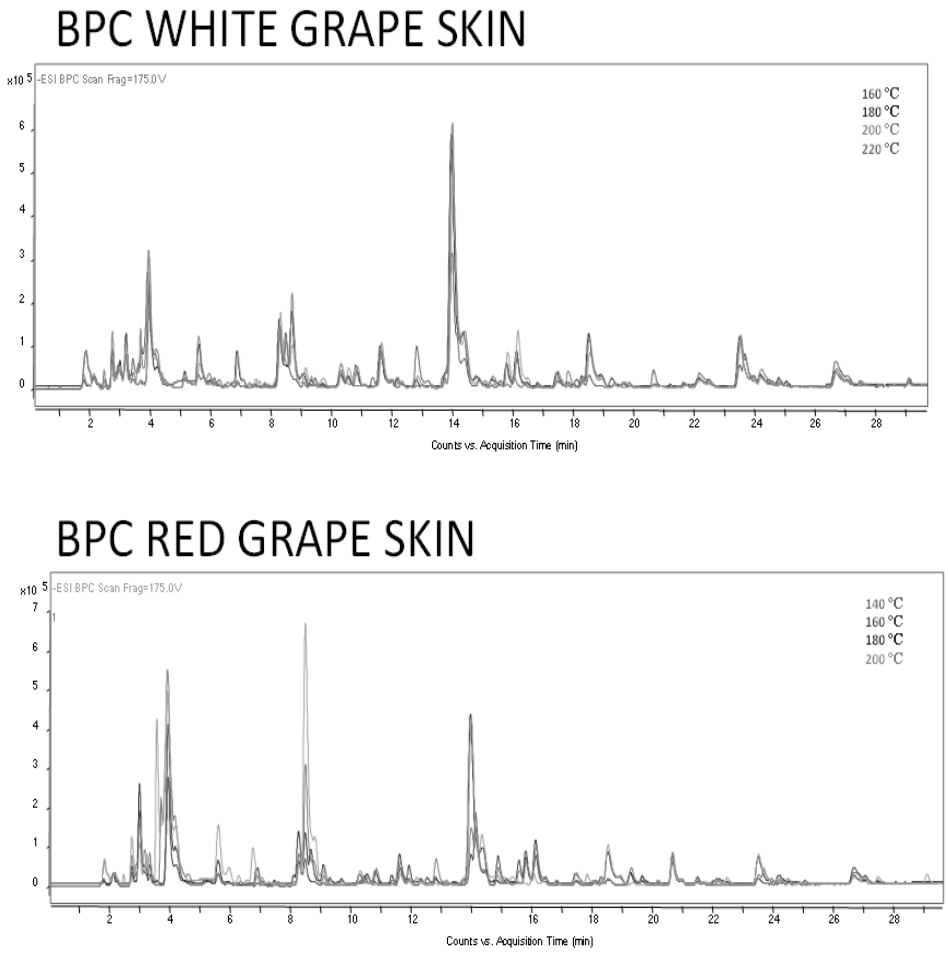
[27] J. Luque-Rodríguez, M.D. Luque de Castro, R. Japón-Luján, Exploitation of Residues from Vineyards, Olive Groves, and Wine and Oil Production to Obtain Phenolic Compounds of High-Added Value, in: Methods of Analysis for Functional Foods and Nutraceuticals, Second Edition, CRC Press, 2008.

[28] M. Corrales, A.F. García, P. Butz, B. Tauscher, J. Food Eng., 90 (2009) 415-421.

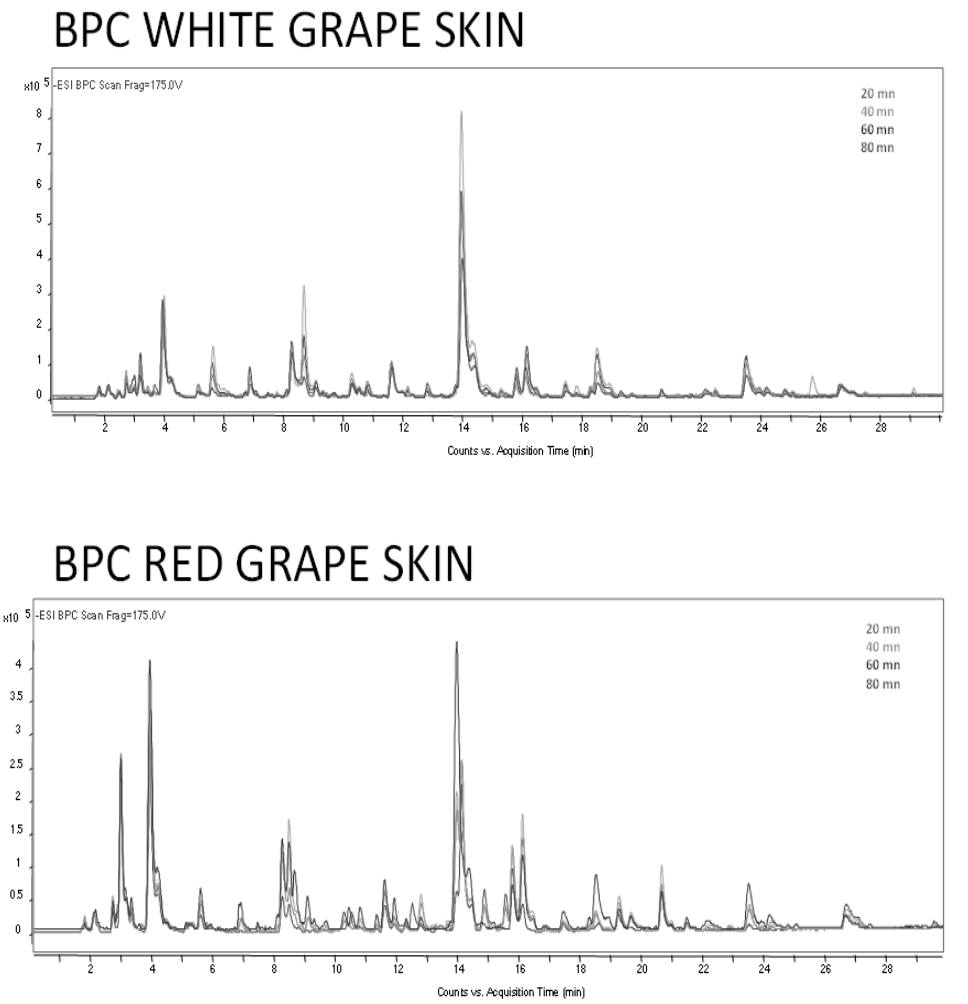
Supplementary material



Supplementary Figure 1. Chromatograms of extracts obtained by the four extraction methods: conventional maceration extraction (MCE), ultrasound-assisted extraction (UAE), microwave-assisted extraction (MAE), and superheated liquid extraction (SHLE) from white-grape skins (A) and red-grape skins (B).



Supplementary Figure 2. Chromatograms from SHL extracts from white and red grape skins subjected to four different extraction temperatures.



Supplementary Figure 3. Chromatograms from SHL extracts from white and red grape skins subjected to four different extraction times.

Capítulo IV:

Characterization of grape seed residues from the ethanol-distillation industry

Analytical methods

Volume 5, Issue 8, 21 February 2013,

Pages 1922-1929

Characterization of grape seed residues the from ethanol-distillation industry

Ángela Peralbo Molina, Feliciano Priego-Capote, María Dolores Luque de Castro*

[1]*Department of Analytical Chemistry, Annex C-3, Campus of Rabanales, 14071 Córdoba, Spain*

[2]*University of Córdoba Agroalimentary Excellence Campus, ceiA3, Campus of Rabanales, 14071 Córdoba, Spain*

[3]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital,*

University of Córdoba, E-14014, Córdoba, Spain

All the authors belong to the 3 institutions.

Characterization of grape seed residues from the ethanol-distillation industry

Ángela Peralbo Molina, Feliciano Priego-Capote, María Dolores Luque de Castro*

Abstract

The aim of the present research was to develop a fast and efficient extraction method for isolation of priced compounds from a raw material with very scant present value (waste from distillery industries which use the cake from red and white wine production) for their subsequent characterization. Superheated liquid extraction (SHLE) was the technique selected (and compared with conventional maceration extraction) to obtain the extracts which were overally characterized by the Folin-Ciocalteu (F-C) and Ferric Reducing Antioxidant Power (FRAP) tests. The extracts were also analyzed by LC-MS in high-resolution mode for qualitative comparison under different optimization conditions. Identification of the extract components was conducted by open databases Plantcyc (www.plantcyc.org) and METLIN Personal Metabolite Database. F-C and FRAP tests reported that the extracts have a high overall concentration of phenols and a high antioxidant activity, respectively (3.4 and 2.7 mg GAE mL^{-1} extract for white and red grape seeds, respectively; 4.8 and 3.8 mg TROLOX mL^{-1} extract for white and red grape seeds, respectively). The variety of identified compounds in the extracts makes them exploitable as additives in the food industry (either as colorants, flavor modifiers or antioxidants), and also in the cosmetics and nutraceutical industries. The relevance of SHL extraction variables has been emphasized with optimization studies.

1. Introduction

Grape seeds are a waste generated in wineries and grape juice industries. Seeds have been well characterized in order to take benefits from this agricultural residue. They contain lipids, proteins, carbohydrates, and, in a less proportion, secondary metabolites.[1] The lipid fraction of grape seeds has a high commercial value although this has been only partially exploited.[2] Grape seed oil is gaining popularity as a culinary oil that has been studied as a possible source of special lipids.[3] The main characteristic of grape seed oil is its high content of unsaturated fatty acids such as linoleic acid (72–76%, w/w)[4] associated to promotion of cardiovascular health by down-regulation of low-density lipoprotein cholesterol and to clearance enhancement.[5] In addition, grape seed oil contains tocopherols, which are also of interest because of their antioxidant properties.[6]

Despite grape seed flour, the residue from seed oil manufacture, has not received much attention, it may be a potential source of natural antioxidants and other healthy bioactive compounds.[7] Among secondary metabolites detected in grape seed, it is worth emphasizing the family of phenolic compounds, which are present in seeds at concentrations ranging between 5 and 8%, depending on the cultivar.[8] Phenols present in grape seeds are mainly simple phenols such as gallic acid, flavonoids including monomeric flavan-3-ols catechin, epicatechin, gallocatechin, epigallocatechin, and epicatechin 3-*O*-gallate; also procyanidin dimmers, trimmers, and more polymerized procyanidins.[9,10] Grape seed extracts of polar components are known by virtue of their antioxidant properties.[11] Extensive research shows that grape seed extracts are endowed with healthy properties because of their oxidation inhibition effect by bonding collagen, promoting youthful skin with higher elasticity and flexibility.[9] The oxidation inhibition power of proanthocyanidins has shown to be 20 times greater than that of vitamin E and 50 times greater than that of vitamin C;[12] and the action of proanthocyanidins to protect the body from sun damage, to improve vision and flexibility in joints, arteries, and cardiac tissue, and also to improve blood circulation by strengthening capillaries, arteries, and veins has been proved.[13]

Despite the healthy properties of the compounds in grape seeds, the residues of wine production are mainly destined to the alcohol industry, in which the remained grape seeds are exhausted through the alcohol isolation process and they have no commercial value.

With the aim of checking if after ethanol distillation grape seeds in the residue still contain profitable compounds, a superheated liquid extraction –SHLE– method has been developed. This way to accelerate extraction was selected as it has proved its efficiency for extraction of natural compounds from different types of vegetable samples such as leaves, fruits, roots or stems.[14-18] In fact the duality of extraction temperature and extractant composition together with the appropriate pressure play a crucial role to success in the leaching of target compounds. The qualitative composition of the extracts thus obtained was planned to be compared with that associated to the conventional maceration extraction –CME–, a soft process which requires long agitation times (at least for 24 h). This comparison based on qualitative and semiquantitative analysis constitutes a way to know both the shortening of the extraction time and potential changes in extracts composition caused for the drastic conditions of SHLE. Characterization of both types of extracts was planned as well as a tentative identification for a final evaluation of the potential of the extracts, with a present nil value. All these qualitative and semiquantitative studies, which globally constituted the aim of this research, could be the key for giving an added value to the nil-priced cake from the ethanol distillation industry.

2. Experimental

2.1. Samples

The raw material used for the present study was the cake from wineries after being subjected to the process for ethanol extraction in the distillation industry, where the process, included as Supplementary Information, is carried out. "Alcoholeras Reunidas, S.A.", Argamasilla de Alba (Ciudad Real, Spain) was

the ethanol distillation industry which provided us with cakes from both red and white grapes. Seeds were separated from skins in both cases and the two types of seeds were dried for 72 h at 35 °C, then milled to get a 40-mesh particle size (less than 0.42 mm diameter).

2.2. Apparatus

Grape seeds were milled by a mechanical grinder (Ball Mill Restch MM301). Superheated liquid extractions were performed by a laboratory-made dynamic extractor, the description of which can be found elsewhere.[16]

A UV-visible molecular absorption Spectrophotometer Thermo Spectronic Helios Gamma (Waltham, MA, USA) was used to monitor the absorbance in the Folin–Ciocalteu (F–C) and Ferric Reducing Antioxidant Power (FRAP) tests. Shaking and centrifugation of the extracts were carried out by an MS2 Minishaker Vortex (IKA, Germany) and a Mixtasel centrifuge (Selecta, Barcelona, Spain), respectively.

All samples were analyzed by an 1200 Series HPLC system (Agilent Technologies, Waldbronn, Germany) equipped with a binary pump, a degasser, a well plate autosampler and a thermostated column compartment, which was coupled to an Agilent 6530 TOF mass spectrometer with a dual electrospray ionization (ESI) source.

2.3. Reagents

Ethanol (96% v/v) PA from Panreac (Barcelona, Spain) and distilled water were used to prepare the different ethanol–water mixtures. LC–MS grade formic acid and acetonitrile (ACN) were from Scharlab (Barcelona, Spain). Deionized water (18 MΩ·cm) was obtained from a Millipore (Bedfore, MA, USA) Milli-Q plus system, and n-hexane for liquid–liquid extractions was from LiChrosolv (Merck, Darmstadt, Germany).

TPTZ (2,4,6-Tris(2-pyridyl)-s-triazine) and Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid) were from Fluka (Buches, Switzerland). The F-C reagent, sodium carbonate, gallic acid and AAPH (2,2'-azobis-2-methyl-propanimidamide dihydrochloride) were from Sigma (St. Louis, USA).

2.4. Superheated liquid extraction

One g of milled grape seeds was placed into the extraction cell installed into the gas chromatograph oven; then, a relative high flow rate (7 mL min^{-1}) was used for 1 min to fill the cell rapidly. To ensure the absence of air inside the extraction cell, the restrictor valve was kept open until the first drop of extractant appeared. At that moment, the restrictor valve was closed and when the preset pressure was reached, the switching valve was closed, the pump was turned off and the oven was switched on. During temperature rising, the switching valve had to be opened at short intervals to prevent the pressure from surpassing the working value. Once the selected temperature and pressure were reached, static extraction was performed for a preset time by closing the valves that allow circulation of the extractant through the cell. Finally, the oven was switched off, the chamber was cooled below the boiling point of ethanol and then, the switching valve and the restrictor valve were switched to enable new extractant to flow through the cell and flush out the extract.

2.5. Conventional maceration extraction (CME)

Two g of grape seeds was extracted with 20 mL 50% (v/v) 0.8% (v/v) HCl aqueous-ethanol at pH= 1 by stirring at 40 °C for 24 h.

2.6. Determination of total phenols by the F-C method

The total amount of phenolic compounds was quantified by the F-C method using gallic acid as standard. With this purpose, a calibration curve was run using solutions of 1, 10, 25, 50, 75 and 100 mg L^{-1} of this acid ($y = 0.0065x + 0.1286$, $R2 = 0.9909$). A 0.5 mL aliquot of extract, 10 mL of distilled water, 1 mL

of F-C reagent and 3 mL of Na_2CO_3 (20%, w/v) were mixed, made to 25 mL with distilled water and heated at 50 °C for 5 min. After heating, the samples were kept at room temperature for 30 min and, finally, the absorbance was measured at 765 nm against a blank solution containing distilled water instead of extract. The concentration of phenolic compounds thus obtained was multiplied by the dilution factor of the extract volume and divided by the amount of grapes seed used. The results were expressed as equivalent to milligrams of gallic acid per mL of grapes seeds extract (mg GAE mL–1).[16]

2.7. FRAP assay

The antioxidant potential (AOP) of grape seed extracts was determined using the FRAP assay, based on reduction of ferric tripyridyltriazin [Fe (III)-TPTZ] complex to ferrous tri-pyridyl triazin [Fe (II)-TPTZ] at low pH by the presence of antioxidants in the target solution. The ferrous complex [Fe (II)-TPTZ] has an intense blue color, which can be monitored at 593 nm. The assay response was standardized against the antioxidant standard Trolox.

2.8. LC-TOF/MS analysis

An injection volume of 20 µL and a flow rate of 1 mL min^{-1} were used. The mobile phases were 0.1% (v/v) formic acid aqueous solution (phase A) and pure acetonitrile (phase B). Separation of the analytes was performed with an Inertsil ODS-2 C18 column (250 mm × 4.6 mm i.d., 5 µm particle size, Análisis Vínicos, Tomelloso, Ciudad Real, Spain) at 25 ºC. The gradient method was as follows: from 96% to 50% A in 60 min, from 50% to 0% A in 5 min. An equilibration time of 8 min was required to assure column conditioning.

The dual ESI source operated in the negative ionization mode using the following conditions: nebulizer gas at 35 psi and drying gas flow rate and temperature at 10 L min^{-1} and 325 ºC, respectively. The capillary voltage was set at 3500 V, while the fragmentor, skimmer and octapole voltages were fixed at 175, 65 and 750 V, respectively. Data were acquired in centroid mode in the

extended dynamic range (2 GHz). Full scan was carried out at 1 spectrum per s within the *m/z* range 100–1700 with subsequent activation of the three most intense precursor ions (only allowed charge, single or double) by MS/MS using a collision energy of 20 eV at 1 spectrum per s within the *m/z* range 100–1700. Before the experiments, the instrument reported mass detection resolution of 25000 FWHM (Full Width at Half Maximum) at *m/z* 112.985587 and 45000 FWHM at *m/z* 966.000725. To assure the desired mass accuracy of recorded ions, continuous internal calibration was performed during analyses with the use of signals at *m/z* 119.0362 (proton abstracted purine) and *m/z* 966.0007 –formate adduct of hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazine.

MassHunter Workstation Data acquisition software (Agilent Technologies) was used to control the instrument. Data were processed using MassHunter Qualitative Analysis software (Agilent Technologies). Extraction of unknown molecular features from raw data was carried out by the Molecular Feature Extraction (MFE) algorithm in MassHunter Qualitative analysis software. The feature extraction algorithm took into account all ions exceeding 1000 counts with a charge state equal to or above one and a feature had to be composed of two or more ions to be valid (e.g. two ions in the isotope cluster). The theoretical formula adjusted to the corresponding isotopic distribution of molecular features was generated with the Molecular Formula Generation software (Agilent Technologies). Using background subtracted data, files in compound exchange format (.cef files) were created for each sample and exported into the Mass Profiler Professional (MPP) software package (version 2.0, Agilent Technologies, Santa Clara, CA, USA) for further processing. In the next step, alignment of RT and *m/z* values was carried out across the sample set using a tolerance window of 0.15 min and 20 ppm, respectively. The resulting files (.cef files) were exported to search molecular entities in the raw data files according to retention times and *m/z* values using the above parameters. This second re-extraction was used to look for missing data. Compound identification was performed using the METLIN Personal Metabolite Database and PlantCyc (http://plantcyc.org/).

3. Results and discussion

A hydroalcoholic extractant was selected to favor isolation of polar and mid-polar metabolites. Ethanol was used as organic solvent avoiding other toxic solvents such as methanol, acetonitrile, chloroform or acetone, because of the potential use of the extracts for human consumption. Apart from that, an acid pH was used to enhance hydrolysis of polymeric structures and release monomeric metabolites that can be easily solubilized. Since the raw material used in this research had previously been exploited for alcohol extraction, selection of an extreme acid pH was mandatory to exhaust it up to near completion. Other extraction conditions used were based on previous SHLE methods found in the literature for isolation of natural products from raw vegetal materials.[16] The pressure was maintained at above 100 bar by closing the inlet and outlet valves of the system to keep the extractant liquid state during the whole process.[19]

3.1. Influence of critical SHLE variables

The two most critical variables influencing the SHLE process were studied for isolation of high-added value compounds from these vegetal residues, namely: the extraction temperature and processing time (the working pressure was adjusted to keep the extractant in liquid state). Apart from that, the extractant pH was also considered in this optimization study. The results obtained after a univariate study of both variables are exposed below.

Extraction temperature. Temperature is the most important parameter influencing the kinetics of mass transfer in SHLE as the increase of this parameter also increases the solubility of compounds to be extracted. Notwithstanding the foregoing, using a high temperature does not always guarantee increased extraction efficiency owing to adverse effects such as thermolability of the compounds. In this study, the extraction temperature was varied from 160 to 220 °C by setting the extraction time at 60 min. Supplementary Figure 1 illustrates the Base Peak Chromatograms (BPC) provided by the

extracts from each raw material at the tested temperatures, showing the differences between the chromatographic profiles they provided. Figure 1 shows the Venn diagrams representing the influence of the extraction temperature on the detection coverage for the two different raw materials. The highest number of features detected in extracts from white and red grape seeds were found at 220 and 180 °C with 125 and 128 features, respectively. The influence of temperature was deduced by comparison of the detection coverage for each temperature as Table 1 shows. As can be seen, the extracts obtained at 180 ºC presented the highest detection coverage by comparison to the extracts obtained at the other three temperatures. Thus, a minimum of 65% of the molecular features detected in the extract obtained at 180 ºC was also detected at other temperature in white grape seeds.

On the other hand, this proportion was considerably increased up to 92.5% in red grape seeds, which reported a high similarity in the qualitative profile of extracts at different temperatures. The extracts obtained at 220 ºC were characterized by qualitative metabolite differences as compared to those isolated at lower temperatures. With these premises, the optimum extraction temperature for white and red grape seeds was 180 °C.

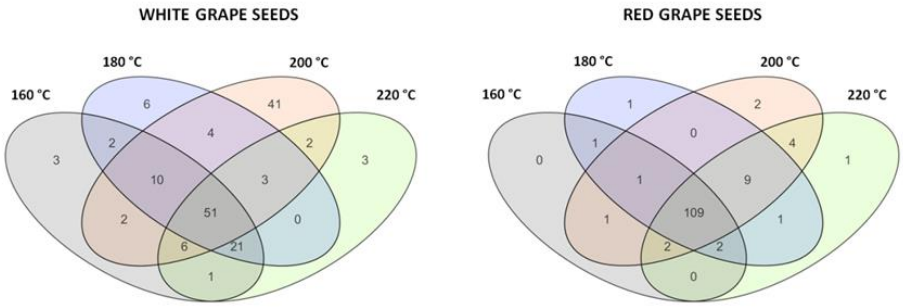


Figure 1: Venn diagrams (obtained after processing the MFs from white- and red grape seed extracts) that show the number of tentative compounds extracted at each extraction temperature.

Table 1. Percentage of molecular extracted at a preset temperature which were also detected at the other temperatures.

WHITE GRAPE SEEDS				
Temperature (°C)/%MF	**160**	**180**	**200**	**220**
160		79.81	51.92	58.65
180	69.75		63.87	72.27
200	46.15	64.96		92.31
220	48.80	68.80	86.40	
RED GRAPE SEEDS				
Temperature (°C)/%MF	**160**	**180**	**200**	**220**
160		97.46	94.07	74.58
180	89.84		92.97	70.31
200	90.24	94.31		72.36
220	93.62	92.55	94.68	

Extraction time. The extraction time in SHLE is usually shorter than that usually required in conventional solid-liquid extraction protocols such as those based on maceration. Apart from that, the extraction time depends on the particular phenomenon which determines the extraction rate. In general, long extraction times provide higher extraction efficiencies. However, continuous exposure to high temperature can promote degradation thus altering the composition of the extracts. For this reason, it is worth developing a kinetics study to set the optimum extraction time. Therefore, 20, 40, 60 and 80 min were the tested extraction times. The results, plotted as Venn diagrams, are shown in Figure 2, demonstrated that an extraction time of 60 min provided the highest number of molecular features for the two matrices (119 and 128 for white and red grape seeds, respectively); for this reason, the optimum extraction time was set at 60 min. Supplementary Figure 2 plots BPCs corresponding to the analysis of extracts obtained at different extraction times from the two types of samples.

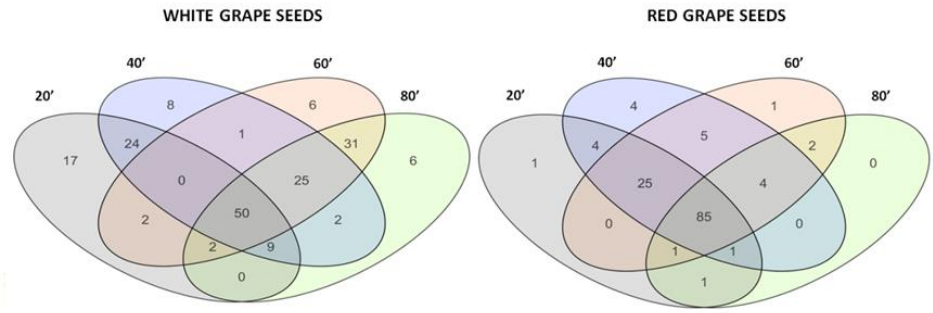


Figure 2. Venn diagrams (obtained after processing the MFs from white- and red grape seed extracts) that show the number of tentative compounds extracted at each extraction time.

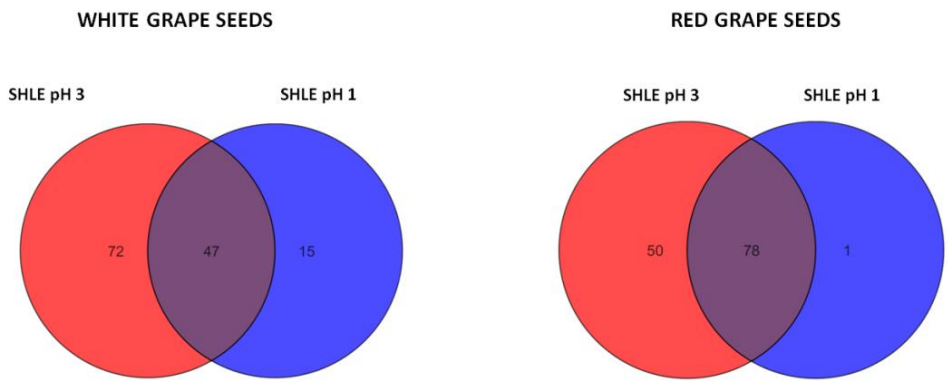


Figure 3. Venn diagrams (obtained after processing the MFs for white- and red grape seed extracts) that show the number of compounds obtained with SHLE at two pH values (1 and 3).

Extractant pH. The pH of the extractant is one other critical variable for isolation of phenolic compounds from raw vegetal materials. Acid pH of extractants favor solubilization of phenolic compounds but also the transference to liquid phases due to hydrolysis of cellular membranes. Two pH values, 1 and 3, were evaluated in SHLE for qualitative comparison of the extracts. Supplementary Figure 3 shows the BPC chromatograms obtained for both types of grapes seeds at the two extraction pHs. As can be seen, the extraction efficiency was superior at pH 3 since most molecular entities detected in the extracts isolated at pH 1 were also detected in the extract obtained at pH 3.

Figure 3 illustrates the Venn diagrams comparing the qualitative profiles reported at both pH values for red and white grape seeds. Tentatively, this could be ascribed to degradation of a representative fraction of molecular entities since 72 and 50 entities were only detected in pH 3 extracts from white and red grape seeds.

3.2. Comparison of extraction techniques

The efficiency of SHLE for isolation of compounds from grape seeds as sample preparation strategy for characterization of this residue was evaluated by comparison to CME as reference extraction technique. Proper comparison required to use the same amount of material, volume and composition of the extractant. This was set at 50:50 (v/v) ethanol–water according to preliminary experiments. Since the same ratio between sample weight and extractant volume was used in both isolation protocols, the extracts obtained were analyzed by LC-TOF/MS in accurate mode with the same dilution pattern. The efficiency of the extraction alternatives was assessed by comparison of the molecular features obtained for each extraction conditions as described in the experimental section. The Venn diagrams thus obtained are shown in Figure 4 for each vegetable raw material used. Different results were found for extracts from red and white grape seeds. Thus, SHLE was clearly the most suited extraction mode for red grape seeds. As can be seen, most entities detected in the extracts obtained by CME (above 96%) were also detected in the extract obtained by SHLE. Concerning white grape seeds qualitative differences were distinguished between extracts provided by CME and SHLE. A representative number of entities, 33 and 55, were exclusively detected in the extracts provided by CME and SHLE, respectively. Attending to these results, SHLE allowed reducing the extraction time from 24 h in CME to 60 min, although qualitative differences were only observed in the case of white grape seeds.

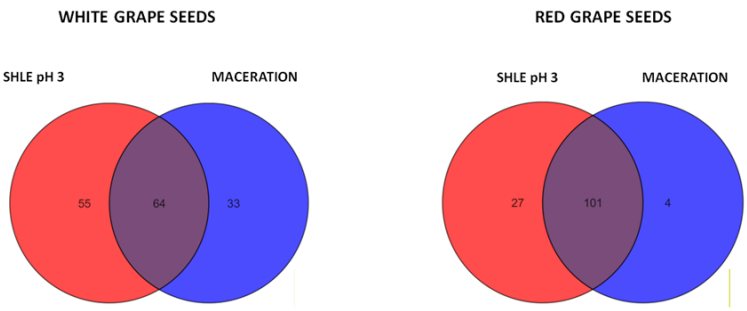


Figure 4. Venn diagrams (obtained after processing the MFs for white- and red grape seeds extracts) that show the number of tentative compounds obtained with SHLE and conventional maceration.

3.3. Total phenolic content by the F-C method and evaluation of the antioxidant potential by the FRAP assay

Two different tests were used to evaluate the potential of the types of seeds used as raw material for providing interesting compounds such as phenols or, more generically, antioxidants. The F-C test revealed that the extracts from white grapes seeds reported higher concentrations of total phenols (with 3.4 mg GAE mL^{-1} of extract for seeds) than those from red grapes seeds, with 2.7 mg GAE mL^{-1} for seeds. Concerning the antioxidant activity, the same trend was observed; therefore, the extract from white grape seeds reported the highest antioxidant activity, with 4.8 mg Trolox mL^{-1} of extract versus red grapes that gave 3.8 mg Trolox mL^{-1}.

3.4. Identification of interesting compounds in seeds from white and red grapes

Once the extraction conditions were defined for isolation of the polar/mid-polar fraction from the two grape seeds obtained after grape pressing and alcohol distillation, characterization of the resulting extracts was demanded for valorization of the residues. Numerous compounds were identified from both types of seeds supported on mass accurate detection with 10 ppm maximum error in monoisotopic mass. Metabolites were identified by searching in PlantCyc

database (www.plantcyc.org) and in a personnel database prepared after review of existing research in the literature about characterization of seeds grapes.[11,20,21] This section was focused on two main families that should be emphasized because of their nutraceutical and oenological interest: phenols and furfural derivatives, which are produced by sugars degradation.

Flavanols occur mainly in the outer coating of seeds, although they can also be detected in some skin cells. The most abundant flavanols in grapes are flavan-3-ols, characterized by a completely saturated heterocycle ring. They can be present as monomers, oligomers and polymeric structures. Among monomers, epigallocatechin and gallocatechin-O-gallate were detected in extracts from both types of seeds, while epicatechin was exclusively found in extracts from red grape seeds. Oligomers and polymers of flavan-3-ols are mainly procyanidins from the B-series, which are composed by catechin, epicatechin and epicatechin-3-O-gallate connected by C4–C8 or C4–C6 bonds. In this study, procyanindin B2 was detected in white grape seeds.

Flavonols are abundant in seeds of fresh grapes mainly as 3-*O*-glucosides of kaempferol, quercetin, myricetin, and isorhamnetin.[22,23] Simple flavonols such as isorhamnetin were detected in extracts from both types of seeds, while kaempferol was only detected in extracts from red grape seeds, and quercetin in the white ones. Quercetin-3-*O*-glucoside, myricetin-7-*O*-glucoside and isorhamnetin-3-*O*-glucoside were detected in extracts from grape seeds. A derivative of quercetin such as astilbin (dihydroquercetin-3-*O*-rhamnoside) was only detected in the extracts from red grape seeds. Attending to these results, the extracts from grape residues still contain flavonoid-type interesting compounds despite the treatment of alcohol distillation that practically eliminated anthocyanins and derivatives in red grape seeds.

Table 2. Tentative metabolites found (√) and not found (χ) in the extracts from grape seeds classified into families.

FLAVONOIDS	THEORETICAL *m/z*	EXPERIMENTAL *m/z*	FORMULA	GRAPE SEED	
				RED	WHITE
ANTHOCYANINS					
MALVIDIN-3-GLUCOSIDE ACETALDEHYDE (VITISIN B)	528.1035	528.1028	$C_{23}H_{25}ClO_{12}$	χ	χ
FLAVANOLS					
EPICATECHIN	290.079	290.0807	$C_{15}H_{14}O_6$	√	χ
CATECHIN	290.079	290.0807	$C_{15}H_{14}O_6$	χ	χ
EPICATECHIN-O-GALLATE	442.09	442.0902	$C_{22}H_{18}O_{10}$	√	√
EPIGALLOCATECHIN	306.074	306.0725	$C_{15}H_{14}O_7$	√	√
GALLOCATECHIN-O-GALLATE	458.0849	458.0867	$C_{22}H_{18}O_{11}$	χ	√
PROCYANIDIN B2	578.1424	578.1436	$C_{30}H_{26}O_{12}$	√	χ
FLAVONOLS					
KAEMPFEROL	286.0477	286.0489	$C_{15}H_{10}O_6$	χ	χ
MYRICETIN	318.0376	318.0395	$C_{15}H_{10}O_8$	χ	√
QUERCETIN	302.0427	302.0427	$C_{15}H_{10}O_7$	√	√
ISORHAMNETIN	316.0583	316.0613	$C_{16}H_{12}O_7$	χ	χ
MYRICETIN-3-O-GLUCOSIDE	480.0904	480.0925	$C_{21}H_{20}O_3$	√	√
QUERCETIN-3-O-GLUCOSIDE	464.0955	464.0964	$C_{21}H_{20}O_{12}$	χ	√
ASTILBIN	450.1162	450.1152	$C_{23}H_{22}O_{33}$	√	χ

NON FLAVONOIDS	THEORICAL *m/z*	EXPERIMENTAL *m/z*	FORMULA	GRAPE SEEDS	
				RED	WHITE
PHENOLIC ACIDS					
CINNAMIC ACID	148.0524	148.0523	$C_9H_8O_2$	χ	χ
CAFFEIC ACID	180.0423	180.0423	$C_9H_8O_4$	√	√
***o*-COUMARIC ACID**	164.0473	164.0476	$C_9H_8O_3$	√	√
FERULIC ACID	194.0579	194.0574	$C_{10}H_{10}O_4$	√	√
GALLIC ACID	170.0215	170.022	$C_7H_6O_5$	√	√
4-HYDROXYBENZOIC ACID	138.0317	138.0317	$C_7H_6O_3$	√	√
PROTOCATECHUIC ACID	154.0266	154.0268	$C_7H_6O_4$	√	√
SYRINGIC ACID	198.0528	198.0527	$C_9H_{10}O_5$	χ	χ
STILBENES					
TRANS-RESVERATROL	228.0786	228.0786	$C_{14}H_{12}O_3$	√	√
α-VINIFERIN	678.189	678.188	$C_{42}H_{30}O_9$	χ	χ
ASTRINGININ	244.0736	244.0736	$C_{14}H_{12}O_4$	χ	√
TRANS-ε-VINIFERIN	454.1416	454.1426	$C_{28}H_{22}O_6$	χ	χ
PICEID	390.1315	390.1331	$C_{20}H_{22}O_8$	√	χ
TRANS-PTEROSTILBENE	256.1099	256.1097	$C_{16}H_{16}O_3$	√	χ
FURANIC					

OTHER PHENOLS	THEORICAL *m/z*	EXPERIMENTAL *m/z*	FORMULA	GRAPE SEEDS	
				RED	WHITE
ETHYL FERULATE	222.0892	222.0902	$C_{12}H_{14}ClO_4$	√	√
ETHYL PROTOCATECHUATE	182.0579	182.0583	$C_9H_{10}O_4$	√	√
PYROCATECHOL	110.0368	110.0368	$C_6H_6O_2$	√	√
PYROGALLOL	126.0317	126.0311	$C_6H_6O_3$	√	√
SYRINGOL	154.063	154.225	$C_8H_{10}O_3$	χ	χ
4-METHYLPYROCATECHOL	124.0524	124.0519	$C_7H_8O_2$	√	√
4-VINYLGUAIACOL	150.0681	150.0682	$C_9H_{10}O_2$	√	√
2-PHENYLACETALDEHYDE	120.0575	120.0578	C_8H_8O	√	√
4-HYDROXYBENZALDEHYDE	122.0368	122.0365	$C_7H_6O_2$	√	√
CONIFERALDEHYDE	178.063	178.0627	$C_{10}H_{10}O_3$	√	√
***p*-HYDROXYBENZALACETONE**	162.0681	162.0683	$C_{10}H_{10}O_2$	χ	√
VANILLIN	152.0473	152.0472	$C_8H_8O_3$	√	√
1-3-4-DIHYDROXYPHENYL-1-PROPANONE	166.063	166.065	$C_9H_{10}O_3$	χ	χ
2,6-DIMETHOXY-4-1-PROPENYL-PHENOL	194.0943	194.0941	$C_{11}H_{14}O_3$	χ	χ
4-ETHOXYPHENOL	138.0681	138.0878	$C_8H_{10}O_2$	χ	√

Phenolic acids in grapevine mainly include gallic, ellagic, and hydroxycinnamic acids. They are usually present as esters, and are also found in free form in grape pomace, probably as a result of partial hydrolysis of esters during extraction. Hydroxycinnamic acids such as caffeic, *p*-coumaric and ferulic acids were detected in extracts from both types of seeds. Gallic acid is one of the main phenolic acids found in grapes, mainly esterified to flavan-3-ols, and also as a constituent of hydrolyzable tannins, which are esters containing one or several molecules of gallic (gallotannins) or ellagic acid (ellagitannins). Gallic acid was found in seeds from both types of grapes. Other phenolic acids such as 4-hydroxybenzoic and protocatechuic acids were also detected in extracts from both types of grape seeds. Concerning phenolic esters, protocatechuate was detected in extracts from both residues.

Stilbenes in grapevine are mainly present as constituents of the woody organs (roots, vine-shoots, and stems), and also as substances of induced production (in leaves and berries) acting as phytoalexins in the mechanisms of grape resistance against certain pathogens such as fungi.[24] Stilbene synthesis in

grapes depends on various viticultural factors such as grape variety, the environment, and cultural practices. Concerning grape variety, red grapes possess higher stilbene levels than white grapes,[25] which also remain in the residue after ethanol distillation. The many stilbenic compounds found in various parts of grapes are derived from *trans*-resveratrol (3,5,40-trihydroxystilbene) and occur as glucosylated derivatives or oligomeric forms of this compound called viniferins. Resveratrol was detected in all extracts from red and white grapes. Apart from them, two other derivatives such piceid (the 3-*O*-glucoside of resveratrol) and astringinin (trans-3,30,4,50-tetrahydroxystilbene) were also detected at trace level in seed extracts. One other minor component such as pterostilbene (the antifungal activity of which is high relative to resveratrol and viniferins)[26] was detected in seed extracts from red grapes.

Other additional phenols were detected in extracts from grapes using SHLE. Thus, hydroxybenzene derivatives such as pyrocatechol and pyrogallol or 4-methylpyrocatechol and 4-vinylguaiacol were detected in seeds from residues from white and red grapes. Phenolic aldehydes such as 2-phenylacetaldehyde, 4-hydroxybenzaldehyde and coniferaldehyde, and ketones such as vanillin (in both cases) and *p*-hydroxybenzalacetone (in white grape seed extracts) were also detected. These compounds deserve to be mentioned because of their characteristic flavor. Finally, ethoxy derivatives were found at trace level in white grape seed extracts.

Concerning furanic compounds, only two compounds were present in the extracts, but at very low concentration. Thus, acetoxymethylfurfural, furfuryl alcohol and 2-furancarboxylic acid ethyl ester were detected in SHL extracts.

Thus, despite the degradation caused by the drastic conditions of ethanol distillation, many interesting compounds are found in the extracts from the waste of this process, which make it a useful matter for more than as a heating source. The variety of the extracted compounds makes them exploitable as additives in the food industry (either as colorants, flavor modifiers or antioxidants), and also in the cosmetics and nutraceuticals industries.

Conclusions

The proposed extraction method offers interesting analytical characteristics for a favorable substitution of the conventional method, making it an advantageous alternative to obtain extracts for the study of the raw material from the distillation process. The drastic reduction of the extraction times with not significant loss of valuable compounds support the use of the proposed method.

The identification of interesting compounds in the extracts shows that despite the degradation caused by the drastic conditions of the distillation process, the waste from this industry constitutes a useful raw matter for more than as a heating source. The variety of identified compounds in the extracts makes them exploitable as additives in the food industry (either as colorants, as flavour modifiers or as antioxidants), and also in the cosmetics and nutraceuticals industries.

Acknowledgements

The Spanish Ministerio de Ciencia e Innovación (MICINN) and FEDER program are acknowledged for financial support through project CTQ2009-07430. F.P.C. is grateful to MICINN for a Ramón y Cajal contract (RYC-2009-03921).

References

1. L. M. de Campos, F. V. Leimann, R. C. Pedrosa and S. R. Ferreira, Bioresour. Technol., 2008, 99, 8413-8420.

2. C. P. Passos, R. M. Silva, F. A. Da Silva, M. A. Coimbra and C. M. Silva, Chem. Eng. J., 2010, 160, 634-640.

3. T. H. Beveridge, B. Girard, T. Kopp and J. C. Drover, J. Agric. Food Chem., 2005, 53, 1799-1804.

4. M. Martinello, G. Hecker and M. C. Pramparo, J. Food Eng., 2007, 81, 60-64.

5. V. Wijendran and K. C. Hayes, Annu. Rev. Nutr., 2004, 24, 597-615.

6. C. Crews, P. Hough, J. Godward, P. Brereton,M. Lees, S. Guiet and W. Winkelmann, J. Agric. Food Chem., 2006, 54, 6261-6265.

7. M. Luther, J. Parry, J. Moore, J. Meng, Y. Zhang, Z. Cheng and L. Yu, Food Chem., 2007, 104, 1065-1073.

8. F. Felice, Y. Zambito, G. Di Colo, C. D'Onofrio, C. Fausto, A. Balbarini and R. Di Stefano, Eur. J. Pharm. Biopharm., 2012, 80, 176-184.

9. Y. Cai, Y. Yu, G. Duan and Y. Li, Food Chem., 2011, 127, 1872-1877.

10. Y. Yilmaz and R. T. Toledo, J. Agric. Food Chem., 2004, 52, 255-260.

11. I. I. Rockenbach, L. V. Gonzaga, V. M. Rizelio, A. E. S. S. Gonçalves, M. I. Genovese and R. Fett, Food Res. Int., 2011, 44, 897-901.

12. J. Shi,J. Yu,J. E. Pohorly and Y. Kakuda, J. Med. Food, 2003, 6, 291-299.

13. E. E. Yilmaz,E. B. Özvural and H. Vural, J. Supercrit. Fluids, 2011, 55, 924-928.

14. J. González-Rodríguez, P. Pérez-Juan and M. D. Luque de Castro, J. Chromatogr. A, 2004, 1038, 3-9.

15. R. Japón-Luján and M. D. Luque de Castro, J. Chromatogr. A, 2006, 1136, 185-191.

16. J. M. Luque-Rodríguez, M. D. Luque de Castro and P. Pérez-Juan, Biores. Technol., 2007, 98, 2705-2713.

17. J. M. Roldán-Gutiérrez, J. Ruiz-Jiménez and M. D. Luque de Castro, Talanta, 2008, 75, 1369-1375.

18. M. P. Delgado-Torre, F. Priego-Capote and M. D. Luque de Castro, J. Agric. Food Chem., 2012, 60, 3409-3417.

19. M. P. Delgado-Torre, C. Ferreiro-Vera, F. Priego-Capote, P. M. Pérez-Juan and M. D. Luque de Castro, J. Agric. Food Chem., 2012, 60, 3051-3060.

20. A. V. Perumalla and N. S. Hettiarachchy, Food Res. Int., 2011, 44, 827-839.

21. I. I. Rockenbach, E. Jungfer, C. Ritter, B. Santiago-Schübel, B. Thiele, R. Fett and R. Galensa, Food Res. Int., 2012, 48, 848-855.

22. V. Cheynier and J. Rigaud, Am. J. Enol. Viticult., 1986, 37, 248-252.

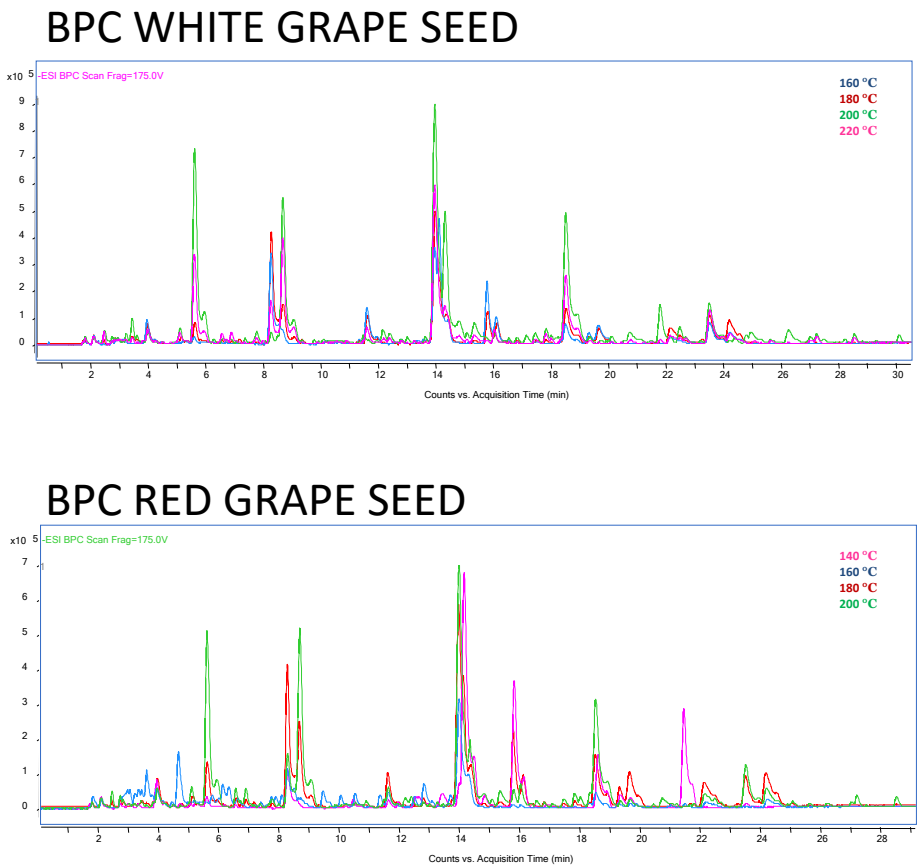
23. V. Freitas, P. Silva Porto, M. Assuncao, A. Cadette-Leite, J. P. Andrade and P. Barbosa, Alcohol Alcoholism, 2004, 39, 303-311.

24. J. M. Luque-Rodríguez, M. D. Luque de Castro and R. Japón-Luján, Exploitation of Residues from Vineyards, Olive Groves, and Wine and Oil Production to Obtain Phenolic Compounds of High-Added Value. In Methods of Analysis for Functional Foods and Nutraceuticals, Second Edition, CRC Press: 2008.

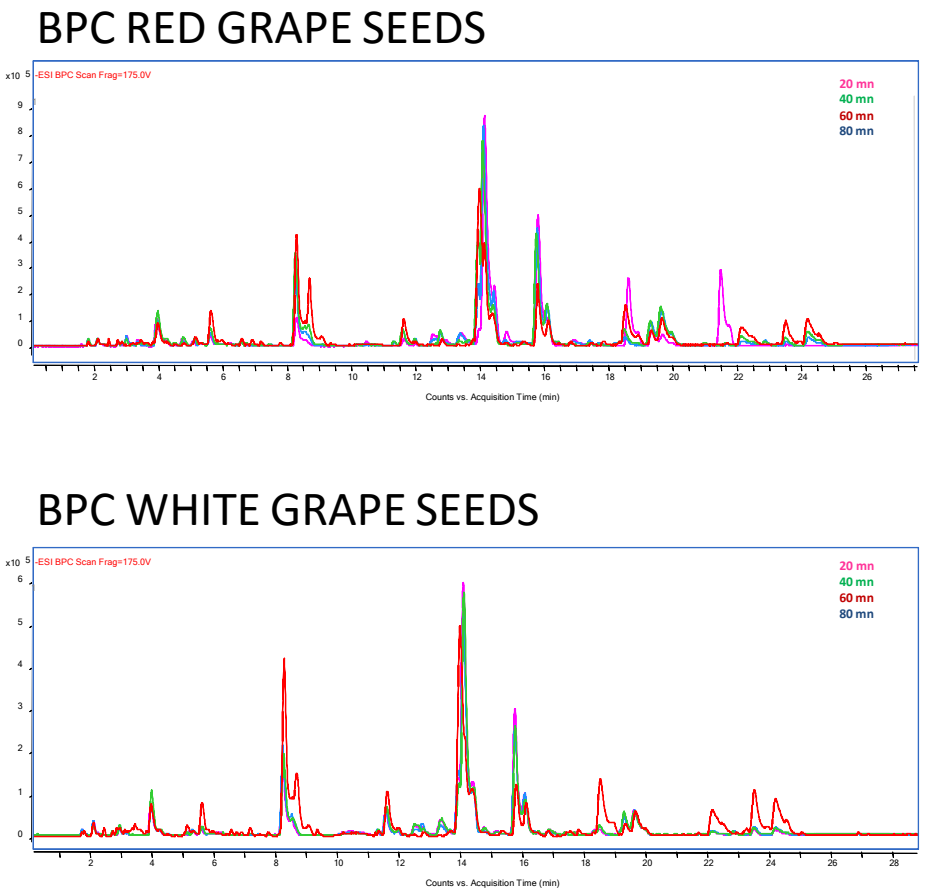
25. L. Bavaresco, Drugs Exp. Clin. Res., 2003, 29, 181-187.

26. P. Langcake, C. A. Cornford and R. J. Pryce, Phytochem., 1979, 18, 1025-1027.

Supplementary material

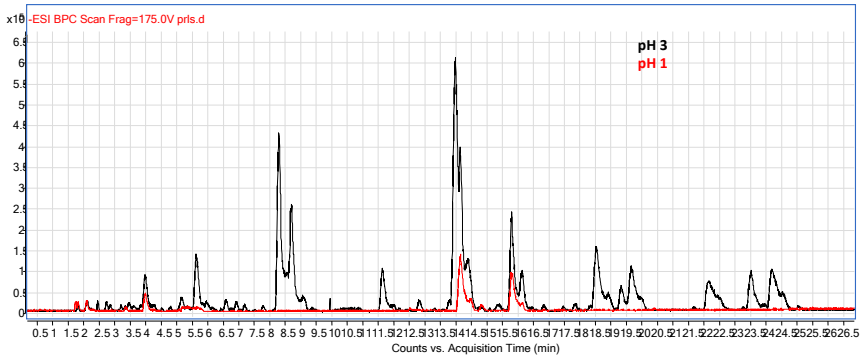


Supplementary Figure 1. Base peak chromatograms obtained by LC-TOF/MS analysis of white and red grape seed extracts obtained at four different temperatures.

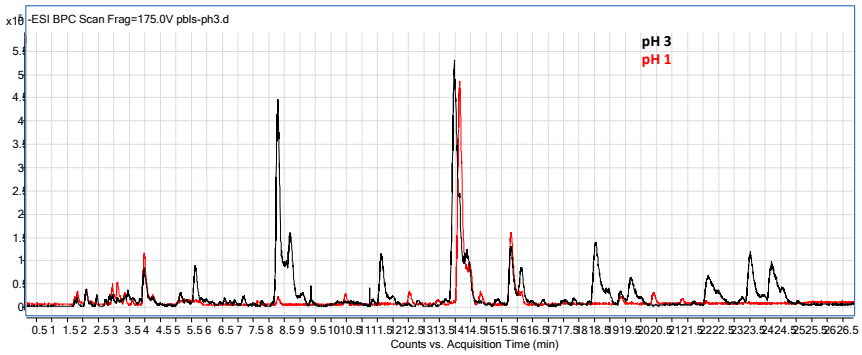


Supplementary Figure 2. Base peak chromatograms obtained by LC-TOF/MS analysis of white and red grape seed extracts obtained at four different extraction times.

BPC RED GRAPE SEEDS



BPC WHITE GRAPE SEEDS



Supplementary Figure 3. Base peak chromatograms obtained by LC-TOF/MS analysis of white and red grape seed extracts obtained at two different pHs.

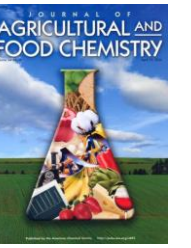
Capítulo V:

Tentative identification of phenolic compounds in olive pomace extracts using liquid chromatography-tandem mass spectrometry with a QqTOF mass detector

Journal of Agricultural and Food Chemistry

Volume 60 (46), 30 October 2012

11542-11550



Tentative identification of phenolic compounds in olive pomace extracts using liquid chromatography-tandem mass spectrometry with a QqTOF mass detector

Ángela Peralbo Molina, Feliciano Priego-Capote, María Dolores Luque de Castro*

[1]*Department of Analytical Chemistry, Annex C-3, Campus of Rabanales, 14014 Córdoba, Spain*

[2]*University of Córdoba Agroalimentary Excellence Campus, ceiA3, Campus of Rabanales, 14014 Córdoba, Spain*

[3]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital,*

University of Córdoba, E-14014, Córdoba, Spain

All the authors belong to the 3 institutions.

Tentative identification of phenolic compounds in olive pomace extracts using liquid chromatography–tandem mass spectrometry with a QqTOF mass detector

Ángela Peralbo Molina, Feliciano Priego-Capote, María Dolores Luque de Castro*

Abstract

The reutilization of agronomical residues is a pending goal for sustainable agriculture. Particular residues in olive oil producing countries are leaves, wastewater and olive pomace. Olive leaves and wastewaters have been previously characterized by isolation of the phenolic fraction. However, olive pomace has not been qualitatively characterized yet as source of phenols. Olive pomace extracts were obtained by using superheated solvent extraction using 50:50 (v/v) water–ethanol as leaching mixture at 160 ºC. The extracts were analyzed by liquid chromatography coupled to tandem mass spectrometry using a QqTOF hybrid mass analyzer (R=25000–45000). Qualitative analysis was supported on measurement of accurate masses for precursor and product ions as well as on their isotopic distribution. Identification was focused on the main families of phenolic compounds present in extra virgin olive oil. The potential of this residue as a rich source of phenols with antioxidant properties has been proved.

Keywords: Phenolic compounds; olive pomace; superheated liquid extraction; tandem mass spectrometry; QqTOF mass analyser.

1. Introduction

The olive oil industry produces vast amounts of a semi-solid by-product after virgin olive oil (VOO) decantation in the two-phase system. The by-product thus generated is a residue known as olive pomace, also known as alperujo. This is formed by a mixture of the liquid and solid phases (alpechín and orujo, respectively), which were the original residues from the previous three-phase process. As no water is added during the two-phase process, oil quality is superior to that provided by the three-phase system since the partition equilibrium enables a higher proportion of olive phenols to remain in the oil fraction. [1]

Once VOO is decanted, additional centrifugation steps or a second extraction process (beating and decanting) leads to second order olive oils characterized by high wax content. The isolation of these low-quality oils is economically viable so, this strategy is implemented in most olive oil factories.[2] This second extraction step has reduced the interest of olive pomace oil isolated with the aid of organic solvents and heat. Currently, olive pomace is used either as fuel in cogeneration plants to produce energy or as organic fertilizer after a composting process.[3] However, the profitability of cogeneration plants is still under doubt, whereas the fertilization industry only absorbs a very small proportion of the produced olive pomace; therefore, a frequent option is dumping the residue in controlled areas due to its high polluting organic load.

It is therefore evident that there is a demand for alternatives capable of taking benefits from olive pomace (either before or after extraction of the low-quality oil), which is characterized by a high content of organic substances such as sugars, tannins, polyalcohols, pectins, lipids and phenols.[4] However, no exhaustive characterization of any form of this residue has been performed to date. One of the most conflictive fractions of olive pomace corresponds to phenolic compounds.[5] Phenols are the major contributors to the antibacterial and phytotoxic activity of black olive-mill residues, which limit their microbial degradability and potential for biogas production.[6] Most phenolic classes present in VOO and olive oil residues are endowed with the features for being free radical

scavengers.[5] It is well known that the antioxidant activity of phenols is based on their capacity, as donors of hydrogen atoms, to neutralize the activity of free radicals.[7] As a general rule, the *o*-dihydroxyl structure characteristic of hydroxytyrosol and secoiridoid derivatives seems to be responsible for the high antioxidant activity of these phenols, followed by 4-*O*-monohydroxy compounds (ligstroside and tyrosol) and 3-O-hydroxy-substituted catechols.[8] With these premises, the properties of olive phenols may turn olive pomace into a cheap source of natural antioxidants in concentrations up to 100 times higher than in virgin olive oils.[4] The presence of these compounds in foods is crucial as they reduce oxidative reactions that decrease both nutritional value and sensory quality.[9]

From a clinical point of view, modest long-term intakes of hydrophilic phenols from olive-tree materials produce a favorable impact on the incidence of cancers and chronic diseases such as cardiovascular disorders and type II diabetes.[10] For example, hydroxytyrosol, a well-known phenol about which the EFSA (European Food Safety Authority) has recently emitted a very favorable opinion,[11] shows better results in scavenging and/or antioxidant capacity tests than common antioxidants such as vitamins C and E or 2,6-di-tert-butyl-4-methylphenol (BTH).[12] Other olive phenols (e.g. tyrosol,[13] verbascoside,[14] and α-taxifolin[15]) provide health benefits, as demonstrated by in vitro activities, in addition to a synergistic effect of their antioxidant capacities when in mixtures.[16] Nevertheless, these studies have shown some limitations when compared with in vivo research. Recent studies on the use of apigenin-7-glucoside[17,18] to fight against Alzheimer's[19] and/or liver diseases[20] support the health effects of this compound and constitute a call for continuing the research on olive phenols.

Research carried out so far for characterization of olive pomace extracts has been mainly focused on panels of target known compounds considered of interest, but global qualitative analysis of these extracts has not been made yet. Therefore, a global profiling of these compounds would be the key for better understanding the biochemical pathways in which they are involved. Similar

studies have been reported for olive leaves and wastewater olive mill residues.[21,22] The research reported here was aimed at obtaining a global metabolic profile of phenols from this unexploited raw material. With this purpose, the extracts obtained by ethanol–water mixtures under superheated conditions were the analytical sample subjected to liquid chromatography separation with tandem mass spectrometry detection using a QqTOF hybrid detector to take benefits from high mass accuracy.

2. Experimental

2.1. Samples

Olive pomace samples from different cultivars obtained at different crop dates in the 2009/2010 season were taken directly from the production line of VOO. These pomace samples were homogenized to prepare a pool representative for qualitative analysis. The mixture was dried in an oven at 30 ºC for 24 h to eliminate the water content (around 60–80%). After drying, the samples were stored at –20 ºC until use.

2.2. Reagents

Ethanol was from Panreac (Barcelona, Spain). LC–MS grade formic acid and acetonitrile were purchased from Scharlab (Barcelona, Spain).[18] MΩ·cm deionized water from a Millipore Milli-Q water purification system was used to prepare both the water–ethanol extractant mixtures and chromatographic mobile phases. The Folin–Ciocalteu (F–C) reagent, sodium carbonate and gallic acid were from Sigma (St. Louis, USA).

2.3. Apparatus and instruments

Superheated liquid extractions (SHLE) were performed by a laboratory-made dynamic extractor[23] consisting of the following units: a) an extractant supplier; b) a high-pressure pump (Shimadzu LD-AC10) which propels the

extractant through the system; c) a switching valve placed next to the pump to develop static extractions; d) a stainless-steel cylindrical extraction chamber (550 mm × 10 mm i.d., 4.3 mL internal volume) where the sample is placed. This chamber is closed at both ends with screws whose caps contain cotton made filters to ensure the sample is not carried away by the extractant; e) a restriction valve to maintain the desired pressure in the system; f) a cooler made of a stainless steel tube (1 m length, 0.4 mm i.d.) and refrigerated with water; g) a gas chromatograph oven (Konix, Cromatix KNK-2000) where the extraction chamber is placed and heated.

The absorbance of the extracts after reaction with the F-C reagent was monitored by a Thermo Spectronic Helios Gamma spectrometer (Waltham, MA, USA). Shaking and centrifugation of the extracts were carried out by means of an MS2 Minishaker Vortex (IKA, Germany) and a Mixtasel centrifuge (Selecta, Barcelona, Spain), respectively.

All samples were analyzed by an Agilent 1200 Series LC system (Agilent Technologies, Waldbronn, Germany) coupled to an Agilent 6540 QqTOF hybrid mass spectrometer with a dual electrospray ionization (ESI) source for simultaneous spraying of a mass reference solution that enabled continuous calibration of detected *m/z* ratios.

2.4. Superheated liquid extraction

Two g of dried olive pomace was placed into the extraction cell installed in the gas chromatograph oven; then, a relative high flow rate (7 mL/min) was used for 1 min to fill the cell rapidly. To ensure the absence of air inside the extraction cell, the restrictor valve was kept open until the first drop of extractant appeared. At that moment, the restrictor valve was closed and when the desired pressure was reached, the switching valve was closed, the pump was turned off and the oven was switched on. During temperature rising (up to 160 ºC), the switching valve had to be opened at short intervals to prevent the pressure from surpassing the working value. Once the selected temperature and pressure were

reached, static extraction was performed for a preset time (5 min). Finally, the oven was switched off, the chamber was cooled below the boiling point of ethanol and the switching valve and the restrictor valve were switched to enable clean extractant to flow through the cell and flush out the extract (collected volume, 10 mL).

2.5. Determination of total phenols by the F-C method.

The total amount of phenolic compounds was quantified by the F-C method using gallic acid as standard. With this purpose, a calibration curve was run using solutions of 1, 5, 10, 25 and 50 mg/L of this acid ($y = 0.0382x + 0.0296$, $R^2 = 0.9963$). A 0.5 mL aliquot of extract, 10 mL of distilled water, 1 mL of F-C reagent and 3 mL of Na_2CO_3 (20%, w/v) were mixed, made to 25 mL with distilled water and heated at 50 °C for 5 min. After heating, the samples were kept at room temperature for 30 min and, finally, the absorbance was monitored at 765 nm against a blank solution containing distilled water instead of extract. The concentration of phenolic compounds thus obtained was multiplied by the dilution factor of the extract volume and divided by the amount of olive pomace used. The results were expressed as equivalent to micrograms of gallic acid per mL of olive pomace extract (µg GAE/mL).[24]

2.6. LC-QqTOF MS/MS analysis

An injection volume of 5 µL and a flow rate of 0.8 mL/min were used. The mobile phases were 0.1% (v/v) formic acid aqueous solution (phase A) and pure acetonitrile (phase B). Separation of the analytes was performed on an Inertsil ODS-2 C18 column (250 mm × 4.6 mm i.d., 5 µm particle, Análisis Vínicos, Tomelloso, Ciudad Real, Spain) at 25 ºC. The gradient method was as follows: 96% A for 4 min, from 96% to 0% A in 36 min, 0% A for 3 min. The dual ESI source operated in negative ionization using the following conditions: nebulizer gas at 35 psi and drying gas flow rate and temperature at 10 L/min and 325 ºC, respectively. The capillary voltage was set at 3500 V, while the fragmentor, skimmer and octapole voltages were fixed at 175, 65 and 750 V, respectively.

The data were acquired in centroid mode in the extended dynamic range (2 GHz). Full scan was carried out at 1 spectrum per s within the *m/z* range 100–1000 with subsequent activation of the three most intense precursor ions (only allowed charge, single or double) by MS/MS using a collision energy of 20 eV at 1 spectrum per s within the *m/z* range 100–1000. An active exclusion window was programmed after one spectrum and released after 0.75 min to avoid repetitive fragmentation of the most intense precursor ions and, in this way, increase the detection coverage. Before the experiments, the instrument reported mass detection resolution of 25000 FWHM (Full Width at Half Maximum) at *m/z* 112.985587, and 45000 FWHM at *m/z* 966.000725. To assure the desired mass accuracy of recorded ions, continuous internal calibration was performed during analyses with the use of signals at *m/z* 119.0362 (proton abstracted purine) and *m/z* 966.0007 –formate adduct of hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazine.

2.7. Data processing

Raw data files were converted to mzData using MassHunter and processed in R statistical language (version 2.15.0, http://www.r-project.org/) using the open-free XCMS (version 1.24, http://metlin.scripps.edu/xcms/index.php) in an automated and unbiased way. This software allows peak finding, alignment and peak picking of mass spectral features resulting in a list of *m/z* precursor ions per retention time with the three MS/MS spectra obtained in each cycle. Parameters for XCMS were taken from the protocol described by Smith *et al.*[25] using the centWave peak finder method designed for high mass accuracy data.[26] The extraction of molecular features was carried out using a prefilter to take at least two ions in each cycle with intensity above 5000. Briefly, peaks were deconvoluted and aligned across samples using a signal-to-noise threshold of 10, a maximum tolerated *m/z* deviation of 5 ppm and a peak width of 10–60 s. The aligned output consisted of accurate mass, retention time and intensity of each peak. The file from this treatment was created in comma separated value files (.csv) for each sample and

exported into the Mass Profiler Professional (MPP) software package (version 2.0, Agilent Technologies, Santa Clara, CA, USA) for further processing. Compound identification was performed using the METLIN Metabolite and MS/MS (http://metlin.scripps.edu/) and PlantCyc Databases (http://plantcyc.org/). The allowed negative precursor ions for identification were formate adducts and deprotonated ions. Dehydratation neutral losses were also allowed.

3. Results and discussion

3.1. Superheated liquid extraction of olive phenols from olive pomace

The use of SHLE for extraction of natural compounds from raw material of plant origin is widely known. SHLE has proved its efficiency for the extraction of natural compounds from different types of vegetables such as leaves, fruits, roots or stems.[27] In SHLE, the duality of extraction temperature and extractant composition plays a crucial role in succeeding in the leaching of target compounds. The extractant composition for isolation of phenolic compounds from vegetal matrices is usually an ethanol-deionized water mixture at ratios typically ranging from 50:50 to 80:20 (v/v) ethanol-water. Japón-Luján *et al.* used 80:20 ethanol-water for extraction of a panel of representative phenols in olive pomace.[18] In the present study, this extraction protocol was used for a characterization study. Thus, the temperature was set at 160 ºC, while the extraction time was programmed for 5 min. Under these conditions, concentrations of phenolic compounds above 500 µg GAE/mL were obtained. Figure 1 illustrates the base peak chromatogram (BPC) provided by the extract. As can be seen, phenolic compounds are mainly eluted in the time window from 2 to 28 min, which fits approximately with a concentration of acetonitrile in the chromatographic gradient of 78%. On the other hand, few peaks are detected in the last part of the chromatogram corresponding to non-polar compounds such as α-tocopherol (*m/z* 429.3759), which eluted at 42.3 min. Once the potential of olive pomace for providing extracts rich in phenolic compounds was defined, characterization of representative families of these compounds was carried out

by tandem mass spectrometry in high accuracy mode. Identification and confirmatory detection was supported on mass accuracy of precursor ion and representative product ions, the structure of which was elucidated, but also on isotopic distribution of signals detected in full scan mode. Table 1 lists the identified compounds organized by families (information about theoretical and experimental *m/z* values as well as characteristic fragment ions are also included).

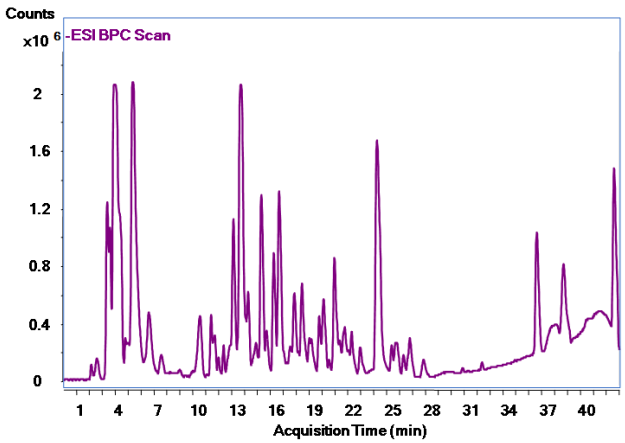


Figure 1. Base peak chromatogram obtained by analysis of an olive pomace extract under the optimum working conditions.

Table 1. Compounds in extracts from olive pomace identified by LC–QqTOF MS/MS analysis

	Formula	*m/z* exp	*m/z* theo	Error (ppm)	t_R	Fragments
Hydroxytyrosol and Tyrosol Derivatives						
Hydroxytyrosol	$C_8H_{10}O_3$	153.0560	153.0557	-1.9601	13.20	123.0450
						105.0339
Hydroxytyrosol glucoside	$C_{14}H_{20}O_8$	315.1069	315.1085	5.2046	13.18	123.0446
						153.0570
Hydroxytyrosol diglucoside	$C_{20}H_{28}O_{13}$	475.1429	475.1457	5.9140	13.28	153.0567
						245.0682

Hydroxytyrosol rhamnoside	$C_{20}H_{34}O_{13}$	481.1943	481.1927	-3.4498	13.35	153.0565 265.3271
Tyrosol	$C_8H_{10}O_2$	137.0608	137.0608	0.0000	17.27	119.0505 111.0098
Tyrosol glucoside	$C_{14}H_{20}O_7$	299.1128	299.1136	2.7749	14.56	119.0505 137.0603
Iridoids Precursors						
Loganin	$C_{17}H_{26}O_{10}$	389.1468	389.1453	-3.9060	13.67	151.0776 113.0248
Loganin glucoside	$C_{23}H_{38}O_{16}$	569.2113	569.2087	-4.4623	13.42	389.1513 313.1332
Loganic acid	$C_{16}H_{24}O_{10}$	375.1322	375.1297	-6.7976	13.90	151.0776 113.0248
Loganic acid glucoside	$C_{22}H_{34}O_{15}$	537.1857	537.1860	0.5957	12.00	375.1432 179.0575
7-Deoxyloganic acid	$C_{16}H_{24}O_9$	359.1368	359.1348	-5.6246	15.04	-
Secologanic acid	$C_{18}H_{26}O_{10}$	401.1477	401.1453	-6.0577	16.42	-
Secologanoside	$C_{16}H_{22}O_{11}$	389.1122	389.1089	-8.4804	15.412	-
Secologanin	$C_{17}H_{24}O_{10}$	387.1268	387.1297	7.4910	13.10	-
Eleanolic acid	$C_{11}H_{14}O_6$	241.0772	241.0790	7.6324	17.53	-
Oleoside	$C_{16}H_{22}O_{11}$	389.1122	389.1089	-8.3524	15.46	121.0660 101.0243 209.0477 345.1253
Oleoside glucoside	$C_{22}H_{32}O_{16}$	551.1643	551.1618	-4.6084	14.89	551.1707 507.1805 209.0480 239.0354
Oleoside diglucoside	$C_{28}H_{42}O_{21}$	713.2205	713.2146	-8.2724	13.50	551.1479 507.1585 713.2029 161.0259
Oleoside riboside	$C_{20}H_{26}O_{15}$	505.1227	505.1195	-6.2955	16.57	389.1166

						505.1298
						345.1256
						121.0663
Oleoside-11-methylester	$C_{17}H_{24}O_{11}$	403.1274	403.1246	-6.9457	16.72	101.0244
						119.0350
						223.0626
						179.0707
Oleoside dimethylester	$C_{18}H_{26}O_{11}$	417.1433	417.1402	7.4316	17.08	-
Secoiridoids and Derivatives						
Oleuropein	$C_{25}H_{32}O_{13}$	539.1776	539.1770	-1.0757	19.95	377.1306
						307.0874
						275.0962
10-Hydroxy-oleuropein	$C_{25}H_{32}O_{14}$	555.1701	555.1719	3.2783	18.31	537.1647
						376.1114
Oleuropein aglycone	$C_{19}H_{22}O_{8}$	377.1440	377.1453	3.4469	16.85	123.0446
						255.0869
Verbascoside	$C_{29}H_{36}O_{15}$	623.1973	623.1981	1.3158	18.48	461.1744
						161.0254
3,4-DHPEA-EDA	$C_{17}H_{20}O_{6}$	319.1207	319.1187	-6.1733	22.54	123.0446
						139.0769
***p*-HPEA-EDA**	$C_{17}H_{20}O_{5}$	303.1235	303.1238	0.9897	23.32	-
3,4-DHPEA-EA	$C_{19}H_{22}O_{8}$	377.1266	377.1242	-6.3905	23.95	123.0446
						255.0869
***p*-HPEA-EA**	$C_{19}H_{22}O_{7}$	361.1300	361.1293	-1.9384	20.29	-
Oleuropein derivative 1	$C_{25}H_{36}O_{13}$	543.2107	543.2083	-4.4182	19.38	377.1522
Oleuropein derivative 2	$C_{25}H_{36}O_{12}$	527.2160	527.2134	-4.9316	20.61	377.1481
Flavonoids						
Rutin	$C_{27}H_{30}O_{16}$	609.1487	609.1461	-4.1698	18.54	301.0391
Apigenin	$C_{15}H_{10}O_{5}$	269.0429	269.0455	9.66049	18.000	-
Luteolin	$C_{5}H_{10}O_{6}$	285.0420	285.0405	-5.2624	23.92	-
Apigenin glucoside	$C_{21}H_{20}O_{10}$	431.0999	431.0984	-3.5259	20.35	265.0450
Luteolin glucoside	$C_{21}H_{20}O_{11}$	447.0970	447.0933	-8.2533	19.27	285.0423

Taxifolin	$C_{15}H_{12}O_7$	303.0527	303.0513	-4.5867	20.68	-
Diosmetin	$C_{16}H_{12}O_6$	299.0584	299.0561	-7.4902	26.53	-
Quercetin	$C_{15}H_{10}O_7$	301.0377	301.0354	-7.6735	24.58	-
Lignans						
Pinoresinol	$C_{20}H_{22}O_6$	357.1345	357.1344	-0.3080	19.47	-
Hydroxypinoresinol	$C_{20}H_{23}O_7$	374.1359	374.1371	3.2074	21.60	-
Acetoxypinoresinol	$C_{22}H_{24}O_8$	415.1400	415.1398	-0.3854	20.10	-
Phenolic Acids						
Shikimic acid	$C_7H_{10}O_5$	173.0486	173.0495	5.2008	14.69	-
Phenylalanine	$C_9H_{11}NO_2$	164.0721	164.0717	-2.4380	14.88	-
Cinnamic acid	$C_9H_8O_2$	147.0452	147.0452	0.0000	13.58	147.0459
p-Coumaric acid	$C_9H_8O_3$	163.0403	163.0401	-1.0427	15.97	119.0503
Caffeic acid	$C_9H_8O_4$	179.0352	179.0350	-1.1171	17.68	135.0449
Protocatechuic acid	$C_7H_6O_4$	153.0193	153.0193	0.0000	14.40	109.0295
Vanillic acid	$C_8H_8O_4$	167.0353	167.0350	-1.9756	13.43	108.0206 123.0446
Ferulic acid	$C_{10}H_{10}O_4$	193.0514	193.0506	-4.1440	16.40	134.0377
Gallic acid	$C_7H_6O_5$	169.0150	169.0142	-4.7333	15.49	125.0246

3.2. Hydroxytyrosol and tyrosol derivatives

Hydroxytyrosol and tyrosol have been identified as two of the major simple phenols in VOO, which is considered the main source of both compounds in the diet. Both compounds are present in VOO mainly as free forms, but also condensed as secoiridoid derivatives.[28] The concentrations of hydroxytyrosol and tyrosol in VOO are usually below 15 μg/g,[29] while in olive pomace these antioxidants have been determined in concentrations around 106 and 18.5 μg/g, respectively.[30] The concentration of phenols in VOO as well as in olive pomace depends on genetic factors such as the cultivar variety, but also on other factors such as olive fruit maturation and agropedoclimatic conditions.[31,32] Glucoside and

acetate derivatives have also been found in other Olea europaea derived materials such as leaves and olive pomace waste after the refining process.[33]

In the case of the olive pomace used in this research (taken directly from the production line after decantation), the extracted ion chromatogram (EIC) for hydroxytyrosol by monitoring the *m/z* ratio at 153.0560 revealed the presence of several derivatives of this metabolite.[34] The peaks at 13.18 and 13.20 min in Figure 2A were clearly identified by MS/MS fragmentation as hydroxytyrosol glucoside (Figure 2B) and free hydroxytyrosol (Figure 2C). Thus, hydroxytyrosol glucoside detected by fragmentation of the precursor ion at *m/z* 315.1069 generated two representative fragments: one of them corresponding to free hydroxytyrosol at *m/z* 153.0570 by cleavage of the hexose unit, and a second one at *m/z* 123.0450, which was assigned to the structure in Figure 2B resulting from the loss of the CH_2OH and CHO groups from the hydroxytyrosol original structure. These two fragments are characteristic for identification of any conjugated hydroxytyrosol present in any olive tree material. From a quantitative point of view, the chromatographic signal attributed to hydroxytyrosol glucoside was more intense than that assigned to free hydroxytyrosol.

Concerning the chromatographic peaks eluted at 13.28 and13.35 min, precursor ions for hydroxytyrosol diglucoside and diramnhoside were detected at *m/z* 475.1429 and 481.1943, respectively. Identifications were confirmed by the corresponding MS/MS spectra, as illustrated in Supplementary Figure 1, which shows the cleavage of the sugar units.

Tyrosol was detected in free form, but also as glucoside derivative with *m/z* 137.0608 and 299.1128 (see Figure 2A). In contrast to hydroxytyrosol, the free form of tyrosol was at the same concentration level as its glucoside derivative. The two MS/MS spectra revealed product ions representative of tyrosol derivatives such as that at *m/z* 119.0505 and the representative neutral loss corresponding to the cleavage of the glucoside unit (Supplementary Figure 2).

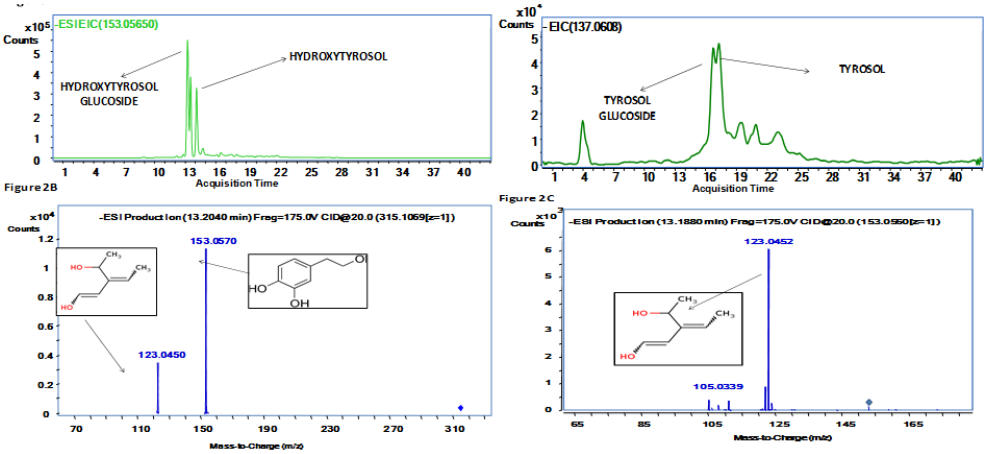


Figure 2. (A) Extracted ion chromatograms from hydroxytyrosol and tyrosol. (B) MS/MS spectrum for hydroxytyrosol glucoside. (C) MS/MS spectrum for hydroxytyrosol.

3.3. Iridoid precursors

The term iridoid is used, in most cases, to name a wide group of monoterpenes as well as glucoside derivatives whose structure may be considered as derived from iridane (*cis*-2-oxabicycle-[4.3.0]-nonane). The secoiridoid compounds, which play a crucial role in the family of olive oil antioxidants, derive from iridoids by opening of the pentacyclic ring. The precursor for biosynthesis of iridoids is mevalonic acid, which leads to different structures by an onwards cyclization step from which the iridane skeleton is formed. Taking into account the relevance of this pathway in the synthesis of one of the major phenolic families, as shows Supplementary Figure 3, the precursors involved in the synthesis of secoridoids may be present in the extracts from olive pomace. Two of these precursors, loganin and loganic acid, were detected in the SHL extracts from olive pomace together with their glucoside derivatives. Figure 3A shows the EIC of loganin obtained by monitoring its precursor ion at *m/z* 389.1468, which revealed the presence of loganin eluting at 13.67 min and a derivative form whose precursor ion was detected at *m/z* 569.2113. This derivative, eluting at 13.42 min, was attributed to loganin glucoside. Tentative identification supported on mass accuracy and isotopic distribution was

confirmed by tandem mass spectrometry. The EICs corresponding to loganic acid and its glucoside, at *m/z* 375.1322 and 537.1857, respectively, can be clearly differentiated at 13.90 and 12.00 min.

Product ions obtained at *m/z* 151.0776 and 113.0248 were characteristic fragments for identification of loganin and loganic acid (Figures 3B and 3C), respectively. These fragments correspond to the loss of $C_9H_{11}O_2$ and $C_5H_8O_3$ for loganin and loganic acid, respectively. The glucoside derivatives also generated characteristic fragments at *m/z* 389.1513 and 375.1432 for loganin and loganic acid conjugates, respectively. These two fragments were obtained by cleavage of the glucose unit (162 Da) (Supplementary Figure 4). It is worth mentioning that both loganin and loganic acid were preferentially detected as free forms.

Apart from these precursors, other compounds involved in the different pathways leading to the synthesis of secoiridoids were also detected. These compounds were 7-deoxyloganic acid, secologanic acid, secologanoside and secologanin, detected at *m/z* 359.1368, 401.1477, 555.1795 and 387.1268, respectively. Secologanin leads to the synthesis of oleoside-11-methyl ester, which is the immediate precursor of secoiridoids. Oleoside-11-methyl ester is formed by glucosidation of eleanolic acid. Both metabolites were clearly identified in the extracts from olive pomace at retention times 16.72 and 17.53, respectively, with precursor ions at *m/z* 403.1274 and 241.0772, respectively. Identification was confirmed for both compounds by MS/MS spectra, as shows Figures 3D and 3E. In addition to these identifications, oleoside and its monoglucoside and diglucoside derivatives were confirmed by MS/MS spectra, which are illustrated in Supplementary Figure 5. Oleoside was detected by a precursor ion at *m/z* 389.1122 eluting at 15.4610 min. The MS/MS spectrum presented characteristic fragments at *m/z* 345.1253 and 209.0477. The former corresponds to the loss of 44 Da, which can be justified by decarboxylation; while the other fragment can be attributed to the loss of a hexose (180 Da).

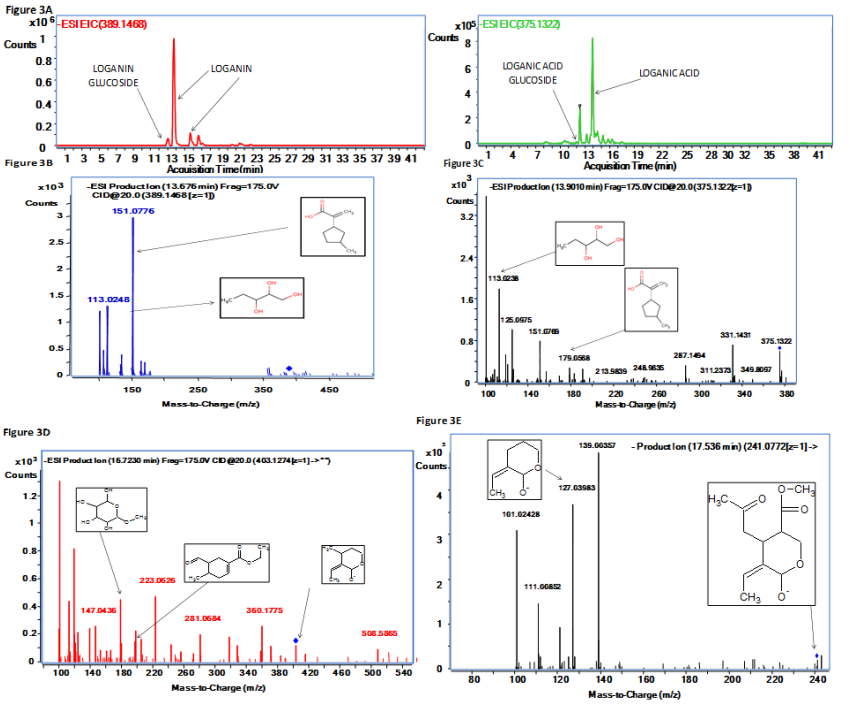


Figure 3. (A) Extracted ion chromatograms for loganin and loganic acid. (B) MS/MS spectrum for loganin. (C) MS/MS spectrum for loganic acid. (D) MS/MS spectrum for oleoside 11-methylester. (E) MS/MS spectrum for eleanolic acid.

3.4. Secoiridoids and derivatives

Secoiridoids are a type of monoterpenes derived from geraniol. The main difference as compared to iridoids is the hydrolysis of the cyclopentanopyran ring.[35] Phenolic compounds classified as secoiridoids are characterized by the presence of either eleanolic acid or eleanolic acid derivatives in their molecular structure.[36] Oleuropein, demethyloleuropein, ligstroside and verbascoside are the most abundant secoiridoid glucosides in olive fruit.[36,37] Several aglycon derivatives of oleuropein (3,4-DHPEA-EA) and ligstroside (*p*-HPEA-EA) as well as their decarboxymethylated aglycone forms (3,4-DHPEA-EDA for oleuropein and *p*-HPEA-EDA for ligstroside) have also been found in olive pulp.[38] The

concentration of these phenols in VOO and olive fruit strongly depends on the season period, but also on agronomical factors.[38]

Among secoiridoids detected in extracts from olive mill waste, oleuropein was one of the most concentrated. This especially well-known compound has been described to posses interesting functional properties including antioxidant, anti-inflammatory, anti-atherogenic, anti-cancer and antimicrobial activities, among others.[39] The precursor ion for oleuropein was detected at *m/z* 539.1776 (Figure 4A). As can be seen in Figure 4, the fragmentation of this ion generated product ions with *m/z* 377.1440 and 275.0962, characteristic of oleuropein structure. The fragmentation scheme of oleuropein enabled identification of these representative fragments by cleavage of the hexose unit (*m/z* 377.1440), while the other ion was assigned to a further fragmentation product of the oleuropein aglycone residue. Oleuropein aglycone, detected at 16.85 min (Figure 4B), was fragmented to generate two intense product ions at *m/z* 153.0924 (corresponding to cleaved hydroxytyrosol) and *m/z* 197.0829 (fit to the rest of the molecular structure after decarboxylation). The precursor ion with *m/z* 319.1207 was associated to 3,4-DHPEA-EDA. The presence of the fragments with *m/z* 123.0826 and 139.0758, corresponding to a fragment of hydroxytyrosol and the dialdehydic subunit released after hydrolysis of hydroxytyrosol, confirmed this identification.

The BPC of the olive pomace extract showed other predominant chromatographic peak ascribed to *m/z* 555.1701 and eluted at 18.31 min (Figure 4C). The main product ion in its MS/MS spectrum was obtained at *m/z* 537.1647 by loss of 18 Da, suggesting the structure of a hydroxyl derivative of oleuropein, which is easily dehydrated by MS activation with nitrogen gas. Its aglycone derivative was detected at *m/z* 376.1114, whose precursor ion reported common product ions to the conjugated secoiridoid, which should be 10-hydroxy-oleuropein.

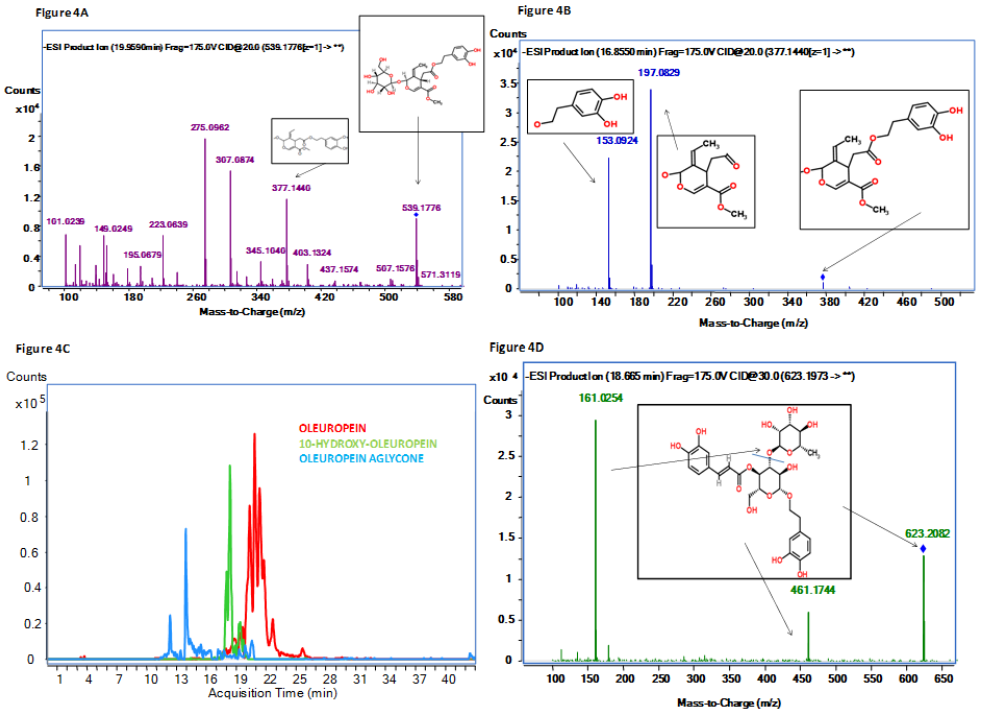


Figure 4. (A) Extracted ion chromatograms from oleuropein and oleuropein aglycone. (B) MS/MS spectrum from oleuropein. (C) MS/MS spectrum from oleuropein aglycone. (D) MS/MS spectrum from verbascoside.

Two other oleuropein derivatives were detected at *m/z* 543.2066 and 527.2120 eluting at 19.18 and 20.50 min, respectively (Supplementary Figure 6). The MS/MS spectra for these two compounds contained ions common to those provided by oleuropein, particularly the intense ions of oleuropein aglycon at *m/z* 377.1522, 151.0771 and 101.0240.

The precursor ion at *m/z* 623.1973, ascribed to verbascoside, was also detected in the olive pomace extracts (Figure 4D). Representative product ions observed in the MS/MS spectrum were in agreement with the fragmentation of this secoiridoid. Thus, the product ion at *m/z* 461.1744 corresponds to the loss of 162 Da, and the *m/z* 161.0254 corroborated the detection of the glucose cleavage.

3.5. Phenols produced by the general phenylpropanoid metabolism

The presence of other simple phenols such as coumaric acid isomers, ferulic acid, vanillic acid or vanillin in VOO has been described.[2] These compounds are synthesized in the general phenylpropanoid metabolism that involves the production of 4-coumaric acid from phenylalanine supplied by the shikimate pathway (Supplementary Figure 7). The analysis reported here revealed the presence of shikimic acid, phenylalanine, cinnamic acid, *p*-coumaric acid, caffeic acid, protocatechuic acid, vanillic acid, ferulic acid and gallic acid, as shows Table 1. For example, the precursor ion at *m/z* 163.0401 was associated to *p*-coumaric acid. The MS/MS spectrum for this precursor ion presents an intense peak at *m/z* 119.0502 that corresponds to the loss of a CHO_2 group. This is the same fragmentation occurring to caffeic acid (*m/z* 179.0352) that results in a great peak at *m/z* 135.0448. The *m/z* 153.0193 was identified as protocatechuic acid, which has a high peak resulting for the loss of a CHO_2 group at *m/z* 109.0295. Vanillic acid (*m/z* 167.0353) presents two high peaks at *m/z* 123.0778 (loss of the CHO_2 group) and *m/z* 108.0205 (loss of an additional OH group). Ferulic acid (*m/z* 193.0514) was also identified thanks to the fragment ion at *m/z* 134.0377, a result of the loss of the C_2H_5O and OH groups. The last one is gallic acid (*m/z* 169.0150) whose MS/MS spectrum presents an ion fragment at *m/z* 125.0246 formed by loss of a CHO_2 group. Supplementary Figure 8 shows the MS/MS spectra for the identified metabolites.

Lignans are phenylpropane dimers also biosynthesized via the phenylpropanoid pathway, in which pinoresinol lariciresinol reductase catalyzes the last steps of lignin production. The structure of lignans also endows these compounds with antioxidant properties. Three lignan metabolites (*viz.* pinoresinol, hydroxypinoresinol and acetoxypinoresinol, eluted at 19, 21 and 20 min, respectively) were also detected in the extracts from olive pomace.

Supplementary Figure 9 shows the MS/MS spectra for pinoresinol and acetoxypinoresinol.

Flavonoids, involved in the phenylpropanoid pathway, are abundant in fruits and vegetables. The dietary intake of this phenolic family is considered important for preventing a wide variety of diseases that involve free radical-mediated damage.[40] Most of the pharmacological effects of flavonoids seem to be associated with their potency as antioxidants.[40] The most relevant flavonoids identified in this research have been rutin, apigenin, luteolin, apigenin glucoside, luteolin glucoside, taxifolin, diosmetin and quercetin. All tandem mass spectra of flavonoides revealed a similar fragmentation scheme. Rutin was detected at *m/z* 609.1487, and its MS/MS spectrum presents a peak at *m/z* 301 due to the loss of glucoside and rhamnoside groups. Apigenin glucoside, detected by fragmentation of the precursor ion at *m/z* 431.0999, generated one representative fragment at *m/z* 269.0424 by cleavage of the hexose unit. Similarly, luteolin glucoside appeared at *m/z* 447.0970 and generated a high peak at *m/z* 285.0422 due to the loss of a hexose unit, as shown in Figure 5; while the fragmentation of aglycone forms of apigenin and luteolin can be deduced from the spectra illustrated in Supplementary Figure 10.

Attending to the profile and content of phenolic compounds in olive pomace, this residue is an exploitable source for isolation of phenols. After this step, the residue could be degraded more easily since the antibacterial properties of phenols have been minimized. This study complements previous studies on olive leaves and other mill residues.

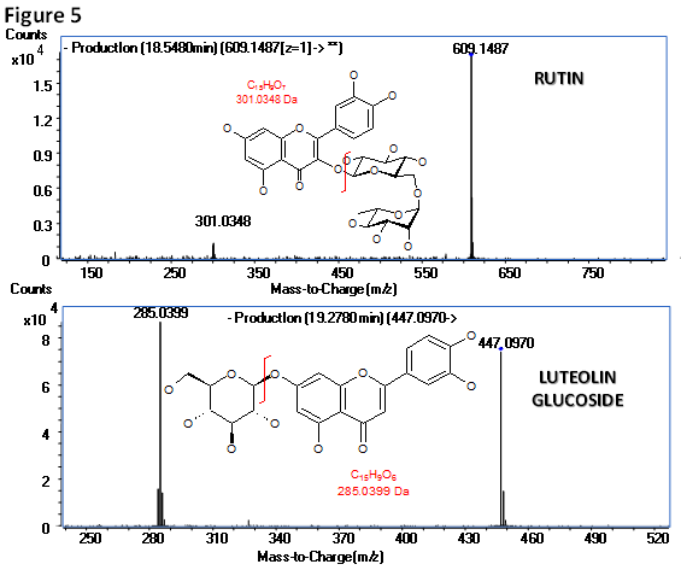


Figure 5. MS/MS spectra for rutin and apigenin-7-glucoside.

Acknowledgement

Núñez de Prado, C.B. is thanked for providing the olive pomace for this study.

References

1. Autority, E. F. S. Scientific Opinion on the evaluation of the substances currently on the list in the annex to Commission Directive 96/3/EC as acceptable previous cargoes for edible fats and oils – Part II of III. EFSA Journal 2012, 5, 150.

2. Lozano-Sánchez, J.; Giambanelli, E.; Quirantes-Pine, R.; Cerretani, L.; Bendini, A.; Segura-Carretero, A.; Fernández-Gutiérrez, A. Wastes generated during the storage of extra virgin olive oil as a natural source of phenolic compounds. J. Agric. Food Chem. 2011, 21, 11491-11500.

3. Sampedro, I.; Aranda, E.; Martín, J.; García-Garrido, J. M.; García-Romera, I.; Ocampo, J. A. Saprobic fungi decrease plant toxicity caused by olive mill residues. Appl. Soil Ecol. 2004, 2, 149-156.

4. Lesage-Meessen, L.; Navarro, D.; Maunier, S.; Sigoillot, J. C.; Lorquin, J.; Delattre, M.; Simon, J. L.; Asther, M.; Labat, M. Simple phenolic content in olive oil residues as a function of extraction systems. Food Chem. 2001, 4, 501-507.

5. Priego-Capote, F.; Ruiz-Jiménez, J.; Luque de Castro, M. D. Fast separation and determination of phenolic compounds by capillary electrophoresis–diode array detection: Application to the characterisation of alperujo after ultrasound-assisted extraction. J. Chromatogr. A 2004, 1-2, 239-246.

6. Celma, A. R.; Rojas, S.; López, F.; Montero, I.; Miranda, T. Thin-layer drying behaviour of sludge of olive oil extraction. J. Food Eng. 2007, 4, 1261-1271.

7. Hinneburg, I.; Damien Dorman, H. J.; Hiltunen, R. Antioxidant activities of extracts from selected culinary herbs and spices. Food Chem. 2006, 1, 122-129.

8. Torres de Pinedo, A.; Peñalver, P.; Morales, J. C. Synthesis and evaluation of new phenolic-based antioxidants: structure–activity relationship. Food Chem. 2007, 1, 55-61.

9. Joubert, E.; de Beer, D. Phenolic content and antioxidant activity of rooibos food ingredient extracts. J. Food Compos. Anal. 2012, 1, 45-51.

10. Owen, R. W.; Giacosa, A.; Hull, W. E.; Haubner, R.; Spiegelhalder, B.; Bartsch, H. The antioxidant/anticancer potential of phenolic compounds isolated from olive oil. Eur. J. Cancer 2000, 10, 1235-1247.

11. European Food Safety Authority. EFSA J., 2011, 4, 2033-2058.

12. Emilia, J. E.; Uwe, W.; Hannelore, D.; Planas, J. M., Chapter 144 - Cancer Chemopreventive Activity of Hydroxytyrosol: A Natural Antioxidant from Olives and Olive Oil. In Olives and Olive Oil in Health and Disease Prevention, Preedy, V. R.; Ross, R.; W., Eds. Academic Press: San Diego, 2010; pp 1295-1300.

13. Fernández-Pachón, M. S.; Villaño, D.; Troncoso, A. M.; García-Parrilla, M. C. Determination of the phenolic composition of sherry and table white wines by liquid chromatography and their relation with antioxidant activity. Anal. Chim. Acta 2006, 1-2, 101-108.

14. Fernández-Orozco, R.; Roca, M.; Gandul-Rojas, B.; Gallardo-Guerrero, L. DPPH-scavenging capacity of chloroplastic pigments and phenolic compounds of olive fruits (cv. Arbequina) during ripening. J. Food Compos. Anal. 2011, 6, 858-864.

15. Pohjamo, S. P.; Hemming, J. E.; Willför, S. M.; Reunanen, M. H. T.; Holmbom, B. R. Phenolic extractives in Salix caprea wood and knots. Phytochemistry 2003, 2, 165-169.

16. Benavente-García, O.; Castillo, J.; Lorente, J.; Ortuño, A.; Del Río, J. A. Antioxidant activity of phenolics extracted from Olea europaea L. leaves. Food Chem. 2000, 4, 457-462.

17. Japón-Luján, R.; Luque-Rodríguez, J. M.; Luque de Castro, M. D. Dynamic ultrasound-assisted extraction of oleuropein and related biophenols from olive leaves. J. Chromatogr. A 2006, 1, 76-82.

18. Japón-Luján, R.; Luque-Rodríguez, J.; Luque de Castro, M. Multivariate optimisation of the microwave-assisted extraction of oleuropein and related biophenols from olive leaves. Anal. Bioanal. Chem. 2006, 4, 753-759.

19. Ansari, M. A.; Abdul, H. M.; Joshi, G.; Opii, W. O.; Butterfield, D. A. Protective effect of quercetin in primary neurons against Aβ(1–42): relevance to Alzheimer's disease. J. Nutr. Biochem. 2009, 4, 269-275.

20. Oh, H.; Kim, D.-H.; Cho, J.-H.; Kim, Y.-C. Hepatoprotective and free radical scavenging activities of phenolic petrosins and flavonoids isolated from Equisetum arvense. J. Ethnopharmacol. 2004, 2-3, 421-424.

21. Herrero, M.; Temirzoda, T. N.; Segura-Carretero, A.; Quirantes, R.; Plaza, M.; Ibañez, E. New possibilities for the valorization of olive oil by-products. J. Chromatogr. A 2011, 1218, 42, 7511-7520.

22. Gortzi, O.; Lalas, S.; Chatzilazarou, A.; Katsoyannos, E.; Papaconstandinou, S.; Dourtoglou, E. Recovery of natural antioxidants from olive mill wastewater using Genapol-X080. J. Am. Oil Chem. Soc. 2008, 2, 133-140.

23. Luque-Rodríguez, J. M.; Luque de Castro, M. D.; Pérez-Juan, P. Dynamic superheated liquid extraction of anthocyanins and other phenolics from red grape skins of winemaking residues. Bioresource Technol. 2007, 14, 2705-2713.

24. Luque-Rodríguez, J. M.; Pérez-Juan, P.; Luque de Castro, M. D. Extraction of polyphenols from vine shoots of Vitis vinifera by superheated ethanol−water mixtures. J. Agric. Food Chem. 2006, 23, 8775-8781.

25. Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal. Chem. 2006, 3, 779-787.

26. Tautenhahn, R.; Bottcher, C.; Neumann, S. Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 2008, 1, 504.

27. Quideau, S. Methods in Polyphenol Analysis. Edited by Celestino Santos-Buelga and Gary Williamson. Angew. Chem. Int. Edit. 2004, 4, 393-394.

28. Vázquez-Martín, A.; Fernández-Arroyo, S.; Cufi, S.; Oliveras-Ferraros, C.; Lozano-Sánchez, J.; Vellón, L.; Micol, V.; Joven, J.; Segura-Carretero, A.; Menéndez, J. A. Phenolic secoiridoids in extra virgin olive oil impede fibrogenic and oncogenic epithelial-to-mesenchymal transition: extra virgin olive oil as a source of novel antiaging phytochemicals. Rejuvenation Res. 2012, 1, 3-21.

29. El Riachy, M.; Priego-Capote, F.; Rallo, L.; Luque de Castro, M. D.; León, L. Phenolic profile of virgin olive oil from advanced breeding selections. Span. J. Agric. Res. 2012, 2, 443-453.

30. Sánchez de Medina, V.; Priego-Capote, F.; Jiménez-Ot, C.; Luque de Castro, M. D. Quality and stability of edible oils enriched with hydrophilic antioxidants from the olive tree: the role of enrichment extracts and lipid composition. J. Agric. Food Chem. 2011, 21, 11432-11441.

31. González-Rodríguez, J.; Pérez-Juan, P.; Luque de Castro, M. D. Superheated liquids for extraction of solid residues from winemaking processes. Anal. Bioanal. Chem. 2003, 7-8, 1190-1195.

32. González-Rodríguez, J.; Pérez-Juan, P.; Luque de Castro, M. D. Use of superheated liquids for the extraction of non-volatile compounds from wood: liquid chromatography studies. J. Chromatogr. A 2004, 1-2, 3-9.

33. Lama-Muñoz, A.; Rodríguez-Gutiérrez, G.; Rubio-Senent, F. t.; Gómez-Carretero, A.; Fernández-Bolaños, J. New hydrothermal treatment of alperujo enhances the content of bioactive minor components in crude pomace olive oil. J. Agric. Food Chem. 2011, 4, 1115-1123.

34. Brenes, M.; García, A.; García, P.; Ríos, J. J.; Garrido, A. Phenolic compounds in Spanish olive oils. J. Agric. Food Chem. 1999, 9, 3535-3540.

35. Pérez, J. A.; Hernández, J. M.; Trujillo, J. M.; López, H., Iridoids and Secoiridoids from Oleaceae. In Studies in Natural Products Chemistry, Atta ur, R., Ed. Elsevier: 2005; 32, Part L, pp 303-363.

36. Garrido Fernández, A.; Fernández Díez, M. J.; Adams, M. R. Table Olives. Chapman and Hall: 2-6 Boundary Row, London SE1 8HN, UK, 1997.

37. Gil, F. S.; Busconi, M.; Da Camara Machado, A.; Fogher, C. Development and characterization of microsatellite loci from Olea europaea. Mol. Ecol. Notes 2006, 4, 1275-1277.

38. Ryan, D.; Prenzler, P. D.; Lavee, S.; Antolovich, M.; Robards, K. Quantitative changes in phenolic content during physiological development of the olive (Olea europaea) cultivar Hardy's Mammoth. J. Agric. Food Chem. 2003, 9, 2532-2538.

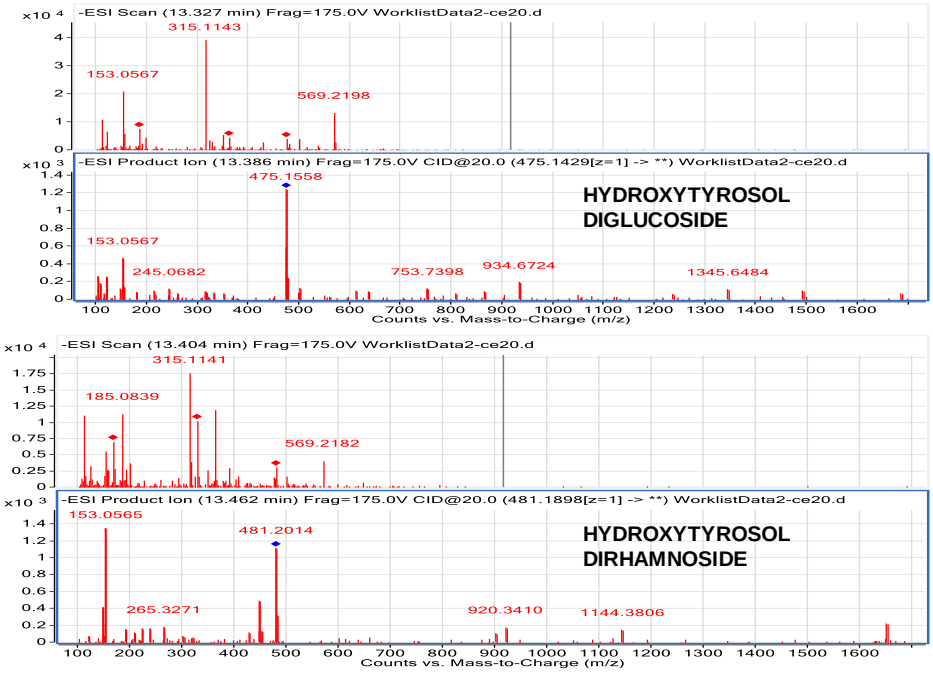
39. Omar, S. H. Oleuropein in olive and its pharmacological effects. Scientia Sci. Pharmaceutica Pharmac. 2010, 2, 133-154.

40. Sugihara, N.; Arakawa, T.; Ohnishi, M.; Furuno, K. Anti- and pro-oxidative effects of flavonoids on metal-induced lipid hydroperoxide-dependent lipid peroxidation in cultured hepatocytes loaded with α-linolenic acid. Free Radical Bio. Med. 1999, 11-12, 1313-1323.

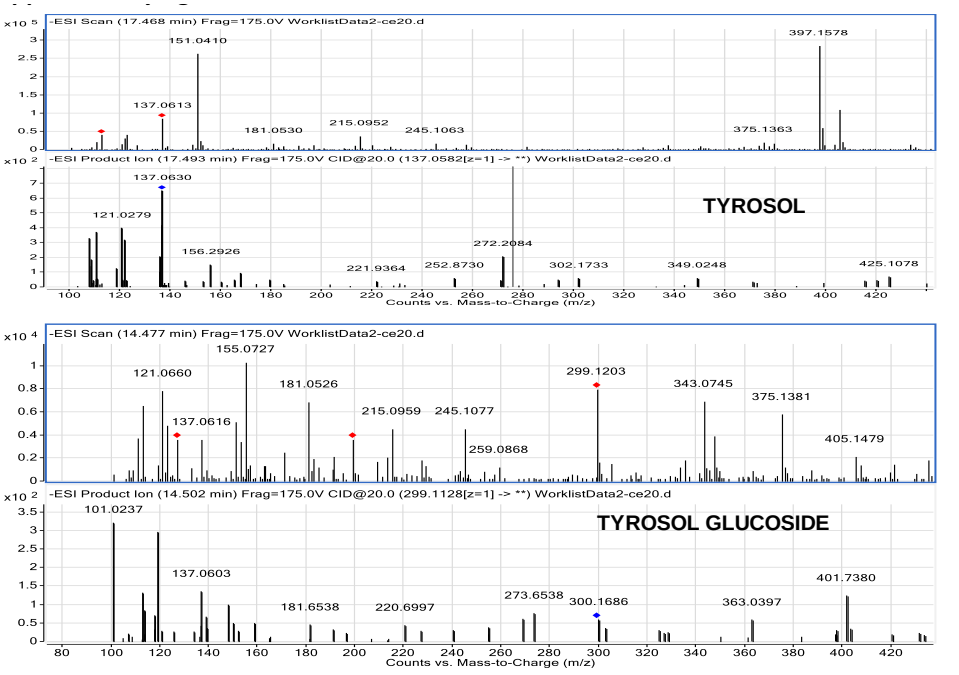
Funding sources

The Spanish Ministerio de Ciencia e Innovación (MICINN) and FEDER program are thanked for financial support through project CTQ2009-07430. F.P.C. is also grateful to the Ministerio de Ciencia e Innovación for a Ramón y Cajal contract (RYC-2009-03921).

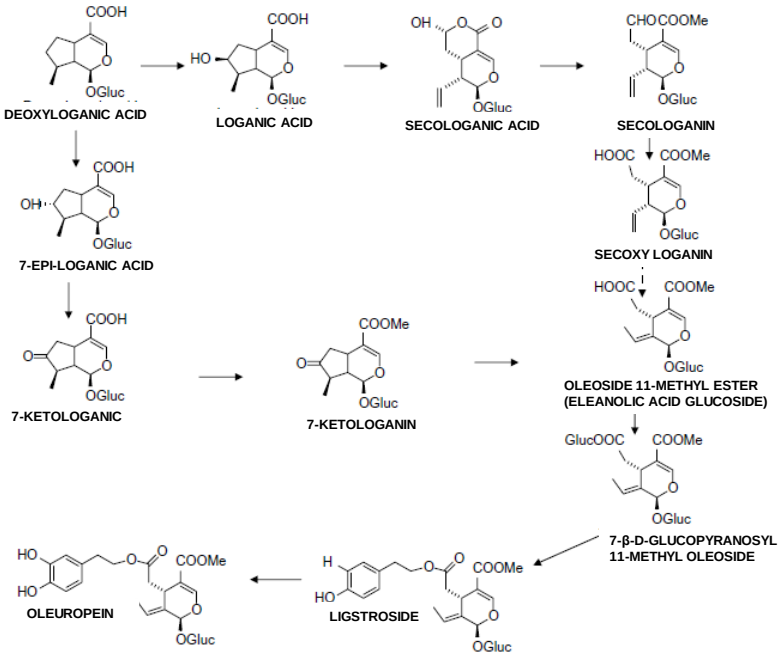
Supplementary material



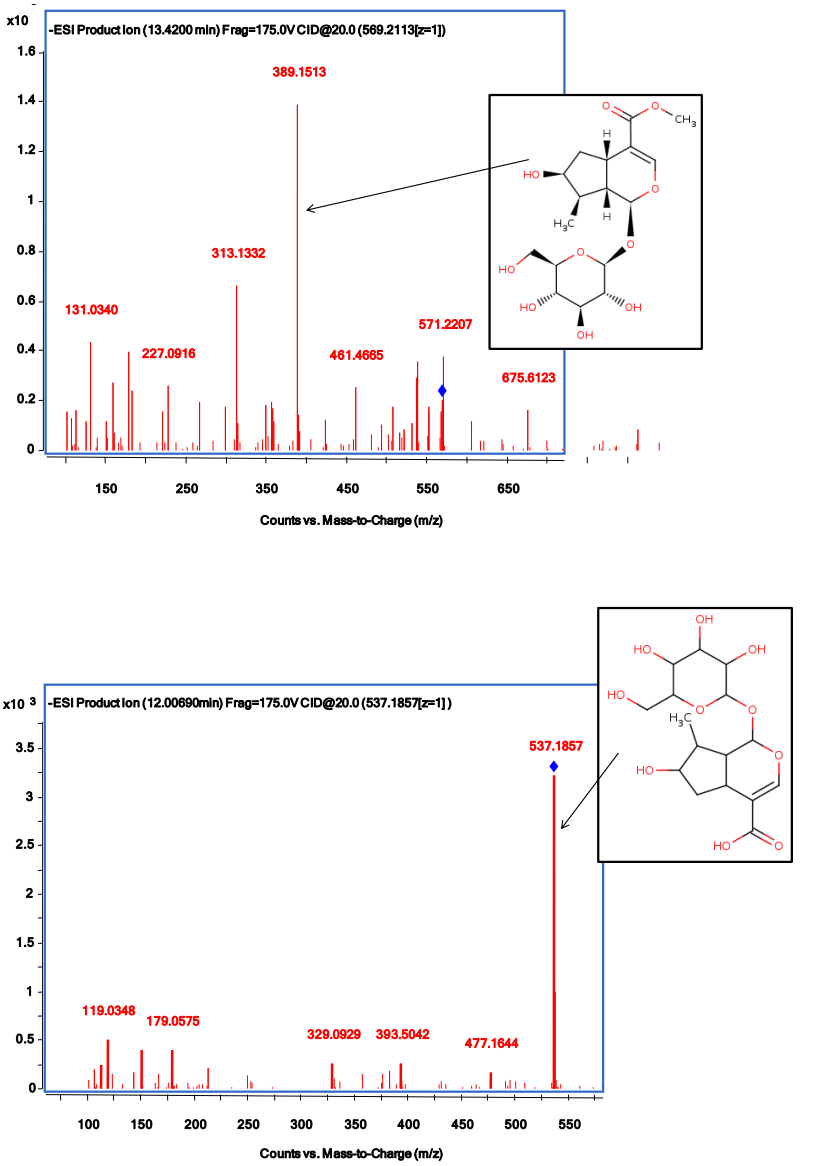
Supplementary Figure 1. MS/MS spectra from hydroxytyrosol diglucoside and hydroxytyrosol dirhamnoside.



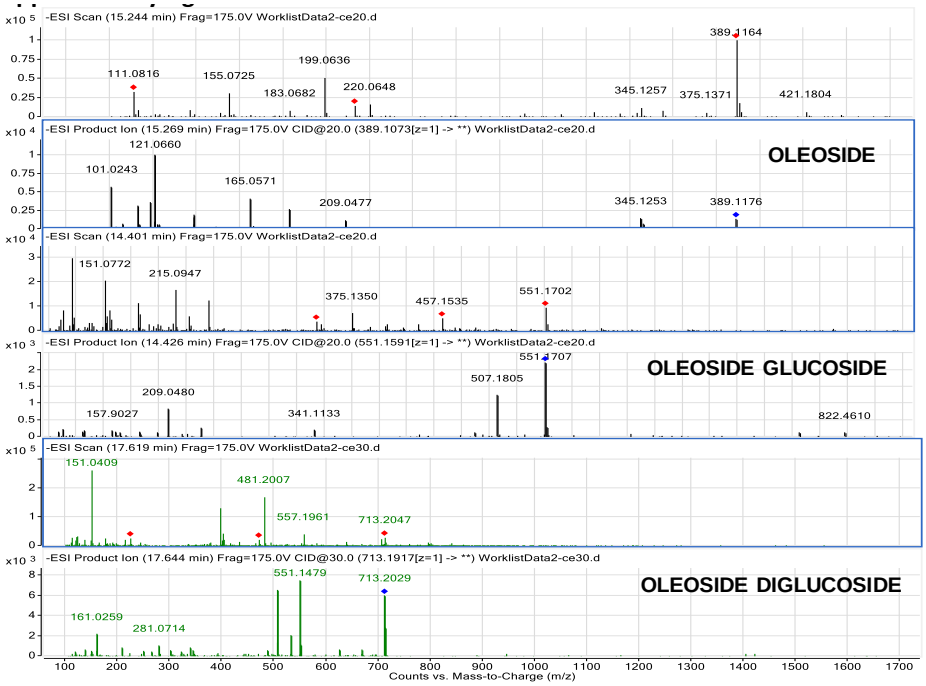
Supplementary Figure 2. MS/MS spectra from tyrosol and tyrosol glucoside.



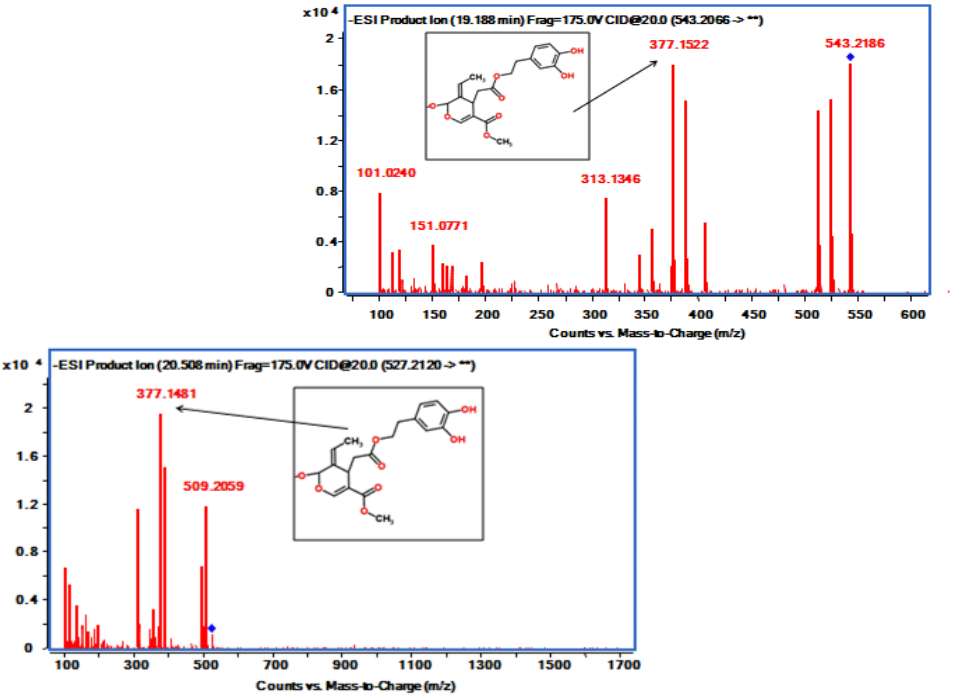
Supplementary Figure 3. Iridoids biosynthetic pathway.



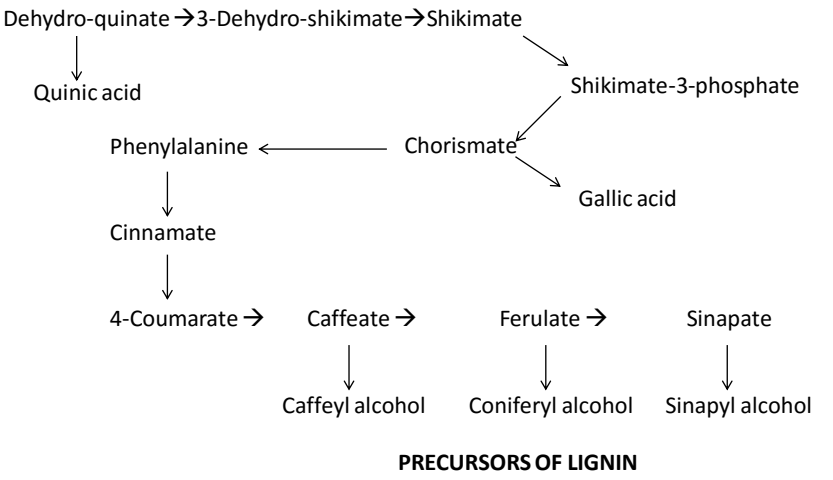
Supplementary Figure 4. MS/MS spectra from (A) loganin glucoside and (B) loganic acid glucoside.



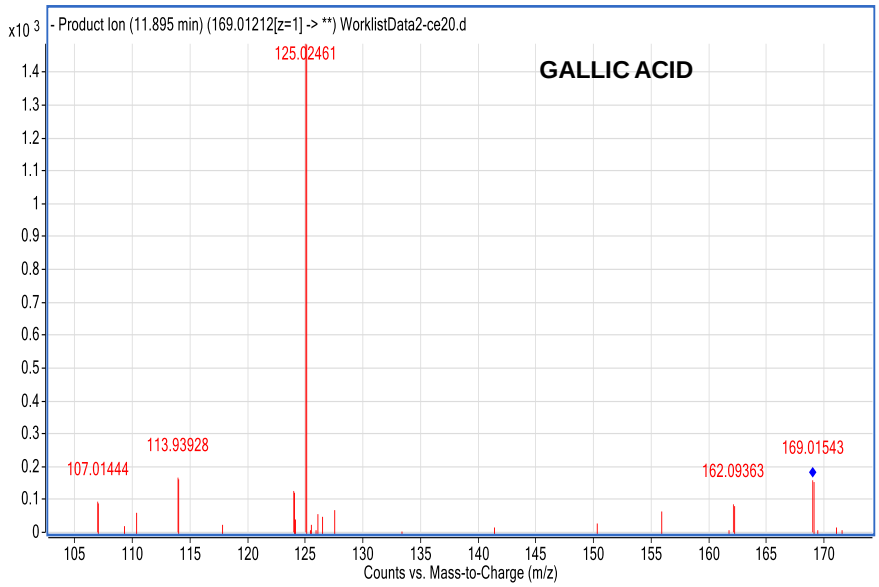
Supplementary Figure 5. MS/MS spectra from oleoside, oleoside glucoside and oleoside diglucoside.

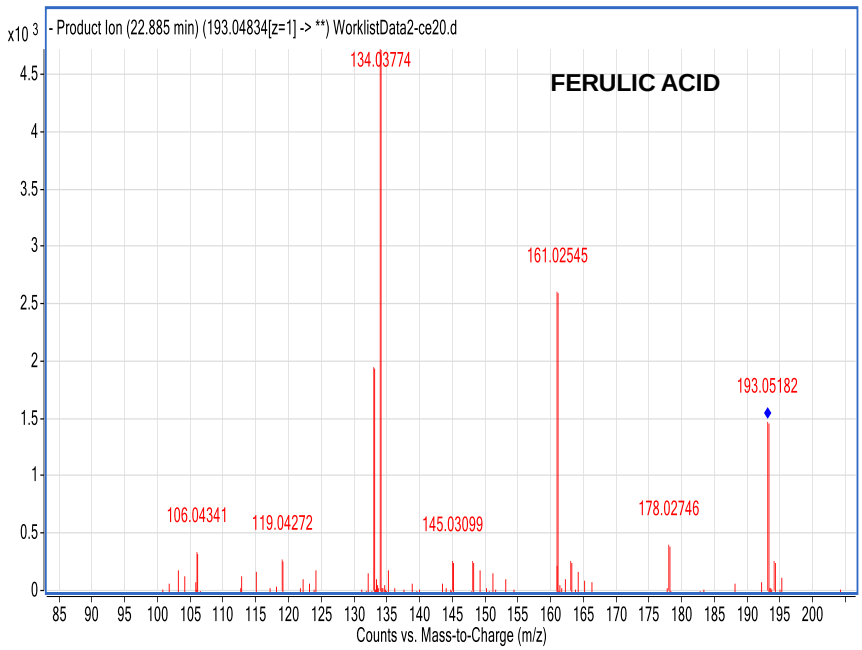


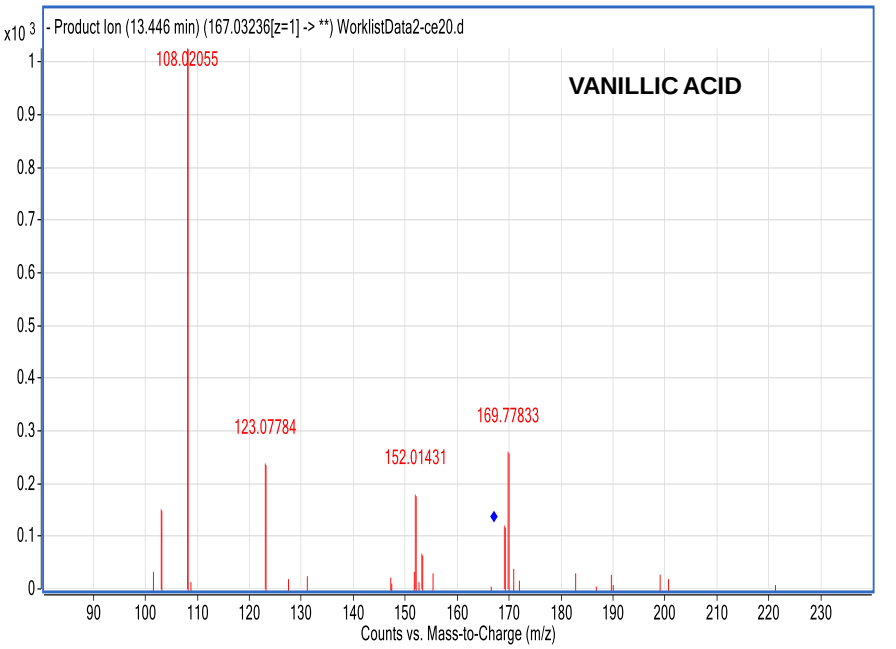
Supplementary Figure 6. MS/MS spectra for oleuropein derivatives.

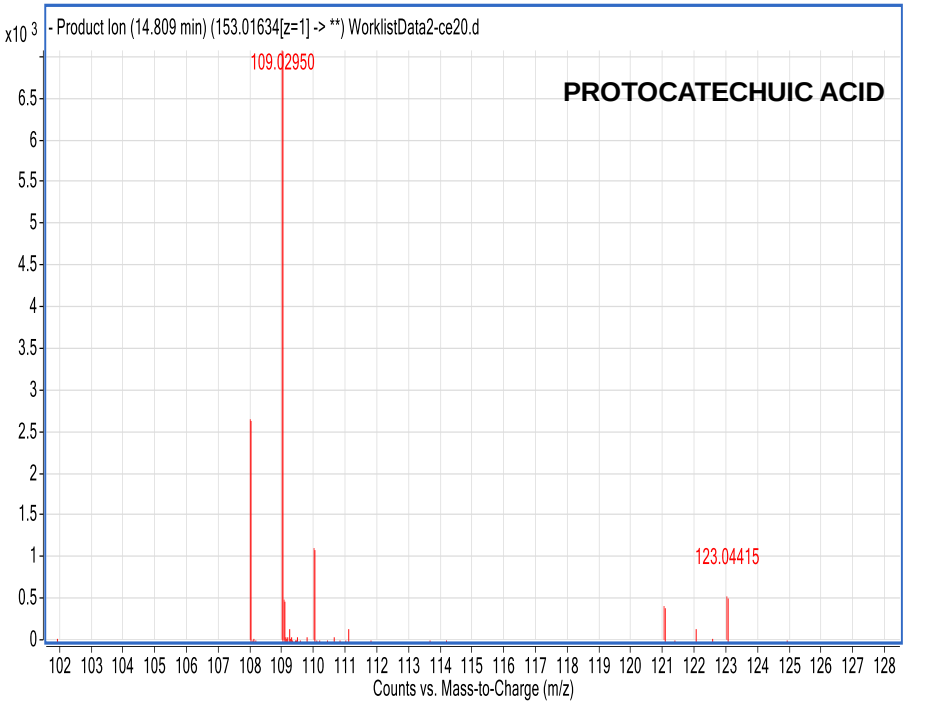


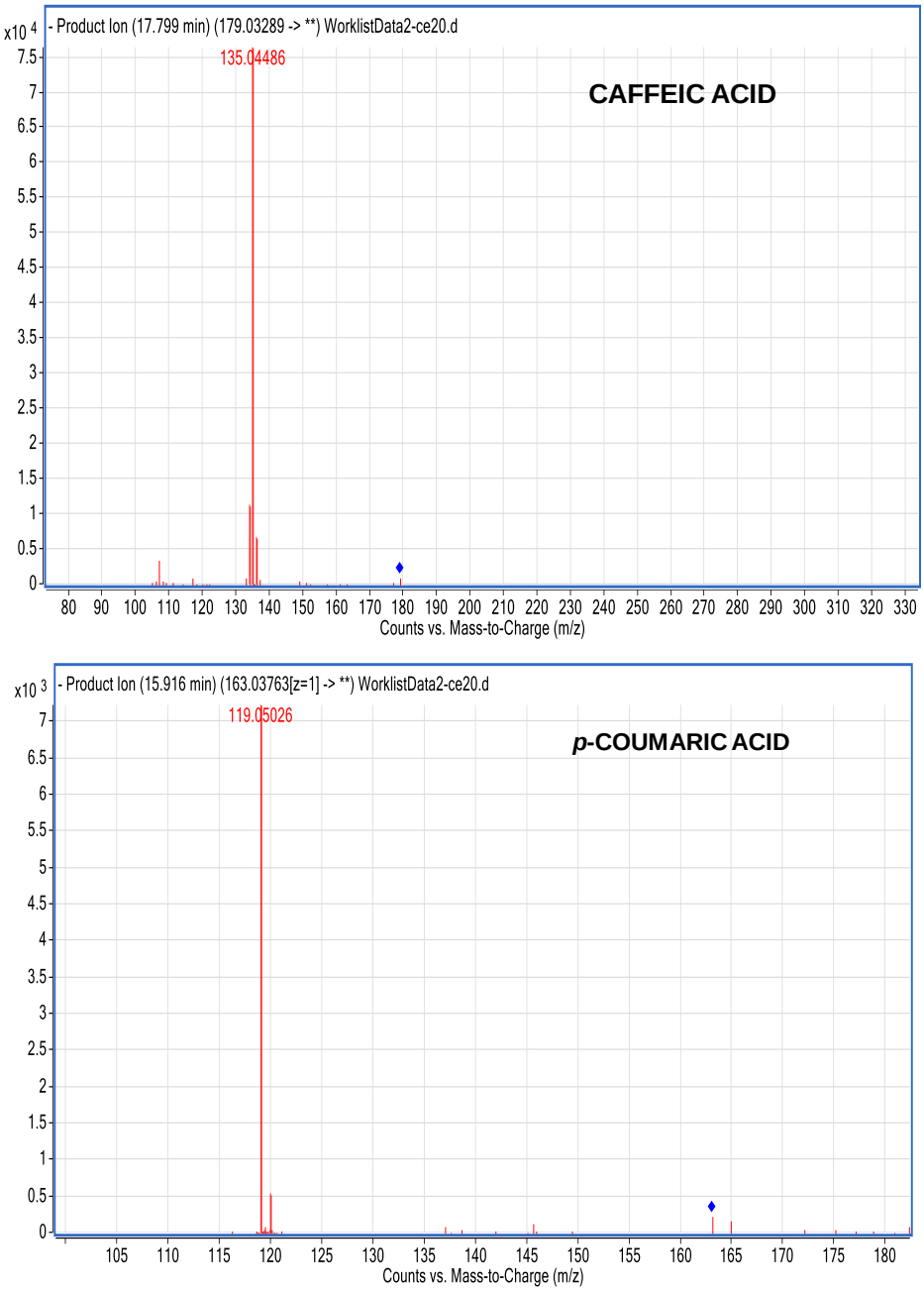
Supplementary Figure 7. Phenylpropanoid biosynthetic pathway.


- Product Ion (11.895 min) (169.01212[z=1] -> **) WorklistData2-ce20.d
x10 3
GALLIC ACID
125.02461
113.93928
107.01444
162.09363
169.01543
Counts vs. Mass-to-Charge (m/z)

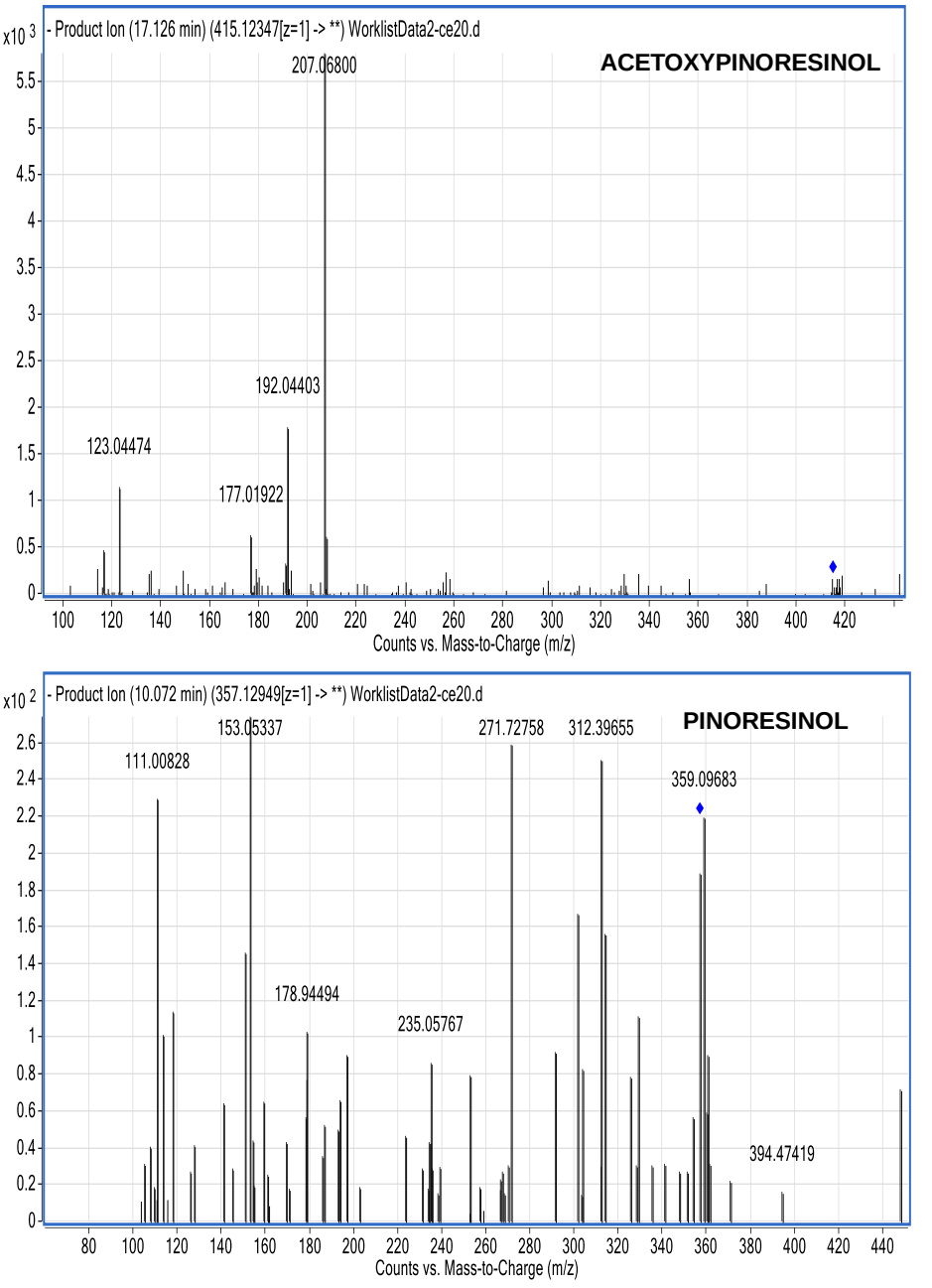

- Product Ion (22.885 min) (193.04834[z=1] -> **) WorklistData2-ce20.d
x10 3
FERULIC ACID
134.03774
161.02545
193.05182
106.04341
119.04272
145.03099
178.02746
Counts vs. Mass-to-Charge (m/z)


- Product Ion (13.446 min) (167.03236[z=1] -> **) WorklistData2-ce20.d
x10 3
108.02055
VANILLIC ACID
123.07784
169.77833
152.01431
Counts vs. Mass-to-Charge (m/z)

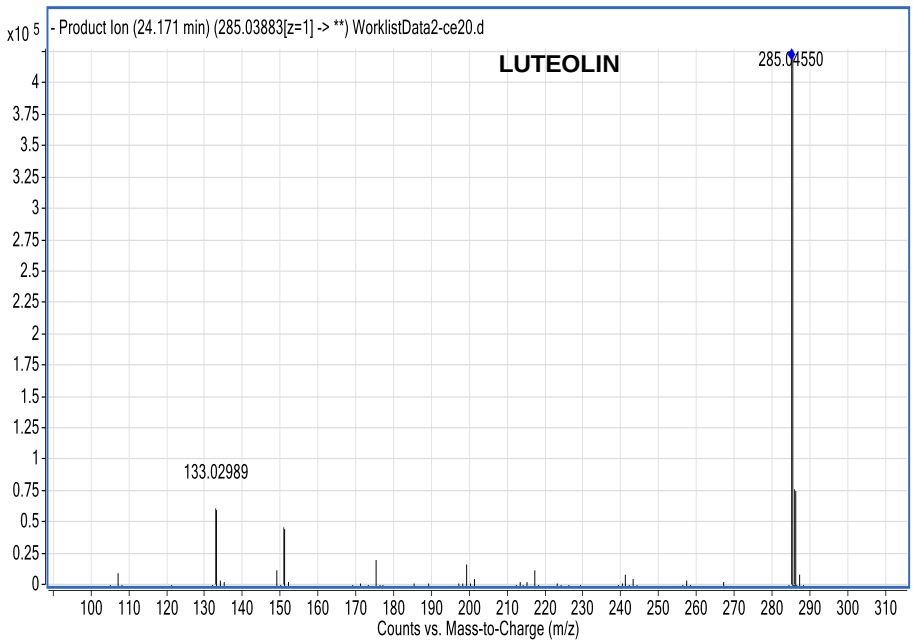

- Product Ion (14.809 min) (153.01634[z=1] -> **) WorklistData2-ce20.d
x10 3
109.02950
PROTOCATECHUIC ACID
123.04415
Counts vs. Mass-to-Charge (m/z)



Supplementary Figure 8. MS/MS spectra from phenolic acids.



Supplementary Figure 9. MS/MS spectra for pinoresinol and acetoxypinoresinol.



Supplementary Figure 10. MS/MS spectrum from luteolin.

PARTE B

Aportaciones al uso de aliento condensado como muestra clínica para la búsqueda de marcadores de cáncer de pulmón

La investigación que se recoge en esta Parte B de la Memoria de Tesis se enmarca en la contribución del grupo de investigación en el que se integra la doctoranda al desarrollo del proyecto “Sistema de detección precoz del cáncer” del IMIBIC (Instituto de Investigación Biomédica de Córdoba), y en el que el grupo participó activamente en la preparación, planificación y presentación. El banco de muestras de aliento condensado creado para el desarrollo del proyecto surtió al grupo, que desarrolló los aspectos metabolómicos del estudio.

Teniendo en cuenta la naturaleza volátil de la mayor parte de los componentes de la muestra, la herramienta analítica idónea para el estudio fue un cromatógrafo de gases conectado a un detector de masas de alta resolución de tiempo de vuelo (GC-TOF/MS), adquirido con cargo al proyecto para esta investigación.

La primera etapa del estudio (obligatoria dado el escaso uso previo del aliento condensado como muestra clínica y la hasta entonces ausencia en este tipo de estudios de un equipo instrumental como el utilizado) con las muestras iniciales del banco fue el desarrollo de un método analítico en el que se optimizó la etapa de preparación de la muestra para extraer la máxima información analítica. Se consiguió así identificar de forma tentativa 51 compuestos y se evaluaron diferentes alternativas de normalización de los datos obtenidos para obviar la contribución de fuentes de variabilidad externas y permitir un análisis comparativo entre grupos de muestras. Todo ello constituye el Capítulo VI de esta Memoria.

La existencia de muestras de pacientes de cáncer de pulmón y de individuos en riesgo de padecerlo, junto con un grupo control compuesto por individuos sanos, dio lugar a un estudio de discriminación entre los grupos, en el que se tuvo en cuenta la procedencia de la muestra a lo largo de las vías respiratorias, ya que el sistema de muestreo así lo permite. Se distinguieron compuestos que permiten discriminar entre los tres grupos de individuos participantes y se compararon los resultados con los de estudios existentes en la bibliografía. Este estudio, recogido como Capítulo VII de la Memoria, representó la base para la siguiente investigación, que se centró en la configuración de

paneles de potenciales marcadores para ayudar en el diagnóstico de cáncer de pulmón, con especial énfasis en la discriminación de pacientes con cáncer de los individuos con factor de riesgo, esencialmente fumadores. Para ello se hizo uso de una herramienta bioinformática que permite la combinación iterativa de los compuestos detectados en aliento condensado con el fin de buscar paneles que ofrezcan la mejor respuesta en términos de sensibilidad, especificidad, o la combinación de ambos parámetros. Esta investigación constituye el Capítulo VIII de esta Memoria de Tesis.

El último capítulo de este bloque se dedica al estudio del efecto del tabaco, principal causa de las enfermedades respiratorias y de la inflamación pulmonar. El mismo tipo de muestra se utilizó para comparar los dos grupos utilizados como control frente al de cáncer de pulmón (individuos sanos y con factor de riesgo) clasificando a los participantes según su hábito en fumadores, ex fumadores (clasificados según el tiempo transcurrido desde que dejaron de fumar) y no fumadores. Los compuestos relevantes para la comparación teniendo en cuenta el hábito, el consumo acumulado y el índice de masa corporal, y la discusión a que dio lugar el estudio se recogen en el Capítulo IX de esta Memoria.

Capítulo VI:

Development of a method for metabolomic analysis of human exhaled breath condensate by gas chromatography-mass spectrometry in high resolution mode

Analytica Chimica Acta

Volume 887, 5 August 2015,

Pages 118-126

Development of a method for metabolomic analysis of human exhaled breath condensate by gas chromatography–mass spectrometry in high resolution mode

Ángela Peralbo-Molina[a,b], Mónica Calderón-Santiago[a,b], Feliciano Priego-Capote[a,b*], Bernabé Jurado-Gámez[b], María Dolores Luque de Castro[a,b*]

[a]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14014, Córdoba, Spain*

[b]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, E-14014, Córdoba, Spain*

Development of a method for metabolomic analysis of human exhaled breath condensate by gas chromatography–mass spectrometry in high resolution mode

Ángela Peralbo Molina, Mónica Calderón-Santiago, Feliciano Priego-Capote, Bernabé Jurado-Gámez, María Dolores Luque de Castro*

Abstract

Exhaled breath condensate (EBC) is a promising biofluid scarcely used in clinical analysis despite its non-invasive sampling. The main limitation in the analysis of EBC is the lack of standardized protocols to support validation studies. The aim of the present study was to develop an analytical method for analysis of human EBC by GC-TOF/MS in high resolution mode. Thus, sample preparation strategies as liquid–liquid extraction and solid-phase extraction were compared in terms of extraction coverage. Liquid–liquid extraction resulted to be the most suited sample preparation approach providing an average extraction efficiency of 77% for all compounds in a single extraction. Different normalization approaches were also compared to determine which strategy could be successfully used to obtain a normalized profile with the least variability among replicates of the same sample. Normalization to the total useful mass spectrometry signal (MSTUS) proved to be the most suited strategy for the analysis of EBC from healthy individuals (n=50) reporting a within-day variability below 7% for the 51 identified compounds and a suited data distribution in terms of percentage of metabolites passing the Skewness and Kurtosis test for normality distribution. The composition of EBC was clearly dominated by the presence of fatty acids and derivatives such as methyl esters and amides, and volatile prenol lipids. Therefore, EBC offers the profile of both volatile and non-volatile components as compared to other similar biofluids such as exhaled

breath vapor, which only provides the volatile profile. This human biofluid could be an alternative to others such as serum/plasma, urine or sputum to find potential markers with high value for subsequent development of screening models.

Keywords: Metabolomics - Profiling - Exhaled breath condensate - Gas chromatography – Exhaled breath condensate metabolome – Mass spectrometry

1. Introduction

The most common biofluids in metabolomics analysis applied to clinical and nutritional studies are blood (serum or plasma) and urine. Nowadays, alternative biofluids such as sweat, tears, saliva or exhaled breath, with a less complex composition, are gaining popularity as they can be easily obtained in a non-invasive manner [1,2]. Among these biofluids, exhaled breath is one of the most accessible samples as it can be obtained without adverse effects, even in children and patients with serious respiratory diseases [3,4]. There are two types of exhaled breath samples, depending on the sampling protocol: exhaled breath vapor (EBV) and exhaled breath condensate (EBC). The former is obtained by collecting the breathing in a suited bag or a similar device. The main fraction of EBV (>99%) comprises nitrogen, oxygen, CO_2, water vapor, and other inert gases, and the remaining fraction is formed by a mixture of volatile organic compounds (VOCs) present at concentration ranges from few $\mu g\ mL^{-1}$ to $pg\ mL^{-1}$ [5,6]. The levels of VOCs such as aldehydes [7,8] and alkane derivatives as 4-methyldecane [9,10] in clinical samples such as urine or EBV have been associated to cancers, lung cancer among them [11]. This type of cancer has been linked to predictive models based on canine detection, ascribed to the presence of VOCs [12].

On the other hand, EBC is collected by cooling the exhaled breath that condensates to give a liquid solution composed by soluble exhaled gases and metabolites of the extracellular lining fluid. Despite the main component of EBC is water, hundreds of different components at trace concentrations can be found in this sample [13-16], ranging from small inorganic ions through large organic molecules to peptides, proteins, surfactants, macromolecules and VOCs [17]. Therefore, EBC is characterized by a more varied composition than EBV.

Concerning the analysis of EBC and EBV, the main difference lies in the sample preparation strategy selected as a function of the physical state of each sample. Solid-phase microextraction (SPME) seems to be a suited technique for treatment of EBV after collection in Tedlar bags and prior to GC-MS analysis of

VOCs [18-20]. One other option is the use of sorbents such as activated carbon to retain sample components, which are further desorbed in the injection unit of GC-MS equipment [21]. On the other hand, the liquid nature of EBC increases the number of alternatives to analyze this sample. Thus, liquid–liquid extraction (LLE) by taking advantage of the polarity of the extractant could be a fast alternative to separate compounds from the aqueous matrix before GC-MS characterization. One other competing alternative could be solid-phase extraction (SPE) that could be used to concentrate the EBC components and remove matrix interferents.

Several target strategies have been used in metabolomics to determine compounds of clinical relevance in EBC samples [17] such as those involved in the NO-oxidation pathway [22], lipid peroxidation products such as isoprostanes and leukotrienes, hydrogen peroxide and cytokines. Concerning untargeted analysis, different platforms have been proposed to seek for markers of different diseases, especially lung diseases. A recent study by Laurentiis *et al.* has proposed NMR to discriminate between patients with smoking related diseases and healthy individuals by some EBC components [23]. The number of patients selected for this study was particularly low and external validation was carried out. Despite the wide variety of methods in the literature to identify or determine metabolites in EBC, a generic characterization of the EBC sample has not so far been carried out, possibly owing to methodological limitations [17], as the absence of normalization protocols. In fact, the variability observed in GC-MS results obtained from targeted and non-targeted metabolomics analysis of EBC was highly dependent on the strategy for normalization. With these premises, the aim of the present study was to develop a protocol for EBC analysis, with special emphasis on sample preparation. For this purpose, LLE and SPE were evaluated by using different extractants and eluting solutions, respectively. The resulting approach was used for tentative profiling of EBC samples from healthy individuals. A normalization study has also been carried out by testing four strategies: the total useful MS signal (MSTUS), the internal standard (IS)

response and the combination of the IS response with the collected EBC volume or the expired EBC volume.

2. Experimental

2.1. Reagents

Cyclohexane, hexane, dichloromethane, acetonitrile and ethyl acetate TraceSELECT® grade from Sigma-Aldrich (St. Louis, USA) were the organic solvents for sample preparation. Deionized water, 18 MΩ·cm, was from a Millipore Milli-Q water purification system (Bedford, USA). A standard mixture containing ten linear alkanes from C10 to C40 designed for performance tests in GC from Sigma-Aldrich was used to establish the retention index (RI) calibration model. Triphenylphospate, also from Sigma-Aldrich, was used as IS.

2.2. Instruments and apparatus

An ECOScreen2 device (FILT Thorax-und Lungen Diagnostik GmbH, Berlin, Germany) was the EBC collector. A centrifugal SPE procedure was carried out by a Mixtasel centrifuge (Selecta, Barcelona, Spain), and homogenization of the extracts by an MS2 Minishaker Vortex (IKA, Germany).

An Agilent 7890A Series GC system coupled to an Agilent 7200 UHD Accurate-Mass QTOF hybrid mass spectrometer equipped with an electron impact (EI) ionization source (Santa Clara, CA, USA) was used. The analytical sample was monitored in high resolution mode.

2.3. Cohort selected for the study

All experiments were carried out in accordance with ethical principles of human medical research (World Medical Association, Helsinki Declaration, 2004). The ethical review board of Reina Sofía University Hospital (Córdoba, Spain) approved and supervised the clinical study (Project “Development of

methods for early cancer detection, December 29, 2011"). The EBC samples were obtained from 50 healthy volunteers early in the morning before breakfast. The individuals of the cohort, aged between 40 and 80, were recruited in the Respiratory Medicine Department. The samples were immediately stored in the biorepository of the hospital at -80 ºC. A pool prepared by mixing 100 µL of EBC from each participant was used for the development of the method.

2.4. EBC collection procedure

Sampling was carried our using a modified ECOScreen2 device that directly condenses and collects the EBC in disposable polyethylene bags. The device operates at -20 ºC to allow controlled collection of condensate from the respiratory tract. The main modification of the sampler was the insertion of a commercial protection filter from Scharlab (Barcelona, Spain) placed over the inlet air valve to avoid the presence of contaminants (derived from cleaning products, cosmetics, perfume, etc.) and particles from the room atmosphere. This filter was daily changed to avoid contamination. Using tidal breathing and a nose-clip during sampling, 15 min were required to collect an EBC volume ranging from 0.5 to 3.0 mL. On the other hand, the expired EBC volume varied from 160 to 699 L. The samples were analyzed within 3 months after collection.

2.5. Sample treatment

Two different experimental protocols were compared to select that providing maximum metabolite coverage. In all cases, 7.5 ng of triphenyl phosphate as IS was added and vortexed with 100 µL of EBC pool prior extraction.

2.5.1. Liquid-liquid extraction

Four different organic solvents (hexane, cyclohexane, ethyl acetate and dichloromethane) were evaluated as extractants. In all cases, 100 µL of EBC pool was vortexed with 100 µL of each extractant in a glass insert at room temperature

for 1 min. To assure the absence of water in the extract to be chromatographed, the vial was subjected to -20 ºC to freeze the aqueous fraction; then, the organic phase was isolated and located into a new glass insert for analysis.

A blank test consisted of using water as sample and performing the whole procedure with each solvent. Additionally, contamination from polyethylene bags was discarded by analysis of water introduced in them as blank, which was subjected to liquid-liquid extraction with the four organic solvents tested in this research.

2.5.2. Centrifugal solid-phase extraction

Clean-up of EBC was carried out by centrifugal SPE with C18 Micro SpinColumnTM systems (Harvard Apparatus, MA, USA). The protocol recommended by the manufacturer was followed and different solvents were tested as eluants. The protocol for centrifugal SPE was as follows: 150 µL of water for solvation, 150 µL of acetonitrile for sorbent conditioning, 150 µL of 4% (v/v) acetonitrile aqueous solution for sorbent equilibration, 100 µL of sample, 100 µL of water and, finally, 100 µL of the tested eluant: hexane, cyclohexane, ethyl acetate or dichloromethane. Similarly to LLE, the vial was subjected to -20 ºC to freeze the aqueous fraction.

A blank test was also prepared using water as sample and performing the whole procedure with each solvent.

2.6. GC-TOF/MS analysis

GC-TOF/MS analyses were performed by EI ionization at 70 eV and controlled by MassHunter Acquisition B.06. The GC separation was carried out with a fused silica DB-5MS-UI 30 m×0.25 mm i.d., 0.25 µm film thickness capillary column. The GC oven temperature program started at 60 °C (1 min held), followed by a temperature ramp of 10 °C min^{-1} to final 300 °C (2 min held). Post-run time was programmed for 4 min up to 310 ºC to assure complete elution

of the injected sample. Pulsed splitless injections of 1 µL of sample were carried out at 250 °C, and ultrapure grade helium was used as carrier gas at 1.0 mL min^{-1}. The interface, ion source and quadrupole temperatures were set at 280, 300 and 200 °C, respectively. A solvent delay of 5.5 min was used to prevent damage of the ion source filament. The TOF detector was operated at 5 spectra s^{-1} in the mass range *m/z* 50–550 and the resolution was 8500 (full width half maximum, FWHM) at *m/z* 501.9706. Mass-spectrometric grade PFTBA (perfluorotri-n-butylamine) was used for daily mass calibration.

2.7. Data processing and statistical analysis

Unknown Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to process all data files obtained by GC-TOF/MS in full scan mode. Treatment of raw data files started by deconvolution of chromatograms to obtain a list of molecular features (MFs) considered as potential compounds defined by the *m/z* value of one representative ion for each chromatographic peak and its retention time. For this purpose, the deconvolution algorithm was applied to each sample by considering all ions exceeding 1500 counts for the absolute height parameter, the accuracy error at 50 ppm and the window size factor at 150 units, respectively. The list of MFs obtained for each analysis was exported as data files in compound exchange format (.cef files). These data files were treated with the Mass Profiler Professional (MPP) software package (version 12.1, Agilent Technologies, Santa Clara, CA, USA) to build the data set including quantitative information (expressed as peak area ratio with the internal standard) of MFs detected in each analysis for further processing.

The data were then processed by alignment of the potential MFs across samples according to their retention time and *m/z* value using a tolerance window of 0.3 min and an accuracy error of 15 ppm. Data pretreatment was based on logarithmic transformation to reduce relatively large differences among the respective MF abundances. The MFs extracted from the analysis of blanks were removed from the dataset. The extraction algorithm confirmed the efficiency of

this filtering step, which was applied to all the sample treatments tested. Stepwise reduction of the MFs number was based on frequency of occurrence by comparing repetitions of the same fraction or extraction procedure. A filter by frequency was set at 100%, thus ensuring detection of each MF in all the injected replicates.

Finally, the MFs from data pretreatment were exported (.cef file) for recursive analysis. For this purpose, the Quantitative Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to reintegrate all potential compounds found in all analyzed samples. The resulting table was exported in comma separated value format (.csv file) and reprocessed with the Mass Profiler Professional (MPP) software package. A filter of 10% for sample variability within replicates was applied to assure the effectiveness of the recursive analysis. Tentative identification of compounds was performed by searching mass spectra on the NIST 11 database taking into account the RI values.

Once the identification step was completed, four different approaches were applied to normalize the data set in terms of methodological variability. Normalization to IS response uses as denominator the area of a compound (TPP) added prior to extraction, whereas normalization to MSTUS uses as denominator the total intensity of the extracted compounds common to all samples. The other two approaches use the volume collected and expired in combination to the internal standard response, so first each sample is divided by its volume (collected or expired) and then by the signal of its internal standard.

3. Results and discussion

As mentioned before, EBC should be considered as a biofluid of potential great interest for clinical analysis because of its non-invasive sampling. However, EBC samples are not ready for direct analysis by GC-MS, which should seem the suited determination approach taking into account the content of VOCs.

Therefore, a sample preparation step is demanded to define suited protocols for as wider as possible analysis and characterization of EBC composition.

3.1. Liquid-liquid extraction

Liquid-liquid extraction can be a good sample preparation strategy prior to GC-MS analysis thanks to its simplicity and efficiency. Because of the chemical diversity of compounds present in EBC, the performance of four different extractants (hexane, cyclohexane, ethyl acetate and dichloromethane) was compared. In all cases, aliquots of the EBC pool were 1:1 extracted with each solvent and the resulting extracts analyzed by GC-TOF/MS. Supplementary Figure 1 shows the base peak chromatograms (BPCs) provided by analysis of each extract. Some surfactants used to clean the EBC collector were detected in the chromatograms as well as other signals ascribed to column bleeding, but they were easily identified and discarded prior to comparison among extracts.

Figure 1 shows the Venn diagram comparing the compounds identified in each extract, which are listed in Supplementary Table 1 as organized by their retention time. As can be seen, extraction by hexane enabled the detection of 51 compounds, and only 27, 25 and 22 compounds were detected in the extracts obtained with cyclohexane, ethyl acetate and dichloromethane, respectively. All the compounds detected in the other extracts were also found in the hexane extract; therefore, hexane is endowed with the best extraction capability for coverage of compounds in EBC.

3.2. Solid-phase extraction

SPE is extensively used to remove and concentrate compounds present in liquid samples at trace level. Therefore, SPE could be an alternative for preparation of aqueous samples prior to GC-MS analysis with the possibility of removing particles and undesired compounds such as inorganic salts or proteins present in EBC. Selection of the appropriate sorbent is crucial to achieve high and reproducible recovery of the target compounds. Due to the retention

properties of C18 material, cartridges packed with this sorbent were selected in this study for centrifugal SPE.

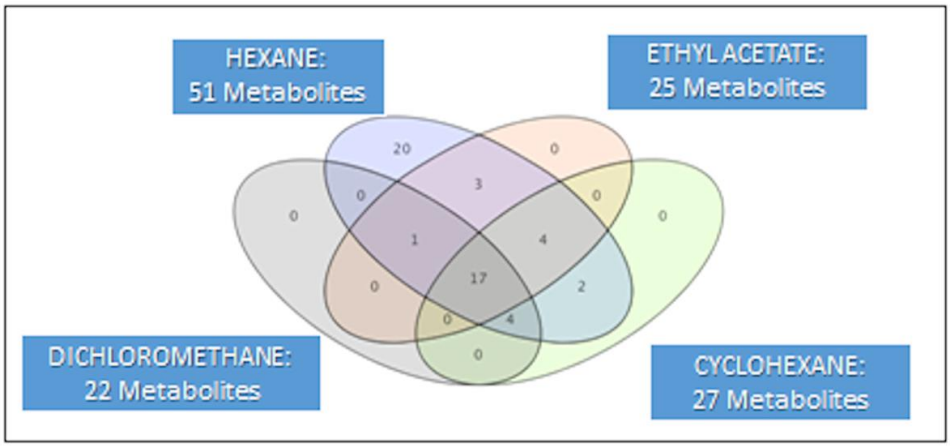


Figure 1. Venn diagram comparing the compounds extracted by LLE using different extractants.

Hexane, cyclohexane, ethyl acetate and dichloromethane were compared as eluents using the protocol described under section 2.5.2. In all cases, the volumes of sample and eluent were equal. Fractions of the eluate were analyzed by GC-TOF/MS with EI ionization. Supplementary Figure 2 shows the BPCs from each eluate after subtracting the blank chromatogram to eliminate signals associated to interferents. Figure 2 shows the Venn diagram that compares the compounds identified in the extract obtained with used eluents. Supplementary Table 2 lists the compounds identified in the eluates from SPE sample preparation. As can be seen, elution using hexane or cyclohexane enabled detection of 39 compounds, meanwhile 23 or 10 compounds were detected with ethyl acetate or dichloromethane, respectively. Therefore, hexane and cyclohexane were the most efficient eluents as they allowed detection of 86.7% of the total coverage achieved by combination of the four eluents (89.7% of the compounds found in the analysis of the eluates provided by hexane or cyclohexane were common).

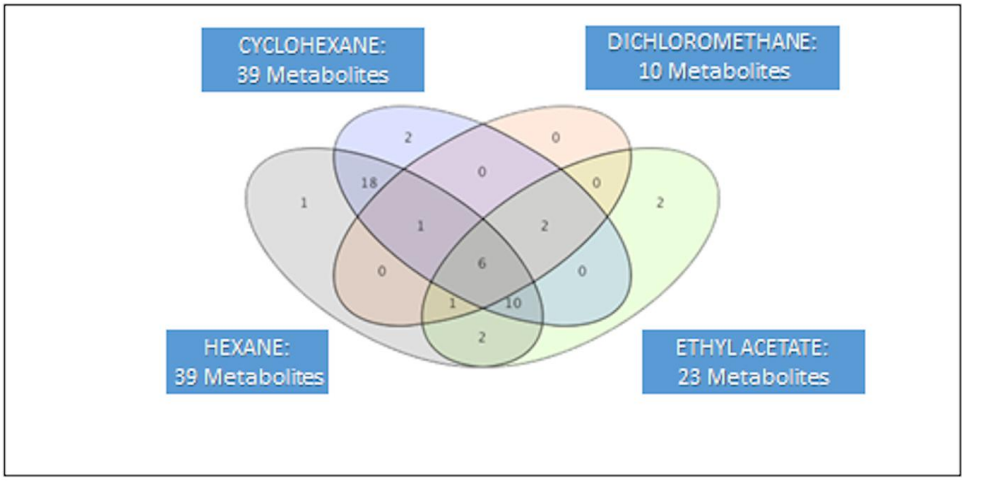


Figure 2. Venn diagram comparing the compounds extracted by SPE using different eluants.

3.3. Comparison of the sample preparation protocols

The two extraction techniques independently tested to maximize detection coverage in EBC were compared and evaluated for implementation in an analytical protocol for analysis of EBC. A common solvent, hexane, was found optimum in both protocols; therefore, comparison between centrifugal SPE and LLE was supported on hexane as eluent or extractant, respectively. Figure 3 shows the Venn diagram comparing the compounds identified in both cases. As can be seen, the number of compounds detected by LLE was clearly higher than by centrifugal SPE, probably because the low polar retention mechanism hinder retention of some sample components. In fact, the analysis of EBC subjected to LLE using hexane as extractant allowed detection of 51 compounds vs 39 compounds detected after centrifugal SPE, all of them detected by LLE. Therefore, LLE with hexane seems to be the suited sample preparation approach for analysis of EBC by GC-TOF/MS.

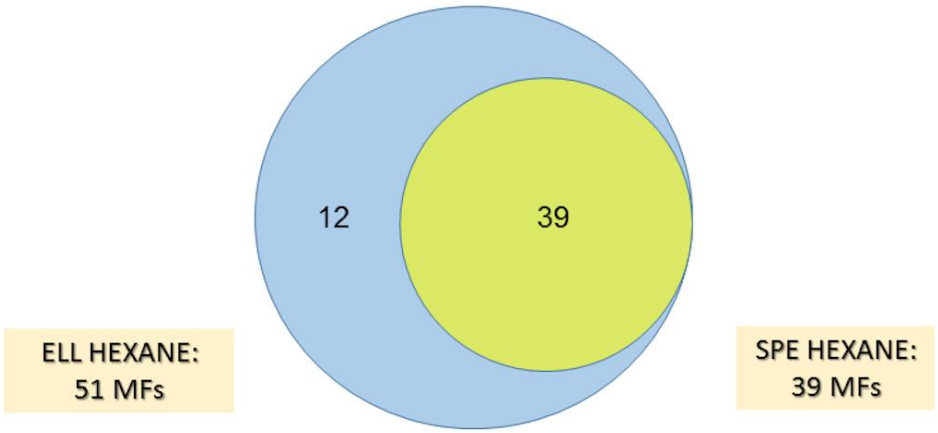


Figure 3. Venn diagrams comparing the compounds extracted by the two sample preparation methods using hexane as extractant or eluant.

Figure 4 shows the comparison between the efficiency of the two extraction approaches for analysis of EBC. Bars correspond to the different compounds defined by their peak area to the IS peak area ratio (vertical axis). As can be seen, some compounds such as palmitic acid, squalene or oleamide provided a better response after LLE, while others as methyl stearate or eicosenamide did it in SPE. No special trends concerning extraction efficiency were found; therefore, LLE provided a more complete coverage in terms of number of tentatively identified compounds, and it should be adopted for preparation of EBC prior to GC-MS analysis.

3.4. Validation study

A validation study of the selected sample preparation protocol was carried out by analysis of EBC as an attempt to assess the methodological variability of the detection step and the stability of the samples. The study was performed by repetitive injection (n=3) of the extract from the EBC from each individual. The autosampler temperature was set at 5 ºC to minimize potential evaporation or degradation, and four different normalization strategies (*viz.*

internal standard response correction; normalization to MSTUS, which uses the total intensity of components that are common to all samples; and a combination of exhaled breath collected volume, exhaled breath expired volume with the internal standard response) were compared to minimize intra-individual variability. The variability was evaluated for the 51 compounds detected by comparing all samples. Supplementary Figure 3 shows the box and whisker plots that allow assessing the variability for each sample in the cohort with the four normalization approaches. The normalization to MSTUS showed a more linear distribution as compared to the resting strategies. A Skewness and Kurtosis test to evaluate if the data set fit a normal distribution was applied to compare the effect of the four strategies. The MSTUS approach and the combination of IS response and expired breath volume were the two alternatives providing the highest percentage of compounds normally distributed according to the Skewness and Kurtosis test, with 76.6 and 80.8%, respectively. On the other hand, the IS response and the combination of IS response and collected EBC volume led to lower values (66 and 12.76%, respectively).

Therefore, the MSTUS and the IS response/expired breath volume could be considered as adequate normalization strategies. To complete this study, the normalization strategy should eliminate the methodological variability. Thus, the strategy providing a normal distribution with the minimum methodological variability would be the most suited approach to be applied. Supplementary Figure 4 compares the methodological variability calculated between replicates for these two strategies standing out that the MSTUS normalization exhibited a better result than the normalization by the IS response/expired breath volume. With the MSTUS approach, the methodological variability, estimated as percentage of relative standard deviation (RSD), ranged from 0.002% to 7%, while the IS response/expired breath volume gave RSD values between 0.028% and 8.6%.

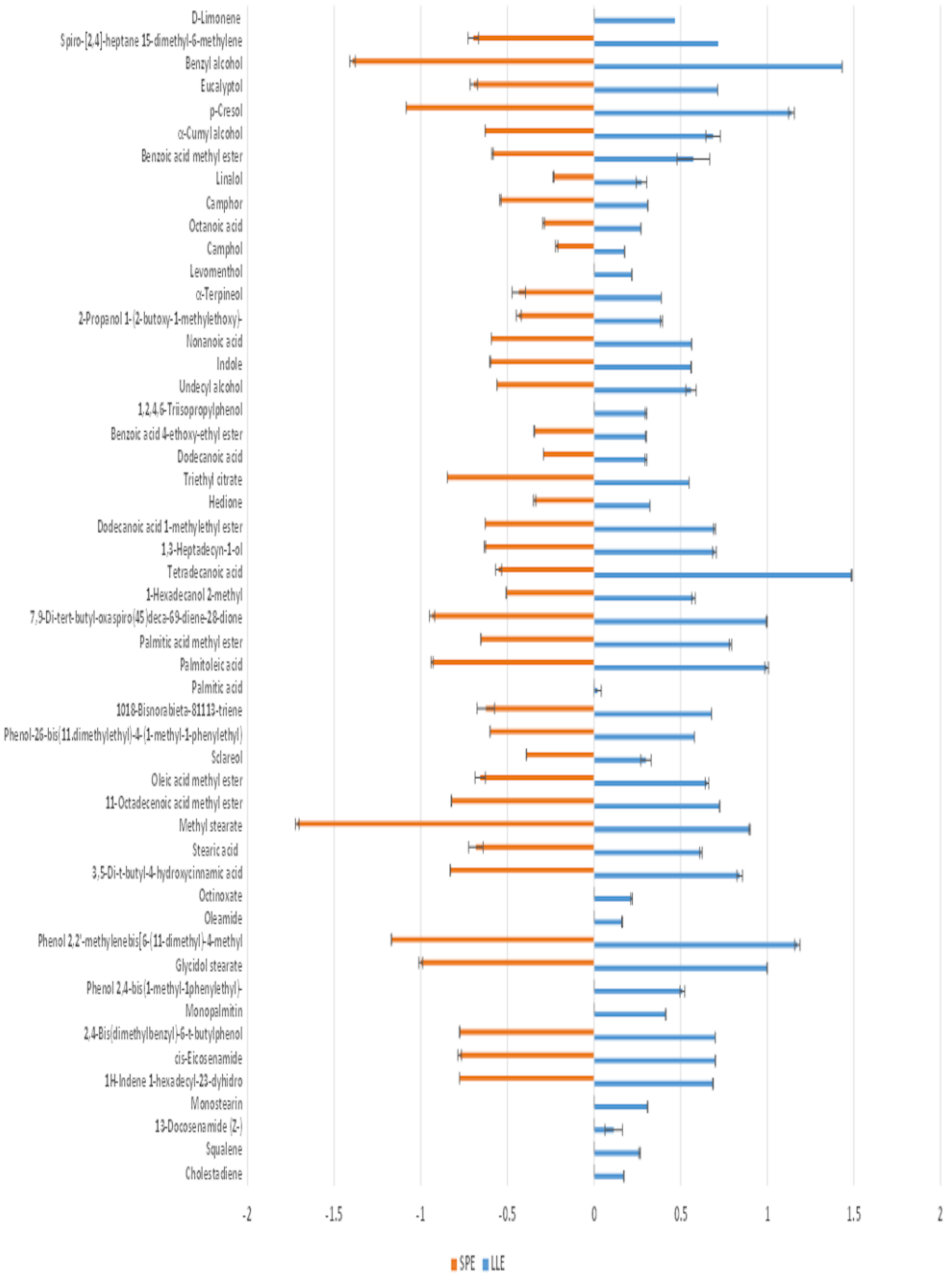


Figure 4. Bars diagram comparing the efficiency of the two extraction approaches, LLE and SPE, for analysis of EBC.

A recovery study was also performed to evaluate the efficiency of the LLE step by three consecutive extractions using the same sample aliquot. The extraction efficiency was estimated by comparing the peak area ratio obtained for each compound in the extracts from extractions 1, 2 and 3 (taking as 100% the overall amount of the target compound from the 3 extractions). The recovery rate for the first extraction ranged between 52 and 95.5% with a mean value of 77%. Additional extraction cycles involved a dilution effect by mixing all fractions and, therefore, promoted a decay of sensitivity. For this reason, a unique extraction cycle was adopted for EBC analysis.

3.5. Identification of compounds detected in EBC by GC-TOF/MS

As mentioned above, few studies on composition of EBC have been carried out since the systematic investigations of exhaled breath started in the last quarter of the twentieth century [24-27]. Pauling *et al.* described for the first time in 1971 volatile organic compounds (VOCs) in human breath [24]. At present, it is well-known that exhaled breath may contain a great variety of VOCs, most of them at the picomolar concentrations. Phillips *et al.* studied the variation in the GC-MS profile of VOCs in breath from fifty healthy individuals, which allowed identifying hydrocarbons, aromatic heterocyclic and homocyclic hydrocarbons, terpenes and derivatives, short carboxylic acids, aldehydes and ketones [21]. It should be emphasized that the protocol used by Phillips *et al.* was mainly focused on VOCs analysis since sampling involved sorbent retention with subsequent desorption into the GC-MS. In our case, only those compounds that condensed at the temperature of the sampler (-20 ºC) would be detected, as the sample collection system is not suitable for on-line coupling to the GC-MS system. For this reason, sample preparation was mandatory to replace water by a solvent compatible with GC-MS analysis.

The protocol involving LLE allowed tentative identification of 51 compounds, which included fatty acids and conjugates, fatty alcohols, fatty amides, prenol lipids as well as some organic compounds, the presence of which

could be attributed to exogenous sources. Identification was firstly carried out by searching mass spectra on the NIST11 database. Only identifications with a match factor and a reverse match factor higher than 700 were considered. The RI values included in the NIST database also supported identifications. An RI calibration model was built by plotting the retention times obtained by analysis of the alkane standard mixture (C10 to C40 with an even number of carbons) with the chromatographic method used in this research and those RI values provided for each alkane by the NIST database. Supplementary Figure 5 shows the RI calibration graph and the equation fitting the model. Then, the RI value was experimentally estimated for each identified compound by using the retention time and the calibration equation. The requirement to accept an NIST identification was that the difference between the experimental RI and the theoretical value provided by the NIST for each target compound should be below 100 units.

The NIST database does not contain high resolution MS information as provided by the TOF detector; thus, a third step was included to validate identification of each compound by using high resolution MS. Thus, the molecular formula for the tentative precursor ion ([M]+) and the most intense product ions for each MF should fit the NIST identification by setting a cut-off value in mass accuracy of 10 ppm. Table 1 lists the identified compounds classified by the chemical family. The great variety of compounds identified in EBC emphasizes the interest of this biofluid for clinical analysis. Most of the identified compounds were fatty acids and conjugates, which represent a highly metabolically active lipid class [28], and were previously described by Martínez-Lozano *et al.* [29], who analyzed directly breath vapor by electrospray ionization and atmospheric pressure chemical ionization in mass spectrometry. Three main saturated fatty acids –*viz.* myristic (C14:0), palmitic (C16:0) and stearic (C18:0)– were identified in EBC extracts. Among them, C16:0 and C18:0, together with oleic acid (C18:1) constitute about 78% of all free fatty acids in the circulatory system [28]. Supplementary Figure 6A shows the mass spectra with the in-source fragmentation pattern for the three fatty acids. The glycerol ester and methyl

ester derivatives of C16:0 and C18:0 were also identified. Despite oleic acid is the most concentrated fatty acid in humans, this compound was only detected as methyl ester and amide derivatives, the fragmentation pattern of which appears in Supplementary Figure 6B. Other fatty amides identified in EBC were erucamide and eicosenamide, previously studied because of their potential as antimicrobial agents [30]. Other fatty acid derivatives detected in EBC were fatty alcohols such as undecyl alcohol, glycidol stearate and 1-hexadecanol, 2-methyl. Short-chain fatty acids such as caprylic, pelargonic and lauric acids, which are formed in the gastrointestinal tract of mammals by microbial fermentation of carbohydrates [31], have also been found in the EBC extract.

Prenol lipids, synthesized from five carbon isoprene units, constitute other family of compounds identified in EBC. Three classes of prenol lipids were detected in this study: monoterpenes, diterpenes and triterpenes. Among the identified monoterpenes are camphol [20], limonene [32], linalool [21], α-cumyl alcohol [21] and α-terpineol [33], which had previously been identified in breath using different protocols.

Other two monoterpenes such as levomenthol and camphor were also identified. Sclareol, which has exhibited antibiotic [34], apoptotic [35] and anti-fungal [36] activities, was the only diterpene detected in EBC. Squalene, the only identified triterpene according to the fragmentation pattern shown in Supplementary Figure 6C, could be considered as an intermediate in the pathway for in vivo formation of human breath isoprene by peroxidation [37], but it is also involved in the synthesis of cholesterol [38]. Other compound involved in the production of cholesterol such as isoprene, previously found in EBV [5], was not detected in EBC, probably due to its low vaporization temperature. Two other lipids such as hedione (a jasmonic acid derivative), and a steroid derivative such as cholestadiene were also identified in EBC, the fragmentation patterns of which are shown in Supplementary Figure 6D.

Table 1. Compounds in extracts from EBC identified by GC-TOF/MS.

Compound Name	Retention Time	Formula	CAS ID	Fragments	Family
Eucalyptol	5.995	C10H18O	470-82-6	154.1361 – $[C10H18O]^+$ 139.1119 – $[C9H15O]^+$ 93.0695 – $[C7H9]^+$	Aliphatic Heteropolycyclic Compounds (Oxanes)
Indole	9.904	C8H7N	120-72-9	117.0558 – $[C8H7N]^+$ 90.0448 $[C7H6]^+$ 74.0145 – $[C6H2]^+$	Aromatic Heteropolycyclid Compounds (Indoles)
Benzoic acid 4-ethoxy-ethyl ester	12.923	C11H14O3	23676-09-7	194.0425 – $[C11H14O3]^+$ 149.0581 – $[C9H9O2]^+$ 121.0269 – $[C7H5O2]^+$	Aromatic Homomonocyclic Compounds (Benzene and Substituted Derivatives)
Benzoic acid methyl ester	6.951	C8H8O2	93-58-3	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic Homomonocyclic Compounds (Benzoic acid Derivatives)
3,5-Di-t-butyl-4-hydroxycinnamic acid	20.155	C17H24O3	22014-01-3	276.1712 – $[C17H24O3]^+$ 261.1479 – $[C16H21O3]^+$ 177.0896 – $[C11H13O2]^+$	Aromatic Homomonocyclic Compounds (Cinnamic Acid Derivatives)
Octinoxate	21.005	C18H26O3	5466-77-3	290.1867 – $[C18H26O3]^+$ 178.0611 – $[C10H10O3]^+$ 161.0578 – $[C10H9O2]^+$	Aromatic Homomonocyclic Compounds (Cinnamic Acid Derivatives)
p-Cresol	6.536	C7H8O	106-44-5	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic Homomonocyclic Compounds (Phenols and Derivatives-Cresol)
Benzyl alcohol	5.963	C7H8O	100-51-6	108.0565 – $[C7H8O]^+$ 91.0535 – $[C7H7]^+$ 79.0533 – $[C6H7]^+$	Aromatic Homomonocyclic Compounds (Primary alcohols)

Cont. Table 1

Dodecanoic acid 1-methylethyl ester	14.97	C15H30O2	10233-13-3	201.1835 – $[C12H25O2]^+$ 157.1203 – $[C9H17O2]^+$ 102.0656 – $[C5H10O2]^+$	Lipids (Fatty Acid Esters)
Palmitic acid methyl ester	17.342	C17H34O2	112-39-0	270.2545 – $[C17H34O2]^+$ 227.1998 – $[C14H27O2]^+$ 143.1048 – $[C8H15O2]^+$	Lipids (Fatty Acid Esters)
Oleic acid methyl ester	19.041	C19H36O2	112-62-9	296.2702 – $[C19H36O2]^+$ 264.2442 – $[18C32HO]^+$ 81.0685 – $[C6H9]^+$	Lipids (Fatty Acid Esters)
Methyl stearate	19.28	C19H38O2	112-61-8	298.2862 – $[C19H38O2]^+$ 255.2315 – $[C16H31O2]^+$ 87.0436 – $[C4H7O2]^+$	Lipids (Fatty Acid Esters)
Octanoic acid	7.939	C8H16O2	124-07-2	129.0894 – $[C7H13O2]^+$ 101.0592 – $[C5H9O2]^+$ 73.0282 – $[C3H5O2]^+$	Lipids (Fatty Acids and Conjugates)
Nonanoic acid	9.375	C9H18O2	112-05-0	129.0900 – $[C7H13O2]^+$ 115.0742 – $[C6H11O2]^+$ 91.0545 – $[C7H7]^+$	Lipids (Fatty Acids and Conjugates)
Dodecanoic acid	13.268	C12H24O2	143-07-7	200.1761– $[C12H24O2]^+$ 171.1372– $[C10H19O2]^+$ 157.1215– $[C9H17O2]^+$	Lipids (Fatty Acids and Conjugates)
Tetradecanoic acid	15.573	C14H28O2	544-63-8	228.2074 – $[C14H28O2]^+$ 129.0905 – $[C7H13O2]^+$ 73.0281 – $[C3H5O2]^+$	Lipids (Fatty Acids and Conjugates)
Palmitoleic acid	17.452	C16H30O2	373-49-9	236.2122 – $[C16H28O]^+$ 98.0710 – $[C6H10O]^+$ 69.0689 – $[C5H9]^+$	Lipids (Fatty Acids and Conjugates)
Palmitic acid	17.697	C16H32O2	57-10-3	227.1997 – $[C14H27O2]^+$ 129.0891 – $[C7H13O2]^+$ 73.0279 – $[C3H5O2]^+$	Lipids (Fatty Acids and Conjugates)

Cont. Table 1

Stearic acid	19.584	C18H36O2	57-11-4	284.2706 – $[C18H36O2]^+$ 129.0908 – $[C7H13O2]^+$ 73.0281 – $[C3H5O2]^+$	Lipids (Fatty Acids and Conjugates)
Glycidol stearate	22.476	C21H40O3	7460-84-6	297.2436 – $[C18H33O3]^+$ 98.0719 – $[C6H10O]^+$ 71.0848 – $[C5H11]^+$	Lipids (Fatty Acids and Conjugates)
Undecyl alcohol	11.25125	C11H24O	112-42-5	111.1157 – $[C8H15]^+$ 83.0844 – $[C6H11]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty Alcohols)
1-Hexadecanol 2-methyl	15.994	C17H36O	2490-48-4	111.1160 – $[C8H15]^+$ 97.1006 – $[C7H13]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty Alcohols)
Oleamide	21.328	C18H35NO	301-02-0	281.2679 – $[C18H35NO]^+$ 126.0914 – $[C7H12NO]^+$ 72.0438 – $[C3H6NO]^+$	Lipids (Fatty Amides)
Eicosenamide	23.01	C20H39NO	10436-08-5	126.0915 – $[C7H12NO]^+$ 72.0439 – $[C3H6NO]^+$ 309.2973 – $[C20H39NO]^+$	Lipids (Fatty Amides)
13-Docosenamide	24.572	C22H43NO	112-84-5	337.3338 – $[C22H43NO]^+$ 126.0916 – $[C7H12NO]^+$ 72.0440 – $[C3H6NO]^+$	Lipids (Fatty Amides)
Monopalmitin	22.57	C19H38O4	542-44-9	299.2577 – $[C18H35O3]^+$ 257.2462 – $[C16H33O2]^+$ 239.2366 – $[C16H31O]^+$	Lipids (Glycerolipids)
Monostearin	24.14	C21H42O4	123-94-4	327.2897 – $[C20H39O3]^+$ 267.2677 – $[C18H35O]^+$ 98.0723 – $[C6H10O]^+$	Lipids (Glycerolipids)
Hedione	14.408	C13H22O3	24851-98-7	83.0479– $[C5H7O]^+$ 97.0623– $[C6H9O]^+$ 226.1566 – $[C13H22O3]^+$	Lipids (Lineolic Acids and Deriatives-Jasmonic Acids)

Cont. Table 1

Sclareol	18.772	C20H36O2	515-03-7	272.2494 – $[C20H32]^+$ 121.0988 – $[C9H13]^+$ 95.0837 – $[C7H11]^+$	Lipids (Prenol Lipids-Diterpenes)
D-limonene	5.938	C10H16	5989-27-5	136.1227 – $[C10H16]^+$ 121.0992 – $[C9H13]^+$ 79.0524 – $[C6H7]^+$	Lipids (Prenol Lipids-Monoterpenes)
α-Cumyl alcohol	6.805	C10H14O	617-94-7	121.0644 – $[C8H9O]^+$ 103.0530 – $[C8H7]^+$ 91.0537 – $[C7H7]^+$	Lipids (Prenol Lipids-Monoterpenes)
Linalol	6.999	C10H18O	78-70-6	136.1218 – $[C10H16]^+$ 93.0679 – $[C7H9]^+$ 71.0844 – $[C5H4]^+$	Lipids (Prenol Lipids-Monoterpenes)
Camphor	7.835	C10H16O	464-48-2	152.1192 – $[C10H16]^+$ 137.0963 – $[C9H13O]^+$ 95.0852 – $[C7H11]$	Lipids (Prenol Lipids-Monoterpenes)
Camphol	8.160999	C10H18O	507-70-0	121.1000 – $[C9H13]^+$ 95.0853 – $[C7H11]^+$ 77.0381 – $[C6H5]^+$	Lipids (Prenol Lipids-Monoterpenes)
Levomenthol	8.251	C10H20O	2216-51-5	138.1379 – $[C10H18]^+$ 95.0837 – $[C7H11]^+$ 81.0683 – $[C6H9]^+$	Lipids (Prenol Lipids-Monoterpenes)
α-Terpineol	8.516	C10H18O	98-55-5	136.1244 – $[C10H16]^+$ 121.1007 – $[C9H13]^+$ 93.0695 – $[C7H9]^+$	Lipids (Prenol Lipids-Monoterpenes)
Squalene	24.835	C30H50	111-02-4	410.3907 – $[C30H50]^+$ 121.0994 – $[C9H13]^+$ 81.0686 – $[C6H9]^+$	Lipids (Prenol Lipids-Triterpenes)
Cholestadiene	25.554	C27H44	747-90-0	368.3437 – $[C27H44]^+$ 247.2412 – $[C18H31]^+$ 147.1141 – $[C11H15]^+$	Lipids (Steroids and Steroids Derivatives)

Cont. Table 1

Triethyl citrate	14.404	C12H20O7	77-93-0	203.0913 – $[C_9H_{15}O_5]^+$ 157.0496 – $[C_7H_9O_4]^+$ 83.0486 – $[C_5H_7O]^+$	Organic Acids and Derivatives (Carboxylic Acids and Derivatives)
Spiro [2,4]heptane-1,5-dimethyl-6-methylene	5.953	C10H16	62238-24-8	136.1225 – $[C_{10}H_{16}]^+$ 121.0990 – $[C_9H_3]^+$ 79.0524 – $[C_6H_7]^+$	Other Organic Compounds
2-propanol, 1-(2-butoxy-1-methylethoxy)-	9.21	C10H22O3	29911-28-2	59.0485 – $[C_3H_7O]^+$ 86.0715 – $[C_5H_{10}O]^+$ 103.0728 – $[C_5H_{11}O_2]^+$	Other Organic Compounds
1,2,4,6-Triisopropylphenol	12.685	C17H26O	08-07-34	220.1822 – $[C_{15}H_{24}O]^+$ 205.1584 – $[C_{14}H_{21}O]^+$ 77.0369 – $[C_6H_5]^+$	Other Organic Compounds
1,3-Heptadecyn-1-ol	15.11	C17H32O	56554-77-9	225.1826 – $[C_{14}H_{25}O_2]^+$ 81.0681 – $[C_6H_9]^+$ 67.0529 – $[C_5H_7]^+$	Other Organic Compounds
2,4-Diphenyl-4-methyl-2(E)-pentene	16.45	C18H20	22768-22-5	236.1567 – $[C_{18}H_{20}]^+$ 143.0809 – $[C_{11}H_{11}]^+$ 91.0513 – $[C_7H_7]^+$	Other Organic Compounds
7,9-Di-tert-butyl-oxaspiro(4,5)deca-6,9-diene-28-dione	17.182	C17H26O2	82304-66-3	175.1104 – $[C_{12}H_{15}O]^+$ 133.0638 – $[C_9H_9O]^+$ 77.0369 – $[C_6H_5]^+$	Other Organic Compounds
10,18-Bisnorabieta-8,11,13-triene	17.999014	C18H26	32624-67-2	242.2007 – $[C_{18}H_{26}]^+$ 227.1790 – $[C_{17}H_{23}]^+$ 143.0864 – $[C_{11}H_{11}O]^+$	Other Organic Compounds
Phenol-2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)	18.709	C23H32O	34624-81-2	324.2438 – $[C_{23}H_{32}O]^+$ 309.2212 – $[C_{22}H_{29}O]^+$ 119.0836 – $[C_9H_{11}]^+$	Other Organic Compounds

Cont. Table 1

Phenol 2,2'-methylenebis[6-(1,1-dimethyl)-4-methyl	21.768	C23H32O2	119-47-1	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other Organic Compounds
Phenol 2,4-bis(1-methyl-1-phenylethyl)-	22.522	C24H26O	2772-45-4	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other Organic Compounds
2,4-Bis(dimethylbenzyl)-6-t-butylphenol	22.57	C28H34O	244080-16-8	386.2617 – $[C28H34O]^+$ 371.2370 – $[C27H31O]^+$ 293.1897 – $[C21H25O]^+$	Other Organic Compounds
1H-indene 1-hexadecyl-2,3-dyhidro	23.928	C25H42	55334-29-7	117.0351 – $[C9H9]^+$ 130.0427 – $[C10H10]^+$ 154.1345 – $[C11H22]^+$	Other Organic Compounds
Surfynol A	11.398	C14H26O2	126-86-3	151.1132 – $[C10H15O]^+$ 133.1021 – $[C10H13]^+$ 91.0545 – $[C7H7]^+$	Surfactant

Other family detected in EBC was that of aromatic homomonocyclic compounds, among which stand benzoic acid derivatives such as benzoic acid methyl ester or benzoic acid 4-ethoxy ethyl ester, and cinnamic acid derivatives such as octylmethoxycinnamate or di-t-butyl-4-hydroxycinnamic acid. Some of these compounds have previously been found in breath, as is the case of benzoic acid 4-ethoxy ethyl ester after analysis of breath by GC-MS [21]. Two of the intermediates in the pathway of toluene metabolism as are benzyl alcohol and *p*-cresol (related to hippuric acid) had been identified by Pierce *et al.* [39], who analyzed exhalation and excretion rates of toluene and metabolites, demonstrating the relationship among these compounds, for some of which Supplementary Figure 7A shows the fragmentation pattern.

Indole, an aromatic heteropolycyclic compound resulting from the catabolism of tryptophan, had also been previously identified in breath [40]. Evidence exists in patients with liver failure that the degradation of this aromatic amino acid is impaired leading to higher levels of free tryptophan [41]. Supplementary Figure 7B shows the fragmentation pattern of indole detected in EBC.

The resting compounds detected in EBC had not been previously identified in this biofluid, although other with similar structures had been reported. This is the case with phenol-2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl) or 1H-indene, 1-hexadecyl-2,3-dyhidro [21] (Supplementary Figure 7C). Other organic compounds such as 1,2,4,6-triisopropylphenol; 10,18-bisnorabieta-8,11,13-triene; 2,4-bis(dimethylbenzyl)-6-t-butylphenol; 2-propanol, 1-(2-butoxy-1-methylethoxy)-; phenol, 2,2'-methylenebis[6-(1,1-dimethyl)-4-methyl; phenol, 2,4-bis(1-methyl-1phenylethyl)-; and spiro [2,4] heptane, 1,5-dimethyl-6-methylene can be added to the list of compounds identified in EBC.

Conclusions

A method for untargeted metabolomics analysis of EBC based on GC-TOF/MS profiling in high resolution mode has been developed. Two sample

preparation alternatives such as LLE and SPE were compared in terms of extraction coverage for the analysis of this biofluid. LLE provided the widest information on the composition of EBC with tentative identification of 51 compounds. Furthermore, four normalization strategies have been compared and the normalization to MSTUS was selected as the optimum to minimize data scattering. Among the identified compounds are fatty acids and other derivatives such as methyl esters and amide derivatives, which are not considered as volatile components. Other representative family of compounds found in EBC was that of prenol lipids such as terpenes, which are volatile components. Therefore, the potential of EBC as biofluid for clinical analysis of both volatile and non-volatile compounds as compared to other samples such as EBV, only useful to analyze volatile compounds, has been highlighted.

Acknowledgements

The Spanish Ministerio de Economía y Competitividad (MINECO) and FEDER programme are gratefully acknowledged for financial support (Project "Development of methods for early cancer detection, December 29, 2011") that also covered A. Peralbo-Molina and M. Calderón-Santiago research contracts. F. Priego-Capote is also grateful to the Ministerio de Ciencia e Innovación (MICINN) for a Ramón y Cajal Contract (RYC-2009-03921). Reina Sofía Hospital (Córdoba, Spain) is also grateful for donors of EBC samples.

References

[1] 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.

[2] L. Zhou, R.W. Beuerman, Tear analysis in ocular surface diseases, Prog. Retin. Eye Res. 31 (2012) 527–50.

[3] G. Hedlin, J. Konradsen, A. Bush, An update on paediatric asthma, Eur. Respir. Rev. 21 (2012) 175–185.

[4] 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.

[5] B. Buszewski, K. Martyna, T. Ligor, A. Amann, Human exhaled air analytics: biomarkers of diseases, Biomed. Chromatogr. 21 (2007) 553–566.

[6] W. Cao, Y. Duan, Breath analysis: potential for clinical diagnosis and exposure assessment, Clin. Chem. 52 (2006) 800–811.

[7] D. Poli, P. Carbognani, M. Corradi, M. Goldoni, O. Acampa, B. Balbi, *et al.*, Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross sectional and nested short-term follow-up study, Respir. Res. 6 (2005) 71.

[8] A. Ulanowska, T. Kowalkowski, E. Trawińska, B. Buszewski, The application of statistical methods using VOCs to identify patients with lung cancer, J. Breath Res. 5 (2011) e046008.

[9] 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.

[10] 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.

[11] K. Matsumura, M. Opiekun, H. Oka, A. Vachani, S.M. Albelda, K. Yamazaki, *et al.*, Urinary volatile compounds as biomarkers for lung cancer: a proof of principle study using odor signatures in mouse models of lung cancer, PLoS One 5 (2010) e8819.

[12] R. Ehmann, E. Boedeker, U. Friedrich, J. Sagert, J. Dippon, G. Friedel, *et al.*, Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon, Eur. Respir. J. 39 (2012) 669-676.

[13] C. Costa, C. Bucca, M. Bergallo, P. Solidoro, G. Rolla, R. Cavallo, Unsuitability of exhaled breath condensate for the detection of herpesviruses DNA in the respiratory tract, J. Virol. Methods 173 (2011) 384-386.

[14] 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.

[15] J. Hunt, Exhaled breath condensate: An evolving tool for noninvasive evaluation of lung disease, J. Allergy Clin. Immunol. 110 (2002) 28-34.

[16] A. Mazzatenta, C. Di Giulio, M. Pokorski, Pathologies currently identified by exhaled biomarkers, Respir. Physiol. Neurobiol. 187 (2013) 128-134.

[17] 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.

[18] A. Bajtarevic, C. Ager, M. Pienz, M. Klieber, K. Schwarz, M. Ligor, *et al.*, Noninvasive detection of lung cancer by analysis of exhaled breath, BMC Cancer 9 (2009) e348.

[19] C. Deng, X. Zhang, N. Li, Investigation of volatile biomarkers in lung cancer blood using solid-phase microextraction and capillary gas chromatography-mass spectrometry, J. Chromatogr. B 808 (2004) 269-277.

[20] A.N. Martin, G.R. Farquar, A.D. Jones, M. Frank, Human breath analysis: methods for sample collection and reduction of localized background effects, Anal. Bioanal. Chem. 396 (2010) 739-750.

[21] M. Phillips, Variation in volatile organic compounds in the breath of normal humans, J. Chromatogr. B 729 (1999) 75–88.

[22] V. de Broucker, S.M. Hassoun, S. Hulo, N. Chérot-Kornobis, R. Nevière, R. Matran, *et al.*, Non-invasive collection of exhaled breath condensate in rats: evaluation of pH, H2O2 and NOx in lipopolysaccharide-induced acute lung injury, Vet. J. 194 (2012) 222–228.

[23] G. de Laurentiis, D. Paris, D. Melck, M. Maniscalco, S. Marsico, G. Corso, *et al.*, Metabonomic analysis of exhaled breath condensate in adults by nuclear magnetic resonance spectroscopy, Eur. Respir. J. 32 (2008) 1175–1183.

[24] L. Pauling, B. Robinson, R. Teranishi, P. Cary, Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography, Proc. Natl. Acad. Sci. U. S. A. 68 (1971) 2374–2376.

[25] S.M. Gordon, J.P. Szldon, B.K. Krotoszynski, R.D. Gibbons, J.O. Neill, Volatile organic compounds in exhaled air from patients with lung cancer, Clin. Chem. 1282 (1985) 1278–1282.

[26] H.J. O'Neill, S.M. Gordon, M.H. O'Neill, R.D. Gibbons, J.P. Szidon, A computerized classification technique for screening for the presence of breath biomarkers in lung cancer, Clin. Chem. 34 (1988) 1613–1618.

[27] G. Preti, J.N. Labows, J.G. Kostelc, S. Aldinger, R. Daniele, Analysis of lung air from patients with bronchogenic carcinoma and controls using gas chromatography-mass spectrometry, J. Chromatogr. B 432 (1988) 1–11.

[28] O. Quehenberger, A.M. Armando, A.H. Brown, S.B. Milne, D.S. Myers, A.H. Merrill, *et al.*, Lipidomics reveals a remarkable diversity of lipids in human plasma, J. Lipid Res. 51 (2010) 3299–3305.

[29] P. Martínez-Lozano, J. Fernández de la Mora, Direct analysis of fatty acid vapors in breath by electrospray ionization and atmospheric pressure ionization-mass spectrometry, Anal. Chem. 80 (2008) 8210–8215.

[30] 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.

[31] Å. Henningsson, I. Björck, M. Nyman, Short-chain fatty acid formation at fermentation of indigestible carbohydrates, Food Nutr. Res. 45 (2001) 165-168.

[32] W. Vautz, J. Nolte, R. Fobbe, J.I. Baumbach, Breath analysis-performance and potential of ion mobility spectrometry, J. Breath Res. 3 (2009) e036004.

[33] J.J.B.N. Van Berkel, J.W. Dallinga, G.M. Möller, R.W.L. Godschalk, E.J. Moonen, E.F.M. Wouters, *et al.*, A profile of volatile organic compounds in breath discriminates COPD patients from controls, Respir. Med. 104 (2010) 557-563.

[34] E. Hayet, B. Fatma, I. Souhir, F.A. Waheb, K. Abderaouf, A. Mahjoub, *et al.*, Antibacterial and cytotoxic activity of the acetone extract of the flowers of Salvia sclarea and some natural products, Pak. J. Pharm. Sci. 20 (2007) 146-148.

[35] N.R. Patel, S. Hatziantoniou, A. Georgopoulos, C. Demetzos, V.P. Torchilin, V. Weissig, *et al.*, Mitochondria-targeted liposomes improve the apoptotic and cytotoxic action of sclareol, J. Liposome Res. 20 (2010) 244-249.

[36] S. van den Brûle, A. Müller, A.J. Fleming, C.C. Smart, The ABC transporter SpTUR2 confers resistance to the antifungal diterpene sclareol, Plant J. 30 (2002) 649-662.

[37] R.A. Stein, J.F. Mead, Small hydrocarbons formed by the peroxidation of squalene, Chem. Phys. Lipids 46 (1988) 117-120.

[38] K. Saito, Y. Shirasago, T. Suzuki, H. Aizaki, K. Hanada, T. Wakita, *et al.*, Targeting cellular squalene synthase, an enzyme essential for

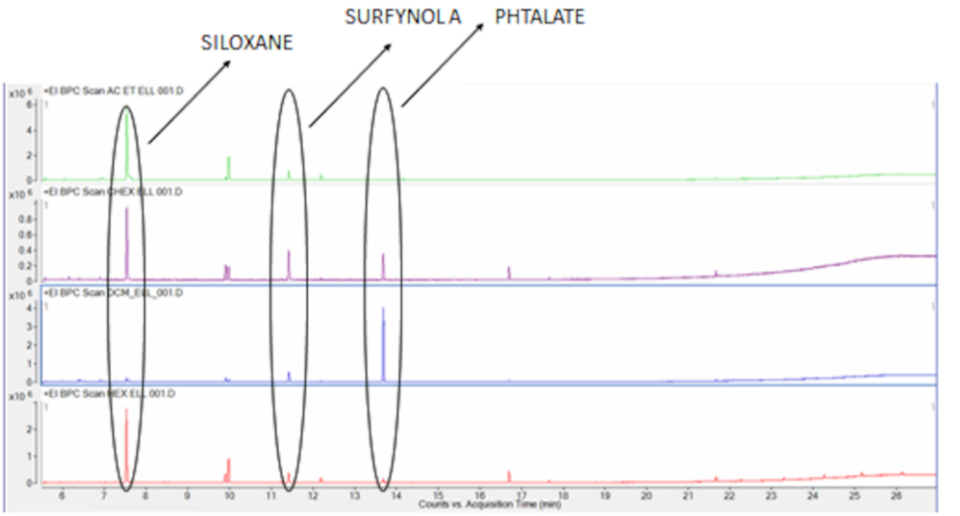
cholesterol biosynthesis, is a potential antiviral strategy against hepatitis C virus. J. Virol. 89 (2015) 2220-2232.

[39] C.H. Pierce, Y. Chen, R.L. Dills, D.A. Kalman, M.S. Morgan, Toluene metabolites as biological indicators of exposure, Toxicol. Lett. 129 (2002) 65–76.

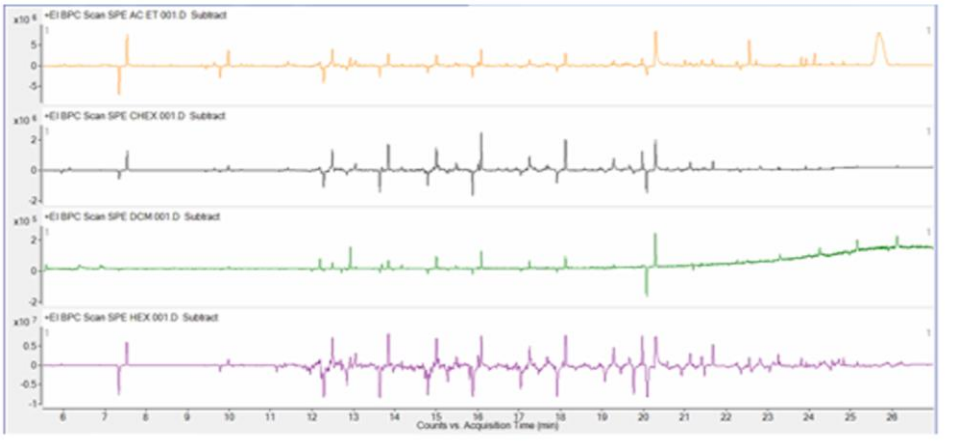
[40] S. Van den Velde, F. Nevens, P. Van Hee, D. van Steenberghe, M. Quirynen, GC-MS analysis of breath odor compounds in liver patients, J. Chromatogr. B 875 (2008) 344-348.

[41] C.H.C. Dejong, M.C.G. van de Poll, P.B. Soeters, R. Jalan, S.W.M. Olde Damink, Aromatic amino acid metabolism during liver failure, J. Nutr. 137 (2007) 1579S–1585.

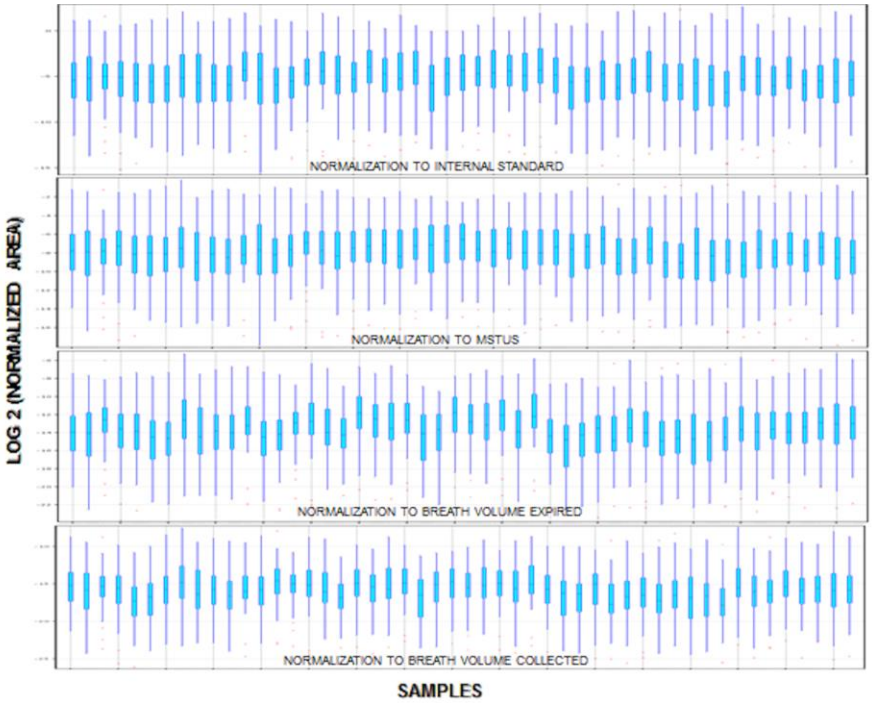
Supplementary material



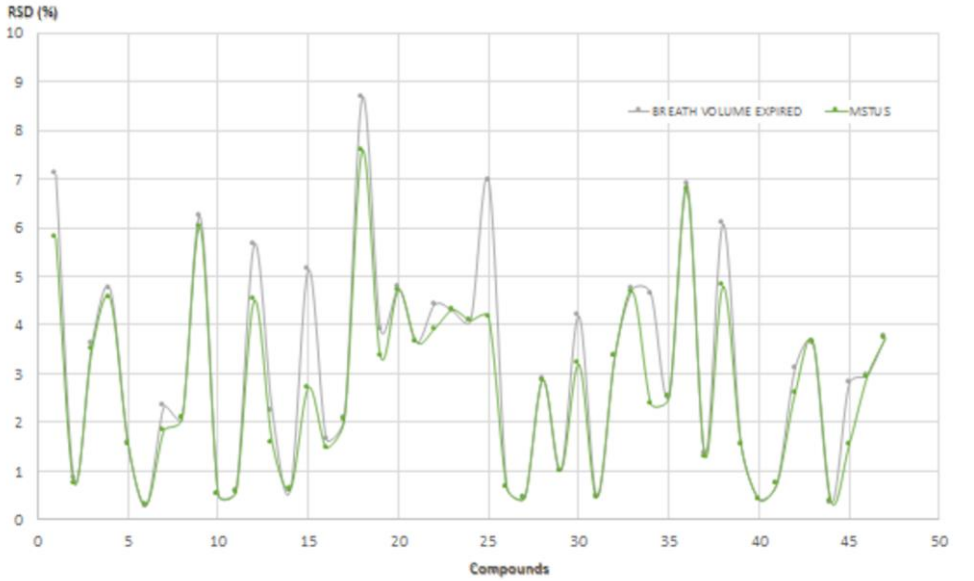
Supplementary Figure 1. BPCs obtained by analysis of EBC using four different extractants.



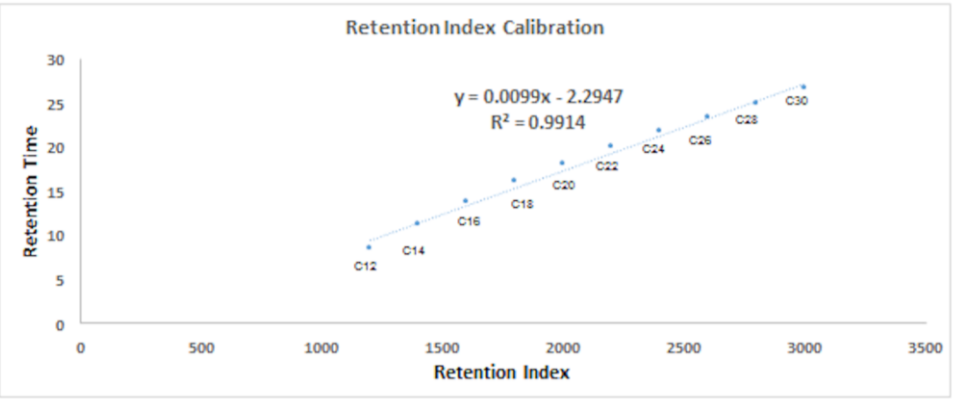
Supplementary Figure 2. BPCs obtained by analysis of each eluate from the SPE procedure after subtracting the blank chromatogram signal.



Supplementary Figure 3. Box and whisker plots showing the variability for each sample in the cohort with the four normalization approaches.

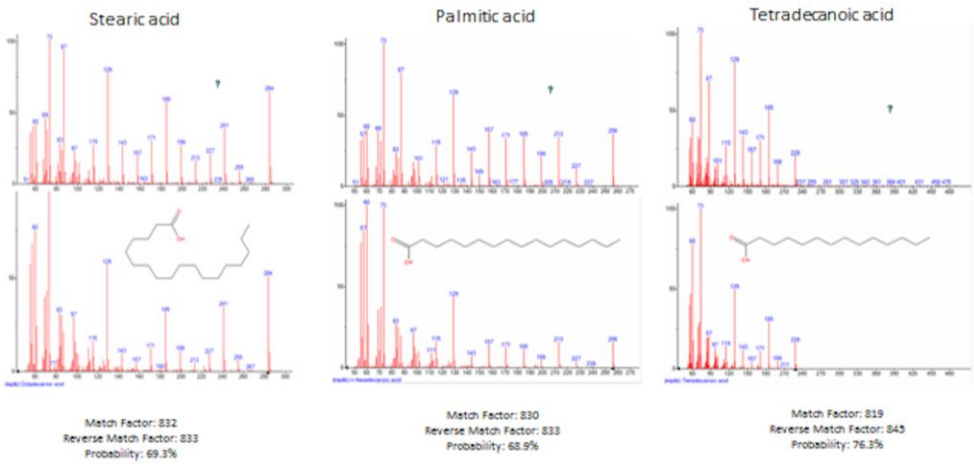


Supplementary Figure 4. Comparison of the methodological variability between replicates for the best two strategies of normalization tested (MSTUS and expired air combined with IS signal).

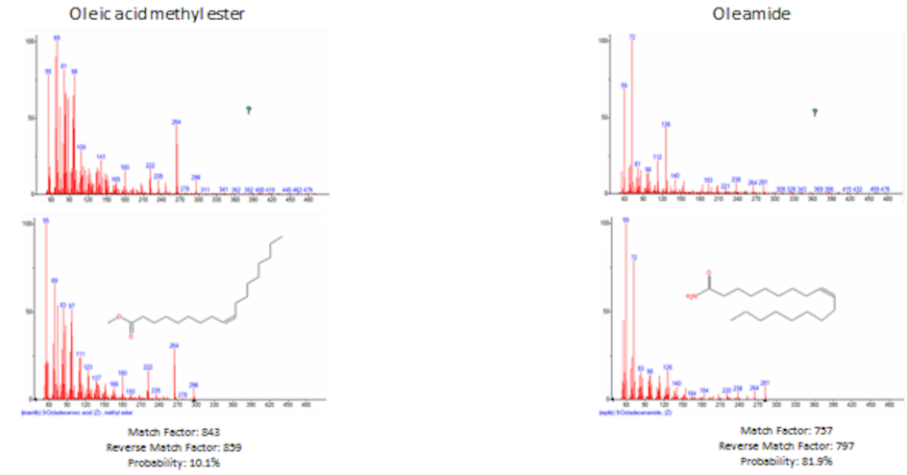


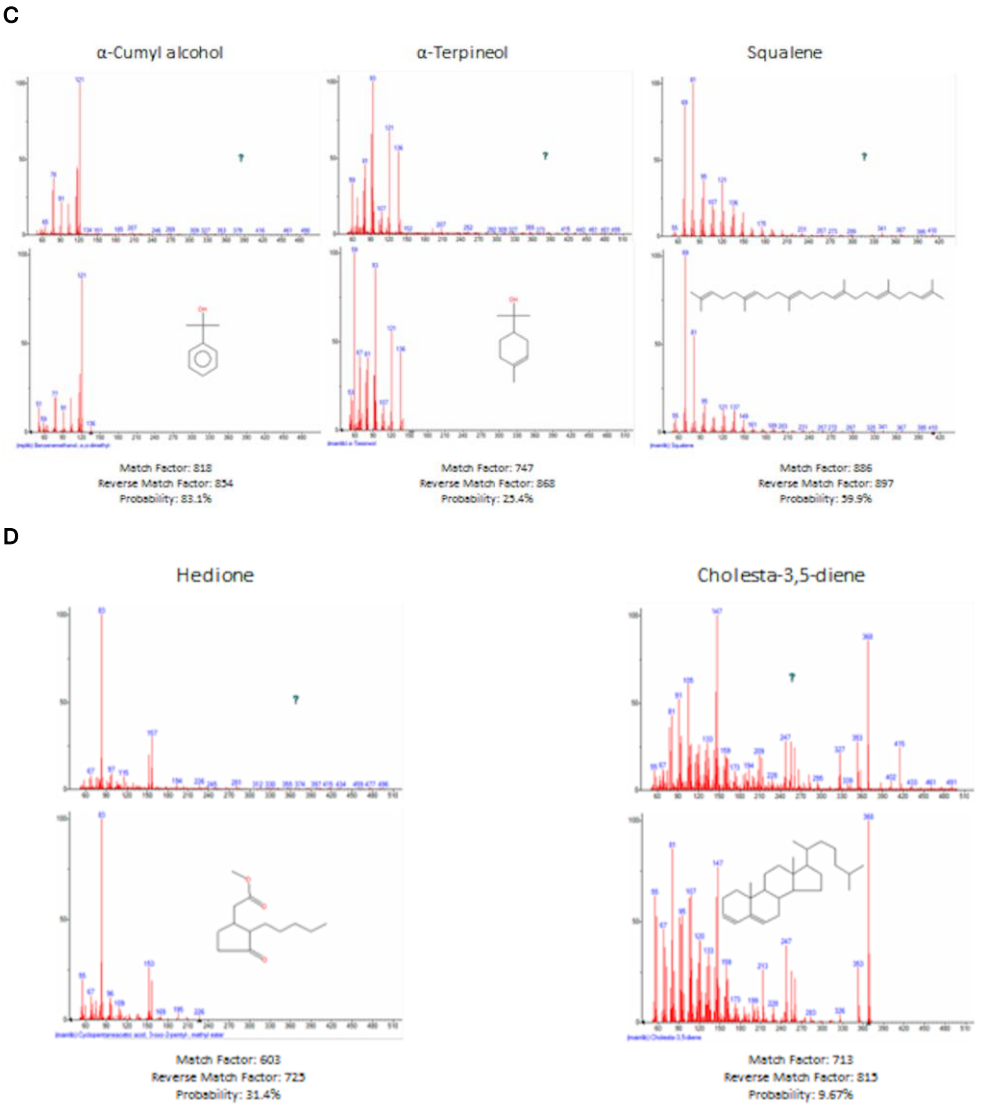
Supplementary Figure 5. RI calibration line obtained by analysis of the alkane standard mixture.

A


Stearic acid
Palmitic acid
Tetradecanoic acid
Match Factor: 832
Reverse Match Factor: 833
Probability: 69.3%
Match Factor: 830
Reverse Match Factor: 833
Probability: 68.9%
Match Factor: 819
Reverse Match Factor: 845
Probability: 76.3%

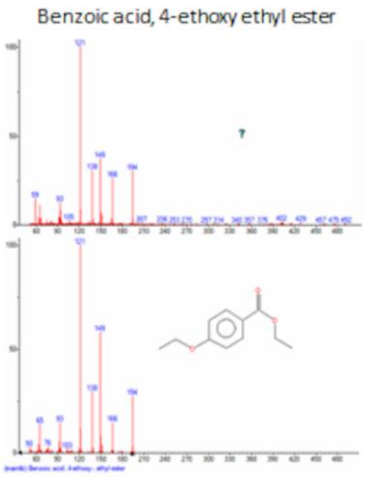
B


Oleic acid methyl ester
Oleamide
Match Factor: 843
Reverse Match Factor: 859
Probability: 10.1%
Match Factor: 757
Reverse Match Factor: 797
Probability: 81.9%

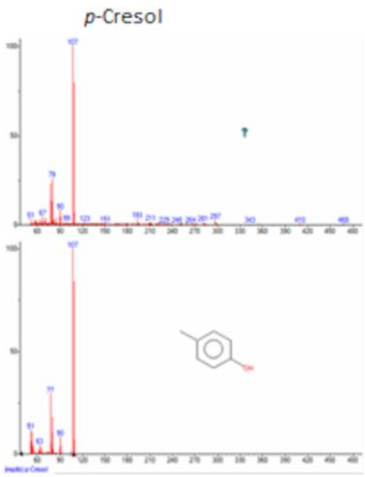


Supplementary Figure 6. Mass spectra with in-source fragmentation pattern for (A) fatty acids, (B) oleic acid derivatives, (C) prenol lipids, (D) other lipids.

A

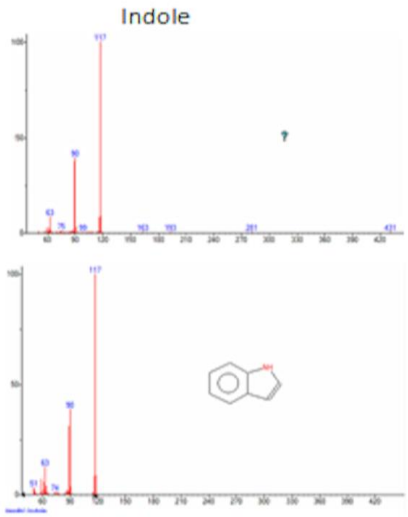


Match Factor: 720
Reverse Match Factor: 830
Probability: 67%



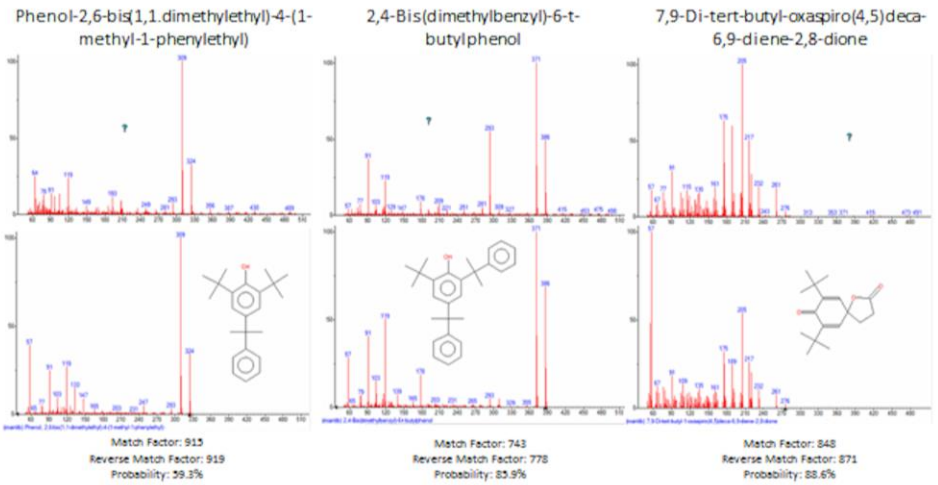
Match Factor: 736
Reverse Match Factor: 844
Probability: 27.6%

B



Match Factor: 913
Reverse Match Factor: 919
Probability: 59.3%

C



Supplementary Figure 7. Mass spectra with in-source fragmentation pattern for (A) intermediates in the pathway of toluene metabolism, (B) indole, (C) other compounds.

Supplementary Table 1. Compounds in each extract from LLE sample preparation identified by GC-TOF/MS.

Compound Name	RT	LLE			
		HEX	CHEX	DCM	ETH AC
Eucalyptol	5.995	✓			
Indole	9.904	✓	✓	✓	✓
Benzoic acid 4-ethoxy-ethyl ester	12.923	✓			
Benzoic acid methyl ester	6.951	✓			
3,5-Di-t-butyl-4-hydroxycinnamic acid	20.155	✓	✓	✓	✓
Octinoxate	21.005	✓			
***p*-Cresol**	6.536	✓	✓		✓
Benzyl alcohol	5.963	✓	✓	✓	✓
Dodecanoic acid 1-methylethyl ester	14.97	✓			
Palmitic acid methyl ester	17.342	✓			
Oleic acid methyl ester	19.041	✓			
Methyl stearate	19.28	✓	✓	✓	

Octanoic acid	7.939	✓	✓		✓
Nonanoic acid	9.375	✓			
Dodecanoic acid	13.268	✓			
Tetradecanoic acid	15.573	✓			✓
Palmitoleic acid	17.452	✓			
Palmitic acid	17.697	✓			
Stearic acid	19.584	✓			
Glycidol stearate	22.476	✓	✓	✓	✓
Undecyl alcohol	11.25125	✓	✓	✓	✓
1-Hexadecanol 2-methyl	15.994	✓			
Oleamide	21.328	✓	✓	✓	✓
Eicosenamide	23.01	✓	✓	✓	
13-Docosenamide	24.572	✓	✓	✓	✓
Monopalmitin	22.57	✓	✓	✓	✓
Monostearin	24.14	✓	✓	✓	✓
Hedione	14.408	✓	✓	✓	
Sclareol	18.772	✓	✓	✓	✓
D-Limonene	5.938	✓	✓	✓	✓
α-Cumyl alcohol	6.805	✓			
Linalol	6.999	✓	✓		✓
Camphor	7.835	✓	✓		✓
Camphol	8.160999	✓			✓
Levomenthol	8.251	✓			✓
α-terpineol	8.516	✓	✓		
Squalene	24.835	✓	✓	✓	✓
Cholestadiene	25.554	✓	✓	✓	✓
Triethyl citrate	14.404	✓	✓	✓	✓
Spiro [2,4]heptane 1,5-dimethyl-6-methylene	5.953	✓	✓	✓	✓
2-Propanol 1-(2-butoxy-1-methylethoxy)-	9.21	✓	✓		
1,2,4,6-Triisopropylphenol	12.685	✓			
1,3-Heptadecyn-1-ol	15.11	✓			

2,4-Diphenyl-4-methyl-2(E)-pentene	16.45	✓			
7,9-Di-tert-butyl-oxaspiro(4,5)deca-6,9-diene-2,8-dione	17.182	✓			
10,18-Bisnorabieta-8,11,13-triene	17.999014	✓	✓	✓	✓
Phenol-2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)	18.709	✓		✓	✓
Phenol 2,2'-methylenebis[6-(1,1-dimethyl)-4-methyl	21.768	✓			
Phenol 2,4-bis(1-methyl-1-phenylethyl)-	22.522	✓	✓	✓	✓
2,4-Bis(dimethylbenzyl)-6-t-butylphenol	22.57	✓	✓	✓	
1H-Indene 1-hexadecyl-2,3-dyhidro	23.928	✓			

Supplementary Table 2. Compounds in each extract from SPE sample preparation identified by GC-TOF/MS.

Compound Name	**RT**	**SPE**			
		HEX	**CHEX**	**DCM**	**ETH AC**
Eucalyptol	5.995	✓	✓		✓
Indole	9.904	✓	✓		
Benzoic acid 4-ethoxy-ethyl ester	12.923	✓			✓
Benzoic acid methyl ester	6.951	✓	✓		✓
3,5-Di-t-butyl-4-hydroxycinnamic acid	20.155	✓	✓		
Octinoxate	21.005				
***p*-Cresol**	6.536	✓	✓		✓
Benzyl alcohol	5.963	✓		✓	✓
Dodecanoic acid 1-methylethyl ester	14.97	✓	✓		
Palmitic acid methyl ester	17.342	✓	✓		
Oleic acid methyl ester	19.041	✓	✓	✓	✓
Methyl stearate	19.28	✓	✓	✓	✓
Octanoic acid	7.939	✓	✓		
Nonanoic acid	9.375	✓	✓		
Dodecanoic acid	13.268	✓			✓
Tetradecanoic acid	15.573	✓	✓		✓

Palmitoleic acid	17.452	✓	✓	✓	✓
Palmitic acid	17.697		✓		
Stearic acid	19.584	✓	✓		
Glycidol stearate	22.476	✓	✓	✓	✓
Undecyl alcohol	11.25125	✓	✓		
1-Hexadecanol 2-methyl	15.994	✓	✓	✓	
Oleamide	21.328		✓	✓	✓
Eicosenamide	23.01	✓	✓	✓	✓
13-Docosenamide	24.572				✓
Monopalmitin	22.57				
Monostearin	24.14				
Hedione	14.408	✓	✓		
Sclareol	18.772	✓	✓		
D-Limonene	5.938				
α-Cumyl alcohol	6.805	✓	✓		
Linalol	6.999	✓	✓		
Camphor	7.835	✓	✓		
Camphol	8.160999	✓	✓		
Levomenthol	8.251		✓	✓	✓
α-Terpineol	8.516	✓	✓		
Squalene	24.835				
Cholestadiene	25.554		✓		
Triethyl citrate	14.404	✓			
Spiro [2,4]heptane 1,5-dimethyl-6-methylene	5.953	✓	✓		
2-Propanol 1-(2-butoxy-1-methylethoxy)-	9.21	✓	✓		
1,2,4,6-Triisopropylphenol	12.685				✓
1,3-Heptadecyn-1-ol	15.11	✓	✓		
2,4-Diphenyl-4-methyl-2(E)-pentene	16.45	✓	✓	✓	✓
7,9-Di-tert-butyl-oxaspiro(4,5)deca-6,9-diene-2,8-dione	17.182	✓	✓		✓
10,18-Bisnorabieta-8,11,13-triene	17.999014	✓	✓		✓

Phenol-2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)	18.709	✓	✓		✓
Phenol 2,2'-methylenebis[6-(1,1-dimethyl)-4-methyl	21.768	✓	✓		✓
Phenol 2,4-bis(1-methyl-1-phenylethyl)-	22.522				
2,4-Bis(dimethylbenzyl)-6-t-butylphenol	22.57	✓	✓		✓
1H-Indene 1-hexadecyl-2,3-dyhidro	23.928	✓	✓		✓

Capítulo VII:

Metabolomics analysis of exhaled breath condensate for discrimination between lung cancer patients and risk factor individuals

Journal of Breath Research

Volume 10 (1), 11 February 2016,

Page 016011

IOPscience

Metabolomics analysis of exhaled breath condensate for discrimination between lung cancer patients and risk factor individuals

Ángela Peralbo-Molina[a,b], Mónica Calderón-Santiago[a,b], Feliciano Priego-Capote[a,b*], Bernabé Jurado-Gámez[b,c], María Dolores Luque de Castro[a,b*]

[a]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14014, Córdoba, Spain*

[b]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, E-14014, Córdoba, Spain*

[c] *Department of Respiratory Medicine, Reina Sofia University Hospital, University of Córdoba, 14004, Córdoba, Spain*

Metabolomics analysis of exhaled breath condensate for discrimination between lung cancer patients and risk factor individuals

Ángela Peralbo Molina, Mónica Calderón-Santiago, Feliciano Priego-Capote, Bernabé Jurado-Gámez, María Dolores Luque de Castro

Abstract

The search for new clinical tests aimed at diagnosing chronic respiratory diseases is a current research line supported on the lack of efficient screening tools and the severity of some of these pathologies, to which alternative biological samples can open a door to new screening tools. A promising biofluid scarcely used for diagnostic purposes is exhaled breath condensate (EBC), the composition of which has been poorly studied. In this research, untargeted analysis of EBC by GC-TOF/MS has been applied to a cohort of patients with lung cancer (n= 48), risk factor individuals (active smokers and ex-smokers, n= 130) and control healthy individuals (no smokers and without respiratory diseases, n= 61). Identical protocol was applied to the two EBC fractions provided by the sampling device (upper and central airways and distal airway) from each individual, which allowed detecting compositional differences between the two EBC fractions. Tentative compounds that contribute to discriminate among the three groups were identified, and a relevant role for lipids such as monoacylglycerols and squalene was found. These results could support the capability of metabolomics to go inside the study of lung cancer.

Keywords: lung cancer, exhaled breath condensate, metabolomics, GC-TOF/MS, untargeted metabolomics

1. Introduction

Chronic respiratory disease is a generic term that encompasses long-term pathological conditions that affect the airways and other lung structures involved in respiration. Among these chronic diseases it is worth emphasizing two pathologies: chronic obstructive pulmonary disease (COPD) and lung cancer, the latter belonging to the list of leading causes of disease-related death in most developed countries [1,2]. The high death rate caused by the most severe respiratory diseases is associated to the low efficiency of screening methods to detect them at early stages [3]. The search for diagnostic tests targeted at early detection of chronic respiratory diseases has intensified in the last decades. Low-dose computed tomography (CT) is at present the most efficient test for lung cancer screening, as it reduces mortality of high-risk population about 20% as compared to chest radiography arm, according to the National Lung Screening Trial [4]. CT has also proved to be useful for COPD evaluation [5].

Apart from tests based on imaging techniques, different “omics” studies have focused on the development of screening tools to diagnose respiratory diseases, mainly searching for potential biomarkers either in tissue or in biofluids. As the use of invasive sampling would impair a screening methodology in terms of time consuming and patient welfare, biofluids sampled in a non-invasive manner are gaining popularity. This is the case with urine or sputum and, more recently, with exhaled breath or sweat [6]. A recent study on sweat to discriminate metabolite patterns for lung cancer screening has resulted in an optimum panel of 5 compounds [7]. The panel provides 80% specificity and 79% sensitivity, which leads to false positive and negative rates around 20% [8]. Concerning exhaled breath, levels of nitrite in epithelial fluid have demonstrated discrimination capability as biomarker since an increase of this anion seems to be associated with cancer [9]. Individual cytokines have also been investigated in exhaled breath from lung cancer patients using enzyme immunoassays (EIA), but the commercially available immunoassays lack of the sensitivity demanded for early detection [10–13]. Other biomarkers (e.g., hydrogen peroxide [14], 8-

isoprostane [14] or pH level [15]) allowed discrimination between COPD patients and healthy controls; however, they seemed not be able to discriminate between controls+COPD and patients with lung cancer, as Antus *et al.* stood out in a recent research [16]. Targeting at VOCs in exhaled breath vapor, Phillips *et al.* developed a mathematical model constituted by 22 compounds (mainly alkanes and benzene derivatives) as primary lung cancer biomarkers. The model reported sensitivity and selectivity values of 71.7% and 66.7%, respectively, even for advanced stages of the disease [17]. This study was later improved by including in the model a set of oxidative stress products excreted in breath [18], thus achieving sensitivity and specificity of 85.1% and 80.5%, respectively. Further on, another study concluded that a two-min breath test can predict pulmonary cancer regardless of histology, stage of the disease or tobacco smoking, with 84.6% sensitivity and 80% specificity [19]. A recent study has provided a 23 VOCs panel in exhaled breath with capability to distinguish between lung cancer patients from controls with 96.5% sensitivity and 97.5% specificity [20]. However, the criteria for classification of endogenous compounds were not restrictive and some of the compounds included in the panel were also detected in the air of the sampling room.

Apart from the direct analysis of exhaled breath using the appropriate interface, this sample can be analyzed after condensation. Exhaled breath condensate (EBC) is the exhaled gas that condensates as a liquid solution by cooling, the analysis of which should allow gaining insight into the composition of extracellular lining fluid and soluble exhaled gases. Despite the main component of EBC is water vapor, hundreds of different components, ranging from small inorganic ions through large organic molecules to peptides, proteins, surfactants and macromolecules at trace concentrations can be found in this sample [21–26].

Regarding the EBC sampling protocol, there are commercial devices that allow collection of two EBC fractions by separating the air coming from the upper and central airways (upper breath, UA) and the distal airway (lower breath, DA). Despite UA is believed to contain compounds without clinical relevance [21], the differences between both types of samples have not been clearly established. In

this research, fractionation of EBC was selected as strategy to compare the composition profiles of UA and DA to find metabolomics differences in lung cancer patients as compared to a risk factor group composed by active smokers. A third group of non-smoker healthy individuals was used as control group. The composition profiles of UA and DA for the three groups were obtained using gas chromatography coupled to mass spectrometry in high-resolution mode (GC-TOF/MS).

2. Experimental

2.1. Cohort selected for the study

The EBC samples were collected from 239 individuals early in the morning before breakfast and stored at -80 ºC in the biorepository of the Reina Sofía Hospital (Córdoba, Spain). All individuals were recruited by the Department of Respiratory Medicine. The cohort included 48 patients diagnosed with lung cancer confirmed cytohistologically after clinical tests based on bronchoscopy, fine-needle biopsy, or video-assisted thoracoscopy. The patients were in the interval 63±7 years, and 94% of them were males. The most frequent diagnostic was squamous cell carcinoma (15 patients, 31.25%), followed by adenocarcinoma (13 patients, 27%), small cell carcinoma (7 patients, 14.5%), and large cell carcinoma (6 patients, 12.5%). Seven individuals (14%) were diagnosed with non-small cell lung cancer without histological classification.

The risk factor group was formed by 130 individuals, all of them active smokers or ex-smokers (<5 years) with a cumulative consumption >20 packages/year. It is worth noting that 47 individuals pertaining to the risk factor cohort were diagnosed with COPD by spirometry (FVC/FEV1 ratio < 0.7). Two reasons supported the inclusion of COPD patients in the risk factor group: firstly, smoking is considered the most important risk factor to develop COPD or lung cancer; secondly, the increase of oxidative stress and the presence of inflammatory cells infiltrated in COPD and lung cancer are common ways to

explain theoretically lung damage. The risk factor group, clinically followed at least for one year, was characterized by ages within the interval 61±8, 82.4% male individuals. The existence of lung cancer in this group was discarded through CT and bronchoscopy. The control group included 61 healthy individuals (age within the interval 60±9 years, 87% male individuals), no active nor passive smokers, without clinical symptoms and with a normal profile set by thorax radiography.

Criteria for exclusion of patients were: a) coexistence of extrapulmonary tumoral pathology or treatment with cytostatic drugs for a different neoplasm; b) diagnostic of neoplasm in the last five years; c) unjustified weight loss (≥7 kg) in the last year; d) severe disorder of any organ with negative influence on the prognostic or avoiding to apply the protocol including IV grade cardiac insufficiency according to the New York Heart Association, advanced hepatic cirrhosis, V stage renal insufficiency with substitutive treatment by hemodialysis or peritoneal dialysis, and lung disease diagnostic not related to smoking, including interstitial pneumopathy, neumonia, tuberculosis, etc. [8].

All experiments were carried out in accordance with the ethical principles of human medical research (World Medical Association, Helsinki Declaration). The ethical review board of Reina Sofía Hospital (Córdoba, Spain) approved and supervised the clinical study.

2.2. EBC collection procedure

The ECOScreen2 device used for sampling directly collects and condenses the EBC in disposable polyethylene bags at -20 ºC. This device offers controlled collection of EBC into two separate bags for physical separation between the air exhaled from the UA and DA [21]. Different valves in the EBC sample allow separating inspiration from expiration and fractioning the exhaled volume according to a threshold volume into the two cavities: UA and DA fractions. This configuration makes saliva contamination highly unlikely [27]. The main modification of the sampler was the insertion of a commercial protection

filter from Scharlab (Barcelona, Spain) over the inlet air valve to avoid the entrance of exogenous organic compounds and particles from the room atmosphere. This filter was periodically changed to avoid saturation.

Tidal breathing and a nose-clip were used during 15 min, time required to collect an average EBC volume of 1.5 ml from the DA and 1 ml from the UA. Only 101 individuals were able to provide enough volume of the two fractions for analysis. The samples were divided into 100 µl aliquots and the vials stored at -80 ºC until analysis. All samples were analyzed within 3 months after collection.

2.3.. Reagents

Hexane TraceSELECT® grade from Sigma-Aldrich (St. Louis, USA) as organic solvent for sample preparation, and an alkane standard mixture (from C10 to C40), also from Sigma-Aldrich, for GC separation tests to establish the retention index (RI) calibration were used. Deionized water (18 mΩ·cm) from a Millipore Milli-Q water purification system was also used.

2.4. Instruments and apparatus

An ECOScreen2 device (FILT Thorax-und LungenDiagnostik GmbH, Berlin, Germany) was used for EBC collection. Homogenization of the extracts was carried out by an MS2 Vortex (IKA, Germany).

An Agilent 7890A Series GC system coupled to an Agilent 7200 UHD Accurate-Mass QTOF hybrid mass spectrometer equipped with an electron impact (EI) source (Santa Clara, CA, USA) was used. The analytical sample was monitored in high resolution mode.

2.5. Sample preparation

Sample preparation consisted of liquid-liquid extraction using hexane as extractant. In all cases, 100 µl aliquots of EBC were 1:1 vortexed with hexane in a glass insert at room temperature for 1 min. Then, the organic phase was isolated

and put into a new glass insert for analysis. To eliminate exogenous interferents, blanks were prepared by using water treated as the samples.

2.6. GC-TOF/MS analysis

GC-TOF/MS analyses were performed by electron impact ionization (EI) mode at 70 eV and controlled by MassHunter Acquisition B.06. The GC separation was carried out by a fused silica DB-5MS-UI 30 m×0.25 mm i.d, 0.25 µm film thickness capillary column. The GC oven temperature program started at 60 °C (1 min held), followed by a temperature ramp of 10 °C min-1 to final 300 °C (2 min held). Post-run time was programmed for 4 min up to 310 ºC to assure complete elution of the injected sample. Pulsed splitless injections of 1 µl of sample were carried out at 250 °C, and ultrapure grade helium was used as the carrier gas at 1.0 ml min^{-1} flow rate. The interface and ion source were set at 280 and 300 °C, respectively. A solvent delay of 5.5 min was used to prevent damage in the ion source filament. The TOF detector was operated at 5 spectra s^{-1} in the mass range m/z 50-550 and the resolution was 8500 (full width half maximum, FWHM) at m/z 501.9706. Mass-spectrometric grade PFTBA (perfluorotri-n-butylamine) was used for daily mass calibration. Tentative identification of metabolites was performed by searching MS spectra on the NIST 11 database taking into account the RI values.

2.7. Data processing and statistical analysis

Unknown Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to process all data obtained by GC¬-TOF/MS in full scan mode. Treatment of raw data files started by deconvolution of potential molecular features (MFs) with the suited algorithm included in the software. For this purpose, the deconvolution algorithm considered all ions exceeding 1500 counts for the absolute height parameter. Additionally, the accuracy error and the window size factor were set at 50 ppm and 150 units, respectively. After extraction of MFs, data files in compound exchange format (.cef files) were created for each sample and exported into the Mass Profiler Professional (MPP)

software package (version 12.1, Agilent Technologies, Santa Clara, CA, USA) for further processing.

In the next step, the data were processed by alignment of the potential MFs according to their retention time and m/z value using a tolerance window of 0.3 min and an accuracy error of 15 ppm. The MFs from the analysis of blanks were removed from the data set of MFs from the EBC samples. The extraction algorithm confirmed the efficiency of this filtering step. Stepwise reduction of the MFs number was based on frequency of occurrence by comparing repetitions of the same group of individuals. A filter by frequency was set at 100%, thus ensuring detection of each MF in all the injected replicates from each group (lung cancer, risk factor group and control healthy individuals).

In the last step, the resulting MFs were exported (.cef file) for recursive analysis. For this purpose, the Quantitative Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to reintegrate all potential compounds found in the analyzed samples. The resulting table was exported in coma separated value format (.csv file) and reprocessed with the Mass Profiler Professional (MPP) software package. A filter to eliminate samples with within-replicates variability above 10% was applied to assure the effectiveness of the recursive analysis. Finally, the data set was normalized by logarithmic transformation of the ratio between the peak area of each molecular feature and the total useful mass spectrometry signal (MSTUS) corresponding to all the samples [26,28].

The resulting data set from each EBC fraction was then subjected to unsupervised and supervised analysis by principal components analysis (PCA) and supported vector machine analysis (SVM). The latter was done using ROCCET (http://www.roccet.ca/ROCCET/), an online toolbox for ROC curve analyses of metabolomics data [29] that uses Monte-Carlo cross-validation (MCCV) to estimate the predictive performance as well as the stability of the selected model performing 50 iterations. In each MCCV, two-thirds of the samples are randomly selected to evaluate the significance of molecular features

and the most important features are selected with different cut-offs to build models that are validated with the remaining 1/3 of the samples.

A paired t-test analysis was applied to compare the two fractions of EBC from the same patient, and an unpaired t-test was used to compare the groups under study using a Bonferroni-Holm multiple testing correction. The SVM analysis was complementarily used to evaluate the predictive response of identified compounds with high discrimination capability.

2.8. Identification of potential MFs detected by GC-TOF/MS

Identification was firstly carried out by searching MS spectra in the NIST11 database. Only identifications with a match factor and a reverse match factor higher than 700 were considered valid. RI values included in the NIST database were also taken into account to support identifications. An RI calibration model was built by comparing the RI values of an alkane standard mixture (composed by alkane between C10 to C40 with an even number of carbons) with the chromatographic method used in this research and the RI values provided by the NIST database. Supplementary figure 1 shows the RI calibration line obtained by this approach. The requirement to accept the NIST identifications was that the difference between the theoretical and the experimental RI obtained by extrapolation in the calibration curve should be within ±100 units.

The NIST database does not contain high resolution MS information as provided by the TOF detector. For this reason, a third step was included to validate identification of each MF by high resolution mass spectrometry. Thus, the molecular formula for the tentative precursor ion [M]+ and the most intense product ions obtained for each MF should fit the NIST identification by setting a cut-off value in mass accuracy of 10 ppm. The KEGG (Kyoto Encyclopedia of Genes and Genomes) database was used in the identification of the main changes occurring in the composition of EBC from individuals pertaining to each class.

Table 1. Compounds identified in EBC by GC-TOF/MS analysis.

Compound Name	Retention Time	Formula	CAS ID	Fragments	Family
Eucalyptol	5.99	$C_{10}H_{18}O$	470-82-6	154.1361 – $[C10H18O]^+$ 139.1119 – $[C9H15O]^+$ 93.0695 – $[C7H9]^+$	Aliphatic heteropolycyclic compounds (Oxanes)
Indole	9.90	C_8H_7N	120-72-9	117.0558 – $[C8H7N]^+$ 90.0448 $[C7H6]^+$ 74.0145 – $[C6H2]^+$	Aromatic heteropolycyclid compounds (Indoles)
Benzoic acid 4-ethoxy-ethyl ester	12.92	$C_{11}H_{14}O_3$	23676-09-7	194.0425 – $[C11H14O3]^+$ 149.0581 – $[C9H9O2]^+$ 121.0269 – $[C7H5O2]^+$	Aromatic homomonocyclic compounds (Benzene and substituted derivatives)
Benzoic acid methyl ester	6.95	$C_8H_8O_2$	93-58-3	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic homomonocyclic compounds (Benzoic acid derivatives)
3,5-Di-t-butyl-4-hydroxycinnamic acid	20.16	$C_{17}H_{24}O_3$	22014-01-3	276.1712 – $[C17H24O3]^+$ 261.1479 – $[C16H21O3]^+$ 177.0896 – $[C11H13O2]^+$	Aromatic homomonocyclic compounds (Cinnamic acid derivatives)
Octinoxate	21.01	$C_{18}H_{26}O_3$	5466-77-3	290.1867 – $[C18H26O3]^+$ 178.0611 – $[C10H10O3]^+$ 161.0578 – $[C10H9O2]^+$	Aromatic homomonocyclic compounds (Cinnamic acid derivatives)
Cresol	6.54	C_7H_8O	106-44-5	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic homomonocyclic compounds (Phenols and derivatives-cresol)

Cont. Table 1

Benzyl alcohol	5.96	C_7H_8O	100-51-6	108.0565 – $[C7H8O]^+$ 91.0535 – $[C7H7]^+$ 79.0533 – $[C6H7]^+$	Aromatic homomonocyclic compounds (Primary alcohols)
Isopropyl laurate	14.97	$C_{15}H_{30}O_2$	10233-13-3	201.1835 – $[C12H25O2]^+$ 157.1203 – $[C9H17O2]^+$ 102.0656 – $[C5H10O2]^+$	Lipids (Fatty acid esters)
Palmitic acid methyl ester	17.34	$C_{17}H_{34}O_2$	112-39-0	270.2545 – $[C17H34O2]^+$ 227.1998 – $[C14H27O2]^+$ 143.1048 – $[C8H15O2]^+$	Lipids (Fatty acid esters)
Oleic acid methyl ester	19.04	$C_{19}H_{36}O_2$	112-62-9	296.2702 – $[C19H36O2]^+$ 264.2442 – $[18C32HO]^+$ 81.0685 – $[C6H9]^+$	Lipids (Fatty acid esters)
Stearic acid methyl ester	19.28	$C_{19}H_{38}O_2$	112-61-8	298.2862 – $[C19H38O2]^+$ 255.2315 – $[C16H31O2]^+$ 87.0436 – $[C4H7O2]^+$	Lipids (Fatty acid esters)
Palmitoleic acid	17.45	$C_{16}H_{30}O_2$	373-49-9	236.2122 – $[C16H28O]^+$ 98.0710 – $[C6H10O]^+$ 69.0689 – $[C5H9]^+$	Lipids (Fatty acids and conjugates)
Palmitic acid	17.69	$C_{16}H_{32}O_2$	57-10-3	227.1997 – $[C14H27O2]^+$ 129.0891 – $[C7H13O2]^+$ 73.0279 – $[C3H5O2]^+$	Lipids (Fatty acids and conjugates)
Stearic acid	19.58	$C_{18}H_{36}O_2$	57-11-4	284.2706 – $[C18H36O2]^+$ 129.0908 – $[C7H13O2]^+$ 73.0281 – $[C3H5O2]^+$	Lipids (Fatty acids and conjugates)
Glycidol stearate	22.48	$C_{21}H_{40}O_3$	7460-84-6	297.2436 – $[C18H33O3]^+$ 98.0719 – $[C6H10O]^+$ 71.0848 – $[C5H11]^+$	Lipids (Fatty acids and conjugates)

Cont. Table 1

Undecanol	11.25	$C_{11}H_{24}O$	112-42-5	111.1157 – $[C8H15]^+$ 83.0844 – $[C6H11]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty alcohols)
1-Hexadecanol 2-methyl	15.99	$C_{17}H_{36}O$	2490-48-4	111.1160 – $[C8H15]^+$ 97.1006 – $[C7H13]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty alcohols)
Oleamide	21.33	$C_{18}H_{35}NO$	301-02-0	281.2679 – $[C18H35NO]^+$ 126.0914 – $[C7H12NO]^+$ 72.0438 – $[C3H6NO]^+$	Lipids (Fatty amides)
11-Eicosenamide	23.01	$C_{20}H_{39}NO$	10436-08-5	126.0915 – $[C7H12NO]^+$ 72.0439 – $[C3H6NO]^+$ 309.2973 – $[C20H39NO]^+$	Lipids (Fatty amides)
Erucamide	24.57	$C_{22}H_{43}NO$	112-84-5	337.3338 – $[C22H43NO]^+$ 126.0916 – $[C7H12NO]^+$ 72.0440 – $[C3H6NO]^+$	Lipids (Fatty amides)
Monopalmitin	22.57	$C_{19}H_{38}O_4$	542-44-9	299.2577 – $[C18H35O3]^+$ 257.2462 – $[C16H33O2]^+$ 239.2366 – $[C16H31O]^+$	Lipids (Glycerolipids)
Monostearin	24.14	$C_{21}H_{42}O_4$	123-94-4	327.2897 – $[C20H39O3]^+$ 267.2677 – $[C18H35O]^+$ 98.0723 – $[C6H10O]^+$	Lipids (Glycerolipids)
Hedione	14.41	$C_{13}H_{22}O_3$	24851-98-7	83.0479– $[C5H7O]^+$ 97.0623– $[C6H9O]^+$ 226.1566 – $[C13H22O3]^+$	Lipids (Lineolic acids and deriativatives-jasmonic acids)
Sclareol	18.77	$C_{20}H_{36}O_2$	515-03-7	272.2494 – $[C20H32]^+$ 121.0988 – $[C9H13]^+$ 95.0837 – $[C7H11]^+$	Lipids (Prenol lipids-diterpenes)
Limonene	5.94	$C_{10}H_{16}$	5989-27-5	136.1227 – $[C10H16]^+$ 121.0992 – $[C9H13]^+$ 79.0524 – $[C6H7]^+$	Lipids (Prenol lipids-monoterpenes)

Cont. Table 1

Cumyl alcohol	6.81	$C_{10}H_{14}O$	617-94-7	121.0644 – $[C8H9O]^+$ 103.0530 – $[C8H7]^+$ 91.0537 – $[C7H7]^+$	Lipids (Prenol lipids-monoterpenes)
Linalool	6.99	$C_{10}H_{18}O$	78-70-6	136.1218 – $[C10H16]^+$ 93.0679 – $[C7H9]^+$ 71.0844 – $[C5H4]^+$	Lipids (Prenol lipids-monoterpenes)
Camphor	7.84	$C_{10}H_{16}O$	464-48-2	152.1192 – $[C10H16]^+$ 137.0963 – $[C9H13O]^+$ 95.0852 – [C7H11]	Lipids (Prenol lipids-monoterpenes)
Camphol	8.16	$C_{10}H_{18}O$	507-70-0	121.1000 – $[C9H13]^+$ 95.0853 – $[C7H11]^+$ 77.0381 – $[C6H5]^+$	Lipids (Prenol lipids-monoterpenes)
Levomenthol	8.25	$C_{10}H_{20}O$	2216-51-5	138.1379 – $[C10H18]^+$ 95.0837 – $[C7H11]^+$ 81.0683 – $[C6H9]^+$	Lipids (Prenol lipids-monoterpenes)
Terpineol	8.52	$C_{10}H_{18}O$	98-55-5	136.1244 – $[C10H16]^+$ 121.1007 – $[C9H13]^+$ 93.0695 – $[C7H9]^+$	Lipids (Prenol lipids-monoterpenes)
Squalene	24.84	$C_{30}H_{50}$	111-02-4	410.3907 – $[C30H50]^+$ 121.0994 – $[C9H13]^+$ 81.0686 – $[C6H9]^+$	Lipids (Prenol lipids-triterpenes)
Cholestadiene	25.56	$C_{27}H_{44}$	747-90-0	368.3437 – $[C27H44]^+$ 247.2412 – $[C18H31]^+$ 147.1141 – $[C11H15]^+$	Lipids (Steroids and steroids derivatives)
Triethyl citrate	14.40	$C_{12}H_{20}O_7$	77-93-0	203.0913 – $[C9H15O5]^+$ 157.0496 – $[C7H9O4]^+$ 83.0486 – $[C5H7O]^+$	Organic acids and derivatives (Carboxylic acids and derivatives)
Spiro [2,4]heptane-1,5-dimethyl-6-methylene	5.95	$C_{10}H_{16}$	62238-24-8	136.1225 – $[C10H16]^+$ 121.0990 – $[C9H3]^+$ 79.0524 – $[C6H7]^+$	Other organic compounds

Cont. Table 1

2-Propanol 1-(2-butoxy-1-methylethoxy)-	9.21	$C_{10}H_{22}O_3$	29911-28-2	59.0485 – $[C3H7O]^+$ 86.0715 – $[C5H10O]^+$ 103.0728 – $[C5H11O2]^+$	Other organic compounds
2,4,6-Triisopropylphenol	12.69	$C_{17}H_{26}O$	08-07-34	220.1822 – $[C15H24O]^+$ 205.1584 – $[C14H21O]^+$ 77.0369 – $[C6H5]^+$	Other organic compounds
1,3-Heptadecyn-1-ol	15.11	$C_{17}H_{32}O$	56554-77-9	225.1826 – $[C14H25O2]^+$ 81.0681 – $[C6H9]^+$ 67.0529 – $[C5H7]^+$	Other organic compounds
2,4-Diphenyl-4-methyl-2(E)-pentene	16.45	$C_{18}H_{20}$	22768-22-5	236.1567 – $[C18H20]^+$ 143.0809 – $[C11H11]^+$ 91.0513 – $[C7H7]^+$	Other organic compounds
7,9-Di-t-butyl-oxaspiro(4,5)deca-6,9-diene-2,8-dione	17.18	$C_{17}H_{26}O_2$	82304-66-3	175.1104 – $[C12H15O]^+$ 133.0638 – $[C9H9O]^+$ 77.0369 – $[C6H5]^+$	Other organic compounds
10,18-Bisnorabieta-8,11,13-triene	17.99	$C_{18}H_{26}$	32624-67-2	242.2007 – $[C18H26]^+$ 227.1790 – $[C17H23]^+$ 143.0864 – $[C11H11O]^+$	Other organic compounds
2,6-Di-t-butyl-4-(2-phenylpropan-2-yl)phenol	18.71	$C_{23}H_{32}O$	34624-81-2	324.2438 – $[C23H32O]^+$ 309.2212 – $[C22H29O]^+$ 119.0836 – $[C9H11]^+$	Other organic compounds
Phenol 2,2'-methylenebis[6-(1,1-dimethyl)-4-methyl	21.77	$C_{23}H_{32}O_2$	119-47-1	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other organic compounds
Phenol 2,4-bis(1-methyl-1-phenylethyl)-	22.52	$C_{24}H_{26}O$	2772-45-4	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other organic compounds

Cont. Table 1

2,4-Bis(dimethylbenzyl)-6-*t*-butylphenol	22.57	$C_{28}H_{34}O$	244080-16-8	386.2617 – $[C28H34O]^+$ 371.2370 – $[C27H31O]^+$ 293.1897 – $[C21H25O]^+$	Other organic compounds
n-Hexadecylindane	23.93	$C_{25}H_{42}$	55334-29-7	117.0351 – $[C9H9]^+$ 130.0427 – $[C10H10]^+$ 154.1345 – $[C11H22]^+$	Other organic compounds

3. Results and discussion

3.1. Comparison of the relative concentration of components in UA and DA samples

Despite fractionation of EBC is supposed to separate exogenous and endogenous components, no previous studies had evaluated the composition of the two EBC fractions. Table 1 lists the 47 compounds identified in this research classified by chemical family. Three compounds were excluded as their origin could be explained only by external sources such as cosmetic products (sclareol and octinoxate) or plastic material used during analysis –7,9-di-tert-butyl-1-oxaspiro(4,5)deca-6,9-diene-2,8-dione. All the identified compounds were present in both UA and DA samples, but some of them showed differences in the relative concentration between fractions. Figure 1 shows the bars diagram comparing the average of the relative concentration for the 44 compounds detected in the samples. As can be seen, most of the identified compounds were found in DA at higher concentrations than in UA. A paired t-test revealed that 15 of these compounds were present at significantly different concentration in the fractions representing both airways (p value<0.05), which means that almost 34% of the identified EBC components presented different concentration levels in UA and DA (Figure 1). As the two EBC fractions gave quantitatively different composition, independent statistical multivariate analysis of each fraction was carried out.

3.2. Individual multivariate analysis of the respiratory airway fractions

The next step was to look for discrimination patterns among the three groups included in this study, which could be slightly distinguished by comparison of representative total ion chromatograms for each group (Supplementary figure 2).

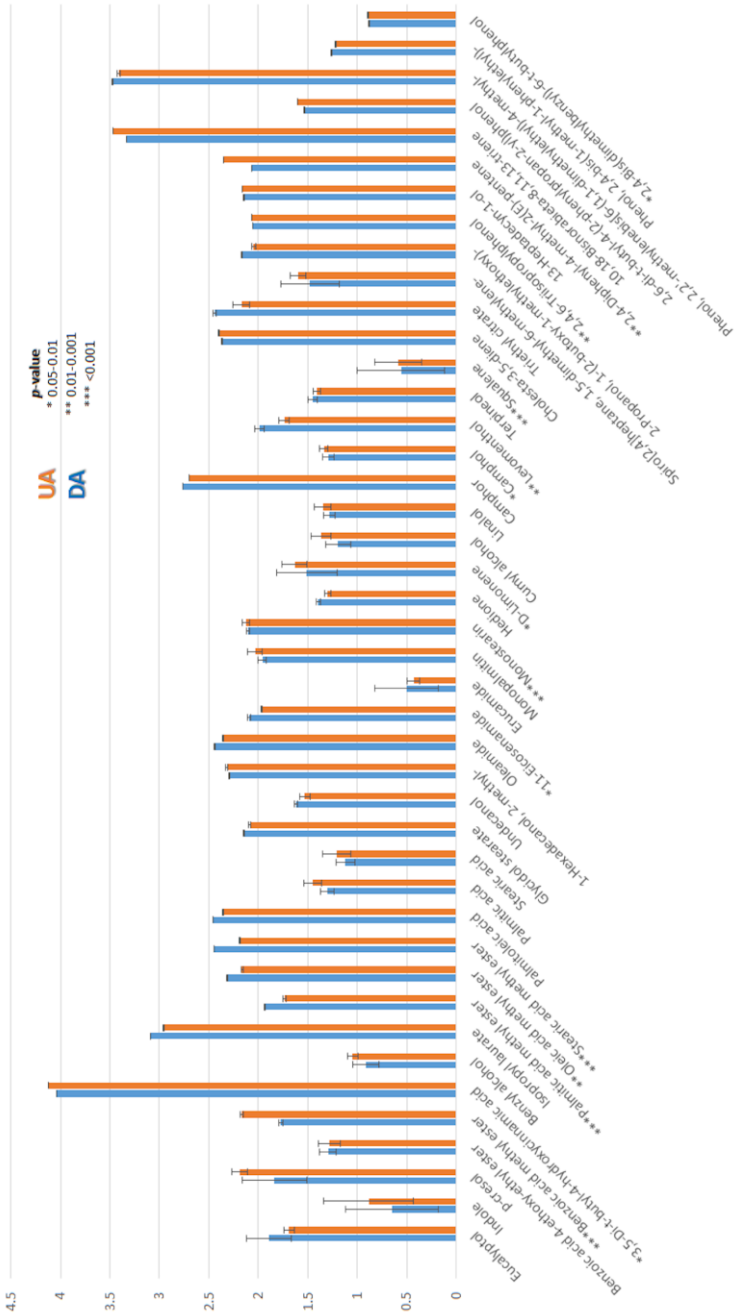


Figure 1. Bars diagram comparing the composition of EBC collected from the upper and central airways (UA) and the distal airway (DA).

Previously, statistical analyses by Fisher's exact and Mann Whitney tests discarded the significance of factors such as age, sex and tobacco cumulative consumption, the latter by comparison of the risk factor and lung cancer cohorts, obtaining in all cases a p-value > 0.05. Unsupervised analysis was applied to find differences in EBC fractions collected from the three groups. Principal component analysis was carried out with the data set including 44 compounds. Owing to the high variability associated to individuals, the three groups under study (lung cancer, risk factor group and healthy individuals) appeared overlapped in the three-dimensional PCA score plots for UA and DA (Supplementary figure 3). For this reason, unsupervised analysis was split into three studies by including only two groups: lung cancer versus risk factor group, lung cancer versus control healthy individuals, and risk factor group versus control healthy individuals. Supplementary figures 4 and 5 illustrate the 3D scores plots for the three cases using the UA and DA, respectively. In all cases, discrimination trends between the two target groups were observed, but without complete separation.

Supervised analysis was applied to find discrimination patterns associated to the diagnostic of lung cancer. As in the previous case, the SVM analysis was split into three studies: lung cancer versus risk factor group, lung cancer versus control healthy individuals, and risk factor group versus control individuals. Figures 2 and 3 illustrate the 2D scores plots for the three independent studies of each fraction. As can be seen, discrimination trends were clearly observed in all instances for the DA samples, while the UA samples only showed a discrimination pattern for comparison between lung cancer patients and control individuals. The percentage of correctly classified samples for the SVM discrimination models performed for UA and DA are listed in Table 2. Lung cancer patients were accurately classified in the model combined with healthy individuals, reporting in the UA fraction a sensitivity of 83.7% and a specificity of 83.3%; while for the DA fraction the sensitivity and specificity were 77.5 and 89.8%, respectively. The discrimination ability decreased when the risk factor group was included in the classification analysis, which is quite logical because it is the intermediate group. Thus, in UA the sensitivity and specificity values for

separation of this group from lung cancer patients were 58.1 and 63.7%, and 63.7 and 69.4% for separation from healthy individuals. These values increased when the DA fraction was the target sample, with 75.5 and 70.5% of sensitivity and specificity for separation of the risk factor group from lung cancer patients and a sensitivity of 79.5% and specificity of 71.4% for discrimination of lung cancer patients from healthy individuals.

Table 2. Correctly classified sample rates for the prediction models developed by SVM analysis from UA and DA information.

Fraction	**UA**			**DA**		
Models	**RF *vs* LC**	**Control vs LC**	**Control vs RF**	**RF *vs* LC**	**Control vs LC**	**Control vs RF**
Sensitivity (%)	58.1	83.7	63.7	75.5	77.5	79.5
Specificity (%)	63.7	83.3	69.4	70.5	89.8	71.4

3.3. Identification of compounds with discriminating capability

Unpaired t-test was applied to identify the most significant compounds (p-value <0.01) contributing to explain the differences observed among the three groups based on DA analysis. Table 3 lists the significant compounds provided by statistical analysis in the three models. p-Values, type of regulation and fold change ratio are included to show the differences in relative concentration of the identified compounds. Eleven compounds were found statistically significant by combination of the three models.

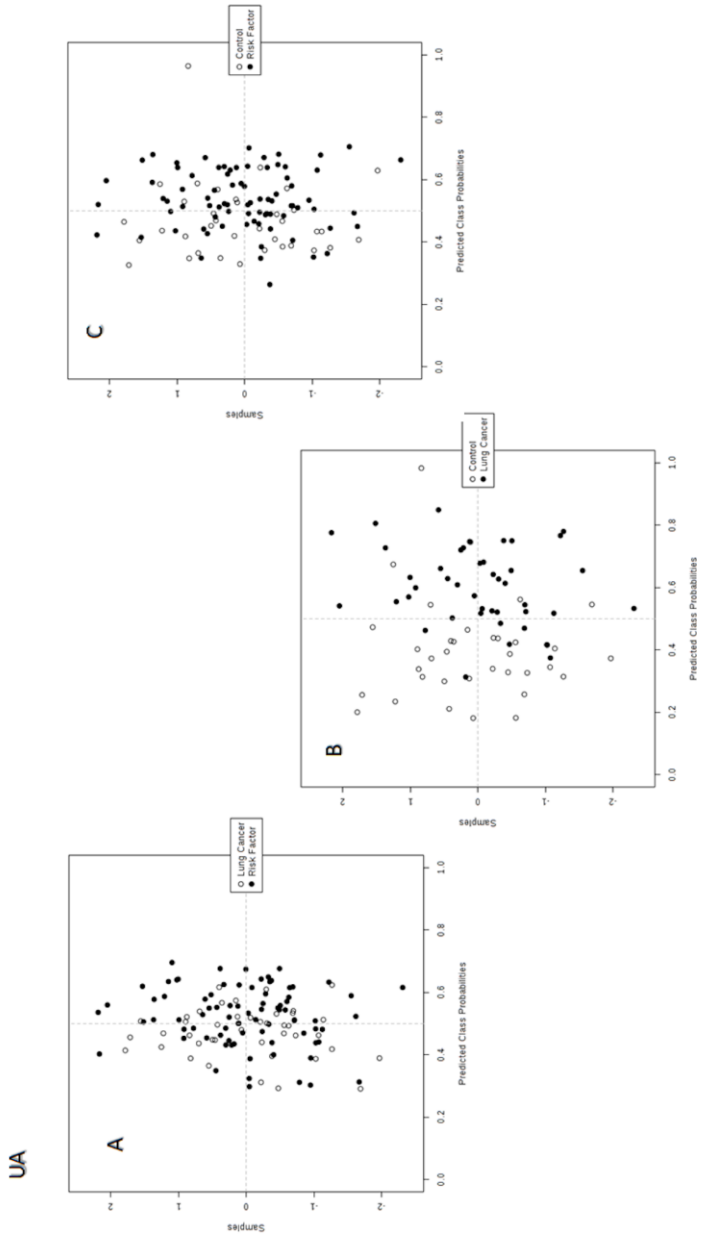


Figure 2. Supported Vector Machine Analysis (SVM) built from the data set obtained after analysis of EBC extracts from upper and central airways comparing (A) lung cancer patients and risk factor individuals, (B) lung cancer patients and control individuals, and (C) control and risk factor individuals.

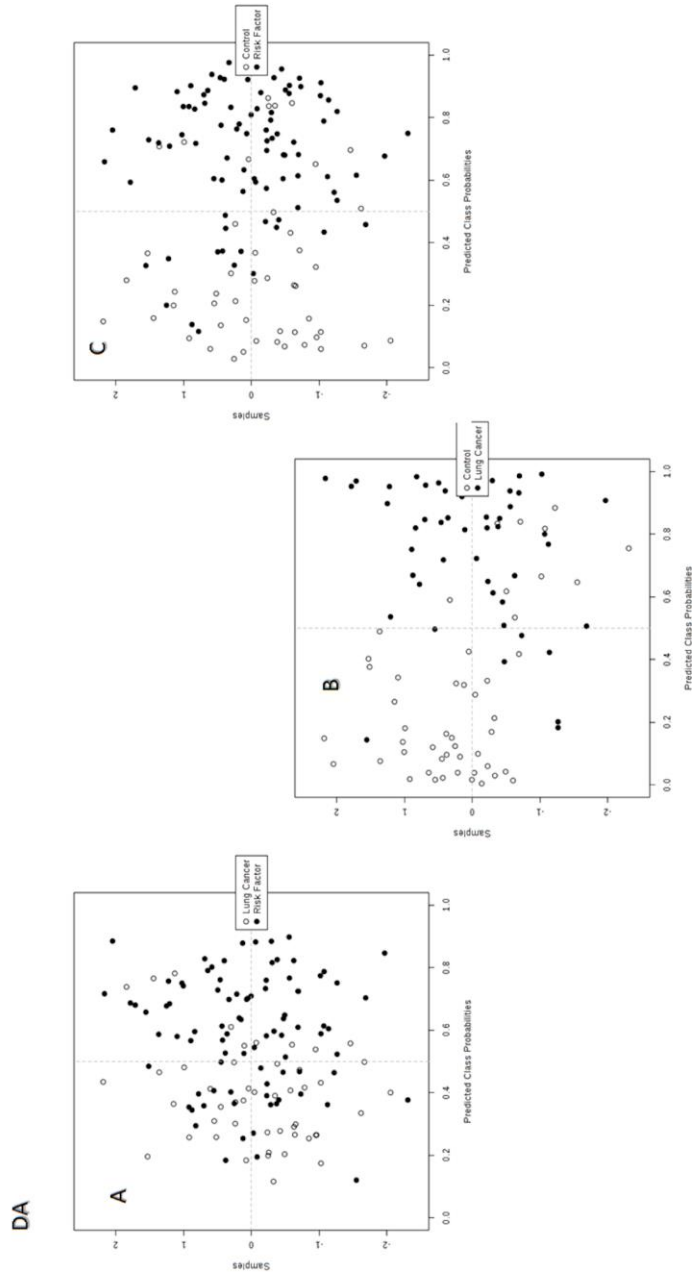


Figure 3. Supported Vector Machine Analysis (SVM) built from the data set obtained after analysis of EBC extracts from distal airway comparing (A) lung cancer patients and risk factor individuals, (B) lung cancer patients and control individuals, and (C) control and risk factor individuals.

According to fold change values, the most relevant changes were found in the models comparing control healthy individuals with the other two groups, risk factor individuals and lung cancer patients. In fact, four compounds were found significant in these two models, which were p-cresol, indole, benzoic acid 4-ethoxy-ethyl ester and 11-eicosenamide. It is also interesting to highlight the presence of two compounds, monoglyceride monostearin and hexadecylindane, in the list of significant compounds in the three models. One other monoglyceride, monopalmitin, was only significant in the two models discriminating risk factor individuals. Finally, some compounds were significant in only one of the three models: triethyl citrate (lung cancer patients versus control individuals), heptadecyn-1-ol and squalene (lung cancer patients versus risk factor individuals) and docosenamide (control versus risk factor individuals).

Table 3. Unpaired t-test analysis using a Bonferroni-Holm multiple testing correction to evaluate the significance of the identified compounds in EBC to discriminate between pairs of groups in this research (lung cancer patients, risk factor individuals and control individuals) and the frequency observed for these compounds by the SVM analysis.

Lung Cancer *vs* Control				
Compound	***p***	**Regulation**	**FC**	**Freq**
***p*-Cresol**	9.00E-04	Down	-3.10466	0.5
Indole	3.51E-04	Down	-2.77727	0.46
Benzoic acid 4-ethoxy-ethyl ester	0.003489	Down	-1.58336	0.84
Triethyl citrate	0.009728	Up	2.233156	0.5
11-Eicosenamide	3.13E-05	Up	2.16227	0.6
n- Hexadecylindane	0.002707	Up	1.952551	0.32
Monostearin	0.003911	Up	2.608156	0.24

Lung Cancer *vs* Risk Factor				
Compound	***p***	**Regulation**	**FC**	**Freq**
13-Heptadecyn-1-ol	0.002017	up	1.4398	0.44
Monopalmitin	0.003899	up	2.630078	0.12
n- Hexadecylindane	4.81E-04	up	1.815379	0.74
Monostearin	1.07E-04	up	3.018592	0.3
Squalene	2.12E-04	up	2.431075	0.86
Control *vs* Risk Factor				
Compound	***p***	**Regulation**	**FC**	**Freq**
***p*-Cresol**	0.009083	down	-2.27088	0.3
Indole	0.004347	down	-2.0566	0.54
Benzoic acid 4-ethoxy-ethyl ester	0.004672	down	-1.35418	0.44
Monopalmitin	1.36E-07	up	5.801463	0.76
11-Eicosenamide	8.68E-05	up	1.799589	0.54
n- Hexadecylindane	3.75E-12	up	3.544619	0.44
Monostearin	1.74E-12	up	7.87296	0.62
13-Docosenamide	5.19E-04	up	1.485012	0.34

The predictive capability of the identified compounds was evaluated by application of the linear SVM algorithm to each discrimination model. As previously indicated, the procedure was repeated 50 times according to the MCCV method explained in the Experimental section. This tool iteratively

combines two-thirds of the samples to evaluate the capability of compounds for sample classification by combination of them to build models with different cut-offs, which are validated on the remaining 1/3 of the samples. The discrimination accuracy and the selection frequency of each compound in all potential combinations provide information about the robustness of these compounds to classify samples in the three groups. The top-15 compounds with the highest selection frequency for each discrimination model are shown in Figure 4, which also shows the average predictive accuracy by considering all combinations. As can be seen, four (13-heptadecyn-ol, hexadecylindane, monostearin and squalene, which fit with the four most significant compounds from a statistical point of view) out of five statistically significant compounds in the model comparing lung cancer patients versus risk factor individuals were in the top-15 compounds with highest selection frequency. The model dealing with lung cancer patients and control healthy individuals only gave three significant compounds (p-cresol, triethyl citrate and 11-eicosenamide) within the list of top-15 with the highest selection frequency. On the other hand, five compounds with the highest selection frequency (indole, benzoic acid 4-ethoxy-ethyl ester, monostearin, monopalmitin and 11-eicosenamide) were statistically significant in the model comparing control and risk factor individuals.

4. Discussion

The selection of the studied cohort was supported on the statistical independence of two main factors such as age and sex, which could influence the representativeness of the study. One interesting result from this study is the compositional comparison between DA and UA by virtue of the used sampler device. Apparently, no qualitative differences were observed, although most of the identified compounds were found in DA at higher concentrations than in UA.

A
Squalene
n-Hexadecylindane
13-Heptadecyn-1-ol
Phenol, 2,4-bis(1-methyl-1-phenylethyl)-
Palmitoleic acid
Dodecanoic acid, 1-methylethyl ester
cis-11-Eicosenamide
Phenol, 2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)-
Benzyl alcohol
a-Monosteanin
Glycidol stearate
Palmitic acid
Phenol, 2,4,6-tris(1-methylethyl)-
Indole
Cumyl alcohol
Selected Frequency (%)
Content
High
Low
B
2,4-Diphenyl-4-methyl-2(E)-pentene
Undecyl alcohol
Dodecanoic acid, 1-methylethyl ester
Glycerol 1-palmitate
Palmitic acid, methyl ester
Palmitic acid
a-Monosteanin
2,4-Bis(dimethylbenzyl)-6-t-butylphenol
a-cumyl alcohol
Palmitoleic acid
Indole
cis-11-Eicosenamide
Levomenthol
Glycidol stearate
a-terpineol
Selected Frequency (%)
Content
High
Low
C
Squalene
Palmitic acid
Dodecanoic acid, 1-methylethyl ester
2,4-Bis(dimethylbenzyl)-6-t-butylphenol
Benzoic acid 4-ethoxy-ethyl ester
Hedione
Undecyl alcohol
2,4-Diphenyl-4-methyl-2(E)-pentene
Cholesta-3,5-diene
Palmitic acid, methyl ester
1-Hexadecanol, 2-methyl-
cis-11-Eicosenamide
Phenol, 2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)-
Phenol, 2,4,6-tris(1-methylethyl)-
Triethyl citrate
Selected Frequency (%)
Content
High
Low

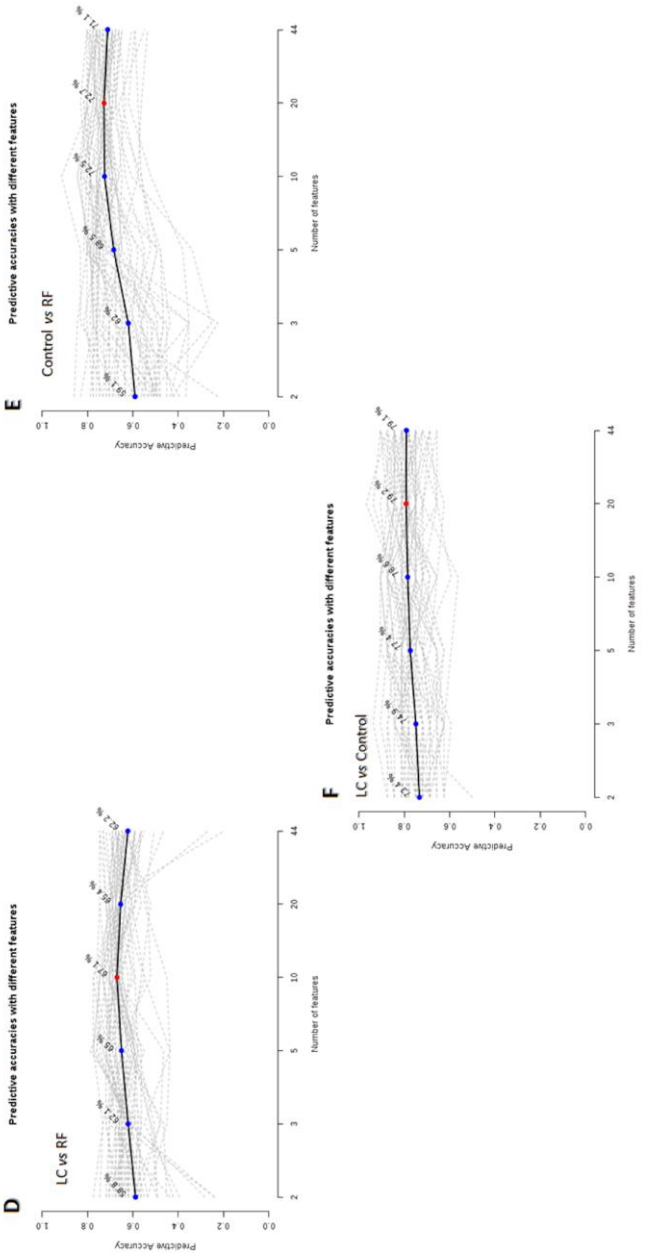


Figure 4. The top-15 compounds with the highest selection frequency for each discrimination model –(A) lung cancer vs risk factor, (B) control vs risk factor, (C) lung cancer vs control– and the average predictive accuracy by considering all combinations –(D) lung cancer vs risk factor, (E) control vs risk factor, (F) lung cancer vs control.

The paired t-test identified seven compounds more concentrated in DA: methyl ester fatty acids (stearic, oleic and palmitic), glycidol stearate, levomenthol, squalene and n-hexadecylindane. On the other hand, other eight compounds –phenol derivatives (2,4-bis(dimethylbenzyl)-6-t-butylphenol, 2,4,6-triisopropylphenol and 3,5-di-t-butyl-4-hydroxycinnamic acid), terpenes (camphol and limonene), benzoic acid methyl ester, 11-eicosenamide and 2,4-diphenyl-4-methyl-2(E)-pentene– were more concentrated in UA samples. The differences reported in the two EBC fractions made advisable independent comparison of the three target cohorts.

As mentioned before, the compounds identified in EBC from the three groups under study (Table 1) were used for multivariate analysis. In unsupervised analysis, both EBC fractions led to a similar discrimination trend; nevertheless, intra-individual and inter-individual variability sources did not allow complete separation of the evaluated groups. In fact, the combination of PC1/PC2/PC3 did not explain above 50% of the total variability contained in the study in the different PCA tests. Concerning supervised analysis, SVM was initially applied to build the prediction models listed in Table 2. From a clinical point of view, the most interesting model to aid in the diagnostic of lung cancer is that comparing lung cancer patients and the risk factor group. Attending to the recognition ability parameter, the DA fraction is characterized by a classification capability higher than the UA fraction. The model provided by DA analysis is well balanced in terms of sensitivity and specificity, being these parameters 75.5 and 70.5%, respectively. A similar situation was found in the comparison between control healthy individuals and risk factor individuals since the sensitivity provided by the UA fraction was lower than that of the DA sample. With these premises, the EBC from the DA was characterized by a higher discrimination capability than that coming from the UA. For this reason, the DA fraction was selected for further studies dealing with identification of significant compounds contributing to explain the observed patterns.

Unpaired t-test was applied to identify the most significant compounds contributing to explain the differences observed among the three groups (Figure 1). Among the five compounds significant in the comparison of lung cancer patients versus the risk factor group, it is worth noting the presence of two saturated monoacylglycerols (monopalmitin and monostearin) and an acyclic triterpenoid (squalene), which is the precursor of sterols, including cholesterol, and bile acids [30]. The presence of squalene in exhaled breath has been widely reported [31]. This compound is structurally similar to isoprene, which is considered one of the most concentrated compounds in human breath [22]. In fact, some authors have proposed that polyisoprenes such as squalene are potential sources of isoprene by peroxidation, one of the mechanisms of the oxidative stress [31]. However, isoprene was not detected in EBC in this research, which could be explained by its high volatility. The differences in the relative concentration of the two monoacylglycerols and squalene in the three groups under study can be visualized in Figure 5. Monopalmitin and monostearin were characterized by different behaviors: monopalmitin was more concentrated in the risk factor group than in lung patients, who also presented higher concentration of this monoacylglycerol than healthy individuals. On the other hand, monostearin offered the inverse profile as the risk factor group gave a lower relative concentration than lung cancer patients. Squalene gave a concentration profile similar to monopalmitin, which could explain the absence of the latter in the top-15 compounds with the highest selection frequency to discriminate lung cancer patients and risk factor individuals. The forth compound (hexadecylindane, an indane derivative) has not been related to any endogenous source and, therefore, it could be ascribed to an exogenous origin. However, Phillips *et al.* also detected indane derivatives in breath [32], and even they selected one of them as biomarker of lung cancer [33]. In this research, hexadecylindane was the second compound with the highest selection frequency (after squalene) in a theoretical model for discrimination of lung cancer patients and risk factor individuals. Concerning 13-heptadecyn-1-ol, no information about its presence in breath or about its biological implication in human processes has previously been reported.

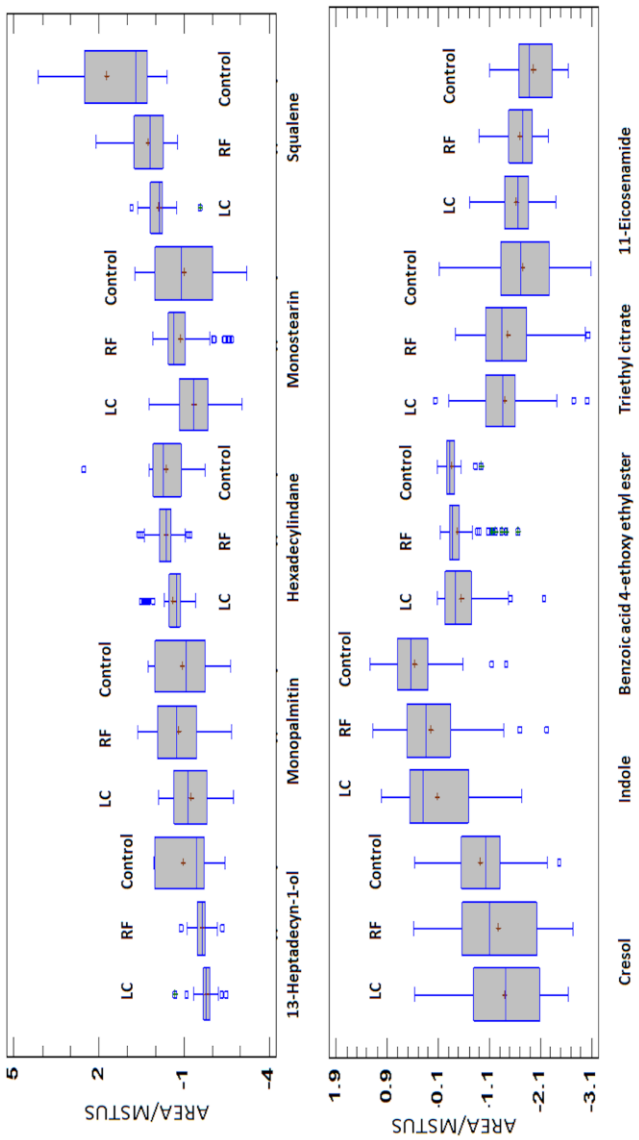


Figure 5. Box-and-whisker plots showing the variability of (A) four metabolites –13-heptadecyn-1-ol, monopalmitin, n-hexadecylindane, monostearin and squalene– in lung cancer patients comparing with all control individuals of the cohort, and (B) four metabolites –11-eicosenamide, p-cresol, indole, benzoic acid 4-ethoxy ethyl ester, triethyl citrate– in control individuals comparing lung cancer patients and risk factor individuals.

Taking the group of healthy individuals as reference, seven compounds resulted significant in the comparison versus lung cancer patients and the risk factor group. Among them, triethyl citrate, which was detected at higher concentration in lung cancer patients and in risk factor individuals than in healthy individuals, has been found in cigarette filters [34]; therefore, its presence in exhaled breath could only be linked to exogenous sources. The SVM model identified triethyl citrate in the list of compounds with the highest selection frequency to discriminate lung cancer patients and risk factor individuals. One phenolic compound (p-cresol) and a phenol derivative (benzoic acid 4-ethoxyethyl ester) were also found at different concentrations in the three groups. p-Cresol was detected at lower concentration in lung cancer patients than in the risk factor individuals, who also reported lower levels than healthy individuals. The ester derivative of benzoic acid was found at lower concentration in lung cancer and risk factor individuals than in the healthy group. Philipps *et al.* identified this compound as a candidate marker in human breath with potential to discriminate lung cancer patients [33]. This result also agrees with that deduced in this research since benzoic acid 4-ethoxy-ethyl ester was the compound with the highest selection frequency in the model comparing lung cancer patients versus control individuals.

One other interesting found compound was indole, involved in the metabolism of tryptophan, particularly in bacteria [35]. Additionally, indole has been found in cigarette smoke; therefore, it could also be associated to tobacco consumption [36]. The relative concentration profile of this compound was marked by a higher concentration in the risk factor group followed by lung cancer patients and, at lower level, in healthy individuals. Finally, one other compound contributing to differentiate healthy individuals from lung cancer patients was eicosenamide, present at lower level in this group as compared to the risk factor group and lung cancer individuals. This fatty amide had not been endogenously reported, but one analog such as oleamide, which was also found in EBC, is structurally related to the endogenous cannabinoid anandamide, involved in many biological functions. Eicosenamide was also characterized by its high

selection frequency in a theoretical model for discrimination of lung cancer patients from control individuals.

In summary, the great variety of compounds identified in EBC emphasizes the interest of this biofluid for clinical analysis. Although identification of volatiles compounds as isoprene (which is detected in breath vapor) has not been possible, non-volatile and medium-volatility compounds have been identified, being most of them of biological interest and previously identified in breath such as fatty acids and conjugates that represent a highly metabolically active lipid class [37,38]; volatile prenol lipid compounds such as camphol [39], limonene [40], linalool [32], α-cumyl alcohol [32] and terpineol [41]; aromatic homomonocyclic compounds as benzoic acid 4-ethoxy ethyl ester [32] or the aromatic heteropolycyclic compound indole [42]. Previous discrimination models based on exhaled breath analysis were characterized by the presence of essentially volatile compounds. With analysis of EBC a more varied composition of compounds with discriminant capability was attained.

In addition, the proposed methodology includes a data normalization strategy and a standardized sampling avoiding saliva and air room contamination. In addition, the used sampler allows separation into two fractions allowing to delete that with greater variability. Therefore, the potential of EBC as biofluid for clinical analysis of both volatile and non-volatile compounds to study respiratory diseases has been evidenced.

Conclusion

As previously exposed, EBC has not been widely exploited in the clinical field despite the advantages associated to its sampling. A method for metabolomics analysis of EBC based on GC-TOF/MS profiling in high resolution mode has been developed using liquid-liquid extraction for sample preparation. Compounds tentatively identified have been used to discriminate among three different groups: individuals diagnosed with lung cancer, risk factor individuals

(including smokers) and healthy control individuals. Among the identified compounds, it is worth mentioning the presence of monoacylglicerol derivatives of two of the four main saturated fatty acids and squalene, which could be considered as an intermediate in the pathway for in vivo formation of human breath isoprene by peroxidation [43] and also involved in the synthesis of cholesterol [44]. The results have shown the potential of EBC as a biofluid to discriminate between lung cancer patients and the risk factor group, which could help in the diagnosis of this disease in the search for a screening method to reduce the use of a confirmatory test to the case of positive response of the former.

Acknowledgements

The Spanish Ministerio de Economía y Competitividad (MINECO), Junta de Andalucía and FEDER programme are gratefully acknowledged for financial support (Projects “Development of methods for early cancer detection, December 29, 2011”, and “Optimización y aplicación de plataformas metabolómicas de análisis de biofluidos no invasivos para la búsqueda de biomarcadores de diagnóstico precoz de cáncen de pulmón, FQM-1602”). A. Peralbo-Molina and M. Calderón-Santiago also thank to these organisms their research contracts. F. Priego-Capote is also grateful to the Ministerio de Ciencia e Innovación (MICINN) for a Ramón y Cajal Contract (RYC-2009-03921). Reina Sofía Hospital and its biorepository node (Córdoba, Spain) is also grateful for EBCsamples.

References

[1] Jemal A, Bray F, Center M M, Ferlay J, Ward E and Forman D 2011 Global cancer statistics CA. Cancer J. Clin. 61 69-90

[2] Burney P *et al.* 2014 Chronic obstructive pulmonary disease mortality and prevalence: the associations with smoking and poverty: a bold analysis Thorax 69 465–73

[3] Jemal A, Center M M, DeSantis C and Ward E M 2010 Global patterns of cancer incidence and mortality rates and trends Cancer Epidemiol. Biomarkers Prev. 19 1893–907

[4] Boiselle P M 2013 Computed tomography screening for lung cancer JAMA 309 1163–70

[5] Mets O M *et al.* 2013 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 59

[6] Beger R D 2013 A review of applications of metabolomics in cancer. Metabolites 3 552–74

[7] Calderón-Santiago M, Priego-Capote F, Jurado-Gámez B and Luque de Castro M D 2014 Optimization study for metabolomics analysis of human sweat by liquid chromatography-tandem mass spectrometry in high resolution mode J. Chromatogr. A 1333 70–8

[8] Calderón-Santiago M, Priego-Capote F, Turck N, Robin X, Jurado-Gámez B, Sanchez J C and Luque de Castro M D 2015 Human sweat metabolomics for lung cancer screening. Anal. Bioanal. Chem.

[9] Liu C Y, Wang C H, Chen T C, Lin H C, Yu C T and Kuo H P 1998 Increased level of exhaled nitric oxide and up-regulation of inducible nitric oxide synthase in patients with primary lung cancer. Br. J. Cancer 78 534–41

[10] Horváth I, Lázár Z, Gyulai N, Kollai M and Losonczy G 2009 Exhaled biomarkers in lung cancer. Eur. Respir. J. 34 261–75

[11] Carpagnano G E, Resta O, Foschino-Barbaro M P, Gramiccioni E and Carpagnano F Interleukin-6 is increased in breath condensate of patients with non-small cell lung cancer. Int. J. Biol. Markers 17 141–5

[12] Bayley D L, Abusriwil H, Ahmad A and Stockley R A 2008 Validation of assays for inflammatory mediators in exhaled breath condensate. Eur. Respir. J. 31 943–8

[13] Dalaveris E, Kerenidi T, Katsabeki-Katsafli A, Kiropoulos T, Tanou K, Gourgoulianis K I and Kostikas K 2009 VEGF, TNF-alpha and 8-isoprostane levels in exhaled breath condensate and serum of patients with lung cancer Lung Cancer 64 219–25

[14] Kostikas K, Papatheodorou G, Psathakis K, Panagou P and Loukides S 2003 Oxidative stress in expired breath condensate of patients with COPD. Chest 124 1373–80

[15] Gessner C, Hammerschmidt S, Kuhn H, Seyfarth H-J, Sack U, Engelmann L, Schauer J and Wirtz H 2003 Exhaled breath condensate acidification in acute lung injury. Respir. Med. 97 1188–94

[16] Antus B and Barta I 2012 Exhaled breath condensate pH in patients with lung cancer. Lung Cancer 75 178–80

[17] Phillips M, Gleeson K, Hughes J M, Greenberg J, Cataneo R N, Baker L and McVay W P 1999 Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. Lancet 353 1930–3

[18] Phillips M, Cataneo R N, Cummin A R C, Gagliardi A J, Gleeson K, Greenberg J, Maxfield R a and Rom W N 2003 Detection of lung cancer with volatile markers in the breath Chest 123 2115–23

[19] Phillips M *et al.* 2007 Prediction of lung cancer using volatile biomarkers in breath Cancer Biomark. 3 95–109

[20] Wang Y, Hu Y, Wang D, Yu K, Wang L, Zou Y, Zhao C, Zhang X, Wang P and Ying K 2012 The analysis of volatile organic compounds biomarkers for lung cancer in exhaled breath, tissues and cell lines. Cancer Biomark. 11 129–37

[21] Kubáň P and Foret F 2013 Exhaled breath condensate: determination of non-volatile compounds and their potential for clinical diagnosis and monitoring. A review Anal. Chim. Acta 805 1–18

[22] Costa C, Bucca C, Bergallo M, Solidoro P, Rolla G and Cavallo R 2011 Unsuitability of exhaled breath condensate for the detection of herpesviruses DNA in the respiratory tract J. Virol. Methods 173 384–6

[23] Dragonieri S, Brinkman P, Mouw E, Zwinderman A H, Carratú P, Resta O, Sterk P J and Jonkers R E 2013 An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls Respir. Med. 107 1073–8

[24] Hunt J 2002 Exhaled breath condensate: An evolving tool for noninvasive evaluation of lung disease J. Allergy Clin. Immunol. 110 28–34

[25] Mazzatenta A, Di Giulio C and Pokorski M 2013 Pathologies currently identified by exhaled biomarkers Respir. Physiol. Neurobiol. 187 128–34

[26] Peralbo-Molina A, Calderón-Santiago M, Priego-Capote F, Jurado-Gámez B and Luque de Castro M D 2015 Development of a method for metabolomic analysis of human exhaled breath condensate by gas chromatography–mass spectrometry in high resolution mode Anal. Chim. Acta 887 118–26

[27] Hoffmeyer F, Raulf-Heimsoth M, Harth V, Bünger J and Brüning T 2009 Comparative analysis of selected exhaled breath biomarkers obtained with two different temperature-controlled devices BMC Pulm. Med. 9 48

[28] Warrack B M, Hnatyshyn S, Ott K-H, Reily M D, Sanders M, Zhang H and Drexler D M 2009 Normalization strategies for metabonomic analysis of urine samples. J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 877 547–52

[29] Xia J, Broadhurst D I, Wilson M and Wishart D S 2013 Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics 9 280–99

[30] Salvioli G, Lugli R and Pradelli J M 1984 Relationships between squalene and cholesterol in bile: Effect of ursodeoxycholic acid administration in patients with radiolucent gallstones Metabolism 33 641–5

[31] Phillips M and Greenberg J 1991 Method for the collection and analysis of volatile compounds in the breath J. Chromatogr. B Biomed. Sci. Appl. 564 242–9

[32] Phillips M 1999 Variation in volatile organic compounds in the breath of normal humans J. Chromatogr. B 729 75–88

[33] Phillips M *et al.* 2008 Detection of lung cancer using weighted digital analysis of breath biomarkers. Clin. Chim. Acta. 393 76–84

[34] Anon 1968 Smoke filters

[35] Nowak A and Libudzisz Z 2006 Influence of phenol, p-cresol and indole on growth and survival of intestinal lactic acid bacteria. Anaerobe 12 80–4

[36] Snook M E, Arrendale R F, Higman H C and Chortyk O T 1978 Isolation of indoles and carbazoles from cigarette smoke condensate Anal. Chem. 50 88–90

[37] Quehenberger O *et al.* 2010 Lipidomics reveals a remarkable diversity of lipids in human plasma J. Lipid Res. 51 3299–305

[38] Martínez-Lozano P and Fernández de la Mora J 2008 Direct analysis of fatty acid vapors in breath by electrospray ionization and atmospheric pressure ionization-mass spectrometry Anal. Chem. 80 8210–5

[39] Martin A N, Farquar G R, Jones A D and Frank M 2010 Human breath analysis: methods for sample collection and reduction of localized background effects Anal. Bioanal. Chem. 396 739–50

[40] Vautz W, Nolte J, Fobbe R and Baumbach J I 2009 Breath analysis-performance and potential of ion mobility spectrometry J. Breath Res. 3 036004

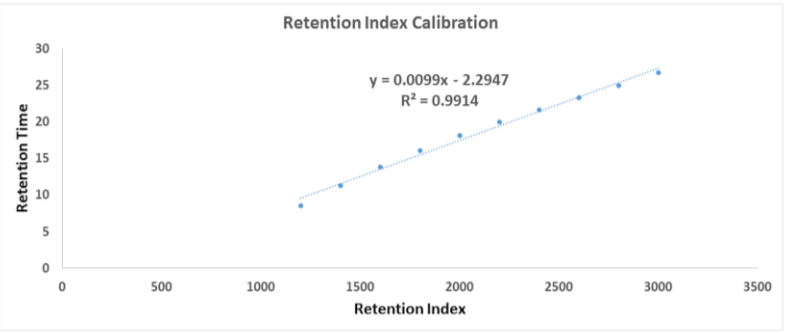
[41] Van Berkel J J B N, Dallinga J W, Möller G M, Godschalk R W L, Moonen E J, Wouters E F M and Van Schooten F J 2010 A profile of volatile organic compounds in breath discriminates COPD patients from controls Respir. Med. 104 557–63

[42] Van den Velde S, Nevens F, Van Hee P, van Steenberghe D and Quirynen M 2008 GC-MS analysis of breath odor compounds in liver patients J. Chromatogr. B 875 344–8

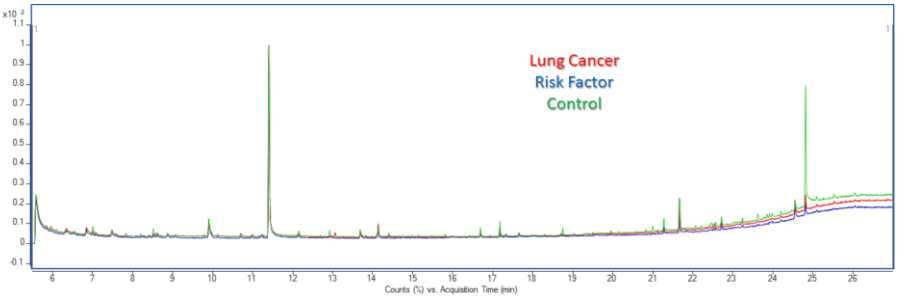
[43] Stein R A and Mead J F 1988 Small hydrocarbons formed by the peroxidation of squalene Chem. Phys. Lipids 46 117–20

[44] Woodward R B and Bloch K 1953 The cyclization of squalene in cholesterol sinthesis J. Am. Chem. Soc. 75 2023–4

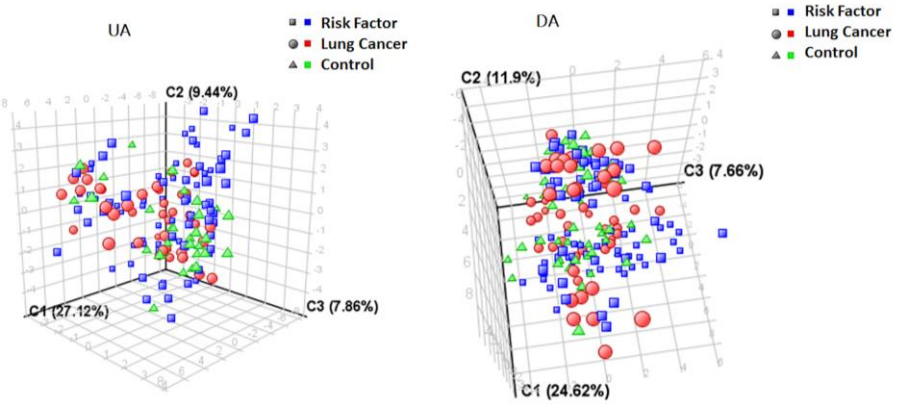
Supplementary material



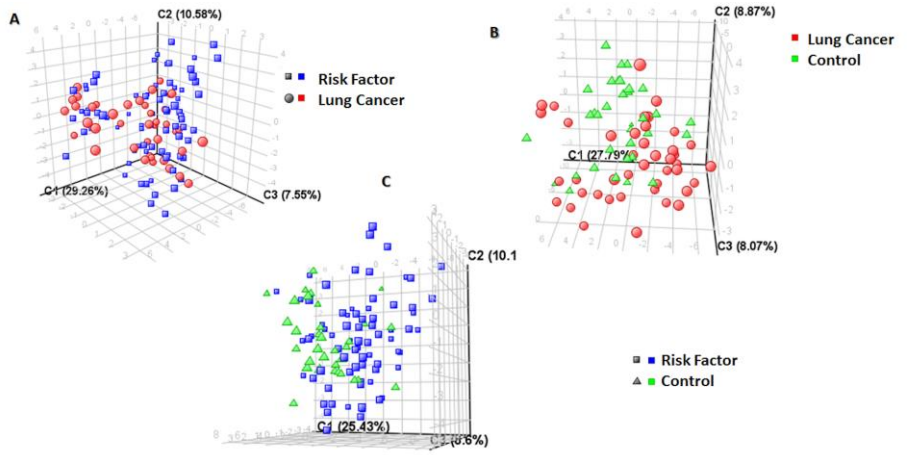
Supplementary figure 1. RI calibration line obtained by analysis of the alkane standard mixture.



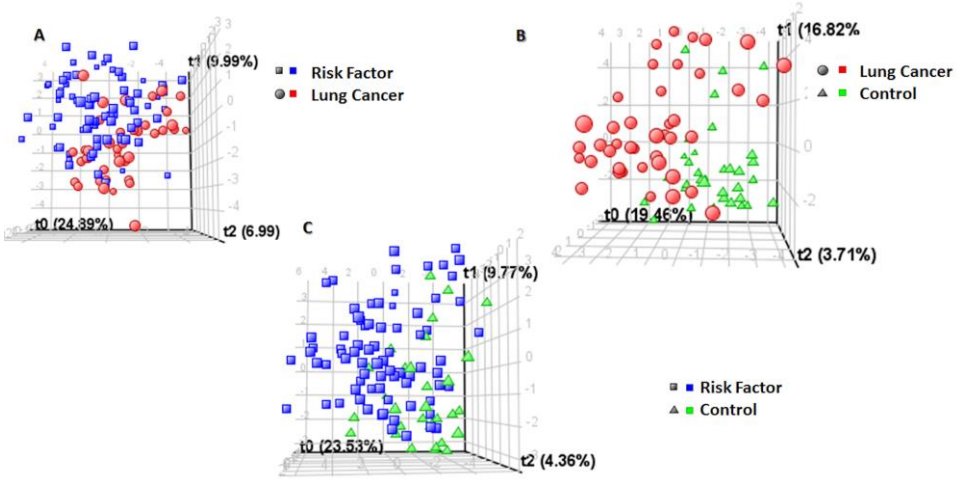
Supplementary figure 2. Total ion chromatogram (TIC) for each group of individuals.



Supplementary figure 3. Principal Component Analysis (PCA) built from the data set obtained after analysis of EBC extracts from both upper and central airways (UA) and distal airways (DA) from lung cancer patients, risk factor individuals and control individuals.



Supplementary figure 4. Principal Component Analysis (PCA) built from the data set obtained after analysis of EBC extracts from upper and central airways comparing (A) lung cancer patients and risk factor individuals, (B) lung cancer patients and control individuals, and (C) control and risk factor individuals.



Supplementary figure 5. Principal Component Analysis (PCA) built from the data set obtained after analysis of EBC extracts from distal airway comparing (A) lung cancer patients and risk factor individuals, (B) lung cancer patients and control individuals, and (C) control and risk factor individuals.

Capítulo VIII:

Identification of metabolomics panels for potential lung cancer screening by analysis of exhaled breath condensate

Journal of Breath Research

Volume 10 (2), 1 June 2016,

Pages 26002-26013

IOPscience

Identification of metabolomics panels for potential lung cancer screening by analysis of exhaled breath condensate

Ángela Peralbo-Molina[a,b], Mónica Calderón-Santiago[a,b], Feliciano Priego-Capote[a,b*], Bernabé Jurado-Gámez[b,c], María Dolores Luque de Castro[a,b*]

[a]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14014, Córdoba, Spain*

[b]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, E-14014, Córdoba, Spain*

[c] *Department of Respiratory Medicine, Reina Sofia University Hospital, University of Córdoba, 14004, Córdoba, Spain*

Identification of metabolomics panels for potential lung cancer screening by analysis of exhaled breath condensate

Ángela Peralbo Molina, Mónica Calderón-Santiago, Feliciano Priego-Capote, Bernabé Jurado-Gámez, María Dolores Luque de Castro

Abstract

Exhaled breath condensate (EBC) is one of the less employed biofluids to search for clinical markers, despite its non-invasive sampling and the potential relationship between its composition and respiratory disease phenotypes such as lung cancer. The advanced stage at which lung cancer is usually detected is the main reason for the high mortality rate of this carcinogenic disease. In this preliminary research, EBC was used as clinical sample to develop a screening tool for lung cancer discrimination from two control groups (with and without risk factor). Three panels of metabolites were configured using the PanelomiX tool to minimize false negatives (specificity) and false positives (sensitivity). The combination of five metabolites led to three panels providing a sensitivity above 77.9%, specificity above 67.5% and the area under the curve (AUC) above 77.5% for the three panels. An additional study was developed as a first approach to study the statistical significance of metabolites at different stages of lung cancer.

Keywords: lung cancer, exhaled breath condensate, metabolomics, GC-TOF/MS, biomarker panel

1. Introduction

Lung cancer is nowadays the neoplastic disease associated to the highest mortality worldwide, being responsible for more deaths than the combination of colon, breast and prostate cancers [1, 2]. The main reason for the high death rate of lung cancer is the low efficiency of screening methods for detection at early stages [3]. The most efficient result has been 20% reduction in mortality rate of a high-risk population by low-dose computed tomography (LDCT), according to the National Lung Screening Trial [4].

The search for new screening tools to diagnose lung cancer, intensified in the last decades, is being also addressed by omics approaches using different biofluids to compare lung cancer patients versus healthy control individuals. The main objective of these studies is to replace present invasive tests such as tissue analysis, which are far from being desirable methods for early detection of lung cancer. New biofluids obtained in a non-invasive manner are being proposed as alternative to classical biofluids such as blood and urine to identify molecular biomarkers for detection of lung cancer. These alternative biofluids are sweat, sputum, bronchoalveolar lavage fluid or exhaled breath (distinction is made between exhaled breath vapor –EBV– and exhaled breath condensate –EBC), characterized by a less complex composition than serum/plasma or urine [5]. Thus, sweat has recently been analyzed with the aim to propose a lung cancer screening panel [6] as did sputum and bronchoalveolar lavage fluid for non-small lung cancer detection [7]. Concerning EBV, Phillips et al. developed a mathematical model based on the analysis of 22 compounds identified in breath (mainly alkanes and derivatives and benzene derivatives) as primary lung cancer biomarker. The model reported sensitivity and selectivity values of 71.7% and 66.7%, respectively, even for advanced stages of the disease [8]. This study was later improved by including in the model a set of oxidative stress products excreted in breath [9], thus achieving sensitivity and specificity of 85.1% and 80.5%, respectively. Further on, another study by the same research group concluded that a two-minute breath test could predict lung cancer regardless

histology, stage of the disease or tobacco smoking, with 84.6% sensitivity and 80% specificity [10]. More recently, a study developed by Wang et al. provided a 23 VOCs panel with capability to distinguish between lung cancer patients and controls with 96.5% sensitivity and 97.5% specificity [11]. Most of the panel components are exogenous compounds detected at low concentrations in the air of the sampling room, and present in the sampled EBV at higher concentrations. This fact shows the main problem arising from the use of exhaled air as sample: the easy contamination from environmental components.

In contrast to EBV, EBC had received scant attention owing to the lack of homogeneity in sampling attributed to old protocols. The present, improved collection protocols allow discriminating the air from different depths of the respiratory system in two separate bags, thus making possible to separate the air exhaled from the upper and central airways to that from the distal airway [12].

Despite the main component of EBC is water vapor, it also contains traces of hundreds of different components, ranging from small inorganic ions through large organic molecules to peptides, proteins, surfactants and macromolecules [12–16]. In addition, the liquid nature of EBC increases the suitability for analysis of this biofluid as compared to EBV. Thus, different liquid–liquid extraction (LLE) protocols taking advantage of the extractant polarity can be used to isolate metabolites before analysis [17].

The aim of the present research was to identify preliminary panels of metabolites with potential discrimination capability for lung cancer screening by analysis of EBC as clinical sample. For this purpose, metabolite profiles were obtained by gas chromatography–mass spectrometry in high resolution mode (GC–TOF/MS) from lung cancer patients, which were compared to those provided by two additional cohorts: a risk factor group formed by active smokers with at least 20 pack-year of exposure and ex-smokers (<5 years without smoking), and a second group including healthy non-smoker individuals. These panels could be used as a first approach to the development of tools with capability to reduce the rate of individuals subjected to invasive intervention or confirmatory test for lung cancer diagnosis, mainly for risk factor individuals.

2. Experimental

2.1. Reagents

Hexane TraceSELECT® grade from Sigma-Aldrich (St. Louis, USA) was used as organic solvent for sample preparation. An alkane standard mixture (from C10 to C40), also from Sigma-Aldrich, was used for GC separation tests to establish the retention index (RI) calibration. Deionized water (18 mΩ·cm) from a Millipore Milli-Q water purification system was also used as blank sample.

2.2. Instruments and apparatus

An ECOScreen2 device (FILT Thorax-und LungenDiagnostik GmbH, Berlin, Germany) was used for EBC collection. Homogenezation of the extracts was carried out by an MS2 Minishaker Vortex (IKA, Germany).

An Agilent 7890A Series GC system coupled to an Agilent 7200 UHD Accurate-Mass QTOF hybrid mass spectrometer equipped with an electron impact (EI) source (Santa Clara, CA, USA) was used. The analytical sample was monitored in high resolution mode.

2.3. Cohort selected for the study

The EBC samples were collected from 256 individuals early in the morning before breakfast, and stored at -80 ºC in the biorepository of the Reina Sofía Hospital (Córdoba, Spain). All the individuals were recruited at the Department of Respiratory Medicine. The cohort included 68 patients diagnosed with lung cancer, all of whom were cytohistologically confirmed after clinical tests based on bronchoscopy, fine-needle biopsy, or video-assisted thoracoscopy. These patients had an average of 61±8 years, 86.4% of them were males and all them were active smokers or ex-smokers with <5 years they quit smoking and a cumulative consumption above 20 packs-year. The most frequent diagnostic was squamous cell carcinoma (20 patients, 29.41%), followed by adenocarcinoma (13 patients, 19.12%), small cell carcinoma (9 patients, 13.23%) and large cell

carcinoma (13 patients, 19.12%). Thirteen individuals (19.12%) were diagnosed with non-small cell lung without histological classification. Despite of the variability in the diagnostic of the individuals included in the lung cancer cohort, all of them were characterized by a severe lung damage. Table 1 includes information about the group, stage, histology and smoking habits.

Table 1. Generic information about each group including means of age, weight, height and body mass index.

Type	Number	Gender (% Male)	Age	% Smokers	Weight	Height	BMI	Stage	
								(%Early)	(%Advanced)
Control	62	87.0%	59±8	0%	81.5±13.0	1.7±0.1	29.0±4.3		
Risk factor	126	84.9%	61±8		82.3±16.4	1.6±0.1	28.7±5.1		
Smoker	60	81.7%	49±8	100%	71.8±14.7	1.4±0.1	25.3±4.2		
COPD	66	87.9%	65±7	100%	83.1±18.2	1.6±0.1	28.9±5.9		
Lung cancer	68	86.4%	61±8	100%	71.0±15.0	1.5±0.1	24.9±4.8	20.6	79.4

The risk factor group was composed by 126 individuals, all of them active smokers or ex-smokers (<5 years without smoking) with at least 20 pack-years of exposure. It is worth noting that 66 individuals pertaining to the risk factor cohort were diagnosed with COPD by spirometry (FVC/FEV1 ratio < 0.7). Three reasons supported the inclusion of COPD patients (all them smokers) within the risk factor group: firstly, smoking is considered the most important risk factor to develop lung cancer or COPD; secondly, increased oxidative stress or presence of inflammatory cells infiltrated in COPD and lung cancer are common ways for theoretical explanation of lung damage (particularly the latter); finally, the suitability of a screening tool should be evaluated under the most real scenario and COPD is a frequent diagnostic in medical units dealing with respiratory diseases. The risk factor group, clinically monitored at least for one year, was characterized by an average 61 ± 8 years with 84.9% of male individuals. Lung cancer was ruled out in this group through CT and bronchoscopy. The control group was formed by 62 healthy individuals with an average of 59 ± 8 years, 87% of them male individuals. All of them were no active nor passive smokers, without either clinical symptoms or abnormalities in the chest radiograph. Independence

of sex and age among groups was verified by Fisher's exact and Mann Whitney tests, respectively.

Criteria for exclusion of patients were: a) coexistence of extrapulmonary tumoral pathology or treatment with cytostatic drugs for a different neoplasm; b) diagnostic of neoplasm in the last five years; c) unjustified weight loss (≥ 10%) during the last year; d) severe disorders of any organ with negative influence on the prognostic or that hinders to apply the protocol (the cases involved included New York Heart Association grade IV heart failure, advanced hepatic cirrhosis, chronic renal (stages 4-5), and lung disease not related to smoking, including interstitial pneumopathy, pneumonia, tuberculosis, etc. [18]).

All experiments were carried out in accordance with the ethical principles of human medical research (World Medical Association, Helsinki Declaration). The ethical review board of Reina Sofía Hospital (Córdoba, Spain) approved and supervised this clinical study.

2.4. EBC collection procedure

The ECOScreen2 sampling device directly collects and condensates the EBC in disposable polyethylene bags. It operates at –20 ºC and offers controlled EBC collection into two separate bags for physical separation between the air contained within upper airway and that of distal airway [12]. The main modification of the sampler was the insertion of a commercial protection filter from Scharlab (Barcelona, Spain) over the inlet air valve to avoid the entrance of exogenous organic compounds and particles from the room atmosphere. This filter was periodically changed to avoid saturation.

Tidal breathing and a nose-clip were used for 15 min, time required to collect an average EBC volume of 1.5 mL from the distal airway. The samples were divided into 100 µL aliquots and the vials stored at –80 ºC until analysis. All samples were analyzed within 3 months after collection.

2.5. Sample preparation

Sample preparation was carried out by liquid-liquid extraction using hexane as extractant as described by Peralbo-Molina et al. [17]. In all cases, 100 µL aliquots of EBC were 1:1 vortexed with hexane in a glass insert at room temperature for 1 min. Then, the organic phase was isolated and put into a new glass insert for analysis. To eliminate exogenous interferences, two water blanks were treated as the EBC samples and then analysed, first of them using water as sample and the other, leaving water inside the collector bag for 15 minutes. Supplementary Figure 1 shows the comparison between a real sample and the latter blank indicating interferences that have been erased.

2.6. GC-TOF/MS analysis

GC-TOF/MS analyses were performed using an Agilent 7890 gas chromatograph system equipped with an electron impact ionization (EI) source fixed at 70 eV and controlled by MassHunter Acquisition B.06. Chromatographic separation was carried out by a fused silica DB-5MS-UI 30 m×0.25 mm i.d, 0.25 µm film thickness capillary column. Pulsed splitless injections of sample (1 µL) were carried out at 250 °C, and ultrapure grade helium was used as carrier gas at 1.0 mL min^{-1} flow rate. The temperature program of the GC oven started at 60 °C (1 min held), followed by a temperature ramp of 10 °C min-1 to final 300 °C (2 min held). Post-run time was programmed for 4 min up to 310 ºC to assure complete elution of the injected sample. The transfer-line and ion source temperatures were set at 280 and 300 °C, respectively. A solvent delay of 5.5 min was used to prevent damage in the ion source filament. The TOF detector was operated at 5 spectra s-1 in the mass range m/z 50-550, and the resolution was 8500 (full width half maximum, FWHM) at m/z 501.9706. Mass spectrometry grade PFTBA (perfluorotri-n-butylamine) was used for daily mass calibration.

2.7. Data processing and statistical analysis

Unknown Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to process all data from GC–TOF/MS in full scan mode. Treatment of raw data files started by deconvolution of potential molecular features (MFs) with the suited algorithm included in the software. For this purpose, the deconvolution algorithm considered all ions exceeding 1500 counts for the absolute height parameter. Additionally, the accuracy error and the window size factor were set at 50 ppm and 150 units, respectively. After extraction of MFs, data files in compound exchange format (.cef files) were created for each sample and exported into the Mass Profiler Professional (MPP) software package (version 12.1, Agilent Technologies, Santa Clara, CA, USA) for further processing.

In the next step, the data were processed by alignment of the potential MFs according to their retention time and m/z value using a tolerance window of 0.3 min and an accuracy error of 15 ppm. The MFs from the analysis of blanks were removed from the data set of MFs corresponding to the EBC samples. The extraction algorithm confirmed the efficiency of this filtering step. Stepwise reduction of the MFs number was based on frequency of occurrence by comparing repetitions of the same group of individuals. A filter by frequency was set at 100%, thus ensuring detection of each MF in all the injected replicates from each group (lung cancer, risk factor group and control healthy individuals).

In the last step, the resulting MFs were exported (.cef file) for recursive analysis. For this purpose, the Quantitative Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to reintegrate all potential compounds found in the analyzed samples. The resulting table was exported in coma separated value format (.csv file) and reprocessed with the Mass Profiler Professional (MPP) software package. A filter to eliminate samples with within-replicates variability above 10% was applied to assure the effectiveness of the recursive analysis. Finally, the data set was normalized by logarithmic transformation of the ratio between the peak area of each molecular feature and

the total useful mass spectrometry signal (MSTUS) corresponding to all the samples [26, 27].

Statgraphics Centurion (XVI version 16.1.18, Statpoint Technologies) was used for statistical analysis, while ROC curves for individual markers and in combined panels were obtained with the PanelomiX toolbox. This toolbox is supported on the iterative combination of biomarkers and thresholds to find a panel providing optimal classification [19].

Identification of compounds was firstly carried out by searching MS spectra on the NIST11 database. Only identifications with both a match factor and a reverse match factor higher than 700 were considered as valid. RI values included in the NIST database were also taken into account to support the identifications. For this purpose, an RI calibration model was built by relating the RT values of an alkane standard mixture (composed by alkane between C10 and C40 with an even number of carbons) using the chromatographic method proposed in this research and the RI values provided by the NIST database. Supplementary Figure 2 shows the RI calibration curve obtained by this approach. The requirement to accept the NIST identifications was that the difference between the theoretical and the experimental RI obtained by extrapolation in the calibration curve should be within ± 100 units. The NIST database does not contain high resolution MS information as provided by the TOF detector. For this reason, a third step was included to validate identification of each MF by using high resolution mass spectrometry. Thus, the molecular formula for the tentative ion [M+] and the most intense fragments obtained for each MF should fit the NIST identification by setting a cut-off value in mass accuracy of 10 ppm. Table 2 lists the 44 identified compounds classified by chemical families.

3. Results and discussion

As mentioned above, the main objective of this research was to propose a panel of metabolites detected in EBC as a potential screening tool to discriminate lung cancer patients. For this purpose, EBC collected from the three groups of individuals was analyzed by GC-TOF to find panels of metabolites with capability of discrimination among them. Thus, a first panel for discrimination between the risk factor cohort and healthy individuals was developed. Then, panels for discrimination between lung cancer patients and the other two groups were also generated. In all cases, the list of all identified metabolites (Table 2) was considered as the variables set.

3.1. Development of panels with capability of discrimination between the risk factor cohort and healthy individuals

The generation of this first panel comparing the risk factor cohort and the healthy group is useful to find patterns that could be associated to smoking or, generally, to lung damage, and also to understand the upcoming comparisons. PanelomiX was used as computational toolbox to configure the panels for discrimination between risk factor individuals and healthy individuals. This tool iteratively combines significant variables and thresholds to find panels of potential markers with good prediction/discrimination capability in terms of sensitivity and specificity [19]. In this research, the configuration of panels was supported on maximum accuracy, so both sensitivity and specificity are maximized. PanelomiX was applied to define panels of five markers with the best classification parameters including an N-Fold cross-validation (with N=10 and 5 repetitions) that reported information on the model stability. To accelerate the panel generation, a pre-filter step was included to limit the number of variables to only 10 compounds by considering 15 thresholds to each one. This pre-filter was based on random forest analysis to identify a preset number of variables, in this case 10 compounds, with the highest frequency of occurrence in the tree branches generated by application of this algorithm [19]. The frequency of the ten

metabolites considered after pre-filtering appears in Figure 1. Indole was the best individual marker giving 75.8% of specificity, 57.9% of sensitivity and an area under the curve (AUC) of 67%. However, the best panel in terms of sensitivity, specificity and AUC did not include indole and it was formed by combination of five metabolites –monopalmitin, monostearin, benzyl alcohol, 2,4-diphenyl-4-methyl-2-E-pentene and *p*-cresol– to provide 90.3% of specificity, 77.9% of sensitivity and 85.1% of AUC. Figure 2A shows the ROC curve of the panel in comparison to the indole curve, while the main characteristics of the panel are shown in Table 3, and cross-validation results are in Supplementary Table 1.

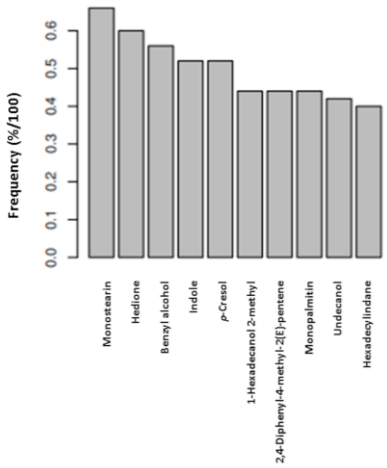


Figure 1. Frequency of the selected markers for the panel performance by comparing control and risk factor individuals.

The relative concentration of the panel compounds is shown in Figure 2B. As can be seen, benzyl alcohol and both monoacylglycerols were more concentrated in the risk factor group; while *p*-cresol and 2,4-diphenyl-4-methyl-2-E-pentene were more concentrated in healthy individuals, despite their previously demonstrated potential connection with tobacco exposure [20].

Table 2. Compounds identified in EBC by GC-TOF/MS analysis.

Compound Name	Retention Time	Formula	CAS ID	Fragments	Family
Eucalyptol	5.99	$C_{10}H_{18}O$	470-82-6	154.1361 – $[C10H18O]^+$ 139.1119 – $[C9H15O]^+$ 93.0695 – $[C7H9]^+$	Aliphatic heteropolycyclic compounds (Oxanes)
Indole	9.90	C_8H_7N	120-72-9	117.0558 – $[C8H7N]^+$ 90.0448 $[C7H6]^+$ 74.0145 – $[C6H2]^+$	Aromatic heteropolycyclid compounds (Indoles)
Benzoic acid 4-ethoxy-ethyl ester	12.92	$C_{11}H_{14}O_3$	23676-09-7	194.0425 – $[C11H14O3]^+$ 149.0581 – $[C9H9O2]^+$ 121.0269 – $[C7H5O2]^+$	Aromatic homomonocyclic compounds (Benzene and substituted derivatives)
Benzoic acid methyl ester	6.95	$C_8H_8O_2$	93-58-3	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic homomonocyclic compounds (Benzoic acid derivatives)
3,5-Di-t-butyl-4-hydroxycinnamic acid	20.16	$C_{17}H_{24}O_3$	22014-01-3	276.1712 – $[C17H24O3]^+$ 261.1479 – $[C16H21O3]^+$ 177.0896 – $[C11H13O2]^+$	Aromatic homomonocyclic compounds (Cinnamic acid derivatives)
Octinoxate	21.01	$C_{18}H_{26}O_3$	5466-77-3	290.1867 – $[C18H26O3]^+$ 178.0611 – $[C10H10O3]^+$ 161.0578 – $[C10H9O2]^+$	Aromatic homomonocyclic compounds (Cinnamic acid derivatives)
Cresol	6.54	C_7H_8O	106-44-5	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic homomonocyclic compounds (Phenols and derivatives-cresol)

Cont. Table 2

Benzyl alcohol	5.96	C_7H_8O	100-51-6	108.0565 – $[C7H8O]^+$ 91.0535 – $[C7H7]^+$ 79.0533 – $[C6H7]^+$	Aromatic homomonocyclic compounds (Primary alcohols)
Isopropyl laurate	14.97	$C_{15}H_{30}O_2$	10233-13-3	201.1835 – $[C12H25O2]^+$ 157.1203 – $[C9H17O2]^+$ 102.0656 – $[C5H10O2]^+$	Lipids (Fatty acid esters)
Palmitic acid methyl ester	17.34	$C_{17}H_{34}O_2$	112-39-0	270.2545 – $[C17H34O2]^+$ 227.1998 – $[C14H27O2]^+$ 143.1048 – $[C8H15O2]^+$	Lipids (Fatty acid esters)
Oleic acid methyl ester	19.04	$C_{19}H_{36}O_2$	112-62-9	296.2702 – $[C19H36O2]^+$ 264.2442 – $[18C32HO]^+$ 81.0685 – $[C6H9]^+$	Lipids (Fatty acid esters)
Stearic acid methyl ester	19.28	$C_{19}H_{38}O_2$	112-61-8	298.2862 – $[C19H38O2]^+$ 255.2315 – $[C16H31O2]^+$ 87.0436 – $[C4H7O2]^+$	Lipids (Fatty acid esters)
Palmitoleic acid	17.45	$C_{16}H_{30}O_2$	373-49-9	236.2122 – $[C16H28O]^+$ 98.0710 – $[C6H10O]^+$ 69.0689 – $[C5H9]^+$	Lipids (Fatty acids and conjugates)
Palmitic acid	17.69	$C_{16}H_{32}O_2$	57-10-3	227.1997 – $[C14H27O2]^+$ 129.0891 – $[C7H13O2]^+$ 73.0279 – $[C3H5O2]^+$	Lipids (Fatty acids and conjugates)
Stearic acid	19.58	$C_{18}H_{36}O_2$	57-11-4	284.2706 – $[C18H36O2]^+$ 129.0908 – $[C7H13O2]^+$ 73.0281 – $[C3H5O2]^+$	Lipids (Fatty acids and conjugates)
Glycidol stearate	22.48	$C_{21}H_{40}O_3$	7460-84-6	297.2436 – $[C18H33O3]^+$ 98.0719 – $[C6H10O]^+$ 71.0848 – $[C5H11]^+$	Lipids (Fatty acids and conjugates)

Cont. Table 2

Undecanol	11.25	$C_{11}H_{24}O$	112-42-5	111.1157 – $[C8H15]^+$ 83.0844 – $[C6H11]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty alcohols)
1-Hexadecanol 2-methyl	15.99	$C_{17}H_{36}O$	2490-48-4	111.1160 – $[C8H15]^+$ 97.1006 – $[C7H13]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty alcohols)
Oleamide	21.33	$C_{18}H_{35}NO$	301-02-0	281.2679 – $[C18H35NO]^+$ 126.0914 – $[C7H12NO]^+$ 72.0438 – $[C3H6NO]^+$	Lipids (Fatty amides)
11-Eicosenamide	23.01	$C_{20}H_{39}NO$	10436-08-5	126.0915 – $[C7H12NO]^+$ 72.0439 – $[C3H6NO]^+$ 309.2973 – $[C20H39NO]^+$	Lipids (Fatty amides)
Erucamide	24.57	$C_{22}H_{43}NO$	112-84-5	337.3338 – $[C22H43NO]^+$ 126.0916 – $[C7H12NO]^+$ 72.0440 – $[C3H6NO]^+$	Lipids (Fatty amides)
Monopalmitin	22.57	$C_{19}H_{38}O_4$	542-44-9	299.2577 – $[C18H35O3]^+$ 257.2462 – $[C16H33O2]^+$ 239.2366 – $[C16H31O]^+$	Lipids (Glycerolipids)
Monostearin	24.14	$C_{21}H_{42}O_4$	123-94-4	327.2897 – $[C20H39O3]^+$ 267.2677 – $[C18H35O]^+$ 98.0723 – $[C6H10O]^+$	Lipids (Glycerolipids)
Hedione	14.41	$C_{13}H_{22}O_3$	24851-98-7	83.0479– $[C5H7O]^+$ 97.0623– $[C6H9O]^+$ 226.1566 – $[C13H22O3]^+$	Lipids (Lineolic acids and deriatives-jasmonic acids)
Sclareol	18.77	$C_{20}H_{36}O_2$	515-03-7	272.2494 – $[C20H32]^+$ 121.0988 – $[C9H13]^+$ 95.0837 – $[C7H11]^+$	Lipids (Prenol lipids-diterpenes)
Limonene	5.94	$C_{10}H_{16}$	5989-27-5	136.1227 – $[C10H16]^+$ 121.0992 – $[C9H13]^+$ 79.0524 – $[C6H7]^+$	Lipids (Prenol lipids-monoterpenes)

Cont. Table 2

Cumyl alcohol	6.81	$C_{10}H_{14}O$	617-94-7	121.0644 – $[C8H9O]^+$ 103.0530 – $[C8H7]^+$ 91.0537 – $[C7H7]^+$	Lipids (Prenol lipids-monoterpenes)
Linalool	6.99	$C_{10}H_{18}O$	78-70-6	136.1218 – $[C10H16]^+$ 93.0679 – $[C7H9]^+$ 71.0844 – $[C5H4]^+$	Lipids (Prenol lipids-monoterpenes)
Camphor	7.84	$C_{10}H_{16}O$	464-48-2	152.1192 – $[C10H16]^+$ 137.0963 – $[C9H13O]^+$ 95.0852 – $[C7H11]$	Lipids (Prenol lipids-monoterpenes)
Camphol	8.16	$C_{10}H_{18}O$	507-70-0	121.1000 – $[C9H13]^+$ 95.0853 – $[C7H11]^+$ 77.0381 – $[C6H5]^+$	Lipids (Prenol lipids-monoterpenes)
Levomenthol	8.25	$C_{10}H_{20}O$	2216-51-5	138.1379 – $[C10H18]^+$ 95.0837 – $[C7H11]^+$ 81.0683 – $[C6H9]^+$	Lipids (Prenol lipids-monoterpenes)
Terpineol	8.52	$C_{10}H_{18}O$	98-55-5	136.1244 – $[C10H16]^+$ 121.1007 – $[C9H13]^+$ 93.0695 – $[C7H9]^+$	Lipids (Prenol lipids-monoterpenes)
Squalene	24.84	$C_{30}H_{50}$	111-02-4	410.3907 – $[C30H50]^+$ 121.0994 – $[C9H13]^+$ 81.0686 – $[C6H9]^+$	Lipids (Prenol lipids-triterpenes)
Cholestadiene	25.56	$C_{27}H_{44}$	747-90-0	368.3437 – $[C27H44]^+$ 247.2412 – $[C18H31]^+$ 147.1141 – $[C11H15]^+$	Lipids (Steroids and steroids derivatives)
Triethyl citrate	14.40	$C_{12}H_{20}O_7$	77-93-0	203.0913 – $[C9H15O5]^+$ 157.0496 – $[C7H9O4]^+$ 83.0486 – $[C5H7O]^+$	Organic acids and derivatives (Carboxylic acids and derivatives)
Spiro [2,4]heptane-1,5-dimethyl-6-methylene	5.95	$C_{10}H_{16}$	62238-24-8	136.1225 – $[C10H16]^+$ 121.0990 – $[C9H3]^+$ 79.0524 – $[C6H7]^+$	Other organic compounds

Cont. Table 2

2-Propanol 1-(2-butoxy-1-methylethoxy)-	9.21	$C_{10}H_{22}O_3$	29911-28-2	59.0485 – $[C3H7O]^+$ 86.0715 – $[C5H10O]^+$ 103.0728 – $[C5H11O2]^+$	Other organic compounds
2,4,6-Triisopropylphenol	12.69	$C_{17}H_{26}O$	08-07-34	220.1822 – $[C15H24O]^+$ 205.1584 – $[C14H21O]^+$ 77.0369 – $[C6H5]^+$	Other organic compounds
1,3-Heptadecyn-1-ol	15.11	$C_{17}H_{32}O$	56554-77-9	225.1826 – $[C14H25O2]^+$ 81.0681 – $[C6H9]^+$ 67.0529 – $[C5H7]^+$	Other organic compounds
2,4-Diphenyl-4-methyl-2(E)-pentene	16.45	$C_{18}H_{20}$	22768-22-5	236.1567 – $[C18H20]^+$ 143.0809 – $[C11H11]^+$ 91.0513 – $[C7H7]^+$	Other organic compounds
7,9-Di-t-butyl-oxaspiro(4,5)deca-6,9-diene-2,8-dione	17.18	$C_{17}H_{26}O_2$	82304-66-3	175.1104 – $[C12H15O]^+$ 133.0638 – $[C9H9O]^+$ 77.0369 – $[C6H5]^+$	Other organic compounds
10,18-Bisnorabieta-8,11,13-triene	17.99	$C_{18}H_{26}$	32624-67-2	242.2007 – $[C18H26]^+$ 227.1790 – $[C17H23]^+$ 143.0864 – $[C11H11O]^+$	Other organic compounds
2,6-Di-t-butyl-4-(2-phenylpropan-2-yl)phenol	18.71	$C_{23}H_{32}O$	34624-81-2	324.2438 – $[C23H32O]^+$ 309.2212 – $[C22H29O]^+$ 119.0836 – $[C9H11]^+$	Other organic compounds
Phenol 2,2'-methylenebis[6-(1,1-dimethyl)-4-methyl	21.77	$C_{23}H_{32}O_2$	119-47-1	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other organic compounds
Phenol 2,4-bis(1-methyl-1-phenylethyl)-	22.52	$C_{24}H_{26}O$	2772-45-4	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other organic compounds

Cont. Table 2

2,4-Bis(dimethylbenzyl)-6-*t*-butylphenol	22.57	$C_{28}H_{34}O$	244080-16-8	386.2617 – $[C28H34O]^+$ 371.2370 – $[C27H31O]^+$ 293.1897 – $[C21H25O]^+$	Other organic compounds
n-Hexadecylindane	23.93	$C_{25}H_{42}$	55334-29-7	117.0351 – $[C9H9]^+$ 130.0427 – $[C10H10]^+$ 154.1345 – $[C11H22]^+$	Other organic compounds

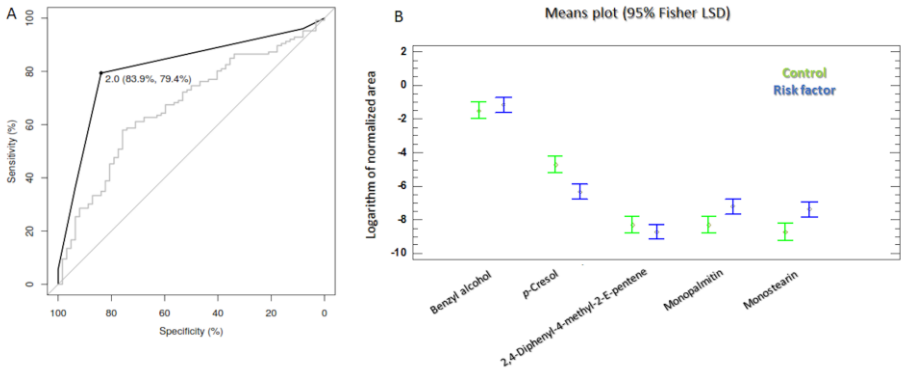


Figure 2. (A) ROC curves of the panel to discriminate between control and risk factor groups (black) and the best individual marker –indole– (grey). (B) Means plot presenting the concentration variability of the five metabolites of the panel –monopalmitin, benzyl alcohol, monostearin, 2,4-diphenyl-4-methyl-2-E-pentene and *p*-cresol– in the two groups.

Table 3. Parameters of the ROC curves obtained for each panel by selecting maximum accuracy.

Panel	% Sensitivity (95% CI*)	% Specificity (95% CI)	% AUC (95% CI)	% Negative Prediction Value	% Positive Prediction Value
Risk factor *vs* Control	77.9 (67.6-86.8)	90.3 (82.3-96.8)	85.1 (78.8–91.4)	66.6	90.90
Lung cancer *vs* Control	82.4 (73.5–91.2)	88.7 (80.6–95.2)	87.7 (81.9–93.6)	82.85	88
Lung cancer *vs* Risk factor	86.8 (77.9-94.1)	67.5 (58.7-75.4)	77.5 (71.3–83.6)	87.63	63.95

*CI: confidence interval.

3.2. Development of panels with prediction capability for lung cancer

In this case, two independent panels were built to discriminate lung cancer patients from the other two cohorts. Both panels would allow minimizing the number of patients subjected to invasive confirmatory tests, thus reducing the probability of false negative cases. On the other hand, the probability of false positive cases is less critical since the certainty could be supported on a

confirmatory test. For this reason, the configuration of panels was also supported on maximum accuracy, maximizing both sensitivity and specificity.

Best results were obtained when confronting lung cancer individuals to healthy controls, which is quite logical because the absence of lung damage and risk factor in the latter. In fact, the combination of only four metabolites led to a panel with 88.7% of specificity, 82.4% of sensitivity and 87.7% of AUC. The four compounds included in this panel were *p*-cresol, eicosenamide, hexadecylindane and cumyl alcohol. The inclusion of an additional compound did not improve these results. Figure 3A shows the ROC curve obtained with this panel in comparison with that provided by *p*-cresol, the individual marker with the highest discrimination capability (74.2% specificity and 77.9% sensitivity, while the AUC was 75%). The main characteristics of the panel are listed in Table 3, while the characteristics of cross-validation are listed in Supplementary Table 1. Regarding the levels of the metabolites included in this panel (Figure 3B), all of them were more concentrated in lung cancer patients than in healthy individuals, except for *p*-cresol, that was present at higher levels in the latter.

Discrimination of lung cancer patients versus the risk factor individuals resulted in one other 5-compounds panel characterized by sensitivity close to 90%, and composed by 2,4-bis-dimethylbenzyl-6-t-butylphenol, monostearin, spiro-2,4-heptane-1,5-dimethyl-6-methylene, 13-heptadecyn-1-ol and methyl stearate. Concretely, parameters obtained for this panel were 67.5 % of specificity, 86.8 % of sensitivity, and an AUC of 77.5%, while for the best individual biomarker [2,6-bis-1,1-dimethylethyl-4-(1-methyl-1-phenylethyl)phenol] were 61.1% of specificity, 70% of sensitivity and an AUC of 66.5%. It is worth mentioning that the main goal for this panel comparing lung cancer patients and risk factor individuals was maximizing sensitivity over specificity in order to prioritize detection of cancer patients. On the other hand, the specificity was acceptable since risk factor individuals identified as cancer patients would be subjected to a confirmatory test. In this sense, this panel would be applied as screening tool. Figure 4A shows the ROC curve for this panel in comparison to

that provided for the best individual biomarker; while Table 3 shows the characteristics of the panel in terms of positive and negative prediction values. Results obtained by cross-validation of this panel are listed in Supplementary Table 1. A comparison between the two groups in terms of concentration of the compounds included in this panel revealed that only monostearin presented less relative concentration in risk factor individuals than in those affected by lung cancer, as Figure 4B shows.

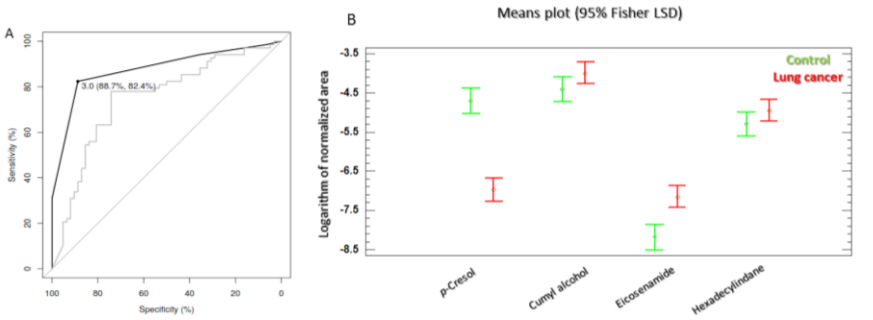


Figure 3. (A) ROC curves of the panel to discriminate between control and lung cancer group (black) and the best individual marker –*p*-cresol – (grey). (B) Means plot showing the variability in concentration of the four metabolites of the panel –*p*-cresol, eicosenamide, hexadecylindane and cumyl alcohol – in the two groups.

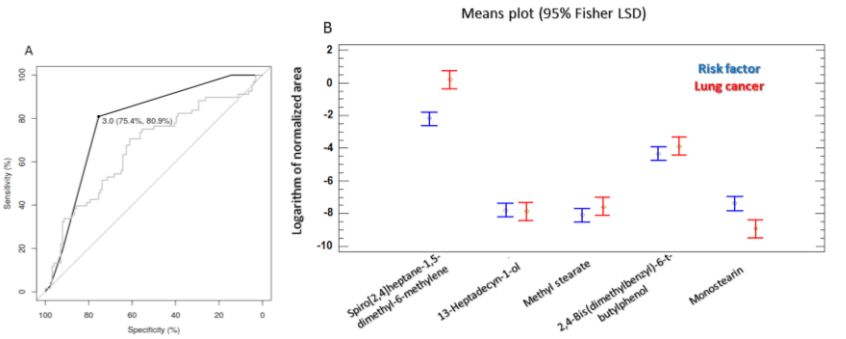


Figure 4. (A) ROC curves of the panel to discriminate between risk factor and lung cancer group (black) and the best individual marker –2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)phenol– (grey). (B) Means plot showing the concentration variability of the five metabolites in the panel –2,4-bis-dimethylbenzyl-6-t-butylphenol, monostearin, spiro-2,4-heptane-1,5-dimethyl-6-methylene, 13-heptadecyn-1-ol and methyl stearate– in all, risk factor and lung cancer individuals of the cohort.

3.3. Biological relevance of the compounds involved in the panels

Once the panels were configured for each case the compounds involved in them were evaluated from a biological perspective. Preliminarily, it is worth emphasizing the presence of compounds that can be associated to smoking habit. This is the case with *p*-cresol and benzyl alcohol, two compounds that can be obtained from toluene [21-23], which is one of the major components of tobacco smoke [24]. In fact, the levels of both compounds were strongly related, as deduced from their relative concentration variability represented in mean plots. As can be seen in Figure 5A, benzyl alcohol was clearly detected at higher levels in lung cancer patients than in the resting individuals, without statistical difference between healthy individuals and the risk factor cohort, while *p*-cresol showed the opposite pattern (Figure 5B). Toluene is mainly metabolized to benzyl alcohol, which is preferentially converted to hippuric acid, and, in a small proportion, to o- and *p*-cresol [25]. The differences in relative concentrations for benzyl alcohol and *p*-cresol in the three groups under study can be explained by considering that an increase of toluene in the body should enhance the production of benzyl alcohol as a preferential pathway. On the other hand, the difference between *p*-cresol and benzyl alcohol production is not so evident for small amounts of toluene. Therefore, the high relative concentration of benzyl alcohol in lung cancer patients as compared to risk factor individuals seems to reveal a difference in the metabolism of toluene between both groups, as risk factor individuals reported concentration of benzyl alcohol similar to healthy individuals. This hypothesis is also supported on the opposite pattern observed for *p*-cresol.

A similar situation occurs with five phenol derivatives such as 2,4-diphenyl-4-methyl-2-E-pentene, 2,4,6-triisopropylphenol, 2,6-bis(1,1-dimethyl ethyl)-4-(1-methyl-1-phenylethyl)phenol, 2,4-bis(1-methyl-1-phenylethyl) phenol and 2,4-bis(dimethylbenzyl-6-t-butylphenol) since their presence in EBC could also be ascribed to smoking habit as they can be directly related to crude oil [26]. One other compound involved in the panels is benzoic acid methyl ester, the increased level of which in lung cancer as compared to the other groups [27] can

also be attributed to smoking, as a product of xylene degradation pathway. The latter is an organic compound present in cigarette smoke.

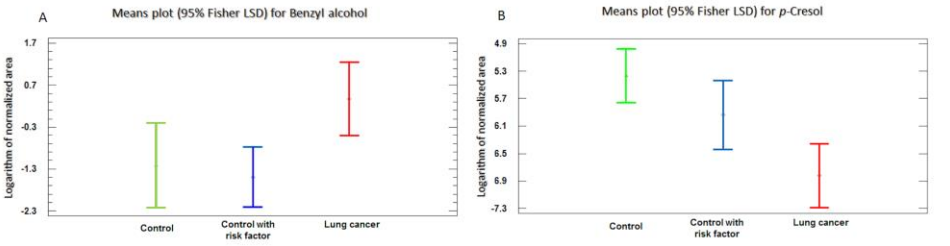


Figure 5. Variability in the concentration of benzyl alcohol (A) and *p*-cresol (B) by comparing lung cancer patients, risk factor and control individuals.

Concerning the rest of compounds in the panels, it should be emphasized the presence of two monoacylglycerols that were previously found by Calderón-Santiago et al. in sweat [6]. These compounds –palmitic acid (C16:0) and stearic acid (C18:0) derivatives–, together with oleic acid (C18:1), constitute about 78% of all free fatty acids in the circulatory system [28].

Other compounds such as cumyl alcohol and indane derivatives have been described in exhaled vapor by Phillips et al. [29], who selected one of the indane derivatives as potential biomarker of lung cancer [30]. On the other hand, the fatty acid derivative eicosenamide –previously studied because of its potential as antimicrobial [31], reducing and stabilizing agent [32]– had not been identified yet in any biofluid; similarly occurs with 13-heptadecyn-1-ol and spiro-2,4-heptane-1,5-dimehtyl-6-methylene, whose presence in breath or biological implication in human processes had not previously been reported.

It is worth emphasizing the proportion of compounds detected in EBC, the origin of which could be attributed to exogenous sources. This is the case with some compounds such as menthol, limonene or eucalyptol, which were not involved in any tentative panel. However, the relationship between the microbiote and EBC composition should not be forgotten since this connection could explain

the presence of some compounds such as indole or phenolic compounds and derivatives. Additionally, the role of the microbiome in EBC composition should be considered. There are considerable gaps about the role of microbiota in carcinogenesis in organs with a substantial bacterial microbiome, as is the case of lungs. However, several findings point out a possible role for bacteria in the promotion of lung cancer, such as the increased bacterial colonization in COPD [33,34], which is also a known risk factor for lung cancer development [35]. Therefore, it is quite important to take into account the contribution of the airway microbiome to explain the observed variability profiles associated to the different groups [36,37].

3.4. Metabolomic differences among lung cancer stages

In a preliminary approach to finding significant metabolites to differentiate among lung cancer stages, patients with the two most common lung cancer diagnostics (squamous cell carcinoma and adenocarcinoma) were selected to compare their EBC composition. These individuals were classified into two groups: one of them of 10 individuals belonging to early stages I and II, and the other of 28 belonging to advanced stages III and IV, which were compared with the risk factor cohort to find discrimination trends in compounds detected in EBC. An ANOVA test was applied to each compound to find statistical differences among the three groups of patients (risk factor group, early stage lung cancer and advanced stage lung cancer) with a *p*-value lower than 0.05. Non-smokers were not considered in this study although they had been included in mean plots to compare the concentration levels of each compound. Six compounds resulted statistically significant among the three groups under study: cumyl alcohol, benzoic acid methyl ester, 2,4,6-triisopropylphenol, 2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)phenol, 2,4-bis(1-methyl-1-phenylethyl)phenol and 2,4-bis(dimethylbenzyl-6-t-butylphenol. Figure 6 shows means plot for these compounds comparing the risk factor group with the two lung cancer subgroups. As can be seen, all compounds were detected at higher concentrations in lung cancer patients at advanced stage. It is worth emphasizing that 5 of these

compounds have a benzenoid structure with alkyl groups which could be related to tobacco consumption or, as previously emphasized, to metabolic variations ascribed to airways microbiome.

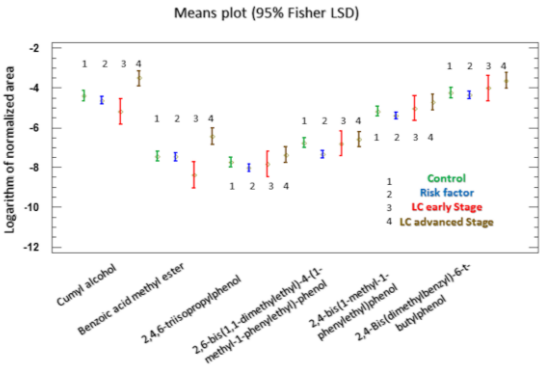


Figure 6. Variability in the concentration of the six statistically significant compounds –cumyl alcohol, benzoic acid methyl ester, 2,4,6-triisopropylphenol, 2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)phenol, 2,4-bis(1-methyl-1-phenylethyl)phenol and 2,4-bis(dimethylbenzyl-6-t-butylphenol– by comparison of risk factor individuals and patients diagnosed with early and advanced stages of lung cancer.

Concerning benzoic acid methyl ester, this compound could also be related to benzyl alcohol as it is metabolized to benzoic acid [38] and, therefore, its presence could be similarly related to smoking as do the four benzene-related compounds [20], although metabolic alterations promoted by the airway microbiome-host interactions should also be considered.

Conclusions

According to the foregoing, one of the main shortcomings of lung cancer diagnostic is the stage at which it is detected, mainly owing to either the invasive character of some of the existing tests or the high cost associated to others, thus making difficult their application with screening purposes to all individuals included in the risk factor population. Thus, more economical screening tests are

demanded for decreasing the number of individuals that would be subjected to confirmatory tests.

EBC has not been widely exploited in the clinical field despite the advantages associated to its sampling. In this research, the potential of EBC as biofluid for implementation in lung cancer diagnosing tools has been pointed out. Prediction models based on panels of metabolites that include monoacylglycerol derivatives, fatty acid methyl esters, monoterpenes and tobacco related compounds have been built to discriminate between patients with lung cancer and control groups with and without risk factor. The high negative predictive value thus obtained indicates the potential use of this biofluid to reduce the number of cases that should be subjected to confirmatory tests.

These preliminary results emphasize the necessity of a large-scale study to validate the proposed panels and assess its applicability to detect lung cancer at early stages. This first approach has also revealed statistically significant differences between early and advanced lung cancer stages, mainly affecting to benzene derivatives.

Acknowledgements

The Spanish Ministerio de Economía y Competitividad (MINECO), Junta de Andalucía and FEDER programme are gratefully acknowledged for financial support (Projects “Development of methods for early cancer detection, December 29, 2011”, and “Optimización y aplicación de plataformas metabolómicas de análisis de biofluidos no invasivos para la búsqueda de biomarcadores de diagnóstico precoz de cáncer de pulmón, FQM-1602”). Xavier Robin et al. are also acknowledged for allowing us to use the PanelomiX tool. A. Peralbo-Molina and M. Calderón-Santiago also thank to these organisms their research contracts. F. Priego-Capote is also grateful to the Ministerio de Ciencia e Innovación (MICINN) for a Ramón y Cajal Contract (RYC-2009-03921). Reina

Sofía Hospital and its biorepository node (Córdoba, Spain) is also grateful for donors of exhaled breath condensate samples.

References

[1] Jemal A, Bray F, Center M M, Ferlay J, Ward E and Forman D 2011 Global cancer statistics CA. Cancer J. Clin. 61 69–90

[2] Burney P et al 2014 Chronic obstructive pulmonary disease mortality and prevalence: the associations with smoking and poverty: a bold analysis Thorax 69 465–73

[3] Jemal A, Center M M, DeSantis C and Ward E M 2010 Global patterns of cancer incidence and mortality rates and trends Cancer Epidemiol. Biomarkers Prev. 19 1893–907

[4] Boiselle P M 2013 Computed tomography screening for lung cancer JAMA 309 1163–70

[5] Álvarez-Sánchez B, Priego-Capote F and Luque de Castro M D 2010 Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation TrAC Trends Anal. Chem. 29 111–9

[6] Calderón-Santiago M, Priego-Capote F, Jurado-Gámez B and Luque de Castro M D 2014 Optimization study for metabolomics analysis of human sweat by liquid chromatography-tandem mass spectrometry in high resolution mode J. Chromatogr. A 1333 70–8

[7] Kim J O, Gazala S, Razzak R, Guo L, Ghosh S, Roa W H and Bedard E L R 2015 Non-small Cell Lung Cancer Detection Using MicroRNA Expression Profiling of Bronchoalveolar Lavage Fluid and Sputum Anticancer Res 35 1873–80

[8] Phillips M, Gleeson K, Hughes J M, Greenberg J, Cataneo R N, Baker L and McVay W P 1999 Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. Lancet 353 1930–3

[9] Phillips M, Cataneo R N, Cummin A R C, Gagliardi A J, Gleeson K, Greenberg J, Maxfield R a and Rom W N 2003 Detection of lung cancer with volatile markers in the breath Chest 123 2115–23

[10] Phillips M et al 2007 Prediction of lung cancer using volatile biomarkers in breath Cancer Biomark. 3 95–109

[11] Wang Y, Hu Y, Wang D, Yu K, Wang L, Zou Y, Zhao C, Zhang X, Wang P and Ying K 2012 The analysis of volatile organic compounds biomarkers for lung cancer in exhaled breath, tissues and cell lines. Cancer Biomark. 11 129–37

[12] Kubáň P and Foret F 2013 Exhaled breath condensate: determination of non-volatile compounds and their potential for clinical diagnosis and monitoring. A review Anal. Chim. Acta 805 1–18

[13] Costa C, Bucca C, Bergallo M, Solidoro P, Rolla G and Cavallo R 2011 Unsuitability of exhaled breath condensate for the detection of herpesviruses DNA in the respiratory tract J. Virol. Methods 173 384–6

[14] Dragonieri S, Brinkman P, Mouw E, Zwinderman A H, Carratú P, Resta O, Sterk P J and Jonkers R E 2013 An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls Respir. Med. 107 1073–8

[15] Hunt J 2002 Exhaled breath condensate: An evolving tool for noninvasive evaluation of lung disease J. Allergy Clin. Immunol. 110 28–34

[16] Mazzatenta A, Di Giulio C and Pokorski M 2013 Pathologies currently identified by exhaled biomarkers Respir. Physiol. Neurobiol. 187 128–34

[17] Peralbo-Molina A, Calderón-Santiago M, Priego-Capote F, Jurado-Gámez B and Luque de Castro M D 2015 Development of a method for metabolomic analysis of human exhaled breath condensate by gas chromatography–mass spectrometry in high resolution mode Anal. Chim. Acta 887 118–26

[18] Calderón-Santiago M, Priego-Capote F, Turck N, Robin X, Jurado-Gámez B, Sanchez J C and Luque de Castro M D 2015 Human sweat metabolomics for lung cancer screening. Anal. Bioanal. Chem.

[19] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C and Müller M 2013 PanelomiX: A threshold-based algorithm to create panels of biomarkers Transl. Proteomics 1 57–64

[20] Kayode R M O and Afolayan A J 2015 Cytotoxicity and effect of extraction methods on the chemical composition of essential oils of Moringa oleifera seeds. J. Zhejiang Univ. Sci. B 16 680–9

[21] Pierce C H, Chen Y, Dills R L, Kalman D A and Morgan M S 2002 Toluene metabolites as biological indicators of exposure Toxicol. Lett. 129 65–76

[22] Kim K-W, Won Y L and Ko K S 2015 Ethnic Differences in the Metabolism of Toluene: Comparisons between Korean and Foreign Workers Exposed to Toluene. Toxicol. Res. 31 25–32

[23] Inoue O, Seiji K, Watanabe T, Nakatsuka H, Jin C, Liu S-J and Ikeda M 1993 Effects of smoking and drinking on excretion of hippuric acid among toluene-exposed workers Int. Arch. Occup. Environ. Health 64 425–30

[24] World Health Organization 2000 Toluene Air Quality Guidelines for Europe ed World Health Organization (Copenhagen, Denmark: WHO Regional Publications, European Series No. 91) pp 1–20

[25] Nakajima T, Wang R-S, Elovaara E, Gonzalez F J, Gelboin H V., Raunio H, Pelkonen O, Vainio H and Aoyama T 1997 Toluene metabolism by cDNA-Expressed human hepatic cytochrome P450 Biochem. Pharmacol. 53 271–7

[26] Liang Y, Nostrand J D Van, Wang J, Zhang X, Zhou J and Li G 2009 Microarray-based functional gene analysis of soil microbial communities during ozonation and biodegradation of crude oil. Chemosphere 75 193–9

[27] Darrall K G, Figgins J A, Brown R D and Phillips G F 1998 Determination of benzene and associated volatile compounds in mainstream cigarette smoke. Analyst 123 1095–101

[28] Quehenberger O et al 2010 Lipidomics reveals a remarkable diversity of lipids in human plasma J. Lipid Res. 51 3299–305

[29] Phillips M 1999 Variation in volatile organic compounds in the breath of normal humans J. Chromatogr. B 729 75–88

[30] Phillips M et al 2008 Detection of lung cancer using weighted digital analysis of breath biomarkers. Clin. Chim. Acta. 393 76–84

[31] Kabara J J, Swieczkowski D M, Conley A J and Truant J P 1972 Fatty acids and derivatives as antimicrobial agents. Antimicrob. Agents Chemother. 2 23–8

[32] Malapermal V, Botha I, Krishna S B N and Mbatha J N 2015 Enhancing antidiabetic and antimicrobial performance of Ocimum basilicum, and Ocimum sanctum (L.) using silver nanoparticles Saudi J. Biol. Sci.

[33] Pragman A A, Kim H B, Reilly C S, Wendt C and Isaacson R E 2012 The lung microbiome in moderate and severe chronic obstructive pulmonary disease. PLoS One 7 e47305

[34] Sethi S and Murphy T F 2008 Infection in the pathogenesis and course of chronic obstructive pulmonary disease. N. Engl. J. Med. 359 2355-65

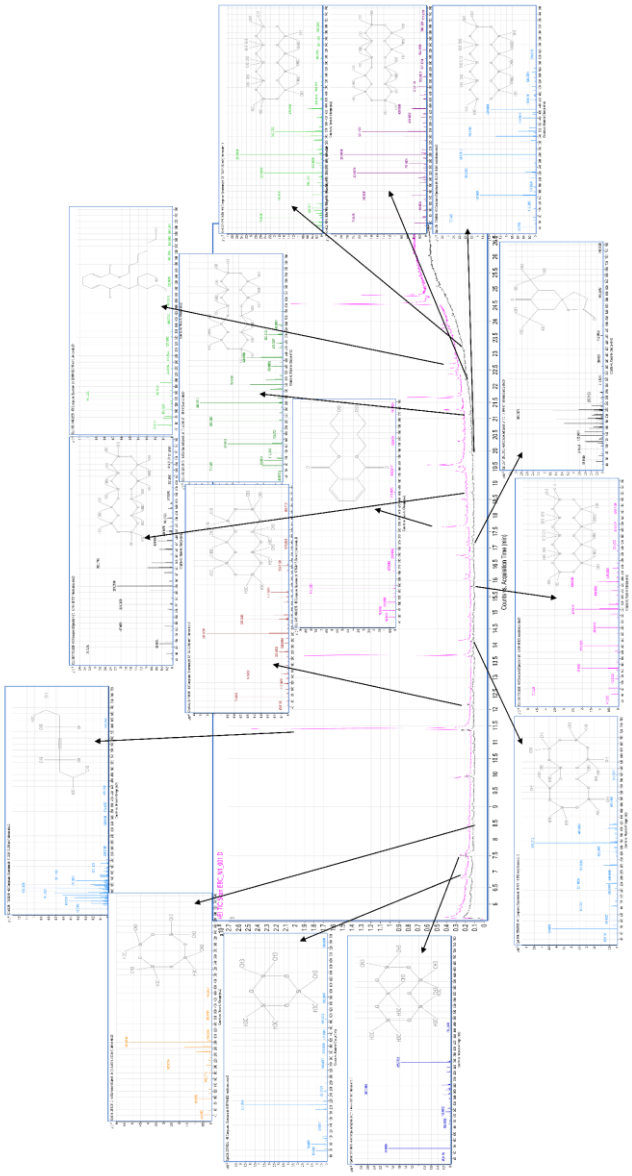
[35] Houghton A M 2013 Mechanistic links between COPD and lung cancer. Nat. Rev. Cancer 13 233–45

[36] Schwabe R F and Jobin C 2013 The microbiome and cancer. Nat. Rev. Cancer 13 800–12

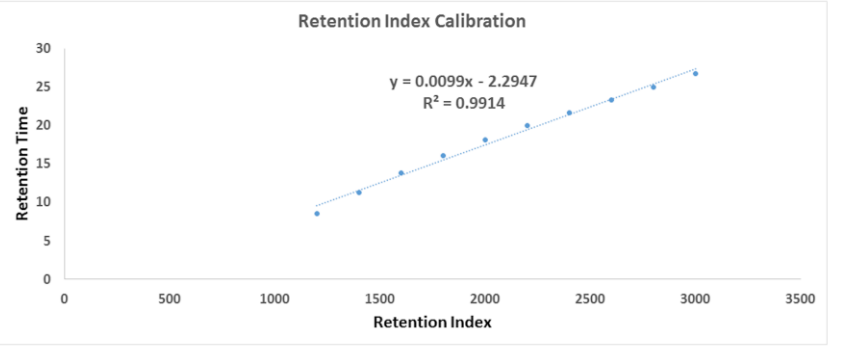
[37] Dickson R P, Erb-Downward J R and Huffnagle G B 2013 The role of the bacterial microbiome in lung disease. Expert Rev. Respir. Med. 7 245–57

[38] Nair B 2001 Final report on the safety assessment of benzyl alcohol, benzoic acid, and sodium benzoate. Int. J. Toxicol. 20 Suppl 3 23-50

Supplementary material



Supplementary Figure 1. Comparison of chromatograms provided by analysis of an EBC sample and a blank sample to identify interferents attributed to the sampling protocol and analysis.



Supplementary Figure 2. RI calibration curve from analysis of the alkane standard mixture.

Supplementary Table 1. Parameters of the ROC curves obtained by cross-validation for each panel with selection of maximum accuracy.

Cross-Validation	% Sensitivity (95% CI)	% Specificity (95% CI)	% AUC (95% CI)
Risk factor *vs* Control	80.9 (70.6-89.7)	62.9 (51.6-74.2)	78.0 (70.3-86.0)
Lung cancer *vs* Control	73.5 (63.2-83.8)	80.6 (69.4-90.3)	81.9 (74.8-89.0)
Lung cancer *vs* Risk factor	77.9 (67.6-86.8)	63.5 (54.8-71.4)	73.7 (66.9-80.5)

Capítulo IX:

Exhaled breath condensate to discriminate individuals with different smoking habits by GC-QTOF

Sent to

Journal of Breath Research

IOPscience

Exhaled breath condensate to discriminate individuals with different smoking habits by GC-QTOF

Ángela Peralbo-Molina[a,b], Mónica Calderón-Santiago[a,b], Feliciano Priego-Capote[a,b*], Bernabé Jurado-Gámez[b,c], María Dolores Luque de Castro[a,b*]

[a]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14014, Córdoba, Spain*

[b]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, E-14014, Córdoba, Spain*

[c] *Department of Respiratory Medicine, Reina Sofia University Hospital, University of Córdoba, 14004, Córdoba, Spain*

Exhaled breath condensate to discriminate individuals with different smoking habits by GC–QTOF

Ángela Peralbo Molina, Mónica Calderón-Santiago, Feliciano Priego-Capote, Bernabé Jurado-Gámez, María Dolores Luque de Castro

Abstract

Smoking is a crucial factor contributing to respiratory diseases and lung inflammation, which are the reason of a high mortality in Europe. Despite the negative impact that tobacco consumption causes on health, few metabolomics studies have compared the composition of biofluids from smoker and non-smoker individuals. Exhaled breath condensate (EBC) is one of the biofluids less employed for clinical studies despite its non-invasive sampling and the foreseeable relationship between its composition and respiratory diseases. EBC was used in this research as clinical sample to compare among three groups of individuals: current smokers (CS), former smokers (FS) and never smokers (NS). Special attention was paid to the cumulative consumption expressed as smoked pack-year. The levels of 12 metabolites found statistically significant among the three groups of individuals were discussed to find an explanation to their altered levels. Significant compounds included monoacylglycerol derivatives, terpenes and other compounds, the presence of which could be associated to the influence of smoking on the qualitative and quantitative composition of the microbiome.

Keywords: Exhaled breath condensate – Metabolomics – GC-TOF/MS – Smoking habits – Tobacco

1. Introduction

Cigarette smoke is a source of toxicants that enter into contact with the microbiome in the respiratory tract perturbing the microbial ecology with direct incidence on immunology, oxygen deprivation or other mechanisms [1,2]. Loss of beneficial species and appearance of new species in the respiratory tract due to smoking can lead to colonization by pathogens and ultimately to diseases. Thus, smoking is a key contributor to respiratory diseases, which are considered the most preventable cause of premature death in the United States [3] and the major factor for development of chronic obstructive pulmonary disease (COPD) and lung cancer [4], two diseases that cause a large part of respiratory deaths in Europe [5].

Smoking cessation has major and immediate health benefits for individuals of all ages. Several studies have highlighted that cardiovascular risk is reduced in smokers who give up cigarettes even after a first myocardial infarction [7,8]. For COPD patients, smoking cessation has shown to be an integral component of rehabilitation courses [4] and also decreases the risk of different cancers, with special emphasis on lung cancer, and stroke [9]. In a recent study, Jerjes et al. demonstrated that reduction/cessation of chronic smoking and drinking tends to significantly reduce mortality in patients with oral cancer [10]. Moreover, other reports concluded that former smokers (FS) live longer than current smokers (CS). For example, persons who quit smoking before age 50 have one-half the risk of dying during the next 15 years compared with continuing smokers [11–13]. In this context, Boué et al. studied the effect of cigarette smoke exposure along with the benefits of its cessation concluding that cigarette smoke exposure induced a series of metabolic as well as adaptive and innate inflammatory responses in murine respiratory tissues that were progressively deactivated by smoking cessation [14]. Nevertheless, for some cancers, especially for adenocarcinoma, the risk remains high in FS compared with non-smokers (NS) [15,16].

Concerning 'omics' approaches, metabolomics seems to be an ideal tool to assess the impact of cigarette smoke on human exposure and health. Notwithstanding the importance of knowing how the metabolism is affected by smoking habit, few metabolomics studies have been targeted at comparing the composition of any biofluid from CS and NS [12,17]. In a recent study, the tobacco-related global metabolome in blood from CS and NS was identified by UPLC-QTOF [18]. Apart from identification of nicotine metabolites, there was a characteristic metabolic profile contributing to discriminate CS and NS individuals. These metabolites, assessed by MS/MS, showed the potential to identify metabolic phenotypes and new metabolites related to cigarette exposure and toxicity [18].

Other relevant study on association of metabolite concentrations with smoking habit showed the reversion of metabolite variations after smoking cessation and demonstrated the results using protein-metabolite networks [12]. The study reported the association of smoking with variations in the concentration of amino acids, ether lipids and glycerophospholipid metabolism. The smoking-related changes in the serum metabolite profile were found reversible after quitting smoking. This fact indicates the remarkable benefits of smoking cessation and provides a link to cardiovascular diseases benefits.

Contrarily to expectations, exhaled breath condensate (EBC) has been scarcely used for metabolomics analysis of the smoking habit [17]. The aim of the present research was to identify metabolites related to this habit by comparison of EBC samples from CS, FS and NS. For this purpose, metabolite profiles were obtained using the methodology developed by Peralbo-Molina et al. based on gas chromatography coupled to mass spectrometry in high resolution mode (GC-TOF/MS) [19] using samples from CS, FS and NS individuals. This study would allow identifying compounds associated to the smoking habit, which are accumulated or not in EBC, and others that either can disappear or appear after quitting the habit.

2. Experimental

2.1. Cohort selected for the study

EBC samples were collected from 119 individuals early in the morning before breakfast, and stored at -80 ºC in the biorepository of the Reina Sofía Hospital (Córdoba, Spain). All the individuals were recruited at the Department of Respiratory Medicine. The cohort included 61 individuals who had never smoked (NS). Table 1 shows the main characteristics of the group expressed as mean value and standard deviation. The target individuals were aged 59±8 years and 77% of them were males. The current smokers (CS) group was formed by 32 individuals 58±7 years old, and 75% of them were males. The former smokers (FS) group included 26 individuals 60±8 years old, and 92% of them were males. Supplementary Table 1 includes data from each individual such as age, sex and cigarettes consumption expressed as smoked pack-year. This parameter is frequently used to rank smoker individuals as a function of the cumulative smoking habit (0, 0-24, 24-45, >45 pack-year). Independence of sex and age among groups was assessed by Fisher's exact and Mann-Whitney tests, respectively.

Table 1. Generic information about each group including means of age and body mass index.

Type	Number	Sex (% male)	Age	BMI
NS	61	77.05	59±8	29±4
FS	26	92.31	60±8	30±4
CS	32	75	58±7	29±5

2.2. Procedure for EBC collection

The ECOScreen2 device (FILT Thorax-und LungenDiagnostik GmbH, Berlin, Germany) used for sampling collects and condenses the EBC in disposable polyethylene bags at -20 ºC and is able to collect EBC into two separate bags for physical separation between the air contained in the upper

airway and that in the distal airway [20]. This configuration makes saliva contamination highly unlikely [21]. The sampler was modified by insertion of a commercial protection filter from Scharlab (Barcelona, Spain) over the inlet air valve to avoid the entrance of exogenous organic compounds and particles from the room atmosphere. This filter was periodically changed to avoid saturation.

In addition, some conditions were taken into account by patients in order to avoid contamination: (1) smokers were advised to avoid the consumption of tobacco at least 12 hours prior to collection; (2) patients attend on an empty stomach, they may take only the prescribed medication with enough water to ensure good hydration; (3) patients should report the medication consumed;(4) inhalers do not have to be inhaled prior to collection, if there were necessary, sampling should be postponed; (5) the collection day, patients must not use colony, makeup or lipstick; (6) patients do not have to perform forced expiratory maneuvers prior to collection.

Tidal breathing and a nose-clip were used for 15 min, time required to collect an average EBC volume of 1.5 ml from the distal airway. The samples were divided into 100 µl aliquots and the vials stored at –80 ºC until analysis. All samples were analyzed within 3 months after collection.

2.3. Reagents

Hexane TraceSELECT® grade from Sigma-Aldrich (St. Louis, USA) was used as organic solvent for sample preparation. An alkane standard mixture (from C10 to C40), also from Sigma-Aldrich, was tested by GC-MS analysis to establish the retention index (RI) calibration. Deionized water (18 mΩ·cm) provided by a Milli-Q water purification system from Millipore (Darmstadt, Alemania) was also used.

2.4. Instruments and apparatus

Homogenization of the extracts was carried out by an MS2 Vortex (IKA, Germany). An Agilent 7890A Series GC system coupled to an Agilent 7200 UHD

Accurate-Mass QTOF hybrid mass spectrometer equipped with an electron impact (EI) source (Santa Clara, CA, USA) was used. The analytical sample was monitored in high resolution mode.

2.5. Sample preparation

Sample preparation consisted of liquid-liquid extraction using hexane as extractant. In all cases, 100 µl aliquots of EBC were 1:1 vortexed with hexane in a glass insert at room temperature for 1 min. Then, the organic phase was isolated and put into a new glass insert for analysis. To eliminate exogenous interferents, blanks prepared with water were analyzed as the samples.

2.6. GC-TOF/MS analysis

GC-TOF/MS analyses were performed by electron impact ionization (EI) mode at 70 eV and controlled by MassHunter Acquisition B.06. The GC separation was carried out by a fused silica DB-5MS-UI 30 m×0.25 mm i.d. 0.25 µm film thickness capillary column. The temperature program of the GC oven started at 60 °C (1 min held), followed by a temperature ramp of 10 °C min-1 to final 300 °C (2 min held). Post-run time was programmed for 4 min up to 310 ºC to assure complete elution of the injected sample. Pulsed splitless injections of 1 µl sample were carried out at 250 °C, and ultrapure grade helium was used as the carrier gas at 1.0 ml min-1. The interface and ion source were set at 280 and 300 °C, respectively. A solvent delay of 5.5 min was used to prevent damage in the ion source filament. The TOF detector was operated at 5 spectra s-1 in the mass range m/z 50-550 and the resolution was 8500 (full width half maximum, FWHM) at m/z 501.9706. Mass-spectrometric grade PFTBA (perfluorotri-n-butylamine) was used for daily mass calibration. Tentative identification of metabolites was performed by searching MS spectra on the NIST 11 database taking into account the RI values.

2.7. Data processing and statistical analysis

Unknown Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to process all data obtained by GC¬–TOF/MS in full scan mode. Treatment of raw data files started by deconvolution of potential molecular features (MFs) with the suited algorithm included in the software. For this purpose, the deconvolution algorithm considered all ions exceeding 1500 counts for the absolute height parameter. Additionally, the accuracy error and the window size factor were set at 50 ppm and 150 units, respectively. After extraction of MFs, data files in compound exchange format (.cef files) were created for each sample and exported into the Mass Profiler Professional (MPP) software package (version 12.1, Agilent Technologies, Santa Clara, CA, USA) for further processing.

In the next step, the data were processed by alignment of the potential MFs according to their retention time and m/z value using a tolerance window of 0.3 min and an accuracy error of 15 ppm. The MFs from the analysis of blanks were removed from the data set of MFs from the EBC samples. The extraction algorithm confirmed the efficiency of this filtering step. Stepwise reduction of the MFs number was based on frequency of occurrence by comparing repetitions of the same group of individuals. A filter by frequency was set at 100%, thus ensuring detection of each MF in all the injected replicates from each of the three groups.

In the last step, the resulting MFs were exported (.cef file) for recursive analysis. For this purpose, the Quantitative Analysis software (version 7.0, Agilent Technologies, Santa Clara, CA, USA) was used to reintegrate all potential compounds found in the analyzed samples. The resulting table was exported in coma separated value format (.csv file) and reprocessed with the Mass Profiler Professional (MPP) software package. A filter to eliminate samples by setting within-replicates variability above 10% was applied to assure the effectiveness of the recursive analysis. Finally, the data set was normalized by logarithmic

transformation of the ratio between the peak area of each MF and the total useful mass spectrometry signal (MSTUS) corresponding to all the samples [19,22].

The resulting data set was then subjected to supervised analysis by partial least squares discriminant analysis (PLS-DA). Cross validation by using an N-fold model was selected as validation approach. With this model, the classes in the input data are randomly divided into N equal parts; N–1 parts are used for training and the remaining part is used for testing. The process is repeated N times, with a different subset being used for testing in an iterative process. Thus, each row is used at least once in training and once in testing, and a confusion matrix is generated. The complete process can be repeated as many times as specified by the number of repetitions. Validation in this research consisted of ten repetitions and a fold number of three.

The software Statgraphics Centurion (XVI version 16.1.18, Statpoint Technologies) and a statistical test were used to study the influence smoking habit and the smoking load on each metabolite. Finally, the platform Metaboanalyst 3.0 (http://www.metaboanalyst.ca/) was used to perform Random Forest Analysis and the ROC curves for the most significant compounds [23].

Identification was firstly carried out by searching MS spectra in the NIST11 database. Only identifications with a match factor and a reverse match factor higher than 700 were considered as valid. RI values included in the NIST database were also taken into account to support identifications. An RI calibration model was built by comparing the RI values of an alkane standard mixture (composed by alkane between C10 to C40 with an even number of carbons) with the chromatographic method used in this research and the RI values provided by the NIST database. Supplementary Figure 1 shows the RI calibration line obtained by this approach. The requirement to accept NIST identifications was that the difference between the theoretical and the experimental RI obtained by extrapolation in the calibration curve should be within ±100 units.

The NIST database does not contain high resolution MS information as provided by the TOF detector. For this reason, a third step was included to

validate identification of each MF by high resolution mass spectrometry. Thus, the molecular formula for the tentative precursor ion [M]+ and the most intense product ions obtained for each MF should fit the NIST identification by setting a cut-off value of 10 ppm in mass accuracy. Table 2 lists the 44 identified compounds classified by chemical families.

The KEGG (Kyoto Encyclopedia of Genes and Genomes) database was used for identification of the main changes occurring in the composition of EBC from individuals pertaining to each class.

3. Results

3.1. Comparison of the composition of EBC from CS and NS individuals

As mentioned above, the present research is focused on detecting modifications in EBC composition associated to smoking habit. With this aim statistical analysis was carried out to compare the composition of EBC from CS and NS individuals. Supervised analysis by PLS-DA was applied to find differences in EBC collected from the two groups with the data set including the 44 identified compounds. This analysis would allow finding discrimination patterns associated to the smoking habit. Figure 1 illustrates the 3D scores plots, in which discrimination trends were clearly observed highlighting the influence of smoking habit on data variability. The percentage of correctly classified samples for the training set of the resulting PLS-DA were 81.3 and 77.1% for CS and NS individuals, respectively; while the validation set gave an overall accuracy of 66.7%. Therefore, there is a variability pattern in the EBC composition associated to the smoking habit.

Table 2. Compounds identified in EBC by GC-TOF/MS analysis.

Compound Name	Retention Time	Formula	CAS ID	Fragments	Family
Eucalyptol	5.99	$C_{10}H_{18}O$	470-82-6	154.1361 – $[C10H18O]^+$ 139.1119 – $[C9H15O]^+$ 93.0695 – $[C7H9]^+$	Aliphatic heteropolycyclic compounds (Oxanes)
Indole	9.90	C_8H_7N	120-72-9	117.0558 – $[C8H7N]^+$ 90.0448 $[C7H6]^+$ 74.0145 – $[C6H2]^+$	Aromatic heteropolycyclid compounds (Indoles)
Benzoic acid 4-ethoxy-ethyl ester	12.92	$C_{11}H_{14}O_3$	23676-09-7	194.0425 – $[C11H14O3]^+$ 149.0581 – $[C9H9O2]^+$ 121.0269 – $[C7H5O2]^+$	Aromatic homomonocyclic compounds (Benzene and substituted derivatives)
Benzoic acid methyl ester	6.95	$C_8H_8O_2$	93-58-3	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic homomonocyclic compounds (Benzoic acid derivatives)
3,5-Di-t-butyl-4-hydroxycinnamic acid	20.16	$C_{17}H_{24}O_3$	22014-01-3	276.1712 – $[C17H24O3]^+$ 261.1479 – $[C16H21O3]^+$ 177.0896 – $[C11H13O2]^+$	Aromatic homomonocyclic compounds (Cinnamic acid derivatives)
Octinoxate	21.01	$C_{18}H_{26}O_3$	5466-77-3	290.1867 – $[C18H26O3]^+$ 178.0611 – $[C10H10O3]^+$ 161.0578 – $[C10H9O2]^+$	Aromatic homomonocyclic compounds (Cinnamic acid derivatives)
Cresol	6.54	C_7H_8O	106-44-5	136.0514 – $[C8H8O2]^+$ 105.0332 – $[C7H5O]^+$ 77.0378 – $[C6H5]^+$	Aromatic homomonocyclic compounds (Phenols and derivatives-cresol)

Cont. Table 2

Benzyl alcohol	5.96	C_7H_8O	100-51-6	108.0565 – $[C7H8O]^+$ 91.0535 – $[C7H7]^+$ 79.0533 – $[C6H7]^+$	Aromatic homomonocyclic compounds (Primary alcohols)
Isopropyl laurate	14.97	$C_{15}H_{30}O_2$	10233-13-3	201.1835 – $[C12H25O2]^+$ 157.1203 – $[C9H17O2]^+$ 102.0656 – $[C5H10O2]^+$	Lipids (Fatty acid esters)
Palmitic acid methyl ester	17.34	$C_{17}H_{34}O_2$	112-39-0	270.2545 – $[C17H34O2]^+$ 227.1998 – $[C14H27O2]^+$ 143.1048 – $[C8H15O2]^+$	Lipids (Fatty acid esters)
Oleic acid methyl ester	19.04	$C_{19}H_{36}O_2$	112-62-9	296.2702 – $[C19H36O2]^+$ 264.2442 – $[18C32HO]^+$ 81.0685 – $[C6H9]^+$	Lipids (Fatty acid esters)
Stearic acid methyl ester	19.28	$C_{19}H_{38}O_2$	112-61-8	298.2862 – $[C19H38O2]^+$ 255.2315 – $[C16H31O2]^+$ 87.0436 – $[C4H7O2]^+$	Lipids (Fatty acid esters)
Palmitoleic acid	17.45	$C_{16}H_{30}O_2$	373-49-9	236.2122 – $[C16H28O]^+$ 98.0710 – $[C6H10O]^+$ 69.0689 – $[C5H9]^+$	Lipids (Fatty acids and conjugates)
Palmitic acid	17.69	$C_{16}H_{32}O_2$	57-10-3	227.1997 – $[C14H27O2]^+$ 129.0891 – $[C7H13O2]^+$ 73.0279 – $[C3H5O2]^+$	Lipids (Fatty acids and conjugates)
Stearic acid	19.58	$C_{18}H_{36}O_2$	57-11-4	284.2706 – $[C18H36O2]^+$ 129.0908 – $[C7H13O2]^+$ 73.0281 – $[C3H5O2]^+$	Lipids (Fatty acids and conjugates)
Glycidol stearate	22.48	$C_{21}H_{40}O_3$	7460-84-6	297.2436 – $[C18H33O3]^+$ 98.0719 – $[C6H10O]^+$ 71.0848 – $[C5H11]^+$	Lipids (Fatty acids and conjugates)

Cont. Table 2

Undecanol	11.25	$C_{11}H_{24}O$	112-42-5	111.1157 – $[C8H15]^+$ 83.0844 – $[C6H11]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty alcohols)
1-Hexadecanol 2-methyl	15.99	$C_{17}H_{36}O$	2490-48-4	111.1160 – $[C8H15]^+$ 97.1006 – $[C7H13]^+$ 69.0691 – $[C5H9]^+$	Lipids (Fatty alcohols)
Oleamide	21.33	$C_{18}H_{35}NO$	301-02-0	281.2679 – $[C18H35NO]^+$ 126.0914 – $[C7H12NO]^+$ 72.0438 – $[C3H6NO]^+$	Lipids (Fatty amides)
11-Eicosenamide	23.01	$C_{20}H_{39}NO$	10436-08-5	126.0915 – $[C7H12NO]^+$ 72.0439 – $[C3H6NO]^+$ 309.2973 – $[C20H39NO]^+$	Lipids (Fatty amides)
Erucamide	24.57	$C_{22}H_{43}NO$	112-84-5	337.3338 – $[C22H43NO]^+$ 126.0916 – $[C7H12NO]^+$ 72.0440 – $[C3H6NO]^+$	Lipids (Fatty amides)
Monopalmitin	22.57	$C_{19}H_{38}O_4$	542-44-9	299.2577 – $[C18H35O3]^+$ 257.2462 – $[C16H33O2]^+$ 239.2366 – $[C16H31O]^+$	Lipids (Glycerolipids)
Monostearin	24.14	$C_{21}H_{42}O_4$	123-94-4	327.2897 – $[C20H39O3]^+$ 267.2677 – $[C18H35O]^+$ 98.0723 – $[C6H10O]^+$	Lipids (Glycerolipids)
Hedione	14.41	$C_{13}H_{22}O_3$	24851-98-7	83.0479– $[C5H7O]^+$ 97.0623– $[C6H9O]^+$ 226.1566 – $[C13H22O3]^+$	Lipids (Lineolic acids and deriativatives-jasmonic acids)
Sclareol	18.77	$C_{20}H_{36}O_2$	515-03-7	272.2494 – $[C20H32]^+$ 121.0988 – $[C9H13]^+$ 95.0837 – $[C7H11]^+$	Lipids (Prenol lipids-diterpenes)
Limonene	5.94	$C_{10}H_{16}$	5989-27-5	136.1227 – $[C10H16]^+$ 121.0992 – $[C9H13]^+$ 79.0524 – $[C6H7]^+$	Lipids (Prenol lipids-monoterpenes)

Cont. Table 2

Cumyl alcohol	6.81	$C_{10}H_{14}O$	617-94-7	121.0644 – $[C8H9O]^+$ 103.0530 – $[C8H7]^+$ 91.0537 – $[C7H7]^+$	Lipids (Prenol lipids-monoterpenes)
Linalool	6.99	$C_{10}H_{18}O$	78-70-6	136.1218 – $[C10H16]^+$ 93.0679 – $[C7H9]^+$ 71.0844 – $[C5H4]^+$	Lipids (Prenol lipids-monoterpenes)
Camphor	7.84	$C_{10}H_{16}O$	464-48-2	152.1192 – $[C10H16]^+$ 137.0963 – $[C9H13O]^+$ 95.0852 – $[C7H11]$	Lipids (Prenol lipids-monoterpenes)
Camphol	8.16	$C_{10}H_{18}O$	507-70-0	121.1000 – $[C9H13]^+$ 95.0853 – $[C7H11]^+$ 77.0381 – $[C6H5]^+$	Lipids (Prenol lipids-monoterpenes)
Levomenthol	8.25	$C_{10}H_{20}O$	2216-51-5	138.1379 – $[C10H18]^+$ 95.0837 – $[C7H11]^+$ 81.0683 – $[C6H9]^+$	Lipids (Prenol lipids-monoterpenes)
Terpineol	8.52	$C_{10}H_{18}O$	98-55-5	136.1244 – $[C10H16]^+$ 121.1007 – $[C9H13]^+$ 93.0695 – $[C7H9]^+$	Lipids (Prenol lipids-monoterpenes)
Squalene	24.84	$C_{30}H_{50}$	111-02-4	410.3907 – $[C30H50]^+$ 121.0994 – $[C9H13]^+$ 81.0686 – $[C6H9]^+$	Lipids (Prenol lipids-triterpenes)
Cholestadiene	25.56	$C_{27}H_{44}$	747-90-0	368.3437 – $[C27H44]^+$ 247.2412 – $[C18H31]^+$ 147.1141 – $[C11H15]^+$	Lipids (Steroids and steroids derivatives)
Triethyl citrate	14.40	$C_{12}H_{20}O_7$	77-93-0	203.0913 – $[C9H15O5]^+$ 157.0496 – $[C7H9O4]^+$ 83.0486 – $[C5H7O]^+$	Organic acids and derivatives (Carboxylic acids and derivatives)
Spiro [2,4]heptane-1,5-dimethyl-6-methylene	5.95	$C_{10}H_{16}$	62238-24-8	136.1225 – $[C10H16]^+$ 121.0990 – $[C9H3]^+$ 79.0524 – $[C6H7]^+$	Other organic compounds

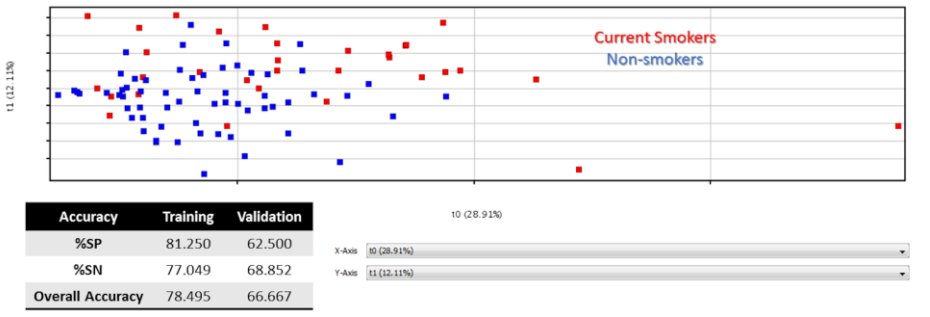
Cont. Table 2

2-Propanol 1-(2-butoxy-1-methylethoxy)-	9.21	$C_{10}H_{22}O_3$	29911-28-2	59.0485 – $[C3H7O]^+$ 86.0715 – $[C5H10O]^+$ 103.0728 – $[C5H11O2]^+$	Other organic compounds
2,4,6-Triisopropylphenol	12.69	$C_{17}H_{26}O$	08-07-34	220.1822 – $[C15H24O]^+$ 205.1584 – $[C14H21O]^+$ 77.0369 – $[C6H5]^+$	Other organic compounds
1,3-Heptadecyn-1-ol	15.11	$C_{17}H_{32}O$	56554-77-9	225.1826 – $[C14H25O2]^+$ 81.0681 – $[C6H9]^+$ 67.0529 – $[C5H7]^+$	Other organic compounds
2,4-Diphenyl-4-methyl-2(E)-pentene	16.45	$C_{18}H_{20}$	22768-22-5	236.1567 – $[C18H20]^+$ 143.0809 – $[C11H11]^+$ 91.0513 – $[C7H7]^+$	Other organic compounds
7,9-Di-t-butyl-oxaspiro(4,5)deca-6,9-diene-2,8-dione	17.18	$C_{17}H_{26}O_2$	82304-66-3	175.1104 – $[C12H15O]^+$ 133.0638 – $[C9H9O]^+$ 77.0369 – $[C6H5]^+$	Other organic compounds
10,18-Bisnorabieta-8,11,13-triene	17.99	$C_{18}H_{26}$	32624-67-2	242.2007 – $[C18H26]^+$ 227.1790 – $[C17H23]^+$ 143.0864 – $[C11H11O]^+$	Other organic compounds
2,6-Di-t-butyl-4-(2-phenylpropan-2-yl)phenol	18.71	$C_{23}H_{32}O$	34624-81-2	324.2438 – $[C23H32O]^+$ 309.2212 – $[C22H29O]^+$ 119.0836 – $[C9H11]^+$	Other organic compounds
Phenol 2,2'-methylenebis[6-(1,1-dimethyl)-4-methyl	21.77	$C_{23}H_{32}O_2$	119-47-1	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other organic compounds
Phenol 2,4-bis(1-methyl-1-phenylethyl)-	22.52	$C_{24}H_{26}O$	2772-45-4	330.1984 – $[C24H26O]^+$ 315.1748 – $[C23H23O]^+$ 237.1263 – $[C17H17O]^+$	Other organic compounds

Cont. Table 2

2,4-Bis(dimethylbenzyl)-6-*t*-butylphenol	22.57	$C_{28}H_{34}O$	244080-16-8	386.2617 – $[C28H34O]^+$ 371.2370 – $[C27H31O]^+$ 293.1897 – $[C21H25O]^+$	Other organic compounds
n-Hexadecylindane	23.93	$C_{25}H_{42}$	55334-29-7	117.0351 – $[C9H9]^+$ 130.0427 – $[C10H10]^+$ 154.1345 – $[C11H22]^+$	Other organic compounds

Statistical analysis by t-test enabled to identify twelve compounds present at significantly different concentrations in CS and NS individuals with p-value below 0.05. Among these compounds, listed in Table 3, it is worth mentioning five that were highly significant, with a p-value below 0.01: indole (p-value 0.0006), undecanol (0.0036) and three phenolic compounds such as p-cresol (0.0065), 2,4,6-tris(1-methylethyl)-phenol (0.0032) and 2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)-phenol (0.0031). Table 3 also includes the fold change ratio for these compounds by considering the two groups of individuals, CS and NS. It is worth mentioning that all the compounds were characterized by a similar concentration profile, with higher relative concentration in EBC from NS individuals as compared to CS, except for one compound, monostearin, that was more concentrated in CS. Special emphasis should be paid in terms of fold change ratio to indole, the most significant compound (p-value<0.001), which, apart from its high significance, led to a fold change of 2.59, being more concentrated in NS individuals. Monostearin, a significant compound (0.05<p-value<0.01), presented the highest fold change value, 2.63, being more concentrated in the CS group. p-Cresol, also highly-significant (0.01<p-value<0.001), reported the third value in terms of fold change ratio (2.30) with a more concentrated level in EBC from NS individuals. The rest of significant compounds in the comparison between CS and NS groups provided fold change ratios from 1.67 to 1.10.



Accuracy	Training	Validation
%SP	81.250	62.500
%SN	77.049	68.852
Overall Accuracy	78.495	66.667

Figure 1. Supervised Analysis by PLS-DA built from the data set obtained after analysis of EBC extracts comparing current smokers and non-smokers.

Table 3. Significant compounds from the statistical analysis by t-test comparing CS and NS individuals.

Compound (NS *vs* CS)	*p*-Value	Fold Change	Regulation
p- Cresol	0.0065	2.30	Up
Camphor	0.0156	1.25	Up
Indole	0.0006	2.59	Up
Undecanol	0.0036	1.51	Up
2,4,6-Tris(1-methylethyl)-phenol	0.0032	1.21	Up
Benzoic acid 4-ethoxy-ethyl ester	0.0410	1.10	Up
2,4-Diphenyl-4-methyl-2(E)-pentene	0.0416	1.30	Up
2,6-Bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)-phenol	0.0031	1.67	Up
Stearic acid	0.0152	1.52	Up
2,4-Bis(1-methyl-1-phenylethyl)-phenol	0.0127	1.26	Up
2,4-Bis(dimethylbenzyl)-6-t-butylphenol	0.0119	1.19	Up
Monostearin	0.0350	2.63	Down

Random Forest Analysis was applied to compare the discrimination accuracy of significant compounds in this study dealing with CS and NS groups. The ranking of compounds according to discrimination accuracy is shown in Figure 2 that enables to identify indole and monostearin as the two compounds with the highest discrimination capability as compared to the rest. The discrimination accuracy of these compounds was also checked by analysis of ROC curves, as shows Figure 3. As can be seen, indole reported the best predictive behavior, with a value of area under the curve (AUC) of 0.758, and specificity/sensitivity values of 0.8/0.7. This good performance can be checked in the box-and-whisker plot with the cut-off value for discrimination between CS and NS groups. Monostearin and p-cresol reported similar AUC values, 0.664 and 0.627, respectively, and close specificity/sensitivity parameters, 0.7/0.6 and 0.6/0.7, respectively. With these premises, it is worth emphasizing the statistical contribution of these three compounds to explain the differences in composition of EBC from CS and NS individuals.

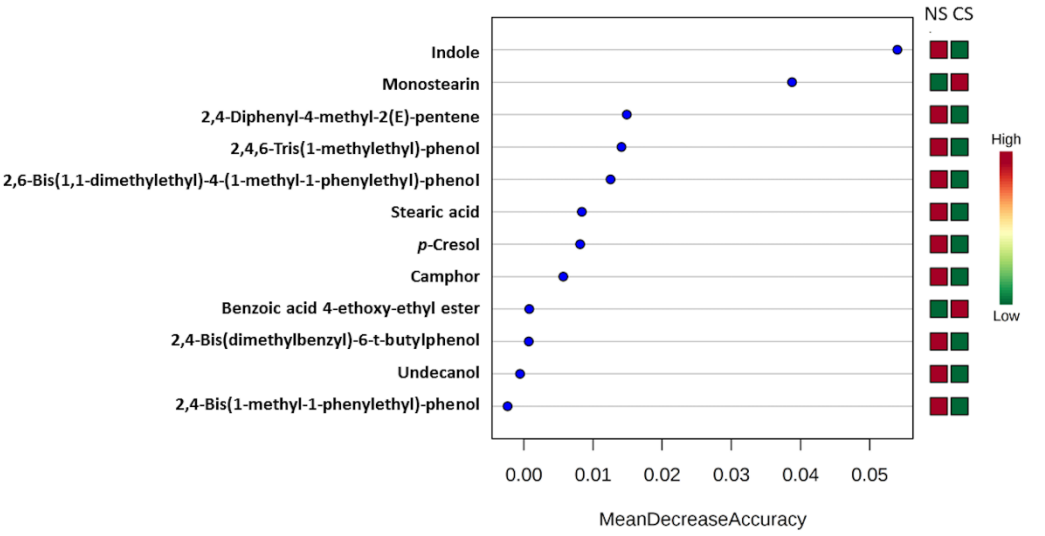


Figure 2. Random Forest Analysis built from the data set obtained after analysis of EBC from CS and NS groups.

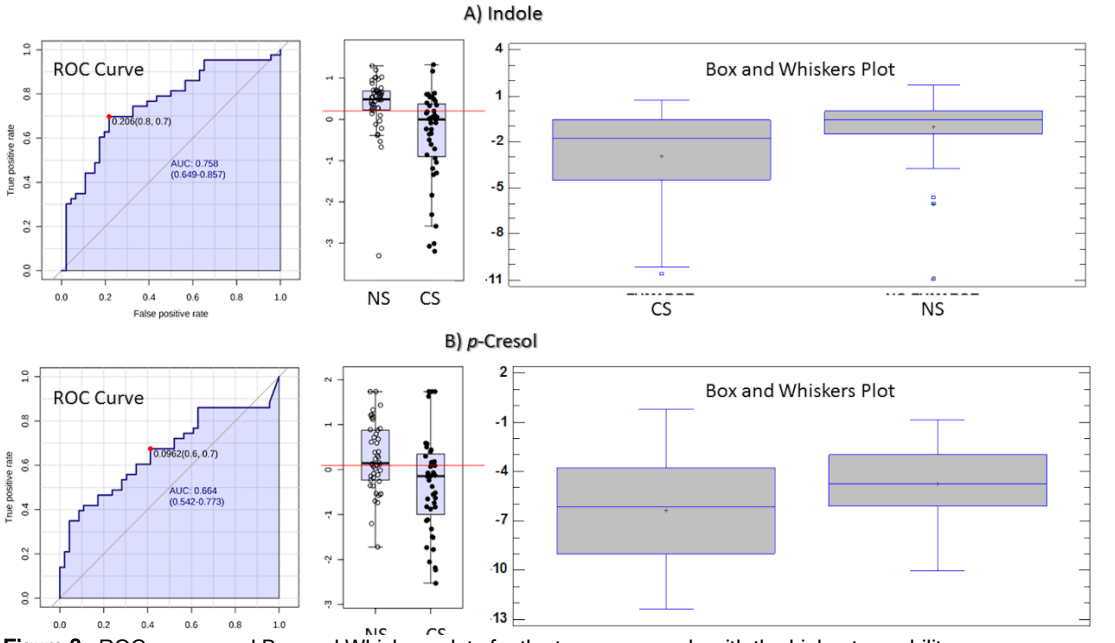


Figure 3. ROC curves and Box and Whiskers plots for the two compounds with the highest capability for discrimination in the comparison between CS and NS groups.

3.2. Comparison of the composition of EBC from current smokers and former smokers

In this step, the EBC composition of CS individuals was compared to that of FS individuals to evaluate the incidence of quitting smoking on EBC composition. Two relevant factors such as the smoking habit and the cumulative consumption, expressed as smoked pack-year, were considered as categorical variables in this study. Concerning the smoking habit, the individuals were divided into two groups –namely, FS (>1 year they quit smoking) and CS. This classification allowed obtaining balanced groups. Three equilibrated groups were also considered to study the effect of the cumulative consumption as smoked pack-year (0–24, 24–45, >45 packs-year). Supplementary Table 1 shows the number of individuals classified in each group. A two-factor ANOVA was applied to detect EBC compositional changes according to the two parameters. The analysis of the smoking habit reported eleven significant compounds with p-values below 0.05, which are listed in Table 4. Among the compounds, it is worth mentioning the presence of six highly-significant (p-value < 0.01) compounds such as camphor (p-value 0.007), undecanol (0.0100), heptadecynol (0.0054), squalene (0.0080) and two phenolic derivatives: 2,4,6-tris(1-methylethyl)-phenol (0.0056) and benzoic acid 4-ethoxy-ethyl ester (0.0048). In all cases, EBC levels for these compounds were higher in FS than in CS, except for squalene that reported the opposite situation being at higher level in EBC from CS. Figure 4 shows the ranking of predictive capability of compounds to discriminate between CS and FS groups. As can be seen, the compounds that reported the highest statistical significance (p-value<0.01) were in the top-7 compounds according to the ranking provided by the Random Forest Analysis. The discrimination accuracy of these compounds can be visualized in Figure 5 that shows the ROC curve for each compound. Thus, heptadecynol and camphor provided AUC values of 0.727 and 0.715, respectively; while specificity/sensitivity were 0.7/0.8 and 0.8/0.7, respectively. The two phenol derivatives benzoic acid 4-ethoxy-ethyl ester and 2,4,6-tris(1-methylethyl)-phenol gave a similar AUC value, while the specificity/sensitivity values were 0.8/0.7 and 0.7/0.7, respectively. Finally,

undecanol and squalene reported AUC values slightly below 0.7, while the specificity/sensitivity values were 0.8/0.6 and 0.6/0.8, respectively.

Table 4. Significant compounds resulted from the statistical analysis by t-test comparing CS and FS individuals.

Compounds (CS *vs* FS)	*p*-Value
Camphor	0.0070
Undecanol	0.0100
2,4,6-Tris(1-methylethyl)-phenol	0.0056
Benzoic acid 4-ethoxy ethyl ester	0.0048
Hedione	0.0419
13-Heptadecyn-1-ol	0.0054
Palmitoleic acid	0.0166
Stearic acid	0.0234
2,4-Bis(dimethylbenzyl)-6-t-butylphenol	0.0270
Eicosenamide	0.0372
Squalene	0.0080

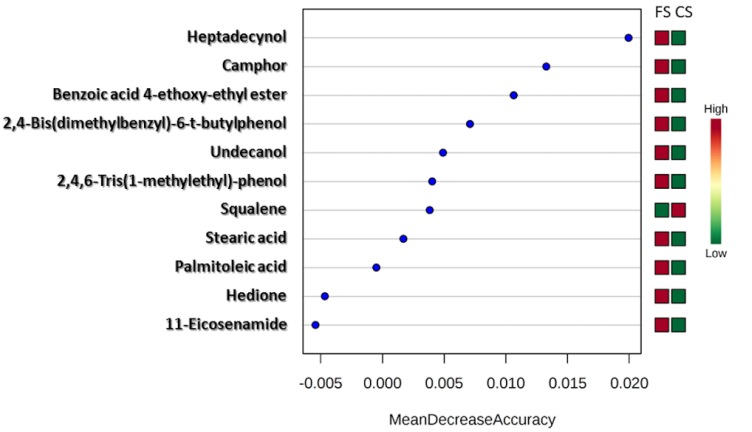


Figure 4. Ranking of predictive capability of compounds to discriminate between CS and FS individuals.

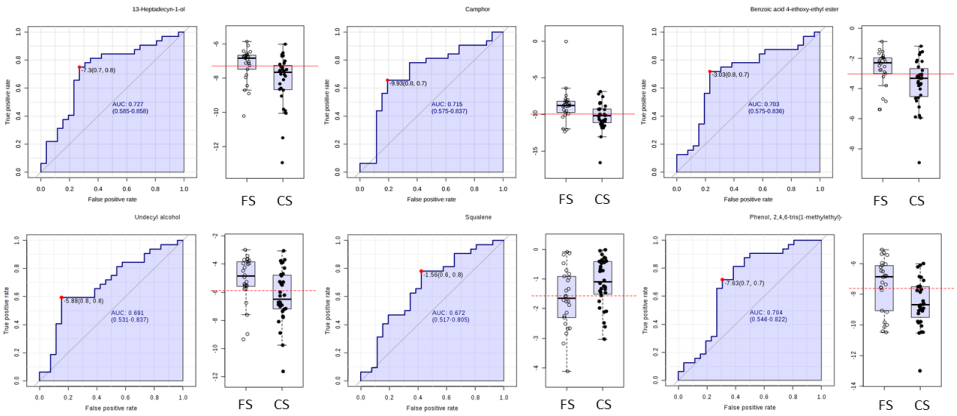


Figure 5. ROC curves for the six compounds with the highest significance and predictive capability to discriminate between CS and FS groups.

Concerning the cumulative consumption, only five compounds were significant by comparing their EBC levels in the three groups: three terpenoids – limonene (0.0124), eucalyptol (0.0071) and levomenthol (0.0209)– together with benzyl alcohol (0.0096) and spiro[2,4]heptane,1,5-dimethyl (0.0114). According to these parameters, only eucalyptol and benzyl alcohol reported statistical differences with 99.9% of confidence. The five significant compounds were characterized by the same pattern, except for levomenthol. Thus, the EBC levels of these compounds were higher in smokers with an intermediate consumption, except for levomenthol that showed similar levels in smokers with intermediate and high cumulative consumption.

4. Discussion

4.1. Comparison of the EBC composition of CS and NS individuals

In addition to be the most relevant risk factor in respiratory diseases and several types of cancer, smoking is also a key risk factor in several diseases such as cardiovascular diseases. As previously emphasized, numerous biochemical changes occurring in CS as compared to NS individuals have been

described, including those supported on metabolomics analysis. However, most of these studies have been focused on conventional biofluids such as serum/plasma, and few of them have been carried out with less conventional samples such as sputum or exhaled breath. The present research was targeted at the analysis of EBC that is obtained in a non-invasive manner and can be easily handled due to condensation in a liquid phase. The first preliminary attempt was to find discrimination patterns in EBC sampled from CS and NS in the controlled cohort of patients enrolled at the Department of Respiratory Medicine of the Reina Sofía Hospital. A PLS-DA analysis allowed finding a discrimination pattern between both groups of individuals. In fact, the PLS accuracy parameters of the resulting model were consistent for the training and validation steps. This result confirmed the existence of differences in the composition of EBC from CS and NS. The differences could be associated to three main changes: (1) alterations of endogenous metabolism –for example, increase of inflammatory markers owing to free radicals formation; (2) exogenous compounds owing to the smoking habit such as those described in the extensive research carried out by Filipiak *et al.* [24]; (3) compounds associated to microbiome-host interactions, which could also be linked to the first group.

Most of the microbiome research carried out to date has focused on bacterial communities in the colon, the site of the largest, most complex and active microbial system in the body. However, polymicrobial communities exist in many other body sites, including the skin [25], the mouth [26,27], the genitourinary [28] and respiratory tracts [29,30]. In each case, bidirectional communication between microbes and the underlying tissue regulates the local environment, including physical characteristics as well as the immune response. Microbiome dysbiosis is strongly associated with certain pathologies although the determinants of these microbial imbalances, particularly in body sites different from colon, are largely unknown.

Targeting at the respiratory tract the disruption of microbial communities has been implicated in dysregulation of local immunity [31], infection

susceptibility [30], and the development of chronic respiratory inflammatory diseases [32]. As an example, specific bacterial species are implicated in the pathogenesis of exacerbations of chronic obstructive pulmonary disease (COPD), which seems to be associated with accelerated loss of lung function in COPD. Recent studies of clinically stable COPD patients have demonstrated a greater diversity of airway microbiota as compared to healthy individuals. Also Huang et al. observed the sputum bacterial composition to be generally stable over the preexacerbation period of clinical stability, but to change at the time of exacerbation, with specific enrichment in not only typical COPD-associated bacterial species (e.g., Haemophilus influenzae) but also in other phylogenetically related species with pathogenic potential such as Gammaproteobacteria, Delta- and Betaproteobacteria [33]. Concurrently, these authors found depleted abundance of other bacteria such as Firmicutes and Actinobacteria phyla whose predicted metagenomes suggest functional capacities to produce a variety of anti-inflammatory compounds such as betalain, flavonoids, macrolides and indole alkaloids. This result is in agreement with EBC levels found in this research for indole, the compound that reported the highest significance (*p*-value 0.0006) comparing CS and NS with a fold change of 2.59, being more concentrated in NS individuals. Therefore, smoking habit leads to the modification of the respiratory microbiome with incidence on the immune response, which is here supported at metabolomics level on EBC concentrations of indole, a metabolite with recognized anti-inflammatory and antimicrobial properties. The alteration of the immune response contributes to explain the susceptibility of CS to respiratory diseases. This biological interpretation gains interest by the predictive behaviour of indole to discriminate CS from NS individuals. As can be seen, the ROC curve for this metabolite reported a high AUC value with relatively high specificity and sensitivity values: 80 and 70%, respectively.

A representative group of compounds was phenols and derivatives, which also were characterized by a high significance in the comparison of their levels in EBC from CS and NS individuals. The most relevant compound was *p*-

cresol (0.0065), involved in the degradation of toluene. Toluene is mainly metabolized to benzyl alcohol or to hydroxytoluene isomers, particularly to *p*-cresol. This compound, also endowed with antioxidant properties by virtue of the phenolic functional group, was also found at higher levels in EBC from NS individuals. This result strongly agrees with that found by Wu *et al.* [34], who found a different oral microbiome composition between current and non-current (former and non) smokers (p<0.001). CS individuals presented lower relative abundance of the phylum Proteobacteria (4.6%) as compared with NS individuals (11.7%). Taxa not belonging to Proteobacteria were also associated with smoking. Thus, the genera Capnocytophaga, Peptostreptococcus and Leptotrichia were depleted, while Atopobium and Streptococcus were enriched in CS compared with NS individuals. Functional analysis from inferred metagenomes showed that bacterial genera depleted by smoking were related to carbohydrate and energy metabolism, and to xenobiotic metabolism such as that of toluene. Particularly, toluene metabolism was more active in the genera of bacteria depleted in CS than in bacteria that are enriched in this group of individuals.

A similar pattern was found for the two alkylphenol derivatives, 2,4,6-tris(1-methylethyl)-phenol (0.0032) and 2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)-phenol (0.0031). These two compounds are structurally similar to a known antioxidant widely used in the food and cosmetic industries: butylated hydroxytoluene (BTH). Both compounds could be formed through the degradation of exogenous compounds by the microbiome, which could explain the differences in EBC levels between CS and NS. The structural analogy between the two alkylphenols and BTH allows deducing an antioxidant activity of these two compounds. There was one other highly significant compound in EBC between CS and NS: undecanol (0.0036), which has not previously been related to the microbiome. Some groups have described the potential formation of fatty alcohols through degradation of polyunsaturated fatty acid (PUFA) metabolites that are considered inflammation markers [35]. Some of these compounds are endowed with anti- or pro-inflammatory properties depending on the PUFA

precursor. In this specific case, undecanol was present at higher level in EBC from NS individuals than in CS, which potentially could explain its formation through PUFA metabolites with anti-inflammatory capability.

Special emphasis should be also put on monostearin, which reported significance ($0.05<p\text{-value}<0.01$) between CS and NS, but experienced the highest fold change value, 2.63, at higher levels in EBC from CS. There are not evidences about relationships of monostearin levels in individuals affected by respiratory diseases or smoking habit as compared to healthy individuals or NS, respectively. Wikoff et al. found significantly lower level of monostearin in malignant lung tissue from early stage adenocarcinoma patients as compared to non-malignant tissue, while monopalmitin reported the opposite situation [36]. Levels in lung tissue and EBC could be correlated due to the absorption/desorption equilibrium during breathing.

4.2. Comparison of EBC composition in current smokers and former smokers

A recent research suggested that the oral microbiome appears to return to its previous state for FS individuals. The microbial balance in individuals who quit smoking for more than 10 years was similar to that observed in NS individuals [1]. However, the results did not reveal the time required to recover the microbiome balance. In the research proposed here an additional aim was to find discrimination patterns in EBC sampled from CS and FS in the controlled cohort of patients enrolled at the Department of Respiratory Medicine of the Reina Sofía Hospital. For this purpose, the smoking load was also taking into account as a categorical factor. Concerning the smoking habit 6 compounds resulted highly significant in the comparison between CS and FS individuals: camphor, undecanol, heptadecynol, squalene and two phenol derivatives –2,4,6-tris(1-methylethyl)-phenol and benzoic acid 4-ethoxy-ethyl ester. Squalene is a naturally occurring polyprenyl compound known for its key role as an intermediate in cholesterol synthesis. Many other polyprenyl compounds

structurally similar to squalene are detected in biofluids and perform critical biological functions such as oxidation inhibition or antimicrobial properties. For example, animals utilize prenyl groups to form the side chain of ubiquinones, among which it is worth emphasizing the role of coenzyme Q10 (the most common form of ubiquinone in the human body) as endogenous antioxidant. Other well-known prenyl derivatives include carotenes, vitamin A, vitamin K, vitamin E and cyclic terpenoid compounds such as camphor, pinene and limonene [37]. Thus, the significant different concentration of camphor and squalene in CS and FS individuals (squalene is more concentrated in CS than in FS and camphor shows the opposite trend) could be attributed to the fact that quitting smoking stimulates the synthesis of compounds with antimicrobial properties like camphor using prenyl groups from squalene. In addition, as discussed above, smoking habit upsets the immune response contributing to explain the susceptibility of CS to respiratory diseases. This biological interpretation is supported on the predictive behavior of both compounds to discriminate between CS and FS. As can be seen, the ROC curves for these metabolites reported a high AUC value with relatively high specificity and sensitivity, 80 and 70% for camphor and 60 and 80% for squalene, respectively. These results could be related to that found by Biedermann et al., who investigated the role of smoking cessation on intestinal microbial composition in 10 healthy smokers undergoing controlled smoking cessation [38]. They found an increase of sequences from Firmicutes and Actinobacteria and a simultaneous decrease of the Proteobacteria and Bacteroidetes fractions after smoking cessation. This behavior was related to the production of a variety of anti-inflamatory and antimicrobial compounds.

Two other highly significant compounds in EBC from CS and FS individuals were undecanol (0.0100) and heptadecynol (0.0054) that have not been previously related to the microbiome. As discussed above, the formation of these fatty alcohols could be related to PUFA metabolites degradation. Both compounds were present at higher level in EBC from FS than in CS individuals,

which potentially could explain their source through PUFA metabolites with anti-inflammatory capability.

One other significant compound was an alkylphenol derivative, 2,4,6-tris(1-methylethyl)-phenol (0.0056) which was discussed in the previous section as a BTH related compound with antioxidant activity. This compound presented higher levels for FS than for CS individuals, supporting the results obtained comparing CS with NS individuals. Special emphasis should be also put on benzoic acid 4-ethoxy ethyl ester, which reported the highest significance between CS and FS individuals (0.0048). This compound has previously been found in exhaled breath [39], but there are not evidences about relationship with smoking habit. The depletion of certain xenobiotic biodegradation pathways in CS suggests important functional losses with potential health consequences [1]. Oral bacteria are first to come into contact with cigarette smoke and may play an important role in degrading the accompanying toxic compounds. Wu et al. observed that functional pathways related to toluene, nitrotoluene, styrene, chlorocyclohexane and chlorobenzene degradation were depleted in CS as well as the cytochrome P450 xenobiotic metabolism [1]. Conversely, polycyclic aromatic hydrocarbon and xylene degradation were enriched in CS individuals. However, Meckenstock et al. concluded in their research that toluene was the responsible for the metabolic inhibition of xylene-degrading organisms [40]. This result could be in agreement with the high level of benzoic acid 4-ethoxy ethyl ester found in EBC from FS individuals. Quitting smoking reduces the presence of toluene levels stimulating the xylene degradation and increasing the level of benzoic acid derivatives.

Concerning the smoking cumulative consumption, benzyl alcohol resulted highly significant (0.0096) as well as eucalyptol (0.0071). Benzyl alcohol is one of the two intermediates in the pathway of toluene metabolism presenting a trend opposite to p-cresol. Therefore, the benzyl alcohol levels increased with smoking load, which agrees with p-cresol EBC levels that were higher for NS than for CS individuals.

Eucalyptol is the bicyclic monoterpene 1,8-cineole, which is transformed by several pathways. Novosphingobium subterranea converts 1,8-cineole initially into 2-endo-hydroxycineole, 2,2-oxo-cineole and 2-exo-hydroxycineole. A cytochrome P450 monooxygenase catalyzes the hydroxylation of 1,8-cineole; thus, the higher levels of eucalyptol with the increase of the smoking load is explained by the cytochrome P450 xenobiotic metabolism, depleted in CS individuals and favoring the accumulation of eucalyptol as a result.

Conclusions

According to the foregoing smoking is one of the main causes of respiratory diseases. In this research, the potential of EBC as biofluid for implementation in the clinical field has been pointed out. Variation of the relative concentration of 12 metabolites (including monoacylglycerol derivatives, monoterpenes and tobacco related compounds) depending on smoking habits and the smoking load has been discussed to discriminate among current smokers, former smokers and non-smokers.

These preliminary results emphasize the necessity of a large-scale study to validate the EBC applicability to detect disorders caused by smoking habits. This first approach has also revealed statistically significant differences among the groups under study, mainly focusing the attention on camphor, indole and p-cresol, compounds which are increased in EBC from non- or former smokers; while monostearin and squalene are decreased.

Acknowledgements

The Spanish Ministerio de Economía y Competitividad (MINECO), Junta de Andalucía and FEDER programme are gratefully acknowledged for financial support (Projects “Development of methods for early cancer detection, December 29, 2011”, and “Optimización y aplicación de plataformas metabolómicas de análisis de biofluidos no invasivos para la búsqueda de biomarcadores de

diagnóstico precoz de cáncer de pulmón, FQM-1602"). A. Peralbo-Molina and M. Calderón-Santiago also thank to these organisms their research contracts. F. Priego-Capote is also grateful to the Ministerio de Ciencia e Innovación (MICINN) for a Ramón y Cajal Contract (RYC-2009-03921). Reina Sofía Hospital and its biorepository node (Córdoba, Spain) are also grateful for donors of exhaled breath condensate samples.

References

[1] Wu J et al 2016 Cigarette smoking and the oral microbiome in a large study of American adults. ISME J doi: 10.1038/ismej.2016.37

[2] Tager I B, Segal M R, Speizer F E and Weiss S T 1988 The natural history of forced expiratory volumes. Effect of cigarette smoking and respiratory symptoms. Am. Rev. Respir. Dis. 138 837–49

[3] Ryu J H, Colby T V, Hartman T E and Vassallo R 2001 Smoking-related interstitial lung diseases: a concise review. Eur. Respir. J. 17 122–32

[4] Tønnesen P et al 2007 Smoking cessation in patients with respiratory diseases: a high priority, integral component of therapy. Eur. Respir. J. 29 390–417

[5] Bartecchi C E, MacKenzie T D and Schrier R W 1994 The human costs of tobacco use (1) N. Engl. J. Med. 330 907–12

[6] Facchini F S, Hollenbeck C B, Jeppesen J, Ida Chen Y-D and Reaven G M 1992 Insulin resistance and cigarette smoking Lancet 339 1128–30

[7] Rosenberg L, Kaufman D W, Helmrich S P and Shapiro S 1985 The risk of myocardial infarction after quitting smoking in men under 55 years of age. N. Engl. J. Med. 313 1511–4

[8] Celermajer D S, Sorensen K E, Georgakopoulos D, Bull C, Thomas O, Robinson J and Deanfield J E 1993 Cigarette smoking is associated

with dose-related and potentially reversible impairment of endothelium-dependent dilation in healthy young adults Circulation 88 2149–55

[9] Lightwood J M and Glantz S A 1997 Short-term Economic and Health Benefits of Smoking Cessation : Myocardial Infarction and Stroke Circulation 96 1089–96

[10] Jerjes W et al 2012 The effect of tobacco and alcohol and their reduction/cessation on mortality in oral cancer patients: short communication. Head Neck Oncol. 4 6

[11] Taylor D H, Hasselblad V, Henley S J, Thun M J and Sloan F A 2002 Benefits of Smoking Cessation for Longevity Am. J. Public Health 92 990–6

[12] Xu T et al 2013 Effects of smoking and smoking cessation on human serum metabolite profile: results from the KORA cohort study. BMC Med. 11 60

[13] Tian X, Tang Z, Jiang J, Fang X, Wu X, Han W, Guan S, Liu H, Diao L and Sun F 2011 Effects of smoking and smoking cessation on life expectancy in an elderly population in Beijing, China, 1992-2000: an 8-year follow-up study. J. Epidemiol. 21 376–84

[14] Boué S et al 2013 Cigarette smoke induces molecular responses in respiratory tissues of ApoE(-/-) mice that are progressively deactivated upon cessation. Toxicology 314 112–24

[15] Ebbert J O 2003 Lung Cancer Risk Reduction After Smoking Cessation: Observations From a Prospective Cohort of Women J. Clin. Oncol. 21 921–6

[16] Halpern M T, Gillespie B W and Warner K E 1993 Patterns of Absolute Risk of Lung Cancer Mortality in Former Smokers JNCI J. Natl. Cancer Inst. 85 457–64

[17] De Laurentiis G, Paris D, Melck D, Montuschi P, Maniscalco M, Bianco A, Sofia M and Motta A 2013 Separating smoking-related diseases using

NMR-based metabolomics of exhaled breath condensate. J. Proteome Res. 12 1502–11

[18] Hsu P-C, Zhou B, Zhao Y, Ressom H W, Cheema A K, Pickworth W and Shields P G 2013 Feasibility of identifying the tobacco-related global metabolome in blood by UPLC-QTOF-MS. J. Proteome Res. 12 679–91

[19] Peralbo-Molina A, Calderón-Santiago M, Priego-Capote F, Jurado-Gámez B and Luque de Castro M D 2015 Development of a method for metabolomic analysis of human exhaled breath condensate by gas chromatography–mass spectrometry in high resolution mode Anal. Chim. Acta 887 118–26

[20] Kubáň P and Foret F 2013 Exhaled breath condensate: determination of non-volatile compounds and their potential for clinical diagnosis and monitoring. A review Anal. Chim. Acta 805 1–18

[21] Hoffmeyer F, Raulf-Heimsoth M, Harth V, Bünger J and Brüning T 2009 Comparative analysis of selected exhaled breath biomarkers obtained with two different temperature-controlled devices BMC Pulm. Med. 9 48

[22] Warrack B M, Hnatyshyn S, Ott K-H, Reily M D, Sanders M, Zhang H and Drexler D M 2009 Normalization strategies for metabonomic analysis of urine samples. J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 877 547–52

[23] Xia J, Sinelnikov I V., Han B and Wishart D S 2015 MetaboAnalyst 3.0--making metabolomics more meaningful Nucleic Acids Res. 43 W251–7

[24] Filipiak W et al 2012 Dependence of exhaled breath composition on exogenous factors, smoking habits and exposure to air pollutants. J. Breath Res. 6 036008

[25] Hannigan G D, Meisel J S, Tyldsley A S, Zheng Q, Hodkinson B P, SanMiguel A J, Minot S, Bushman F D and Grice E A 2015 The human skin double-stranded DNA virome: topographical and temporal diversity, genetic

enrichment, and dynamic associations with the host microbiome. MBio 6 e01578-15

[26] Dewhirst F E, Chen T, Izard J, Paster B J, Tanner A C R, Yu W-H, Lakshmanan A and Wade W G 2010 The human oral microbiome. J. Bacteriol. 192 5002–17

[27] Wade W G 2013 The oral microbiome in health and disease. Pharmacol. Res. 69 137–43

[28] Brotman R M, Shardell M D, Gajer P, Fadrosh D, Chang K, Silver M I, Viscidi R P, Burke A E, Ravel J and Gravitt P E 2014 Association between the vaginal microbiota, menopause status, and signs of vulvovaginal atrophy. Menopause 21 450–8

[29] Neish A S 2014 Mucosal immunity and the microbiome. Ann. Am. Thorac. Soc. 11 Suppl 1 S28–32

[30] Vissers M, de Groot R and Ferwerda G 2014 Severe viral respiratory infections: are bugs bugging? Mucosal Immunol. 7 227–38

[31] Hansel T T, Johnston S L and Openshaw P J 2013 Microbes and mucosal immune responses in asthma. Lancet (London, England) 381 861–73

[32] Marsland B J and Gollwitzer E S 2014 Host-microorganism interactions in lung diseases. Nat. Rev. Immunol. 14 827–35

[33] Huang Y J, Sethi S, Murphy T, Nariya S, Boushey H A and Lynch S V 2014 Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. J. Clin. Microbiol. 52 2813–23

[34] Wu A H, Paganini-Hill A, Ross R K and Henderson B E 1987 Alcohol, physical activity and other risk factors for colorectal cancer: a prospective study. Br. J. Cancer 55 687–94

[35] Patterson E, Wall R, Fitzgerald G F, Ross R P and Stanton C 2012 Health implications of high dietary omega-6 polyunsaturated Fatty acids. J. Nutr. Metab. 2012 539426

[36] Wikoff W R et al 2015 Metabolomic markers of altered nucleotide metabolism in early stage adenocarcinoma. Cancer Prev. Res. (Phila). 8 410–8

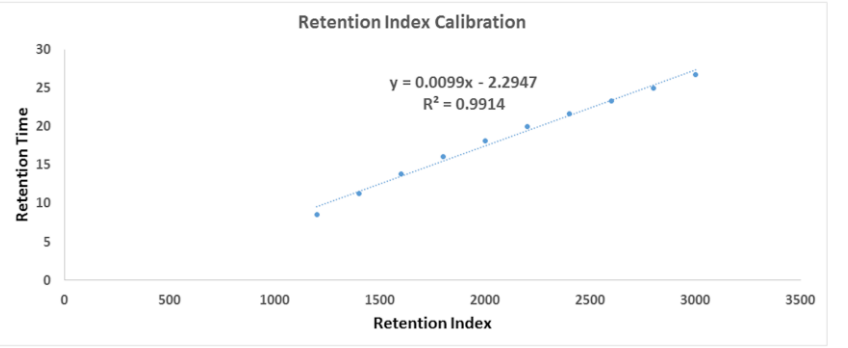
[37] Kelly G S 1999 Squalene and its potential clinical uses. Altern. Med. Rev. 4 29–36

[38] Biedermann L et al 2013 Smoking cessation induces profound changes in the composition of the intestinal microbiota in humans. PLoS One 8 e59260

[39] Phillips M 1999 Variation in volatile organic compounds in the breath of normal humans J. Chromatogr. B 729 75–88

[40] Meckenstock R U, Warthmann R J and Schäfer W 2004 Inhibition of anaerobic microbial o-xylene degradation by toluene in sulfidogenic sediment columns and pure cultures. FEMS Microbiol. Ecol. 47 381–6

Supplementary material



Supplementary Figure 1. RI calibration curve from analysis of the alkane standard mixture.

Supplementary Table 1. Data from each individual such as age, sex and cigarettes consumption expressed as smoked pack-year.

CODE	Gender	Age	Weight	Heigth	BMI	Smoking Habits	Smoking Load
OV14-01-052	Female	51	52	1.65	19.1	CS	25-44
OV14-04-002	Male	57	61.5	1.73	20.55	NS	0
OV13-04-071	Female	54	57.7	1.66	20.94	NS	0
OV14-01-106	Female	51	57	1.62	21.72	NS	0
OV13-01-020	Male	67	63	1.7	21.8	CS	>45
OV13-04-202	Male	48	74.5	1.81	22.74	NS	0
OV14-01-082	Male	51	60	1.62	22.86	CS	>45
OV14-01-049	Male	59	69	1.72	23.32	FS	>45
OV13-04-114	Male	44	71.5	1.74	23.62	CS	25-44
OV13-01-100	Male	57	76	1.79	23.72	NS	0

OV13-04-138	Male	52	77	1.8	23.77	CS	25-44
OV13-01-119	Male	70	73	1.75	23.84	FS	25-44
OV14-01-071	Male	60	65	1.65	23.88	NS	0
OV14-01-117	Male	61	77	1.79	24.03	NS	0
OV14-01-070	Male	51	78	1.8	24.07	NS	0
OV14-01-075	Female	66	55	1.5	24.44	CS	25-44
OV14-01-105	Male	52	75	1.75	24.49	CS	0-24
OV14-01-079	Male	54	82	1.83	24.49	FS	25-44
OV13-04-215	Male	58	67.5	1.66	24.5	NS	0
OV14-01-094	Male	56	70	1.69	24.51	NS	0
OV13-04-166	Male	50	71	1.7	24.57	NS	0
OV13-04-060	Male	51	77	1.77	24.58	FS	>45
OV14-01-058	Female	57	67	1.65	24.61	NS	0
OV13-04-214	Male	51	72	1.71	24.62	NS	0
OV13-04-120	Male	56	76.5	1.76	24.7	CS	0-24
OV13-04-088	Male	63	66.5	1.64	24.72	FS	0-24
OV14-01-026	Male	69	70	1.68	24.8	CS	25-44
OV13-01-130	Male	56	61	1.56	25.07	CS	0-24
OV13-04-113	Male	43	73.5	1.7	25.43	NS	0
OV13-01-074	Male	60	71	1.67	25.46	CS	>45
OV14-01-033	Female	71	66	1.6	25.78	NS	0
OV13-04-008	Female	60	68	1.62	25.91	FS	0-24

OV13-04-076	Male	55	78.4	1.73	26.2	NS	0
OV14-01-044	Male	54	85	1.8	26.23	FS	25-44
OV13-01-099	Male	55	76	1.7	26.3	NS	0
OV13-01-110	Male	66	76	1.7	26.3	CS	0-24
OV13-04-058	Female	55	70	1.63	26.35	NS	0
OV13-04-077	Male	65	81.7	1.76	26.38	FS	0-24
OV13-04-004	Female	65	68	1.6	26.56	NS	0
OV13-04-075	Female	70	69	1.61	26.62	NS	0
OV13-01-116	Male	64	82	1.75	26.78	CS	>45
OV13-04-132	Male	46	79.9	1.72	27.01	CS	0-24
OV13-04-081	Male	53	80	1.72	27.04	NS	0
OV13-04-103	Male	67	65	1.55	27.06	NS	0
OV13-01-095	Male	56	85	1.77	27.13	CS	25-44
OV14-01-076	Male	61	84	1.75	27.43	NS	0
OV14-01-091	Male	60	73	1.63	27.48	NS	0
OV14-01-034	Male	72	75	1.65	27.55	NS	0
OV14-01-029	Female	63	62	1.5	27.56	FS	0-24
OV14-01-077	Male	68	80	1.7	27.68	FS	25-44
OV13-04-123	Male	45	79.5	1.69	27.84	NS	0
OV13-04-186	Male	74	70	1.58	28.04	NS	0
OV14-01-100	Male	62	83	1.72	28.06	CS	0-24
OV13-01-014	Female	72	63	1.49	28.38	NS	0

OV13-04-154	Male	56	87.5	1.75	28.57	NS	0
OV14-01-087	Male	58	78	1.65	28.65	NS	0
OV14-01-080	Male	54	78	1.65	28.65	CS	0-24
OV14-01-098	Male	51	88	1.75	28.73	FS	0-24
OV14-01-001	Male	75	88	1.75	28.73	FS	>45
OV13-04-003	Male	56	69.3	1.55	28.84	NS	0
OV13-01-030	Female	53	75	1.61	28.93	CS	25-44
OV13-04-054	Female	47	77	1.63	28.98	NS	0
OV13-04-227	Male	71	88	1.74	29.07	NS	0
OV13-04-087	Male	66	85	1.71	29.07	CS	25-44
OV13-01-036	Male	57	105	1.9	29.09	CS	>45
OV13-04-160	Male	56	89	1.74	29.4	NS	0
OV13-01-129	Male	55	75	1.59	29.67	NS	0
OV14-01-084	Male	59	94	1.78	29.67	NS	0
OV13-01-091	Female	63	78	1.62	29.72	CS	25-44
OV13-04-129	Male	52	83	1.66	30.12	FS	25-44
OV14-04-017	Male	57	83.6	1.66	30.34	NS	0
OV14-01-028	Male	59	91	1.72	30.76	NS	0
OV14-01-115	Male	50	102	1.82	30.79	NS	0
OV13-04-106	Male	53	94.5	1.75	30.86	NS	0
OV14-01-086	Male	59	93	1.73	31.07	NS	0
OV13-04-108	Male	44	87	1.67	31.2	FS	0-24

OV13-04-007	Female	59	81	1.61	31.25	NS	0
OV13-04-082	Male	58	101.5	1.8	31.33	CS	25-44
OV13-04-016	Female	52	49	1.25	31.36	CS	0-24
OV14-01-103	Male	62	93	1.72	31.44	CS	>45
OV13-04-194	Male	53	100	1.78	31.56	NS	0
OV13-04-099	Male	53	102.3	1.8	31.57	NS	0
OV13-04-216	Male	47	87.5	1.66	31.75	NS	0
OV13-04-070	Male	61	86.5	1.65	31.77	NS	0
OV14-01-090	Male	51	94	1.72	31.77	CS	25-44
OV13-01-024	Male	59	96	1.73	32.08	NS	0
OV13-04-199	Male	77	89.5	1.67	32.09	NS	0
OV14-01-064	Female	58	80	1.57	32.46	CS	25-44
OV13-01-154	Male	64	114	1.87	32.6	FS	25-44
OV14-01-036	Male	56	100	1.75	32.65	FS	25-44
OV13-04-122	Male	74	88	1.64	32.72	NS	0
OV13-01-113	Male	75	87	1.63	32.74	CS	>45
OV13-01-019	Female	59	85	1.61	32.79	NS	0
OV13-04-161	Male	65	80	1.56	32.87	NS	0
OV14-01-030	Male	68	74	1.5	32.89	NS	0
OV14-01-041	Male	60	94	1.69	32.91	NS	0
OV14-01-099	Male	57	100	1.74	33.03	NS	0
OV14-01-122	Male	65	82	1.57	33.27	CS	>45

OV14-01-032	Female	68	78	1.53	33.32	NS	0
OV13-04-084	Male	62	93	1.67	33.35	FS	0-24
OV13-04-126	Male	72	92	1.66	33.39	FS	0-24
OV13-04-064	Male	61	99	1.72	33.46	FS	>45
OV13-04-101	Male	54	97	1.7	33.56	FS	>45
OV14-01-119	Male	73	95	1.68	33.66	FS	>45
OV14-01-078	Male	52	100	1.72	33.8	FS	>45
OV14-01-024	Male	69	100	1.72	33.8	FS	>45
OV14-04-001	Male	72	104	1.75	33.96	NS	0
OV13-01-120	Female	49	94	1.66	34.11	CS	0-24
OV13-04-095	Male	59	112.5	1.81	34.34	FS	0-24
OV14-04-003	Male	63	100	1.7	34.6	NS	0
OV13-04-131	Male	49	100	1.7	34.6	FS	0-24
OV13-04-100	Female	63	88.5	1.57	35.9	NS	0
OV13-04-093	Male	58	102.5	1.68	36.32	CS	0-24
OV13-04-065	Female	67	91	1.58	36.45	NS	0
OV13-04-090	Male	67	105.5	1.68	37.38	FS	0-24
OV14-01-081	Female	56	95	1.57	38.54	CS	25-44
OV13-04-042	Male	54	113	1.68	40.04	NS	0
OV13-04-173	Male	52	106	1.61	40.89	NS	0
OV14-01-121	Male	59	128	1.7	44.29	CS	>45

DISCUSIÓN DE LOS RESULTADOS

La normativa actual de la Universidad de Córdoba referente a la presentación de la Memoria de la Tesis Doctoral, en la modalidad en la que se incluyen los artículos (publicados o próximos a su publicación) como tales, requiere una discusión conjunta de los resultados de la investigación realizada, cuya estructura depende de su homogeneidad.

La mayor parte de la presente Memoria de Tesis está constituida por la investigación desarrollada, que tiene como denominador común la metabolómica, aplicada por un lado a la explotación de residuos agroalimentarios y por otro al diagnóstico precoz de enfermedades.

La Memoria se ha dividido en dos partes en función tanto del objetivo como del tipo de muestra utilizado, de modo que la Parte A tiene como denominador común la caracterización de residuos de la industria agroalimentaria mediterránea (vid/vino y olivo/aceite), mientras que la Parte B, eminentemente clínica, se dedica a la búsqueda de biomarcadores de cáncer de pulmón en aliento condensado.

Con el anterior criterio, en la Parte A se recoge el desarrollo de diferentes protocolos de extracción de compuestos de interés y su identificación ¬-y cuantificación cuando se ha considerado conveniente-, orientados al aprovechamiento de estos residuos. El objetivo final de la investigación fue obtener información de su composición para poder decidir qué componentes son los que pueden y deben explotarse. El escaso o nulo valor actual de los residuos objeto del estudio hace más interesante la obtención de productos de alto valor añadido.

La Parte B recoge la investigación sobre el perfil metabólico (el análisis global más utilizado en metabolómica clínica cuyo objetivo es abarcar la detección de la mayor parte posible de los componentes de la muestra y facilitar su identificación) del aliento y el estudio de grupos específicos o familias de compuestos de interés clínico por su implicación en rutas biológicas cruciales que pueden alterarse o desregularse como consecuencia de ciertos estados patológicos.

Los resultados obtenidos en cada una de las dos secciones se discuten a continuación, no sin antes comentar algunos aspectos instrumentales sobre espectrometría de masas de alta resolución como técnica para la detección común a las dos secciones. Las propiedades analíticas de esta técnica, en términos de sensibilidad, selectividad, exactitud, precisión y resolución, la convierten en la herramienta de mayor potencial para el análisis metabolómico, tanto en muestras vegetales como clínicas. Entre los diferentes tipos de espectrómetros de masas el más utilizado para análisis global es el QTOF gracias a su alta resolución (exactitud en la relación m/z), velocidad de barrido y sensibilidad a un coste competitivo. Por estas razones constituye el detector más adecuado para análisis cualitativo y para la cuantificación relativa, principalmente en metabolómica clínica debido a la complejidad de las biomuestras. Los espectrómetros de masas son especialmente adecuados cuando se acoplan a sistemas de separación cromatográficos. La naturaleza de la técnica de separación a utilizar (principalmente cromatografía de gases o de líquidos) depende de las propiedades químicas de los metabolitos en estudio. Cuando la complejidad de la muestra es elevada, la estrategia óptima pasa por la integración del análisis mediante ambas técnicas cromatográficas acopladas a espectrometría de masas con el fin de incrementar la capacidad de detección de metabolitos y, por tanto, el nivel de información generada a partir del análisis de las muestras.

Parte A. Aportaciones al aprovechamiento de residuos: Extracción de compuestos de alto valor añadido e identificación de componentes de desechos.

La Parte I recoge el desarrollo de diferentes estrategias de extracción y caracterización orientadas a la evaluación de una posible explotación de los residuos procedentes de la agricultura y de la industria agroalimentaria mediterráneas (vid/vino y olivo/aceite) como fuente de compuestos fenólicos, principalmente. Esta parte engloba 5 capítulos en los que la investigación que se recoge tuvo un doble objetivo: (1) conocer la situación actual descrita en la bibliografía de la explotación, real y potencial, de estos residuos y, (2) aportar nuevas formas de obtención e identificación de los compuestos existentes en

dichos residuos con el fin último de un mejor aprovechamiento de estos materiales de escaso o nulo valor actual.

Capítulos I y II

Estos capítulos recogen los estudios bibliográficos realizados por la doctoranda y que le proporcionaron la base para los estudios experimentales que se recogen en los tres siguientes capítulos de esta Parte A de la Memoria.

El Capítulo I evalúa en forma de revisión bibliográfica crítica el estado actual del aprovechamiento de los residuos agroalimentarios más abundantes en la cuenca mediterránea derivados de los dos cultivos tradicionales: olivos y viñedos. Se estudian los sarmientos y las hojas de vid (frescas y senescentes), las lías, los hollejos y las pepitas, por un lado, y las hojas de olivo y el alperujo, por otro. El potencial de estos residuos como fuente de fenoles se describe en este Capítulo, poniendo de manifiesto las numerosas propiedades beneficiosas de estos compuestos para la salud humana, especialmente por sus propiedades antioxidantes.

Esta revisión permitió conocer las metodologías empleadas en las investigaciones publicadas en los últimos años teniendo en cuenta factores como el tiempo necesario para que se complete la extracción, el volumen de extractante, la responsabilidad respecto al uso de extractantes no tóxicos, la eficiencia de la extracción y su coste, atendiendo tanto a la materia prima como a los compuestos que se pretenden extraer.

El Capítulo II de la Memoria es a su vez uno de los que conforman un libro multiautor en el que se exploran las propiedades de la uva en general y de los compuestos fenólicos que contiene en particular. Se discuten en profundidad tanto la producción mundial de uva, como sus múltiples usos. Además, se realiza una descripción completa de cada una de las partes que componen la uva, así como del contenido (absoluto y relativo) de los compuestos fenólicos presentes en cada una de ellas, clasificados a su vez por clases y subclases con el fin de facilitar la explicación de sus propiedades saludables. Se utiliza por tanto el contenido

fenólico como información analítica para proporcionar una visión crítica de: (a) los diferentes tipos de residuos disponibles como fuente de fenoles; (b) las técnicas de extracción que se emplean para obtener fenoles a partir de las uvas, comparando además su eficiencia, rapidez e implicación del usuario; (c) los métodos disponibles para obtener información analítica sobre una clase concreta de fenoles; y (d) las tecnologías que se utilizan actualmente para identificar y cuantificar fenoles.

La redacción de estos dos capítulos proporcionó a la doctoranda una puesta al día sobre la materia en cuestión, con la que adquirió una base sólida en la que soportar la investigación a desarrollar en los Capítulos III-V, así como la formación para la interpretación crítica de la investigación publicada por otros autores.

Capítulos III y IV

La comparación de distintas técnicas de extracción asistidas por diferentes formas de energía para la obtención de extractos polares ricos en compuestos fenólicos –importantes en enología– a partir de un residuo de residuo (el que resulta de la obtención de etanol de los desechos de la industria vinícola) es la investigación que se recoge en estos dos capítulos. Se utilizaron separadamente los hollejos y las pepitas de uvas blancas y tintas para estudiar las posibilidades de estas materias primas, muy diferentes.

Con el objetivo de establecer las posibilidades de aprovechamiento de estos residuos, se utilizaron y compararon tres métodos para mejorar/acelerar la extracción mediante la asistencia de ultrasonidos, microondas y presión+temperatura altas. Como método de extracción de referencia se empleó el tradicional basado en agitación a temperatura ambiente durante 24 horas. Las condiciones de extracción más adecuadas para cada tipo de residuo se determinaron en base al número de entidades moleculares detectadas en cada extracto como parámetro de respuesta. El método de extracción convencional junto con el asistido por presión y temperatura altas fueron los que presentaron

mejores resultados, aunque el primero requiere 24 horas de extracción frente a los 60 minutos necesarios por el segundo. La eficiencia de la extracción de este último método mejoró significativamente (aumentando el número de entidades moleculares detectadas frente al método convencional) al disminuir a 3 el pH del extractante pasando de 87 a 119 y de 105 a 128 para las pepitas de uva blanca y tinta respectivamente; de 267 a 401 y de 318 a 368 para los hollejos de uva blanca y tinta respectivamente.

El siguiente paso para optimizar la extracción de compuestos de interés consistió en evaluar la temperatura (entre 140 y 220 ºC) y el tiempo de proceso (entre 20 y 80 min) sobre la eficiencia de extracción. La temperatura de extracción óptima para los hollejos de uva blanca y ambas variedades de pepitas fue 180 ºC, mientras que 160 ºC fue la temperatura óptima para extraer los hollejos de uva tinta. Se demostró que temperaturas superiores a las óptimas proporcionaban extractos con composiciones cada vez más diferentes debidas a la degradación de los compuestos extraídos. En cuanto al tiempo de extracción, 60 minutos fueron necesarios en todos los casos para extraer el mayor número posible de compuestos y, por tanto, un extracto más rico en compuestos de interés.

Además, el contenido fenólico total en los extractos obtenidos, que proporcionó el método de Folin-Ciocalteu, se utilizó como parámetro para comparar las dos variedades de uva de las que procedían los residuos. La comparación de la capacidad antioxidante de los extractos mediante el método FRAP dio resultados consistentes con los obtenidos mediante el método de Folin-Ciocalteu, ya que los residuos procedentes de la uva blanca, que fue la de mayor contenido fenólico, resultó ser también la variedad con mayor capacidad antioxidante; lo que confirma la solidez de los resultados obtenidos.

Tras la optimización completa del proceso de extracción, se llevó a cabo una caracterización completa de los extractos. La obtención de los datos necesarios para la caracterización completa mediante LC-TOF/MS requirió 65 min.

Los compuestos fenólicos identificados pueden clasificarse de forma básica en flavonoides y no flavonoides. Entre los primeros destacan:

(i) La antocianina vitisina B conocida como acetaldehído malvidina-3-glucósido. El resto de antocianinas que normalmente están presentes en las uvas tintas no se detectaron en los extractos, lo que indica que esta familia de compuestos se degradó en el proceso de destilación industrial.

(ii) Los flavonoides catequina, epicatequina y galato, sus derivados y un isómero de la procianidina B también se detectaron.

(iii) Los flavonoles, que estaban presentes en mayor número en los extractos de los residuos de la uva roja, se detectaron como derivados libres y 3-glucósidos.

En cuanto a los fenoles no flavonoides, se detectó una gran variedad de ácidos fenólicos como ácido caféico, ácido gálico y ácido protocatéquico. Algunos estilbenos y otros compuestos fenólicos como ésteres de etilo (ferulato de etilo y protocatecuato etílico), monómeros de lignina y derivados (pirocatecol, pirogalol y siringol), aldehídos (coniferaldehÍdo) y cetonas (vainillina), entre otros, también se identificaron.

En cuanto a los derivados furánicos, la eliminación de los azúcares que se lleva a cabo previamente en la industria justifica el escaso número de derivados del furano identificados (entre ellos el alcohol furfurílico).

Los resultados, recogidos en los Capítulos III y IV, constituyen un sólido soporte para la caracterización de estos “residuos de residuos” y muestran que ambas materias primas son fuentes de compuestos de interés para las industrias alimentaria y nutracéutica, a pesar de la degradación causada por las condiciones drásticas del proceso de destilación.

Capítulo V

La investigación recogida en el Capítulo V, realizada en el contexto del análisis metabolómico global, se centró en un desecho indeseable de la industria

aceitera: el alperujo. El objetivo de este estudio fue la caracterización de los metabolitos polares para su aprovechamiento integral. La primera etapa fue la extracción de esta amplia familia de metabolitos utilizando presión+temperatura altas. La medida de los fenoles totales en el extracto se realizó mediante el método de Folin-Ciocalteu, mientras que su perfil fenólico se obtuvo por LC-QTOF MS/MS con el fin de aumentar el nivel de selectividad en la identificación de compuestos.

El análisis cualitativo de los compuestos en los extractos se centró de forma especial en aquellos compuestos fenólicos que se encuentran en el aceite de oliva virgen. Se identificó un total de 49 compuestos atendiendo a las medidas de masa exacta del ión precursor y de los iones productos más representativos, así como su distribución isotópica. Cabe destacar compuestos como hidroxitirosol/tirosol y sus derivados, precursores de iridoides, secoiridoides y sus derivados, flavonoides, lignanos y ácidos fenólicos.

El hidroxitirosol y tirosol son dos de los fenoles simples más abundantes en el VOO (aceite de oliva virgen, del inglés virgin olive oil). Ambos compuestos suelen encontrarse en su forma libre pero también formando diferentes derivados. Estos antioxidantes se encuentran además en mayor concentración en el alperujo que en el VOO. En el alperujo empleado en este estudio se identificaron además de estos dos compuestos, algunos de sus derivados, como el glucósido en el caso de los dos compuestos y el diglucósido y el ramnósido del hidroxitirosol

Los precursores de iridoides son un amplio grupo de monoterpenos implicados en la síntesis de los secoiridoides, una familia de compuestos importante para el VOO. Se identificaron el ácido logánico y la loganina junto con sus derivados glucósidos, además de otros compuestos implicados en la síntesis de secoiridoides como el ácido 7-deoxilogánico, el ácido secologánico, secologanosido y la secologanina, que conduce a la síntesis del éster metílico del oleasido, precursor inmediato de los secoiridoides. Este compuesto es el derivado glucósido del ácido eleanólico, compuestos identificados también en este trabajo. Además de estos compuestos, el oleasido y sus derivados mono y diglucósido también fueron identificados. En cuanto a los secoiridoides, la oleuropeína, la demetiloleuropeína, el ligtrósido y el verbascósido fueron también identificados

además de varios derivados de la oleuropeina y del ligtrósido, destacando entre ellos la oleuropeína aglicona.

Finalmente se identificaron numerosos fenoles simples como los ácidos cumárico, ferúlico, vainíllico, shikímico, cinámico, cafeico, protocatéico y gálico, además de la fenilalanina, implicada en la ruta de los fenilpropanoides, dando lugar a la síntesis de algunos de los ácidos anteriores. En esta ruta también están involucrados los lignanos, aunque sólo tres metabolitos de lignanos (viz. Pinoresinol, hydroxypinoresinol, y acetoxypinoresinol) se detectaron en el extracto de alperujo.

Los flavonoides, implicados también en la ruta de los fenilpropanoides, son una familia importante por sus propiedades beneficiosas para la prevención de una amplia variedad de enfermedades. En este trabajo, se han identificado numerosos flavonoides como la rutina, apigenina, luteolina, apigenina glucósido, luteolina glucósido, taxifolina, diosmetina, y la quercetina.

Se abrió así una vía al aprovechamiento de estos compuestos de los que se pudo conocer su concentración relativa, consiguiendo un extracto con valor añadido y un residuo libre de antioxidantes que presenta unas características óptimas para la producción de gas, compost, etc., es decir, para cualquier tipo de aprovechamiento en el que la presencia de antioxidantes suponga un problema o una ralentización en la obtención del producto.

Parte B. Aportaciones sobre el uso de aliento condensado como muestra clínica para la búsqueda de marcadores de cáncer de pulmón

Esta sección se centra en la aplicación del análisis metabolómico global a estudios clínicos, poniendo una vez más de manifiesto la efectividad de la estrategia en esta área. El análisis global en metabolómica puede dividirse en dos categorías genéricas en función del objetivo de la investigación: La primera se refiere a la búsqueda de biomarcadores y la segunda a la interpretación para explicar el efecto de un cambio en una determinada ruta metabólica o en un sistema biológico. Las dos categorías han sido objeto de la investigación que se

recoge en esta sección, abarcando una de las áreas de aplicación orientada a la medicina personalizada: diagnóstico, pronóstico y tratamiento de enfermedades. Se ha considerado para ello una enfermedad clave, el cáncer de pulmón, uno de los tipos de cáncer que causa más muertes al ser el segundo más comúnmente diagnosticado y en una etapa más tardía de su desarrollo.

Para el análisis metabolómico global se requieren detectores de masas de alta resolución con el fin de obtener una buena exactitud en la medida de masas con la que se consiga la identificación del máximo número posible de metabolitos detectados en el biofluido en estudio. Un equipo TOF es de gran utilidad en estos casos porque satisface las necesidades del análisis global, resolución espectral y buen nivel de sensibilidad que proporcionan suficiente información para la identificación de los metabolitos. En este bloque se utilizó como técnica analítica de determinación un GC-TOF/MS que representa una herramienta ideal para determinar la composición del EBC debido a la combinación de la alta resolución que se consigue con el cromatógrafo de gases con la alta resolución espectral asociada al detector TOF. La GC-MS destaca por su gran reproducibilidad y robustez lo que soporta la existencia de bases de datos con información espectral de una amplia variedad de compuestos, como por ejemplo la NIST, que además de presentar el patrón de fragmentación, proporciona información del índice de retención, necesario para confirmar la identificación.

La selección de la GC-MS también viene avalada por el tipo de muestra seleccionada para este estudio, el aliento condensado. Es conveniente destacar la escasez de aplicaciones desarrolladas que impliquen el análisis de aliento condensado a pesar de ser una muestra que se puede obtener de forma no invasiva, relativamente fácil y con un volumen suficiente para abordar el análisis. Además, el aliento condensado es un biofluido que presenta una matriz no compleja en comparación con biofluidos convencionales como el caso de sangre u orina.

La principal limitación en el análisis de este biofluido es la falta de protocolos estandarizados que soporten la utilización de este biofluido en clínica. Por ello, el primer estudio de este bloque fue el desarrollo de un método para el análisis de aliento condensado mediante GC-TOF/MS en modo de alta resolución.

Para la preparación de la muestra se compararon la LLE y la SPE. En ambos casos se probaron 4 disolventes distintos: hexano, ciclohexano, diclorometano y acetato de etilo. Comparando el número de metabolitos que proporcionó cada uno de los extractos, la LLE con hexano resultó ser la estrategia que permitió obtener un perfil más completo de los metabolitos existentes en aliento condensado, detectándose 51 metabolitos. Además, todos los compuestos detectados en el eluato de SPE con hexano (también el mejor eluyente para la aplicación de la técnica en este caso), estaban entre los existentes en el extracto obtenido por LLE. Con objeto de conocer la eficiencia de los métodos basados en cada una de las técnicas, se calculó este parámetro para ambos, resultando el método basado en LLE el más adecuado, ya que proporcionó mejor respuesta para la mayoría de los compuestos tentativamente identificados, por lo que finalmente se seleccionó esta estrategia como óptima.

Una vez optimizada la etapa de preparación de muestra, se compararon también diferentes estrategias de normalización de los datos obtenidos para determinar cuál de ellas proporcionaba un perfil normalizado con la menor variabilidad entre réplicas de la misma muestra. La normalización basada en la suma total de señal útil (MSTUS, del inglés “total useful mass spectrometry signal”) fue la más adecuada para el análisis de 50 muestras de aliento condensado pertenecientes a individuos sanos, mostrando una variabilidad en el mismo día inferior al 7% para los 51 compuestos identificados, así como una adecuada distribución de los datos en términos de porcentaje de metabolitos que presentaban una distribución normal al aplicar el test de Skewness y Kurtosis. Finalmente, se llevó a cabo un estudio de recuperación para evaluar la eficacia de la LLE realizando tres extracciones consecutivas de la misma alícuota de aliento

condensado. La recuperación para la primera extracción proporcionó un valor medio del 77%. Los dos ciclos de extracción adicional provocaron una disminución de la sensibilidad causada por el efecto de dilución al mezclar las fracciones; por lo que se adoptó como óptimo un único ciclo de extracción para el análisis de aliento condensado.

La caracterización de la muestra puso de manifiesto una composición heterogénea claramente dominada por la presencia de ácidos grasos y derivados como ésteres metílicos, amidas y lípidos prenoles volátiles. Entre los ácidos grasos y derivados identificados, destacan el ácido mirístico (C14:0), el palmítico (C16:0) y el esteárico (C18:0), que junto con el ácido oleico (C18:1), constituyen aproximadamente el 78% de todos los ácidos grasos libres en el sistema circulatorio. Los derivados monoglicéridos del ácido palmítico y del esteárico también se identificaron. El ácido oleico no se identificó en su forma libre pero si se detectaron el metil ester derivado y la oleamida. Entre los derivados de ácidos grasos, se identificaron también la erucamida y la eicosenamida, importantes por sus propiedades antimicrobianas. También se detectaron alcoholes grados, como el undecanol, el estearato de glicidol, el derivado metilado del hexadecanol además de ácidos grasos de cadena corta como el caprílico, el pelargónico y el láurico.

Los prenoles, sintetizados a partir de unidades isopreno, constituyen otra familia de compuestos identificados en aliento consensado. Se detectaron tres clases de lípidos prenólicos: monoterpenos, diterpenos y triterpenos. Entre los monoterpenos, se identificaron compuestos como el borneol, el limoneno, el linalol, el alcohol cumílico y el terpineol, que previamente se habían detectado en aliento. Otros monoterpenos detectados fueron el alcanfor y el levomentol. El esclareol fue el único diterpeno detectado en el aliento y el escualeno, el único triterpeno. Éste último está implicado en la síntesis del colesterol y en la síntesis de unidades isopreno. Otros dos lípidos, la hediona (derivado del ácido jasmónico) y un derivado de esteroide, el colestadieno, también se identificaron en este trabajo.

Otra familia detectada en el aliento es la de los compuestos homomonocíclicos aromáticos, entre los que destacan el metil ester y el 4-etoxi-etil-ester derivados del ácido benzoico, el metoxicinamato de octilo y el ácido di-t-butyl-4-hidroxicinámico. Además, se detectaron dos de los compuestos intermedios del metabolismo del tolueno, el alcohol bencílico y el p-cresol, y un compuesto aromático relacionado con el catabolismo del triptófano, el indol. Finalmente, otros compuestos derivados del fenol que previamente no se habían identificado en aliento como el 2,4,6-triisopropilfenol fueron identificados

Por tanto, estos resultados, presentados en el Capítulo VI, ponen de manifiesto que este biofluido puede ser una alternativa a otros como suero/plasma, orina o esputo para la búsqueda de nuevos tests en el ámbito clínico para el diagnóstico de enfermedades respiratorias crónicas. Esta área presenta un vacío en cuanto a herramientas de cribado se refiere; por lo que nuevos modelos de cribado podrían ayudar en la clasificación de algunas patologías.

Capítulo VII

Con el objetivo de conseguir nuevos modelos de cribado, el método desarrollado se aplicó a muestras provenientes de una cohorte formada por pacientes con cáncer de pulmón (48), individuos con factor de riesgo (130), en este caso fumadores, y un tercer grupo compuesto por individuos sanos (61). De los 51 compuestos identificados en la investigación anterior, se excluyeron 3 (octinoxate, sclareol and 7,9-di-tert-butyl-oxaspiro(4,5)deca-6,9-diene-28-dione) debido a su procedencia claramente exógena. En este caso, se aprovechó la capacidad del muestreador para separar el aliento proveniente de las vías superiores del de las inferiores; por lo que se obtuvieron dos muestras de cada individuo, con las que estudiar la potencial diferente composición de estas fracciones, así como la capacidad de predicción de cada una de ellas. Los resultados mostraron que la composición de las dos fracciones es bastante similar, difiriendo más en la concentración de metabolitos que en la composición.

Para estudiar la capacidad de predicción de cada una de las fracciones se obtuvieron modelos supervisados mediante análisis discriminante de mínimos cuadrados parciales (PLS-DA del inglés Partial Least Squares Discriminant Analysis). Los individuos con cáncer de pulmón fueron correctamente clasificados frente a los controles en ambos modelos, presentando para la vía superior (upper airway) una sensibilidad del 83.7 % y una especificidad del 83.3% mientras que para la inferior (distal airway) se obtuvieron una sensibilidad y especificidad del 77.5 y 89.8% respectivamente. Los resultados empeoraron al comparar a los individuos con cáncer de pulmón y los controles con el grupo con factor de riesgo, obteniendo para la vía superior una sensibilidad y especificidad del 58.1 y 63.7% para los individuos con cáncer y un 63.7 y un 69.4% para los controles. Estos valores mejoraron para la vía inferior, obteniendo 75.5 y 70.5% de sensibilidad y especificidad respectivamente para los individuos con cáncer de pulmón y un 79.5 y 71.4 % para los controles. Este estudio demostró que la mejor fracción para discriminar entre individuos con cáncer de pulmón, individuos con factor de riesgo y sanos es la inferior.

Se identificaron los compuestos significativamente diferentes entre los tres grupos de individuos, resultando significativos lípidos como los monoacilgliceroles y el escualeno. Estos resultados soportan la capacidad de la metabolómica para profundizar en el estudio del cáncer de pulmón empleando aliento condensado como biofluido.

Capítulo VIII

La capacidad de discriminación observada soportó el desarrollo de paneles de metabolitos presentes en aliento condensado que pudieran usarse como método de cribado para el cáncer de pulmón discriminando tanto frente a individuos con factor de riesgo como frente a individuos sanos. Para ello se aplicó el método desarrollado exclusivamente a la fracción de aliento que proviene de la vía inferior y empleando los 48 metabolitos identificados (una vez eliminados del total de 51 los 3 compuestos exógenos). La cohorte estaba formada por pacientes con cáncer de pulmón (68), individuos con factor de riesgo (126) e individuos

sanos (62). Los paneles se crearon mediante la herramienta PanelomiX, que permite minimizar los falsos negativos y positivos, maximizando la especificidad y sensibilidad, respectivamente. Se compararon los grupos dos a dos, de forma que se obtuvieron 3 paneles compuestos por 5 metabolitos cada uno, y que presentaban una sensibilidad superior al 77.9% y una especificidad mayor del 67.5% así como un área bajo la curva superior al 77.5%. Esta área puede interpretarse como la probabilidad de que, ante un par de individuos, uno enfermo y el otro sano, la prueba los clasifique correctamente. De este modo, la sensibilidad representa el porcentaje de verdaderos positivos (proporción de individuos enfermos clasificados correctamente) y la especificidad, el porcentaje de verdaderos negativos (proporción de individuos sanos clasificados correctamente). Por lo tanto, para un método de cribado, se prima la sensibilidad por encima de la especificidad, pues los falsos positivos se pueden clasificar correctamente en el análisis confirmatorio mientras que un falso negativo genera un problema mayor.

Dos de los compuestos que resultaron significativos pertenecen a la ruta de degradación del tolueno, uno de los componentes del humo de tabaco. De hecho, los niveles de ambos compuestos presentaron una estrecha relación, aunque un comportamiento opuesto, detectándose niveles más altos de alcohol bencílico que de p-cresol en individuos con cáncer de pulmón. El tolueno se metaboliza principalmente a alcohol bencílico, que se convierte preferentemente en ácido hipúrico; y en una pequeña proporción da lugar a metabolitos como el o- y el p-cresol. La diferencia de los niveles entre estos dos compuestos en los tres grupos en estudio puede explicarse considerando que un aumento de tolueno en el cuerpo debe traducirse en un aumento de la producción de alcohol bencílico (por la relación preferente). Del mismo modo, la diferencia de concentración entre ambos compuestos no se verá tan potenciada para pequeñas cantidades de tolueno. Sin embargo, la comparación de los niveles de estos compuestos para los dos grupos de controles estudiados (con y sin factor de riesgo) no refleja cambios tan significativos, lo que sugiere que son los individuos con cáncer de pulmón los que presentan diferencias en el metabolismo del tolueno.

Un comportamiento similar se observa en cinco compuestos derivados del fenol (2,4-diphenyl-4-methyl-2-E-pentene, 2,4,6-triisopropylphenol, 2,6-bis(1,1-dimethylethyl)-4-(1-methyl-1-phenylethyl)-phenol, 2,4-bis(dimethylbenzyl-6-t-butylphenol), ya que estos compuestos se relacionan directamente con el hábito de fumar debido a su presencia en el humo del tabaco. También en este caso, el éster metílico del ácido benzoico, incluido en uno de los paneles, es un producto de degradación del xileno, un compuesto orgánico presente en el humo del tabaco, por lo que también se relaciona con el hábito de fumar.

Respecto al resto de los compuestos presentes en los paneles, cabe destacar que dos monoacilgliceroles fueron previamente detectados en sudor por investigadores del grupo al que pertenece la doctoranda; y otros, como el alcohol cumílico o el derivado del indano, también se detectaron por otros autores en la fracción volátil del aire exhalado.

Además, también hay que resaltar que ninguno de los compuestos detectados en aliento condensado que pueden atribuirse a fuentes exógenas (como el mentol o el limoneno) está presente en los paneles obtenidos.

Estos resultados preliminares requieren un estudio a gran escala para validar los paneles propuestos con un doble objetivo: reducir el número de individuos que tengan que someterse a un test confirmatorio y detectar el cáncer de pulmón en una etapa lo más temprana posible.

Un primer paso en esa detección temprana del cáncer de pulmón fue el estudio de las diferencias significativas entre estadios de la enfermedad. Para ello, los pacientes con las dos patologías de cáncer de pulmón más común (carcinoma de células escamosas y adenocarcinoma) y se clasificaron en dos grupos, uno de ellos (10 individuos) con estadios I y II y el otro (28 individuos) con estadios avanzados III y IV. Estos dos grupos junto con el grupo de individuos con factor de riesgo se sometieron a un test estadístico para encontrar tendencias de discriminación en los compuestos detectados en el aliento.

Seis compuestos –alcohol cumílico, ácido benzoico metil éster, 2,4,6-triisopropifenol, 2,6-bis (1,1-dimetiletil) -4- (1-metil-1-feniletil) fenol, 2,4-bis (1-

metil-1-feniletil) fenol y 2,4-bis (dimetilbencilo-6-t-butilfenol– resultaron significativos por su mayor concentración en individuos con cáncer en estadios avanzados. Cinco de ellos podrían estar relacionados directamente con el consumo del tabaco, cuatro de ellos por su estructura bencenoide y el metil ester derivado del ácido benzoico por su relación con el alcohol bencílico, lo que demuestra que estas alteraciones de la microbiota habitual de las vías respiratorias deben ser consideradas y estudiadas en profundidad.

Capítulo IX

El requerido estudio de las alteraciones de la microbiota de las vías respiratorias exclusivamente debidas al hábito de fumar se realizó utilizando muestras de donantes fumadores, ex fumadores y no fumadores.

El primer objetivo de la presente investigación se centra en la detección de modificaciones en la composición del aliento condensado asociadas al hábito de fumar. Con este propósito el análisis estadístico se llevó a cabo para comparar la composición de aliento condensado de fumadores y no fumadores. El análisis supervisado por PLS-DA permitió encontrar diferencias entre los dos grupos con un alto porcentaje de muestras correctamente clasificadas presentando una precisión global del 78.5 y del 66.7% en las etapas de entrenamiento y validación respectivamente. Por lo tanto, existe un patrón de variabilidad en la composición del aliento condensado asociado al hábito de fumar. El análisis estadístico mediante el test de la t permitió identificar doce compuestos presentes en concentraciones significativamente diferentes entre individuos fumadores y no fumadores con un valor de p inferior a 0,05. Entre estos compuestos cabe destacar cinco que eran altamente significativos, con un valor de p por debajo de 0,01: indol (valor p 0,0006), undecanol (0,0036) y tres compuestos fenólicos tales como p-cresol (0,0065), 2,4,6-tris (1-metiletil) fenol (0,0032) y 2,6-bis (1,1-dimetiletil) -4-(1-metil-1-feniletil) fenol (0,0031); y uno por su alto valor de cambio, la monoestearina. Estos compuestos se caracterizan por un perfil de concentración similar, con una mayor concentración relativa en el aliento condensado de

individuos no fumadores en comparación con los fumadores, a excepción de un compuesto, monoestearina, que estaba más concentrado en fumadores.

El indol fue el compuesto que presentó mayor significancia con un valor de p de 0.0006 y con un valor de cambio de 2.59 estando, como se ha dicho anteriormente, a mayor concentración en la población no fumadora. Esta variación está relacionada con la disminución en la abundancia de bacterias Firmicutes y Actinobacterias phyla en poblaciones afectadas por alguna enfermedad respiratoria. Estas bacterias se conocen por su capacidad para producir una variedad de compuestos antiinflamatorios como flavonoides y los alcaloides del indol. Por lo tanto, el hábito de fumar conduce a la modificación del microbioma de las vías respiratorias con incidencia directa sobre la respuesta inmune afectando a la liberación de metabolitos con propiedades antiinflamatorias y antimicrobianas reconocidas como el indol. La alteración de la respuesta inmune contribuye a explicar la susceptibilidad de los fumadores a las enfermedades respiratorias. Estos resultados quedan además soportados por el alto poder predictivo del indol para discriminar entre los dos grupos de interés presentando una curva ROC con 80 y 70% de especificidad y sensibilidad, respectivamente.

Entre los derivados de fenoles, el compuesto más relevante fue el p-cresol (0,0065), que participan en la ruta de degradación de tolueno como se ha comentado anteriormente. El tolueno se metaboliza principalmente a alcohol bencílico o a isómeros del hidroxitolueno, en particular a p-cresol. Este compuesto, dotado de propiedades antioxidantes, también fue encontrado en niveles más altos en no fumadores. Este resultado está relacionado con la disminución en la abundancia relativa de phylum Proteobacteria, Capnocytophaga, Peptostreptococcus y Leptotrichia en fumadores en comparación con los no fumadores, estando estos géneros bacterianos relacionados con los hidratos de carbono y el metabolismo energético, y con el metabolismo de xenobióticos, tales como la de tolueno. En particular, el mecanismo del tolueno se activa con la presencia de estas bacterias.

Un patrón similar se encontró para los dos derivados 2,4,6-trisisopropilfenol (0,0032) y 2,6-bis(1,1-dimetiletil)-4-(1-metil-1-feniletil)-fenol

(0,0031). Estos dos compuestos son estructuralmente similares a un antioxidante ampliamente conocido utilizado en las industrias de alimentos y cosméticos: el hidroxitolueno butilado (BTH, del inglés butylated hydroxytoluene). Ambos compuestos podrían haberse formado por la degradación del microbioma de compuestos exógenos, lo que podría explicar las diferencias en los niveles entre fumadores y no fumadores. La analogía estructural entre los dos alquilfenoles y el BTH permite deducir una actividad antioxidante de estos dos compuestos.

Otro compuesto altamente significativo entre fumadores y no fumadores fue el undecanol (0,0036), el cual no se ha relacionado con el microbioma. Algunos grupos han descrito la formación de alcoholes grasos a través de la degradación de los metabolitos de ácidos grasos poliinsaturados (PUFA, del inglés polyunsaturated fatty acid) que se consideran marcadores de la inflamación. Algunos de estos compuestos están dotados de propiedades anti o pro-inflamatorias, dependiendo del precursor PUFA. En este caso específico, undecanol estuvo presente en un nivel más alto en fumadores que en no fumadores, lo que podría explicar su formación mediante metabolitos PUFA con capacidad antiinflamatoria.

Conviene poner énfasis en la monoestearina, que con una significancia menor que los hasta ahora descritos (0,05 <valor de p <0,01), presentó el mayor valor de cambio entre fumadores y no fumadores (2.63). El tejido pulmonar de pacientes con adenocarcinoma presenta menor nivel de monoestearina que el tejido sano, mientras que monopalmitina informó la situación opuesta. Los niveles en el tejido pulmonar y en el aliento condensado podrían estar correlacionados debido al equilibrio de absorción/desorción durante la respiración.

El segundo objetivo de este estudio fue demostrar que el microbioma oral vuelve a su estado anterior para los exfumadores comparándolos con los fumadores, teniendo en cuenta además la carga tabáquica. El análisis estadístico mediante el test de la t permitió identificar 6 compuestos con una significancia inferior a 0.05 comparando fumadores y ex fumadores: alcanfor, undecanol, heptadecynol, escualeno y dos derivados fenólicos, 2,4,6-triisopropylfenol y el ácido benzoico-4-etoxi-etiléster.

El escualeno es un polímero del isopreno de origen natural conocido por su papel clave como intermedio en la síntesis de colesterol. Muchos otros derivados del isopreno estructuralmente similares al escualeno realizan funciones biológicas críticas tales como la inhibición de la oxidación o la generación de otros derivados con propiedades antimicrobianas. Por lo tanto, la concentración significativamente diferente del alcanfor y el escualeno en fumadores y exfumadores (escualeno está más concentrada en fumadores que en exfumadores y alcanfor muestra la tendencia opuesta) podría atribuirse al hecho de que el dejar de fumar estimula la síntesis de compuestos con propiedades antimicrobianas como alcanfor utilizando grupos isopreno del escualeno. Además, como se discutió anteriormente, el hábito de fumar altera la respuesta inmune que contribuye a explicar la susceptibilidad de los fumadores a las enfermedades respiratorias. Esta interpretación biológica se apoya en el comportamiento predictivo de ambos compuestos para discriminar entre fumadores y exfumadores. Como se puede observar, las curvas ROC para estos metabolitos presentaron alta especificidad y sensibilidad, 80 y 70% para el alcanfor y el 60 y 80% para el escualeno, respectivamente.

Otros dos compuestos altamente significativos entre individuos fumadores y no fumadores fueron undecanol (0,0100) y heptadecynol (0,0054) que, como se discutió anteriormente, la formación de estos alcoholes grasos podría estar relacionado con la degradación de los PUFA. Ambos compuestos estaban presentes en un nivel más alto en exfumadores que en los individuos fumadores, lo que podría explicar su fuente a través de PUFAs con capacidad antiinflamatoria.

El derivado fenólico 2,4,6-triisopropilfenol (0,0056) se discutió anteriormente como un compuesto relacionado con BTH con actividad antioxidante. Este compuesto presenta niveles más altos en exfumadores que en fumadores, apoyando los resultados obtenidos comparando fumadores con no fumadores. El ácido benzoico 4-etoxi-etilester, que registró el mayor valor de significancia entre fumadores y exfumadores (0,0048), podría resultar de la ruta de degradación del xileno. Esta ruta es inhibida por la presencia de altas dosis de

tolueno, lo cual puede explicar el alto nivel de este metabolito en los exfumadores, ya que dejar de fumar reduce la presencia de niveles de tolueno estimulando la degradación de xileno y aumentando el nivel de derivados de ácido benzoico.

En cuanto a la carga tabáquica, el alcohol bencílico resultó altamente significativo (0,0096), así como el eucaliptol (0,0071). El alcohol bencílico es uno de los dos productos intermedios en la vía del metabolismo de tolueno que presentan una tendencia opuesta al p-cresol. Por lo tanto, el nivel de alcohol bencílico aumenta con la carga tabáquica, lo cual concuerda con los niveles de p-cresol que eran más altos para no fumadores que para los fumadores.

Finalmente, el eucaliptol es el monoterpeno 1,8-cineol bicíclico, cuya hidrólisis está catalizada por una enzima monooxigenasa del citocromo P450, por lo tanto, los niveles más altos de eucaliptol con el aumento de la carga tabáquica se explica por el metabolismo de xenobióticos del citocromo P450, empobrecido en individuos fumadores y favoreciendo la acumulación de eucaliptol como resultado.

De acuerdo con lo anterior, el tabaquismo es una de las principales causas de las enfermedades respiratorias. En esta investigación recogida en el Capítulo IX de esta memoria, se ha vuelto a poner de manifiesto la aplicabilidad del aliento condensado en el campo clínico.

Estos resultados preliminares ponen en manifiesto la necesidad de un estudio a gran escala para validar la aplicabilidad de aliento condensado para detectar los trastornos causados por el hábito de fumar. Este primer enfoque también ha puesto de manifiesto diferencias estadísticamente significativas entre los grupos en estudio, principalmente centrándose en el alcanfor, indol y p-cresol, compuestos que se incrementan en el aliento condensado de los fumadores y exfumadores, mientras que monoestearina y escualeno disminuyen.

Los resultados generales obtenidos en la parte B de esta Memoria han revelado la importancia del estudio del aliento condensado como muestra clínica

para el estudio de numerosas enfermedades respiratorias ya que permite detectar diferencias sólo teniendo en cuenta el hábito de fumar. Además, ha demostrado tener un gran potencial en la discriminación entre pacientes con y sin cáncer de pulmón, permitiendo también diferenciar entre estadios de la propia enfermedad. El siguiente paso de esta investigación pasaría por seguir reclutando muestras con el fin de enriquecer la base de datos con un mayor número de pacientes, de forma que, por ejemplo, la consideración de los estadios tuviera mayor peso. Además, el análisis puede hacerse dirigido empleando los compuestos que han resultado significativos.

CONCLUSIONES

La investigación que constituye esta Tesis se ha centrado en el análisis metabolómico global en dos áreas: vegetal y clínica, lo que ha puesto de manifiesto la utilidad de la espectrometría de masas y su capacidad para la innovación metodológica, así como para la obtención de información de la que han derivado las siguientes conclusiones:

- En el estudio de la uva y sus componentes ha permitido desarrollar metodologías para la caracterización de "residuos de residuos", que resultan tras la destilación alcohólica del orujo de uva generado en la elaboración del vino, poniendo de manifiesto que:

(i) El perfil de compuestos obtenido a partir de la misma materia prima resulta muy diferente en función del tipo de energía que se utilice para acelerar/mejorar la etapa de extracción.

(ii) La extracción con líquidos sobrecalentados y el método convencional basado en maceración proporcionan un mayor número de compuestos en comparación con la extracción asistida por ultrasonidos o por microondas.

(iii) La disminución del pH del extractante aumenta la eficacia de la extracción con líquidos sobrecalentados frente a la convencional, además de reducir significativamente el tiempo requerido (60 min en lugar de 24 h).

(iv) Se extraen más compuestos polares que apolares como consecuencia tanto del agente de extracción empleado como del alto contenido de compuestos polares en los residuos.

(v) La variedad de compuestos tentativamente identificados en los extractos convierte a estos en potenciales aditivos en la industria alimentaria (ya sea como colorantes, como los modificadores del

sabor o como antioxidantes), y también en las industrias de cosméticos y nutracéuticos.

(vi) Todo lo anterior permite concluir de este estudio que, a pesar de la degradación causada por las condiciones drásticas del proceso de destilación, existen muchos compuestos de interés en los extractos de estos desechos, lo que los convierte en un material útil para un mejor uso que ser fuente de calor.

- El desarrollo de la metodología más apropiada para la caracterización de alperujo, residuo generado tras la elaboración de aceite de oliva, demostrando que:

(i) Los extractos obtenidos utilizando presión+temperatura altas contienen un gran número de fenoles (corroborado tanto con la medida de fenoles totales en el extracto por el método de Folin-Ciocalteu, como por el perfil fenólico obtenido mediante LC-QTOF).

(ii) Los 49 compuestos identificados en estos extractos también se encuentran en el aceite de oliva virgen. Entre ellos cabe destacar hidroxitirosol/tirosol y sus derivados, los precursores de iridoides, secoiridoides y sus derivados, flavonoides, lignanos y ácidos fenólicos.

(iii) Se ha conseguido un extracto con alto valor añadido y un residuo libre de antioxidantes que presenta unas características óptimas para la producción de gas, compost, etc., es decir, para cualquier tipo de aprovechamiento en el que la presencia de antioxidantes suponga un problema o una ralentización en la obtención del producto.

- Las innovaciones desarrolladas en análisis global en el área clínica permiten concluir que:

(i) Los biofluidos poco convencionales como el aliento condensado deben considerarse en investigación clínica. Las ventajas del aliento condensado como muestra clínica están soportadas en su muestreo no invasivo y en su composición simple.

(ii) El análisis de aliento condensado mediante GC-TOF/MS tras una simple extracción líquido-líquido posibilita la identificación de una amplia variedad de familias de metabolitos, lo que lo convierte en una opción muy adecuada para obtener una instantánea del metaboloma de este biofluido.

(iii) La normalización basada en la suma total de la señal útil (MSTUS) es una opción muy apropiada para obtener una adecuada distribución de los datos además de una mínima variabilidad eliminando fuentes de variabilidad externa.

(iv) La comparación de los metabolitos existentes en el aliento condensado de individuos sanos, con factor de riesgo y con cáncer de pulmón muestra las diferencias en el metabolismo de estos grupos, según el análisis llevado a cabo con la plataforma GC-TOF/MS.

(v) Se ha constatado que la fracción inferior (vía respiratoria distal) del aliento condensado es la más adecuada para la discriminación entre individuos con cáncer de pulmón, con factor de riesgo e individuos sanos.

(vi) Se han propuesto tres paneles de marcadores para la predicción del cáncer de pulmón que se caracterizan por una sensibilidad superior al 77.9% y una especificidad mayor del 67.5%, así como por un área bajo la curva superior al 77.5 %. Dos compuestos de estos paneles corresponden a la ruta de degradación del tolueno, uno de los componentes del tabaco.

(vii) Se han identificado seis compuestos directamente relacionados con el tabaco que se encuentran a mayor concentración en individuos con cáncer que en individuos con riesgo de padecerlo o en individuos sanos.

(viii) Se han identificado 12 compuestos significativos en la comparación del perfil del aliento condensado de fumadores y no fumadores, y 6 en la comparación de exfumadores y fumadores. En ambos test estadísticos aparecen un compuesto fenólico, el 2,4,6-triisopropylfenol y un alcohol graso, el undecanol. Además, entre los primeros cabe destacar el indol, un derivado de ácido graso (monoestearina) y dos compuestos fenólicos tales como p-cresol, y 2,6-bis (1,1-dimetiletil) -4- (1-metil-1-feniletil) fenol. En el segundo caso alcanfor, heptadecynol, escualeno y un derivado fenólico, el ácido benzoico-4-etoxi-etiléster.

(ix) Se ha relacionado el cambio significativo en concentración de ciertos compuestos con la alteración del microbioma de las vías respiratorias causada por el hábito de fumar. La alteración tiene una incidencia directa en la respuesta inmune afectando a la producción de metabolitos con propiedades anti-inflamatorias y antimicrobianas.

ANEXOS

Anexo I

Investigación en la que ha participado la doctoranda y ha dado lugar a publicaciones

1. Cholesterol oxidation products in milk: processing formation and determination
***M. Calderón-Santiago**, A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro*
European Journal of Lipid Science and Technology 114 (2012) 687-694

---

2. Sunlight exposure increases the phenolic content in postharvested white grapes. An evaluation of their antioxidant activity in Saccharomyces cerevisiae
*José Peinado, Nieves López de Lerma, **Angela Peralbo-Molina**, Feliciano Priego-Capote, Cristina de Castro, Brian McDonagh*
Journal of Functional Foods 5 4 (2013) 1566-1575

---

3. Influence of the collection tube on metabolomic changes in serum and plasma
*MA López-Bascón, F Priego-Capote, **A Peralbo-Molina**, M Calderón-Santiago, MD Luque de Castro*
Talanta 150 (2016) 681-689

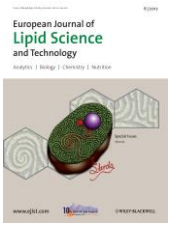
---

4. Ageing red wines by oak chips or vine-shoots: a comparison of gas chromatography-mass spectrometry
M. Pilar Delgado-Torre, A. Peralbo-Molina, P. Pérez-Juán, F. Priego-Capote, M. D. Luque de Castro
Sent to Analytical Methods

European Journal of Lipid Science and Technology

Volume 114, Issue 6, June 2012

pages 687-694,



Cholesterol oxidation products in milk: Processing formation and determination

*M. Calderón-Santiago, A. Peralbo-Molina, F. Priego-Capote, M. D. Luque de Castro**

Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, and Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, E-14014, Córdoba, Spain

Abstract

Cholesterol oxidation products (oxysterols) are commonly present in foods derived from animals. The main limitations for analysis of these products are associated with the long sample preparation protocols that require saponification. In this research, a protocol was optimized for determination of oxysterols in milks under the process to obtain given commercial milk products. UHT milks, powder milks, condensed milk, and evaporated milk were selected for this study. The levels of the different oxysterols are discussed attending to the type of milk. UHT and skimmed powder milks provide the highest concentration of oxysterols. Particularly, high concentrations were found for 7β-hydroxycholesterol (up to 205.6 ng/mL), 5-cholesten-3β-ol-7-one (up to 21.8 ng/mL), cholesterol 5α,6α-epoxide (up to 14.9 ng/mL), and 25-hydroxycholesterol (up to 5.1 ng/mL). An increase in cholesten-3β-ol-7-one and cholesterol 5α,6α-epoxide, resulting from the skimming process, was also observed. Additionally, two different heating modes, microwave-assisted and conventional heating, have been mimicked to study their influence on the formation of oxysterols in milk. Both heating protocols

contributed to the formation of oxysterols but short periods of microwave irradiation (60 s) can lead to levels of oxysterols similar to conventional heating for 5 min.

Practical applications: The results of the research described in this article, supported by previous studies in the field, allow to advice about the strong influence of microwave heating of milk on the formation of toxic oxides from the lipid fraction. Although conventional heating can be more tedious and time consuming, it seems to be healthier.

Journal of Functional Foods

Volume 5, Issue 4, October 2013,

Pages 1566-1575

Sunlight exposure increases the phenolic content in postharvested white grapes. An evaluation of their antioxidant activity in *Saccharomyces cerevisiae*

José Peinado[a,b] , Nieves López de Lerma[c], Angela Peralbo-Molina[b, d], Feliciano Priego-Capote[b, d], Cristina de Castro[e], Brian McDonagh[a, f]

[a] Department of Biochemistry and Molecular Biology, University of Córdoba, E-14071 Córdoba, Spain

[b] Institute of Biomedical Research Maimónides (IMIBIC), ReinaSofía Hospital, University of Córdoba, E-14071 Córdoba, Spain

[c] Department of Agricultural Chemistry, University of Córdoba, E-14071 Córdoba, Spain

[d] Department of Analytical Chemistry, University of Córdoba, E-14071 Córdoba, Spain

[e] Department of Molecular Biology, University of León, E-24071 León, Spain

[f] Department of Musculoskeletal Biology, Institute of Ageing and Chronic Disease (IACD), University of Liverpool, L69 3GA, United Kingdom

Abstract

Sunlight exposure is a traditional way to dry postharvested white Pedro Ximenez grapes, and the dried grapes are the basis of a sweet fortified wine elaborated in

the South of Spain. In this paper, we have studied the effect of seven days of exposure of grapes to sunlight. The phenolic content in the skins and flesh has been determined. The transient induction of phenolics has been detected and identified in skins by HPLC and MS/MS. Maximum induction was after 2 days exposure with an increase in the levels of quercetin-3-O-glucoside, engeletin (dihydrokaempferol-3-O-rhamnoside) and astilbin (taxifolin 3-O-rhamnoside). We have evaluated the antioxidant and protective effects of the phenolic extracts of these grapes in Saccharomyces cerevisiae after exposure to H_2O_2. Phenolic extracts reduced the basal intracellular level of peroxides in a concentration dependent manner. There was a corresponding significant reduction in carbonylated proteins in treated and control cells and increased survival of yeast cells exposed to H_2O_2. Our results indicate that sundried grapes display a high antioxidant capacity resulting in a decrease in the oxidative cellular environment.

Talanta

Volume 150, 1 April 2016,

Pages 681-689

Influence of the collection tube on metabolomic changes in serum and plasma

M.A. López-Bascón[a, b, c], F. Priego-Capote[a, b, c], , , A. Peralbo-Molina[a, b, c], M. Calderón-Santiago[a, b, c], M.D. Luque de Castro[a, b, c]

[a]Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain

[b]University of Córdoba Agroalimentary Excellence Campus, ceiA3, Campus of Rabanales, 14071, Córdoba, Spain

[c]Maimónides Institute for Research in Biomedicine of Córdoba, Reina Sofía University Hospital, University of Córdoba

Abstract

Major threats in metabolomics clinical research are biases in sampling and preparation of biological samples. Bias in sample collection is a frequently forgotten aspect responsible for uncontrolled errors in metabolomics analysis. There is a great diversity of blood collection tubes for sampling serum or plasma, which are widely used in metabolomics analysis. Most of the existing studies dealing with the influence of blood collection on metabolomics analysis have been restricted to comparison between plasma and serum. However, polymeric gel tubes, which are frequently proposed to accelerate the separation of serum

and plasma, have not been studied. In the present research, samples of serum or plasma collected in polymeric gel tubes were compared with those taken in conventional tubes from a metabolomics perspective using an untargeted GC-TOF/MS approach. The main differences between serum and plasma collected in conventional tubes affected to critical pathways such as the citric acid cycle, metabolism of amino acids, fructose and mannose metabolism and that of glycerolipids, and pentose and glucuronate interconversion. On the other hand, the polymeric gel only promoted differences at the metabolite level in serum since no critical differences were observed between plasma collected with EDTA tubes and polymeric gel tubes. Thus, the main changes were attributable to serum collected in gel and affected to the metabolism of amino acids such as alanine, proline and threonine, the glycerolipids metabolism, and two primary metabolites such as aconitic acid and lactic acid. Therefore, these metabolite changes should be taken into account in planning an experimental protocol for metabolomics analysis.

Sent to

Analytical Methods

Ageing red wines by oak chips or vine-shoots: a comparison of gas chromatography-mass spectrometry

M. Pilar Delgado-Torre, [a,b,c] *A. Peralbo-Molina,* [a,b,c] *P. Pérez-Juán,*[d] *F. Priego-Capote,*[*a,b,c] *and M. D. Luque de Castro**[a,b,c]

[a]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain*

[b]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, E-14071, Córdoba, Spain*

[c]*University of Córdoba, Agroalimentary Excellence Campus, ceiA3, Campus of Rabanales, 14071, Córdoba, Spain.*

[d]*Bodega y Viñedos Vallebravo, Sierra de Segura, 23350, Puente de Génave, Jaén, Spain*

Abstract

Changes in the composition of red wine promoted by contact with either vine-shoots or oak chips have been studied by gas chromatography-mass spectrometry (GC-MS). The wine was aged in the presence of each of the materials for 3 or 9 months, and the chromatographic profiles of the different

samples were compared among them and with control wine which was not in contact with any of the ageing agents. This strategy allowed obtaining a representative profile of compounds present in each sample. In general, at short ageing times both materials affected similarly the composition of wine by increasing the levels of compounds responsible for floral, sweet and smoky attributes with respect to control wines. After 9 months ageing, wines in the presence of oak chips were statistically more concentrated in compounds responsible for herbaceous notes, floral aroma, green, smoky and sweet taste than wines aged in the presence of vine-shoots and control wines. Finally, identification of the compounds representative of the differences was carried out; thus showing vine-shoots as a feasible alternative or complement to traditional ageing methods. This is the first time that wines aged in the presence of vine-shoots are characterized.

Anexo II

Revisión en la que la doctoranda participa como co-autor

Journal of Chromatography B

Volume 879, Issues 17–18, 2011,

Pages 1189–1195

The role of ultrasound in analytical derivatizations

M.D. Luque de Castro, F. Priego-Capote, A. Peralbo-Molina

Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain

Abstract

Ultrasound is a type of energy that until recently was rarely used for analytical purposes. In recent years, work in chemical and industrial fields alerted analytical chemists to the great potential of ultrasonic energy to accelerate or improve different steps of the analytical process. One of these steps is derivatization: depolymerization, redox, hydrolysis, esterification, alkylation and complex formation are examples of derivatization reactions, all of which are significantly improved with the aid of ultrasound. This review discusses the valuable characteristics of ultrasound and its influence on a number of derivatization reactions is discussed in this review.

Anexo III

Patente en la que la doctoranda participa como co-autor

No Patente: P201531301

BIOMARCADORES EN AIRE EXHALADO PARA EL DIAGNÓSTICO, CLASIFICACIÓN Y SEGUIMIENTO DEL CÁNCER DE PULMÓN

Bernabé Jurado Gámez, María Dolores Luque de Castro, Feliciano Priego Capote, Mónica Calderón Santiago, Ángela Peralbo Molina

Los autores de la presente invención han analizado la concentración de los distintos metabolitos del aire exhalado en individuos que no padecen cáncer de pulmón y en individuos que padecen cáncer de pulmón. 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. Esto, entre otras cosas, permitiría hacer un cribado inicial para diferenciar aquellos individuos que serían susceptibles de ser sometidos a otras pruebas diagnósticas, más agresivas o caras, y/o confirmar o apoyar el diagnóstico mediante otras pruebas. Así pues, la presente invención proporciona un método de obtención de datos útiles para la clasificación, diagnóstico y seguimiento de individuos con cáncer de pulmón.

Anexo IV

Comunicaciones a congresos y reuniones

1. Determination of oxysterols in dairy products and study of their formation under heating conditions

M. Calderón-Santiago, A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro

7th Lipidomics Congress (Lipids in all states)

Anglet-Biarritz (Francia), 2010

Tipo de evento: Poster en Congreso **Ámbito:** Internacional

---

2. Determinación de oxisteroles en derivados lácteos mediante cromatografía de gases con detección por espectrometría de masas y estudio de la influencia de la energía microondas en su formación

M. Calderón-Santiago, A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro

XII Reunión del Grupo Regional Andaluz de la Sociedad Española de Química Analítica

Córdoba, España, 2010

Tipo de evento: Poster en Congreso **Ámbito:** Nacional

---

3. Towards a comprehensive exploitation of residues from wine production

A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro

The International Conference On Natural Products

Castres (France), 2011

Tipo de evento: Poster en Congreso **Ámbito:** Internacional

---

4. Towards a comprehensive exploitation of residues from wine production

A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro

V Jornadas de Divulgación de la Investigación en Biología Molecular, Celular, Genética y Biotecnología

Universidad de Córdoba, Córdoba, España, 2012

Tipo de evento: Póster en congreso **Ámbito:** Autonómico

5. Optimization of superheated liquid extraction for the characterization of phenolic compounds in alperujo by using LC–QqTOF MS/MS

A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro

XXIII Reunión Nacional de Espectroscopía – VII Congresso Ibérico de Espectroscopía

Córdoba, España, 2012

Tipo de evento: Poster en Congreso **Ámbito:** Nacional

---

6. Aprovechamiento de desechos de la agricultura y de la industria agroalimentaria andaluzas

M.D. Luque de Castro, F. Priego-Capote, V. Sánchez de Medina-Banea, P. Delgado de la Torre, A. Peralbo-Molina,

XII Reunión del Grupo Regional Andaluz de la Sociedad Española de Química Analítica

Málaga, España, 2012

Tipo de evento: Poster en Congreso **Ámbito:** Nacional

---

7. Optimization of superheated liquid extraction for the characterization of phenolic compounds in alperujo by using LC–QqTOF MS/MS

A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro

International Congress on "Green Extraction of Natural Products"

Université d'Avignon et des Pays de Vaucluse, France, 2013

Tipo de evento: Poster en Congreso **Ámbito:** Internacional

---

8. Optimization of a protocol to obtain superheated liquid extracts enriched with preferred phenols from olive pomace

A. Peralbo-Molina, F. Priego-Capote, M.D. Luque de Castro

XVIII Reunión de la Sociedad Española de Química Analítica

Úbeda, España, 2013

Tipo de evento: Póster en congreso **Ámbito:** Internacional

9. Sample preparation for untargeted metabolomics analysis of exhaled breath condensate

A. Peralbo-Molina, M.A. Fernández-Peralbo, F. Priego-Capote, M.D. Luque de Castro

IV Jornadas de Jóvenes Investigadores en Biomedicina

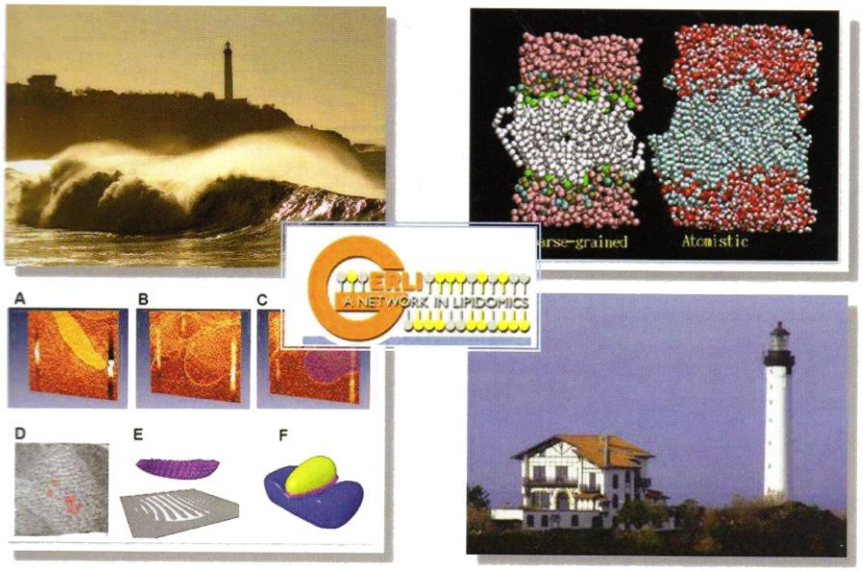
Córdoba, España, 2013

Tipo de evento: Poster en Jornada **Ámbito:** Regional

Groupe d'Etude et de Recherche en Lipidomique

METABOLOMIC STUDY OF CORONARY LESIONS

7TH LIPIDOMICS CONGRESS
« LIPIDS IN ALL STATES »



3rd-6th October 2010 Anglet-Biarritz, France

P8 - Determination of oxysterols in dairy products and study of their formation under heating conditions

Mónica Calderón-Santiago[1], Ángela Peralbo-Molina[1], Feliciano Priego-Capote[1] and M. D. Luque de Castro[1]

[1]Department *of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain*

Email: b42casam@uco.es

Cholesterol, because of its chemical structure, is a compound sensitive to oxidation, generating a series of derived products that are collectively termed oxysterols. Oxysterols can be formed in processing of foods under different physico-chemical conditions (heating, cooling, etc.). The clinical interest aims at oxysterols can be absorbed in the small intestine and incorporated into chylomicrons. Additionally, oxysterols have shown different toxic effects, some of which are alteration of the structure and function of the cellular membranes or changes in the activity and expression of enzymes involved in cholesterol biosynthesis [1]. All these characteristics of oxysterols make monitoring the presence of them in foods a necessity from the standpoint of food safety. Furthermore, the oxidation mechanism is not well known and must be clarify [2].

Oxysterols determination in food products is a challenge in analytical chemistry due to the low concentration of these compounds in processed foods and the complexity of the lipid fraction. In this research, a method for determination of oxysterols in milk-derived products was developed. Additionally, a study about the manufacture process on the formation of oxysterols was carried out. This study was based on elucidating the influence of microwave energy *versus* conventional heating (both used in manufacture processes) on their formation: the higher influence of the former on degradation leads to alert about the importance of microwave energy (time and power) on the quality of the dairy products of daily intake.
After mimicking the processing of products, treated samples were prepared by liquid–liquid extraction following saponification in order to isolate oxysterols present in the unsaponifiable fraction. This fraction was cleaned by solid-phase extraction with aminopropyl cartridges in order to remove interferences and preconcentrate oxysterols. Finally, oxysterols were derivatized to trimethylsilyl ethers for subsequent analysis by CG-MS in single ion monitoring (SIM) mode for selective determination of the oxysterols panel.

References:

1. Valenzuela, A., Sanhueza, J., Nieto, S. (2002) Cholesterol oxides (oxisterols): factors conditioning their formation, biological effects and content in foods, *Rev. Chil. Nutr. 29*, 116.
2. Smith, L. (1996) Review of progress in sterol oxidation, Lipids *31*, 453–488

REUNIÓN DEL GRUPO REGIONAL ANDALUZ DE LA
SOCIEDAD ESPAÑOLA DE QUÍMICA ANALÍTICA

XII GRASEQA 2010

Córdoba, 10 y 11 de junio

LIBRO DE RESÚMENES

UNIVERSIDAD Ð CORDOBA

GRASEQA 2010

SC-37

Determinación de oxisteroles en derivados lácteos mediante cromatografía de gases con detección por espectrometría de masas y estudio de la influencia de la energía microondas en su formación

M. Calderón-Santiago[a,b], A. Peralbo-Molina[a,b], F. Priego-Capote*[a,b], M.D. Luque de Castro[a,b]

[a]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain*
[b]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, E-14071, Córdoba, Spain*

Los procesos de oxidación en la industria alimentaria son indicadores nutricionales y de calidad. Los oxiesteroles o productos derivados de la oxidación del colesterol constituyen un ejemplo claro de este tipo de procesos, ya que exhiben actividades biológicas indeseables tales como citotoxicidad o la aparición de mutaciones. Los oxiesteroles han demostrado ser biológicamente activos en experimentos *in vitro* actuando como ligandos de los receptores X del hígado (*liver X receptors,* LXRs) que regulan la síntesis enzimática de colesterol en la ruta SREBP-2 y aceleran la degradación de la enzima HMG-CoA reductasa. A pesar de esta notable actividad *in vitro*, su papel *in vivo* es todavía objeto de debate. Teniendo en cuenta la presencia de colesterol en productos alimentarios de origen animal tales como leche, carne o huevos, la formación de oxiesteroles puede activarse mediante el procesado de alimentos a altas temperaturas o su almacenamiento prolongado.

La determinación de oxiesteroles en productos alimenticios supone un reto en química analítica debido a la baja concentración a la que se encuentran y a la presencia de interferencias procedentes de la fracción lipídica. En este caso se realizó un estudio de determinación de oxisteroles en distintos derivados lácteos y de la influencia de la energía de microondas en su formación. Para ello, se llevó a cabo un proceso de simulación de la formación de estos compuestos en un hipotético procesado del alimento. El proceso analítico se inició con la extracción líquido–líquido de la fase lipídica seguido de su saponificación con el fin de aislar los oxiesteroles contenidos en la fracción insaponificable. Esta fracción se trató mediante extracción en fase sólida utilizando cartuchos empaquetados con aminopropilo con el fin de eliminar interferencias y preconcentrar los oxiesteroles. Finalmente, se llevó a cabo la derivatización de los oxisteroles mediante sililación para su posterior análisis mediante CG-MS en modo SIM ("Single Ion Monitoring") con el fin de determinar de forma selectiva el perfil de los siete oxiesteroles seleccionados para este estudio.

SEIGNEMENT PERIEUR SUD DU TARN

Université de Toulouse

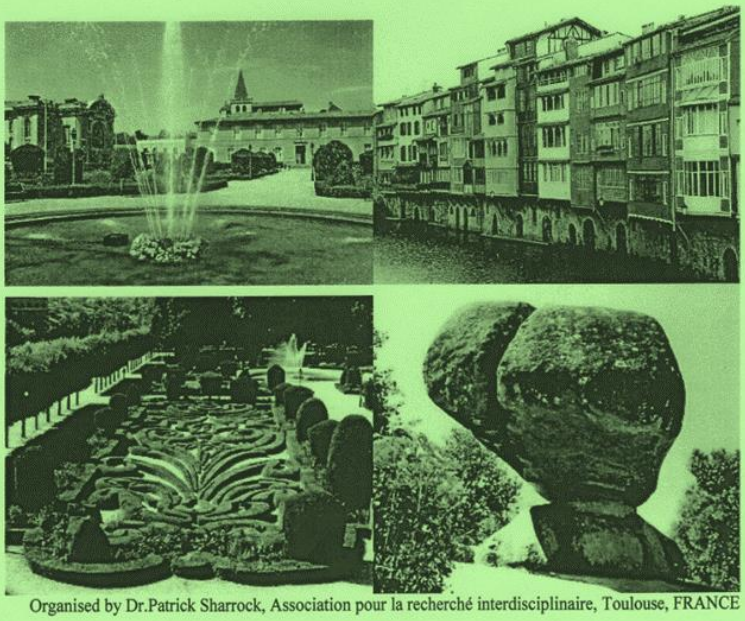
Université Paul Sabatier TOULOUSE III

RÉGION MIDI-PYRÉNÉES

The International Conference on Natural Products

24-28 Mai 2011, Castres, FRANCE

ISIS, Centre Universitaire J.F.Champollion



Organised by Dr.Patrick Sharrock, Association pour la recherché interdisciplinaire, Toulouse, FRANCE

P1-P

Towards a comprehensive exploitation of residues from wine production

A. Peralbo-Molina, F. Priego-Capote, M. D. Luque de Castro
Department of Analytical Chemistry, C-3 Annex to Marie Curie Building, Campus of Rabanales, University of Córdoba, 14071 Córdoba, Spain

The Mediterranean agrifood industry has traditionally been characterized by one cultivation–one manufactured product (*e.g.* vineyard–wine, olive tree–oil). Almost all by-products from vineyards (*i.e.* vine-shoots, leaves) and wine production (*i.e.* wine-pomace, lees), and also from olive trees (*i.e.* leaves, stems) and oil production (*i.e.* alperujo —semisolid residue from the two-phase extra-virgin olive-oil) are low- or nil-priced materials with high-contamination charge in cases such as alperujo.

Presently, there is a growing trend to take profit from these residues as they are rich in compounds of high interest for the food, cosmetic or pharmaceutical industries because their healthy properties as antioxidants, colorants and scavengers, among others.

In the case of the vineyard–wine binomial, an incipient interest has appeared about the extraction of oil and tannins from seeds and colorants from grape skin; thus exploiting wine pomace in a small proportion. Nevertheless, the major part of the wine pomace is used to obtain alcohol by using high temperatures. After this process, the remaining residue (RR) is mainly used as heating source.

The objectives of the present research, using RR from red and white wines, were as follows:

(i) To check several extraction methods (*i.e.* stirring for 24 h at 40 °C; ultrasound assistance either with the probe immersed in the solid–liquid system or in an external water bath surrounding the flask containing the solid–liquid system; microwave-assistance; high pressure and temperature assistance —superheated liquid extraction, SHLE) separately for seeds and skin using an 1:1 ethanol–water mixture at pH 3 in all cases.

(ii) To select the best extraction method —SHLE—, after which the temperature of the system was in-depth optimized between 120 and 220 °C, and the extraction kinetics studied at the optimal temperatures.

(iii) To quantify the total content of phenol compounds by using both the Folin–Ciocalteu method and obtaining the overall profile of the extracted compounds by liquid chromatography–diode array detection.

(iv) To identify the compounds in the extract by liquid chromatography–quadrupole/time-of-flight.

A number of high-interest compounds both naturally existing in seeds and skin and formed as a consequence of the high temperatures for ethanol production open the door to full exploitation of this non-valuable material.


V JORNADAS DE DIVULGACIÓN DE LA INVESTIGACIÓN EN BIOLOGÍA MOLECULAR, CELULAR, GENÉTICA Y BIOTECNOLOGÍA
Córdoba, 27-28 marzo 2012
Organiza
Departamento de Bioquímica y Biología Molecular
Universidad de Córdoba
Autores:
Juan Jurado Carpio
Víctor Manuel Luque Almagro
Ángel Llamas Azúa
Rosario Blanco Portales

Grupo BIO128

Panel 11

TÍTULO: La enzima con cofactor de Molibdeno CrARC tiene una actividad dependiente de Zn.
AUTORES: Alejandro Chamizo-Ampudia, Aurora Galván, Emilio Fernández, Ángel Llamas.
PRESENTADO EN: XXXIV Congreso Nacional de la SEBBM; Barcelona 2011.

Panel 12

TÍTULO: Fotoproducción de Hidrogeno en *Chlamydomonasreinhardtii*.
AUTORES: JoséLuis Jurado-Oller, David González-Ballester, Aurora Galván, Emilio Fernández.
PRESENTADO EN: XXXIVCongreso Nacional de la SEBBM; Barcelona 2011.

Grupo FQM227

Panel 13

TÍTULO: Tailor-made enriched oil with olive phenols: improvement of stability/ quality properties and health benefits.
AUTORES: Verónica Sánchez de Medina, Feliciano Priego Capote, Carlos Ferreiro Vera, María Dolores Luque de Castro.
PRESENTADO EN: 5th International Conference on Polyphenols and Health; Sitges, Barcelona 2011

Panel 14

TÍTULO: Towards a comprehensive exploitation of residues from wine production.
AUTORES: Ángela Peralbo Molina, Feliciano Priego Capote, María Dolores Luque de Castro.
PRESENTADO EN: The International Conferenceon Natural Products. Castres; Francia 2011

Grupo BIO202

Panel 15

TÍTULO: Hak1p es el transportador de potasio de alta afinidad en *Hansenulapolymorpha*.
AUTORES: Maria C alvarez, Elisa Cabrera, Yusé Martín, José M. Siverio, José Ramos.
PRESENTADO EN: XXXIV Congreso Nacional de la SEBBM; Barcelona 2011.

Panel 16

TÍTULO: Estudio proteómico del mutante trk1,2 de *Saccharomyces cerevisiae* en condición limitante de potasio.
AUTORES: Samuel Gelis, Inmaculada Redondo, Casimiro Barbado, Jesús Jorrín, José Ramos.

XXIII REUNIÓN NACIONAL – VII CONGRESSO IBÉRICO
ESPECTROSCOPÍA
CÓRDOBA 2012 (XXIII RNE – VII CIE)

ACTAS DE LA XXIII RNE – VII CIE

LIBRO DE COMUNICACIONES

UNIVERSIDAD D CÓRDOBA

CÓRDOBA, 17 AL 20 DE SEPTIEMBRE DE 2012

Optimization of superheated liquid extraction for the characterization of phenolic compounds in alperujo by using Lc–qqtof ms/ms

Ángela Peralbo Molina[1,2], Feliciano Priego-Capote[1,2], María Dolores Luque de Castro[1,2]

[1]Department of Analytical Chemistry, Annex Marie Curie Bulding, Campus of Rabanales, 14071, University of Córdoba, Spain
[2]Maimónides Institute of Biomedical Research (IMIBIC), Reina Sofía University Hospital, University of Córdoba, E-14071, Córdoba, Spain

The olive oil industry produces vast amounts of waste from the production line. The waste generated after oil decantation in the two-phases system is a semi-solid residue, known as olive pomace or alperujo. This consists of a mixture of the liquid and solid phases (alpechin and orujo, respectively), which were the original residues formed in the previous three-phases process. Research carried out so far for valorization of olive pomace extracts has been mainly focused on panels of target known compounds considered of interest, but global characterization of these extracts has not been made yet. Therefore, a global profiling of these compounds would be of interest for a better understanding of the biochemical pathways in which they are involved. Similar studies have been reported for olive leaves and other residues.

A simple and rapid method for the extraction of phenolic compounds from olive pomace using superheated liquid extraction (SHLE) has been developed. The experimental variables that affect SHLE, such as solvent type and composition, extractant temperature and extraction time, have been optimized. The Folin-Ciocalteu test was used for evaluation of the extraction conditions. Figure 1 illustrates the base peak chromatogram (BPC) obtained by LC–QqTOF-MS/MS under the optimum extraction conditions (75:25 ethanol–deionized water; pH 3; 160 °C). As can be seen, phenolic compounds are mainly eluted in the time window from 2 to 28 min which fits approximately with a concentration of acetonitrile in the chromatographic gradient of 78%. On the other hand, few peaks are detected in the last part of the chromatogram corresponding to non-polar compounds.

Once defined the potential of olive pomace for isolation of phenolic compounds, characterization of representative families of these compounds was carried out by tandem mass spectrometry in high accuracy mode. Identification and confirmatory detection was supported on mass accuracy of precursor ion and representative product ions whose structure was elucidated, but also on isotopic distribution of signals detected in full scan mode.

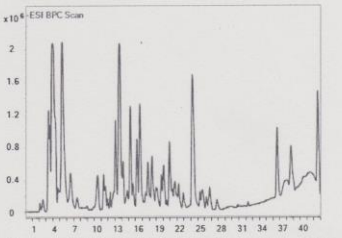


Figure 1. LC–QqTOF BPC obtained under the optimum extraction conditions.

Complete characterization started with identification of hydroxytyrosol and tyrosol as well as their major derivatives (glucoside derivatives), followed by identification of larger secoiridoid precursors (loganin, loganic acid and its derivatives). At this point, secoiridoids were exhaustively studied. Figure 2 shows the chromatograms used for identification of oleuropein and its aglycone form, one of the most important metabolites of this family of compounds. This especially well-known compound has been described to posses interesting functional properties including antioxidant, anti-inflammatory, anti-atherogenic, anti-cancer and antimicrobial activities, among others. The precursor ion for oleuropein was detected at *m/z* 539.2019. As can be seen in the figure, the fragmentation of this ion generated product ions with *m/z* 377, 307 and 275, which are characteristic of the oleuropein structure. The fragmentation scheme of oleuropein enabled identification of these representative fragments by cleavage of the hexose unit (*m/z* 377), while the other ions were assigned to further fragmentation products of the oleuropein aglycone residue.

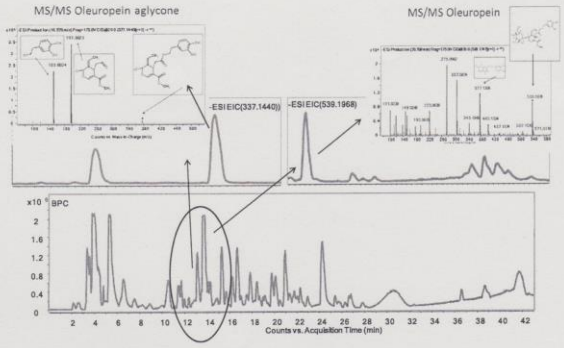


Figure 2. Extracted ion chromatograms for m/z corresponding to oleuropein and its aglycone form after analysis of olive pomace extract by LC-QqTOF. MS/MS spectra enabled to confirm the presence of these compounds.

Other families of compounds such as flavonoids (luteolin and apigenin), phenolic acids (such as caffeic and ferulic acids) and pinoresinols were also extensively studied.

Attending to these results, a complete characterization of the phenolic fraction present in olive pomace has been carried out by combination of an exhaustive sample preparation technique such as superheated liquid extraction and a high resolution technique such as liquid chromatography coupled to tandem mass spectrometry. The high accuracy of the hybrid QqTOF analyzer was crucial to attain confidence levels in the identification step. This characterization study increases considerably the possibilities of this unexploited raw material that can be a valuable source of phenolic compounds with high antioxidant potential.

References

1. Aragón, J.M. and Palancar, M.C. (2001) *Present and Future of Alpeorujo*, Editorial Complutense, S.A.
2. Ouni, Y., et al. *Food Chemistry*, **127**(3), 1263-1267 (2011)
3. Omar, S.H., *Scientia Pharmaceutica*, **78**(2), 133-154 (2010)

International Congress

on "Green Extraction of Natural Products"

Université d'Avignon et des Pays de Vaucluse, 16-17 April 2013

International Congress on
Green Extraction of Natural Products - GENP
Avignon, France, April 16 – 17, 2013.

Optimization of Superheated Liquid Extraction for the Characterization of Phenolic Compounds in Olive Pomace by LC–QqTOF MS/MS

A. Peralbo-Molina, F. Priego-Capote, M. D. Luque de Castro*

[1]Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, 14071, University of Córdoba, Spain
[2]Maimónides Institute of Biomedical Research (IMIBIC), Reina Sofía University Hospital, University of Córdoba, E-14071, Córdoba, Spain

The Mediterranean agrifood industry has been dominated by olive-tree cultivars and olive-oil production. The traditional one cultivation-one manufactured product characteristic in this case is experiencing a change because olive leaves have demonstrated to be a source of compounds of enormous healthy properties. In addition, the residue from oil production (i.e. olive pomace —a semisolid residue from the two-phase virgin olive oil production) is being converted from a nil-priced material with high-contamination charge in a source of compounds even of higher interest than leaves. A number of industries, among them food, cosmetics or pharmaceutical industries are taking profit of the compounds from this raw material because their healthy properties as antioxidants and colorants, among others.

The aim of this research was to study the composition of extracts obtained from olive pomace for the potential use of their components either as food additives or nutraceuticals using a green extraction techniques, such as that assisted by high temperature + pressure, which have been compared with the conventional extraction maceration to show the advantages provided by this auxiliary energies (lower extractant consumption and shorter ime). Once defined the potential of the residues, characterization of repre milies of these compounds was carried out by tandem mass Identification and confirmation were supported on mass accuracy of precursor ion and representative product ions the structure of which was elucidated, but also on isotopic distribution of signals detected in full scan mode. This characterization study increases the possibilities of this unexploited raw material that can be a valuable source of compounds characteristics of the unsaponifiable fraction of virgin olive oil, endowed with high antioxidant potential.

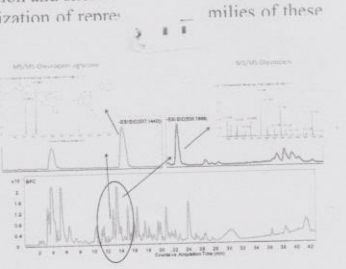


Figure 1. Extracted ion chromatograms for m/z corresponding to oleuropein and its aglycone form after analysis of olive pomace extract by LC-QqTOF. MS/MS spectra have enabled to confirm the presence of these compounds.

XVIII REUNIÓN DE LA SOCIEDAD ESPAÑOLA DE QUÍMICA ANALÍTICA

SEQA 2013

VI REUNIÓN DE LA SOCIEDAD ESPAÑOLA DE ESPECTROMETRÍA DE MASAS

Úbeda, 16-19 junio 2013

SEQA

S.E.E.M

ALI-21

OPTIMIZATION OF A PROTOCOL TO OBTAIN SUPERHEATED LIQUID EXTRACTS ENRICHED WITH PREFERRED PHENOLS FROM OLIVE POMACE

Peralbo-Molina, Ángela, Priego-Capote, Feliciano, Luque de Castro, María Dolores

[1]Department of Analytical Chemistry, Annex Marie Curie Building. Campus of Rabanales, University of Córdoba, Córdoba, Spain.
[2]University of Córdoba Agroalimentary Excellence Campus, ceiA3.
[3]Maimónides Institute of Biomedical Research (IMIBIC), Reina Sofía University Hospital,
qa1lucam@uco.es

The Mediterranean agrifood industry has been dominated by olive-tree cultivars and olive-oil production. The traditional one cultivation-one manufactured product characteristic in this case is experiencing a change because olive leaves have demonstrated to be a source of compounds of enormous healthy properties. In addition, the residue from oil production (i.e. olive pomace or alperujo —a semisolid residue from the two-phase virgin olive oil production) is being converted from a nil-priced material with high-contamination charge into a source of compounds even of higher interest than leaves. A number of industries —among them food, cosmetics or pharmaceutical industries— are taking profit from the compounds from this raw material because of their healthy properties as antioxidants and colorants, among others.

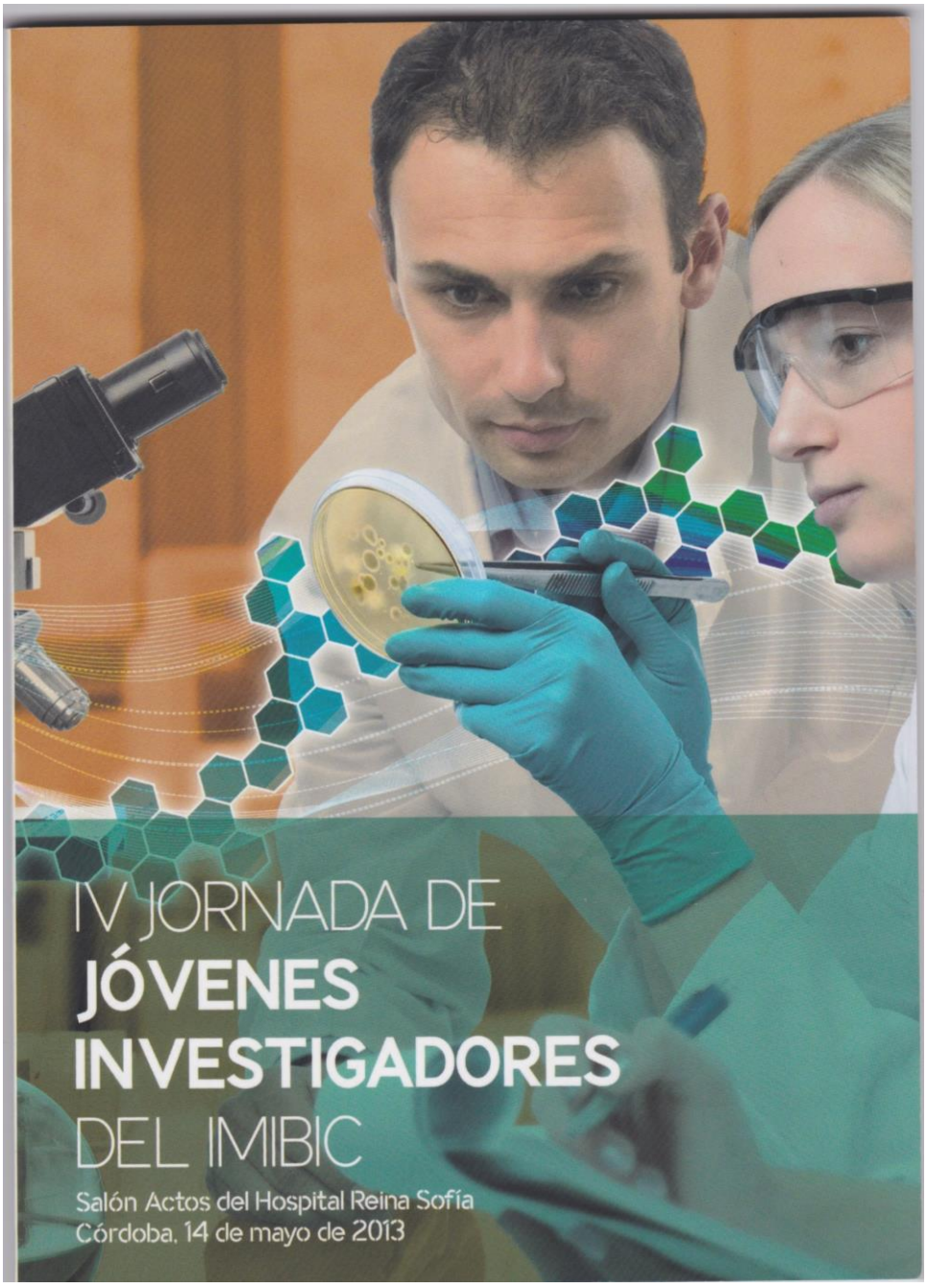
Olive pomace was subjected to superheated liquid extraction (SHLE) with the aim of studying the influence of the main variables involved on the extraction of phenolic compounds (viz. temperature, pH and extractant composition). The extracts obtained in this optimization study were monitored by the Folin–Ciocalteu test in terms of total phenol concentration to evaluate the extraction efficiency. Apart from that, the composition of the extracts was profiled by LC–QqTOF MS/MS in high resolution mode. The data thus obtained were compared by PCA and ANOVA, which showed the extractant composition and temperature as the critical variables for isolation of phenols. The highest overall extraction efficiency was obtained by 75:25 (v/v) ethanol–water at 160 ºC.

After extraction optimization, characterization of compounds of the different families was carried out by tandem mass spectrometry in high accuracy mode. Identification and confirmation were supported on mass accuracy of the precursor ion and representative product ions the structure of which was elucidated, but also on isotopic distribution of signals detected in full scan mode.

A subsequent optimization was aimed at achieving extracts enriched in panels of given phenols with recognized nutraceutical properties. In this way, alperujo can be used to provide extracts of preselected families of valuable compounds characteristic of the unsaponifiable fraction of virgin olive oil.

References

[1] Aragón, J.M. and Palancar, M.C. (2001) Present and Future of Alpeorujo, Editorial Complutense, S.A.
[2] Ouni, Y., et al. Food Chemistry, 127(3), 1263-1267 (2011)
[3] Omar, S.H., Scientia Pharmaceutica, 78(2), 133-154 (2010)


IV JORNADA DE
JÓVENES
INVESTIGADORES
DEL IMIBIC
Salón Actos del Hospital Reina Sofía
Córdoba, 14 de mayo de 2013

05 SAMPLE PREPARATION AND UNTARGETED METABOLOMICS ANALYSIS OF EXHALED BREATH CONDENSATE

Author/es: Ángela Peralbo-Molina, Mª Auxiliadora Fernández-Peralbo, José Ruiz-Jiménez, Feliciano Priego-Capote, Mª Dolores Luque de Castro, Bernabé Jurado-Gámez
IMIBIC Group: D04. Metabolómica. Identificación de componentes bioactivos.
Poster

Exhaled breath condensate (EBC) has been widely used for analysis of inflammatory and oxidative stress markers in humans and it enables the study of the early effects of different diseases or exposures on the lung and upper airways. EBC is usually sampled by cooling expired air using a condenser (–20 °C) and it is composed by volatile compounds and liquid aerosol droplets. The latter are constituted by water and a mixture of semivolatile and nonvolatile compounds.

Gas chromatography (GC) and liquid chromatography (LC) coupled to mass spectrometry (MS) were used for analysis of EBC samples. In this study, the former was selected for untargeted metabolomics analysis. Due to the low concentration of the analytes in the samples and the incompatibility between the sample matrix and the selected analytical technique, sample preparation must be considered as the key step of the analytical process. Two sample separation techniques such as liquid–liquid extraction and solid-phase extraction were evaluated, operating under optimal conditions, in order to find the most adequate for sample preparation. In addition, as volatility of the target analytes was required for analysis, most of the semivolatile and nonvolatile compounds were transformed into volatile derivatives via derivatization reactions such as silylation, methylation or acylation. Identification was achieved by comparing the GC retention indices and the mass spectra provided by the detector with the NIST-MS and the Golm metabolome database reference libraries. The reliability of the identified compounds was ensured by the use of spectral match (match ≥ 875) and retention index values (I-difference ≤ 50).

The developed methodology was applied to the analysis of EBC samples collected from healthy volunteers. The number of identified compounds in the samples was higher than 450. Differences in the composition of the analyzed samples due to the sample preparation were evaluated using statistical tools.

Keywords: Exhaled breath condensated, Untargeted metabolomics analysis, Sample preparation, Gas Chromatography- Mass Spectrometry.

NOTES:

Abreviaturas

ACN, acetonitrilo

ANOVA, análisis de varianza

APCI, ionización química a presión atmosférica

ARNm, ARN mensajero

AUC, área bajo la curva

BMI, índice de masa corporal

BPC, cromatograma pico de base

BSTFA, bis-(trimetilsilil)-fluoroacetamida

CE, electroforesis capilar

CI, ionización química

DAD, detector de red de diodos

EBC, aire exhalado condensado o aliento

EFSA, Autoridad Europea de Seguridad Alimentaria

EI, ionización por impacto electrónico

EIC, cromatograma de iones extraídos

ESI, ionización por electrospray

F-C, Folin-Ciocalteu

FDA, Administración de Alimentos y Medicamentos

FWHM, anchura total a la mitad del máximo

GC, cromatografía de gases

HMDB, Base de datos del metaboloma humano

HPLC, cromatografía líquida de alta resolución

IS, estándar interno

IR, infrarrojos

IT, trampa de iones

KEGG, Kyoto Enciclopedia de Genes and Genomas

LC, cromatografía de líquidos

LLE, extracción líquido-líquido

NMR, resonancia magnética nuclear

MAE, extracción asistida por microondas

METLIN, Metabolites and Tandem MS Database

MetOH, metanol

MF, entidad molecular

MPP, mass profiler professional

MS, espectrometría de masas

m/z, relación masa carga

NIST, National Institute of Standards and Technology

pAUC, área parcial bajo la curva

PCA, análisis de componentes principales

PLS-DA, análisis de mínimos cuadrados parcialmente discriminante

PUFA, ácidos grados poliinsaturados

QqQ, triple cuadrupolo

QTOF, cuadrupolo tiempo de vuelo

ROC, receiver-operating characteristic

RRLC, cromatografía de resolución rápida

RT, tiempo de retención

SD, desviación estándar

SHLE, extracción con líquidos sobrecalentados

SPE, extracción en fase sólida

SPME, microextracción en fase sólida

SRM, monitorización de reacciones seleccionadas

TIC, corriente total de iones

TMCS, trimetilclorosilano

TOF, tiempo de vuelo

UAE, extracción asistida por ultrasonidos

UPLC, cromatografía líquida de ultra alta resolución

VOC, compuestos orgánicos volátiles

IMIBIC
INSTITUTO MAIMÓNIDES DE INVESTIGACIÓN BIOMÉDICA DE CÓRDOBA

UNIVERSIDAD D CÓRDOBA