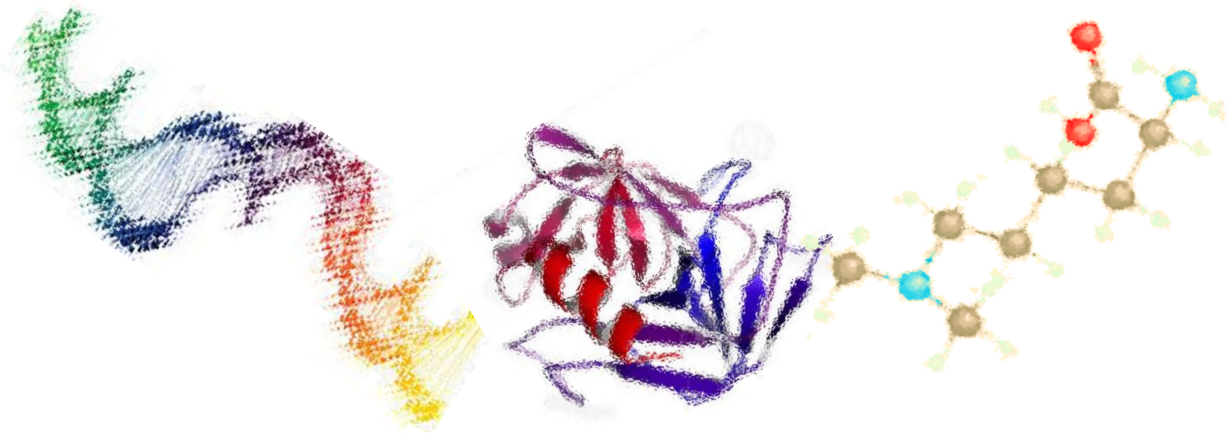
Universidad de Córdoba
FACULTAD DE CIENCIAS
Departamento de Química Analítica

METABOLOMICS CONTRIBUTIONS TO TARGETED AND UNTARGETED CLINICAL ANALYSIS BY CHROMATOGRAPHY AND MASS SPECTROMETRY

APORTACIONES DE LA METABOLÓMICA AL ANÁLISIS CLÍNICO ORIENTADO Y GLOBAL MEDIANTE CROMATOGRAFÍA Y ESPECTROMETRÍA DE MASAS



María Auxiliadora Fernández Peralbo

Tesis Doctoral

Córdoba 2017

TITULO: *Aportaciones de la metobolómica al análisis clínico orientado y global mediante cromatografía y espectrometría de masas*

AUTOR: *María Auxiliadora Fernández Peralbo*

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

APORTACIONES DE LA METABOLÓMICA AL ANÁLISIS CLÍNICO ORIENTADO Y GLOBAL MEDIANTE CROMATOGRAFÍA Y ESPECTROMETRÍA DE MASAS

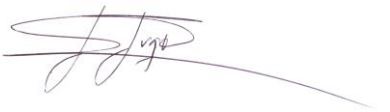
METABOLOMICS CONTRIBUTIONS TO TARGETED AND UNTARGETED CLINICAL ANALYSIS BY CHROMATOGRAPHY AND MASS SPECTROMETRY

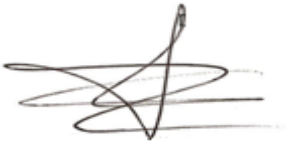
María Auxiliadora Fernández Peralbo

Córdoba, 2017

APORTACIONES DE LA METABOLÓMICA AL ANÁLISIS CLÍNICO ORIENTADO Y GLOBAL MEDIANTE CROMATOGRAFÍA Y ESPECTROMETRÍA DE MASAS

Los Directores,





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

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

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

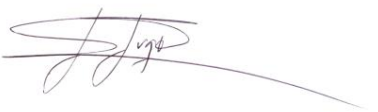
Fdo. María Auxiliadora Fernández Peralbo
Licenciada en Química

María Dolores Luque de Castro, Catedrática Emérita, 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 María Auxiliadora Fernández Peralbo, con el título "**Aportaciones de la metabolómica al análisis clínico orientado y global mediante 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, a 30 de mayo de 2017.



Fdo. María Dolores Luque de Castro

Fdo. Feliciano Priego Capote

Mediante la defensa de esta Memoria se pretende optar a la mención de **Doctorado Internacional**, habida cuenta de que la doctoranda reúne los requisitos exigidos para tal mención, a saber:

1. Informes favorables de dos doctores pertenecientes a Instituciones de Enseñanza Superior de otros países:
   - Prof. Dra. Marta Sousa Silva, Chemistry and Biochemistry Center, University of Lisbon, Portugal.
   - Prof. Dr. José González Rodríguez, School of Chemistry, College of Science, University of Lincoln, UK
2. Uno de los miembros del tribunal que ha de evaluar la Tesis pertenece a otro centro de Enseñanza Superior de otro país:
   - Prof. Dr. Carlos Cordeiro, Department of Chemistry and Biology, University of Lisbon, Portugal.
3. La exposición y defensa de parte de esta Tesis se realizarán en una lengua diferente a la materna: inglés.
4. Estancia de tres meses en un centro de investigación de otro país:
   - Center for Proteomics and Metabolomics of the Leiden University Medical Center (LUMC), Leiden, Países Bajos, bajo la supervisión del Prof. Dr. Oleg Mayboroda.

TÍTULO DE LA TESIS: Aportaciones de la metabolómica al análisis clínico orientado y global mediante cromatografía y espectrometría de masas

DOCTORANDA: María Auxiliadora Fernández Peralbo

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 María Auxiliadora Fernández Peralbo manifestó desde su inserción en el grupo de investigación del que forma parte un especial interés por el área clínica que, junto a la metabolómica ha constituido el eje fundamental de la investigación que ha realizado y que constituye su Tesis Doctoral.

Los tres aspectos que ha abarcado su Tesis enmarcada, por tanto, en la metabolómica clínica han sido estudios bibliográficos para conocer en profundidad el uso en metabolómica de muestras clínicas poco utilizadas en esta disciplina (Sección I) y una parte experimental que ha abarcado las dos estrategias más comunes en metabolómica: el análisis orientado (Sección II) y el global (Sección III).

(i) Los estudios bibliográficos, amplios y críticos, se centraron en el pretratamiento, preparación y uso de dos tipos de muestras clínicas: orina y condensado de aire exhalado. La orina es una de las muestras más utilizadas en metabolómica clínica y nutricional por su obtención de forma no invasiva y la información que puede aportar, reflejo directo de la parte final del metabolismo. Por otro lado, el condensado de aire exhalado es una muestra que había recibido hasta la fecha un interés en metabolómica que no correspondía a las posibilidades que podían deducirse de sus características: obtención no invasiva y matriz no compleja.

(ii) La investigación en análisis metabolómico orientado ha abarcado un estudio sobre el uso de estándares marcados con isótopos estables (deuterio en este caso), de enorme interés para asegurar una información cuantitativa de calidad cuando se utiliza la espectrometría de masas, y su aplicación al análisis de eicosanoides derivados de los ácidos omega-6 en suero de pacientes afectados por enfermedades coronarias. Un método para el análisis cuantitativo de paclitaxel y sus principales metabolitos en suero, plasma y tejido de enfermas de cáncer de ovario y su aplicación completa esta sección.

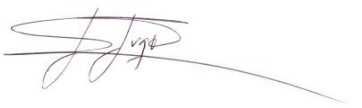
(iii) Un estudio para la simplificación del tratamiento de los datos provenientes de los modos de ionización positiva y negativa evitando duplicidad y mejorando los resultados finales encabeza la Sección C, al que sigue el desarrollo de un método de preparación de la muestra de aire exhalado condensado, con identificación tentativa de 49 de los metabolitos que lo componen. La última parte de esta sección está dedicada a la búsqueda en orina de biomarcadores de cáncer de próstata: dos capítulos que ponen de manifiesto las posibilidades de la metabolómica para proporcionar biomarcadores con excelente sensibilidad y especificidad.

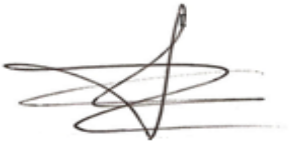
Consideramos que la investigación desarrollada por María Auxiliadora Fernández Peralbo y recogida en la Memoria reúne la originalidad, innovación y calidad requeridas.

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

Córdoba, 29 de mayo de 2017

Firma de los directores





Fdo.: M. D. Luque de Castro

Fdo.: F. Priego Capote

INFORME SOBRE EL FACTOR DE IMPACTO DE LAS PUBLICACIONES DE LA TESIS

TÍTULO DE LA TESIS: APORTACIONES DE LA METABOLÓMICA AL ANÁLISIS CLÍNICO ORIENTADO Y GLOBAL MEDIANTE CROMATOGRAFÍA Y ESPECTROMETRÍA DE MASAS

DOCTORANDA: MARÍA AUXILIADORA FENANDEZ PERALBO

PUBLICACIÓN	FI	DECIL/CUARTIL
Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. **TrAC**, 2012, 41, 75–85.	6.351	D1 2/75 Analytical Chemistry
Analytical methods based on exhaled breath for early detection of lung cancer. **TrAC**, 2012, 38, 13–20.	6.351	D1 2/75 Analytical Chemistry
Stable isotopic internal standard correction for quantitative analysis of hydroxyeicosatetraenoic acids (HETEs) in serum by on-line SPE–LC–MS/MS in selected reaction monitoring mode. **Talanta**, 2014, 126, 170–176.	3.545	Q1 11/74 Analytical Chemistry
Targeted analysis of omega-6-derived eicosanoids in human serum by SPE-LC-MS/MS for evaluation of coronary artery disease. **Electrophoresis,** 2013, 34 (19), 2901–2909.	3.161	Q1 16/76 Analytical Chemistry

LC–MS/MS quantitative analysis of paclitaxel and its major metabolites in serum, plasma and tissue from women with ovarian cancer after intraperitoneal chemotherapy. **Journal of Pharmaceutical and Biomedical Analysis,** 2014, 91, 131–137.	2.979	Q1 16/74 Analytical Chemistry
MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes. **Metabolomics,** 2016, 12 (3), 1–12.	3.661	Q1 40/133 Endocrinology and Metabolism
Study of exhaled breath condensate sample preparation for metabolomics analysis by LC–MS/MS in high resolution mode. **Talanta,** 2015, 144, 1360–1369.	4.035	Q1 9/75 Analytical Chemistry
Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC-QTOF: upstream information on other omics. **Scientific Reports,** 2016, 6, 38243	5.282	Q1 7/62 Multidisciplinary Sciences
Urinary metabolite panels for prostate cancer detection. Enviado a **Cancer letters**	5.99	Q1 23/213 Oncology

ÍNDICE

INDEX

OBJETIVOS

El **objetivo global** de la Tesis Doctoral fue realizar contribuciones en el ámbito del análisis clínico a través de estrategias metabolómicas tanto orientadas como globales en diferentes tipos de muestras biológicas mediante técnicas cromatográficas y espectrometría de masas. Estas técnicas han sido ampliamente utilizadas en metabolómica por su gran versatilidad y excelentes niveles de sensibilidad, selectividad, precisión, exactitud y resolución. Para la consecución de este objetivo global se contó con equipos analíticos de última generación con los que diseñar metodologías adecuadas para cada objetivo concreto: sistemas automáticos de extracción en fase sólida (SPE), cromatógrafos de líquidos (LC) acoplados a analizadores de masas (MS) de triple cuadrupolo (QqQ) para análisis cuantitativo, o híbridos de cuadrupolo–tiempo de vuelo (QTOF) para análisis cualitativo o semicuantitativo.

De este objetivo global derivaron los siguientes **objetivos genéricos** cuyo desarrollo ha dado lugar a los tres bloques en los que se ha dividido esta Memoria de Tesis: objetivo 1, revisar exhaustiva y críticamente la bibliografía para conocer el estado del tratamiento de muestra y de los métodos de análisis de dos biofluidos utilizados para el desarrollo de esta Tesis. Objetivo 2, optimizar métodos de análisis orientado para mejorar la cuantificación tanto de compuestos con potencial biomarcador como de fármacos y sus metabolitos para su aplicación en el diagnóstico y seguimiento de enfermedades o tratamientos. Objetivo 3, analizar de forma global biofluidos para (a) estudiar y optimizar el tratamiento de una muestra escasamente utilizada en el área clínica (el aire exhalado condensado – EBC), (b) identificar metabolitos con potencial predictivo para ayudar en el diagnóstico del cáncer de próstata utilizando la orina como muestra, y (c) mejorar y acelerar el tratamiento de datos a través de herramientas quimiométricas desarrolladas para combinar en una única matriz los datos obtenidos mediante ionización positiva y negativa en espectrometría de masas.

De estos tres objetivos generales derivaron los siguientes objetivos concretos que se recogen en los diferentes capítulos que componen esta Memoria de Tesis:

- Conocer y discutir de forma crítica la situación actual del tratamiento de la orina como muestra analítica, previo al análisis tanto global como orientado mediante LC acoplada a MS en modo de alta resolución. La orina como biofluido constituye una de las muestras de mayor interés en investigación clínica debido a su fácil y no invasivo muestreo, además de su composición que refleja el catabolismo del individuo. Una amplia visión sobre este tema dio lugar a una revisión que constituye el Capítulo 1 de esta Tesis Doctoral.
- Revisar los métodos existentes en la literatura para la detección temprana del cáncer de pulmón utilizando aire exhalado, con especial énfasis en aquéllos que utilizan el olfato de perros entrenados y en los que usan diferentes estrategias ómicas, ha dado como resultado la publicación que recoge el Capítulo 2.

Los objetivos concretos en investigación experimental se dividieron en función de la estrategia analítica utilizada. Así, el objetivo general 2, relacionado con el análisis orientado recoge los siguientes **objetivos concretos**:

- Evaluar la influencia de un estándar interno marcado isotópicamente al incluirlo en el análisis cuantitativo de ácidos hidroxieicosatetraenoicos en suero. El estudio del efecto matriz a tres concentraciones diferentes y de cómo el estándar interno corrige este efecto dio como resultado el Capítulo 3.
- Desarrollar una plataforma analítica para la cuantificación de compuestos omega 6 derivados de eicosanoides acoplando en línea la etapa de SPE con la separación cromatográfica y la detección por QqQ fue otro de los objetivos cumplidos (Capítulo 4). La aplicación de esta plataforma a una cohorte formada por pacientes afectados por enfermedad arterial coronaria y el tratamiento estadístico de los niveles de estos metabolitos en suero para establecer una gradación de diferentes

lesiones coronarias como la angina estable, el síndrome coronario agudo sin elevación del segmento T y el infarto agudo de miocardio, fue otro de los objetivos que se recogen en este capítulo.

- Diseñar y aplicar un método para la cuantificación del fármaco de quimioterapia paclitaxel y de sus principales metabolitos en tejido, suero y plasma procedente de mujeres con cáncer de ovario tras un tratamiento con quimioterapia intraperitoneal fue un objetivo cumplido que recoge el Capítulo 5.

El objetivo general 3, sobre análisis metabolómico global, comprende los siguientes objetivos concretos:

- Desarrollar una herramienta para la combinación de matrices de datos adquiridos mediante MS en los modos de ionización positivo y negativo, mejorando y acelerando de esta manera el tratamiento de datos en el análisis metabolómico no orientado fue un objetivo cuyo desarrollo constituye el Capítulo 6.

- Estudiar diferentes métodos de preparación de muestra para el análisis global de EBC mediante LC–QTOF dio lugar al desarrollo y aplicación del método que se recoge en el Capítulo 7. Este método puede ser de interés en el estudio de enfermedades respiratorias.

- Buscar compuestos significativos en la orina que permitan discriminar pacientes con cáncer de próstata avanzado de individuos con biopsia negativa y dilucidar las rutas metabólicas alteradas como consecuencia de este tipo de cáncer fue el estudio que dio lugar al Capítulo 8.

- Configurar paneles que integren metabolitos con potencial biomarcador para el diagnóstico de cáncer de próstata basados en la discriminación de pacientes afectados por esta enfermedad e individuos con biopsia negativa usados como control fue el último de los objetivos concretos. La evaluación estadística de los paneles configurados permite seleccionar aquéllos con mayor potencial priorizando la sensibilidad, la especificidad

o la combinación de ambos parámetros. Esta investigación constituye el Capítulo 9 de la Tesis.

La formación de la futura doctora, **objetivo último** de una tesis doctoral, ha incluido el máster en "Química Fina" con el número de créditos correspondientes, así como todas las etapas implicadas en el cumplimiento de los requisitos exigidos para optar a la mención de Doctorado Internacional. Paralelamente a la investigación recogida en la parte principal de la Memoria, se ha pretendido una formación más amplia de la doctoranda con la realización de otras actividades que se recogen como anexos, tales como:

Anexo I: Investigación relacionada con la Tesis, realizada en colaboración con la Unidad de Cirugía Oncológica y de Páncreas, que ha dado lugar a la preparación de un artículo –enviado para su publicación al International Journal of Hyperthermia– en el que se recoge la aplicación del método desarrollado en el Capítulo 5 a una cohorte de pacientes de cáncer de ovario.

Anexo II: Capítulos de libro sobre temas no relacionados con la Tesis, como las aplicaciones de la energía de microondas en las disciplinas ómicas y en la extracción de grasas y aceites.

Anexo III: Solicitud de patente de paneles de biomarcadores en orina para el cribado de pacientes en el diagnóstico de cáncer de próstata.

Anexo IV: Comunicaciones orales y en cartel en 13 conferencias, 11 nacionales y 2 internacionales.

OBJECTIVES

The **overall objective** of this Doctoral Thesis was to contribute to the development of both targeted and untargeted metabolomics strategies in clinical analysis for application to different types of biological samples using chromatography and mass spectrometry techniques. This equipment has been widely used in metabolomics thanks to its versatility and excellent levels of sensitivity and selectivity, precision, accuracy and resolution. The achievement of the overall objective was supported on the use of last-generation analytical equipment as a function of each concrete objective: automated systems for solid-phase extraction (SPE), liquid chromatographs (LC) coupled to mass analyzers (MS) of the triple quadrupole (QqQ) type for quantitative analysis, or hybrid quadrupole time-of-flight (QTOF) for qualitative or semiquantitative analysis.

The following **general objectives** resulted from the overall objective, each of them corresponding to one of the parts into which this Thesis-Book is divided: general objective 1, to revise in an exhaustive and critical way the literature on both sample treatment and analysis of two different human biofluids used to develop the planned research. General objective 2, to optimize targeted analysis methods for quantification of compounds with biomarker potential, and also of drugs and their metabolites with a view on application to diagnose and monitor diseases or treatments. General objective 3, to analyze comprehensively biofluids with the aims of (a) studying and optimizing preparation of a sample scantly used in the clinical area (exhaled breath condensate —EBC), (b) identifying metabolites with predictive potential to help in the diagnostic of prostate cancer using urine as sample, and (c) improving and accelerating data treatment by chemometric tools developed for combination in a unique data set of data obtained by positive and negative ionization modes in mass spectrometry.

From the general objectives derived the following **concrete objectives** each of which given place to one of the different chapters that constitute this Thesis-Book:

-To know and critically discuss the present situation of the treatment of urine as analytical sample, prior to analysis –either global or targeted– by LC coupled to MS in high resolution mode. Urine as biofluid constitutes one of the samples of higher interest in clinical research thanks to its easy and non-invasive sampling, in addition to reflexing the individual catabolism. A comprehensive vision on this subject gave place to a review that constitutes Chapter 1 of this Thesis-Book.

-To revise critically the methods in the literature for the early detection of lung cancer using exhaled breath —with special emphasis on both those that use trained dogs and those based on different omics strategies— gave the matter of Chapter 2.

The concrete objectives on experimental research were divided as a function of the analytical strategy. Thus, the general objective 2, related to targeted analysis, involves the following concrete objectives:

-To evaluate the influence of an isotopic labeled internal standard included in the quantitative analysis of hydroxyeicosatetraenoic acids in serum samples. The study of both the matrix effect at different concentrations and the way the internal standard corrects this effect resulted in Chapter 3.

-To develop an analytical platform to quantitate omega 6 eicosanoid derivatives by on-line coupling the SPE step with the chromatographic separation and detection by QqQ was other of the fulfilled objectives (Chapter 4). Application of this platform to a cohort constituted by patients with coronary disease and the statistical treatment of the levels of these compounds in serum to establish a gradation of the different coronary lesions such as stable angina, acute coronary syndrome without elevation of the T

segment, and acute myocardial infarction was other of the objectives included in this chapter.

-To design and apply a method for quantitation of paclitaxel —a drug for chemotherapy— and its main metabolites in tissue, serum and plasma from patients with ovarian cancer after treatment with intraperitoneal chemotherapy was the fulfilled objective that constitutes Chapter 5.

General objective 3, on untargeted metabolomic analysis, was formed by the following concrete objectives:

-To develop a tool to combine matrices from the data acquired by MS in the positive and negative ionization modes, thus improving and accelerating data treatment in untargeted metabolomics analysis was the objective that gave place to Chapter 6.

-To study different sample preparation methods for subsequent overall analysis of EBC by LC–QTOF resulted in the development and application of the method in Chapter 7. This method can be of great interest in the study of respiratory diseases.

-To search for compounds in urine that allow discriminating patients with aggressive prostate cancer and individuals with negative biopsy and elucidate the metabolic pathways altered as a consequence of this type of cancer was the study shown in Chapter 8.

-To make up panels to integrate metabolites with biomarker potential for prostate cancer diagnostic based on discrimination of patients suffering this disease and individuals with negative biopsy as control was the last of the fulfilled concrete objectives. The statistical evaluation of the panels allows selection of those with higher potential by prioritizing sensitivity, specificity or a combination of both parameters. This research constitutes Chapter 9 of this Thesis-Book.

The formation of the future PhD, which is the **final objective** of a Doctoral Thesis, has also included the master on "Fine Chemistry" , in which the PhD student developed the mandatory courses. In parallel to the above mentioned tasks and to the research in the main part of the Thesis-Book, a wider formation of the PhD student has been sought by development of other activities summarized below as annexes:

-Annex I: Research related to the Thesis, developed as a collaboration with the Oncologic and Pancreatic Surgery Unit of the University Hospital Reina Sofía, resulted in a manuscript sent for potential publication to the International Journal of Hyperthermia, which contains the application of the method developed in Chapter 5 to a cohort of ovarian cancer patients.

-Annex II: Book chapters on subjects non-related with the Thesis matter, as are the application of microwave energy in omics disciplines and in the extraction of fats and oils.

-Annex III: A patent on panels of biomarkers in urine for patients screening in the diagnostic of prostate cancer.

-Annex IV: Oral and poster communications in 13 meetings (11 national and 2 international).

INTRODUCCIÓN

INTRODUCTION

What is intended to show in this introduction is the knowledge of the PhD student on the material and nonmaterial tools she has managed for development of the research mandatory to attain her PhD degree: (i) the state-of-the-art of metabolomics as the omics on which is supported all the research presented in this Thesis-Book, without forgetting its relationship with the other great omics, very useful in the developed research. (ii) The mass spectrometry (MS) instruments, with special emphasis on those she has managed and the each time more indispensable use of stable isotopic labeled internal standards (SIL-ISs). (iii) The chemometric approaches mandatory for treatment of the data provided by MS equipment. The gaps existing in these approaches have been in some case overcome by the research in this Thesis-Book.

Metabolomics

1.1. Definition of terms

The late twentieth century, more than three decades ago, was notorious for revolutionizing the life sciences with the emergence and subsequent expansion of the terminology "-omics" which derived from the Latin suffix “ome” meaning mass or many.

Omics studies involve the measurement of large numbers of parameters, typically genes (genomics), mRNA (transcriptomics), proteins (proteomics), or metabolites (metabolomics), resulting in the primary omics. The measure of specific families such as lipids, glycans, toxics, foreign substances, or even interactions among them, provides the secondary omics (lipidomics, glycomics, toxicomics, xenometabolomics, and interactomics, respectively), all them as parts of primary omics.

The term metabolome and the derived omics were not used until 2002, when Fiehn defined metabolomics as a global and comprehensive analysis in which all metabolites of a biological system are identified and quantified [1]. However, the term metabolome first appeared in the literature in 1998, when Oliver *et al.* [2] measured the change in the relative concentrations of metabolites as the result of deletion or overexpression of a gene.

The terms "metabolomics" and "metabonomics" were coined almost simultaneously to define the same discipline, but in the plant and clinical/pharmaceutical areas, respectively [3,4]. Moreover, the authors working on this omics consider that the goal of metabolomics is holistic quantitative and qualitative analysis of all metabolites in a given biological system, whereas the goal of metabonomics is the comparison of metabolite levels in a given system in response to a particular stimulus or treatment [5]. Although some differences in the concept remain, there is now an overlap in methodologies and the two terms have been used interchangeably by scientists [6-9] who are at present more fond of metabolomics. In the last decade, other ways have appeared in the literature to define and to classify metabolomic studies. Beyoğlu and Idle (2013) combined all of them in a broad definition of the discipline as "the global and unbiased study of small molecules (<1 kDa) in a biofluid, tissue, organ or organism" [10].

The information provided by metabolomics would lead to the complete metabolome, which is defined as the quantitative complement of low-molecular-weight metabolites (thus excluding proteins) present in a biological fluid, cell or organism under a given set of physiological conditions or under different perturbations (*e.g.*, genetic variations, pathological states or responses to external stimuli) [11]. The metabolome, and therefore the metabolic state of a living organism, can be affected by various intrinsic factors (age, health status, reproductive status, etc.) and extrinsic (nutrients, artificial compounds such as pesticides, fertilizers, environmental pollutants, etc.). Therefore, the study of the metabolome is of great importance to know the state of an organism and how it is affected by the environment. In the case of cells and tissues, the metabolome comprises the endometabolome (all intracellular metabolites) and exometa-

bolome (all metabolites that are excreted to the growth medium or extracellular fluid).

On the other hand, metabolism is an extensive network of metabolic reactions where the products of one reaction are the reactants of a subsequent reaction. Therefore, a metabolite is defined as "any intermediate or end product of metabolism, usually restricted to small molecules that are not genetically encoded". Depending on the metabolic pathways in which the metabolites are involved and their function, they can be classified into primary metabolites (those that are directly involved in the growth, development and reproduction of the organism) or secondary metabolites (those not directly involved in the processes of growth, development and reproduction, but generally they have a key biological function as, for example, defense against predators, parasites or diseases, competition among species, facilitating reproduction processes or in cellular signaling mechanisms).

In the great diversity of chemical structures that comprise the metabolome, there are endogenous and exogenous metabolites; the former (including amino acids, organic acids, nucleic acids, fatty acids, sugars, vitamins, cofactors, pigments, antibiotics, etc.) are naturally produced by an organism, and the latter (such as drugs, environmental toxins, food additives, and other xenobiotics) come from the interaction with the outside.

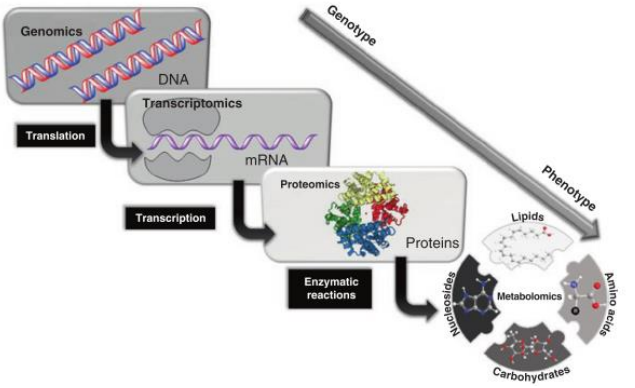
The present Thesis has focused on metabolomics from a clinical perspective. The clinical approach to metabolomics pursues either finding diagnostic markers for a disease or a response to a treatment, as well as understanding the biochemical basis of such situations. Perhaps one of the most suggestive aspects of metabolomics from a clinical point of view is the potential identification of new biomarkers that contribute to overcome the current diagnostic tests, surpassing them in selectivity, specificity and even allowing to distinguish phases of a disease or its earlier detection.

1.2. Relationship of metabolomics with other great omics and secondary omics

The structure and dynamics of cells and organisms functions, rather than the characteristics of isolated parts of them, must be examined to understand biology at the system level. The integration of the data from the different omics constitutes what is known as systems biology, term coined by Nicholson and Wilson [12] to better understand the functioning of a given biological system. At present, systems biology is more frequently known as integrative omics [13].

Figure 1 schematizes both the traditional central dogma of molecular biology, where the flow of information goes from genes to transcripts to proteins, finishing with the site where enzymes act on metabolism (Figure 1.A), and the omics organization (Figure 1.B), where the flow of biological information ends in the metabolites and is finally reflected in the phenotype.

A)



B)

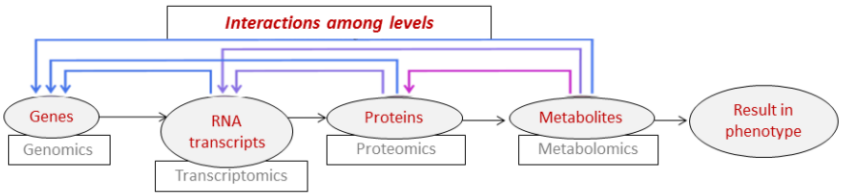


Figure 1. A) Traditional central dogma of molecular biology. B) General scheme of systems biology.

A fundamental aspect that differentiates systems biology from the central dogma of molecular biology is that the latter is governed by unidirectionality in the expression of information from the genes of a cell or organism (*viz.*, DNA) to its transcription at mRNA and its translation to proteins that catalyze metabolic reactions, as shows Figure 1.A. Systems biology breaks this traditional view by revealing the interaction of the different levels, the flow of information that takes place not only downstream but also upstream and explains the phenotype of an individual as the direct reflection of its metabolism. Also, lower levels can provide information of higher levels that justify the situation of the former, as outlined in Figure 1.B.

The metabolome is the level of the functional cascade that best reflects the physiological state, being not only the most sensitive to any change, either internal or external, but its ultimate expression, since the metabolites are the real active regulating agents of homeostasis [14]. Hence, metabolomics is one of the most powerful bioanalytical strategies as it allows obtaining a picture of the metabolites of an organism in the course of a biological process, being considered as a phenotyping tool, thus justifying its usefulness.

In contrast to proteins or genes, which are activated by post-translational or epigenetic modifications, respectively, the metabolites result from cellular activity and serve as direct indicators of the biochemical activity of the cell in each time. In fact, physiological changes resulting from gene deletion or overexpression are amplified through the hierarchy of the transcriptome and proteome and are, therefore, more easily measurable through the metabolome, even when changes in metabolic fluxes are negligible [15]. Thus, small changes in the expression level of individual proteins have large effects on the concentrations of intermediary metabolites. As a result, the metabolome is more sensitive to overall perturbations than the transcriptome or the proteome, because the activities of metabolic pathways are reflected more directly in the concentrations of metabolites than in those of the relevant enzymes or, indeed, those of their encoding mRNAs [16].

In brief, metabolomics has been developed into a more integrative approach than other omics approaches. Thus, it is not surprising that metabolic studies are increasingly being used in clinical diagnosis since the metabolic profile closely reflect the total cellular state. Serve as a measure of the relevance of the metabolomics a quote from James Watson, one of the discoverers of the DNA structure, who, in a recent interview, said: "If I had to do a doctorate right now, I would do it in metabolomics" [17]. Thus, the use of metabolomics has grown exponentially in the last decade. It has been the omics on which the research that has resulted in this Thesis is supported.

Nowadays, metabolomics is a mature discipline with a status similar to that of other major omics due to the relevance of the metabolome in the biological context. Figure 2 shows, in a comparative way, the evolution of the publications on the great omics until early 2017.

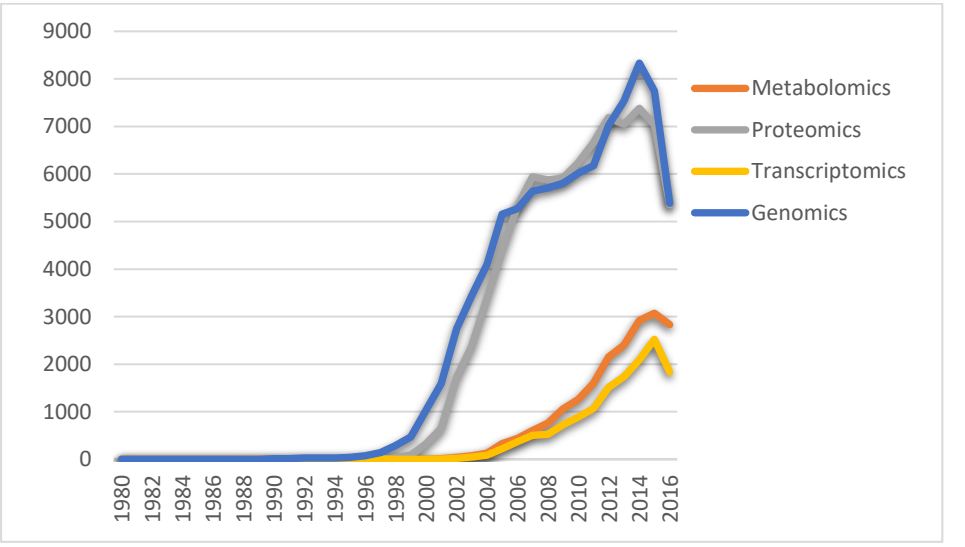


Figure 2. Evolution of the number of publications on the great omics since 1980 to 2017 (data collected from Scopus database 20/02/2017)

Despite the great interest shared by many branches of biology and biotechnology on this new omics science, it must be admitted that metabolomics has not evolved as rapidly as genomics and proteomics. This is due to several aspects

that are determined by the physico-chemical characteristics of the small organic molecules. Unlike genes, mRNAs and proteins, all of which are bio-polymers that encode information from a sequence of well-known monomers (or residues) — *e.g.*, nucleotides and amino acids—, the metabolites are chemical entities that do not originate from a transfer of information between residues within the cell. The great success in the characterization of genes, mRNA and proteins is a direct consequence of technologies and bioinformatics tools capable of amplifying and subsequently characterizing the sequence of nucleotides and amino acids, respectively, in these biopolymers. Metabolomics, however, aims to detect, quantify and elucidate the structure of the metabolites, which are characterized by a great physicochemical diversity in their molecular structures [18]. The great diversity of molecules is reflected in a wide range of polarities, molecular weights, functional groups, stability and chemical reactivity, among other key properties [19,20]. This leads inevitably to use multiple platforms and analytical configurations that maximize the coverage of the analyzed metabolome, something that does not occur in genomics or proteomics experiments. However, metabolomics presents positive characteristics that promote its use and establish a difference with other omics, as discussed below:

-The existence of robust and stable analytical platforms for the analysis of a significant number of metabolites, as a result of research in this area during the last two decades of the twentieth century, prior to the appearance of omics as such.

-The analytical platforms used in metabolomics studies that are characterized by excellent analytical and biological reproducibility with a low cost of analysis per sample, as opposed to the costs of any other major omics application. This fact facilitates application to a high number of samples, and extensive studies in cohorts with a significant biological variability; thus providing wider and well-supported information.

-There are fewer metabolites than genes and proteins. According to a recent study, the predictable number of metabolites in humans is only 40,000 (29,000

of them endogenous [21]), while the number of genes, transcripts and proteins can reach 25,000, 300,000 and 1,000,000, respectively [22]. In the case of genomics, the inclusion of gene modifications —the subject of epigenomics— greatly increases the number of components [23]. Therefore, metabolomics studies lead to a reduction in the complexity of the data, thus simplifying the analysis and increasing the probability of detecting significant changes that reflect the biological problem under study [24].

- The technology involved in metabolomics is generic, as a given metabolite —unlike a transcript or protein— is the same in every organism that contains it [25].

- The ability to integrate data obtained with different techniques (integrated metabolomics) [26] applied to samples from different sources is other of its advantages.

- One other benefit is that, although some environmental perturbations or genetic manipulations may not cause changes in transcriptomics and proteomics levels, they will have significant effects on the concentrations of numerous individual metabolites due to the "downstream" character of metabolomics.

Negative or unsolved aspects of metabolomics that make it more complex than other omics are the following:

- The necessity for multiple analytical platforms as a result of the enormous variety of metabolites that ranges from non-polar lipids to ionic compounds; from relatively big non-polar molecules such as cholesterol to very small ions such as ammonium.

- The concentrations at which the metabolites appear in the organisms cover more than nine orders of magnitude (from picomoles to millimoles) and make necessary the application of very different techniques both for sample preparation and for analysis of the metabolites in question, with the researcher's experience playing a key role in selecting the most suitable ones as a function of

both the nature of the metabolites to be determined and that of the sample matrix.

- The absence of enough standardized methods, as a consequence of being the last of the primary omics.

Metabolomics can be applied to almost any type of matrix and can be focused on a multitude of case studies. Thus, the wide thematic diversity in metabolomics has led to a series of subdisciplines that, depending on their extension, can be classified in primary and secondary. Among the former are lipidomics, xenometabolomics, clinical metabolomics and nutrimetabolomics or nutritional metabolomics. Secondary subdisciplines of metabolomics are pharmacometabolomics, environmental metabolomics, cardiometabolomics, secretomics, cosmetobolomics, microbial metabolomics and many others that have emerged and will continue to emerge in view of the importance of the applications of metabolomics in new sectors.

Clinical metabolomics is aimed at evaluating the health status of an individual as well as the risk of disease by any of the different metabolic analysis of biofluids or tissues, which are influenced by their genetic, epigenetic, environmental exposures, diet and behavior.

The metabolome of an organism changes in response to stimuli and environmental conditions. This does not happen with its genome or its transcriptome, which only change when a genetic mutation occurs, being practically constant throughout its life. Diseases affect metabolism and, as a result, produce changes of sufficient duration to be detected by measurements of the metabolites [27]. Examples of this ability include schizophrenia, depression, Alzheimer's, cardiovascular disease, hypertension, type 2 diabetes, liver, ovarian, breast, lung, or Huntington's disease [28]. These diseases are characterized by changes (decrease–increase) in the concentration of dozens of metabolites with respect to homeostasis. The analysis of the metabolite profile can provide information of these pathologies, very useful to combat them.

Pharmacometabolomics [29] and nutrimetabolomics [30] are two of the metabolomic subdisciplines that, together with the study of the microbiome and epigenome, belong to the clinical field constituting one of the basis of the personalized medicine and, therefore, they are of interest within the framework of this Thesis. Personalized medicine seeks to improve tailoring and timing of preventive and therapeutic measures by utilizing biological information and biomarkers at the level of molecular disease pathways, genetics, proteomics as well as metabolomics [31]. The definition of these secondary metabolomics disciplines is below.

Pharmacometabolomics is an emerging field that complements genomics, transcriptomics, proteomics, and epigenomics "systems biology" approaches to drug development and contributes to a comprehensive understanding of drugs effects, predicting their efficacy or toxicity in a given individual by taking into account both intrinsic and extrinsic contributions to interindividual variation in drug response. An example of pharmacometabolomics applications to oncology in the literature is a metabolomic profiling study of changes resulting from the exposure of MCF7 breast cancer cells to docetaxel, that showed varied effects at high and low doses [32]. Although it is not a targeted therapy approach, it is an example of an approach that could be used to stratify patients to receive the appropriate dose of chemotherapy. Thus, the field of pharmacometabolomics, which not only takes into account drug effects but can also use data on drug metabolism, has enormous potential for improving personalized medicine approaches.

Nutrimetabolomics refers to the study of the dose and temporary changes in compounds of low molecular weight in response to dietary treatments, and constitutes, together with nutrigenomics, the basis on which personalized nutrition —which can provide huge benefits to the population— is intended to be supported.

Metabolomics could help in nutrition assessment through three main pathways: dietary biomarker discovery, diet-related diseases exploration, and

dietary intervention assessment. Several studies have used metabolomics tools to identify specific markers of consumption of certain foods, validate nutritional questionnaires, or evaluate in humans extracts with nutraceutical potential [33–35].

Nutrimetabolomics can expedite our ability to identify metabolic diseases that are influenced by nutrients and to develop targeted diet-based treatments. Self-reporting dietary assessment methods such as food frequency questionnaires, recalls, and weighed food records are intrinsically associated with errors, including energy intake under-reporting, recall errors, and portion size assessment approximations. Therefore, to enhance the accuracy of dietary intake assessments, dietary biomarkers based on nutrimetabolomics constitute a promising tool. In acute intervention studies, participants consume specific foods and then biofluids are collected and analyzed to characterize related biomarkers. In cohort studies, high and low consumers of a chosen food are selected; then, a comparison of their metabolic profiles may lead to the discovery of potential biomarkers. Dietary factors alter metabolism via numerous and diverse mechanisms that tightly control tissue-specific responses as well as whole-body metabolism and homeostasis. This regulation often occurs through changes in the transcriptome, proteome, and metabolome. Thus, unbiased systems-based nutrimetabolomics studies can expedite our ability to identify metabolic diseases influenced by nutrients and to develop personalized diet-based treatments.

The growing development of metabolomics studies is expanding our knowledge of the human microbiome metabolic landscape and the impact of both the microbiota and pathogenic bacteria on human health. It is known that the human gut microbiome exhibits an important metabolic activity that has an impact on health and disease. Microbiome-related metabolomics has been used to explore a range of clinical conditions [36]. In this regard, it is generally considered useful to characterize both fecal/gut metabolome and microbiome in combination with the host's systemic metabolome, such as blood, urine, or saliva. Different examples have been reported in the literature with various applications [37]. For instance, Janson *et al.* correlated the fecal metabolome with blood and urine

metabolome to understand better peripheral markers of gut microbial metabolism [38].

Finally, it should be mentioned that metabolomics could help in the understanding of epigenetics effects. Epigenetics mechanisms give place to biochemical changes mainly using small molecules that affect the chromatin structure leading to either gene expression extinction (*e.g.*, methylation) or activation (*e.g.*, acetylation). Thus, metabolomics can track the related small molecule changes defined by the epigenome [39, 40].

1.3. Metabolomics strategies

Metabolomics analysis encompasses different strategies the nature of which depends on the objective of the study and previous knowledge of the biological problem [41].

-*Targeted analysis*, defined as a quantitative analysis (concentrations are determined) or semiquantitative analysis (relative signals are registered) of a few metabolites and/or substrates of metabolic reactions that might be associated to common chemical classes or linked to selected metabolic pathways [42]. In this metabolomic approach there is a previous knowledge of the metabolites involved in the biological process under study, which allows for the most appropriate analytical technique and a selective sample preparation depending on the nature and abundance of the target metabolites, enabling a sufficient sample cleanup as well as preconcentration, typically using solid phase extraction (SPE) or liquid–liquid extraction (LLE). A mass spectrometer is the detector of choice —particularly the triple quadrupole (QqQ) approach— in targeted metabolomics because of its sensitivity. This was the strategy and instrument selected to develop the studies in Chapters 3-to-5 for the analysis of a small number of compounds: eicosanoids from the serum of patients with cardiac diseases (Chapter 4); the study of the inclusion of an isotopically-labeled internal standard and its effect on the quantification of the compounds studied in chapter 4 (Chapter 3); and that of the chemotherapeutic drug paclitaxel and its main

metabolites in tissue, serum and plasma from women with ovarian cancer in Chapter 5.

-*Untargeted analysis,* based primarily on the qualitative or semiquantitative analysis of the largest possible number of metabolites from a diversity of chemical and biological classes contained in a biological specimen [43]. Detection of this wide range of metabolites can be done through a single analytical platform or a combination of complementary platforms, mainly based on nuclear magnetic resonance (NMR) or mass spectrometry (MS), along with gas chromatography (GC), liquid chromatography (LC) or, less commonly, capillary electrophoresis (CE). In these studies, the relative concentrations of the analyzed metabolites are generally calculated and their variations between two or more system situations are studied. In untargeted approaches, the sample preparation step kept as simple as possible to obtain the widest metabolite coverage. Non-selective sample pretreatments such as protein precipitation (PP) for plasma/serum and simple dilution for urine samples (referred to as dilute-and-shoot) are generally used [44,45]. Besides, the data set obtained by applying these methods is very extensive, so it requires treatments with advanced chemometric approaches for conversion into manageable signals and, finally, interpretable results. The signals require annotation using either available experimental libraries, structural elucidation by *in silico* fragmentation tools or experimental identification. Thus, this strategy allows identification of new metabolites that are involved in metabolic pathways and, therefore, the knowledge of these routes. For the research developed and collected in section C of this Thesis-Book this strategy was selected to obtain as much information as possible from the compounds present in urine and serum (Chapter 6), in exhaled breath condensate (Chapter 7), and in urine of patients with prostate cancer (Chapters 8 and 9).

-*Metabolomics fingerprinting*, defined by Fiehn as a high throughput, fast methodology for analysis of biological samples that provides fingerprints for sample classification and screening [46]. Fingerprinting is not focused on particular metabolite(s); therefore, it allows the discovery of novel metabolic pathways disturbed by the disease or preset stimulus.

-*Metabolomics footprinting*, term proposed in 2005 by Kell *et al.* to refer to the study of metabolites in extracellular fluids, also known as exometabolome or secretome [47]. As in the previous strategy, the purpose is mainly to obtain information for classification of groups of samples or for screening.

Each of these strategies has its own inherent advantages and disadvantages, but they can be highly complementary when used in combination.

The detection techniques used also depend on the chosen strategy. For example, to obtain a metabolomic fingerprint the detection technique should allow direct and rapid analysis of the sample. NMR, MS (depending on the complexity of the sample) and, to a lesser extent, the infrared and Raman spectroscopies, are mainly used in this context. A separation technique is generally applied in both targeted and untargeted analysis prior to individual detection of the metabolites for their quantification or identification, being GC or LC the most used in this area. Then, MS is the commonest detector in metabolomics, both for targeted and untargeted analysis, as it provides a high spectral resolution —and therefore a great accuracy in the measurement of the *m/z* ratio—, or an excellent sensitivity, depending of the given MS approach.

1.4. Present challenges in metabolomics

A proper knowledge of clinical metabolomics challenges is essential to plan research that generates adequate and reliable results for appropriate biological interpretation. Some of the major challenges that face the metabolomics community —and therefore offer potential opportunities— are below.

1. One of the main challenges in metabolomics is the analysis of the entire metabolome in biological systems. The volume of metabolites to analyze can be very large and present different molecular structures (lipids, carbohydrates, nucleic acids, amino acids, organic acids, steroids, peptides and others) in biofluids and tissues in a wide dynamic range of concentrations, so that some metabolites cannot be detected through current experimental methods. As occurs with instru-

mentation, sample preparation and data analysis continue to advance; therefore, achievement of this challenge in the near future is foreseeable.

2. With consideration to the high complexity of biological mixtures, the vast majority of MS analysis methods involve prior separation using LC, GC, or less commonly CE. However, the fast-growing number of applications using continuously evolving separation methods and protocols presents opportunities and challenges. Clearly, advances in chromatographic methods enabled efficient separation of metabolites and enhanced the number of detected metabolites. A major challenge, however, is the inability to compare and correlate the results of such studies on the same or similar samples obtained by independent research groups. This is a major bottle-neck to the growth of the field.

3. The confident assignment of metabolites is critical for the advancement of metabolomics. Thus, identification of 'unknown' metabolites presents a major challenge for reaching a clinical grade in metabolomics. To the best of our knowledge, there is no currently available software with the ability to fully and smoothly automate the whole metabolomics process, particularly the identification step. Moreover, metabolite identification is a prerequisite for absolute quantitation, especially in MS-based methods using labeled isotopes. Therefore, more efforts are urgently needed to overcome this problem by reliable metabolomics databases, increasing data accessibility and data comparability.

4. Regarding clinical utility and application of novel biomarkers, one of the major challenges in metabolomics is validation of the compounds statistically significant in small sets of well-selected samples, in a big cohort. While there are numerous screening studies in metabolomics research producing potential biomarkers, most of the identified biomarkers have failed to replace existing clinical tests. To become a clinically approved test, a potential biomarker should be confirmed and validated using thousands of individuals and should be reproducible, specific, and sensitive. The reproducibility is assured by validating the biomarkers in other samples, preferably from an independent cohort.

5. The real promise of clinical metabolomics is to establish which biochemical pathways are modified in response to a disease in humans or animal models. This provides the ability to monitor disease progression and the response to treatments. Once a set of specific metabolites has been successfully identified with the highest confidence, the biological interpretation of the data represents a major challenge.

6. Also a key challenge is the comparison and correlation of the results from studies on the same samples obtained by different omics approaches. Indeed, in the large majority of clinical metabolomics-based studies, the candidate biomarkers do not show any obvious functional or pathway associations at a first glance. It is only by fusion with other multiple omics data —thus is, data from genomics, proteomics and/or transcriptomics— merged with systems biology and mechanistic approaches (*e.g.*, genome-scale metabolic modeling) that relevant pathways and biomolecular mechanisms may be brought to light. Moreover, the huge amount of metabolome data demands novel bioinformatic strategies. Reducing the data processing time from hours to minutes (if not seconds), and developing more multifunctional tools allowing a user to start with raw data and end up with biologically meaningful results is another critical need.

7. Other limitations are largely associated with experimental design and the information obtained from different samples. One consideration when interpreting the results of metabolomics studies on human samples is whether or not the data come from matched and truly comparable cases. It is a challenge to perform a scientifically accurate 'fair test' between human subjects, but effort made to match diseased and control groups with respect to sample size and epidemiology of subjects provides the possibility for a more reliable comparison between groups. This can be difficult if availability of samples is limited due to small numbers of healthy volunteers to act as control subjects or small numbers of diseased subjects, particularly for rare conditions. Otherwise, the potential confounding factors between cases and controls including sex, age, ethnicity, body mass index (BMI), diet, lifestyle (*e.g.*, smoking habits, physical activity, etc.), health status or medications/xenobiotics should be as similar as possible.

The lipidome is particularly affected by the sex (higher plasma concentrations for triglycerides, ceramides, gangliosides and phospholipids in male), BMI, medications (*e.g.*, statins) and the diet [48]. Some of these confounders, such as age and sex, are easily controlled but a full control and knowledge of all factors is difficult.

8. A common challenge for LC– and GC–MS measurements is to avoid the inter-batch variation in large-scale studies. Batch effects are the main obstacle for validation of reproducibility data. In addition, this drawback is a potential confounder in association studies that may cause spurious associations. Such technical heterogeneity makes meta-analysis and metabolomics data fusion with other omics data a challenging task.

9. While MS is intrinsically highly sensitive, and therefore capable of detecting early biomarkers, problems with reproducibility arising from the chromatography of biofluids and factors such as matrix effects and/or ion suppression still present considerable challenges. ESI allows for an efficient ionization of a large range of analytes presenting diverse physicochemical properties and masses, and a straightforward hyphenation to both LC and CE. However, this ionization approach may suffer from significant matrix effects, in other words, ion suppression or enhancement.

10. Since each biofluid can potentially present its own set of challenges in terms of preparation and analysis in metabolomics, efforts have been made to start addressing this issue. Use of standard operating (SOPs) describing sampling procedures (matrix, sampling site, containers, storage conditions), as well as standardized protocols for sample pretreatment and analytical performance as regular injection of quality controls (QCs), blanks, labelled IS; sample randomization are required to minimize data variability.

11. Other factors such as sample preparation, sample matrix, and carryover effects also contribute to data variability. To circumvent these challenges, there is a need to move away from the conventionally used relative metabolite concentration measurements to more reliable absolute concentration determi-

nations, which will then be independent of the analytical platforms, methods, and protocols used. This approach is not straightforward for MS, as it requires suitable internal standards and proper calibration, but it is very important. Such efforts not only enable comparison and correlation of the vast data accumulated in the literature, but the results are more meaningful as, ultimately, clinical translation of metabolite-based biomarker technology requires biomarkers to be measured and validated in absolute concentration.

2. Separation and detection in metabolomics

The high number of metabolites in any vegetal or animal sample makes mandatory implementation of an appropriate separation of the sample components before being subjected to detection.

2.1. Separation techniques used in metabolomics

The most used separation techniques in metabolomics (commonly prior to a high resolution–high sensitivity detector) are GC or LC depending on the nature of the target metabolites. In addition to improved sensitivity and resolution power of the analysis, chromatographic separation provides extra information (retention time) that facilitates identification of the metabolites, especially in dealing with untargeted metabolomic strategies.

Chromatography is set of separation techniques where the components of the sample are distributed between a mobile phase (a liquid in LC, an inert gas in GC, and a supercritical fluid —SF— in the case of SFC) and a stationary phase (the column packing). The separation takes place as a function of the different affinities of the sample constituents towards each phase. GC allows the separation of volatile compounds, such as, for example, aromatic compounds, but also compounds that can become volatile after derivatization.

In LC the most determining component in the separation is the column, whose characteristics define the type and mode of the chromatography. Reverse

phase columns, mainly those packed with silica (C18 or C8), are characterized by strong interaction with low polarity compounds, whereas normal phase columns, such as HILIC, are more effective for polar compounds, thus covering a wide range of metabolites with very different features. The combination of both GC and LC provides a higher level of information since there are metabolites that are not volatile and do not yield volatile products, do not ionize adequately at atmospheric pressure or are thermolabile. Therefore, both should be used complementary to get the best coverage of metabolites in the biological sample. However, the majority of metabolites to be separated by GC requires a laborious sample preparation step (hydrolysis, derivatization); thus, its use for clinical diagnostic and large-scale studies can be limited. The increase types of materials that provide new possibilities to LC are replacing GC applications, particularly in the field of high-throughput targeted metabolomics due to the easier sample preparation and the shorter analysis times.

2.2. Detection equipment in metabolomics with special emphasis on that used in the present research

Metabolomics analysis demands high sensitivity for detecting metabolites present at low concentrations in biological samples, high-resolution power for identifying the metabolites and wide dynamic range to detect metabolites with variable concentrations in complex matrices. MS spectrometry is the most extensively used analytical technique that fulfils these requirements.

The fundamentals of MS is the differential displacement of ionized molecules through vacuum by applying an electrical field. Simplistically, a mass spectrometer consists of a sample inlet, an ion source, a mass analyzer, a detector and a data system. The sample inlet has the function of introducing the sample molecules into the ion source, where they become ionized. The ion source generates gas-phase ions via an ionization technique, the mass analyzer separates the ions according to their mass-to-charge ratio (m/z), and the detector generates

an electric current from the incident ions that is proportional to their abundances [49].

The required ionization of the target analytes prior to MS detection can be produced by different devices. The most common ionization sources in metabolomics are electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI) and atmospheric pressure photoionization (APPI). Among them, ESI is the most widely used for untargeted metabolomics, mainly because it produces a soft ionization (*viz*, ions are generated with little or no fragmentation, which can help to identify unknown metabolites), and also because it can ionize an extensive polarity range of compounds. This ionization source has been extensively applied in clinical metabolomics and it is particularly useful when coupled to LC as ionization occurs at atmospheric pressure [50]. In addition, ESI does not require derivatization of the target analytes, and ionizes a large mass range of compounds from non-volatile to polar compounds.

ESI is a process that creates or transfers intact ions from the solution to the gas phase at atmospheric pressure. In the coupling of an ESI source to LC, the liquid sample from the chromatography column is introduced through the mobile phase into a capillary tube applying a high potential difference (±3–6 kV) between the capillary tube and the counter electrode in positive or negative ionization modes The electric field causes an accumulation of charge at the liquid surface at the end of the capillary, which break the spray of highly charged droplets. This nebulization is usually assisted by a flow of gaseous nitrogen (the nebulizer gas) flowing through a tube coaxial to the main capillary, aiding in the formation of the aerosol. The drying gas evaporates the solvent in the droplets causing them to shrink. When the repelling forces come close to the cohesion forces, there is a coulombic explosion. The droplets undergo series of ruptures resulting in smaller and smaller droplets which will continue to undergo successive evaporation and explosion processes till the repulsive columbic forces are great enough to overcome the forces of the surface tension of the droplet (Rayleigh limit). The fission process continues until each droplet contains only one ion at the end. The evaporation of the remaining solvent leads to formation

of single ions with multiple charges (z) to the gas phase, even though metabolomics has interest in $z = 1$. Once the molecules have been ionized, they are transported to the mass analyzer via an electric or magnetic field.

ESI is a soft ionization technique, the ions produced are mainly derived from the incorporation or transfer of one or more protons to the molecules, although it may also originate ions corresponding to adducts with other species present in the mobile phase.

Mass analyzers can be used either alone or combined. The combination can be made between the same type of mass analyzer or between different mass analyzers (hybrid instruments) and is called tandem MS (MS/MS). In this approach, the ions that arrive at the first mass analyzer (precursor ions) are isolated, subsequently fragmented, and finally the fragment ions are separated according to their m/z in a second mass analyzer and detected. For some types of mass analyzers, the number of mass analysis steps can be increased (*i.e.*, the fragment ions can be re-fragmented and further detected. In this case, the experiment is termed multiple-stage MS (MS^n), where n refers to the number of mass analysis steps).

Tandem MS and multiple-stage MS improve identification of a given molecule, because not only the molecular ions are detected but also the fragments generated from precursor ions.

The main performance characteristics of a mass analyzer are [51–53]:

(a) Mass accuracy of the measured m/z provided by the mass analyzer, directly related to the mass resolving power and stability of the instrument.

(b) Mass resolving power or ability of an MS to provide a specified value of mass resolution (*i.e.*, the instrument generates distinct signals for two ions with a small m/z difference).

(c) Mass range or limits of m/z within which an MS can detect ions or is operated to record a mass spectrum.

(d) Transmission efficiency or ratio of the number of ions reaching the detector and the number of ions leaving the mass analyzer, related to the sensitivity of the mass spectrometer (*i.e.*, the minimal concentration of a compound leading to a peak intensity greater than a specified signal-to-noise ratio).

(e) Scan speed or rate at which the analyzer measures over a certain mass range.

(f) Scan cycle time or the time required to obtain a mass spectrum, also known as duty cycle.

MS alone can be used in a metabolomics analysis, being this performance known as direct infusion mass spectrometry (DIMS). The major drawbacks of this approach are ion suppression effects —which make difficult the identification/quantification of metabolites with lower concentrations or some metabolite classes that do not ionize as well as others—, and the requirement of a high-resolution mass analyzer, which increases the cost of the analysis.

Mass analyzers to be coupled to LC or GC are similar. However, the different physical state of the eluate from the column requires a different insertion into the detection system. Therefore, whilst in GC the eluate is already a gas, in LC the eluate is a liquid that must be evaporated before being ionized, as explained above in the description of the ESI process.

The particularity that guides the choice of the instrument is the type of metabolomic analysis to be performed, either targeted or untargeted. As described in section 1.3, the goal of targeted metabolomics is to perform a quantitative analysis of specific metabolites (or a defined set of metabolites). The main features for an MS to be used in targeted metabolomics are transmission efficiency, scan cycle time and scan speed. In this case, single quadrupole (Q), triple quadrupole (QqQ), quadrupole ion trap (QIT) and Orbitrap (OT) are the most employed instruments.

The main demanded features for an MS to be used for untargeted metabolomics are mass resolution power, mass range and mass accuracy. In this case,

time of flight (TOF), quadrupole time of flight (QTOF), Fourier transform ion cyclotron resonance (FTICR) and OT are the most used.

The methods developed in the experimental part of this Thesis have been based on a chromatographic separation by LC and subsequent detection by MS. The mass detectors used are described below.

-The QqQ mass spectrometer, of which a scheme is shown in Figure 3 [54]. It consists of an ion source followed by ion optics that transfer the ions to the first quadrupole —a device formed by four parallel rods to which specific direct current and radio frequency voltages are applied. The rods filter out all ions except those of one or more particular *m/z* values as determined by the applied voltage, which is variable, so that ions with other *m/z* values are allowed to pass through.

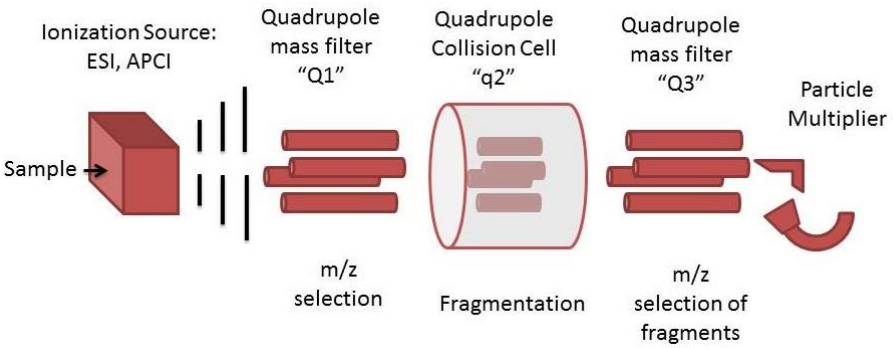


Figure 3. General scheme of a triple quadrupole mass spectrometer.

Afterwards, the selected ions reach a collision cell where they are fragmented. The collision cell —typically known as the second quadrupole— is actually a hexapole filled with an inert gas such as nitrogen or argon. The fragment ions formed in the collision cell are then sent to the third quadrupole for a second filtering step to enable isolation and subsequent examination of multiple precursor-to-product ion transitions. This operational way is called selected reaction monitoring mode (SRM). Since the fragment ions are pieces of the precursor, they represent portions of the overall structure of the precursor molecule. Due to the low-mass accuracy achieved with respect to other mass

analyzers, triple quadrupole spectrometers are preferably used for targeted analysis, as they allow quantification with high sensitivity and selectivity in the SRM mode, actually the most sensitive operational mode for the triple quadrupole MS instrument.

-*The QTOF mass spectrometer*, together with the time-of-flight (TOF) instrument are the most used mass analyzers for untargeted metabolomics as they provide data acquisition over a wide mass range with high mass accuracy and resolving power.

Figure 4 shows the diagram of the instrument used in the research in this Thesis [55]. The QTOF mass analyzer is based on the same configuration as the QqQ but replacing the last quadrupole by an acceleration tube as mass analyzer (usually an orthogonal configuration) to filter out ions according to the equation of kinetic energy. The QTOF can operate in MS mode with the TOF as scanning tool, by taking benefit from the high mass accuracy, or in MS/MS mode for structural elucidation. This hybrid mass analyzer offers better selectivity than triple quadrupoles, meanwhile sensitivity is considerably lower. On the other hand, thanks to its good mass accuracy (below 2 ppm) highly reliable identification can be achieved, thus allowing its use for global metabolic profiling.

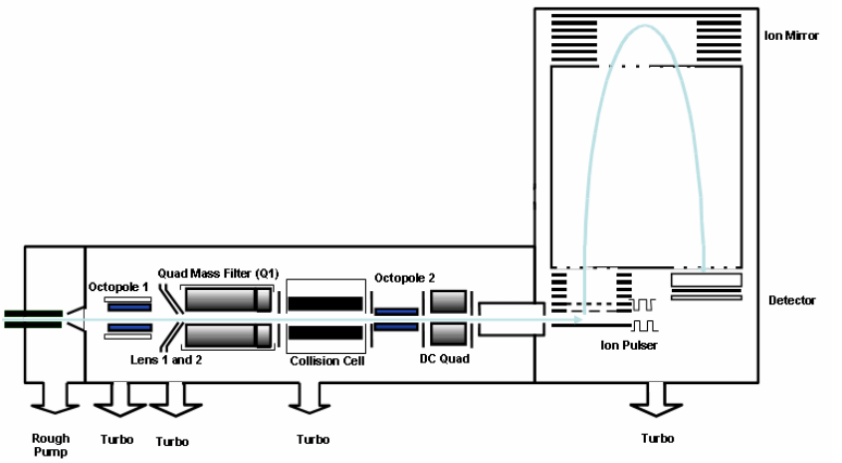


Figure 4. General scheme of a quadrupole time-of-flight mass spectrometer.

2.3. SIL-IS and other approaches used in metabolomics

As previously mentioned, LC–ESI-MS/MS is one of the most frequently used approaches for quantifying endogenous metabolites in biosamples [56]. However, it has to be pointed out that the main drawback of this hyphenated equipment is the ESI interface, which tends to suffer from matrix effects (either ion suppression or enhancement [57]) leading to inaccurate quantifications, especially in dealing with complex matrixes. Ion suppression is caused by the presence of (endogenous or exogenous) coeluting components in the ion source of the MS, which has been attributed to numerous mechanisms including competition for "charges" between analytes and interfering compounds or a change of viscosity and/or surface tension of the droplets in the ion source [58,59].

The two approaches commonly used for accurately quantifying endogenous metabolites are the standard addition method (SAM) and the stable isotope labeled-internal standard (SIL-IS) method [60]. The former works by serial addition of target standards of the analytes into specific sample aliquots to establish the standard addition curve, which is an effective method of calibrating matrix effects caused by quantification errors. The major disadvantage of the SAM is that each sample needs to be analyzed by several runs to obtain the standard addition curve, which makes this approach impractical for clinical analysis where a large batch of samples need to be analyzed.

The SIL-IS method is currently the gold standard for quantifying metabolites. SIL-ISs are chemically and structurally similar to the target analytes with specific atoms in the analytes replaced by their corresponding isotopes, such as deuterium for hydrogen, ^{13}C for ^{12}C, ^{15}N for ^{14}N, or ^{18}O for ^{16}O. The use of SIL-ISs normalizes the MS intensity of analytes to their isotopic analogues and, therefore, effectively compensates for the matrix effect, ion suppression from other coeluting analytes, and variations caused by sample preparation, injection, and instrument parameters [61–66]. Because the physicochemical properties of

an IL-IS are almost the same as the target analyte, the signal changes caused by MEs can be effectively calibrated by the SIL-IS method [67–69].

Theoretically, since the SIL-IS is almost identical in structure to (and coelutes with) the analyte, the degree of ionization suppression or enhancement caused by the coeluting matrix components should be the same for the SIL-IS and analyte. Therefore, while the absolute response may be affected, the analyte-to-IS peak area ratio should be unaffected and the analytical method should be accurate, precise and rugged. The mass difference between the target analyte and SIL-IS has to be more than 3 mass units to avoid signal contribution. The ^{13}C- or ^{15}N-labeled are often more appropriate as IS than the ^{2}H-labeled ones, because the physical properties of ^{12}C and ^{13}C are more similar as compared to ^{1}H and ^{2}H. However, synthesizing the non-deuterated SIL may be expensive and impractical.

In the absence of SIL standards, structural analogues are used as the second best choice; but their use may result in poor quantification performance, particularly in analyzing metabolites present in a complex matrix [70]. Structural analogues are often not coeluted with the analyte of interest, and therefore, they would experience levels of matrix effect different from that of the analyte. As a result, the relative signal intensities of a metabolite and its analogue may not reflect their concentration ratio in the sample.

In targeted analyses, the use of internal standards, specially SIL-ISs, is recommended to improve precision and to handle matrix effects. However, in broad metabolic profiling approaches this is not a practical tool because the number of metabolites is large, they are chemically too diverse to afford a common labeling approach, and many of them may not even be known.

Other tools from which metabolomics takes advantage are SOPs and QC samples. The former could, to some extent, prevent heterogeneity and avoid misinterpretations [71–74]. The use of standard QC and metrics normalizes intralaboratory and interlaboratory metabolomics measurement variations [75,76], reinforced by applying consistent statistical correction methods and appropriate computational tools to overcome some technical variations [77,78].

3. Data treatment

3.1. Generalities

Chemometrics can be defined as "the chemical discipline that uses mathematical, statistical, and other tools employing formal logic, to design or select optimal measurement procedures and experiments, and to provide maximum relevant chemical information by analyzing chemical data" [79]. This discipline has been fundamental for the development of metabolomics, while growing with it. From design of experiments, through data preprocessing and processing, to data analysis, chemometrics tools are used to design, process, visualize, explore and analyze metabolomics data. The huge amount of data generated by the entire set of metabolites in metabolomics analysis are evaluated by biostatistics and bioinformatics tools. Chemometrics has been heavily used in the research developed in this Thesis and it is here discussed in a clinical metabolomics context.

3.2. Data preprocessing

From data acquisition to statistical analysis, metabolomics data need to undergo several preprocessing and pretreatment steps (see Figure 5), which are crucial for data quality and interpretation of the results. In this sense, the use of SIL-IS already mentioned above is involved in the tasks of pretreatment of the data, preparing them to be comparable.

The purpose of data preprocessing is the transformation of the raw data (those provided by the analytical instrument that are exclusive of its trademark) into clean data (those that have a universal format) for processing by any software of data processing. Some representative examples of the tools used in this step are as follows:

-Extraction of the potential metabolites (usually called potential molecular features). The term feature has been defined to be a bounded, two-dimensional

LC–MS signal consisting of a chromatographic peak (*i.e.,* retention time) and an MS signal (*m/z* value).

-Alignment: inline order of the signals that correspond to the same metabolites in chromatographic registers. The retention times in chromatography are subjected to variability between analyzers, so it is necessary to correct the retention times to align the signals corresponding to the same metabolites, thus allowing combining data from different samples.

-Remove redundant information: ions with identical elution profiles and related *m/z* values (representing different adducts or isotopes of the same compound) are grouped as a unique feature to remove redundant information.

-Missing values: the management of missing values is also a key step in data preprocessing. Commonly, the missing values are automatically replaced by zero by the software during peak detection. The missing values can also be replaced by noise-signal levels or a fix number which is far below the smallest signal [80].

As described in the section on current metabolomics challenges (section 1.4.), one of the major limitations of MS is the long time consumed in metabolomics analysis to process the extensive raw data sets generated, mainly in untargeted analysis. Moreover, one of the main shortcomings of LC–MS approaches is that it produces two independent data sets (one per polarity) that requires two parallel data analyses to identify significant metabolites associated with the target biological effect, which entails to spend twice the amount of time. In this sense, Chapter 6 of this Thesis-Book overcomes this shortcoming by developing a tool to combine the data matrices from the positive and negative ionization modes in MS. Thus, redundant information is eliminated, and the time spent in data processing is shortened as a result.

3.3. Data pretreatment

Once data preprocessing is complete, the results must be summarized in a data table to start data pretreatment. This table contains the analytical results

distributed in *m* files (observations, samples) and *n* columns (variables, frequencies, *m/z* values, peak areas).

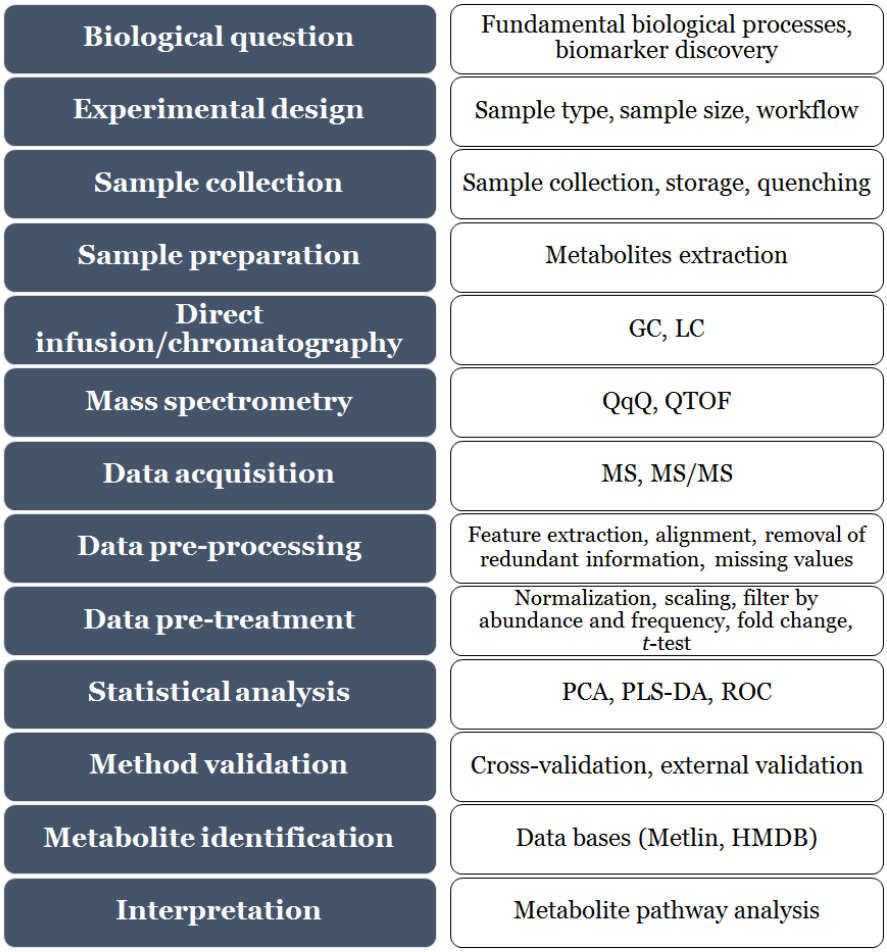


Figure 5. Metabolomics workflow in MS.

The next step is to apply mathematical operations to facilitate comparison among samples, which will be the objective of data pretreatment. The rows and the columns must be as comparable as possible after all the processing steps. For this, the effects of experimental and instrumental variables are aimed to be reduced, whereas the relevant biological variations are aimed to be enhanced. Data

pretreatment may affect the final results of the metabolomics study [81]. In fact, pretreatment can make the difference between a useful model and no model at all. Prior to data analysis, the dataset is commonly pretreated to transform the data to be suitable for analysis by using the following tools:

- Normalization: the main objective of this approach is to remove unwanted variation among samples allowing quantitative comparison of them. In metabolomics, the samples under study are biofluids in most cases, and they exhibit differences in the concentration of metabolites due to variable dilution factors for different samples. For example, metabolite concentrations in different urine samples may differ because of the different amount of water as the solvent. Therefore, the concentrations of the metabolites will reflect dilution instead of the changes in metabolic responses. In order to remove such variations among samples, a normalization factor should be computed for each row of the data table. Different normalization approaches can be applied, from normalization using a single or multiple internal/external addition, total area normalization, or standard to logarithmic transformation and more complex alternatives such as probabilistic quotient normalization and quantile normalization [82–86].

- Scaling: the scaling methods are data pretreatment approaches that divide each variable by a different factor in order to treat all variables equally, regardless of their intensity. One of the most common factors employed is the mean of each variable, although Z-transform can also be used in case of combining data obtained from different platforms or methodologies [87].

- Reducing batch and drift effects: LC–MS instruments are prone to batch and time drift effects owing to changes in instrument sensitivity and intensity, among others. Targeted methods correct batch and drift effects with the inclusion of SIL-IS (as mentioned in section 2.3.) calculating the ratio between the target compound and the internal standard. For untargeted LC–MS, several strategies have been used to correct for these effects: inclusion of periodic QC samples (pooled from all or from a group of samples in the experimental set) that are expected to yield the same results along time; addition of internal standards to

the samples (heavily labelled and/or not occurring in the samples) representing different compound classes, and experimental design using paired samples (when having samples with and without effect).

Untargeted metabolomics experiments generate large multivariate data sets that require some filters to facilitate further statistical analysis by reducing the number of potential molecular features. Thus, only potential entities with high intensity, high significance with respect to a factor of interest, or entities that appear with high frequency in the samples batch would pass the filter. As a result, different filters have been created to fulfill different requirements, namely:

- Filter by abundance: this tool allows removing all potential molecular features below a certain abundance.

- Filter by frequency: the application of this filter eliminates entities present below a certain percentage of samples from each group.

- Fold change analysis: this tool compares the level of each potential molecular feature among the groups under study and retains only the entities that show a preset change.

- Filter based on univariate hypothesis test: only entities with a *p*-value below 0.05 or 0.01 pass this filter.

- Filter by Volcano plot: this approach is a combination of analysis of variance and fold change analysis.

3.4. Data analysis

After data pretreatment, the data matrix can be analyzed by multivariate or univariate methods to distinguish different diseases or subtypes and defining candidate biomarkers. Metabolomics studies typically generate large amounts of data that complicates the use of univariate statistical analysis. However, if the concentration of a particular metabolite is found to be significantly altered through multivariate analysis, univariate analysis can be used to test the

statistical significance of the change. This typically involves the use of the Student's *t*-test or one-way analysis of variance (ANOVA). A more useful statistical approach is multivariate analysis, which can be applied to reduce large volumes of data into a few dimensions for classification and prediction of outcomes. Statistical analysis strategies employed in metabolomics are designed according to the approach selected, being slightly different for targeted and untargeted analysis.

Targeted metabolomics analysis often requires statistical methods to be used during the development of the optimized analytical methodology for the determination of the compounds of interest. Thus, statistical procedures as response surface or screening design are employed to optimize experimental parameters to obtain the best conditions for determination of the target compounds [88,89]. The rest of statistical analyses employed in the process of developing a new methodology for targeted metabolomics are based on simple approaches such as simple regression, used to obtain the calibration curves for metabolites quantitation, or simple calculations, used to characterize the analytical method (reproducibility and repeatability, mainly).

On the other hand, untargeted metabolomics experiments generate large multivariate data sets that constitute a great challenge from a statistical point of view. Thus, reliable and robust approaches are needed to handle and extract the relevant information from the vast amount of generated data.

The most commonly used multivariate methods include unsupervised methods such as principal component analysis (PCA), and supervised methods such as partial least square discrimination analysis (PLS-DA). In unsupervised strategies, no prior knowledge of the data is built into the statistical model. In contrast, supervised learning techniques require a training set of data to be used for prediction and classification. It is worth noting that the predictive power of multivariate analysis comes from patterns in the data, and identification of specific metabolites is not necessary to discriminate different data classes.

PCA is a non-supervised method, as it contains no assumptions on the data. It is used as a visualization and exploration tool at the start of any analysis to detect trends, groups and outliers. PCA attempts to uncover hidden internal structures by decomposing the data matrix into principal components (latent variables or latent structures), describing the maximal variance of the data [90]. This allows representing the structured variance in the data by a smaller number of (latent) variables preserving as much of the relevant information as possible, while discarding the noise, thus making it appropriate for dimensional reduction. Mathematically, the first principal component is the line that better approximates the data, in the least-squares sense. Thus, it represents the direction of the largest variance in the dataset or, in other words, the direction in which the variance of the samples coordinates is maximized. The popularity of this technique stems from the easy graphical interpretation of the data, with clustering between data points based on class-type (age, sex, etc.) often observed. The vast majority of metabolomics studies involves PCA as a first exploratory step [91–95], whereas univariate statistical tests such as the Student's *t*-test are used to identify the relevant variables after exploratory analysis.

The Student's *t*-test, the one-way analysis of variance or the non-parametric equivalents can be used to identify statistical differences between samples of distinct classes [96]. The predictive power of each variable is assessed by finding statistically significant differences between the mean intensity values of a given signal, of which the calculated *p*-value is a straightforward indicator. Such procedures are easily understandable but their use is rather limited in dealing with thousands of highly correlated variables. False positives (type I error) are likely to occur when performing multiple comparisons. Procedures such as the Bonferroni of the Benjamini correction have been introduced to address this issue [97]. The vast majority of metabolomics-dedicated software provides statistical hypothesis testing [98,99]. Many studies focus on identifying biomarkers and though mainly cite *p*-values from various tests as their results. However, the *p*-value only indicates that the amount of a given compound is statistically different among the investigated groups. It does not allow con-

clusions about degree of overlap, direction (positive or negative) or specificity. Therefore, plots of actual values (for instance box-whisker plots) should supplement the *p*-values, so readers can better form their own opinion about the usefulness of the proposed biomarker. On the other hand, fold change (FC) is commonly used to demonstrate the metabolic alterations between different groups that provide additional information beyond hypothesis tests. PCA is usually followed by a supervised learning tool, such as PLS-DA, which attempts to maximize the covariance between the independent and dependent variables to discriminate among samples. This method often discriminates among classes better than PCA does. Moreover, the lack of class information when determining the principal components in PCA can lead to discrimination of samples based on factors non-related to the classes of interest.

PLS is a supervised multivariate linear regression method similar in concept to PCA, which finds the relations between two matrices (data X and response Y), by maximizing the covariance of their latent variables. Supervised learning considers each object with respect to an observed response and includes regression and classification problems depending on the output type under consideration (*i.e.*, a numerical value in the first case and a class label in the second). This classification aims at producing general hypotheses based on a training set of examples that are described by several variables and identified by known labels corresponding to the existing classes. A PLS-based classification therefore has the property that it builds data structures with an intrinsic prediction power, by maximizing the covariance between the data and the class assignment. In the context of classification, PLS-DA is performed to sharpen the partition between groups of observations, in such a way that a maximum separation among classes is obtained. The model can then be analyzed to understand which variables carry the class-separating information [100]. However, overfitting of the model should be noted especially when too many metabolic features are enrolled compared with limited sample numbers. Therefore, cross validation and external validation are necessary steps to handle the high dimensional dataset from LC–MS. Nevertheless, this also requires a separate set

of data to test the predictions of the model, which can sometimes be problematic when not enough independent data to build both training and validation sets are available. This can be circumvented by using a data-splitting method such as cross-validation, where the data are continually split into training and validation sets and the predictions of the data sets are averaged [101].

While the above validation methods can provide information of the quality of the models, prediction of external data can elevate to a higher level our confidence in the model quality. Depending on the objective of the experiment, more confidence can be deposited in a model that can predict samples that were acquired or processed in different times or machines, by different operators, etc. and were not used for model construction.

The statistical analyses described above are not sufficient to acquire detailed biological understanding, needed for biomarker discovery. In contrast to metabolomics studies focused on deciphering biological processes, where interesting metabolites are found *post hoc*, in biomarker studies metabolite selection should be performed a priori rather than *post hoc*. That is, biomarker selection must be performed before deriving a definitive multivariate predictive model. Furthermore, whereas long lists of metabolites or large multivariate models are quite useful for understanding pathways and biological processes, they are not ideal for developing cost-effective biomarker tests. Rather, a short list of 1-to-10 biomarkers is mathematically much more robust and far more practical for clinical testing purposes. While pattern discovery methods such as unsupervised clustering or PCA are useful for discovering novel biological processes, they are not ideal for biomarker discovery. Instead, supervised machine learning algorithms (PLS), or multivariate regression models are used for modelling the discriminatory relationship between a binary dependent variable and one or more explanatory variables.

The selection of potential biomarkers by univariate statistical significance test is inappropriate, as often metabolites that are not significant in isolation can produce clear and reproducible discrimination when combined into a single

multivariate model. For this reason, iterative methodologies for biomarker panel development, like the employed in the PanelomiX bioinformatics tool [102], are recommended.

On the other hand, biomarker panels are designed to discriminate with an optimal sensitivity/specificity but in most of the cases the parameter employed for measuring the clinical utility of a panel developed by PLS, for example, is the R^2 (multiple correlation coefficient). Despite R^2 of the model and R^2 of its cross-validation can certainly be used as part of the biomarker selection process, they provide very little transparency and are not a readily interpretable indication of the clinical utility of a given model. Additionally, most clinicians are not familiar with this style of model evaluation, being the receiver operating characteristic (R) curve analysis the standard method for describing and assessing the performance of medical diagnostic tests with binary classification [103].

Chapters 8 and 9 of the present Thesis are devoted to the search of biomarkers panels for the diagnosis of prostate cancer. For this purpose, two multivariate methods have been used: random forest (RF) analysis, to identify preset number of compounds with the highest frequency of occurrence in the tree branches generated by application of this algorithm, and ROC curves analysis to check the discrimination accuracy of the target compounds. RF analysis constructs models comprising an ensemble of classification trees, each formed on a random subset of features, from a random subset of samples, selected (with bootstrapping) from the training population. Test samples are evaluated by every tree in the forest, and class prediction is based on the majority vote of the ensemble. RF analysis can be applied to comparing the discrimination accuracy of significant compounds in case-control studies, usually checked by analysis of ROC curves.

ROC curves consider the frequency with which the test produces true positives, true negatives, false positives and false negatives. These values are summarized into the proportion of actual positives that are correctly classified as

positive (sensitivity) and the proportion of actual negatives that are correctly classified as negative (specificity). An example of ROC curve is shown in Figure 6.

For ease of interpretation, sensitivity can be considered as the probability of true positives (the positive result of the test from a subject that has an actual positive outcome), and specificity can be considered as the probability of false positives (the negative result of the test from a subject that has an actual negative outcome).

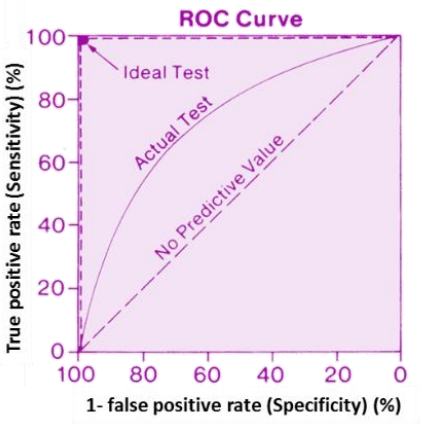


Figure 6. ROC curve plot including the curve described for an ideal test, and a test without predictive value.

As the sensitivity and specificity of a test can vary depending on the biomarker decision boundary, the best way to observe how this decision threshold affect sensitivity and specificity is through a ROC curve. Furthermore, the ROC curve is a non-parametric measure of biomarker utility, being widely considered the most objective and statistically valid method for biomarker performance evaluation [103,104].

There are different informatics tools for ROC curve generation and the parameter that often summarized the curve is the area under the curve (AUC). This parameter can be interpreted as the probability that a diagnostic test or a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. If all positive samples are ranked before negatives ones (example of a perfect classifier), the AUC is 1.0. An AUC of 0.5 is equivalent to

randomly classifying subjects as either positive or negative (classifier without utility). However, using the whole area under a ROC curve may not be appropriate in some cases and the partial AUC is most useful when only certain regions of the ROC space (high sensitivity or high specificity) are of particular interest [105].

References

[1] O. Fiehn, Metabolomics—the link between genotypes and phenotypes. Plant Mol. Biol. 48(1-2) (2002)155–171.

[2] S.G. Oliver, M.K. Winson, D.B. Kell, F. Baganz, Systematic functional analysis of the yeast genome, Trends Biotechnol. 16 (1998) 373–378.

[3] G. Theodoridis, H.G. Gika, I.D. Wilson, LC–MS-based methodology for global metabolite profiling in metabonomics/metabolomics, Trends Anal. Chem. 27 (2008) 251–260.

[4] J.K. Nicholson, J.C. Lindon, Systems biology: metabonomics, Nature 455 (2008) 1054–1056.

[5] J.K. Nicholson, J.C. Lindon, E. Holmes, 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 11 (1999) 1181–1189.

[6] V.S. Gomase, S.S. Changbhale, S.A. Patil, K.V. Kale, Metabolomics, Curr. Drug Metabol. 9 (1) (2008) 89–98.

[7] S. Rochfort, Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research, J. Nat. Prod. 68 (12) (2005) 1813–1820.

[8] U. Christians, J. Albuisson, J. Klawitter, J. Lawitter, Metabolonomics: screening for biomarkers of kidney disease, in: biomarkers in kidney disease. L.E Charles, (Ed.), Academic Press, San Diego (2011).

[9] D. Ryan, K. Robards, Metabolomics: The greatest omics of them all?, Anal. Chem. 78 (23) (2006) 7954–7958.

[10] D. Beyoğlu, J.R. Idle, Metabolomics and its potential in drug development, Biochem. Pharmacol. 85 (2013) 12–20.

[11] M.J. Nunes de Paiva, H. Costa Menezes, Z. de Lourdes Cardeal, Sampling and analysis of metabolomes in biological fluids, Analyst 139 (2014) 3683–3694.

[12] J.K. Nicholson, Global systems biology, personalized medicine and molecular epidemiology, Mol. Syst. Biol., 52 (2006) 1–6.

[13] A.R. Joyce, B.O. Palsson, The model organism as a system: integrating omics data sets, Nat. Rev. Mol. Cell Biol. 7 (2006) 198–210.

[14] A. Zhang, A. H. Sun, H. Xu, S. Qiu, X. Wang, Cell metabolomics, OMICS 17 (10) (2013) 495–501.

[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, Trends Anal. Chem. 29 (2010) 111–196.

[16] 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(7) (2005) 557–65.

[17] https://www.webofstories.com/play/james.watson/97

[18] D.A. Dias, O.A.H. Jones, D.J. Beale, B.A. Boughton, D. Benheim, K. A. Kouremenos, J.L. Wolfender, D.S. Wishart, Current and future perspectives on the structural identification of small molecules in biological systems, Metabolites 6(4) (2016) 46–75.

[19] W.B. Dunn, N.J. Bailey; H.E. Johnson, Measuring the metabolome: current analytical technologies, Analyst 130 (5) (2005) 606—625.

[20] S.G. Oliver, M.K. Winson, D.B. Kell, F. Baganz, Systematic functional analysis of the yeast genome, Trends Biotechnol., 16 (1998) 373–378.

[21] D. Wishart, R. Mandal, A. Stanislaus, M. Ramirez-Gaona, Cancer metabolomics and the human metabolome database, Metabolites, 6 (2016) 10–27.

[22] D.S. Wishart, T. Jewison, A.C. Guo, M. Wilson, C. Knox, Y. Liu, HMDB 3.0-The human metabolome database in 2013, Nucleic Acids Res. 41 (2013) 801–807.

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

[24] 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) 451–459.

[25] D.B. Kell, Metabolomic biomarkers: search, discovery and validation, Expert Rev. Mol. Diagn. 7(4), (2007) 329–333.

[26] Q. Feng, Z. Liu, S. Zhong, R. Li, H. Xia, Integrated metabolomics and metagenomics analysis of plasma and urine identified microbial metabolites associated with coronary heart disease, Sci. Rep. 6 (22525) (2016) 1–14.

[27] K. Dettmer, P.A. Aronov, B. D. Hammock, Mass spectrometry-based metabolomics, Mass Spectrom. Rev. 26(1) (2007) 51–78.

[28]R. Kaddurah-Daouk, B Kristal, Methods for drug discovery, disease treatment, and diagnosis using metabolomics- US Patent 7,005,255 (2006).

[29] T. Burt, S. Nandal, Pharmacometabolomics in early-phase clinical development, Clin. Transl. Sci. 9 (2016) 128–138.

[30] R. Llorach, M. Garcia-Aloy, S. Tulipani, R. Vazquez-Fresno, C. Andrés-Lacueva, Nutrimetabolomic strategies to develop new biomarkers of intake and health effects, J. Agric. Food Chem. 60(36) (2012) 8797–8808.

[31] S. Bekri, The role of metabolomics in precision medicine, Expert Rev. Precis. Med. Drug. Dev., 1(6) (2006) 517–532.

[32] M. Bayet-Robert, D. Morvan, P. Chollet, C. Barthomeuf, Pharmacometabolomics of docetaxel treated human MCF7 breast cancer cells provide evi-

dence of varying cellular responses at high and low doses, Breast Cancer Res. Treat. 120 (2010) 613–26.

[33] M. Castell, S. Saldaña-Ruiz, M. J. Rodríguez-Lagunas, À. Franch, F. J. Pérez-Cano, Second international congress on chocolate and cocoa in medicine held in Barcelona, Spain, 25–26th September 2015, Nutrients 7(12) (2015) 9785–9803.

[34] L. Brennan, Metabolomics in nutrition research: current status and perspectives, Biochem. Soc. Trans. 41(2) (2013) 670–673.

[35] J. Astle, J.T. Ferguson, J.B. German, G.G. Harrigan, N.L. Kelleher, T. Kodadek, B.A. Parks, M.J. Roth, K.W. Singletary, C.D. Wenger, G.B. Mahady, Characterization of proteomic and metabolomic responses to dietary factors and supplements, J. Nutr. 137(12) (2007) 2787–2793.

[36] R.D. Beger, W. Dunn, M.A. Schmidt, Metabolomics enables precision medicine: "A white paper, community perspective", Metabolomics 12(9) (2016) 1–15.

[37] A.J. Baumler, V. Sperandio, Interactions between the microbiota and pathogenic bacteria in the gut, Nature 535(7610) (2016) 85–93.

[38] J. Jansson, B. Willing, M. Lucio, Metabolomics reveals metabolic biomarkers of Crohn's disease, PLoS One 4(7) (2009) e6386.

[39] T.M. Hardy, T.O. Tollefsbol, Epigenetic diet: impact on the epigenome and cancer, Epigenomics 3(4) (2011) 503–518.

[40] D.H. Roukos, Next-generation sequencing and epigenome technologies: potential medical applications, Expert Rev. Med. Devices 7(6) (2010) 723–726.

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

[42] K.T. Myint, K. Aoshima, S. Tanaka, T. Nakamura, Y. Oda, Quantitative profiling of polar cationic metabolites in human cerebrospinal fluid by reversed-phase nanoliquid chromatography/mass spectrometry, Anal. Chem 81 (2009) 1121–1129.

[43] G. Theodoridis, H.G. Gika, I.D. Wilson, LC–MS-based methodology for global metabolite profiling in metabonomics/metabolomics, Trends Anal. Chem. 27 (2008) 251–260.

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

[45] 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.

[46] E. Allard, D. Bäckström, R. Danielson, 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) 8946–8955.

[47] 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) 557–565.

[48] T. Hyotylainen, M. Oresic, Bioanalytical techniques in nontargeted clinical lipidomics, Bioanalysis 8(4) (2016) 351–364.

[49] E. de Hoffmann, V. Stroobant (Eds), Mass spectrometry–principles and applications. John Wiley & Sons, England (2007).

[50] Z. Lei, D.V. Huhman, L.W. Sumner, Mass spectrometry strategies in metabolomics, J. Biol. Chem. 286(29) (2011) 25435–25442.

[51] K.K. Murray, R.K. Boyd, M.N. Eberlin, G.J. Langley, L. Li, Y. Naito, Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013), Pure Appl. Chem. 85 (2013) 1515–609.

[52] E. de Hoffmann, V. Stroobant (Eds), Mass spectrometry–principles and applications. John Wiley & Sons, England (2007).

[53] A. Makarov, M. Scigelova, Coupling liquid chromatography to orbitrap mass spectrometry, J. Chromatogr. A 1217 (2010) 3938–3945.

[54] J. Ni, J. Rowe, Microdosing assessment to evaluate pharmacokinetics and drug metabolism using liquid chromatography–tandem mass spectrometry technology, in: topics on drug metabolism. J. Paxton (Ed.) Intech (2012).

[55] Agilent Technologies. QTOF manual (Components) (accessed on 13th May 2014).

[56] B. Zhou, J.F. Xiao, L. Tuli, H.W. Ressom, LC-MS-based metabolomics, Mol. Biosyst. 8(2) (2012) 470–481.

[57] C. Ghosh, C.P. Shinde, B.S. Chakraborty, Influence of ionization source design on matrix effects during LC–ESI–MS/MS analysis, J. Chromatogr. B 893-894 (2012) 193–200.

[58] J.P. Antignac, K. de Wasch, F. Monteau, H.D. Brabander, F.O. Andre, B.L Bizec, The ion suppression phenomenon in liquid chromatography mass spectrometry and its consequences in the field of residue analysis, Anal. Chim. Acta 529(12) (2005) 129–136.

[59] W.M.A. Niessen, P. Manini, R. Andreoli, Matrix effects in quantitative pesticide analysis using liquid chromatography–mass spectrometry, Mass Spectrom. Rev. 25(6) (2006) 881–899.

[60] E. Gorrochategui, J. Jaumot, S. Lacorte, R. Tauler, Data analysis strategies for targeted and untargeted LC–MS metabolomic studies: overview and workflow, Trends Anal. Chem. 82 (1) (2016) 425–442.

[61] E. Stokvis, H. Rosing, J. Beijnan, Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not?, Rapid Commun. Mass Spectrom. 19 (2005) 401–407.

[62] H. Liang, R. Foltz, M. Meng, P. Bennet, Ionization enhancement in atmospheric pressure chemical ionization and suppression in electrospray ionization

between target drugs and stable-isotope-labeled internal standards in quantitative liquid chromatography/tandem mass spectrometry, Rapid Commun. Mass Spectrom. 17 (2003) 2815–2821.

[63] T. M. Annesley, Ion suppression in mass spectrometry, Clin. Chem. 49 (2003) 1041–1044.

[64] F. Magni, S. Pereira, M. Leoni, G. Grisenti, G. Kienle, Quantitation of cyclosporin A in whole blood by liquid chromatography/stable isotope dilution electrospray ionization mass spectrometry, J. Mass Spectrom. 36 (2001) 670–676.

[65] P. Jacob, M. Wilson, L. Yu, J. Mendelson, R.T. Jones, LC–MS has become an important tool for metabolome, Anal. Chem. 74 (2002) 5290–5296.

[66] Z. Kuklenyik, D.L. Ashley, A.M. Calafat, quantitative detection of trichloroacetic acid in human urine using isotope dilution high-performance liquid chromatography–electrospray ionization tandem mass spectrometry, Anal. Chem. 74 (2002) 2058–2063.

[67] J.M. Wu, R. Wiegand, P. Lo Russo, J. Li, A stable isotope-labeled internal standard is essential for correcting for the interindividual variability in the recovery of lapatinib from cancer patient plasma in quantitative LC–MS/MS analysis, J. Chromatogr. B, 941 (2013) 100–108.

[68] W.Y. Jian, R.W. Edom, Y.D. Xu, J. Gallagher, N.D. Weng, Potential bias and mitigations when using stable isotope labeled parent drug as internal standard for LC–MS/MS quantitation of metabolites, J. Chromatogr. B 878 (2010) 3267–3276.

[69] G. Haubl, F. Berthiller, R. Krska, R. Schuhmacher, Suitability of a fully C-13 isotope labeled internal standard for the determination of the mycotoxin deoxynivalenol by LC–MS/MS without clean up, Anal. Bioanal. Chem. 384 (2006) 692–696.

[70] E. Stokvis, H. Rosing, L. López-Lázaro, J. Schellens, J. Beijnan, Biomed. Chromatogr. 18 (2004) 400–402.

[71] E. Kolker, V. Ozdemir, L. Martens, Toward more transparent and reproducible omics studies through a common metadata checklist and data publications, OMICS 18(1) (2014) 10–14.

[72] J.D. Tenenbaum, S.A. Sansone, M. Haendel, A sea of standards for omics data: sink or swim?, JAMIA 21(2) (2014) 200–203.

[73] S. Chitayat, J.F. Rudan, Phenome centers and global harmonization, in: metabolic phenotyping in personalized and public healthcare. E. Holmes, J.J. Nicholson, A.W. Darzi, J.C. Lindon, (Eds.), Academic Press, Boston, (2016).

[74] P. Rocca-Serra, R.M. Salek, M. Arita, Data standards can boost metabolomics research, and if there is a will, there is a way, Metabolomics 12(1) (2016) 1–13.

[75] W.B. Dunn, I.D. Wilson, A.W. Nicholls, D. Broadhurst, The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans, Bioanalysis 4(18) (2012) 2249–2264.

[76] M. Walzer, L.E. Pernas, S. Nasso, qcML: an exchange format for quality control metrics from mass spectrometry experiments, Mol. Cell Proteomics 13(8) (2014) 1905–1913.

[77] H.J. Issaq, T.J. Waybright, T.D. Veenstra. Cancer biomarker discovery: opportunities and pitfalls in analytical methods, Electrophoresis 32(9) (2011) 967–975.

[78] P. Jonsson, A. Wuolikainen, E. Thysell, Constrained randomization and multivariate effect projections improve information extraction and biomarker pattern discovery in metabolomics studies involving dependent samples, Metabolomics 11(6) (2015) 1667–1678.

[79] D.L. Massart, S.N. Deming, Y. Michotte, L. Kaufman, B.G.M. Vandeginste, (Eds.), Chemometrics: a text-book, Elsevier Sciences Ltd., New York, (1988).

[80] P. Yin, D. Wan, C. Zhao, J. Chen, X. Zhao, W. Wang, X. Lu, S. Yang, J. Gu, G. Xu, A metabonomic study of hepatitis B-induced liver cirrhosis and hepato-

cellular carcinoma by using RP-LC and HILIC coupled with mass spectrometry, Mol. Biosyst. 5 (2009) 868–876.

[81] J. Godzien, M. Ciborowski, S. Angulo, C. Barbas, From numbers to a biological sense: how the strategy chosen for metabolomics data treatment may affect final results. A practical example based on urine fingerprints obtained by LC–MS, Electrophoresis 34 (2013) 2812–2826.

[82] F. Dieterle, A. Ross, G. Schlotterbeck, H. Senn, Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics, Anal. Chem. 78(13) (2006) 4281–4290.

[83] B.M. Bolstad, R.A. Irizarry, M. Åstrand, T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias, Bioinformatics 19(2) (2003) 185–193.

[84] K.A. Veselkov, L.K. Vingara, P. Masson, S.L. Robinette, E. Want, J.V. Li, Optimized pre-processing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery, Anal Chem., 83(15) (2011) 5864–5872.

[85] M. Sysi-Aho, M. Katajamaa, L. Yetukuri, M. Orešič, Normalization method for metabolomics data using optimal selection of multiple internal standards, BMC Bioinf. 8(1) (2007) 1–17.

[86] S. Bijlsma, I. Bobeldijk, E.R. Verheij, R. Ramaker, S. Kochhar, I.A. Macdonald, Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation, Anal Chem. 78(2) (2006) 567–574.

[87] R.A. van der Berg, H.C.J. Hoefsloot, J.A. Westerhuis, A.K. Smilde, M.J. van der Werf, Centering, scaling, and transformations: improving the biological information content of metabolomics data, BMC Genomics 7 (2006) 1–15.

[88] M. Calderón-Santiago, J.M. Mata-Granados, F. Priego-Capote, J.M. Quesada-Gómez, M.D. Luque de Castro, Analytical platform for verification and

quantitation of target peptides in human serum: application to cathelicidin, Anal. Biochem. 415 (2011) 39–45.

[89] M.P. Delgado-Torre, C. Ferreiro-Vera, F. Priego-Capote, P.M. Pérez- Juan, M.D. Luque de Castro, Comparison of accelerated methods for the extraction of phenolic compounds from different vine-shoot cultivars, J. Agric. Food Chem. 60 (2012) 3051–3060.

[90] J. Boccard, J.L. Veuthey, S. Rudaz, Knowledge discovery in metabolomics: an overview of MS data handling, J. Sep. Sci. 33 (2010) 290–304.

[91] H. Idborg-Björkman, P.O. Edlund, O.M. Kvalheim, I. Schuppe-Koistinen, S.P. Jacobsson, Screening of biomarkers in rat urine using LC/electrospray ionization-MS and two-way data analysis, Anal. Chem. 75 (2003) 4784—4792.

[92] H.J. Major, R. Williams, A.J. Wilson, I.D. Wilson, A metabonomic analysis of plasma from Zucker rat strains using gas chromatography/mass spectrometry and pattern recognition, Rapid Commun. Mass Spectrom. 20 (2006) 3295–3302.

[93] K.M. Pierce, J.L. Hope, J.C. Hoggard, R.E. Synovec, A principal component analysis based method to discover chemical differences in comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry (GC × GC–TOFMS) separations of metabolites in plant samples, Talanta, 70 (2006) 797–804.

[94] D. Bylund, J. Samskog, K.E. Markides, S.P. Jacobsson, Classification of lactate dehydrogenase of different origin by liquid chromatography–mass spectrometry and multivariate analysis, J. Am. Soc. Mass Spectrom. 14 (2003) 236–240.

[95] P.S. Belton, I.J. Colquhoun, E.K. Kemsley, I. Delgadillo, P. Roma, M.J. Dennis, M. Sharman, E. Holmes, J.K. Nicholson, M. Spraul, Application of chemometrics to the 1H NMR spectra of apple juices: discrimination between apple varieties, Food Chem. 61 (1998) 207–213.

[96] M.C. Wiener, J.R. Sachs, E. Deyanova, N.A. Yates, Differential mass spectrometry: a label-free LC–MS method for finding significant differences in complex peptide and protein mixtures, Anal. Chem. 76 (2004) 6085–6096.

[97] J.P. Shaffer, Multiple hypothesis testing, Ann. Rev. Psychol. 46 (1995) 561–584.

[98] O. Vorst, C.H.R. De Vos, A. Lommen, R.V. Staps, R.G.F. Visser, R.J. Bino, R.D. Hall, A non-directed approach to the differential analysis of multiple LC-MS-derived metabolic profiles, Metabolomics 1 (2005) 169–180.

[99] N.J. Serkova, M. Jackman, J.L. Brown, T. Liu, R. Hirose, J.P. Roberts, J.J. Maher, C.U. Niemann, Metabolic profiling of livers and blood from obese rats, J. Hepatol. 44 (2006) 956–962.

[100] M. Daszykowski, B. Walczak, D.L. Massart, Projection methods in chemistry, Chemomet. Intell. Lab. Syst. 65 (2003) 97–112.

[101] G. Schlotterbeck, A. Ross, F. Dieterle, H. Senn, Metabolic profiling technologies for biomarker discovery in biomedicine and drug development, Pharmacogenomics 7 (2006) 1055–1075.

[102] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J. Sanchez, M. Müller, PanelomiX: a threshold-based algorithm to create panels of biomarkers, Translant. Proteomics 1 (2013) 57–64.

[103] N.A. Obuchowski, L.M. Lieber, F.H. Wians, ROC curves in clinical chemistry: uses, misuses, and possible solutions, Clin. Chem. 50 (2004) 1118–1125.

[104] K. Soreide, Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research, J. Clin. Pathol. 62 (2009) 1–5.

[105] S.D. Walter, The partial area under the summary ROC curve, Stat. Med. 24 (2005) 2025–2040.

Herramientas y equipos analíticos

Analytical tools and equipments

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

1. Muestras

Los tipos de muestras biológicas que se han estudiado en la investigación desarrollada han sido suero (Capítulos 3, 4, 5 y 6), plasma (Capítulo 5), tejido intraperitoneal (Capítulo 5), aire exhalado condensado (Capítulo 7) y orina (Capítulos 6, 8 y 9).

2. Sistemas automáticos y/o continuos de preparación de muestra

En los Capítulos 3 y 4 de la Memoria se recoge el uso de un sistema automático, especialmente indicado para para llevar a cabo la SPE de forma automatizada: el sistema Symbiosis, que se usó para el pretratamiento de muestra en el análisis de eicosanoides (Capítulo 4) y para estudiar la influencia de un estándar interno isotópico en el análisis de un grupo de esos eicosanoides (Capítulo 3). Este sistema, además de trabajar en modo dinámico, permite la elución con la fase móvil cromatográfica de los compuestos retenidos. Trabajar a alta presión posibilita la conexión en línea de este sistema con el conjunto cromatográfo/detector, consiguiéndose así la automatización completa del método analítico. El equipo está compuesto de 5 módulos: un automuestreador refrigerado , de enorme interés

para muestras termoinestables; una unidad de extracción en fase sólida, dos bombas de alta presión dispensadoras de disolventes, de 2 mL de capacidad, y una bomba de HPLC.

3. Sistemas discontinuos de preparación de muestra

Se han empleado sistemas discontinuos se han empleado tanto en análisis orientado como no orientado. Dada la naturaleza no selectiva, semi- o no cuantitativa de las plataformas analíticas empleadas a lo largo de la Tesis para análisis metabólico global, se hizo especial hincapié en el desarrollo de métodos con mínima o nula preparación de muestra. Por tanto, se evitó usar sistemas continuos de pretratamiento de la muestra. En general, el pretratamiento de la muestra se basó en etapas de agitación, centrifugación, dilución o precipitación de proteínas. Por otro lado, la homogeneización de tejido intraperitoneal se llevó a cabo con un homogeneizador previa adición de cloruro de sodio y la extracción y preconcentración del analito (paclitaxel) y sus principales metabolitos en suero, plasma y tejido se realizó mediante extracción líquido–líquido con metil-*ter*butil-éter (MTBE), según se recoge en el Capítulo 5. La separación–determinación de los compuestos presentes en aire exhalado condensado mediante LC–QTOF requirió diferentes pretratamientos de muestra basados en la precipitación de las proteínas, la extracción en fase sólida usando μ-SpinColumn (un sistema de SPE diseñado para pequeños volúmenes de muestra y en el que se utiliza la fuerza centrífuga para hacer pasar los disolventes a través del cartucho) y la liofilización de la muestra (ver Capítulo 7).

4. Detección con separación cromatográfica previa

Los métodos desarrollados en la parte experimental de esta Tesis Doctoral se han basado en una separación cromatográfica mediante LC y la posterior detección basada en espectrometría de masas.

El análisis orientado, que se utiliza en los Capítulos 3, 4 y 5, —con el uso de estándares internos para el caso de los eicosanoides y de paclitaxel y sus metabolitos— se llevó a cabo mediante LC y posterior detección por MS/MS utilizando un cromatógrafo Agilent 1200 Series LC equipado con una bomba binaria, un desgasificador, un automuestreador y un compartimento de columna termostatizados, y un espectrómetro de masas Agilent 6410 de triple cuadrupolo con una fuente de ionización por electrospray (ESI) con tecnología "Jet Stream". En las plataformas dedicadas a la obtención del perfil metabolómico de muestras de orina (Capítulos 6, 8 y 9), suero (Capítulo 6) y aire exhalado condensado (Capítulo 7) se utilizó un equipo HPLC Agilent 1200 Series acoplado a un detector de masas de tiempo de vuelo de alta resolución, Agilent 6540, realizando la detección en el modo MS/MS. En todos los casos se empleó una columna de fase reversa C18 para la separación de los compuestos. También en todos ellos se usó el software MassHunter para la adquisición de espectros y el análisis cuali-cuantitativo de los datos.

5. 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 la optimización de los métodos analíticos como para el tratamiento de datos multivariantes.

El tratamiento de datos se realizó con distintos programas informáticos:

a) Alineamiento de entidades moleculares en análisis no orientado de los datos obtenidos utilizando un cromatógrafo de líquidos acoplado a un detector de tiempo de vuelo:

- 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 el software Mass Profinder de Agilent para realizar las etapas de extracción de entidades potenciales y su alineamiento en una única etapa.

b) 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.

-MetaboAnalyst: herramienta web (http://www.metaboanalyst.ca) que permite realizar estudios analíticos en metabolómica de alto rendimiento para el análisis e interpretación de datos metabolómicos.

c) Evaluación de la capacidad de predicción de metabolitos o paneles con potencial marcador: esta evaluación se ha realizado a través de gráficos de Random Forest y de las curvas ROC, que representan especificidad frente a sensibilidad. Para su obtención se han usado las herramientas gratuitas de acceso online Metaboanalyst (http://www.metaboanalyst.ca/) y pROC (http://web.expasy.org/pROC/).

d) Diseño de modelos de predicción y creación de paneles de marcadores: en este caso se emplearon tanto el MassProfiler Professional como Metaboanalyst, que hacen posible la creación de modelos por PLS-DA, como el software PanelomiX, que permite crear paneles de marcadores y evaluar su capacidad de predicción/discriminación.

e) Diseño de paquetes estadísticos: El software de análisis de datos de código abierto R a través de su interfaz gráfica R-studio permite diseñar paquetes estadísticos utilizando diferentes funciones.

6. Bases de datos

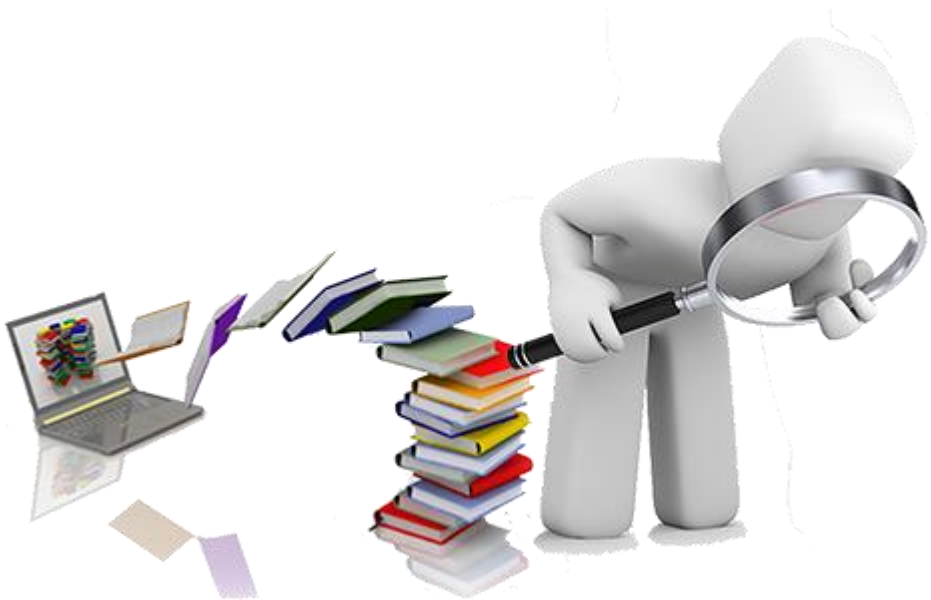
Aunque queda un largo camino por recorrer, existen varias bases de datos de metabolitos a disposición del usuario que contienen información para la identificación y caracterización de muchos de los compuestos presentes en diversas matrices biológicas. Entre ellas, las bases de datos 'Metabolites and Tandem MS Database' (METLIN), 'Human Metabolome Database' (HMDB) y 'MassBank' contienen información sobre la naturaleza de los compuestos, así como sobre su ionización y fragmentación en espectrometría de masas, lo que ha permitido la identificación de compuestos presentes en los biofluidos analizados (suero, orina y aire exhalado condensado) a partir de los datos espectrales obtenidos mediante espectrometría de masas en los capítulos 6, 7, 8 y 9 de esta Memoria. Además, se ha uutilizado la información biológica disponible en la HMDB y las rutas biosintéticas de la base de datos 'Kyoto Encyclopedia of Genes and Genomes' (KEGG) para la interpretación de resultados obtenidos en los Capítulos 7, 8 y 9.

Parte experimental

Experimental part

SECTION I

Bibliographical studies



An aspect of paramount importance when starting any research is an in-depth and critical knowledge about what has been done/published in the field of interest. This information requires getting inside each piece of information and mine what is enough assessed in it for subsequent comparison with similar/dissimilar results from other sources, obtainment of proper conclusions and proposition of subsequent steps to be developed. Enough rigor in doing these tasks may give place to a review publication in a journal, the impact of which will be a function of the importance of the target subject. This section of the Thesis-Book contains the two review publications the PhD student wrote during the development of her Thesis.

Before working with urine samples to search for metabolomics prostate cancer biomarkers, the author of this Thesis-Book developed an in-depth study of the literature in dealing with preparation of this biofluid prior to its metabolomics analysis by high resolution mass spectrometry. Thus, she classified the steps involved in the complex process into urine sample-pretreatment steps that encompass from sampling to storage (*viz.,* sampling time and collection, quenching, preservative addition, volume correction, pH adjustment, deproteinization —by centrifugation, normal filtration, ultrafiltration, ultrasound-assisted filtration and freezing/thawing cycles), and urine sample-preparation steps from thawing to insertion into the separation device and/or detector (*viz.* dilution, enzymatic hydrolysis, sample clean-up and preconcentration —by solid-phase extraction, liquid–liquid extraction, evaporation, and lyophilization). All these steps are critically discussed in Chapter 1.

Prior to enter into the complex world of exhaled breath condensate to work with it, the mandatory mining into the literature on the analytical methods involving this biofluid and lung cancer gave to the PhD student a wide overview on: (i) the research based on each of the great omics looking for lung cancer

biomarkers; (ii) the design of chemical-sensor systems to detect lung-cancer breath biosignatures; (iii) the use of trained dogs to detect this cancer; (iv) the proposal for research looking for detection of the breath components at the very low concentration detected by dogs. Chapter 2 discusses all these aspects.

CHAPTER 1

Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis

TrAC, Trends in Analytical Chemistry 41 (2012) 75–85

Trends in Analytical Chemistry

41 (2012) 75–85

Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis

M.A. Fernández-Peralbo [a,b], M.D. Luque de Castro [a,b]

[a]*Department of Analytical Chemistry, University of Córdoba, Annex Marie Curie Building, Campus of Rabanales, 14071 Córdoba, Spain.*

[b]*Maimónides Institute for Research in Biomedicine of Córdoba (IMIBIC), Reina Sofía University Hospital, University of Córdoba, 14071 Córdoba, Spain.*

Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis

M.A. Fernández-Peralbo, M.D. Luque de Castro

Abstract

Steps preceding introduction of the analytical sample into individual separation equipment and/or mass-spectrometry (MS) detectors for targeted or untargeted metabolomics studies have not generally received the required attention, but particularly in dealing with urine samples. These steps have been classified into urine sample-pretreatment steps (*i.e.*, those encompassing from sampling to storage) and urine sample-preparation steps (*i.e.*, from thawing to insertion into the separation device and/or detector). We first discuss the steps constituting the former group (sampling time and collection, quenching, preservative addition, volume correction, pH adjustment, deproteinization –by centrifugation, normal filtration, ultrafiltration, ultrasound-assisted filtration and freezing/thawing cycles), and of those forming the latter group (dilution, enzymatic hydrolysis, sample clean-up and preconcentration –by solid-phase extraction, liquid–liquid extraction, evaporation, and lyophilization). We then reach conclusions that open a door to future, necessary research to improve the field of metabolomics on urine samples.

Keywords Mass spectrometry; Metabolomics; Preconcentration; Sample preparation; Sample pretreatment; Separation; Solid-phase extraction; Targeted analysis; Untargeted analysis; Urine sample

1. Introduction

Metabolomics can be defined as the comprehensive analytical approach to study all low-molecular-weight species (typically defined as <1000 Da or <1500 Da) present in a given biological system of interest. Recently, Ryan et al. [1] seem to lean towards a definition that encompasses characterization and quantitation of all metabolites in a sample, but with the qualification that metabolomics considers both spatial resolution (due to cellular compartmentalization of metabolism) and temporal changes in response to environmental stimuli (the latter should be extended to xeno-stimuli).

At present, metabolomics analysis encompasses different strategies depending on the situation under study [2], namely:

(1) targeted analysis [3,4], which aims at qualitative and quantitative study of one metabolite or, more frequently, a small group of chemically similar metabolites;

(2) global metabolomics profiling [5,6] or untargeted analysis, which allows detection of a broad range of metabolites by using a single analytical platform or a combination of complementary analytical platforms –based on GC–MS, LC–MS, capillary electrophoresis (CE)-MS or NMR– to obtain a comprehensive profile of the metabolome;

(3) metabolomics fingerprinting [7,8], a high-through-put, rapid methodology for analysis of biological samples that provides fingerprints for sample classification and screening; and,

(4) metabolomics footprinting, which aims at the study of metabolites in extracellular fluids, also known as exometabolome or secretome [9,10].

Obviously, the complexity of sample preparation for each strategy is different increasing from (4) to (1). The most recent trend leans to simplification by establishing distinction only between targeted and untargeted analyses, which require quite different sampling and sample-preparation approaches [11].

Untargeted approaches establish that the significant metabolites are by definition unknown prior to analysis, as are their physico-chemical characteristics, which usually encompass a wide variability range. The main aims are therefore to keep intact the sample as much as possible and to avoid fractionation and/or losses; nevertheless, potentially important diagnostic differences might remain undetected due to confounding factors (*e.g.,* instability of the analytical system, chemical or instrumental noise, and biochemical differences that are unrelated to the condition under study). By contrast, in targeted analysis, the physico-chemical characteristics of the metabolites are known and an exhaustive separation of them from the matrix is usually required for interferent-free quantitation. An ideal sample-preparation method for untargeted metabolomic analysis of biological samples should be:

(1) unselective;

(2) simple and fast with the minimal number of steps;

(3) reproducible; and,

(4) incorporate a metabolism-quenching step, if required [5].

In other words, the overall goal of the sample-preparation procedure to implement this strategy is to transform the sample reproducibly into a format that is compatible with the given analytical platform while maintaining the original metabolite composition of the sample as much as possible. By contrast, an ideal sample-preparation method for targeted analysis should:

(1) be as selective as possible to the metabolite(s) under study;

(2) be reproducible;

(3) allow implementation of preconcentration and/or clean-up steps, if required; and,

(4) incorporate a metabolism-quenching step, if required.

One of the critical steps in metabolomics is sample selection, as the results generated will depend on its suitability. Specifically, plasma, serum and urine have

traditionally been used for prognosis or diagnosis of many diseases, as they are easily collected, reflect directly the global state of an individual and allow the biological response to drug therapy to be monitored. In addition, plasma and urine provide complementary information about the state of an organism [12]. Thus, plasma gives an "instantaneous" read-out of the metabolic state at the time of collection, and its composition directly reflects catabolic and anabolic processes occurring in the whole organism. However, urine provides an "averaged" pattern of easily-excreted polar metabolites discarded from the body as a result of catabolic processes. A typical biological fluid sample is expected to contain >1000 metabolites even under the most conservative accounts, with the number expected to increase further as analytical methods become more sensitive. However, the human urinary metabolome is potentially vaster, if one includes all endogenous metabolites and exogenous entities that can be in it, accidentally or not. Urine is composed of over 95% water, plus sodium, ammonia, phosphate, sulfate, urea, creatinine, proteins and products processed by the kidney and liver, including drugs and metabolites. Among the analytical advantages of urine over serum are that the urine:

(1) it can be obtained in large quantities by non-invasive sampling, and sampling repetition is not a problem; and,

(2) it requires less complex sample preparation because it contains lower protein amounts because of the relatively small size and higher thermodynamic stability of urinary peptides and proteins.

By contrast, the number of conjugated compounds can be higher than in serum [13,14].

Once the type of fluid has been selected, the choice of a sample-preparation method plays an extremely important role in metabolomics, because it affects the observed metabolite profile and data quality, as a result. Without a deep understanding of the capabilities and the limitations of the sample-preparation method used in a given study, the accuracy of biological interpretation of collected data may be compromised, leading to erroneous conclusions. Duportet *et al.*

showed that the extraction of the same quenched samples using four different extraction methods could result in significantly different metabolite levels and contradictory biological interpretation of active metabolite pathways [15].

Because the use of urine in the metabolomics field is not as frequent as would be foreseeable, this article calls attention to the importance of this easy-to-obtain biofluid because of its high potential, to which researchers in metabolomics should give an in depth view for proper exploitation. A wider use of urine in the metabolomics field constitutes a challenge, as, traditionally, urine has been used to measure one or two components in a wide range of areas (pharmaceutical, biochemical, forensic and therapeutic drug monitoring), far from the complete characterization of the entire metabolome regardless of molecular size, which includes all endogenous metabolites and exogenous entities and covers a wide dynamic range from nM, in the case of acylcarnitines, to mM for glucose [16]. This difficult task requires analytical techniques that are sensitive and specific, capable of measuring all analytes, even at trace levels, an aspect that has been widely discussed in reviews on the subject [17,18]. Not so much attention has been paid to sample preparation of urine, which, as with other samples [19], constitutes a critical step and the key to proper exploitation of cutting-edge analytical instruments used subsequently.

The different steps in urine-sample preparation encompass pretreatment (*i.e.,* basic, preliminary operations that are mainly developed before storage and include sampling, quenching, centrifugation or filtration, and addition of stabilizers, if required) and sample preparation [including all steps preceding introduction of the analytical sample either into the detector or the equipment for individual separation (for chromatographic or capillary electrophoretic steps)]. When this equipment is involved in the overall method, it is so closely related to the detector that sometimes distinction between the tasks in each step may be difficult to distinguish clearly [20].

2. Urine sampling and sample pretreatment

Depending on the aim of the research, both sampling time and sampling procedure can be very different, as can be subsequent steps.

2.1. Sampling time

Urine samples are usually collected as random samples [21], timed samples [22] or 24-h samples [23]; even longer sampling times may be used for pharmacokinetics studies [24].

Random sampling is usually performed at any time of day (*i.e.,* diurnal variation of excretion is not taken into account), while timed samples are needed to study time-related trends to catalogue metabolites with high diurnal variation in different species and to aid in the search for true biomarkers. This sampling mode is mandatory when the aim is excretion monitoring (*e.g.,* a given food or drug, cosmetic, or contaminant in nutrimetabolomics and xenometabolomics studies, respectively). In this case, control or blank (*viz.,* urine before intake) is desirable for proper comparison of the situation before intake with respect to the given analyte(s). Then, kinetic sampling is generally preferred prior to targeted analysis (if the single evolution of the target compounds is planned to be monitored), or untargeted analysis (if the overall effect after intake is pursued).

A typical example of timed samples is the nutritional study based on targeted analysis of phenols from olive oil through urine [25], in which the samples were collected in special urine-collecting plastic bottles during the wash-out period (first void spot urine in the morning) and at three different timed collection periods (0–2, 2–4, and 4–6 h) after oil ingestion. Storage at −20 °C followed measurement of urinary volume. No pH measurement or filtration was carried out before analysis.

However, 24-h sampling collection is preferred to eliminate large variability in metabolite profiles obtained in shorter collection periods, when an overall status of the individual is the aim of the analysis. Less common is sampling urine day and night that allows researchers to deduce that there may be different gut microbiota activities in healthy people during day and night urine, clearly reflected by the

presence of given biomarkers. The results of this study indicated the obvious change in structure and activity of gut microbiota with the circadian rhythm [26]. Therefore, some metabolites are related to circadian rhythms and them plus dietary components need to be differentiated with time-of-day urine collection [27], detail that is lost with 24-h sampling. Nevertheless, although 24-h samples are cumbersome, the excretion picture they provide is generally more complete.

2.2. Urine collection

Targeted analysis requires a procedure for only sample collection and a container that does not degrade the compounds of interest. Unlike serum or plasma containers, those for urine are not endowed with special characteristics or reagents, usually bare polypropylene containers of the required volume suffice. As the physico-chemical characteristics of the analytes are known, to keep them under optimal conditions is an easy task.

It is more complicated to prevent deterioration in untargeted analysis, as the behavior of a number of compounds in the sample is unknown. In this case, extreme care is sometimes adopted, as is the case of providing the subjects with insulated ice packs in which they were asked to store the samples immediately until they were received in the laboratory for proper storage, if required, prior to analysis [28].

An additional shortcoming is the need to avoid losses of metabolites due to non-specific adsorption to surfaces of containers in which the samples are collected, stored or processed. Sample treatment with additives to increase metabolite solubility and/or to minimize interaction with container surfaces is essential in this case. The most common additives are surfactants, as shown in a recent study from Silvester and Zang [29]. The authors showed that compounds in urine with lipophilic character (*e.g.,* quaternary amines in the absence of any additive) underwent up to 35% adsorptive losses –phenomenon that can be avoided in the presence of surfactant additives. Nevertheless, a side effect of surfactant addition can be ionization suppression in the subsequent MS analysis, which should be minimized by using an isotopically-labeled internal standard (IS).

2.3. Urine quenching is a doubtful necessary step

One of the most contradictory metabolomics requirements is the need for metabolism quenching, which is routinely incorporated in cell and plant metabolomics, but it is not clear that it is necessary in urine samples. The rationale that metabolism quenching is unimportant for the majority of biofluid studies is accepted in dealing with stable metabolites, but such approach is questionable for labile metabolites prone to degradation or conversion.

Metabolism quenching, when applied, aims to stop metabolic processes through use of low temperatures (cold solvent addition, freezing in liquid nitrogen), addition of acid or fast heating [30,31]. To complicate matters even further, application of a quenching step can inadvertently cause degradation or loss of some metabolites —*e.g.,* acidic treatments can cause severe degradation of some metabolites, poor overall metabolite coverage and poor compatibility with the subsequent MS methods [32].

Supporters of urine quenching (*e.g.,* van de Merbel [33]) reported that, under ambient conditions, storage of urine can cause depletion of easily oxidizable species (*e.g.,* ascorbic acid, adrenaline, noradrenaline and dopamine). Also, they observed that enzymatic conversion of various species occurred rapidly after sample collection (*e.g.,* conversion of cytidine into uridine or adenosine into inosine). Monitoring of substrates for enzymes (*e.g.,* lutamyltransferase, alkaline phosphatase and urease) has also provided useful insight into residual enzymatic activity in urine [34]. When a necessary quenching step is omitted, residual enzymatic activity can significantly change metabolite levels prior to analysis, causing poor correlation between the observed metabolome and the true metabolome at the sampling time.

However, Dunn *et al.* [35] examined urine stability in a pilot metabolomic study of urine (and serum) samples from 40 healthy volunteers. Samples were either processed immediately by freezing at −80 °C or stored at −4 °C for 24 h before being frozen, and then thawed and profiled by gas chromatography time-of-flight mass spectrometry (GC–TOF–MS) revealing more than 700 unique

metabolite peaks. Over 200 peaks were detected in any one sample. Differences between samples stored in the two ways were assessed by Wilcoxon rank sum test and Principal Components Analysis (PCA) methods, and no statistically significant differences in the metabolic profile of samples stored under the two conditions were found. It is worth mentioning that analytical variance (from replicate analyses of the same sample) was of the same magnitude as the variance between samples stored under the two conditions. Extension of this study to analysis of fresh samples should be very convenient in order to know if degradation occurs fast enough to remain undetected, as is the case of well-known metabolic processes with time scales <1 s (for example: ATP, glucose-6-phosphate). The accurate determination of such metabolites necessitates an extremely rapid quenching step, which can be very difficult to implement for biological samples within an appropriate time scale [2].

2.4. Preservative addition

Although protease inhibitors are often added to biological samples to preserve protein forms intact, their use in urine samples is doubtful due to the variable pH, dilution, and denaturation of proteases in urine [33,34]. The addition of preservatives (*e.g.*, boric acid or sodium azide) or inclusion of filtration have been used to prevent bacterial contamination. Boric-acid addition was avoided in the urine collected in the UK Biobank because of the potential formation of chemical complexes in the samples [36]. Despite azide not forming protein complexes, its use was also carefully considered but rejected because of the difficulty in maintaining consistency and quality of sample handling in the assessment centers, the potential for variable dilutions factors through different initial sample volumes, health and safety concerns over use of this toxic compound and the potential impact on unknown metabolites [37]. A recent comparative study of sodium azide and 0.20-μm filtration showed the latter to be superior for bacterial removal [34].

2.5. Volume correction by normalization

Urine volume can vary widely based upon water consumption and other physiological factors. As a result, the concentrations of endogenous metabolites in urine vary widely and normalizing for these effects is necessary, particularly in random sampling. Approaches based on urine volume, creatinine concentration, osmolality, and components that are common to all samples ("total useful MS signal") are strategies successfully used for normalization in both targeted and untargeted urine analysis. In general, volume correction is useful if the biomarker being monitored is variable in its time of excretion (*e.g.,* glycosaminoglycans) [38].

There is no universally accepted procedure for correction of the urine volume, as each procedure has restrictions.

Measurement of protein/creatinine ratio is one of the most frequent normalization procedures. It was shown to be useful to predict 24-h protein excretion with reasonable accuracy in patients with low levels of protein excretion, but is unreliable in patients with high protein excretion [39].

Osmolality, representative of the concentration of solutes in a fluid, is determined by measuring freezing-point depression. Used for normalization of urine in clinical settings, it has been reported to offer advantages over normalization by creatinine concentration [40].

A recent approach is MS total useful signal (MSTUS), which uses the total intensity of components that are common to all samples, thus avoiding xenobiotics and artifacts that would be inappropriate for measures of urine concentration [40].

Comparison of the three approaches leads to the recommendation to use two of them in order to facilitate detection of statistically significant changes in the endogenous metabolite profile in urine samples. Together with protein/creatinine ratio, osmolality can be used in cases where it is possible to get this independent measurement, and MSTUS or a related approach can be used when osmolality measurements are inconvenient or invalid [40].

It is worth pointing out that storage and transport of urine samples for doping analysis is not subjected to specific conservation protocols (*e.g.,* cooling,

quenching, or volume correction) [41], so it is assumed that no changes of the targeted metabolites are produced from sampling to analysis –an assumption that seems difficult to check exhaustively for all types of dopants.

2.6. pH adjustment of urine samples

Controversial results on pH adjustment of urine samples can be found in the literature, mostly due to the different purposes for including this step. Sediment dissolution, improvement of retention in solid-phase extraction (SPE) or better transfer in liquid–liquid extraction (LLE) are, or are not, favored by pH adjustment, depending on the target analytes, the pH adopted and the type of sediment to be dissolved, and sorbent or organic phase used [42]. By way of example, optimal retention and clean-up of β2- agonists from urine on Bond-Elut Certify cartridges prior to GC–MS was achieved at pH 9.5 [43]. By contrast, the best retention of β-blockers, hormones and lipid regulators on Oasis-HLB cartridges, also prior to GC–MS, was obtained at pH 7.0 [44].

Enzymatic reactions (*viz.,* hydrolysis to verify the presence of the conjugated forms of polyphenols and their metabolites [45]) require the pH to be adjusted for optimum performance of the biocatalyst.

2.7. Separation of a precipitate (deproteinization)

Deproteinization, an essential step for blood serum or plasma because their high protein content (60–80 g/L), is not considered mandatory in urine samples because the protein concentration in urine is significantly less (0.5–1.0 g/L), but still considerable [46]. When this step is applied –as a sample pretreatment or a preparation step– 1:5 dilution in 50% acetonitrile has been used [47]. Nevertheless, it worth pinpointing that scant information on sample preparation is sometimes given, as in the study by Laiakis *et al.*, in which nothing was said about urine sampling or storage sample, but only that deproteinization by 1:5 dilution in 50% acetonitrile with 2 mM debrisoquine sulfate and 30 mM 4-nitrobenzoic acid as internal controls was made before analysis [48]. By contrast, detailed information is given on UPLC–TOF-MS analysis and data treatment. The

sample–precipitant medium ratio varied widely in the range 1:1–1:9, usually with addition of the IS in the precipitant medium [49].

Centrifugation or filtration is also a usual step in dealing with urine samples to remove materials in suspension (*e.g.,* calculi, cellular components or proteins) (in this case after appropriate chemical treatment). Conventional filtration is performed on a cellulose membrane (0.45 μm pore size), prior to addition of a stabilizer (*e.g.,* sodium azide). Mild precentrifugation (1 000–3 000 rcf for 5 min) immediately after blood sample collection is recommended in order to remove cellular components effectively and to minimize cell breakage that can occur with high centrifugation speeds. By contrast, this recommendation has not yet been widely adopted in urine metabolomics, so various study designs still employ high-speed centrifugation. Recently, it was found that 0.20 μm filtration is superior to centrifugation or sodium-azide addition in preventing bacterial growth during storage [34].

Ultrafiltration is a simple sample-pretreatment procedure, by which a sample is filtered through a special filter, which only allows molecules of certain molecular weight to pass through, and the filtration is achieved through application of pressure or through centrifugation. Filters with molecular weight cut-offs of 3 kDa, 10 kDa, and 30 kDa are commonly available. The use of a 3 kDa filter can physically separate small molecular-weight species from proteins or other macromolecules present in a given sample, and is thus useful for metabolomics. For example, for erythrocyte extraction, Darghouth *et al.* proposed the use of a cellulose column to separate red blood cells from leukocytes, centrifugation, cell-pellet extraction using a boiling water method for 3 min, followed by centrifugation and two-step ultrafiltration, first using a 30 kDa device followed by a 10 kDa device for 20 min at room temperature [50]. While ultrafiltration has scarcely been applied to urine samples, ultrasound filtration (known as filterless filtration [20]) has never been applied to urine samples.

2.8. Stability of urine samples using storage and freeze–thaw cycles

Sample-storage conditions in metabolomics in general, particularly urine samples, had not previously received the deserved attention (*e.g.*, compared with drug bioanalysis, in which the investigation of the effect of sample-storage conditions and freeze–thaw cycles on the determination of analytes forms is a basic part of the analytical method validation required by regulatory authorities, such as the US Food and Drugs Administration (FDA) [51]). Nevertheless, as metabolomics matures, sample storage is receiving proper attention.

Inadequate metabolism quenching, employed in most common strategies for blood and urine samples, puts a great emphasis on sample storage at low temperatures. Storage at −80 °C is usually preferred, but −20 °C is also used in targeted analysis when this temperature is low enough to ensure stability of the target metabolites.

On long-term stability of urine samples for untargeted analysis, it is recommended to avoid freeze–thaw cycles and rapidly freeze and store pre-aliquoted samples at −80 °C to minimize potential degradation as much as possible [52] – despite aliquoted samples seeming to be stable in the autosampler at 4 °C for up to 20 h [53]. Nevertheless, storage at even −80 °C was found to influence metabolite levels and pH, so some authors recommended a 1-week minimum for storage to achieve consistent degradation of urinary components across the cohort [54]. In this way, comparison across samples is possible, despite the analyzed metabolome not entirely reflecting the true metabolome, as potential metabolites of interest may not be present at the time of analysis. In addition, peaks due to degradation may be misidentified as possible features of interest.

In cases where the amounts of sample may be limited and irreplaceable, some idea of extended sample stability in the autosampler is essential. Samples may also need to be analyzed more than once, sometimes separated by several months. In this situation, knowledge of sample stability both in the instrument during analysis and on storage is key for the production of valid data. With this aim, Gika *et al.* [11] developed a urine-stability study prior to untargeted analysis that encompassed short-term stability in an autosampler at 4 °C, stability of frozen samples (storage

at both –20 °C and –80 °C) and in freeze–thaw cycles. The authors did not find obvious differences in the profiles of urine samples stored at–20 °C and –80 °C by either LC or UPLC–MS, at least on the basis of PCA. However, this does not necessarily mean that, after 6 months in the freezer, the samples are identical to the original samples at the time of collection, but that, to the level described by PCA, as it would be in a typical metabolomic study, they do not differ. In robust models, moderate differences in the amounts of a few analytes would not alter the relative projection of the samples. This means that although the relative content of some compounds may change on storage of the urine samples, the position of the samples in the scores plot may be unchanged. This conclusion is in general agreement with Lauridsen *et al.* [55], who applied similar PCA to urinary NMR data, and also with Maher *et al.* [12], who applied both PLS and OPLS-DA to analyze NMR data for long-term stability of human urine. All three groups recognized that biological variability is the major source of variation and observed minimal sample differentiation on the NMR profiles for 3 or 6 months of storage. Saude and Sykes [56] reported, in a study based on targeted analysis of the urine of two individuals, that urine that underwent eight freeze–thaw cycles (–80 °C) had an intermediate degree of metabolite change compared to urine stored at room temperature or frozen at (–80 °C).

In general terms, it can be said for untargeted metabolomics studies that: (1) storage of human urine for up to 6 months at –20 °C or below is suitable for metabolomics studies;

(2) the number of freeze–thaw cycles (up to 9) does not seem to affect sample integrity; and,

(3) short-term storage in a fridge (0 °C to –4 °C) or a cooled autosampler for up to 48 h still seems to provide useable samples that provide meaningful results.

Nevertheless, these statements do not mean that all the components of the sample are stable under the given conditions, so rigorous specific stability studies are essential for targeted analysis.

3. Urine-sample preparation

This key part of an analytical method usually comprises different steps, all under the umbrella that they take place in between thawing the sample and inserting it into the instrument or the equipment for providing individual identification and/or quantitation in reaching the detector. So, after development of some or all steps described in Section 2, one or several steps in the present section can be essential, advisable or unnecessary with targeted or untargeted analysis in metabolomics research involving urine samples.

Direct sample insertion, also known as direct infusion, into the detector (usually MS detectors and for untargeted analysis), with sample dilution [57,58] or without dilution [65,66], involves shortcomings mainly related to signal suppression. Application of sample-dilution approach, also known as "dilute and shoot", is only feasible when the levels of targeted metabolites are relatively high and the matrix components do not co-elute or otherwise interfere with the ionization of the analytes [57]. There have been a number of studies on application of LC–MS in direct urine analysis by "dilute and shoot". Most used the column-switching approach [58] and only a few used direct urine injection with no sample preparation other than dilution [59,60]. Except for the above strategy, urine-sample preparation can require one or several of the following steps.

3.1. Enzymatic hydrolysis

Biotransformation processes involve changes in chemical and physical properties of endogenous and exogenous compounds, which, in the case of "phase II reactions" [61], yield hydrophilic products, thereby greatly promoting excretion of the target compounds, mainly in urine. Examples of endogenous compounds undergoing this type of biotransformation are hormones –*e.g.,* female steroid hormones, which are present in urine almost exclusively as conjugated metabolites (*e.g.,* glucuronides, sulfates, diglucuronides, disulfates and sulfoglucorodines) [62]. An example of exogenous compounds are sunscreen agents –*e.g.,* 2-ethylhexyl 4-(N,N-dimethylamino)benzoate (EDP)–, present in urine as glucuronide and sulfate derivatives [63]. Also in the vegetal kingdom,

biotransformation processes yield conjugated compounds, so, food intake also gives rise to subsequent excretion of both conjugated and non-conjugated compounds, depending on the subsequent transformation into the body, as is the case with phenols analyzed in urine after virgin-olive-oil intake for understanding their metabolism in humans [45].

Conventional enzymatic hydrolysis of the conjugated compounds by the proper biocatalyst (*e.g.,* β-glucuronidase and/or sulfatase) requires different times, depending on the target aglycone. While EDP required overnight incubation [63], female hormones required 18-h incubation [64] and olive phenols only 4-h incubation [45]. One alternative to enhance enzymatic hydrolysis is assistance with ultrasonic energy [64], still an emerging field. The mechanism through which ultrasound enhances enzymatic reactions has not been elucidated; nevertheless, cavitation seems to promote the reaction rates by increasing analyte transport from the bulk to the enzyme surface, which leads to effective contact between substrate and active sites of the enzyme [65]. The time for enzymatic hydrolysis of female steroid hormones in urine has been dramatically shortened (from 18 h to 30 min) by applying ultrasound, while the reaction efficiency was similar, or even higher, for all the hormones under study [64]. No modification caused by assistance with auxiliary energy, as compared to that obtained with the conventional protocol, was observed. Similar results (*i.e.,* shortening of hydrolysis time from incubation overnight to 50 min with no modification as compared with the conventional enzymatic step) was found for EDP and metabolites in urine [63].

Studies on acceleration of hydrolysis steps in the absence of biocatalysts using ultrasound of variable frequency (tunable transducers) is a pending goal of the authors. Success in this study would involve reduction of time and costs of this step. Addition of an IS, if required, is usually performed at this stage [45].

3.2. Sample clean-up and preconcentration

A step for sample clean-up and/or preconcentration is usually essential in dealing with biological samples in general and urine in particular prior to targeted analysis, but it can also be required for untargeted analysis. Selection of a specific

technique for this task depends upon the analytes and their concentrations, the sample matrix, the sample size and instrumental technique. Among the techniques used to implement this step the most common is SPE in its different formats, but also LLE to a lesser extent. Much less frequent is the use of evaporation for determination of very diluted metabolites, as potential interferents are also concentrated by this step. It is worth pointing out that addition of an IS, if required, is usually performed after this step when no enzymatic step is previously developed.

3.2.1. Solid-phase extraction

This very frequent step in metabolomics analysis can be implemented in manual, semi-automatic or automatic ways. In the automatic case, on-line coupling with the individual separation step (chromatographic or CE) can be implemented.

Depending on the sample volume applied on the cartridge–volume of eluant ratio, significant pre-concentration of the target metabolites can be achieved. Proper selection of the sorbent phase sometimes requires compromise between selectivity and metabolite coverage.

The best option for complete use of the eluate from a cartridge is on-line connection with LC for subsequent individual separation. Commercial equipment with this aim usually works at high pressure, so the pure [66] or modified [67] mobile phase is also the eluant to remove the retained compounds from the sorbent and lead them to the head of the chromatographic column, so all participate in the subsequent identification and/or quantitation step. In addition, there is no metabolite–atmosphere contact during the overall process, so alteration by light and/or air is avoided. This has been the way selected for clean-up and preconcentration of free and glucuronide-conjugated female steroid hormones in urine prior to LC–EI/tandem MS [68].

Comparative studies of the performance of SPE *versus* precipitation with methanol or with acetonitrile similar to those on human plasma [69] have not been carried out on urine. Also the use of complementary SPE for enhancing detection

coverage in urine is a pending goal, as is the ability to enrich low-abundance metabolites routinely missed by other methodologies.

Solid-phase microextraction (SPME) is also used for sample preparation, mainly for clean-up, but also for preconcentration despite the non-exhaustive extraction it provides. For the latter reason, it is more appropriate for global metabolomics profiling in which detection of diluted urine components is not the goal. Its easily automated handling makes this alternative very useful in routine, untargeted metabolomics analysis [70]. Also, targeted analysis of urinary volatile organic metabolites as potential cancer biomarkers has been carried out by SPME in combination with GC–MS [71].

3.2.2. Liquid–liquid extraction

Despite it being commonly established that LLE is to be avoided in metabolomics in general [72], this step is used in dealing with urine samples, mainly for separation of non-polar or low-polar analytes. Mostly, the preconcentration factor achieved is not high. Ethyl acetate or ethyl acetate–diethyl were recently used for extraction of synthetic cannabinoids after enzymatic deconjugation [73].

Meanwhile, for monohydroxylated polycyclic aromatic hydrocarbons (OH-PAHs), the preferred immiscible solvent was a pentane–toluene mixture [74], which significantly improved the sensitivity of the method, compared with the use of SPE. Despite the generally wider use of SPE in clinical analysis, this sample clean-up is not good enough to deal with the extremely low level of OH-PAHs, due to matrix interferences.

3.2.3. Partial, total evaporation or lyophilization for reconstitution

Several studies incorporate some of these steps as part of sample preparation –steps that can be developed in different ways depending on the final aim of the analytical process.

The main objectives pursued in implementing partial, total evaporation or lyophilization are:

(1) pre-concentration;

(2) change of solvent to ensure compatibility with the individual separation–detection step, particularly with MS detectors;

(3) improved sample stability during storage; and,

(4) subsequent analysis by ambient MS.

Nevertheless, the benefits of these steps and whether they help to improve metabolite coverage have not been well established in untargeted studies. Potential losses of hydrophobic and volatile species and increased ionization suppression due to pre-concentration may result in lower metabolite coverage. Furthermore, the use of higher strengths of organic solvents to improve resolubilization of hydrophobic metabolites eliminates the second advantage, while the long times needed to perform this step reduce sample throughput. The pending question is therefore whether this step is truly necessary and beneficial in metabolomics analysis.

Partial evaporation is sometimes used in untargeted analysis in order to increase metabolomics coverage. This step is unusual in targeted analysis –at least a subsequent separation step is developed–as partial evaporation yields more concentrated target metabolites, but also more concentrated interferents.

Total evaporation usually requires a subsequent step for reconstitution, which is implemented differently depending on the final aim of the analytical process:

(i) Untargeted analysis makes essential total solubilization of the dry residue, which is not achieved with any reconstitution solvent, as some solvents yield poor solubilization of some metabolites [75]. The composition of the reconstitution solvents used varied greatly from 5% acetonitrile [75], to 50% and 75% acetonitrile [76], and even 100% acetonitrile [77]. For reconstitution, Lawton *et al.* used 10% methanol with 0.1% formic acid [78], while Evans *et al.* tested 100% aqueous and 10% methanol [75]. Poor peak shape was observed for early eluting compounds using 10% methanol, and both methods failed to resolubilize relatively

hydrophobic species (*e.g.,* diacyl-glycerols, triacylglycerols or diacylphospholipids). The use of acidic (pH 3) and basic (pH 8) reconstitution solvents enhanced the detection of ionizable compounds [75]. By contrast, Bruce *et al.* advocated the use of 80% methanol as reconstitution solvent, but used very low injection volumes (2 μL) to ensure good peak shape for early eluting compounds [79].

(ii) Targeted analysis requires selective reconstitution of given metabolites, which may involve partial reconstitution of interferents –probably creating the necessity of a subsequent selective separation step– or incomplete reconstitution of target metabolites, giving rise to erroneous quantitation.

(iii) Separate comments cover ambient MS, a rapidly growing field that provides fast, direct analysis of solid-sample surfaces or liquid samples introduced on a suitable surface [80], which has been successfully applied to targeted analysis of urine samples. Different ambient ionization MS techniques –*e.g.,* desorption electrospray ionization (DESI) [81], direct analysis in real time (DART) [82], and desorption atmospheric pressure photo-ionization (DAPPI) [83]– have been used in dealing with urine samples. The very simple preparation step (application of urine in sample spots of few μL and either left to dry or add a reagent [81]) allowed development of fast methods that are not exempt from interferents, mainly ionization suppression [80,82]. A way to overcome interferent effects in some instances is *in situ* derivatization [81].

3.3. Use of internal standards

A step characteristic of sample preparation is the addition of ISs as a way of avoiding or minimizing ion-suppression/enhancement effects caused by co-eluting compounds (*e.g.,* drugs, metabolites, matrix components, impurities, and degradation products). This step is also desirable when total evaporation and reconstitution are performed, and essential in partial evaporation steps prior to targeted analysis. These effects are generally considered as the Achilles heel of MS analysis of biological samples, particularly urine samples [84]. Different approaches have been suggested to reduce and/or to correct for the effects (*e.g.,* to optimize sample preparation or chromatographic separation, to split the LC-flow

after chromatographic separation and before MS analysis, to use negative ionization mode or APCI instead of ESI, to use micro or nano LC, sample dilution, decrease of sample injection volume or inclusion of stable-isotope labeled ISs – SIL-ISs). Combinations of two or more of the approaches listed above have also been suggested, mostly including SIL-IS [85].

The most common (and usually the less expensive) SIL-ISs are deuterated forms of the target analytes, which in some cases also allow sample preparation to be simplified by avoiding the SPE step [86]. Several substitutions of ^{1}H by ^{2}H in the target molecule are desirable to ensure the presence of the isotope in the monitored fragment. Tsikas has reported the *de novo* synthesis of deuterated SIL with 3-to-7 ^{2}H atoms [87] for target amino acids. The author found the approach especially useful for amino acids for which no SIL analogues are commercially available.

Nevertheless, in many LC separations, a ^{2}H-labeled IS has slightly different retention time from the analyte. To show the different behavior of deuterated IS, compared with IS containing heavier isotopic elements, Berg and Strand compared the use of ^{13}C- and ^{2}H-labeled ISs [85] for UPLC–MS/MS analysis of drugs (amphetamine and methamphetamine) in urine samples. The ^{13}C-labeled ISs co-eluted with their analytes under different chromatographic conditions, while the ^{2}H-labeled ISs and their analytes were slightly separated. Improved ability to compensate for ion-suppression effects, increased linear range and a better correlation between UPLC–MS/MS analyses and GC-MS analyses of unknown samples were observed when the ^{13}C-labeled ISs were used, with correction of the ion-suppression effect up to 70–80%. Unfortunately, there are few ^{13}C-labeled IS commercially available.

A simple procedure (possibly not always with good enough performance) has been proposed as a substitute when SIL-ISs are not available, as is the case with recent synthetic cannabinoids (SCs). It involves using a series of urine samples produced by a drug-naïve individual, who has smoked a given amount of a herbal blend known to contain SCs. The absolute amount of compounds present in the

herbal blend used was unknown, so it was impossible to state the exact dose. Urine was collected prior to use (0 h), and up to 65 h after smoking (a total of 15 samples). The levels of SCs and those of creatinine were determined by LC–MS/MS and the data obtained as the ng SC/mg creatinine suggested that, at an LLOQ of 100 pg/mL, a detection window of 2–3 days is achievable, following a single use of a blend containing SC [73].

4. Conclusions

The steps prior to introduction of the analytical sample into individual separation equipment and/or detector have so far not generally received the required attention, particularly in dealing with urine samples. Inappropriate use or development of sample-preparation steps is perfectly reflected in subsequent steps. Thus, losses of metabolites can occur due to metabolite co-precipitation with proteins and/or poor solubility in the selected extraction solvent, and non-removed compounds lead to ionization suppression in MS [17,18]. All these produce errors in both targeted and untargeted analysis –negative errors in targeted analysis and absence in untargeted analysis. Furthermore, the use of methods with poor sample clean-up can significantly reduce column life-time and cause contamination of ion source, thus requiring frequent MS cleaning.

Concerted efforts to standardize, to improve and to understand urine sample preparation protocols further in both untargeted and targeted metabolomics studies are still needed, although significant progress has been made to understand the effect of sample preparation on both observed metabolite profiles and quantitation. As such, increased validation and understanding of existing sample-preparation methods and the emergence of new sample-preparation methods based on existing or new techniques for metabolomics are necessary before the potential of metabolomics can be fully exploited.

The main aspects to be dealt with are as follows:

(1) studies on sample stability as a function of the metabolites to be determined in targeted analysis, or the threshold for impact in chemometric results in dealing with profiling studies;

(2) the influence of medication, drugs of abuse, or habits (*e.g.,* smoking) in metabolomics profiling;

(3) acceleration–automation of some slow-manual steps (*e.g.,* those involving enzymes); and,

(4) in-depth studies on the possibilities of dried urine spots.

Acknowledgments

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

References

[1] D. Ryan, K. Robards, P.D. Prenzler, M. Kendall, Anal. Chim. Acta 684 (2011)17.

[2] W.B. Dunn, D.I. Ellis, Trends Anal. Chem. 24 (2005) 285.

[3] K.T. Myint, K. Aoshima, S. Tanaka, T. Nakamura, Y. Oda, Anal. Chem. 81 (2009)1121.

[4] H. Yoshida, J. Yamazaki, S. Ozawa, T. Mizukoshi, H. Miyano, J. Agric. Food Chem. 57 (2009) 111.

[5] G. Theodoridis, H.G. Gika, I.D. Wilson, Trends Anal. Chem. 27 (2008) 251.

[6] C. Bottcher, E. von Roepenack-Lahaye, E. Willischer, D. Scheel, S. Clemens, Anal. Chem. 79 (2007) 1507.

[7] E. Allard, D. Bäckström, R. Danielson, P.J.R. Sjöberg, J. Bergquist, Anal. Chem. 80 (2008) 8946.

[8] B. Kapur, S. Hershkop, G. Koren, V. Gaughan, Ther. Drug Monit. 21 (1999) 243.

[9] N.N. Kaderbhai, D.I. Broadhurst, D.I. Ellis, R. Goodacre, D.B. Kell, Comp. Funct. Genom. 4 (2003) 376.

[10] D.B. Kellet, M. Brown, H.M. Davey, W.B. Dunn, I. Spasic, S.G. Oliver, Nat. Rev. Microbiol. 3 (2005) 557.

[11] H.G. Gika, G.A. Theodoridis, I.D. Wilson, J. Chromatogr. A 1189 (2008) 314.

[12] A.D. Maher, S.F.M. Zirah, E. Holmes, J.K. Nicholson, Anal. Chem. 79 (2007) 5204.

[13] B. Álvarez Sánchez, F. Priego Capote, J. Ruiz Jiménez, M.D. Luque de Castro, J. Chromatogr. A 1207 (2008) 46.

[14] Z. León, A. Chisvert, I. Tarazona, A. Salvador, Anal. Bioanal.Chem. 398 (2010) 831.

[15] X. Duportet, R.B.M. Aggio, S. Carneiro, S.G. Villas-Boâs, Metabolomics 8 (2012) 410.

[16] G. Duranti, S. Boenzi, C. Rizzo, L. Ravà, V. Di Ciommo, R.Carrozzo, M.C. Meschini, D.W. Johnson, C. Dionisi-Vici, Clin. Chim. Acta 398 (2008) 86.

[17] J.M. Halket, D. Waterman, A.M. Przyborowska, R.K.P. Patel, P.D. Fraser, P.M. Bramley, J. Exp. Botany 56 (2005) 219.

[18] D. Vuckovic, Anal. Bioanal. Chem. 403 (2012) 1523.

[19] M.D. Luque de Castro, J.L. Luque-García, Acceleration and Automation of Solid Sample Treatment, Elsevier, Amsterdam, The Netherlands, 2002.

[20] M.D. Luque de Castro, F. Priego-Capote, Analytical Applications of Ultrasound, Elsevier, Amsterdam, The Netherlands, 2006.

[21] C. Núñez, M.A. Ortiz Apodaca, Eur. J. Clin. Chem. Clin. Biochem. 32 (1994) 461.

[22] G. Bieniek, Int. Arch. Occup. Environ. Health 70 (1997) 334.

[23] J. Le, C. Perier, S. Peyroche, F. Rascle, M.A. Blanchon, R. Gonthier, J. Frey, A. Chamson, Amino Acids 17 (1999) 315.

[24] R.A. Blum, T.J. Comstock, D.A. Sica, R.W. Schultz, E. Keller, P. Reetze, H. Bockbrader, D. Tuerck, J.A. Busch, P.A. Reece, A.J. Sedman, Clin. Pharmacol. Ther. 56 (1994) 154.

[25] R. García-Villalba, A. Carrasco-Pancorbo, E. Nevedomskaya, Anal. Bioanal. Chem. 398 (2010) 463.

[26] X. Wang, H. Lv, G. Zhang, J. Sep. Sci. 31 (2008) 2994.

[27] C.M. Slupsky, K.N. Rankin, J. Wagner, H. Fu, D. Chang, A.M. Weljie, E.J. Saude, B. Lix, D.J. Adamko, S. Shah, R. Greiner, B.D. Sykes, T.J. Marrie, Anal. Chem. 79 (2007) 6995.

[28] M.C. Walsh, L. Brennan, E. Pujos-Guillot, J.L. Sébédio, A. Scalbert, A. Fagan, D.G. Higgins, M.J. Gidney, Am. J. Clin. Nutr. 86 (2007) 1687.

[29] S. Silvester, F. Zang, J. Chromatogr. B 893–894 (2012) 134.

[30] J.M. Büscher, D. Czernik, J.C. Ewald, U. Sauer, N. Zamboni, Anal. Chem. 81 (6) (2009) 2135.

[31] R. El Rammouz, F. Létisse, S. Durand, J.C. Portais, Z.W. Moussa, X.Fernández, Anal. Biochem. 398 (2) (2010) 169.

[32] J.L. Griffin, R.A. Kauppinen, J. Proteome Res. 6 (2) (2007) 498.

[33] N.C. van de Merbel, Trends Anal. Chem. 27 (2008) 924.

[34] P. Bernini, I. Bertini, C. Luchinat, P. Nincheri, S. Staderini, P. Turano, J. Biomol. NMR 49 (2011) 231.

[35] W.B. Dunn, D. Broadhurst, D.I. Ellis, M. Brown, A. Halsall, S. O´Hagan, I.Spasic, A. Tseng, D.BKell, Int. J.Epidemiol.37(2008)23.

[36] P. Gupta, R. Gokhale, B.S. Kirschner, J. Pediatr. Gastroenterol. Nutr. 33 (2001) 450.

[37] Y.L. Cheng, L.G. Rider, J.D. Bell, R.M. Summers, L.S. Zemel, R.M.

Rennebohm, M.H. Passo, J. Hicks, F.W. Miller, D.L. Scott, Arthritis Care Res. 53 (2005) 565

[38] C. Lane, M. Brown, J. Nelly, G. Mangos, Nephrology 11 (206) 245.

[39] N.D. Ferrante, H.S. Lipscomb, Clin. Chim. Acta 30 (1970) 69.

[40] B.M. Warrack, S. Hnatyshyn, K.H. Ott, H. Zhang, M. Sanders, Proc. 54th ASMS Conf. Mass Spectrom. Allied Topics, Seattle, WA, USA, 2006.

[41] J.G. Kuenen, W.N. Konings, Accredit. Qual. Assur. 15 (2010) 133.

[42] A. Jiye, Q. Huang, G. Wang, W. Zha, B. Yan, H. Ren, S. Gu, Y. Zhang, Q. Zhang, F. Shao, L. Sheng, J. Sun, Anal. Biochem. 379 (2008) 20.

[43] R. Ventura, L. Damasceno, M. Farre´, J. Cardoso, J. Segura, Anal. Chim. Acta 418 (2000) 79.

[44] A. Azzouz, E. Ballesteros, J. Chromatogr. B 891–892 (2012) 12.

[45] R. García-Villalba, A. Carrasco-Pancorbo, E. Nevedomskaya, O.A. Mayboroda, A.M. Deelder, A. Segura-Carretero, A. Fernández-Gutiérrez, Anal. Bioanal. Chem. 398 (2010) 463.

[46] M.S. Chang, Q. Ji, J. Zhang, T.A. El-Shourbagy, Drug Dev. Res. 68 (2007) 107.

[47] F. Michopoulos, L. Lai, H. Gika, G. Theodoridis, I. Wilson, J. Proteome Res. 8 (2009) 2114.

[48] E.C. Laiakis, G.A.J. Morris, A.J. Fornace, PLoS One 5 (2010) e12655.

[49] R. Turner, L.K. Stamp, A.J. Kettle, J. Chromatogr. B 891–892 (2012) 85.

[50] D. Darghouth, B. Koehl, G. Madalinski, et al., Blood 117 (2011) e57.

[51] US Food and Drug Administration (FDA), FDA Guidance for Industry, Bioanalytical Method Validation, FDA Centre for Drug Valuation and Research (CDER), Silver Spring, MD, USA, May 2001.

[52] W.B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis- McIntyre, N.

Anderson, M. Brown, J.D. Knowles, A. Halsall, J.N. Haselden, A.W. Nicholls, I.D. Wilson, D.B. Kell, R. Goodacre, Nat. Protoc. 6 (2011) 1060.

[53] H.G. Gika, G.A. Theodoridis, J.E. Wingate, I.D. Wilson, J. Proteome Res. 8 (2007) 3291.

[54] E.J. Saude, B.D. Sykes, Metabolomics 3 (2007) 19.

[55] M. Lauridsen, S.H. Hansen, J.W. Jaroszewski, C. Cornett, Anal. Chem. 79 (2007) 1181.

[56] S. Cubbon, T. Bradbury, J. Wilson, J. Thomas-Oates, Anal. Chem. 79 (2007) 8911.

[57] J. Henion, E. Brewer, G. Rule G, Anal. Chem. 70 (1998) 650A.

[58] S.M.R. Stanley, H.C. Foo, J. Chromatogr. B 836 (2006) 1.

[59] M.H. Spyridaki, P. Kiousi, A. Vonaparti, P. Valavani, V. Zonaras, M. Zahariou, E. Sianos, G. Tsoupras, C. Georgakopoulos, Anal. Chim. Acta 573–574 (2006) 242.

[60] E. Gustavsson, M. Andersson, N. Stephanson, O. Beck, J. Mass Spectrom. 42 (2007) 881.

[61] A. Parkinson, B. Olgivie, in: C. Klaassen (Editor), Casarett and Doull's Toxicology, 7th Edition, McGraw–Hill, New York, USA, 2008.

[62] R. Raftogianis, C. Creveling, R. Weinshilboum, J Natl. Cancer Inst. Monogr. 27 (2007) 113.

[63] Z. León-González, F. Ferreiro-Vera, F. Priego-Capote, M.D. Luque de Castro, J. Chromatogr. A 1218 (2011) 3013.

[64] B. Álvarez-Sánchez, F. Priego-Capote, M.D. Luque de Castro, Analyst (Cambridge, UK) 134 (2009) 1416.

[65] V.G. Migues, E. Tátar, I. Virág, C. Zang, Y. Jao, G. Záray, Food Chem. 105 (2007) 1718.

[66] B. Álvarez-Sánchez, F. Priego-Capote, J.M. Mata-Granados, M.D. Luque de

Castro, J. Chromatogr. A 1217 (2010) 4688.

[67] M. Calderón-Santiago, F. Priego-Capote, M.D. Luque de Castro, J. Pharm. Biomed. Anal. 70 (2012) 476.

[68] B. Álvarez-Sánchez, F. Priego-Capote, M.D. Luque de Castro, J. Chromatogr. A 1207 (2008) 46.

[69] F. Michopoulos, L. Lai, H. Gika, G. Theodoris, I. Wilson, J. Proteome Res. 8 (2009) 2114.

[70] C.L. Silva, M. Passos, J.S. Cámara, Br. J. Cancer 105 (2011) 1894.

[71] C.R. Daniel, K.L. Schwartz, J.S. Colt, Br. J. Cancer 105 (2011) 1096.

[72] J.M. Halket, D. Waterman, A.M. Przyborowska, R.K.P. Patel, P.D. Fraser, P.M. Bramley, J. Exper. Botany 56 (410) (2005) 219.

[73] A.D. de Jager, J.V. Warner, M. Henman, W. Ferguson, A. Hall, J. Chromatogr. B 897 (2012) 22.

[74] R. Fan, R. Ramage, D. Wang, J. Zhou, J. She, Talanta 93 (2012) 383.

[75] A.M. Evans, C.D. DeHaven, T. Barrett, M. Mitchell, E. Milgram, Anal. Chem. 81 (2009) 6656.

[76] K.D. Sheikh, S. Khanna, S.W. Byers, A.J. Fornace, A.K. Cheema, J. Biomol. Techn. 22 (2011) 1.

[77] Y. Minami, T. Kasukawa, Y. Kakazu, M. Ligo, M. Sugimoto, S. Ikeda, A. Yasui, G.T.J. can der Horst, T. Soga, H.R. Ueda, Proc. Natl. Acad. Sci. USA 106 (2009) 9890.

[78] K.A. Lawton, A. Berger, M. Mitchell, K.E. Milgram, A.M. Evans, L. Guo, R.W. Hanson, S.C. Kalhan, J.A. Ryals, M.V. Milburn, Pharmacogenomics 9 (2008) 383.

[79] S.J. Bruce, I. Tavazzi, V. Parisod, S. Rezzi, S. Kochhar, P.A. Guy, Anal. Chem. 81 (2009) 3285.

[80] E.S. Chernetsova, G.E. Morlock, Bioanal. Rev. 3 (2011) 1.

[81] G. Huang, H. Chen, X. Zhang, R.G. Cooks, Z. Ouyang, Anal. Chem. 79 (2007) 8327.

[82] E.S. Chernetsova, G.E. Morlock, Mass Spectrom. Rev. 30 (2011) 875.

[83] N.M. Suni, P. Lindfors, O. Laine, P. Östman, I. Ojanperä, T. Kotiaho, T.J. Kauppila, R. Kostiainen, Anal. Chim. Acta 699 (2011) 73.

[84] P.J. Taylor, Clin. Biochem. 38 (2005) 328.

[85] T. Berg, D.H. Strand, J. Chromatogr. A 1218 (2011) 9366.

[86] Y. Suzuki, Y. Endo, M. Ogawa, S. Yamamoto, A. Takeuchi, T. Nakagawa, N. Onda, J. Chromatogr. B 877 (2009) 3743.

[87] D. Tsikas, J. Chromatogr. B 877 (2009) 2308

CHAPTER 2

Analytical methods based on exhaled breath for early detection of lung cancer

TrAC, Trends in Analytical Chemistry 32 (2012) 13–20

Trends in Analytical Chemistry

32 (2012) 13–20

Analytical methods based on exhaled breath for early detection of lung cancer

M.A. Fernández-Peralbo [a,b], M.D. Luque de Castro [a,b]

[a]*Department of Analytical Chemistry, University of Córdoba, Annex Marie Curie Building, Campus of Rabanales, 14071 Córdoba, Spain.*

[b]*Maimónides Institute for Research in Biomedicine of Córdoba (IMIBIC), Reina Sofía University Hospital, University of Córdoba, 14071 Córdoba, Spain.*

Analytical methods based on exhaled breath for early detection of lung cancer

M.A. Fernández-Peralbo, M.D. Luque de Castro

Abstract

Lung cancer is among the most common malignancies and among the most difficult to treat, mainly due to the lack of symptoms, which means that the vast majority of tumors to be diagnosed are at an advanced stage. After reviewing the literature on this subject, with special emphasis on both canine smell and the potential of omics biomarkers for detecting lung cancer, we discuss the role of analytical chemists in the new steps to be taken for early detection. We also comment on the steps to be developed (from sampling to full validation of the found biomarkers).

Keywords Biomarker; Canine smell; Early detection; Exhaled breath condensate (EBC); Lung cancer; Omics; Sampling; Sensor; Validation; Volatile organic compound (VOC)

1. Introduction

Lung cancer is among the most common malignancies, and remains, at the same time, among the most difficult to treat. Thus, the number of deaths from lung cancer each year surpasses those of the other most common causes of cancer-related mortality combined [1], mainly due to the lack of symptoms, which means that the vast majority of tumors are diagnosed at an advanced stage. Despite recent developments, the statistics are sobering: lung cancer is by far the most common cancer worldwide in men and women, with incidence and mortality rates of 35.5 and 31.2 per 100,000, respectively [2]. Less than 15% of individuals diagnosed with lung cancer are alive five years following diagnosis. When lung cancer is detected in early stages, surgical resection alone can achieve five-year survival rates as high as 50% [3].

Because lung cancer is almost always asymptomatic in the early stages, screening is the only systematic means of finding lung tumors at the surgically curable stage [4–8]. For earlier diagnosis of lung cancer, screening programs in asymptomatic, high-risk population groups have been studied by several technologies, including cytology of the sputum [9,10], circulating tumor biomarkers [11,12], blood proteomic patterns [13,14], chest tomography [15,16], nuclear magnetic resonance (NMR) [17], and other techniques or combinations (*e.g.,* chest X-ray, and low-dose computed tomography —CT), which have been the subject of much study and debate over the past several decades. Each approach has limited diagnostic specificity as currently applied [18,19], so that identifying particularly high-risk individuals for application of these candidate early-disease-detection strategies may allow leveraging of their performance.

Chest radiographs and sputum cytology frequently provide false negatives and fail to detect many early-stage cases as a result [3]. Also, CT scans can detect lesions as small as 1 mm in diameter but have many false positives that lead to unnecessary biopsies [20] unless a positron-emission-tomography scan is used for confirmation [21].

Concerning individually considered biomarkers, they have not shown suitability for use as screening tests, and it is not yet clear what panel of tests used together will provide sufficient diagnostic sensitivity and specificity [22,23]. Screenings based on visceral epithelia by non-invasive sampling for risk assessment in asymptomatic subjects poses anatomic challenges. Meanwhile, the use of urine, which is very easy to obtain as a sample, has recently provided promising results when interrogated by NMR [24].

Early detection of cancer is highly desirable since it frequently reduces treatment toxicity and ensures longer patient survival. Very often, however, the limited capabilities of today's diagnostic methods probably lead to high cancer mortality.

Exhaled breath contains aerosols and vapors that can be collected for non-invasive analysis of physiologic and pathologic processes in the lung. To capture the breath for assay, exhaled air is passed through a cooled, condensing apparatus, which is also available as a hand-held, disposable device. The result is an accumulation of condensed fluid that is referred to as exhaled breath condensate (EBC). Predominantly derived from water vapor, EBC has dissolved within it aqueous, soluble, non-volatile compounds. The technique has attracted broad research interest, and there is significant literature describing its utility in procuring small metabolites for the investigation of inflammatory lung diseases [25,26], thus opening a door in this area to a wide use of metabolomics. Because breath collection can be performed even in very severe patients and also repeated within short intervals, breath testing is considered to be a potentially ideal candidate for screening purposes. Although the number of studies using EBC has exponentially increased in the past decade, only a few have investigated samples from lung-cancer patients.

It is worth distinguishing between EBC and volatile organic compounds (VOCs) in breath. Despite VOCs also belonging to exhaled breath, only the volatile fraction (*e.g.,* acetonitrile, benzene, toluene, ethyl benzene and xylenes, pentane, isoprene and methyl-pentane, octane, decane, and styrene) is considered to be in it, so, as an analytical sample, it is very different from EBC.

At present, only canine scent has been able to detect lung cancer at an early stage using EBC [27]. Despite these procedures being distrusted by scientists, the finding opens a door to hope for this type of screening. Given the dramatic sensitivity of dogs' sense of smell, analytical chemists must pursue similar sensitivity based on both sample preparation and cutting-edge measurement equipment for a more rational, confident and wide information about the target individual. Thus, we first discuss the contribution of dog scent; then, advances in this area based on the three great omics and the use of sensors arrays that has been pursued. We finish with what could be the step to be taken by analytical chemists to contribute to this hard, but key, task of paramount importance to society.

2. Canine scent detection

In addition to other, less frequent, uses of dogs to study a safe, effective means of locoregional delivery of chemotherapy to the lung [28], the hypothesis that dogs may be able to detect malignant tumors on the basis of odor, first put forward by Williams and Pembroke in a letter to the *Lancet* in 1989 [29], has been a matter of study of several research teams.

A preliminary study was developed using methods normally used in canine olfactory detection of drugs and explosives. In this study, two dogs demonstrated reliable localization of melanoma-tissue samples hidden on the skin of healthy volunteers. One dog (A) then "confirmed" clinically suspected (and subsequently biopsy-proven) diagnoses of melanoma in five patients. In a sixth patient, this dog "reported" melanoma at a skin location for which initial pathological examination was negative, despite clinical suspicion.

The results obtained provided some evidence that there are volatile clues released from melanoma tissue that allow lesion localization by the canine olfactory system, but also demonstrated the need to generate new information on both the chemicals released from the body during disease states and the limits of detection (LODs) of the canine olfactory system. Progress in these areas is

foreseeable, subject to selection of those diseases for which the greatest diagnostic benefit is predicted. For a given disease, it was concluded that the relative merits of biological (*e.g.*, canine olfactory system) *versus* non-biological sensing could be determined only through rigorous proficiency testing that incorporates a range of objective endpoints [30].

Dogs have also been trained to distinguish patients with bladder cancer on the basis of urine odor more successfully than would be expected by chance alone. This suggests that tumor-related volatile compounds are present in urine, imparting a characteristic odor signature distinct from those associated with secondary effects of the tumor (*e.g.*, bleeding, inflammation, and infection) [31].

A number of individual biomarkers are unsuitable for use in screening tests. Also, it is uncertain whether a given test panel is sensitive and specific enough for the accurate diagnosis of lung or breast cancer [22,23]. These shortcomings led us to explore the use of canine scent detection for this purpose —dogs are known to possess an extraordinary scenting ability and to be able to detect substances at concentrations as low as a few parts-per-trillion [32]. The authors of the study supported their premises on the previous uses described in the literature [29–31]. The previous applications have shown that the canine olfactory system can not only detect substances at very low concentration levels, but also discriminate among substances in complex chemical mixtures (*e.g.*, exhaled human breath).

After appropriate training of both gender dogs, their responses to breath samples were classified as:

Correct response:

(1) indicated by a sitting or lying down response directly in front of a sample station containing a cancer sample (a true positive in sensitivity calculations); and

2) sniffing but not indicating on a control sample (true negative).

Incorrect responses were:

(1) indicating on a control sample (false positive);

(2) sniffing but not indicating on a cancer sample (false negative); and,

(3) hesitation, an incomplete reaction either toward cancer or control samples (either false-positive or false-negative depending on whether hesitation was on a cancer or control sample).

Application of this biochemical tool (comprising five dogs of both genders) to cases in a study (double-blinded testing of breath samples from 28 lung-cancer patients and 6 breast cancer patients —in both cases at four stages of disease— and 17 controls) showed that, among lung-cancer patients and controls, overall sensitivity of canine scent detection compared to biopsy-confirmed conventional diagnosis was 0.99 (95% CI, 0.99, 1.00) and overall specificity 0.99 (95% CI, 0.96, 1.00). The sensitivity thus obtained compares favorably with that of chest X-ray in detecting early-stage lung cancer, and so does the specificity with that of a CT scan in excluding lung cancer.

In the case of breast cancer, the sensitivity and the specificity compared favorably with that of mammography. It is worth noting that, for both cancers, sensitivity and specificity were remarkably similar across all four stages of disease. This fact demonstrated the presence of enough concentration of biomarkers in breath for precocious diagnostic of lung cancer and opened a door to precocious diagnostic of lung cancer by analytical equipment after lowering the detection and quantitation limits of the corresponding methods, performing dramatic preconcentration steps, or both [27].

Key aspects that decrease the importance of canine scent for lung-cancer detection are as follows:

(1) the authors of the study overestimated specificity as they used only healthy controls (rather than a broad spectrum of subjects that included, for example, those with bronchitis or emphysema as controls for lung cancer);

(2) despite the authors measuring for potential confounding by other odors, their dogs may nevertheless have detected and responded to odors associated with

cancer (*e.g.*, inflammation, infection, or necrosis), rather than to cancer specifically;

(3) current smokers varied between lung-cancer patients and controls. Although after adjusting for confounding by smoking high sensitivity and specificity remain, the design should have included specific recruitment strategies to avoid differences in smoking status between patients and controls;

(4) because there was no statistically significant difference between patients and controls in the most recently eaten meal (the most recent eating of foods that may affect breath such as fish, spicy foods, alcohol, lamb, pork, coffee, tea, and garlic) whether significant differences between patients and controls with respect to odors on the breath associated with dietary factors could interfere with training and scenting accuracy of dogs was unknown; and,

(5) other factors (*e.g.*, age, smoking, diabetes, and dental infection) could also play a role in the dog's response [27].

Recently, a patent on this matter registered:

(1) a method for conditioning a dog for recognizing a lung-cancer illness of a human; and,

(2) a method for preparing a sample for recognizing lung-cancer illness using a dog [33].

In the light of these and similar studies, two key questions are crucial:

(1) How confident we will be with a dog response on which our life could depend?

(2) Accepting as real that the dogs, with that very superior smell sense, are capable of identifying one, several or many compounds in breath, how can analytical methods approach the dog LODs?

While the former question is very personal, the latter requires as starting points:

(1) a review of the state-of-the-art of screening methods for precocious detection of lung cancer based on breath analysis, which are clearly insufficient;

(2) a review of the state-of-the-art of present analytical instrumentation and equipment for sample preparation, particularly for preconcentration steps; and,

(3) the search not for individual biomarkers of lung cancer, but for the EBC profile of the individual under the systems biology umbrella in order to distinguish both the type of lung cancer and the status of the disease.

3. Omics in the search for biomarkers of lung cancer

The use of recent, highly-sensitive technologies in sample analysis has considerably facilitated assessment of EBC samples [34]. The fact that EBC contains a number of compounds of very different natures and complexity (*e.g.*, ions, metabolite —*e.g.*, ammonia, hydrogen peroxide, isoprostanes, lactate, leukotrienes, nitrogen oxides, peptides, prostaglandins, and tromboxanes— and other molecules, including adenosine, ATP, and various cytokines [35]) corroborates that the three higher omics could be the basis for early lung-cancer detection. The questions are whether one of them suffices to provide discriminating biomarkers or several should be involved (systems biology, also known as omics integration), and analytical technology used to obtain proper information.

Proteomics and metabolomics require the use of innovative technologies —*e.g.*, mass spectrometry (MS), gas chromatography with MS (GC–MS), and ion-mobility spectrometry (IMS)—, which offer great potential for the field of exhaled biomarker profiling [4], but others should be successfully involved —*e.g.*, liquid chromatography (LC) with tandem MS (LC–MS^2), time-of-flight (TOF)-MS, matrix-assisted laser-desorption ionization (MALDI)-TOF or Orbitrap analyzers and stable-isotope technologies [36].

Biomarkers in exhaled breath have been used to elucidate the mechanisms behind some diseases and also to facilitate decision making in the clinical field by using a variety of approaches, including the determination of individual biomarkers and the recognition of signal patterns provided by unknown

substances. Although pattern recognition is statistically challenging, it provides a powerful means for analyzing samples containing a variety of components.

3.1. Genomics and lung cancer

DNA-related alterations were found by Gessner *et al.* [37] in EBC samples from lung-cancer patients. They detected p53 mutations in most EBC samples from the patients, but not in those from healthy controls. Also, they identified K-ras gene mutations in EBC samples from patients with lung cancer [38]. Similarly, EBC samples from patients with lung cancer were found to exhibit microsatellite instability to a great extent relative to control subjects [39,40]. The microsatellite profile of DNA from EBC was consistent with that for lung-cancer tissue from each patient.

Gene-promoter hypermethylation is recognized as a crucial component in lung-cancer initiation and progression [41], and suggestive correlations with smoking and lung-cancer case-control status have been found to depend on individual genes [42].

Also, according to one hypothesis [43], VOCs produced by tumors and detected by dogs are the products of major histocompatibility complex (MHC) genes. These human leukocyte antigen (HLA) molecules have soluble isoforms that are present in blood, urine and sweat, and MHC-dependent odor components can be detected by an electronic nose [44]. There is an association between changes in HLA expression and cancer, suggesting that the HLA-associated smell prints of human cancer could easily serve as olfactory clues. However plausible, the hypothesis was not tested; rather, these authors supported their assumptions with two pieces of evidence, namely:

(1) human body odor is genetically determined by MHC via soluble, detectable isoforms present in sweat; and,

(2) changes in HLA expression and cancer are closely related.

Tumor changes are often associated with low expression of classical HLA class I molecules (HLA-A, HLA-B and HLA-C). Likewise, cancer is associated with

high expression of non-classical HLA class I molecules (HLA-G, HLA-E). According to these authors, this evidence suggests that these HLA-associated olfactory clues of human cancer should be easy to analyze (*e.g.*, with an "electronic nose") in order to facilitate very early, reliable detection and diagnosis of cancer.

3.2. Proteomics

The total protein content of EBC has never been compared between patients with lung-cancer and control subjects. Analyses in this context have been restricted to cytokines, which, in biological samples, can be determined in two different ways, namely by:

(1) detecting individual molecules with specific immunoassays; or,

(2) using a sensor array (*i.e.,* proteomics).

The latter has been deemed a powerful tool for detecting lung cancer in blood. Simultaneously assessing five selected serum-cytokine markers allowed patients in a study to be discriminated from control subjects with 90% specificity and sensitivity [45].

In raw EBC samples, interleukin 6 (IL-6) was measured in patients and healthy control subjects [46]. The results reported meant that the concentration of this glucoprotein was significantly higher in patients with lung cancer. The concentration of IL-2 together with that of tumor necrosis factor (TNF)-α is also significantly increased in raw EBC samples of patients with lung cancer compared with healthy control subjects [47]. Also, increased levels of TNF-α and vascular endothelial growth factor (VEGF), together with isoprostane, have been found in lung-cancer patients [48].

The only study using proteomics to investigate EBC in lung cancer was that reported by Kullmann *et al.* in 2008 [49]. The authors circumvented the shortcoming of the expected low concentration of cytokines in EBC by freeze-drying, resuspending and pooling the samples for antibody microarray analysis of 120 cytokines. Each cytokine on the array gave a signal in both groups. Some 10 cytokines, including eotaxin, fibroblast growth factors, IL-10 and macrophage

inflammatory protein-3, were detected with a two-fold difference between groups. However, the results were not validated by using an alternative method for protein detection. Also, pooling the samples may have masked individual protein patterns. In any case, this pilot study was the first testifying that proteomics on EBC may be an effective means for detecting lung cancer, even though its potential for clinical practice remains unassessed.

3.3. Metabolomics

Most of molecules in exhaled breath are metabolites (particularly when VOCs are the target analytes), so this discipline together with the bottom-up approach could be the key to providing accurate information for discrimination of lung cancer affected individuals in the earliest stages of the disease.

The analysis of exhaled air for VOCs has been judged a highly promising approach to discriminating between healthy subjects and patients with lung cancer ever since the earliest report by Gordon *et al.* on this subject in 1985 [50]. Using a custom breath-collection technique and computer-assisted GC–MS afforded the identification of various VOCs associated with the disease in exhaled air from lung-cancer patients [51]. However, the selected volatiles exhibited sufficient diagnostic power in the GC–MS profiles to enable nearly accurate distinction of control subjects from cancerous patients in a limited patient population only.

The term "breathalyzer" was first used in 1999 by Phillips *et al.* in connection with lung cancer [52]. A combination of 22 VOCs in breath samples from patients with an abnormal chest radiograph afforded discrimination between individuals with and without lung cancer. Poli *et al.* showed a combination of 13 VOCs enabled accurate classification of cases, including lung cancer, chronic obstructive pulmonary disease, asymptomatic smokers and healthy subjects into well-defined groups by using conventional diagnostic methods [53]. The results of this research have not been assessed.

Poli *et al.* demonstrated that most of the VOCs in lung-cancer patients were higher at the baseline in comparison with controls, and this difference was still evident 1 month after surgery and 3 years later. Therefore, exhaled VOCs are not

specific biomarkers for lung cancer —they are seemingly not produced by tumor tissue alone— but rather an epiphenomenon of lung-cancer development, probably resulting from the chronic load and burden of VOCs in lung tissue as a whole. The results of factor and discriminant analyses were used to classify VOCs into three different categories, namely:

(1) smoke- and environment-related substances (*e.g.*, acetonitrile, benzene, toluene, ethyl benzene, and xylenes);

(2) endogenous substances involved in inflammation, steroid genesis and methylation (*e.g.*, pentane, isoprene, and methyl pentane); and,

(3) a heterogeneous group of uncertain origin including octane, decane, styrene and pentamethyl heptane [54].

Recent studies support that exhaled breath contains thousands of different VOCs, most of them at the picomolar/litre level (10^{-12} mol/L or particles per trillion). In normal subjects, more than 3000 different VOCs can be detected. However, only 20–30 of these VOCs are present in all subjects —principally isoprene, alkanes, methyl alkanes and benzene derivatives [55].

Another proposed way to investigate the role of given VOCs in lung cancer is the use of isotope-labeled compounds. In addition to endogenous compounds observed in exhaled breath, the use of exogenous precursors has been proposed to test the effect of mutations in signaling pathways on the biochemistry of cancerous cells using isotope-labeled precursors [56–59]. At present, such breath tests use $^{13}CO_2$ as a "target compound", measuring the isotopic ratio of $^{13}CO_2/^{12}CO_2$ after ingestion of the precursor. Despite a promising future that has been predicted for measurement of the target compounds by proton-transfer reaction (PTR)-TOF-MS, photoionization-MS or GC–MS, there has so far been no significant research using this approach.

Concerning EBC samples, Wehinger *et al.* used PTR-MS to investigate the exhaled breath of primary lung-cancer patients. They demonstrated that concentrations of mass-to-charge ratio 31 (tentatively identified as formaldehyde)

and mass-to-charge ratio 43 (tentatively identified as isopropanol) are increased in lung-cancer patients compared with healthy controls [60].

However, mass-to-charge ratio 69 (tentatively identified as isoprene) turned out to be decreased in the exhaled breath of lung-cancer patients. This decrease in isoprene concentration as measured by PTR-MS was later confirmed by Bajtarevic *et al.* [61] in a much larger cohort. But, Kischkel *et al.* demonstrated that data analysis of concentration patterns of lung-cancer patients requires careful attention and that a proper, generally accepted normalization method for the concentrations of volatiles has not yet been found [62], despite recent advances on mathematical models [63,64]. One of the aspects that hinders normalization relates to the sampling method, which hardly provides reproducible and comparable quantitative results from EBC analysis. Furthermore, the influence of ventilation and blood flow can be considerable [65–68] and has not yet been studied in enough detail. Real-time analysis of exhaled breath, as pioneered by King *et al.* [65], is of great importance for understanding these factors. Last, but not least, are issues in statistics, data processing and data interpretation (including design, effect of disease incidence in groups under study, ratio of independent measurements and parameters used for data analysis) [69–71].

Bajtarevic *et al.* also examined the concentration patterns of 100 volatiles in patients with lung cancer by GC–MS with solid-phase microextraction [61]. In this way, they succeeded in discriminating cancerous patients from healthy controls with 80% sensitivity and 100% specificity. Consequently, these authors were the first to apply the validated identification of peaks based on calibration mixtures (starting from pure compounds).

Also, a sub-discipline of metabolomics, ionomics, has been considered of interest in lung-cancer detection, as the level of chromium, a heavy metal usually associated with occupational lung diseases, was elevated in EBC from patients with lung cancer but without known exposure to heavy metals [72].

4. Sensors: a magic word for clinicians

Non-analytical chemists in general, and clinicians in particular, are reluctant to use MS techniques (whether or not coupled to chromatographic equipment) because they "are expensive and require expert interpretation" [73,74]. For these reasons, they are fonder of techniques involving smaller, less expensive equipment, which is simple to handle and able to be miniaturized, particularly sensors or sensor arrays [75–77].

As established by Amann *et al.* [34], a given sensor system for clinical analysis typically uses an array of sensors, each with slightly different reactivities, allowing the system to respond to a large number of analytes. Sensor systems can respond to individual analytes, or complex mixtures of analytes such as those found in the breath. According to the authors [34], an ideal breath-analysis chemical-sensor system should:

(1) be sensitive to a wide range of analytes at very low concentrations;

(2) be insensible to the presence of water vapor because exhaled breath is fully humidified;

(3) respond rapidly and differently to small changes in concentration;

(4) provide a consistent output that is specific to a given exposure;

(5) return to its baseline state rapidly when not in contact with the analytes; and,

(6) be simple and inexpensive enough to manufacture large numbers of disposable units.

There have been several studies examining the potential for chemical-sensor systems to detect lung-cancer-breath biosignatures. The studies differ in:

(1) the type of sensor system used;

(2) the number of subjects studied;

(3) the stage distributions of the lung-cancer patients;

(4) the inclusion of individuals at risk of developing cancer and healthy controls;

(5) the means of interfacing the study subjects with the sensor system;

(6) the portion of exhaled breath studied;

(7) the means of controlling for ambient volatiles;

(8) the pattern-recognition statistical techniques that were used to analyze the sensor output; and,

(9) the means of validating the accuracy of the identified biosignatures.

The most salient studies on these sensors reported to date range between:

(1) the electronic nose that comprised eight quartz-microbalance gas sensors coated with different metalloporphyrins, described by Di Natale *et al.* [77], which found 100% classification of lung-cancer-affected patients, while 94% of the referents were correctly classified, by applying a partial least squares-discriminant; and,

(2) the colorimetric sensor array reported by Mazzone *et al.,* based on stronger dye–analyte interactions than those that cause simple physical adsorption (such as occurs with many other sensor platforms), thus causing color changes in the red, green and blue spectrum, with development of a prediction model that provided a sensitivity of only 73.3% and a specificity of 72.4% [78]; in addition, colorimetric sensors are not able to identify the specific compounds in a mixture.

Gold nanoparticle-based sensors [79], and dog-scent sensors [27–33] were also among those described.

The main problems of the sensors described so far are scant populations for testing the performance, always comprising healthy individuals as controls *versus* patients with an advanced development of lung cancer, and absence of validation in all cases. Other limitations include loss of sensitivity in the presence of water vapor or high concentrations of a single component, sensor drift and the inability to provide absolute calibration, relatively short life of some sensors, necessity to

do considerable method-development work for each specific application, and the inability to obtain quantitative data.

5. New analytical steps in early detection of lung cancer suggested by canine scent detection

In the light of the scant success so far achieved in the search for an approach to precocious detection of lung cancer, we consider that analytical chemists have the key to providing tools to clinicians in order to advance in the right direction.

It seems clear from canine scent detection that very low concentrations of biomarkers exist in EBC. Thus, the question is: "Are analytical chemists able to develop sample-preparation steps with enough clean up and mainly with enough preconcentration factors that, together with the high sensitivity of cutting-edge detection equipment, allow detection (better quantification) of these biomarkers detected by dogs?'. The steps to be taken in this direction should involve:

(1) to check the sampling systems (breath condensers) and design new devices in which this step is simultaneous with preconcentration by coating the sampler with one proper sorbent (or better, a mixture of proper sorbents). Sampling time (either a snapshot or looking at the dynamics during a full night or during a challenge) [80].

(2) To develop preconcentration methods, mainly based on solid-phase extraction, preferably working with on-line connection to the detector (TOF, ion-trap-MS, Orbitrap), either directly or through LC or GC equipment, in order to avoid losses or degradation of the analytes.

Then, in close collaboration with clinicians:

(3) to obtain fingerprinting of a number of samples from:

(i) healthy individuals, as control, involving smokers and non-smokers;

(ii) high-risk population of lung-cancer;

(iii) individuals with other types of cancers;

(iv) patients after lung-cancer resection; and,

(v) individuals with different respiratory diseases;

(4) to carry out an in-depth chemometrics classification of the fingerprinting in order to obtain the profile corresponding to lung-cancer biomarkers. It is not foreseeable to achieve a single lung-cancer biomarker;

(5) to develop analytical quantitative methods based on cutting-edge equipment (*e.g.*, triple-quadrupole mass spectrometers) to quantify the biomarkers, thus making it possible to evaluate the degree of development of lung cancer, and to establish reference values of the biomarkers;

(6) full, exhaustive validation of the biomarkers by conducting large multi-center studies; and,

(7) to address in an appropriate way the assessment of the biomarkers found in a really prospective way. In short, if early cancer recognition is accomplished, a large number of healthy individuals at risk (*e.g.*, smokers and ex-smokers) would have to be followed up as a certain number of them develop cancer.

After development of the planned research, the results can establish the role of each of the great omics in providing biomarkers, thus making way for the use of a single omics or a system biology (integrated omics) approach. Depending on these results and the type and/or complexity of the biomarkers obtained, it may, or may not, be possible to develop small equipment for point-of-care applications.

References

[1] A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, M.J. Thun, CA Cancer J. Clin. 59 (2009) 225.

[2] H.P. Chan, C. Lewis, P.S. Thomas, Lung Cancer 63 (2009) 164.

[3] F.V. Fossella, R. Komaki, J.B. Putnam, Lung Cancer, Springer-Verlag, New York, USA, 2003.

[4] I. Horváth, Z. Lár, N. Gyulai, M. Kollai, G. Losonczy, Eur. Respir. J. 34 (2009) 261.

[5] J.L. Mulshine, F. Scott, Chest 107 (Suppl. 6) (1995) 280S.

[6] R.A. Smith, V. Cokkinides, H.J. Eyre, Cancer J. Clin. 53 (2003) 27.

[7] P.B. Bach, D.E. Niewoehner, W.C. Black, Chest 123 (Suppl. 1) (2003) 83S.

[8] E. Edell, S. Lam, H. Pass, Y.E. Miller, T. Sutedja, T. Kennedy, G. Loewen, R.L. Keith, A. Gazdar, J. Thorac. Oncol. 4 (2009) 49.

[9] F.L. Johnson, B. Turic, R. Kemp, B. Palcic, R. Sussman, K.G. Voelker, E. Robinette, Chest 125 (2004) 157S.

[10] M. Varella-García, J. Kittelson, A.P. Schulte, K.O. Vu, H.J. Wolf, C. Zeng, F.R. Hirsch, T. Byers, T. Kennedy, Y.E. Miller, R.L. Keith, W.A. Franklin, Cancer Detect Prev. 28 (2004) 244.

[11] S. Helmig, J. Schneider, Expert Rev. Mol. Diagn. 7 (2007) 555.

[12] P.B. Yildiz, Y. Shyr, J.S. Rahman, N.R. Wardwell, L.J. Zimmerman, B. Shakhtour, W.H. Gray, S. Chen, M. Li, H. Roder, D.C. Liebler, W.L. Bigbee, J.M. Siegfried, J.L. Weissfeld, A.L. González, M. Ninan, D.H. Johnson, D.P. Carbone, R.M. Caprioli, P.P. Massion, J. Thorac. Oncol. 2 (2007) 893.

[13] C.M. Maciel, M. Junqueira, M.E. Paschoal, M.T. Kawamura, R.L. Duarte, M.G. Carvalho, G.B. Domont, J. Exp. Ther. Oncol. 5 (2005) 31.

[14] T. Kikuchi, D.P. Carbone, Respirology 12 (2007) 22.

[15] S. Diederich, Cancer Imaging 8 (Suppl. A) (2008) S24.

[16] J. Aisner, Clin Cancer Res 13 (2007) 4951.

[17] M.P. Lichy, P. Aschoff, C. Plathow, A. Stemmer, W. Horger, C. Mueller-Horvat, G. Steidle, M. Horger, J. Schafer, S.M. Eschmann, B. Kiefer, C.D. Claussen, C. Pfannenberg, H.P. Schlemmer, Invest. Radiol. 42 (2007) 605.

[18] T. Blanchon, J.M. Brechot, P.A. Grenier, G.R. Ferretti, E. Lemarie, B. Milleron, D. Chagué, D. Laurent, Y. Martinet, C. Beigelman-Aubry, F. Blanchon, M.P. Revel, S. Friard, M. Rémy-Jardin, M.Vasile, N. Santelmo, A. Lecalier, P. Lefébure, D. Moro-Sibilot, J.L. Breton, M.F. Carette, C. Brambilla, F. Fournel, A. Kieffer, G. Frija, A. Flahault, Lung Cancer 58 (2007) 50.

[19] P.B. Bach, G.A. Silvestri, M. Hanger, J.R. Jett, Chest 132 (2007) 69S.

[20] S. Dammas, E.F. Patz, P.C. Goodman, Lung Cancer 33 (2001) 11. [21] S.S. Shim, K.S. Lee, B.T. Kim, M.J. Chung, E.J. Lee, J. Han, J.Y. Choi, O.J. Kwon, Y.M. Shim, S. Kim, Radiology 236 (2005) 1011.

[22] T.C. Kennedy, F.R. Hirsch, Lung Cancer 45 (Suppl. 2) (2004) S21.

[23] W. Hilbe, S. Dirnhofer, R. Greil, E. Woll, Eur. J. Cancer Prev. 13 (2004) 425.

[24] J. Carrola, C.M. Rocha, A.S. Barros, A.M. Gil, B.J. Goodfellow, I.M. Carreira, J. Bernardo, A. Gomes, V. Sousa, L. Carvalho, I.F. Duarte, J. Proteome Res. 10 (2011) 221.

[25] S.A. Kharitonov, P.J. Barnes, Chest 130 (2006) 1541.

[26] P.J. Barnes, B. Chowdhury, S.A. Kharitonov, H. Magnussen, C.P. Page, D. Postma, M. Saetta, Am. J. Respir. Crit. Care. Med. 174 (2006) 6.

[27] M. McCulloch, T. Jezierski, M. Broffman, A. Hubbard, K. Turner, T. Janeck, Integrative Cancer Therapies 5 (1) (2006) 30.

[28] A.E. Hershey, I.D. Kurzman, L.J. Forrest, C.A. Bohling, M. Stonerook, M.E. Placke, A.R. Imondi, D.M. Vail, Clin. Cancer Res. 5 (1999) 2653.

[29] H. Williams, A. Pembroke, Lancet 1 (1989) 734.

[30] D. Pickel, G.P. Manucy, D.B. Walker, S.B. Hall, J.C. Walker, Appl. Animal Behaviour Sci. 89 (2004) 107.

[31] C.M. Willis, S.M. Church, C.M. Guest, W.A. Cook, N. McCarthy, A.J. Bransburry, M.R.T. Church, J.C.T. Church, Brit. Med. J 329 (2004) 712.

[32] D.B. Walker, P. Cavnar, J. Taylor *et al.*, Paper, Meeting Assoc. Chemorecept. Sci., April 2003, Sarasota, FL, USA.

[33] G. Wolfgang, European Patent EP 2177153 (A1), 2010.

[34] A. Amann, M. Corradi, P. Mazzone, A. Mutti, Expert Rev. Mol. Diagn. 11 (2011) 207.

[35] I. Horváth, J. Hunt, P.J. Barnes, Eur. Respir. J. 26 (2005) 523.

[36] R.C.A. Schellekens, F. Stellaard, H.J. Woerdenbag, H.W. Friklink, J.G.W. Kosterink, Br. J. Clin. Pharmacol. 72 (2011) 879.

[37] C. Gessner, H. Kuhn, K. Toepfer, S. Hammerschmidt, J. Schauer, H. Wirtz, Lung Cancer 43 (2004) 215.

[38] C. Gessner, B. Rechner, S. Hammerschmidt, H. Kuhn, G. Hoheisel, U. Sack, P. Ruschpler, H. Wirtz, Lung Cancer 68 (2010) 177.

[39] A. Mangia, F. Carpagnano, G. Stea, A. Susca, G. Di Gioia, M. De Lena, A. Paradiso, Am. J. Respir. Crit. Care Med. 172 (2005) 738.

[40] G.E. Carpagnano, M.P. Foschino-Barbaro, A. Spanevello, O. Resta, F. Carpagnano, G. Mule, R. Pinto, S. Tommasi, A. Paradiso, Am. J. Respir. Crit. Care Med. 177 (2008) 337.

[41] X. Chen, F. Xu, Y. Wang, Y. Pan, D. Lu, P. Wang, K. Ying, E. Chen, W. Zhang, Cancer 110 (2007) 835.

[42] W. Han, T. Wang, A.A. Reilly, S.M. Keller, S.D. Spivack, Respiratory Res. 10 (2009) 86.

[43] S.C. Balseiro, H.R. Correia, Medical Hypotheses 66 (2006) 270.

[44] S. Montag, M. Frank, H. Ulmer, D. Wernet, W. Gopel, H.G. Rammensee, Proc. Natl. Acad. Sci. USA 98 (2001) 9249.

[45] L. Zhong, S.P. Coe, S. Stromberg, N.H. Khattar, J.R. Jett, E.A. Hirschowitz, J. Thoracic Oncol. 1 (2006) 513.

[46] W. Jungraithmayr, C. Frings, G. Zissel, A. Prasse, B. Passlick, E. Stoelben, Respirology 13 (2008) 1022.

[47] G.E. Carpagnano, A. Spanevello, C. Curci, F. Salerno, G.P. Palladino, O. Resta, G. Di Gioia, F. Carpagnano, M.P. Foschino Barbaro, Oncol. Res. 16 (2007) 375.

[48] E. Dalaveris, T. Kerenidi, A. Katsabeki-Katsafli, T. Kiropoulos, K. Tanau, C.I. Gourfoulianis, K. Kostikas, Lung Cancer 64 (2009) 219.

[49] T. Kullmann, I. Barta, E. Csiszér, B. Antus, I. Horváth, Pathol. Oncol. Res. 14 (2008) 481.

[50] S.M. Gordon, J.P. Szidon, B.K. Krotoszynski, R.D. Gibbons, H.J. O´Neill, Clin. Chem. 31 (1985) 1278.

[51] G. Preti, J.N. Labows, J.G. Kostelc, S. Aldinger, R. Daniele, J. Chromatogr. B 432 (1988) 1.

[52] M. Phillips, K. Gleeson, J.M. Hughes, *et al.*, Lancet 353 (1999) 1930.

[53] D. Poli, P. Carbognani, M. Corradi, *et al.*, Respir. Res. 6 (1) (2005) 71.

[54] D. Poli, M. Goldoni, A. Caglieri, *et al.*, Acta Biomed. 79 (2008) 64.

[55] M. Phillips, J. Herrera, S. Krishnan, *et al.*, J. Chromatogr. B 729 (1999) 75.

[56] H.H. Ezzeldin, E.P. Acosta, L.K. Mattison, J. Fourie, A. Modak, R.B. Diasio, J. Breath Res. 3 (2009) 047002.

[57] L.K. Mattison, H. Ezzeldin, M. Carpenter, A. Modak, M.R. Johnson, R.B. Diasio, Clin. Cancer Res. 10 (2004) 2652.

[58] A. Modak, J. Breath Res. 3 (2009) 040201.

[59] A.S. Modak, J. Breath Res. 4 (2010) 017002.

[60] A. Wehinger, A. Schmid, S. Mechtcheriakov, *et al.*, Int. J. Mass Spec. 265 (2007) 49.

[61] A. Bajtarevic, C. Ager, M. Pienz, *et al.*, BMC Cancer 9 (2009) 348.

[62] S. Kischkel, W. Miekisch, A. Sawacki, *et al.*, Clin. Chim. Acta 411 (2010) 1637.

[63] J. King, K. Unterkofler, G. Teschl, *et al.*, J. Math. Biol. 63 (2011) 959.

[64] J. King, K. Unterkofler, G Teschl, *et al.*, J. Breath Res. 6 (2012) 016005.

[65] J. King, A. Kupferthaler, K. Unterkofler, *et al.*, J. Breath Res. 3 (2009) 027006.

[66] K.A. Cope, M.T. Watson, W.M. Foster, S.S. Sehnert, T.H. Risby, J. Appl. Physiol. 96 (2004) 1371.

[67] J. King, P. Mochalski, A. Kupferthaler, *et al.*, Physiol. Meas. 31 (2010) 1169.

[68] J. King, H. Ko, K. Unterkofler, *et al.*, J. Theor. Biol. 267 (2010) 626.

[69] R.A. Poldrack, J.A. Mumford, SCAN 4 (2009) 208.

[70] E. Vul, C. Harris, P. Winkielman, H. Pashler, Perspect. Psycholog. Sci. 4 (2009) 274.

[71] K. Fiedler, Perspect. Psycholog. Sci. 6 (2011) 163.

[72] M. Goldoni, A. Caglieri, M. Corradi, D. Poli, M. Rusca, P. Carbognani, A. Mutti, Int. Arch. Occup. Environ. Health 81 (2008) 487.

[73] X. Chen, F.J. Xu, Y. Wang, *et al.*, Cancer 110 (2007) 835.

[74] T. Ligor, M. Ligor, A. Amann, *et al.*, J. Breath Res. 2 (2008) 046006.

[75] P.J. Mazzone, T. Mekhail, Curr. Oncol. Rep. 9 (2007) 265.

[76] R.F. Machado, D. Laskowski, O. Deffenderfer, *et al.*, Am. J. Respir. Crit. Care Med. 171 (2005) 1286.

[77] C. Di Natale, A. Macagnano, E. Martinelli, *et al.*, Biosens. Bioelectron. 18 (2003) 1209.

[78] P.J. Mazzone, J. Hammel, R. Dweik, *et al.*, Thorax 62 (2007) 565.

[79] G. Pen, U. Tisch, O. Adams, M. Hakim, N. Shehada, Y.Y. Broza, S. Billan, T. Abdah-Bortnyak, A. Ruten, H. Haick, Nature Nano- technol. 4 (2009) 669.

[80] Z.L. Borrill, K. Roy, D. Sing, Eur. Respir. 32 (2008) 472.

SECTION II

Contributions to targeted metabolomics analysis

This Section II is composed by different contributions to targeted metabolomic analysis, a controversial strategy in the opinion of metabolomics detractors.

The research that constitutes Chapter 3 is an example of strategies to overcome the shortcomings caused by a complex sample such as human serum in the analysis of hydroxyeicosatetranoic acids (HETEs), metabolites from essential polyunsaturated ω-6 fatty acids: (i) on-line SPE to remove as many as possible interferents with minimum losses of the target metabolites and maximum preconcentration effect; and, (ii) use of a deuterated stable isotopic labeled internal standard (SIL-IS) to compensate for ionization suppression effects and for other chemical and instrumental variability sources.

An extended application of the above developed method is the matter of Chapter 4, in which the target metabolites were also eicosanoids from ω-6 fatty acids such as series-2 prostaglandins, thromboxane B2, hydroxyeicosatetraenoic acids, and hydroxyl-octadecadienoic acids. They were determined in serum from patients diagnosed with coronary artery disease and healthy individuals by SPE–LC–MS/MS. The chemometrical comparison of the data allowed discriminating those treated with aspirin and also the statistical evaluation of the patients according to the severity of the coronary lesion stratified as stable angina, non-ST-elevation acute coronary syndrome, and acute myocardial infarction.

A targeted method for determination of the antineoplastic drug paclitaxel and its main metabolites in serum and tissue by LC–ESI-MS/MS (detection limits within 0.03–0.15 ng/mL for serum and 0.07–0.62 ng/g for tissue) is the matter of Chapter 5. The exhaustive optimization and the low detection limits make the method a useful tool for application to samples from women suffering from ovarian peritoneal carcinomatosis, after hyperthermic intraperitoneal

intraoperative chemotherapy treatment by monitoring the concentration of the drug in these patients as strategy to evaluate the toxicity and efficiency of the treatment.

CHAPTER 3

Stable isotopic internal standard correction for quantitative analysis of hydroxyeicosatetraenoic acids (HETEs) in serum by on-line SPE–LC–MS/MS in selected reaction monitoring mode

Talanta 126 (2014) 170–176

Talanta

126 (2014) 170–176

Stable isotopic internal standard correction for quantitative analysis of hydroxyeicosatetraenoic acids (HETEs) in serum by on-line SPE–LC–MS/MS in selected reaction monitoring mode

M.A. Fernández-Peralbo [a,b], C. Ferreiro Vera [a,b], F. Priego-Capote [a,b],

M.D. Luque de Castro [a,b]

[a]*Department of Analytical Chemistry, University of Córdoba, Annex Marie Curie Building, Campus of Rabanales, 14071 Córdoba, Spain.*

[b]*Maimónides Institute for Research in Biomedicine of Córdoba (IMIBIC), Reina Sofía University Hospital, University of Córdoba, 14071 Córdoba, Spain.*

Stable isotopic internal standard correction for quantitative analysis of hydroxyeicosatetraenoic acids (HETEs) in serum by on-line SPE–LC–MS/MS in selected reaction monitoring mode

M.A. Fernández-Peralbo, C. Ferreiro Vera, F. Priego-Capote,

M.D. Luque de Castro

Abstract

The influence of the inclusion of a stable isotopic labeled internal standard (SIL-IS) on the quantitative analysis of hydroxyeicosatetranoic acids (HETEs) in human serum is evaluated in this research. A solid-phase extraction–liquid chromatography–tandem mass spectrometry (SPE–LC–MS/MS) platform, one of the preferred approaches for targeted analysis of biofluids through the selected reaction monitoring (SRM) operational mode, was used to determine HETEs. These compounds were chosen as targeted metabolites because of their involvement in cardiovascular disease, cancer and osteoporosis. 15HETE-d8 was chosen as internal standard to evaluate matrix effects. Thus, the physico-chemical properties of the SIL-IS were the basis to evaluate the analytical features of the method for each metabolite through four calibration models. Two of the models were built with standard solutions at different concentration levels, but one of the calibration sets was spiked with an internal standard (IS). The other two models were built with the serum pool from osteoporotic patients, which was spiked at different concentrations with the target analytes. In this case, one of the serum calibration sets was also spiked with the IS. The study shows that the IS allowed noticeable correction of matrix effects for some HETE isomers at certain concentration levels, while accuracy was decreased at low concentration (15 ng/mL) of them. Therefore, characterization of the method has been properly completed at different concentration levels.

Keywords Internal standard; Stable isotope; LC–MS/MS; On-line SPE; HETE

1. Introduction

Liquid chromatography coupled to mass spectrometry (LC–MS/MS) using a triple quadrupole spectrometer (QqQ) is considered one of the best options for targeted analysis in biofluids. Methods based on selected reaction monitoring (SRM) ensures unrivaled sensitivity and selectivity both in metabolomics and proteomics studies [1,2]. One important benefit associated to the SRM mode is robustness, mainly limited by potential decreased precision caused by ionization suppression effects, particularly in electrospray ionization (ESI), which is the most widely used ionization mode. Ionization suppression may be caused by high concentration of endogenous matrix components that co-elute with the target analytes [3]. Matrix effects can be eliminated or minimized by using a highly pure internal standard (IS) [4–6] with physico-chemical properties similar to that of the target analyte and generating a response signal distinguishable from that of the analyte of interest. Apart from ionization suppression effects, selection of a suitable IS compensates for other variability sources induced by sample preparation steps such as dilution, concentration by solvent evaporation and/or analytes adsorption, instrumental variations including injection volume as well as variations in response over time [7,8]. In fact, the ideal IS should behave as the target analytes during the entire analytical protocol [3].

There are two different approaches in dealing with the selection of the best IS: a structural or chemical analog related to the analyte or a stable isotopic labeled internal standard (SIL-IS) which contains stable isotopes such as ^{13}C, ^{15}N, ^{18}O or ^{2}H in its molecular structure [9]. SIL-forms of a given analyte are presently the most commonly used IS in small-molecule bioanalysis using LC–MS/MS [10]. Although ^{2}H is the cheapest and most frequently used isotope, other stable isotopes such as ^{13}C, ^{15}N or ^{18}O may be more suitable since deuterium and hydrogen have greater differences in their physical properties than, for example, ^{12}C and ^{13}C [11,12]. However, the number of studies based on the use of non-deuterated IS is lower than those involving deuterated IS for targeted analysis of human biological samples by LC–MS/MS [13,14].

The complexity in the analysis of biological samples arises from multicomponent methods, in which a number of labeled IS similar to the target compounds would be required to avoid ionization suppression effects ascribed to certain chemical structures [15]. However, these IS are not always feasible owing to both the high cost of the labeled compounds and the number of analytes to be determined, which can encompass of the order of 10, 50 or even 100 [16]. Theoretically, ionization suppression or enhancement effects caused by co-eluted matrix components should affect equally to the SIL-IS and analyte. It is, however, well-known that partial chromatographic separation of the analyte and its corresponding deuterated IS often occurs due to a small change in lipophilicity when exchanging hydrogen by deuterium [17]. How noticeable the effect is depends on factors such as the number of substituted atoms, size of the molecule, efficiency of the column, retention mechanisms and retention time. As an example, Wang *et al.* have shown that the MS response of carvedilol was significantly more affected than that of the SIL–IS carvedilol when they co-eluted within a region with severe ion suppression [11]: the accuracy and precision of the method were severely affected as a result.

An alternative to the use of IS to overcome matrix effects is supported on implementation of advanced sample preparation methods [18,19] with capability to suppress the contribution of matrix effects. In this sense, automated solid-phase extraction (SPE) has proved a high efficiency as high accuracy and precision can be achieved under optimum operation conditions. A number of studies involving on-line SPE coupled to LC–MS/MS analysis have been reported for determination of different families of compounds [20]. In this research, the differences found by using or not an SIL-IS in SPE–LC–MS/MS for quantitative analysis of hydroxy-eicosatetraenoic acids (HETEs) in human serum are revealed. The target metabolites were 8, 11, 12 and 15HETEs, while the selected IS was a stable isotopic hydroxyeicosatetraenoic acid (HETE) isomer. Thus, the physico-chemical properties of the IS were the basis for evaluation of the analytical features of the method for each metabolite.

2. Materials and methods

2.1. Reagents

Deionized water (18mΩ·cm) from a Millipore Milli-Q water purification system was used to prepare all aqueous solutions. The standards used —all from Cayman Chemicals (Ann Arbor, MI, USA)— were hydroxyeicosatetraenoic acid isomers 8HETE (8-hydroxy-5Z,9E,11Z,14Z-eicosatetraenoic acid), 11HETE (11-hydroxy-5Z,8Z,12E, 14Z-eicosatetraenoic acid) 12HETE (12-hydroxy-5Z,8Z,-10E,14Z-eicosatetraenoic acid) and 15HETE (15-hydroxy-5Z,8Z,11Z,13E-eicosate-traenoic acid). 15HETE-d8 (eight deuterium atoms at 5, 6, 8, 9, 11, 12, 14, and 15 positions) was used as IS for quantitative analysis of the target analytes. Formic acid and acetonitrile (ACN) from Scharlab (Barcelona, Spain) were used for chromatographic separation. All chemicals were LC grade and used without further purification.

2.2. Blood extraction and serum isolation

Venous blood was collected in evacuated sterile serum tubes without additives (Vacutainer, Becton Dickinson, Franklin Lakes, NJ, USA) from healthy donors and post-menopausal women diagnosed with osteoporosis. Blood samples were incubated for 30 min at room temperature to allow coagulation, then centrifuged at 2 000 × g for 15 min at 4 °C to isolate the serum fraction, and processed within 2 h after collection. Serum was placed in plastic ware tubes and stored at −80 °C until analysis. All steps from blood extraction to analysis were performed in compliance with the guidelines dictated by the World Medical Association Declaration of Helsinki (2004), which were supervised by the Ethical Review Board (ERB) of Reina Sofía Hospital (Córdoba, Spain) that approved the experiments. Individuals selected for this study were informed to obtain consent prior to sample extraction.

A serum pool was prepared with serum isolated from osteoporotic women for development of the calibration models and validation steps. Aliquots from this pool were spiked with the target metabolites at three concentrations (15, 35 and 55

ng/mL) according to the normal serum content described in the Human Metabolome Database (www.hmdb.ca). Non-spiked aliquots were also analyzed to elucidate the natural content of these metabolites in the original pool. The samples from healthy donors were subjected to the protocol as independent samples.

2.3. Instruments

SPE was performed with an automated workstation Symbiosis-Pharma (Spark Holland, Emmen, The Netherlands) equipped with an autosampler Reliance (Spark Holland) furnished with an 100 μL sample loop and a refrigerated stacker sample compartment. Supplementary Fig. 1 shows the scheme of the different units forming the instrumental configuration. The SPE workstation is endowed with a unit for SPE cartridges exchange—automatic cartridge exchanger (ACE) and two high-pressure dispensers (HPD) for SPE solvents delivery. Peek tubes of 0.25 mm i.d. (VICI, Houston, Texas, USA) were used for all connections. The ACE unit included two clamp valves. The Sparklink 3.10 SP3 software was used to control the system. Polaris C18-A cartridges (8 μm particle size, 10 × 2.0 mm^2, Agilent) were used as sorbent material in the SPE step. Chromatographic separation was performed by an Agilent (Palo Alto, CA, USA) 1200 Series chromatograph equipped with a Mediterranea Sea C18 analytical column (3 mm particle size, 150 mm × 4.6 mm, Teknokroma, Barcelona, Spain) thermostated at 25°C. Detection was carried out by an Agilent 6460 triple quadrupole mass spectrometer equipped with a Jet Stream Technology electrospray ion source.

2.4. Analytical protocol for determination of the target analytes

One hundred microliters of serum 1:1 diluted with water and spiked with 30 ng/mL 15HETE-d8 as internal standard was injected in the analytical system. Then, the SPE–LC–MS/MS arrangement developed the following sequence of automatic operations. The sample preparation process started with solvation of the SPE sorbent with methanol (4 mL), conditioning and equilibration with water containing 0.01% formic acid (3 mL). Then, the sample was loaded into the cartridge with water containing 0.01% formic acid (2 mL). Under these conditions, the target compounds were retained in the sorbent, which was washed with 20%

acetonitrile aqueous solution (1.4 mL) to remove interferents. The chromatographic step started by switching the left clamp valve to elute the retained analytes into the chromatographic column for 3 min. The gradient used for LC separation of the compounds increased from 40% acetonitrile (maintained for 11 min) to 80% of acetonitrile in 1 min (maintained for 8 min). A 10 min post-time was necessary to equilibrate the column prior to the next injection.

Mass spectrometry detection was performed in negative ionization mode with an electrospray source (ESI) set at 4 kV capillary voltage, 350 °C source temperature and 40 psi pressure nebulizer. Nitrogen as desolvation gas was flowed at 10 mL/min. The target eicosanoids were monitored in the selected reaction monitoring mode (SRM) for highly-selective and sensitive determination in a complex matrix such as serum.

The entire analytical process was completed within 33 min, while chomatographic separation was completed in 20 min. Therefore, the automated approach enabled chromatographic analysis of one sample overlapped with preparation of the next sample.

2.5. Quantitative analysis

Stock standard solutions of 15HETE, 11HETE, 8HETE and 12HETE (500 ng/mL) were prepared in ethanol and stored at –20 °C. Ten calibration levels were prepared by both diluting standard solutions and spiking human serum with standard solutions at known concentrations. The concentration of target analytes ranged from 0.125 to 100 ng/mL. 15HETE-d8 (at 30 ng/mL) was used as IS to obtain the calibration model for HETEs. The calibration curves were obtained by plotting the ratios of peak area of HETEs to that of IS, and subtracting the natural content of HETEs in serum. The calibration plots of HETEs were used for quantitation.

3. Results and discussion

3.1. Preliminary aspects

HETEs were chosen for their biochemical interest since these metabolites are obtained by oxidation of polyunsaturated fatty acids (mainly 20-carbon fatty acids) which play a regulation role in inflammation [21] and, therefore, they have proved to be involved in different pathological states such as cardiovascular diseases [22], certain types of cancer [23] or osteoporosis [24]. Taking into account the relationship between the inflammatory response and osteoporosis, a serum pool prepared from osteoporotic patients was used in this study to check the suitability of the internal standard for quantitative analysis of HETEs. This evaluation was supported on the analytical features estimated for each metabolite. Optimum MS parameters such as voltage for efficient filtration of precursor ions in the first quadrupole, collision energy to fragment precursor ions by collision-induced dissociation (CID) and product ions selected as quantitation for each analyte are listed in Table 1. There was a unique transition with detection capability for each HETE. However, this transition was highly selective for each isomer because the fragmentation occurred differently for each of them. As Fig. 1 shows, fragmentation of precursor ions for each HETE isomer depended on the position of the hydroxyl substituent, which enabled to differentiate the isomers by the SRM transition.

Table 1. Optimal values of the parameters involved in quantitation of HETEs by SRM mode.

Analyte	**tr** *(min)*	**Quantitation transition**	**Qualification transition**	**Collision energy** *(eV)*	**Dwell time** *(ms)*	**Voltage MS1** *(eV)*
15(S)HETE-d8	16.8	327.2→ 182.2	327.2→ 264.2	15	55	95
15HETE	16.9	319.2→ 175.2		10	55	80
11HETE	17.2	319.2→ 167.2		15	55	80
8HETE	17.4	319.2→ 155.2		10	55	80
12HETE	17.6	319.2→ 179.2		10	55	80

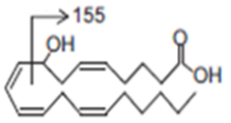
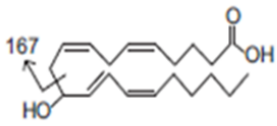
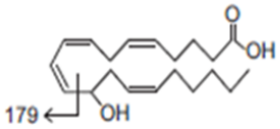
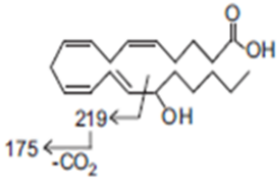
Compound	Precursor *(m/z)*	Product *(m/z)*	Structure
8HETE	319.2	155.2	
11HETE	319.2	167.2	
12HETE	319.2	179.2	
15HETE	319.2	175.2	

Fig. 1. HETEs structure with the corresponding transition.

3.2. Optimization of SPE and on-line elution to the liquid chromatograph

The complexity of biological samples demands for suitable sample preparation protocols able to remove potential interferents prior to chromatographic separation. The main limitation of previously reported methods involving extraction of eicosanoids from biological samples is the use of manual protocols. Generally, these protocols start with a liquid–liquid extraction (LLE) step, preconcentration by off-line SPE or simple evaporation and, finally, reconstitution prior to chromatographic analysis. These off-line approaches usually involve analyte losses, degradation of intermediate metabolites and time consuming protocols. Two extraction procedures such as LLE and off-line SPE have already been evaluated in a research for analysis of eighteen eicosanoids in human and monkey plasma [25]. Overally, the off-line SPE method yielded better recovery (>55%) and reproducibility (<21%) than the LLE method and, for this reason, SPE seems to be

more suited than LLE. However, automation of the SPE step could improve the analytical properties of the method with on-line separation and determination by LC–MS/MS. On-line SPE–LC–MS/MS was optimized in this research and an IS was used to evaluate the benefits, if present, of its implementation. Thus, the influence of the SPE sorbent, loading solvent, washing solvent and elution time — directly related to analytes retention, interferents removal and recovery, respectively— were studied by a univariate approach due to their discontinuous character. Additionally, loading and washing volumes and flow rate were set by a multivariate response surface design consisting of 16 experiments and 2 central points. Table 2 summarizes the variables optimized, the range studied and the optimum values.

Table 2. Optimization of the main variables involved in the SPE step.

Variable	Range tested	Optimum value
SPE sorbent	Hysphere CN, C2, C8 (EC), C18 HD, Polaris C18-A, Resin GP and Resin SH	Polaris C18-A
Sample dilution	1:1–1:5	1:1
Loading solvent		
Organic composition	From 0% to 50% of methanol and acetonitrile	0%
Acidification	From 0% to 10% of formic acid and acetic acid	0.01% formic acid
Volume	0.5–3 mL	2
Flow rate	0.2–3 mL/min	0.5
Washing solvent		
Organic composition	From 0% to 50% of methanol and acetonitrile	20% acetonitrile
Acidification	From 0% to 10% of formic acid	0% formic acid
Volume	0–3 mL	1.4
Flow rate	0.2–3 mL/min	1.3
Elution time	0.5–3.5 min	3 min

3.3. Calibration models

Four calibration models were developed under the optimum working conditions to evaluate the performance and stability of the complete approach for quantitation of HETEs. Two of the models were built with standard solutions, with and without spiking the calibration solutions with the IS. The other two models were built with the serum pool from osteoporotic patients, which was spiked at different concentrations with the target analytes (n=10 calibration levels), also with and without spiking with the IS. Non-spiked aliquots of the same pool were also analyzed to obtain the natural content of each analyte in the serum pool, which was subtracted for preparation of the calibration curves. The results from these four calibration models are shown in Table 3 (the calibration equations and regression coefficients, R^2), and Fig. 2 (the calibration curves). The slopes, intercepts and regression coefficients of the calibration curves prepared with standard solutions were characterized by a good fitting. However, strong matrix effects were detected in the calibration models prepared with serum when their intercepts were compared with those obtained by standard solutions. This phenomenon was particularly evident in the calibration models without IS normalization. As Table 3 shows, the ratio of slopes from calibration equations in serum and standard solutions approached to the theoretical value (1.0) when the IS was considered. The no inclusion of IS in the calibration model led to critical differences between the use of standard solutions and serum, which could be clearly attributed to matrix effects. Therefore, the similarity of calibration models prepared with standard solutions and serum was higher when IS was used as compared to its no inclusion. This fact is indicative of a noticeable correction effect thanks to the IS. The similarity between both calibration models with spiked IS was as follows: 8HETE (ratio 0.98), 11HETE (1.06), 15HETE (1.12) and 12HETE (1.31). The correction effect, estimated by subtracting the ratios obtained with and without IS, was as follows: 12HETE (0.21), 15HETE (0.19), 11HETE (0.14) and 8HETE (0.13). The subtraction of the ratios showed a higher IS correction for 15HETE and 12HETE than for 8HETE and 11HETE. This high correction effect could be explained by chemical similarity to the IS, 15HETE-d8, since the hydroxyl group is

farther from position 15 in 11HETE and 8HETE metabolites (see Fig.1).

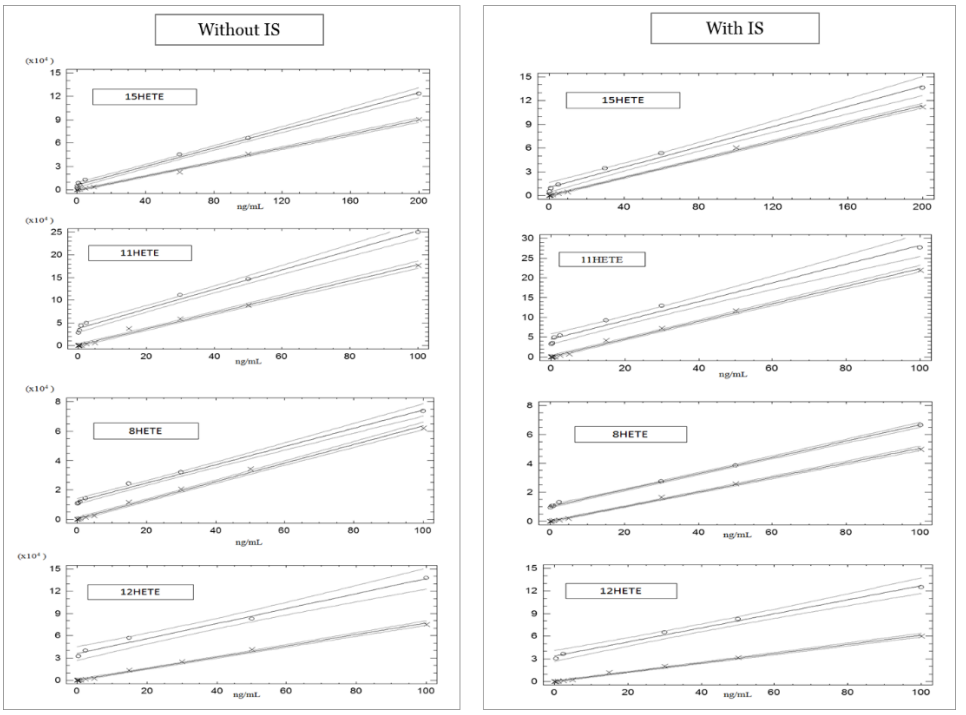


Fig. 2. Calibration curves obtained for HETES using standard solutions and spiked serum with and without IS. (x = Standard solutions, o = Serum).

3.4. Validation of the models

3.4.1. SPE efficiency

The efficiency of the SPE step was evaluated with a dual cartridge configuration based on the in-serial coupling of two cartridges through which the sample was sequentially circulated helped by the loading solution. Supplementary Fig. 2 shows the modified scheme for this configuration. In this way, the analytes non-retained in the first cartridge could be retained in the second to estimate the retention efficiency. Then, the material retained in both cartridges was sequentially eluted to the analytical column to compare the results obtained from both cartridges and, thus, to evaluate the efficiency of the SPE step.

This study was carried out with a serum pool spiked with the target analytes

Table 3. Parameters of the calibration models prepared with standard solutions and a serum pool.

Analyte	t_r		Standard solutions		Serum		Slope	
			Calibration equation	Regression coefficient R^2	Calibration equation	Regression coefficient R^2	Ratio (serum slope/standard solution slope)	Ratios subtraction
15HETE	16.9	With IS	y= 0.0568x + 0.0109	0.999	y=0.0638x + 1.0475	0.993	1.12	0.19
		Without IS	y=450.54x – 562.28	0.997	y=590.94x + 6403.6	0.995	1.31	
11HETE	17.2	With IS	y=0.2221x + 0.1338	0.996	y=0.2366x + 4.5293	0.985	1.06	0.14
		Without IS	y=1775.7x – 21.197	0.998	y=2145.2x + 38048	0.993	1.20	
8HETE	17.4	With IS	y=0.063x + 0.0463	0.996	y=0.062x + 1.2301	0.995	0.98	0.13
		Without IS	y=503.54x + 96.339	0.998	y=561.32x + 10434	0.998	1.11	
12HETE	17.6	With IS	y=0.0767x + 0.0464	0.996	y=0.1008x + 3.6031	0.990	1.31	0.21
		Without IS	y=610.69x + 398.6	0.995	y=930.45x + 33850	0.993	1.52	

at three concentrations (15, 35 and 55 ng/mL). Fig. 3 shows the retention efficiencies, expressed as percentage, obtained for each analyte as a function of the spiked concentration, with values above 90.0% for all target analytes, which is indicative of an efficient retention in the first cartridge. The values of relative standard deviation were below 2.0%; therefore, the SPE protocol provided high retention efficiency for the target compounds in the first cartridge.

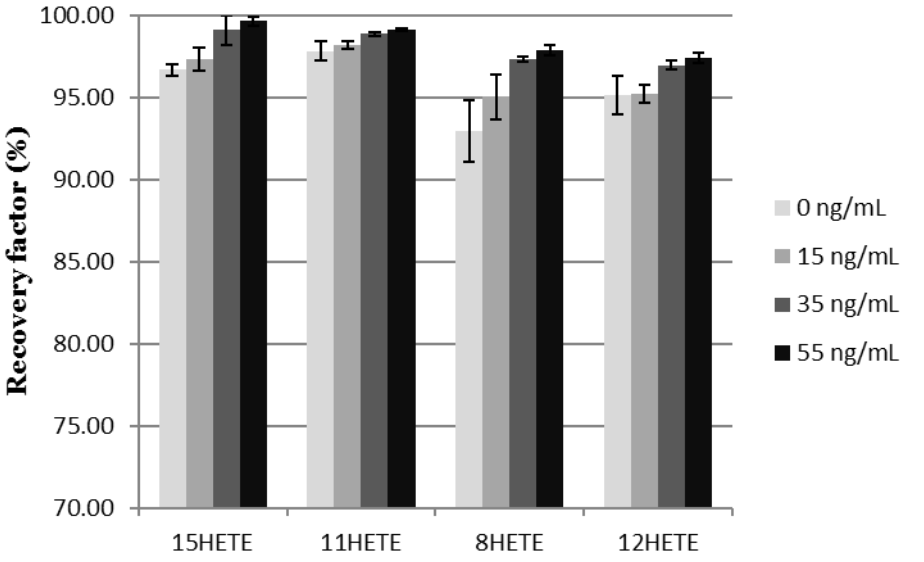


Fig. 3. Recovery for each analyte calculated with a two-cartridge configuration.

3.4.2. Sensitivity

The sensitivity of the method was evaluated by estimation of the detection and quantitation limits both with standard solutions and serum samples spiked at low concentrations. This analytical feature was assessed by estimation of the limits of detection (LOD) and quantitation (LOQ) for the analytes by injecting dilution series of each in serum to obtain the concentrations which provided signals three and ten times the background noise, respectively. As shows Table 4, the LODs in the calibration model prepared with spiked serum ranged between 75 and 151 pg/mL, while the LOQs were from 0.25 to 0.5 ng/mL. On the other hand, the sensi-

tivity was clearly improved when standard solutions were used to prepare the calibration model, except for 15HETE, as shown in Table 4. Matrix effects justify the loss of sensitivity, which was especially relevant for 12HETE. This compound was the last eluted from the analytical column (17.6 min) and, therefore, the most seriously affected by ionization suppression effects owing to lipophilic substances. Comparison with previous methods based on off-line SPE for the same analytes revealed an improvement in sensitivity terms by using the on-line configuration. In fact, LODs (from0.075 to 0.151 ng/mL) were lower than those obtained by methods based on off-line SPE (range from 0.084 to 0.448 ng/mL) [26].

Table 4. Limits of detection and quantitation for each analyte (expressed as pg/mL) obtained by the calibration models prepared with standard solutions and serum.

Analyte	Standard solutions		Serum	
	LODs	LOQs	LODs	LOQs
15HETE	75	250	75	250
11HETE	38	125	75	250
8HETE	38	125	75	250
12HETE	38	125	151	250

3.4.3. Precision

The precision of the method was evaluated by an experimental strategy intended to estimate intra-day variability (by repetitive analyses within the same day) and inter-days variability (by analyses for three consecutive days), as recommended by the Food and Drug Administration (FDA) [27]. The study was carried out with serum aliquots spiked at three concentrations and analyzed in triplicate. Precision, calculated as relative standard deviation (RSD) and expressed as percentage, was evaluated with and without internal standard correction. Table 5 shows the values obtained for the three experimental sets as a function of the spiked concentration. While no substantial effects on intra-day variability were

observed by IS correction when added at 15 or 35 ng/mL, an increased variability occurred when the IS was at 55 ng/mL. Inter-days variability was clearly improved when the IS was at 15 ng/mL, slightly improved for 11HETE, 8HETE and 12HETE at 35 ng/mL IS and deteriorated at 55 ng/mL. Nevertheless, the RSDs obtained by IS correction were always below 9.0%, which is a quite acceptable value in clinical analysis.

3.4.4. Accuracy

Accuracy estimation is crucial to evaluate matrix effects in methods involving sample preparation protocols, especially those based on SPE steps. For this purpose, serum aliquots spiked at three concentrations were analyzed using the instrumental configuration based on a single cartridge. Non-spiked serum aliquots were also analyzed in this study to take into account the presence of the target metabolites in the serum pool. Accuracy was calculated by the recovery factor using serum calibration curves with and without IS. The results obtained by both models are listed in Table 6. As can be seen, the recovery factor was significantly improved when calibration models including the IS were used. In fact, the ionization suppression is significant when IS correction is not applied. For HETE isomers the recovery factors were above 83% at 35 and 55 ng/mL, except for 12HETE which yielded a 72% recovery factor at the highest spiked concentration of IS. At 15 ng/mL IS the recovery factors were below 78.5% for all HETE isomers.

With these premises, it can be said that the IS allows noticeable correction of matrix effects (*viz.,* ionization suppression and decreased SPE retention efficiency) for some HETE isomers thus improving the analytical features of the method. Therefore, quantitation of 8HETE, 11HETE and 15HETE was properly supported on analytical parameters at medium concentrations of them (35 and 55 ng/mL), while the analysis of 12HETE was characterized by a low recovery factor at the three tested concentration levels. On the other hand, quantitation at low concentration of eicosanoids was characterized by a decrease of the recovery factors below 78.5%. Therefore, the implementation of correction factors would be required for

accurate quantitation at specific concentration ranges if a given accuracy is demanded.

Table 5. Intra-day and inter-day variability for each analyte in serum (at three different concentrations), expressed as % RSD, with and without IS.

Analyte	Intra-day variability (%)						Inter-day variability (%)					
	15 ng/mL		35 ng/mL		55 ng/mL		15 ng/mL		35 ng/mL		55 ng/mL	
	Without IS	With IS	Without IS	With IS	Without IS	With IS	Without IS	With IS	Without IS	With IS	Without IS	With IS
15HETE	4.2	4.1	1.3	2.3	3.6	5.7	13.6	7.0	5.7	6.3	3.8	8.4
11HETE	7.9	7.2	2.9	2.7	2.6	4.1	14.9	7.6	7.2	5.7	3.5	7.4
8HETE	2.4	2.9	7.7	5.9	2.6	1.0	12.7	5.2	8.9	5.9	4.9	7.4
12HETE	5.3	4.4	1.8	2.9	3.1	3.0	9.8	5.9	4.5	4.3	3.1	4.9

Table 6. Accuracy estimation for each HETE usin the single-cartridge configuration in serum spliked at three concentrations. The given values correspond to the recovery factor calculated for each added concentration.

Analyte		Recovery factor (%)		
		15 ng/mL	35 ng/mL	55 ng/mL
15HETE	Without IS	43.3	53.9	61.8
	With IS	73.0	85.2	83.5
11HETE	Without IS	39.4	61.4	72.1
	With IS	72.5	97.3	93.3
8HETE	Without IS	41.7	55.7	72.2
	With IS	72.6	87.7	95.5
12HETE	Without IS	54.6	50.0	63.3
	With IS	78.5	76.1	72.4

4. Conclusion

The effect of 15HETE-d8 as SIL-IS in the quantitative analysis of HETEs has been checked by SPE–LC–MS/MS. Calibration models prepared with serum aliquots revealed important matrix effects for HETE isomers, which were confirmed by accuracy estimation. Nevertheless, quantitative analysis of these metabolites can be assessed by application of the recovery correction factor and using an internal standard. In summary, the SIL-IS selected in this research was crucial for quantitative analysis of the target HETEs by SPE–LC–MS/MS, which demand for application of a recovery correction factor to compensate for matrix effects affecting retention of analytes in the sorbent and ionization suppression.

Acknowledgments

The Spanish Ministerio de Economía y Competitividad (MINECO) and FEDER Program are thanked for financial support through Project (CTQ2012-

37428). F. Priego-Capote is also grateful to the MICINN for a Ramón y Cajal Contract (RYC-2009-03921).

References

[1] E. Stokvis, H. Rosing, J.H. Beijnen, Rapid Commun. Mass Spectrom. 19 (2005) 401–407.

[2] F. Xie, T. Liu, W.J. Quian, V.A. Petyuk, R.D. Smith, J. Biol. Chem. 286 (2011) 25443–25449.

[3] A. Bergeron, M. Furtado, F. Garofolo, Rapid Commun. Mass Spectrom. 23 (2009) 1287–1297.

[4] B.K. Matuszewski, J. Chromatogr. B 830(2) (2006) 293–300.

[5] S. O'Halloran, K.F. lett, Clin. Chem. 54(8) (2008) 1386–1389.

[6] A. Tan, S. Hussain, A. Musuku, R. Massé, J. Chromatogr. B 877 (27) (2009) 3201–3209.

[7] R.K. Boyd, Rapid Commun. Mass Spectrom. 7 (1993) 257–271.

[8] H. Rosing, W.Y. Man, E. Doyle, A. Bult, J.H. Beijnen, J. Liq. Chromatogr. Rel. Technol. 23 (2000) 329–354.

[9] A.V. Eeckhaut, K. Lanckmans, S. Sarre, I. Smolders, Y. Michotte, J. Chromatogr. B 877 (23) (2009) 2198–2207.

[10] K.J. Bronsema, R. Bischoffa, N.C. van de Merbel, J. Chromatogr. B 893-894 (2012) 1–14.

[11] S. Wang, M. Cyronak, E. Yang, J. Pharm. Biomed. Anal. 43(2) (2007) 701–707.

[12] C. Briscoe, M. Stiles, D. Hage, J. Pharm. Biomed. Anal. 44 (2007) 484–491.

[13] K. Guo, C. Ji, L. Li, Anal. Chem. 79(22) (2007) 8631–8638.

[14] C.M. Chavez-Eng, M.L Constanzer, B.K. Matuszewski, J. Chromatogr. B 767(1) (2002) 117–129.

[15] F. Hernández, J. Sancho, O. Pozo, Anal. Bioanal. Chem 382(4) (2005) 934–946.

[16] A. Zhang, H. Sun, P. Wang, Y. Han, X. Wang, J. Proteomics 75(4) (2012) 1079–1088.

[17] S.S. Iyer, Z.P. Zhang, G.E. Kellogg, H.T. Karnes, J. Chromatogr. Sci. 42 (7) (2004) 383–387.

[18] C. Muller, P. Schafer, M. Stortzel, S. Vogt, W. Weinmann, J. Chromatogr. B 773(1) (2002) 47–52.

[19] S. Souverain, S. Rudaz, J.L. Veuthey, J. Chromatogr. A 1058 (1-2) (2004) 61–66.

[20] M. Calderón-Santiago, F. Priego-Capote, J.G. Galache-Osuna, M.D. Luque de Castro, J. Pharm. Biomed. Anal. 70 (2012) 476–484.

[21] D.W. Gilroy, Int. J. Biochem. Cell Biol. 42(4) (2010) 524–528.

[22] S. Gleim, J. Stitham, W.H. Tang, K.A. Martin, J. Hwa, Cell Mol. Life Sci. 69(20) (2012) 3361–3380.

[23] E.R. Greene, S. Huang, C.N. Serhand, D. Panigrahy, Prostaglandins Other Lipid Mediat. 96(1–4) (2011) 27–36.

[24] M. Almeid, E. Ambrogini, L. Han, S.C. Manolagas, E.L. Jilka, J. Biol. Chem. 284(40) (2009) 27438–27448.

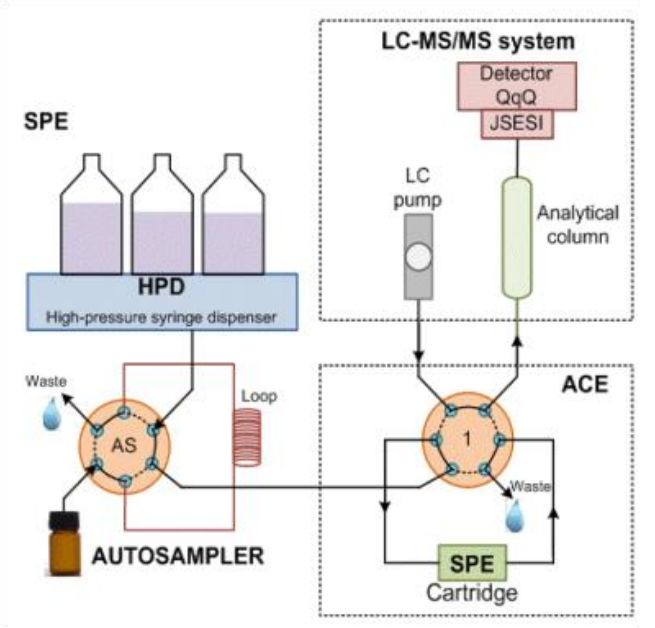
[25] D. Mangal, C.E. Uboh, L.R. Soma, Rapid. Commun. Mass Spectrom. 25 (2011) 585–598.

[26] K Strassburg, A.M.L. Huijbrechts, KA. Kortekaas, J.H. Lindeman, T.L. Pedersen, A. Dane, R. Berger, A. Brenkman, T. Hankemeier, J. van

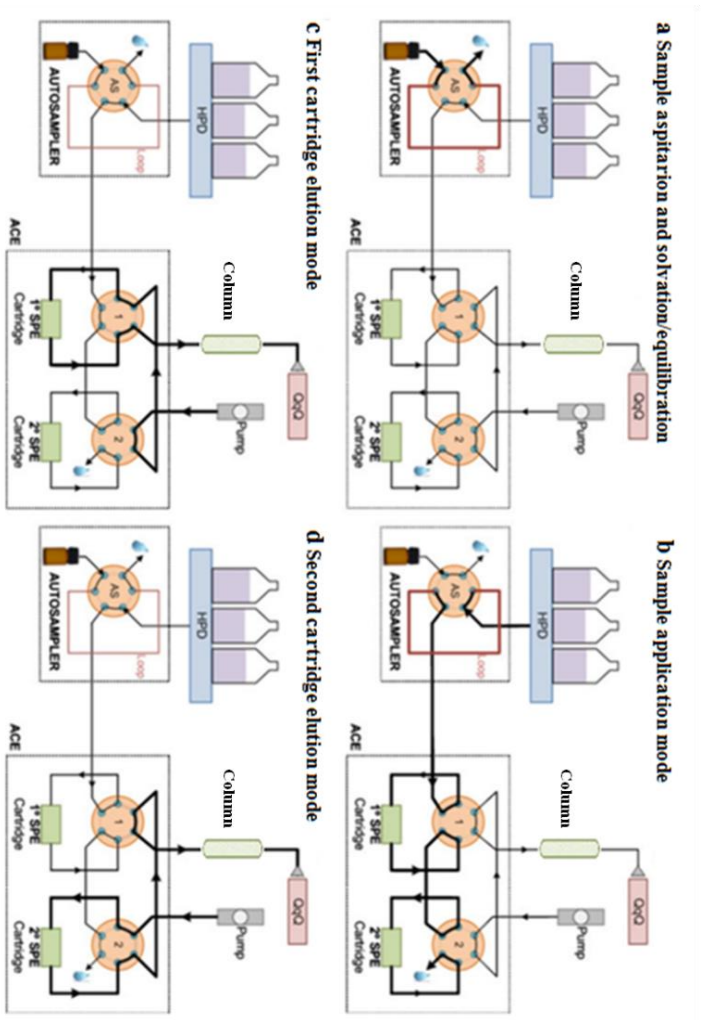
Duynhoven, E. Kalkhoven, J.W. Newman, R.J. Vreeken, Anal. Bioanal. Chem. 404 (2012) 1413–1426.

[27] Guidance for Industry Bioanalytical Method Validation, U.S. Department of Health and Human Services, Food and Drug Administration, 2001.

Supplementary material



Supplementary Fig. 1. Scheme of the device based on an SPE workstation coupled to an LC–MS/MS system. Automatic Cartridge Exchange (ACE), Electro Spray Ionization module (JSESI), High Pressure Dispenser (HPD), Triple Quadrupole Mass Detector (QqQ).



Supplementary Fig. 2. Scheme of the dual cartridge configuration with the different stages.

CHAPTER 4

Targeted analysis of omega-6-derived eicosanoids in human serum by SPE-LC-MS/MS for evaluation of coronary artery disease

Electrophoresis 34 (2013) 2901–2909

Electrophoresis

34 (2013) 2901–2909

Targeted analysis of omega-6-derived eicosanoids in human serum by SPE-LC-MS/MS for evaluation of coronary artery disease

María Auxiliadora Fernández Peralbo[1,2], Feliciano Priego-Capote[1,2], José Gabriel Galache-Osuna[3], María Dolores Luque de Castro[1,2]

[1]*Department of Analytical Chemistry, Campus of Rabanales, University of Córdoba, Córdoba, Spain*

[2]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, Córdoba, Spain*

[3]*Department of Cardiology, Hemodynamic and Interventional Cardiology, University Hospital Miguel Servet, Zaragoza, Spain*

Targeted analysis of omega-6-derived eicosanoids in human serum by SPE-LC-MS/MS for evaluation of coronary artery disease

M. A. Fernández-Peralbo, F. Priego-Capote, J. G. Galache-Osuna, M. D. Luque de Castro

Abstract

A targeted approach has been applied to quantitative analysis of eicosanoids derived from omega-6 fatty acids in serum from individuals diagnosed with coronary artery disease (CAD). The target metabolites were series-2 prostaglandins, thromboxane B2, hydroxyeicosatetraenoic acids, and hydroxyoctadecadienoic acids. The method was based on SPE–LC–MS/MS in selected reaction monitoring mode for highly selective and sensitive determination of the target eicosanoids. The combination of SPE and LC–MS/MS involved the benefits from both direct analysis of serum without a step for protein precipitation and fully automation of the analysis. The method allowed comparison of omega-6-derived eicosanoids in serum from patients diagnosed with CAD and from control individuals. The effect of treatment with aspirin on the profile of the target compounds was evaluated through its incidence on the different pathways. Finally, the serum levels of the target metabolites in patients diagnosed with CAD were also statistically examined according to the severity of the coronary lesion stratified as stable angina, non-ST-elevation acute coronary syndrome, and acute myocardial infarction.

Keywords Arachidonic acid; Coronary artery disease; Eicosanoids; Hydroxyeicosatetraenoic acids; Metabolomics

1. Introduction

Linoleic acid (C18:2n-6) and arachidonic acid (AA, C20:4n-6) are essential polyunsaturated fatty acids which belong to omega-6 series that act as substrates for a set of oxygenase enzymes such as cyclooxygenases (COX), lipoxygenases (LOXs), and cytochrome P450 oxidase (CYP). These enzymes promote the formation of eicosanoids which encompass several families such as prostaglandins (PGs), prostacyclins, thromboxanes (TXs), leukotrienes, lipoxins, and hydroperoxy- and hydroxy-fatty acids [1]. The clinical relevance of eicosanoids is their involvement in a great variety of biological processes, among which the inflammatory response is worth emphasizing [2]. Thus, COX-derived eicosanoids such as PGs and TXs play key roles in the regulation of the cardiovascular function [3] and tumor progression [4]. In contrast, LOX-derived eicosanoids such as leukotrienes (LTs) and hydroxyeicosatetraenoic acids (5, 12 and 15HETE) are considered important mediators of inflammation, asthma, cardiovascular disease (CVD), and cancer [5]. Epoxidation of AA by CYP enzymes generates 5,6-, 8,9-, 11,12-, and 14,15-epoxyeicosatrienoic acids (EETs), which have shown a reduction effect of ischemia–reperfusion injury, to exert strong anti-inflammatory effects on the vasculature, and be potent vasodilators in the coronary circulation [6]. These pathways (schemed in Supporting Information Fig. 1) are the target of well-known drugs for treatment of inflammation, pain, asthma, allergies, and cardiovascular disorders.

Coronary artery disease (CAD) is a complex disease of the arterial walls in which the inflammatory cascade seems to be especially involved. In this disease, a sclerotic plaque of lipids and fibrous tissue is deposited over time in the walls of arteries, often leading to rupture and thrombus formation. Occlusion of the artery lumen can lead to ischemia by decrease of oxygen supply to tissues. Such vascular lesions create a depot for circulating lipids, prompting an immune response, and developing feedback amplification of inflammatory mediators further enhancing material deposition [7]. As the sclerotic plaque remains relatively innocuous while stable, the onset of a thrombotic event is highly unpredictable in both occurrence and severity. Chest pain is usually the first symptom of the ischemia, which,

depending on frequency and persistence, can be caused at different stages in the evolution of atherosclerotic coronary disease [8]. These stages can be classified into stable chest angina [9], also known as chronic ischemic cardiopathy; and acute coronary syndrome, which involves atherothrombosis by partial or total occlusion of arteries [10]. In turn, the acute coronary syndrome encompasses unstable angina, non-ST-elevation acute coronary syndrome (NSTE-ACS) [11], and acute myocardial infarction (AMI) [12]. Unstable angina is characterized by the same symptoms as stable angina but with a greater frequency and intensity; while NSTE-ACS is based on formation of a thrombus but without necrosis of myocardial cells (therefore necrosis markers are not elevated), and AMI involves necrosis.

Traditional cardiovascular risk factors (*viz.,* hypercholesterolemia, obesity, hypertension, smoking, diabetes mellitus, gender, and age) remain the most secure predictors of CAD [13]. An invasive method such as catheterization is the current tool employed for diagnosis of CAD. However, it is desirable to find tests with capability to aid clinicians for prognosis/diagnosis of atherosclerosis and, particularly, to prevent episodes of AMI. Due to the involvement of the inflammatory cascade, eicosanoids should be considered important regulators in CAD. Thus, drugs that directly or indirectly modulate these signals, including COX and LOX inhibitors, have proven to play major roles in the atherothrombotic process [13].

LC–MS/MS is an excellent tool to monitor products of the eicosanoid degradation pathways as reported in the literature [14–17]. This arrangement is especially useful because, apart from avoiding derivatization, the analysis can be automated by on-line sample preparation and determination, particularly when sample preparation is based on SPE. Additionally, the combination of SPE and LC–MS/MS in on-line configuration minimizes or eliminates the involvement of the analyst in sample preparation tasks. For these reasons, an automated SPE–LC–MS/MS configuration [14] was selected to analyze eicosanoid metabolites from omega-6 fatty acids in human serum from individuals diagnosed with CAD. In this way, serum levels of eicosanoids in these patients could be compared with those obtained in control individuals, but also among three groups of patients: those diagnosed with stable angina, with NSTEMI and those affected by AMI.

2. Materials and methods

2.1. Materials, reagents, and chemicals

Deionized water (18 mΩ·cm) from a Millipore Milli-Q water purification system was used to prepare all aqueous solutions. All standards used in this research were from Cayman Chemicals (Ann Arbor, MI, USA): PGs PGE2 (7-[3-hydroxy-2-(3-hydroxyoct-1-enyl)-5-oxo-cyclopentyl]hepta-5-enoic acid), PGE1 (9-oxo-11α,15S-dihydroxy-prost-13E-en-1-oic acid), PGF2α (9α,11α,15S-trihydroxy-prosta-5Z,13E-dien-1-oic acid), and15-ketoPGF2α (9α,11α-dihydroxy-15-oxo-prosta-5Z,13E-dien-1-oic acid); TX B2 (9α,11,15S-trihydroxythromba-5Z,13E-dien-1-oic acid, TBX2); hydroxyoctadecadienoic acid isomers 9-HODE (9- hydroxy-10E,12Z-octadecadienoic acid) and 13-HODE (13-hydroxy-9Z,11E-octadecadienoic acid); and HETE isomers 5HETE (5-hydroxy-6E,8Z,11Z,14Z-eicosatretraenoic acid), 8HETE (8-hydroxy-5Z,9E,11Z,14Z-eicosatetraenoic acid), 11HETE (11-hydroxy-5Z,8Z,12E,14Z-eicosatetraenoic acid), 12HETE (12-hydroxy-5Z,8Z,10E,-14Z-eicosatetraenoic acid), and 15HETE (15-hydroxy-5Z,8Z,11Z,13E-eicosatetraenoic acid); 15HETE-d8 (eight deuterium atoms at the 5, 6, 8, 9, 11, 12, 14, and 15 positions) was used as internal standard. Formic acid and ACN from Scharlab (Barcelona, Spain) were used for chromatographic separation. All chemicals were LC grade and used without further purification.

2.2. Cohort selected for the study

A total of 45 patients affected by a significant coronary artery lesion as atherosclerosis were recruited at the Cardiology Unit of Miguel Servet Hospital (Zaragoza, Spain). They were selected after evaluation through cardiac catheterization which resulted in angiographic stenosis with a reduction of the arterial lumen≥70%. These patients were clinically diagnosed by three different manifestations such as stable angina, NSTE-ACS, or AMI. On the other hand, 16 serum samples from healthy individuals were used as controls. The main characteristics of the cohort are shown in Table 1.

Table 1. Features of the cohort under study.

Characteristic	**Samples** (n=61)	**Controls** (n=16)	**Coronary lesion patients** (n=45)	**Fisher's exact test** *(p-value)*
Age	59.86 ± 17.7	34.7 ± 6.0	68.2 ± 11.1	–
Male gender, n (%)	43 (70.5)	5 (31.3)	38 (84.4)	0.0002**
Smoking habit, n (%)	24 (39.3)	2 (12.5)	22 (48.9)	0.0158*
Obesity, n (%)	10 (16.4)	0 (0.00)	10 (22.2)	0.0506*
h-cholesterol level, n (%)	23 (37.7)	0 (0.00)	23 (51.1)	0.0002**
Diabetes, n (%)	7 (11.5)	0 (0.00)	7 (15.5)	0.174
Characteristic	**Samples** (n=30)	**AMI** (n=15)	**Stable Angina patients** (n=15)	**Fisher's exact test** (*p-value)*
Age	67.6 ± 9.6	69.2 ± 10.5	65.9 ± 8.6	–
Male gender, n (%)	27 (90)	13 (86.6)	14 (93.3)	1
Smoking habit, n (%)	16 (53.3)	7 (46.6)	9 (60)	0.7152
Obesity, n (%)	6 (20)	4 (26.6)	2 (13.3)	0.6513
h-cholesterol level, n (%)	17 (56.6)	9 (60)	8 (53.3)	1
Diabetes, n (%)	6 (20)	4 (26.6)	2 (13.3)	0.6513
Characteristic	**Patients** (n=30)	**AMI** (n=15)	**Unstable Angina** (n=15)	**Fisher's exact test** (*p-value)*
Age	69.4 ± 12.1	69.2 ± 10.5	69.6 ± 13.8	–
Male gender, n (%)	24 (80)	13 (86.6)	11 (73.3)	0.6513
Smoking habit, n (%)	24 (43.3)	7 (46.6)	6 (40)	1
h-cholesterol level, n (%)	15 (50)	9 (60)	6 (40)	0.4661
Diabetes, n (%)	5 (16.6)	4 (26.6)	1 (6.6)	0.3295

* *p*- value between 0.05 and 0.01.

***p*- value lower than 0.001.

2.3. Blood extraction and serum isolation

Venous blood was collected in evacuated sterile serum tubes without additives (Vacutainer, Becton Dickinson, NJ, USA) and incubated for 10min at room temperature to allow coagulation. Then, the tubes were centrifuged at 2000 × *g* for 15 min at 4 °C to isolate the serum fraction (processing within 1 h after collection). Serum was placed in plastic ware tubes and stored at −80 °C until analysis.

All steps from blood extraction to analysis were performed in compliance with the guidelines dictated by the World Medical Association Declaration of Helsinki

(2004), and supervised by specialized personnel from Miguel Servet Hospital. The individuals selected for this study were previously informed to obtain consent.

2.4. Instruments and apparatus

Hyphenated SPE was performed with an automated robotized workstation Symbiosis-Pharma (Spark Holland, Emmen, The Netherlands) equipped with an autosampler Reliance (Spark Holland) furnished with an 100 µL sample loop and a refrigerated stacker sample compartment (Supporting Information Fig. 2). The SPE workstation is endowed with a unit for SPE cartridge exchange—automatic cartridge exchanger and two high-pressure dispensers (HPD) for SPE solvents delivery. Peek tubes of 0.25 mm id (VICI, Houston, Texas, USA) were used for all connections. The automatic cartridge exchanger unit includes three switching valves, one of them equipped with a T-rotor, and two clamps. The Sparklink 3.10 SP#3 software was used to control the system. Polaris C18-A cartridges (8 µm, 10 × 2.0 mm, Spark Holland) were used in the SPE step. Chromatographic separation was carried out by an Agilent (Palo Alto, CA, USA) 1200 Series chromatograph composed by a binary pump and a Mediterranea Sea C18 analytical column (3 µm, 150 mm × 4.6 mm, Teknokroma, Barcelona, Spain) that was thermostated at 25 °C. Detection was carried out by an Agilent 6410 Triple quadrupole mass spectrometer, equipped with a Jet Stream Technology electrospray ion source.

2.5. Analytical protocol for determination of the target analytes

One hundred microliters of serum 1:1 diluted with water and spiked with 15HETE-d8 at 30 ng/mL is injected in the analytical system. Then, the SPE–LC–MS/MS analysis is initiated with the following sequence of automatic operations. First, the sample preparation process starts with solvation of the stationary phase of the cartridge with methanol (4 mL), conditioning and equilibration with water containing 0.01% formic acid (3 mL), and then, the sample is loaded into the cartridge with water containing 0.01% formic acid (2 mL). Under these conditions, the target compounds are retained in the sorbent, and 20% ACN aqueous solution (1.4 mL) is used as washing solution to remove non-retained serum components. The chromatographic step starts by switching the left clamp valve and eluting for

3 min the content of the cartridge into the chromatographic column. The gradient used for LC separation of the compounds goes from 40% ACN (maintained for 11 min) to 80% ACN in 1 min (maintained for 8 min) pumped at 1.0 mL/min. A post time of 10 min was necessary to equilibrate the column prior to the next injection.

MS detection is performed in negative ESI mode with 4 kV capillary voltage, 350 °C source temperature and 40 psi pressure nebulizer. Nitrogen as desolvation gas is flowed at 10 mL/min. The selected voltage for efficient filtration of precursor ions in the first quadrupole, collision energy to fragment the precursor ions by CID and generate the quantification and qualifier ions in the selected reaction monitoring SRM) mode and the dwell time were optimized previously. Table 2 shows the optimum values found in this study. HETE and HODE compounds only present a transition with detection capability, which is selective for each isomer because fragmentation is related to the position of the double bond. The entire analytical process is completed within 33 min. The automated system enables the chromatographic analysis of one sample while the next sample is prepared.

2.6. Statistical analysis

Statgraphics Centurion (XV.I version, Statpoint Technologies), R (URL http://www.R-project.org), pROC (1.5.1 version, URL http://web.expasy.org/pROC/) and Mass Profiler Professional (MPP) software package (version 2.0, Agilent Technologies, Santa Clara, CA, USA) were used for data analysis.

Fisher's exact test was used to define the selected cohort according to categorical variables such as age, smoking habit, obesity, cholesterol level, and diabetes. The Sapphiro-Wilks test was used to check if the levels of eicosanoids in serum showed or not normal distributions in the cohort selected for the study.

Table 2. Optimal values parameters used for the quantification of eicosanoids by MRM mode.

Compound Name	**TR** *(min)*	**Quantitation transition** *Precursor (m/z)-> Product (m/z)*	**Qualification transition** *Precursor (m/z)-> Product (m/z)*	**Collision Energy** *(eV)*	**Dwell time** *(ms)*	**Voltage MS1** *(eV)*
Thromboxane	4.2	369.3->195.3	369.3->169.3	10	75	80
PGE3	4.95	349.3->313.3		10	75	90
			349.3->269.3	15	75	80
PGF2a	5.1	353.3->309.3	353.3->193.3	25	75	90
PGE2	6.6	351.3->271.3		10	50	70
			351.3->189.3	15	50	80
PGE1	7	353.3->317.3	353.3->235.3	10	50	70
15-Keto-PGF2a	7.2	351.3->315.3	351.3->219.3	10	50	80
PGD2	8.1	351.3->233.3		5	50	90
13-HODE	16.75	295.2->195.2		15	55	80
9-HODE	16.75	295.2->171.2		15	55	80
15(*S*)-HETE-8*d*	16.85	327.3->264.3	327.3->182.3	15	55	95
15-HETE	16.95	319.2->175.2		10	55	80
11-HETE	17.2	319.2->167.2		15	55	80
8-HETE	17.4	319.2->155.2		10	55	80
12-HETE	17.55	319.2->179.2		10	55	80

ANOVA test was applied to evaluate significance of target metabolites characterrized by a normal distribution. For those metabolites with a concentration profile noncharacterized by a normal distribution normalization was applied by a Box-Cox transformation prior to ANOVA. Apart from that, a fold-change analysis was applied to correlate the significance of metabolites with relative change in concentration levels.

A receiver-operating characteristic (ROC) curve was determined for each individual predictor by setting as threshold predicting the occurrence of a coronary lesion with specificity>90%. Partial ROC area under the curve (pAUC) [18,19] were calculated by adapting algorithms previously reported [20]. pAUCs were restricted between 90 and 100% specificity, considering an efficient predictor in clinical practice should be able to identify clearly at least nine out of ten patients with a favorable prognosis for negative tests.

3. Results and discussion

3.1. Characterization of the SPE–LC–MS/MS method

The analytical signal used in all instances was the ratio between the peak area provided by the target metabolite and that of the internal standard to minimize potential matrix effects and other variability sources. Sensitivity was evaluated because of its relevance in this method aimed at quantifying low concentrated metabolites, as are eicosanoids, in a biofluid with a complex matrix such as serum. This analytical feature was assessed by estimation of LODs and LOQs for each analyte in SRM chromatograms obtained from serum samples on the basis of a minimal accepted value of S/N ratio of 3 and 10, respectively. For metabolites that were non-detected in serum, the sensitivity was estimated by spiking serum samples at low concentrations (15, 35, and 55 ng/mL). As Table 3 shows, the LODs ranged between 15 pg/mL and 8.6 ng/mL, while LOQs were from 52 pg/mL to 28.6 ng/mL.

Another analytical feature of interest was the precision of the method, which was studied by analysis of three replicates of human serum samples spiked with the target metabolites at three known concentrations (15, 35, and 55 ng/mL). The repeatability, evaluated by analysis of five replicates of each concentration, ranged between 3.04 and 22.32% (expressed as RSD). The reproducibility, estimated by analysis of five replicates of each concentration in consecutive days, ranged between 4.33 and 25.16%.

Finally, the method was characterized in terms of accuracy, which was evaluated by relative comparison of spiked and non-spiked serum samples. A dual configuration with two on-line coupled cartridges was designed to calculate the accuracy of the method expressed as recovery factor. In this configuration, two cartridges are put in serial so the sample passes sequentially through both. Then, the retained metabolites are sequentially eluted from both cartridges to the LC–MS/MS system for independent analysis. In this way, it is possible to differentiate the amount of target metabolites retained in each of the cartridges and to estimate the recovery factor. This was calculated by dividing the amount of each analyte retained in the first cartridge into the total amount retained in both cartridges. The recovery factors ranged from 84.9 to 100%, which evaluates the retention efficiency of the SPE sorbent under the optimum operating conditions. Apart from

that, this test also enabled to evaluate the clean-up efficiency by washing matrix components, which could interfere in the MS detection. For this purpose, the concentration of the target analytes was calculated in the spiked samples and the found values were compared with spiked concentrations after subtraction of background concentrations. The accuracy results calculated in this way and expressed as percentage ranged from 67.4 to 104.59%.

Table 3. Characteristics of the method.

Analyte	**LOD** *(ng/mL)*	**LOQ** *(ng/mL)*	**Recovery** *(%)*	**Repeatability** *(RSD %)*	**Reproducibility** *(RSD %)*
Thromboxane	0.53	1.77	99.12	11.52	12.71
PGF2α	8.57	28.58	104.59	20.82	25.16
PGE2	0.02	0.068	84.75	18.49	18.89
PGE1	0.015	0.052	81.32	12.07	14.16
15-Keto-PGF2α	0.05	0.166	67.40	20.28	22.88
PGD2	1.67	5.58	87.90	22.32	22.41
15-HETE	2.04	9.64	87.29	1.13	6.35
11-HETE	0.0165	0.055	86.20	2.67	5.79
8-HETE	0.0261	0.087	92.01	4.68	5.95
12-HETE	0.044	0.1469	88.32	3.04	4.39
13-HODE	0.025	0.084	85.20	4.78	5.89
9-HODE	0.028	0.091	84.62	3.98	4.75

3.2. Characteristics of the selected cohort

The cohort selected for this study was influenced by different variability factors associated to individuals. Thus, the average age was 68.2 ± 11.1 years with 84.4% of male individuals. CAD is strongly influenced by risk factors such as hypercholesterolemia, smoking habit, obesity, and diabetes, among others. Table 1 shows the incidence of these factors on the cohort studied here. Thus, 48.9% of individuals were active smokers, 22.2% were obese, 51.1% were affected by hypercholesterolemia, and 15.5% by diabetes. However, it is worth mentioning that all individuals were diagnosed with CAD by a clinical test such as catheterization. With this preliminary report, a first data treatment was aimed at evaluating the influence of these variability factors on the cohort selected for this study. This was carried out by the Fisher exact test, which allowed analyzing the statistical weight of these factors to compare eicosanoid levels between control individuals and coronary disease patients. This test revealed the prevalence of four factors (see

Table 1): smoking habit (p=0.0158) and, specially, gender (p=0.0002) and hypercholesterolemia (p=0.0002). The gender significance is characterized by a greater presence of male individuals in the population of injured individuals. Hypercholesterolemia is a high-significant risk factor in the occurrence of CAD. In fact, high cholesterol levels are rather associated to the presence of coronary lesions due to deposition of lipids as atheroma plaques in arterial walls [21]. In addition, hypercholesterolemia enhances the synthesis of TX [22].

Concerning the smoking habit, there are some studies which pointed out the connection between smoking habit and CVDs [23]. Smoking is not only a risk factor to trigger coronary diseases [23], but its influence is also enhanced when combined with other risk factors such as obesity and high cholesterol level [24,25]. The synthesis of LOX metabolites in granuloma is increased in smokers, who also show significant qualitative and quantitative differences in the profile of eicosanoids [26]. It is well known that these compounds modulate a series of pathways with influence on hypertension [27], tachycardia [28], reduced cardiac output [29], cholesterol accumulation [30], vessel damage [31], and clot formation [32]. Also, HETE levels are elevated in chronic smokers with no further increase by acute smoking [33]. Smoking modifies the production of each eicosanoid class and occurs predominantly through an increased inflammatory state caused by the combined effects of inhaled mediators and oxidative stress [2].

Within patients with a coronary lesion, three different groups were set according to the diagnostic: stable angina, NSTE-ACS, and AMI. The three groups were constituted by the same number of individuals. The Fisher exact test did not show statistical contribution of these risk factors, as shown in Table 1. No quantitative data (concentration of glycated hemoglobin, cholesterol level, or smoking frequency) were available for these factors to allow stratifying these categorical variables. However, there are studies that associate risk factors to the progress of CVDs. Teixeira *et al.* concluded that whole body weight and central fatness are correlated with several CVD risk factors early in life [34]. Haffner demonstrated a relationship between progression of CVD and diabetes [35]. Fortunately, most of these cardiovascular risk factors are modifiable; therefore,

timely implementation of lifestyle changes and pharmacological approaches can play a pivotal role in prevention of chronic disease [35].

3.3. Comparison of control individuals versus coronary lesion patients: univariate analysis

Serum levels of the target eicosanoids in healthy individuals and coronary lesion patients were compared to check the significance of these compounds in each category, as shown in Table 4. Although the method allowed quantitative analysis of 12 metabolites, only 5 eicosanoids were above the quantitation limit in a representative number of samples of the cohort. Fig. 1 compares SRM chromatograms from a patient diagnosed with CAD and from a control individual. The study of controls *versus* patients diagnosed with a coronary artery lesion showed that no metabolites but 11HETE were characterized by a parametric distribution according to the Sapphiro-Wilks test. For this reason, ANOVA was applied after Box-Cox transformation of the data set for normalization. This test emphasized the significance of TX and 11HETE. TX, produced by COX action, is widely involved in a number of CVDs, owing to its acute and chronic effects on vasoconstriction [36], endothelial adhesion molecule expression [37] and cell migration [38], proliferation [39], and hypertrophy [40]. 11HETE —produced from AA in endothelial cells, together with other metabolites, by COX or CYP activity— influences vascular tone, leukocyte function, and platelet aggregation [41]. The present statistical study was completed by relative quantitation of the target metabolites using a fold-change algorithm. The cut-off value was set at 2.0. Thus, TX was the only metabolite reporting a relative concentration change of 14.66, which means a highly significant effect. On the other hand, 11HETE experienced relative concentration changes of 1.18; therefore, its levels —expressed as average concentrations— were similar in both groups of individuals.

3.4. Comparison of patients diagnosed with CAD: univariate analysis

This study was focused on patients diagnosed with CAD. They were grouped into three classes according to the clinical diagnostic provided by specialized personnel (*viz.,* stable angina, NSTE-ACS, and AMI) and each group compared with

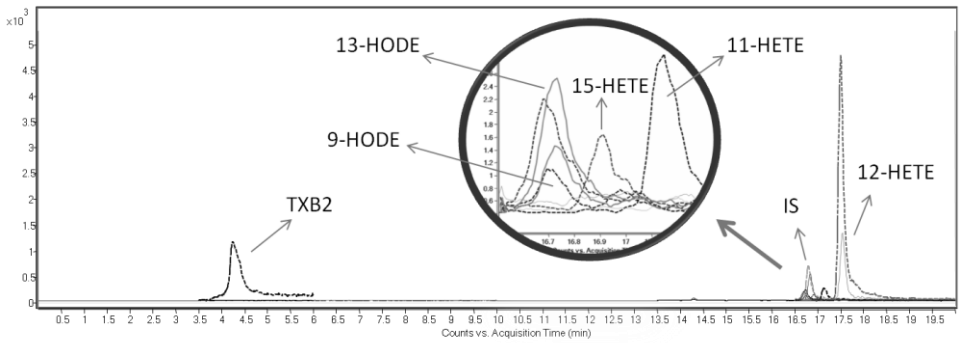


Fig. 1. Selected reaction monitoring chromatograms for a patient diagnosed with coronary artery disease (continuous line) and a control individual (segmented line).

each of the others to evaluate the discrimination capability of the target eicosanoids. In the AMI/NSTE-ACS comparison no statistical significance was detected for any metabolite by application of the ANOVA test: the profile of monitored eicosanoids did not allow discrimination between AMI and NSTE-ACS individuals. On the other hand, 12HETE was the only significant metabolite to explain differences between individuals who suffered an AMI episode and those diagnosed with stable angina with a significance of 95%, as shown in Table 4. 12HETE eicosanoid is a 12-LOX metabolite of AA involved in the inflammatory cascade. This monohydroxy fatty acid derivative is a vasoconstrictor agent that contributes to increase blood pressure and, therefore, it is involved in hypertension [42]. As synthesized by an LOX enzyme, the concentration of this compound is not affected by aspirin treatment. This compound was more concentrated in patients who suffered AMI (2.5-fold change), which is supported on previous studies demonstrating the formation of 12HETE during ischemia [43].

3.5. Evaluation of the prediction capability of monitored eicosanoids

The significance shown in statistical tests for some of the target eicosanoids was assessed by visualization of box and whisker plots, where normalized values of the peak area ratio were used. The compounds included in this study were TX, 11HETE, and 12HETE. The variability in the concentration of TXB2 was explained by the occurrence of CAD. This compound acts as the most powerful platelet aggre-

Table 4. Significance of compounds by different statistical ests.

Controls *vs* coronary lesion patients

ANALYTE	Saphiro Wilks (*p*-value)	ANOVA (*p*-value)	ANOVA after box-cox transformation (*p*-value)	Fold change
TXB2 [a]	0,002	-	**0**	14.66
13 HODE [a]	0	-	0.9272	0.99
9 HODE [a]	0	-	0.7345	0.98
11 HETE	0,067	**0.0006**	-	1.18
12 HETE [a]	1,110E-15	-	0.1386	1.53

AMI *vs* stable angina

ANALYTE	Saphiro Wilks (*p*-value)	ANOVA (*p*-value)	ANOVA after box-cox transformation (*p*-value)	Fold change
13 HODE [a]	0,00001	-	0.2443	1.17
9 HODE	0,06118	0.5162	-	1.1
12 HETE [a]	1,8102E-8	-	**0.0355**	2.5

AMI *vs* NSTE-ACS

ANALYTE	Saphiro Wilks (*p*-value)	ANOVA (*p*-value)	ANOVA after box-cox transformation (*p*-value)	Fold change
13 HODE [a]	0,0004	-	0.6691	1.07
9 HODE [a]	0,0017	-	0.3334	1.14
11 HETE	0,4555	0.9937	-	1
12 HETE [a]	8,129 E-7	-	**0.0452**	2.24

[a]Non-normal distribution of the data.

Bold type = metabolites with (95-99) % of significance

gating agent. However, the box and whisker plot for TX showed increased levels in controls as compared to patients with coronary lesion. The opposite effect can be

explained because most patients had taken aspirin, which inhibits COX enzymatic activity and reduces blood TX levels with a platelet anti-aggregating effect, even at low doses [44]. Box and whisker plots in Fig. 2 explain this behavior because the levels of TX and 11HETE were higher in control individuals than in coronary lesion patients.

TXB2, *t*-test (*p*-value=4.96*10^{-5}) 11HETE, *t*-test (*p*-value=9.52*10^{-5})

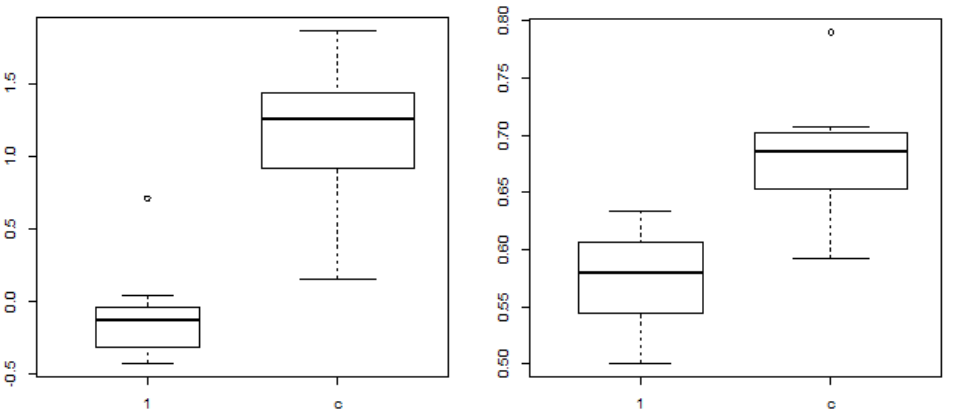


Fig. 2. Box-and-whiskers plots for TXB2 and 11HETE (0=control individuals, 1=patients with coronary artery disease).

Patients affected by AMI *versus* stable angina provided the box and whisker plot for 12HETE shown in Fig. 3. As can be seen, the cohort formed by patients after an AMI episode presented higher 12HETE levels than patients with stable angina, who are characterized by partial occlusion of the artery lumen. As previously mentioned, 12HETE possesses a vasoconstriction effect and is produced during ischemia [43]; therefore, its increase in individuals after a heart attack is justified.

The significant metabolites were additionally assessed by ROC curves with estimation of the pAUC. Evaluation of a predictor by the total AUC is not recommended when the performance test only takes place in high-specificity or high-sensitivity regions [45]. For this reason, the pAUC parameter is more suited

since it is restricted to specific regions of the curve. In this study, pAUCs were restricted between 90 and 100% specificity since an efficient predictor in clinical practice should be able to identify 90% of patients with a favorable prognosis for negative tests. In this study, only 12HETE was evaluated by one-factor ROC curves for discrimination of coronary lesion patients with AMI and stable angina, as shown in Fig. 4. The variability in the concentration of 12HETE in serum was explained by the occurrence of an AMI episode in coronary lesion patients. Evaluation of 12HETE by ROC curves for AMI patients and stable angina patients provided 90% specificity and 36% of sensitivity. These results are supported on the pAUC parameter, which was 1.4%. Once again, the result is consistent with all the studies previously developed [42, 43].

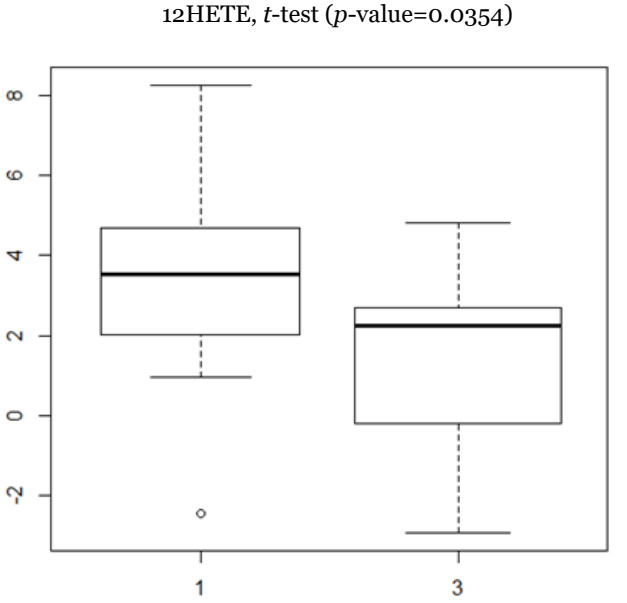


Fig. 3. Box-and-whiskers plots for 12HETE in AMI *versus* stable angina (1=AMI, 3=stable angina).

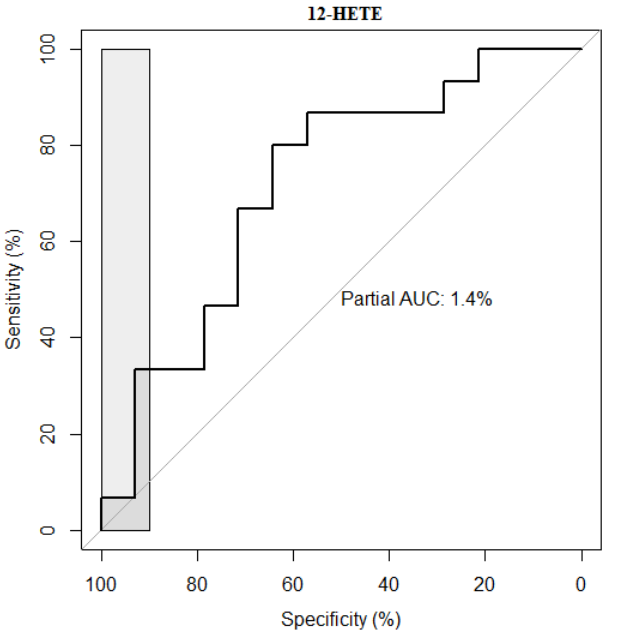


Fig. 4. ROC curve for 12HETE as predictor to discriminate individuals after an acute myocardial infarction episode from those affected by stable angina.

4. Concluding remarks

The inflammatory response of certain eicosanoids to the risk factors of CAD is the reason why these compounds should be analyzed. SPE–LC–MS/MS is one of the best analytical platforms for the analysis of eicosanoids in serum. Despite COX inhibition by aspirin intake in injured patients, HETEs and HODEs are generated by the action of the other two enzymatic pathways (LOX and CYP), which metabolize AA and linoleic acid. This research shows that 12HETE levels presented significant differences in patients who had suffered an AMI as compared to patients with stable angina. These results can help to diagnose and prevent a possible AMI in patients already injured.

Acknowledgements

The Spanish Ministerio de Ciencia e Innovación (MICINN) and FEDER Program are thanked for financial support through Project CTQ2009-07430. F. Priego-Capote is also grateful to the MICINN for a Ramón y Cajal Contract /RYC-2009-03921). Hospital Miguel Servet (Zaragoza, Spain) is also thanked for donors serum samples.

References

[1] Fischer, S., Chromatographia 1986, 22, 416–420.

[2] Scott, G., Kathleen, A.M., Jeremiah, S., John, H., Wai, H. T., Cell. Mol. Life Sci. 2012, 69, 3361–3380.

[3] Iniguez, M.A., Cacheiro-Llaguno, C., Cuesta, N., Diaz-Muñoz, M., Fresno, M., Arch. Physiol. Biochem2008, 114, 201–209.

[4] Wang, M. T., Honn, K.V., Nie, D., Cancer Metastasis Rev. 2007, 26, 525–534.

[5] Peters-Golden, M., Henderson Jr., W. R., N. Engl. J. Med. 2007, 357, 1841–1854.

[6] Node, K., Huo, Y., Ruan, X., Yang, B., Becker, M. S., Ley, K., Zeldin, D. C., Liao, J. K., Science 1999, 285, 1276–1279.

[7] Fuster, V., Moreno, P. R., Fayad, Z. A., Corti, R., Badimon, J. J., J. Am. Coll. Cardiol. 2005, 46, 937–954.

[8] Tatum, J. L., Jesse, R. L., Kontos, M. C., Nicholson, C. S., Schmidt, K. L., Roberts, C. S., Omato, J. P., Ann. Emerg. Med. 1997, 29, 116–125.

[9] Fox, K., García, M. A., Ardissino, D., Buszman, P., Camici, P. G., Crea, F., Daly, C., de Backer, G., Hjemdahl, P., López-Sendón, J., Morais, J., Pepper,

J., Sechtem, U., Simoons, M., Thygesen, K., Rev. Esp. Cardiol. 2006, 59, 919–970.

[10] Libby, P., Circulation 2001, 104, 365–372.

[11] Bassand, J. P., Hamm, C.W., Ardissino, D., Boersma, E., Budaj, A., Fernández-Avilés, F., Fox, K. A. A., Hasdai, D., Magnus-Ohman, E., Wallentin, L., Wijns, W., Eur. Heart J. 2007, 28, 1598–1660.

[12] Savonitto, S., Ardissino, D., Granger, C. B., Morando, G., Prando, M. D, Mafrici, A., Cavallani, C., Melandri, G., Thompson, T. D., Vahanian, A., Magnus-Ohman, E., Califf, R. M., Van de Werf, F., Topol, E. J., JAMA 1999, 281, 707–713.

[13] Gleim, S., Stitham, J., Tang,W. H., Martin, L. A., Hwa, J., Cell. Mol. Life Sci. 2012, 69, 3361–3380.

[14] Ferreiro-Vera, C., Mata-Granados, J. M., Priego-Capote, F., Quesada-Gómez, J. M., Luque de Castro, M. D., Anal. Bioanal. Chem. 2011, 399, 1093–1103.

[15] Ferreiro-Vera, C., Mata-Granados, J. M., Priego-Capote, F., Quesada-Gómez, J. M., Luque de Castro, M. D., J. Chromatogr. A. 2011, 1218, 2848–2855.

[16] Strassburg K., Huijbrechts A. M., Kortekaas K. A., Linde- man J. H., Pedersen T. L., Dane A., Berger R., Brenkman A., Hankemeier T., van Duynhoven J., Kalkhoven E., NewmanJ.W., Vreeken R. J., Anal. Bioanal. Chem. 2012, 404, 1413–1426.

[17] Yang J., Dong H., Hammock B.D., Curr Opin Lipidol. 2011, 22, 197–203.

[18] Fawcett, T., Pattern Recogn. Lett. 2006, 27, 861–874.

[19] Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., Müller, M., BMC Bioinformat. 2011, 12, 77–84.

[20] Carpenter, J., Bithell, J., Stat.Med. 2000, 19, 1141–1164.

[21] Gundu, H. R. R., Sampath, P., Prostag. Leukotr. Ess. 1996, 54, 155–166.

[22] Reilly, M. P., Praticò, D., Delanty, N., DiMinno, G., Tremoli, E., Rader, D., Kappor, S., Rokach, J., Lawson, J., FitzGer-Ald, G. A., Circulation 1998, 98, 2822– 2828.

[23] Weintraub, W. S., Klein, L. W., Seelaus, P. A., Agar- wal, J. B., Helfant, R. H., Am. J. Cardiol. 1985, 55, 669–672.

[24] Stewart, S. T., Cutler, D. M., Rosen, A. B., N. Engl. J.Med. 2009, 361, 2252–2260.

[25] Bazzano, L. A., He, J., Muntner, P., Vupputuri, S., Whelton, P. K., Ann. Intern. Med. 2003, 138, 891–897.

[26] Andreas, E., Elke, K., Serge, P.V. D., Helmut, S., Robert, B., Arch. Oral Biol. 2012, 57, 1133–1140.

[27] Jia, Z., Zhang, A., Zhang, H., Dong, Z., Yang, T., Circ. Res. 2006, 99, 1243–1251.

[28] Sokolowska, B., Dropinski, J., Rzeszutko, M., Szczeklik, W., Sanak, M., Szczeklik, A., Int. J. Cardiol. 2010, 145, 625–626.

[29] Mozes, T., Zijlstra, F. J., Heiligers, J. P., Saxena, P. R., Bonta, I. L. Prostag. Leukotr. Ess. 1991, 42, 209–216.

[30] Gesquiere, L., Loreau, N., Blache, D., Free Radical Biol. Med. 2000, 29, 181–190.

[31] Makheja, A. N., Mol. Cell. Biochem. 1992, 111, 137–142.

[32] O'Brien, J. J., Ray, D. M., Spinelli, S. L., Blumberg, N., Taubman, M. B., Francis, C.W., Wittlin, S. D., Phipps, R. P., Prostag. Oth. Lipid M. 2007, 82, 68–76.

[33] Seet, R. C. S., Lee, C. Y. J., Loke, W. M., Huang, S. H., Huang, H., Looi, W. F., Chew, E. S., Quek, A. M. L., Lim, E. C. H., Halliwell, B., Free Radical Biol. Med. 2011, 50, 1787–1793.

[34] Teixeira, P., Sardinha, L., Going, S. B., Lohman, T. G., Obesity Res. 2001, 9, 432–442.

[35] Haffner, S.M., Obesity 2006, 14, 121–127.

[36] Ding, X., Murray, P. A., Am. J. Physiol. Lung Cell. Mol. Physiol. 2005, 289, 825–833.

[37] Kobayash, T., Tahara, Y., Matsumoto, M., Iguchi, M., Sano, H., Murayama, T., Arai, H., Oida, H., Yurugi- Kobayashi, T., Yamashita, J. K., Katagiri, H., Majima, M., Yokode, M., Kita, T., Narumiya, S., J. Clin. Invest. 2004, 114, 784–794.

[38] Nie, D., Lamberti, M., Zacharek, A., Li, L., Szekeres, K., Tang, K., Chen, Y., Honn, K. V., Biochem. Biophys. Res. Commun. 2000, 267, 245–251.

[39] Cheng, Y., Austin, S. C., Rocca, B., Holler, B. H., Coffman, T. M., Grosser, T., Lawson, J. A., FitzGerald, G. A., Science 2002, 296, 539–541.

[40] Ali, S., Davis, M. G., Becker, M. W., Dorn, G. W., J. Biol. Chem. 1993, 268, 17397–17403.

[41] Richard, J. R., Physiol. Rev. 2002, 82, 131–185.

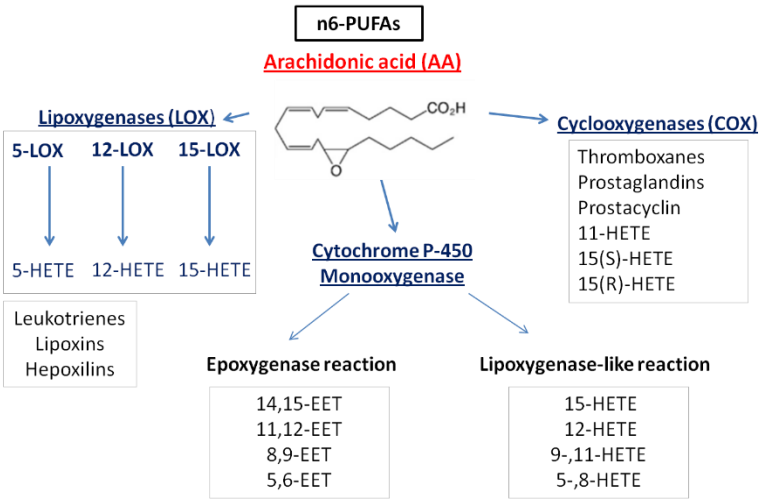
[42] González-Núñez, D., Claria, J., Rivera, F., Poch, E., Hypertension 2001, 37, 334–338.

[43] Morykwas, M. J., Duell, E. A., Jennings, D. A., Ledbetter, M. S., White, W. L., Argenta, L. C., Inflammation 1992, 16, 251–258.

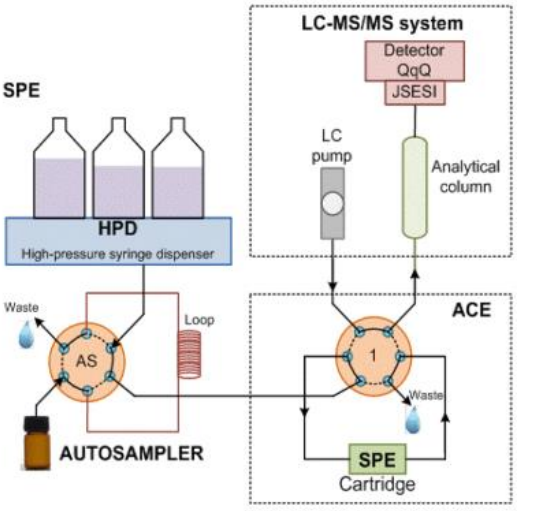
[44] Capra, V., Back, M., Barbieri, S. S., Camera, M., Tremoli, E., Rovati, G. E.,Med. Res. Rev. 2012, 23, 364–438.

[45] Stranjalis, G., Korfias, S., Psachoulia, C., Kouyialis, A., Sakas, D. E., Mendelow, A. D., Acta Neurochir. 2007, 149, 231–238.

Supplementary material



Supplementary Fig. 1. Scheme of the arachidonic acid pathway.



Supplementary Fig. 2. Scheme of the device based on an SPE workstation coupled to an LC–MS/MS system. Automatic Cartridge Exchange (ACE), Electro Spray Ionization module (JSESI), High Pressure Dispenser (HPD), Triple Quadrupole Mass Detector (QqQ).

CHAPTER 5

LC–MS/MS quantitative analysis of paclitaxel and its major metabolites in serum, plasma and tissue from women with ovarian cancer after intraperitoneal chemotherapy

Journal of Pharmaceutical and Biomedical Analysis 91 (2014) 131–137

Journal of Pharmaceutical and Biomedical Analysis

91 (2014) 131–137

LC–MS/MS quantitative analysis of paclitaxel and its major metabolites in serum, plasma and tissue from women with ovarian cancer after intraperitoneal chemotherapy

M.A. Fernández-Peralbo [a,b], F. Priego-Capote [a,b], M.D. Luque de Castro[a,b], A. Casado-Adam [b], A. Arjona-Sánchez [b], F.C. Muñoz-Casares [b]

[a]*Department of Analytical Chemistry, Campus of Rabanales, University of Córdoba, Córdoba, Spain*

[b]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, Córdoba, Spain*

LC–MS/MS quantitative analysis of paclitaxel and its major metabolites in serum, plasma and tissue from women with ovarian cancer after intraperitoneal chemotherapy

M.A. Fernández-Peralbo, F. Priego-Capote, M.D. Luque de Castro, A. Casado-Adam, A. Arjona-Sánchez, F.C. Muñoz-Casares

Abstract

A method for determination of the antineoplastic drug paclitaxel and its main metabolites (*viz.,* 6α-hydroxypaclitaxel and p-3´-hydroxypaclitaxel) at the sub-ng/mL level is here presented. Sample preparation consisted of a liquid–liquid extraction step for cleanup and preconcentration of the target analytes prior to chromatographic analysis by tandem mass spectrometry detection (LC–ESI-MS/MS). The determination step was optimized by selected reaction monitoring (SRM) mode for highly selective identification and sensitive quantitation of paclitaxel and its metabolites in human serum, plasma and tissue. The detection limits were in the range 0.03–0.15 ng/mL for serum and 0.07–0.62 ng/g for tissue, with intra-day variability range from 0.5 to 2.7%, expressed as relative standard deviation. The method was applied to determine paclitaxel and its metabolites in serum and tissue from 13 women suffering from ovarian peritoneal carcinomatosis, after hyperthermic intraperitoneal intraoperative chemotherapy (HIPEC) treatment. The method reported here can be considered a suited tool to monitor the concentration of this drug in patients subjected to HIPEC as strategy to evaluate the toxicity and efficiency of this treatment.

Keywords: Paclitaxel; HIPEC; LC–MS/MS; Metabolomics; Ovarian cancer

1. Introduction

Paclitaxel is an antineoplastic drug used for chemotherapy treatment of ovarian, breast and lung cancer [1]. It was discovered in 1968 by isolation from the bark of *Taxus brevifolia*, a species of yew tree originally from Pacific Northwest in North America [2]. However, the clinical use of paclitaxel is restricted because yew trees grow extremely slowly and the content of paclitaxel is too low for large-scale production. Currently, paclitaxel precursors as 10-deacetylbaccatin III [3], which comes from the leaves of different species of yew, are used to obtain sufficient amount of paclitaxel for clinical use by semi-synthetic methods.

Microtubules, the major structural components in cells, are the target of a large and diverse group of natural anticancer drugs [4]. In the case of paclitaxel, it binds to the β subunit of tubulin and promotes microtubule polymerization and stabilization as a result. Thus, paclitaxel blocks mitosis and induces apoptosis of malignant tumor cells [5,6].

Often, chemotherapy is a systemic treatment that involves intravenous or oral administration of drugs. Intraperitoneal chemotherapy (IP) is applied to certain ovarian cancers where the chemotherapy drug is injected through a catheter directly into the abdominal cavity and thus, a high dose of drug can be administered [7]. In this way, toxicity can be reduced as systemic administration of lower concentrations of drug is possible [8].

Paclitaxel has been successfully used in the last few years for treatment of advanced ovarian cancer by administration in hyperthermic intraperitoneal intraoperative chemotherapy (HIPEC) [9]. Presently, radical cytoreductive surgery by peritonectomy procedures and HIPEC is the best option for treatment of ovarian peritoneal carcinomatosis, not only to eliminate the macroscopic tumoral disease, but also the residual microscopic aspect as the cause of frequent recurrences in these patients. In this context, paclitaxel has shown excellent pharmacokinetic behavior for intraperitoneal administration, in addition to a salient activity against ovarian carcinoma and acceptable tolerance; characteristics that have promoted the use of this antineoplastic drug [10,11].

Paclitaxel is metabolized in the human liver by cytochrome P450 (CYP) enzymes [12]. CYP2C8 metabolizes paclitaxel to 6α-hydroxypaclitaxel (6α-OHP) and CYP3A4 to p-3´-hydroxypaclitaxel (C3´-OHP). Because of the multiple side effects produced by paclitaxel, quantitation of this drug together with its main metabolites is of major importance since information on the pharmacokinetics of paclitaxel is thus obtained. In this way both the doses and the exposure time can be optimized to consolidate therapy, prolong survival and hinder disease progression.

Different analytical techniques have been used for determination of paclitaxel in biological samples, including capillary electrophoresis–UV [13], immunoassay [14], LC–UV [15], LC–MS/MS [16–18] and LC–SSI–ITMS (liquid chromatography–ion trap mass spectrometry with a sonic spray ionization interface) [19] or column-switching LC–MS/MS has also been used for the determination of paclitaxel and its metabolites in human plasma without any pretreatment [20]. The most common sample preparation strategies for extraction and preconcentration of drugs and their metabolites from biological samples are based on protein precipitation, liquid–liquid extraction (LLE) and off-line or on-line solid-phase extraction (SPE) [21–23].

An LC–MS/MS platform in selected reaction monitoring (SRM) mode was chosen to support the present clinical pharmacological study on the simultaneous quantitation of paclitaxel, 6α-OHP and C3´-OHP after a simple sample preparation step based on LLE. The platform allows taking benefits from the combination of high selectivity, sensitivity and reproducibility. The final aim was to provide a method to determine paclitaxel and its metabolites in different biological samples (serum, plasma and tissue) to monitor the evolution of the drug in post-chemotherapy application.

2. Experimental

2.1. Study design, sampling and storage

The selected cohort was part of a prospective clinical study performed in 13 patients diagnosed with peritoneal carcinomatosis from epithelial ovarian cancer who underwent radical surgery–peritonectomy, achieving an optimal R0–R1 cytoreduction—microscopic tumor residues (R0) or macroscopic tumor residues <1 cm (R1)—followed by HIPEC. All steps involving selection of patients, sampling and storage were developed according to the guidelines dictated by the World Medical Association Declaration of Helsinki in 2004. The ethical review board of Reina Sofía Hospital (Córdoba, Spain) approved and supervised the clinical study.

2.1.1. Inclusion criteria

The inclusion criteria were as follows: (1) age ranging between 18 and 75 years; (2) histopathologic confirmation of peritoneal carcinomatosis from epithelial ovarian cancer (stage IIIc-FIGO); (3) Karnosfsky index >70 or performance status ≤2; and (4) informed consent form filled by all patients.

2.1.2. Exclusion criteria

The exclusion criteria were as follows: (1) unfulfillment of inclusion criteria; (2) extra-abdominal metastasis or stage IV FIGO (liver, lung, bone, etc.); (3) concomitance of other malignant neoplasm; (4) renal, hepatic or cardiovascular dysfunction; (5) intolerance during treatment; and (6) refusal to participate in the study.

2.1.3. Sampling and storage

All patients suffering peritoneal carcinomatosis from epithelial ovarian cancer and subjected to radical surgery–peritonectomy were then intraperitoneal administrated 60 mg/m^{-2} paclitaxel per 2 L of 1.5% dextrose at continuous hyperthermic perfusion (41–42 °C). A drastic cytoreduction was observed in all cases, leaving a small and well-perfused area of infiltrated peritoneum (residual tumor ≤1 cm), from which tissue biopsies were taken before and after intraperitoneal chemotherapy for 1 h. Approximately 0.7 g of tissue was introduced in Eppendorf tubes and preserved at −80 °C until analysis.

Blood samples were obtained before, immediately after and one h after instillation of intraperitoneal chemotherapy. In all instances 6 mL of blood was introduced in two types of tubes, one containing EDTA K3 as anticoagulant (for plasma) and the other without it (for serum). The tubes were centrifuged at 1 200 × *g* for 12 min and the serum or plasma thus obtained was stored at −80 °C until analysis.

2.2. Materials and standards

Paclitaxel, 6α-OHP and docetaxel as internal standard (IS) were from Cayman Chemical Company (Ann Arbor, MI, USA). C3´-OHP was from Sigma–Aldrich (Schnelldorf, Germany). All stock and working solutions were prepared in methanol and stored at −20 °C. The primary stock solutions were prepared at concentrations 1 μg/mL for 6α-OHP and C3´-OHP and 2.5 μg/mL for paclitaxel and docetaxel. Working solutions were prepared by appropriate dilution with methanol of the corresponding stock solution. LC–MS grade acetonitrile, MS grade formic acid (Scharlab, Barcelona, Spain) and deionized water (18 mΩ·cm) from a Millipore Milli-Q water purification system (Millipore, Bedford, MA, USA) were used for sample preparation and to prepare the chromatographic phases. Methyl *tert*-butyl ether (MTBE) from Scharlab was used as extractant.

2.3. Instrumentation

Chromatographic separation was performed by an Agilent (Palo Alto, CA, USA) 1200 Series LC system. The LC chromatograph was interfaced to an Agilent 6460 Triple Quad LC/MS detector from Agilent Technologies (Palo Alto, USA) equipped with a Jet Stream Technology electrospray ion source, also from Agilent. Nitrogen was supplied by a high purity generator from CLAN Tecnológica (Sevilla, Spain) and used as the source gas. Nitrogen (99.999%) from Carburos Metálicos (Córdoba, Spain) was used as collision gas.

2.4. Sample preparation

The three types of samples (serum, plasma and tissue) were thawed and vortexed for 30 s. For preparation of the peritoneal tissue homogenate, tissue from

tumor was weighed, gradually added to 1.5 mL 0.9% sodium chloride solution, and homogenized using a high-speed homogenizer (Shanghai Kingdom Biochemical Instrument, China). Sample volumes of 200 μL were mixed with 15 μL of IS (final concentration 150 ng/mL) in an Eppendorf tube. Paclitaxel, metabolites, and IS were extracted from the samples with MTBE. Following addition of 500 μL of MTBE, each tube was vortexed for 1 min, then centrifuged at 2 000 × *g* for 1 min. The liquid phase was transferred to other tube and dried in an Eppendorf vacuum concentrator. The dried residue was reconstituted with 50 μL of methanol. After centrifugation at 2 000 × *g* for 1 min, the clean liquid phase was transferred to the autosampler in a glass vial with insert, from which 5 μL were injected for analysis. Three independent replicates of each sample were analyzed.

2.5. LC–MS/MS conditions

A Mediterranea Sea C18 analytical column (3 μm, 50 mm × 4.6 mm, Teknokroma, Barcelona, Spain) thermostated at 25 °C was used for separation. 5 μL of sample was injected in each chromatographic run. Mobile phases, delivered at 0.6 mL/min, consisted of 0.1% formic acid either in water (A) or in acetonitrile (B). The chromatographic gradient was as follows: 4% B 0.5 min followed by a gradient to 100% B in 7 min and an isocratic step at 100% B for 5 min. A post-run of 5 min was included to equilibrate the column.

Mass spectral data were obtained in positive electrospray mode (ESI+) in SRM mode. Operating conditions as regards quantitation of SRM transitions were optimized as discussed below. Two ionization agents (acetic acid and formic acid), and two volatile salts (ammonium acetate and ammonium formate) at different concentrations were tested to compare the influence of pH on ionization. Formic acid led to the highest ionization efficiency by comparing signal-to-noise ratios and, therefore, to the best sensitivity by setting at the concentration 0.1% (v/v).

2.6. Preparation of calibration curves

Two pools of human serum and tissue were prepared from patients sampled after surgery peritonectomy and prior to intraperitoneal chemotherapy. The absence of paclitaxel or metabolites in these two pools of blank biofluids was

proved by application of the complete protocol. Calibration curves were prepared for each biofluid by spiking aliquots of both pools in a concentration range from 0.125 to 100 ng/mL in serum and from 0.21 to 450 or 750 ng/mL in peritoneal tissue (5 levels, n=3). Docetaxel (at 150 ng/mL) was used as IS to obtain the calibration model for paclitaxel, 6α-OHP and C3´-OHP. The calibration plots of paclitaxel and its metabolites were used for their quantitation. The calibration curve run with serum was also used to quantitate the target analytes in plasma. Spiked matrices, which were analyzed using the proposed method, were also used for recovery, precision, and accuracy studies.

3. Results and discussion

The study started with optimization of the MS detection step and appropriate chromatographic separation of the target analytes using standards in all instances. Then, the sample preparation step was optimized using the pools of the target samples (serum and tissue) taken prior to intraperitoneal chemotherapy. Then, the method was characterized by setting analytical features and, finally, validated by application to the given samples.

3.1. ESI-MS/MS detection

MS parameters such as voltage for efficient filtration of precursor ions in the first quadrupole, collision energy to fragment precursor ions by collision-induced dissociation (CID), and product ions selected for quantitation of each analyte were optimized. In the same way, the conditions of the electrospray ionization source were also optimized. All optimization parameters are listed in Table 1. The most abundant adduct ion for paclitaxel, 6α-OHP, C3´-OHP and docetaxel was $[M+Na]^+$, while the major product ions for all them come from the rupture of ester bonds, as shows Fig. 1. Quantitation was performed using SRM of the transitions *m/z* 876.0 → *m/z* 308.0 for paclitaxel, *m/z* 892.0 → *m/z* 607.0 for 6α-OHP, *m/z* 892.0 → *m/z* 324.0 for C3´-OHP, and *m/z* 830.0 → *m/z* 304.0 for docetaxel.

Table 1. Optimal values for ESI-MS/MS detection.

A) MS/MS conditions					**B) ESI conditions**		
Analyte	**t_R** *(min)*	**Voltage MS1** *(V)*	**Collision energy** *(eV)*	**Quantitation transition** *Precursor ion -> Product ion (m/z)*	**Variable**	**Tested range**	**Optimal value**
C3´-OHP	7.1	170	20	892.0 → 324.0	Gas temperature (ºC)	200-350	350
					Gas flow (ml/min)	5-13	11
6α-OHP	7.3	170	20	892.0 → 607.0	Nebulizer pressure (psi)	30-45	45
					Sheath gas temp (ºC)	100-400	400
Docetaxel	7.4	250	20	830 .0→ 304.0	Sheath gas flow (ml/min)	8-12	12
					Capillary voltage (V)	2500-4000	4000
Paclitaxel	7.7	170	30	876.0 → 308.0	Nozzle voltage (V)	300-2000	500

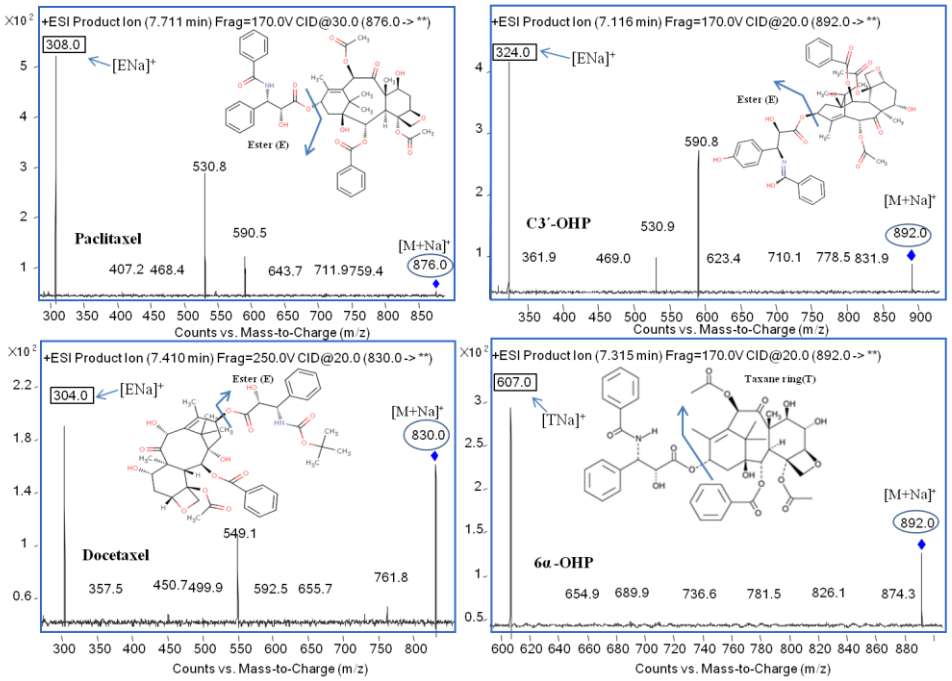


Fig. 1. MS/MS spectra of precursor ions from paclitaxel, 6α-OHP, C3´-OHP and docetaxel in positive electrospray ionization mode.

3.2. Sample preparation

Preparation of samples is a critical step for accurate and reliable LC–MS/MS assays. Initially, removal of proteins from serum was tested using methanol as treatment prior to LC–MS/MS analysis, which resulted in a lack of sensitivity as compared to standard solutions. For this reason, LLE was then chosen for serum/plasma and tissue because this technique not only removes potential interferents but also allows concentrating the target analytes by evaporation of the solvent. MTBE was adopted as extractant because of its high extraction efficiency and low interference, as found in the literature [24].

3.3. Linearity of calibration curves and lower limits of quantitation

The calibration curves were established by plotting peak area ratios of paclitaxel and its metabolites to IS *versus* nominal concentration. A least-squares

linear regression weighted by the reciprocal of the concentration was applied to generate a calibration curve with correlation coefficients higher than 0.99 (Table 2). The sensitivity of the method was evaluated by estimation of the detection and quantitation limits both with serum and tissue samples spiked at low concentration levels. This analytical feature was assessed by estimation of the limits of detection (LOD) and quantitation (LOQ) for each analyte by injecting dilution series of each in serum and tissue to obtain the concentrations which provided signals three and ten times the background noise, respectively, as shows Table 2. Both the detection and quantitation limits were lower in serum than in tissue for all analytes. 6α-OHP showed higher LODs and LOQs in serum and tissue than paclitaxel or C3´-OHP. As can be appreciated in Fig. 2, SRM chromatograms associated to serum pool spiked with the target analytes at 1 ng/mL are differentiated from the background signal ascribed to the analysis of non-spiked serum.

Table 2. Characteristics of the method.

Sample	Analyte	Linear range *(ng/ml)*	Coefficient of regression (R^2)	Limit of detection *(ng/ml)*	Limit of quantitation *(ng/ml)*
Serum	Paclitaxel	0.125-100	0.9952	0.03	0.125
	6α-OHP	0.5-100	0.9980	0.15	0.5
	C3´-OHP	0.125-100	0.9946	0.03	0.125
Tissue	Paclitaxel	0.21-750	0.9970	*0.07	*0.26
	6α-OHP	0.5-450	0.9975	*0.18	*0.62
	C3´-OHP	0.21-450	0.9981	*0.07	*0.26

*ng/g

3.4. Precision and accuracy

The precision and accuracy of the method was evaluated by an experimental strategy intended to estimate intra-day variability (by triplicate analyses within the

same day) and inter-days variability (by triplicate analyses for two consecutive days) of serum and tissue aliquots spiked at two concentrations (10 and 20 ng/mL in serum, 40 and 150 ng/mL in tissue). The precision was expressed as percentage

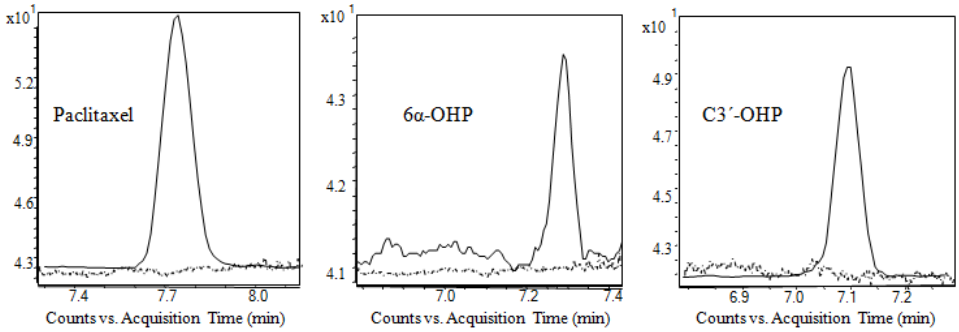


Fig. 2. Chromatograms from serum from patients prior to HIPEC treatment. Continuous line: serum spiked with the target analytes spiked at 1 ng/mL. Dotted line: non-spiked serum.

of relative standard deviation (%RSD) as shown in Table 3. Accuracy was also assessed for each spiked concentration (10 and 20 ng/mL in serum, 40 and 150 ng/mL in tissue) by calculating the percentage deviation from the theoretical added concentration. The results of accuracy in both types of samples are listed in Table 3. The precision and accuracy of the assay for tissue and serum were within acceptable limits. RSD values for intra-day and inter-days variability were below 3.2 and 12.0%, respectively, while accuracy values were higher than 85%, which supports the quantitative capability of the method.

3.5. Application of the method

The method was applied to quantitation of paclitaxel and its major metabolites in serum and tissue of 13 women diagnosed with ovarian peritoneal carcinomatosis and plasma of one of them. As mentioned above, serum and plasma samples were obtained before, immediately after and one hour after being treated with HIPEC while biopsy tissue was sample before and 1 h after HIPEC. Three biological replicates of each sample were analyzed.

Paclitaxel and its metabolites were not detected in the samples extracted prior to chemotherapy treatment. As Table 4 shows, paclitaxel metabolites were neither

Table 3. Intra-day and inter-days variability and accuracy of the method in peritoneal tissue homogenate and serum spiked with the target analytes.

Sample	Analyte	Spiked concentration *(ng/mL)*	Intraday (n=3)		Inter-days (n= 2 days)	
			Precision *(RSD%)*	**Accuracy** *(%±SD)*	**Precision** *(RSD%)*	**Accuracy** *(%±SD)*
Serum	Paclitaxel	10	2.2	98.2 ± 2.1	5.1	94.5 ± 4.8
		20	2.4	105.9 ± 2.5	2.8	107.3 ± 3
	6α-OHP	10	2.5	85.3 ± 1.7	3.2	88.6 ± 2.1
		20	0.7	90.9 ± 0.5	12.0	98.1 ± 7.7
	3C'-OHP	10	3.2	93.6 ± 2.4	5.2	88.8 ± 4.2
		20	1.6	104.8 ± 1.4	5.4	91.0 ± 4.4
Tissue	Paclitaxel	40	1.2	107.2 ± 1.4	1.9	108.7 ± 2.1
		150	1.3	86.9 ± 1.1	2.3	85.3 ± 2.0
	6α-OHP	40	2.5	104.8 ± 2.7	2.6	106.0 ± 2.6
		150	1.4	86.07 ± 1.2	1.6	85.5 ± 1.4
	3C'-OHP	40	1.8	115.1 ± 2.0	2.6	116.6 ± 3.1
		150	0.7	98.6 ± 0.7	1.1	98.1 ± 1.1

Table 4. Levels of paclitaxel and metabolites found in samples from the cohort under study expressed as ng/mL in serum and ng/g in tissue.

Patient	SERUM						TISSUE		
	POSTCHEMOTERAPHY *(immediately after HIPEC)*			POSTCHEMOTERAPHY *(1 h after HIPEC)*			POSTCHEMOTERAPHY *(immediately after HIPEC)*		
	Paclitaxel	**6α-OHP**	**C3´-OHP**	**Paclitaxel**	**6α-OHP**	**C3´-OHP**	**Paclitaxel**	**6α-OHP**	**C3´-OHP**
1	9.44 ± 1.25	0.65 ± 0.17	0.74 ± 0.05	6.19 ± 0.61	0.81 ± 0.31	0.92 ± 0.02	1898.0 ± 83.7	<LOD	<LOD
2	9.00 ± 0.09	2.09 ± 0.95	0.92 ± 0.21	10.13 ± 0.86	5.65 ±0.16	0.87 ± 0.17	836.3 ± 14.9	<LOD	<LOD
3	6.20 ± 0.15	3.52 ± 0.48	<LOQ	13.67 ± 0.59	2.34 ± 0.55	0.98 ± 0.24	2551.3 ± 139.4	<LOD	<LOD
4	9.99 ± 1.19	0.91 ± 0.09	0.68 ± 0.04	5.44 ± 0.21	0.80 ± 0.18	0.68 ± 0.03	471.2 ± 20.1	<LOD	<LOD
5	22.50 ± 0.37	1.00 ± 0.10	0.94 ± 0.06	19.62 ± 1.11	1.57 ± 0.33	1.37 ± 0.01	158.6 ± 13.4	<LOD	<LOD
6	9.36 ± 0.21	0.51 ± 0.08	0.68 ± 0.01	7.76 ± 0.29	0.86 ± 0.11	0.78 ± 0.01	246.5 ± 19.7	<LOD	<LOD
7	14.79 ± 0.38	0.81 ± 0.04	0.67 ± 0.01	12.97 ± 0.86	1.25 ± 0.19	0.81 ± 0.02	731.3 ± 37.1	<LOD	<LOD
8	14.38 ± 0.47	3.15 ± 0.28	1.07 ± 0.08	13.73 ± 0.59	7.75 ± 0.19	1.94 ± 0.11	382.1 ± 41.0	<LOD	<LOD
9	19.00 ± 2.88	1.23 ± 0.41	0.64 ± 0.01	23.86 ± 0.16	2.18 ± 0.22	0.82 ± 0.06	1152.1 ± 50.5	<LOD	<LOD
10	14.23 ± 0.51	<LOD	0.92 ± 0.07	13.59 ± 1.91	<LOD	0.99 ± 0.09	513.7 ± 33.7	<LOD	<LOD
11	16.73 ± 0.41	0.63 ± 0.02	1.09 ± 0.06	14.85 ± 0.78	1.32 ± 0.32	1.47 ± 0.04	254.4 ± 14.2	<LOD	<LOD
12	22.92 ± 1.87	1.45 ± 0.31	0.70 ± 0.02	15.82 ± 1.32	2.35 ± 0.42	0.87 ± 0.02	582.7 ± 84.6	<LOD	<LOD
13	11.24 ± 0.24	0.80 ± 0.03	0.69 ± 0.02	6.33 ± 0.32	<LOQ	0.93 ± 0.04	166.5 ± 22.1	<LOD	<LOD
13 (plasma)	12.05 ± 0.77	0.58 ± 0.11	0.74 ± 0.02	6.69 ± 0.63	0.67 ± 0.09	0.27 ± 0.04			

<LOD = values below the detection limit; <LOQ = values below the quantitation limit

detected in tissue after chemotherapy, an expected fact since paclitaxel hardly metabolized in peritoneal tissue [25,26]. However, the concentration of paclitaxel in tissue was, in all instances, significantly higher than in serum. In most cases, the concentration of paclitaxel in serum decreased one hour after HIPEC as compared to levels found immediately after intraperitoneal chemotherapy. It is worth noting that the concentration of 6α-OHP was slightly higher than that of C3´-OHP in most cases; therefore, paclitaxel seems to be preferentially metabolized to 6α -OHP. The concentration of both metabolites increased after 1 h; however, the increase of C3´-OHP was less significant than that observed for 6α-OHP.

The potential suitability of the method for determination of the target analytes in plasma as biological sample was evaluated using this sample from patient number 13. Fig. 3 shows the chromatograms obtained from serum, plasma and tissue from this patient. The differences in the concentration of paclitaxel in both blood samples and at the two sampling times were within the experimental error of the method; therefore, the method could also be appropriate for determination of this drug in plasma.

4. Conclusions

The proposed method is sensitive and selective for the simultaneous determination of paclitaxel and its metabolites in serum, plasma or peritoneal tissue. The simple LLE procedure and short LC–MS/MS run time make the method appropriate for routine analysis of large number of samples as a tool for monitoring HIPEC treatments with paclitaxel. No significant interference was caused by endogenous compounds. Application of the method to quantitation of paclitaxel and its metabolites in serum and peritoneal tissue of 13 women suffering from ovarian peritoneal carcinomatosis, after HIPEC treatment, shows its potential to monitor the evolution of this drug. Therefore, future research on toxicity of paclitaxel and its metabolites, dosage to be applied and metabolism kinetics could be supported on the method here reported.

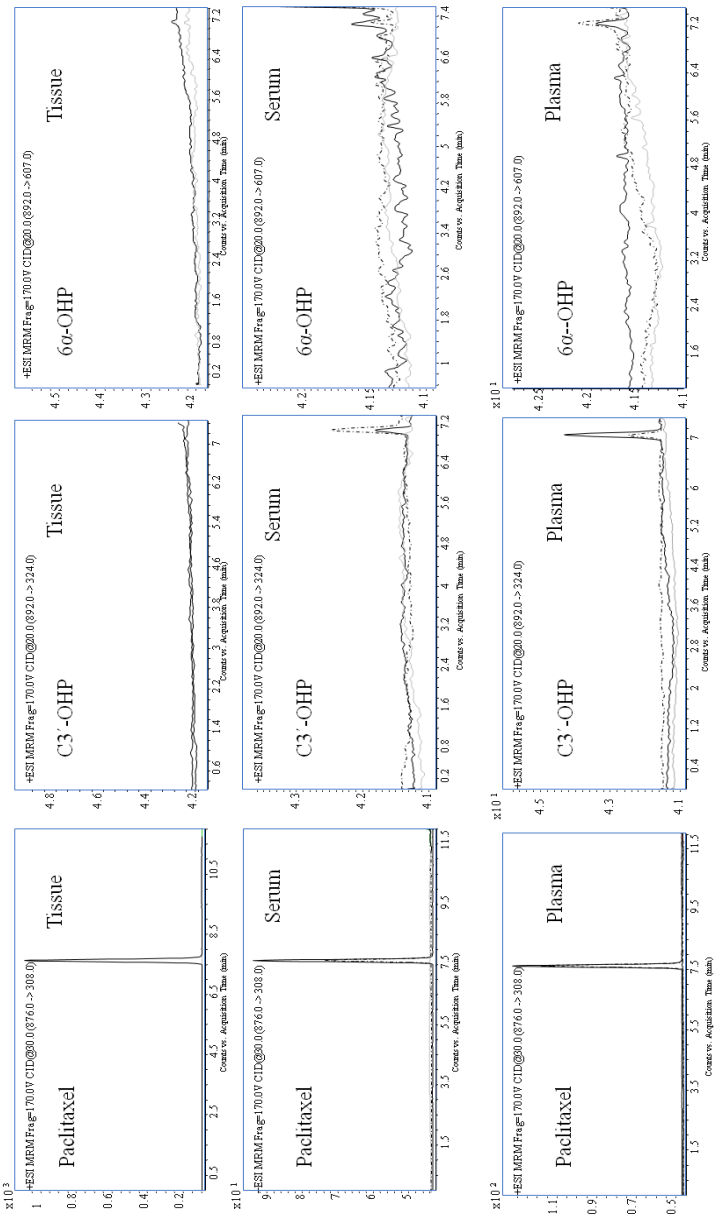


Fig. 3. Chromatograms from paclitaxel and its metabolites in serum, plasma and ovarian tissue from patient number 13 (gray line: prior to HIPEC, black line: immediately after HIPEC, dotted line: 1-h after HIPEC).

Acknowledgements

The authors thank the Spanish Ministerio de Economía y Competitividad (MEC) and FEDER programme for financial support through Project CTQ2012-37428. F.P.C. is also grateful to the Spanish Ministerio de Ciencia e Innovación (MICINN) for a Ramón y Cajal contract (RYC-2009-03921).

References

[1] M. Markman, Intraperitoneal antinelastic drug delivery: rationale and results, Lancet Oncol. 4 (2004) 277–283.

[2] E. de Bree, P.A. Theodoropoulos, H. Rosing, J. Michalakis, J. Romanos, J.H. Beijnen, D.D. Tsiftsis, Treatment of ovarian cancer using intraperitoneal chemotherapy with taxanes: from laboratory bench to bedside, Cancer Treat. Rev. 32 (6) (2006) 471–482.

[3] A. Zhiri, M. Jaziri, Y. Guo, R. Vanhaelen-Fastré, M. Vanhaelen, J. Homès, K. Yoshimatsu, K. Shimomura, Tissue cultures of Taxus baccata as a source of 10-deacetylbaccatin III, a precursor for the hemisynthesis of taxol, Biol. Chem. 376 (10) (1995) 583–586.

[4] M.A. Jordan, Mechanism of action of antitumor drugs that interact with microtubules and tubulin, Curr. Med. Chem. Anticancer Agents 2 (1) (2002) 1–17.

[5] T.H. Wang, H.S. Wang, Y.K. Soong, Paclitaxel-induced cell death: where the cell cycle and apoptosis come together, Cancer 88 (2000) 2619–2628.

[6] S.B. Horwitz, Mechanism of action of taxol, Trends Pharmacol. Sci. 13 (1992) 134–136.

[7] M. Markman, J.L. Walker, Intraperitoneal chemotherapy of ovarian cancer: a review, with a focus on practical aspects of treatment, J. Clin. Oncol. 24 (6) (2006) 988–994.

[8] S.B. Howell, in: D.S. Alberts, M.C. Clouser, L.M. Hess (Eds.), Selection of Drugs for Intraperitoneal Chemotherapy for Ovarian Cancer, Springer, Berlin, 2010, pp. 77–88.

[9] J.H. Bae, J.M. Lee, K.S. Ryu, Y.S. Lee, Y.G. Park, S.Y. Hur, W.S. Ahn, S.E. Namkoong, Treatment of ovarian cancer with paclitaxel- or carboplatin-based intraperitoneal hyperthermic chemotherapy during secondary surgery, Gynecol. Oncol. 106 (1) (2007) 193–200.

[10] E. de Bree, H. Rosing, D. Filis, J. Romanos, M. Melisssourgaki, M. Daskalakis, M. Pilatou, E. Sanidas, P. Taflampas, K. Kalbakis, J.H. Beijnen, D.D. Tsiftsis, Cytoreductive surgery and intraoperative hyperthermic intraperitoneal chemotherapy with paclitaxel: a clinical and pharmacokinetic study, Ann. Surg. Oncol. 15 (4) (2008) 1183–1192.

[11] F.C. Muñoz-Casares, S. Rufián, M.J. Rubio, C.J. Díaz, R. Díaz, A. Casado, A. Arjona, M.C. Muñoz-Villanueva, J. Muntané, The role of hyperthermic intraoperative intraperitoneal chemotherapy (HIPEC) in the treatment of peritoneal carcinomatosis in recurrent ovarian cancer, Clin. Transl. Oncol. 11 (11) (2009) 753–759.

[12] S.J. Gardiner, E.J. Begg, Pharmacogenetics, drug-metabolizing enzymes, and clinical practice, Pharmacol. Rev. 58 (3) (2006) 521–590.

[13] G. Hempel, D. Lehmkuhl, S. Krumpelmann, G. Blaschke, J. Boos, Determination of paclitaxel in biological fluids by micellar electrokinetic chromatography, J. Chromatogr. A 745 (1996) 173–179.

[14] J.G. Leu, B.X. Chen, P.B. Schiff, B.F. Erlanger, Characterization of polyclonal and monoclonal anti-taxol antibodies and measurement of taxol in serum, Cancer Res. 53 (1993) 1388–1391.

[15] X. Tong, J. Zhou, Y. Tan, Liquid chromatography/tandem triple–quadrupole mass spectrometry for determination of paclitaxel in rat tissues, Rapid Commun. Mass Spectrom. 20 (2006) 1905–1912.

[16] T. Nguyen, J. Eshraghi, G. Gonyea, R. Ream, R. Smith, Studies on factors influencing stability and recovery of paclitaxel from suspension media and cultures of Taxus cuspidata cv Densiformis by high-performance liquid chromatography, J. Chromatogr. A 911 (2001) 55–61.

[17] P. Li, B.J. Albrecht, X. Yan, M. Gao, H.R. Weng, M.G. Bartlett, A rapid analytical method for the quantification of paclitaxel in rat plasma and brain tissue by high-performance liquid chromatography and tandem mass spectrometry, Rapid Commun. Mass Spectrom. 27 (19) (2013) 2127–2134.

[18] W. Zhang, G.E. Dutschman, X. Li, Y.C. Cheng, Quantitation of paclitaxel and its two major metabolites using a liquid chromatography–electrospray ionization tandem mass spectrometry, J. Chromatogr. B 879 (22) (2011) 2018–2022.

[19] H. Gréen, K. Vretenbrant, B. Norlander, C. Peterson, Measurement of paclitaxel and its metabolites in human plasma using liquid chromatography/ion trap mass spectrometry with a sonic spray ionization interface, Rapid Commun. Mass Spectrom. 20 (14) (2006) 2183–2189.

[20] H. Yamaguchi, A. Fujikawa, H. Ito, N. Tanaka, A. Furugen, K. Miyamori, N. Takahashi, J. Ogura, M. Kobayashi, T. Yamada, N. Mano, K. Iseki, Quantitative determination of paclitaxel and its metabolites, 6α-hydroxypaclitaxel and p-3´-hydroxypaclitaxel, in human plasma using column-switching liquid chromatography/tandem mass spectrometry, Biomed. Chromatogr. 27 (4) (2013) 539–544.

[21] R. Dams, M.A. Huestis, W.E. Lambert, C.M. Murphy, Matrix effect in bio-analysis of illicit drugs with LC–MS/MS: influence of ionization type, sample preparation, and biofluid, J. Am. Soc. Mass Spectrom. 14 (11) (2003) 1290–1294.

[22] X.R. Naxing, L. Fan, M.J. Rieser, T.A. El-Shourbagy, Recent advances in high-throughput quantitative bioanalysis by LC–MS/MS, J. Pharm. Biomed. Anal. 44 (2) (2007) 342–355.

[23] L. Nováková, H. Vlcková, A review of current trends and advances in modern bioanalytical methods: chromatography and sample preparation, Anal. Chim. Acta 656 (2009) 8–35.

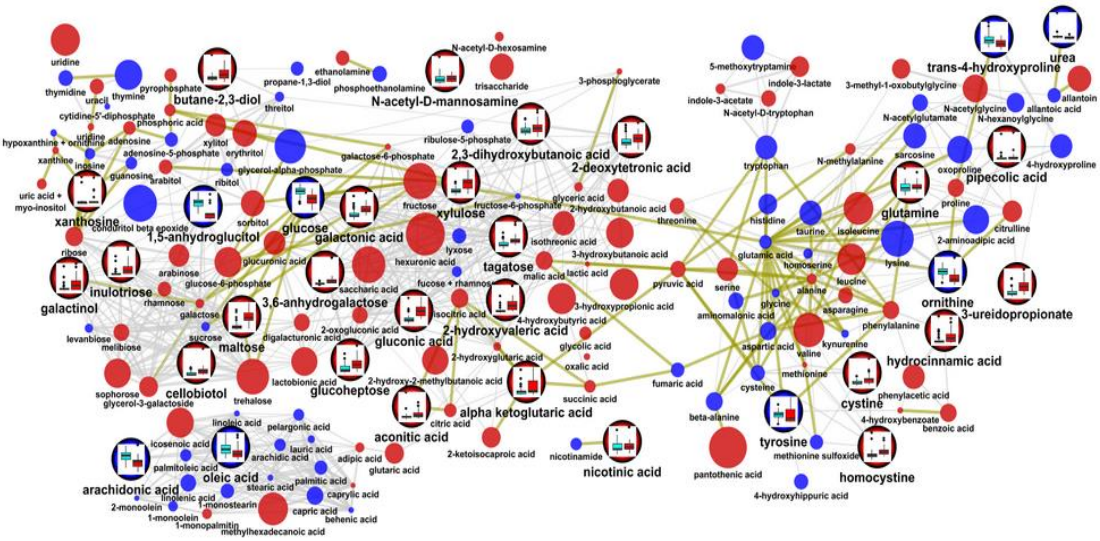
[24] G. Basileo, M. Breda, G. Fonte, R. Pisano, C.A. James, Quantitative determination of paclitaxel in human plasma using semi-automated liquid–liquid extraction in conjunction with liquid chromatography/tandem mass spectrometry, J. Pharm. Biomed. Anal. 32 (4–5) (2003) 591–600.

[25] F. Mohamed, P. Marchettini, O.A. Stuart, P.H. Sugarbaker, Pharmacokinetics and tissue distribution of intraperitoneal paclitaxel with different carrier solutions, Cancer Chemother. Pharmacol. 52 (2003) 405–410.

[26] A. Sparreboom, O. van Tellingen, W.J. Nooijen, J.H. Beijnen, Tissue distribution, metabolism and excretion of paclitaxel in mice, Anticancer Drugs 7 (1) (1996) 78–86.

SECTION III

Contributions to untargeted metabolomics analysis



Untargeted analysis is the strategy providing the maximum information level in metabolomics. It usually precedes to targeted metabolomics analysis by profiling to reveal the metabolites that must be monitored in the given samples and, thus, to know the state and/or evolution of the cohort under study. The location of untargeted metabolomics analysis in an overall research involving targeted and untargeted strategies is always controversial: the former is simpler because involves a smaller number of metabolites —therefore, it must be located before—, but the latter always precedes targeted analysis as the way to know the metabolites to be analyzed. The higher complexity of untargeted analysis predominated in this case and location of this strategy as the last was decided.

The research in Chapter 6 solves one of the problems caused by the versatility of LC–MS of allowing the two (positive and negative) ionization modes: the production of two independent data sets (one per polarity) that require two parallel data analyses. The problem was overcome by developing a tool to combine the data matrices for the positive and negative ionization modes by filtering off the molecular features of a given metabolite in whichever ionization mode they are detected with the less sensitivity. In this way redundant information is avoided and a single data set is obtained, as assessed by application to data sets obtained from the analysis of human serum and urine.

Despite untargeted metabolomics analysis usually requires a sample preparation less exhaustive than targeted analysis, this step can be complicated when involving exhaled breath condensate (EBC) because of the low concentration of its components, and the necessity for removing potential interferents. The exhaustive study that constitutes Chapter 7 shows lyophilization as the best, providing an 80-times preconcentration factor with respect to the original sample, and making possible the identification of 49 EBC metabolites. In addition to sample preparation optimization, the chapter discusses the nature of

the tentatively identified compounds among which amino acids, fatty acids, fatty amides, fatty aldehydes, sphingoid bases, oxoanionic compounds, imidazoles, hydroxy acids, and aliphatic acyclic acids are the most remarkable.

The search for clinical biomarkers for prostate cancer (PCa) diagnosis that replace or complement the PSA test is the subject of the last two chapters of this Thesis-Book: Chapters 8 and 9. The former is devoted to compare the biochemical pathways in PCa patients to those of negative biopsy control individuals through urine samples. A partial least squares discriminant analysis (PLS-DA) model, applied to 28 significant metabolites tentatively identified in urine, was characterized by 88.4 and 92.9% of sensitivity and specificity, respectively. The different concentration of the significant metabolites in PCa patients as compared with control individuals allows connection among the biochemical pathways in which they are involved, thus supporting, once again, the role of metabolomics in upstream processes.

Chapter 9, as the final chapter of the Thesis-Book, contains a more in-depth chemometric treatment of the data from the study in the previous chapter developed by Random Forest Analysis and ROC curves. Finally, using the information from selected metabolites as potential biomarkers and the data from the PSA analysis, 5-variables panels (5 metabolites or 4 metabolites+PSA) were selected for diagnostic of PCa. Maximizing sensitivity and specificity, a screening tool was characterized by high sensitivity and selective detection of true positive cases (sensitivity and specificity, 97.7% and 100%, respectively). Biochemical pathways of target metabolites also provide in this case upstream information.

CHAPTER 6

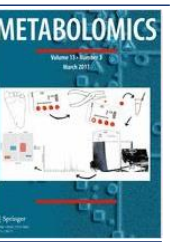
MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes

Metabolomics 12:43 (2016) 1–12

Springer

Metabolomics

12:43 (2016) 1–12



MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes

M. Calderón-Santiago[1,2], M.A. Fernández-Peralbo[1,2], F. Priego-Capote[1,2], M. D. Luque de Castro[1,2]

[1]*Department of Analytical Chemistry, Campus of Rabanales, University of Córdoba, Córdoba, Spain*

[2]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, Córdoba, Spain*

MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes

Mónica Calderón-Santiago, María A. Fernández-Peralbo, Feliciano Priego-Capote, María D. Luque de Castro

Abstract

The wide chemical diversity of metabolites present in biological samples hinders full characterization of their metabolomic profile by using a single analytical technique. However, LC–MS is a competitive platform for obtaining a comprehensive description of a wide range of metabolites by using a combination of the two ionization modes, positive and negative, the suitability of which for this purpose depends on the chemical structure of the particular metabolites. One of the main shortcomings of this approach is that it produces two independent data sets (one per polarity) that require two parallel data analyses to identify significant metabolites associated to the target biological effect. As a result, analyzing the data is a time-consuming step of metabolomic analysis. In this work, this shortcoming was overcome by developing a tool to combine the data matrices for the positive and negative ionization modes in MS. The most critical function of the tool is filtering off the molecular features of a given metabolite in whichever ionization mode they are detected with the less sensitivity in order to avoid using redundant information and complicating the statistical treatment as a result. The software allows the two data sets to be merged into a single one that can be statistically processed in the conventional metabolomics workflow. Its efficiency was assessed by application to data sets obtained from the analysis of human serum and urine.

Keywords Mass spectrometry; LC–MS; Metabolomics; Untargeted analysis; R package

1. Introduction

Metabolomics is the 'omics' concerned with the identification and/or quantitation of small molecules or metabolites (generally<1000 Da) present in biofluids, cells or organisms. There are three different ways of approaching metabolomics commonly known as "fingerprinting", "targeted profiling analysis" and "untargeted profiling analysis" (Álvarez-Sánchez *et al.* 2010). Untargeted analysis, also known as "profiling analysis", focuses on identifying the greatest possible number of metabolites. However, the wide chemical variety of metabolites comprising the metabolome, and their widely ranging concentrations preclude characterization of the whole metabolome with a single analytical tool. One of the most widely used analytical platforms for this purpose is liquid chromatography coupled with mass spectrometry (LC–MS), which affords unbiased detection of a large number of metabolites in biological samples (Vuckovic 2012; Xiao *et al.* 2012; Zhang *et al.* 2012; Zhou *et al.* 2012). One of the most salient features of this platform is the ionization mode; in fact, MS can operate in the positive and negative modes, the better choice in each case being dictated by the structure of the target metabolites. Thus, some metabolites containing easily protonated functional groups such as acylcarnitines with a trimethylamine group can only be detected in the positive ionization mode (Kivilompolo *et al.* 2013). On the other hand, compounds with easily deprotonated functional groups such as fatty acids are mainly detected in the negative ionization mode. In fact, the least abundant fatty acids are rarely or never detected in the positive ionization mode (Aslan *et al.* 2013; Sommer *et al.* 2006). For this reason, untargeted studies generally require duplicate analysis of all samples —one with each ionization mode— to increase metabolite coverage (Arbulu *et al.* 2015; Gika *et al.* 2014; Xiao *et al.* 2012). Duplicate analysis of the whole sample set produces two independent data sets that may contain information inputs corresponding to common metabolites (Chen *et al.* 2012; Fraser *et al.* 2014; Xiao *et al.* 2012). The fact that some metabolites including amino acids, short-chain carboxylic acids, purines and phosphatidyl-ethanolamines can be detected in both ionization modes precludes combining the two data sets into a single matrix because the resulting data set will contain

redundant information on those metabolites detected in both modes. As a result, analysis of the data frequently requires using independent workflows to process the two sets (positive and negative ionization) and an extended processing time (Chen *et al.* 2012; Evans *et al.* 2014; Fraser *et al.* 2014). Also, data can only be interpreted after significant metabolites detected in each ionization mode have been independently identified.

Recent advances in fast, robust switching polarity technology have enabled the simultaneous acquisition of positive and negative ionization data in a single analysis, thus shortening processing times. However, this approach has proved inefficient in untargeted analysis because it decreases resolution and sensitivity through a combination of cycles for MS scanning and MS/MS fragmentation in the two polarities. As a result, it has been exclusively used for targeted analysis of panels of metabolites (Gilbert-López *et al.* 2012; Herrera-Herrera *et al.* 2011). Identifying molecular features detected in the two ionization modes prior to statistical analysis would allow one to use a single data set consisting of known metabolites only (Calderón-Santiago *et al.* 2015; Evans *et al.* 2014). However, this is most often rather difficult owing to the high complexity of the process, which involves identifying hundreds or even thousands of molecular entities prior to analyzing the data. For this reason, the most frequent practice is to restrict identification to those statistically significant metabolites contributing to explaining the target variability sources. In this work, a tool for combining data obtained in the positive and negative ionization modes in LC–MS was developed. The package, named MSCombine, affords statistical analysis of a single data set, thereby avoiding the need for two parallel analyses —one per polarity. For this purpose, the open source data analysis software R is used to configure the algorithms in MSCombine. After all potential molecular features or entities have been extracted from raw data files obtained by analyzing samples in the two ionization modes, the resulting matrices are used to identify the molecular entities corresponding to the same tentative metabolite with provision for mass accuracy and retention time. Then, the molecular features associated with each metabolite are validated and the variables with the lowest sensitivity in the original data sets

removed. Finally, the two sets are merged into the final data set, which will contain the variables detected in both the positive and the negative ionization mode. The applicability of the proposed package was demonstrated by using it to process data sets obtained by analyzing two types of samples commonly encountered in metabolomics studies: urine and serum.

2. Materials and methods

2.1. Chemicals

Chromatographic mobile phase B was prepared in LC–MS optimum grade acetonitrile from Fisher Scientific (Madrid, Spain). MS grade formic acid from Scharlab (Barcelona, Spain) was used as ionization agent in LC–QTOF analyses, and deionized water (18 mΩ·cm) supplied by a Milli-Q water purification system from Millipore (Bedford, MA) was used to prepare the chromatographic aqueous phase and the sample solutions. Ammonium formate from Scharlab (Barcelona, Spain) and HPLC grade methanol were used to prepare urine and serum samples, respectively.

2.2. Instruments and apparatus

A Sorvall Legend Micro 21R centrifuge from Thermo Scientific (Waltham, MS) was used to centrifuge samples. Also, a 1200 Series LC system coupled to a 6540 UHD Accurate-Mass QTOF hybrid mass spectrometer equipped with a dual electrospray (ESI) source, all from Agilent (Santa Clara, CA), was used to obtain MS spectra. The chromatographic eluates were monitored by tandem mass spectrometry in the high resolution mode.

2.3. Sample collection and pretreatment

Venous blood was collected from three healthy donors in evacuated sterile serum tubes (Vacutainer, Becton–Dickinson, Franklin Lakes, NJ, USA) containing no additives and incubated at room temperature for 10 min to facilitate coagulation. Then, the tubes were centrifuged at 2000 × *g* at 4 °C for 15 min to isolate

the serum fraction (processing within 2 h after collection). Finally, a serum pool was prepared and split into 24 aliquots prior to storage at –80 °C until analysis.

Urine was collected from 112 donors; 60 were on ciprofloxacin and the other 52 were used as controls. Urine was collected after fasting for at least 10 h and aliquoted prior to storage at –80 °C until analysis. All individuals were males, the average age being 68 yr (51–94) for those on the drug and 66 years (51–89) for the controls.

All steps from serum and urine sampling to analysis were performed in compliance with the guidelines dictated by the World Medical Association Declaration of Helsinki of 2004 and supervised by specialist staff of the Maimonides Biomedical Research Institute (Córdoba, Spain). The study was approved by the Ethics Committee of the Reina Sofia University Hospital. Informed consent was previously obtained from all participants.

2.4. Sample preparation and untargeted analysis

A total of 112 urine samples and 24 serum samples were analyzed by LC–QTOF MS/MS in both ionization modes. Methodological details including sample preparation and chromatographic analysis methods are specified below.

2.4.1. Urine pretreatment

100 μL aliquots of urine were unfrozen slowly by immersion in an ice bath. Then, samples were shaken for 30 s and centrifuged at 4 °C at 13.800 × *g* for 5 min. Finally, supernatants were 1:3 (v/v) diluted with 5 mM ammonium formate at pH 6, and the vial placed in the chromatograph autosampler for subsequent analysis.

2.4.2. LC–QTOF MS/MS urine analysis

Chromatographic separation was performed by using a Mediterranean C18 reversed phase analytical column (50 mm × 0.46 mm i.d., 3 μm particle size) from Teknokroma (Barcelona, Spain), which was thermostated at 25 °C. The mobile phases were water (phase A) and ACN (phase B) both containing 0.1% formic acid as ionization agent. The LC pump was programmed to use a flow rate of 0.6 mL/min and the following elution gradient: 2% phase B as initial mobile phase (1

min) and a ramp to 100% phase B (16 min). A post-time of 5 min was used to regain the initial conditions for the next analysis. Thus, the total analysis time per sample was 22 min (including postprocessing), which is quite acceptable for metabolomics studies frequently dealing with large batches of samples. The injected volume was 5 µL and the injector needle was washed 10 times with 20% ACN between injections. Also, the needle seat back was flushed with 20% ACN at 4 mL/min for 10 s to avoid cross contamination. The autosampler was kept at 4 °C to increase sample stability.

The settings of the electrospray ionization source, which was operated in the positive and negative ionization modes, were as follows: capillary voltage ±3.5 kV, fragmentor voltage 130 V, N_2 pressure in the nebulizer 40 psi; N_2 flow rate and temperature as drying gas 12 L/min and 325 °C, respectively. The instrument was calibrated and tuned as recommended by the manufacturer. MS/MS data were acquired in both polarities using the centroid mode at a rate of 2 spectra/s in the extended dynamic range mode (2 GHz). Accurate mass spectra over the *m/z* range 60–1200 were acquired in the MS scan mode, and MS/MS was performed with automated selection of two precursor ions per cycle and an exclusion window of 0.25 min after two consecutive selections of the same precursor. All samples were injected in triplicate, and the collision energy was set at 10, 20 or 30 V for the whole run in the three replicates of each sample.

The QTOF spectrometer provided a typical resolution of 15000 FWHM (Full Width at Half Maximum) at *m/z* 118.086255 and 30000 FWHM at *m/z* 922.009798. To ensure the required mass accuracy for the recorded ions, the spectrometer was subjected to continuous internal calibration during analyses by using the signals at *m/z* 121.0509 (protonated purine) and 922.0098 [protonated hexakis(1H,1H,3H-tetrafluoropropoxy)-phosphazine or HP-921] in the positive ionization mode, and those at *m/z* 119.0362 (deprotonated purine) and 966.0007 (formate adduct of HP-921) in the negative ion mode.

2.4.3. Serum pretreatment

Each serum pool aliquot of 100 µL was unfrozen slowly by immersion in an

ice bath. Then, 200 µL of methanol was added to each aliquot for protein precipitation, vortexed for 10 min and centrifuged at 4 °C at 13,800 × g for 5 min.

2.4.4. LC–QTOF MS/MS serum analysis

Chromatographic separation was performed at 25 °C using a C18 reversed phase analytical column (Mediterranea, 100 mm × 4.6 mm i.d., 3 µm particle size) from Teknokroma (Barcelona, Spain). The mobile phases used were water (phase A) and ACN (phase B) both containing 0.1% formic acid as ionization agent. The LC pump was programmed to operate at a flow rate of 0.8 mL/min in accordance with the following gradient elution sequence: 3% phase B for 1 min; phase B ramp from 3 to 100% from min 1 to min 25; and hold at 100% for 3 min. The injected volume was 5 µL and the injector needle was washed 10 times with 80%ACN between injections. Also, the needle seat back was flushed with 80% ACN at 4 mL/ min for 10 s to avoid cross contamination. The autosampler was kept at 4 °C to increase sample stability.

The electrospray ionization source was operated in the positive and negative ionization modes using the following settings: capillary and fragmentor voltage ±3.5 kV and ±175 V, respectively; N_2 nebulizer gas pressure 40 psi; and N_2 drying gas flow rate and temperature 10 L/min and 325 °C, respectively. The instrument was calibrated and tuned as recommended by the manufacturer. MS and MS/MS data were collected in the centroid mode, using a rate of 2.0 spectra per second in the extended dynamic range mode (2 GHz). Accurate mass spectra in auto MS/ MS mode were acquired over the *m/z* range 60–1100, and MS/MS *m/z* range 31–1100. The instrument gave typical resolution 15000 FWHM at *m/z* 118.086255 and 30000 FWHM at *m/z* 922.009798. Mass accuracy in recorded ions was assured by continuous internal calibration following the indications used for urine analysis. The auto MS/MS mode was configured with 2 maximum precursors per cycle and an exclusion window of 0.25 min after 2 consecutive selections of the same precursor ion. The collision energy selected was different in the three replicates of each analysis (10, 20, and 40 V) to increase the level of MS/MS information.

2.5. Data processing and pretreatment

All data obtained by LC–QTOF in the MS/MS mode were processed with the MassHunter Workstation software package (B.07.00 Qualitative Analysis and B.06.00 Profinder, Agilent Technologies, Santa Clara, CA). Potential molecular features from all injections performed in each ionization mode were extracted and aligned by using the recursive feature extraction algorithm in the MassHunter Profinder. This algorithm initially deconvolutes chromatograms and aligns features across the selected sample files in terms of mass accuracy and retention time; then, it uses the mass and retention time of each molecular feature for recursive targeted extraction. This two-step procedure reduces the number of false negatives and false positives in feature extraction. The target parameters for feature extraction included a threshold of 1500 counts for the monoisotopic peak and a maximum charge state of 2. In addition, the isotopic distribution of valid features had to be defined by two or more ions, with a peak spacing tolerance of 10.0 ppm. Adduct formation in the positive (+Na, +K, $+NH_4$) and negative ionization mode (+HCOO, +Cl) was also used together with protonated and deprotonated ions to identify features associated with the same metabolite. The presence of dimers and trimers was also taken into account. Features were aligned by using a tolerance window of 0.30 min and a mass accuracy of 10 ppm for retention time and *m/z* value, respectively, across all data files.

The minimum absolute height required for feature extraction in the recursive step was set at 3 000 counts for the sum of all peaks of the isotopic distribution of each molecular entity, which should be fulfilled in at least 75% of samples pertaining to each class. Background contribution was removed by subtracting entities linked to plasticizers, solvent impurities and other contaminants after analysis of a blank. The resulting data set for each polarity was exported as a *.csv* file containing the quantitative response, as peak area, of each potential entity in each sample, and the main characteristics of the entities —retention time and neutral mass, on the assumption that the detected ions were $[M+H]^+$ and $[M-H]^-$. Data pretreatment was based on baselining and normalization by quantile and

logarithmic transformation to reduce relatively large differences among molecular feature abundances.

2.6. MSCombine package architecture

MSCombine was developed and implemented by using the open source data analysis software R through its graphical user interface R-studio. The aim was to compile a single data set combining information from the positive and negative ionization data sets after removal of redundant information associated with molecular entities detected in both polarities.

Fig. 1 shows the architecture of MSCombine, while the detailed performance of the package is included as Supporting Information S1. First, the data sets containing information about potential entities in the positive and negative ionization modes exported from raw data files are input as a .csv file into R. Each data set should include two types of information. Thus, three columns contain descriptive information about molecular features such as entity identifier, calculated neutral mass and retention time (RT). The set also contains the quantitative response of each molecular entity found in the different samples —peak area included in this case. A third file containing a database with combinations between pairs of ions that could be detected in both ionization modes is also input into the R workspace. This database includes the mass difference in high resolution mode for each pair of ions in terms of neutral mass or *m/z* value depending on the software used for extraction of molecular features. Thus, this package can be implemented in metabolomics pipeline using either a specific software such as Qualitative Analysis from Agilent or universal packages such as XCMS. By way of example, Table 1 shows the previous bits of information for the most common adducts found with electrospray ionization. Once the three files have been compiled, the *Findcommon* function is executed specifying a mass and RT tolerance value in order to correlate the two data sets —positive and negative ionization. A new data set containing a list of potential entities detected in both data sets that could fit the same metabolite is thus obtained. The table also shows the compound identifier, detected adduct or ion, mass and retention time, mean peak

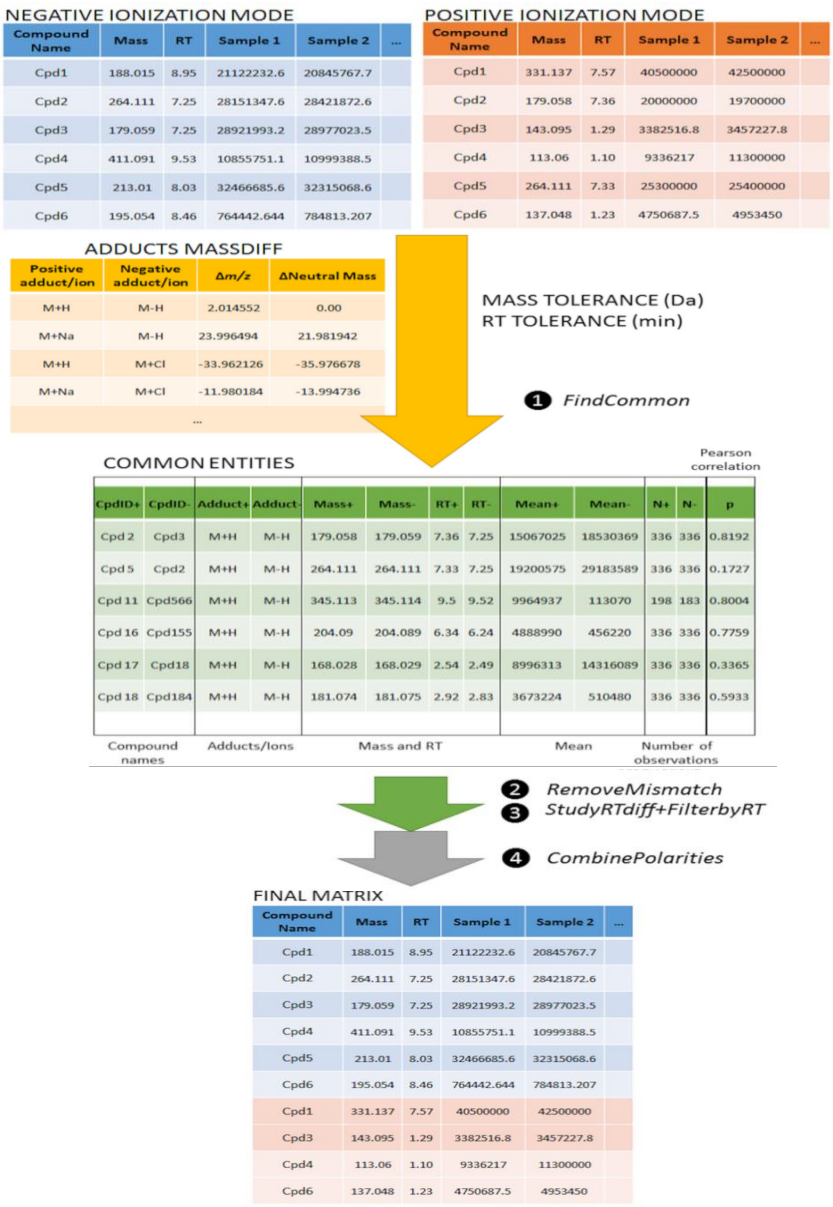


Fig. 1. Architecture of MSCombine, which uses the following functions: (1) FindCommon, (2) RemoveMismatch, (3) StudyRTdiff and FilterbyRT, and (4) CombinePolarities.

area, and number of samples in which the entities have been detected in each polarity, together with the Pearson *p* value for correlation between samples in the

two polarities. The table is exported as a *.csv* table into the working directory, where it can be opened and viewed before the second step is started.

Table 1. Example database listing mass differences between pairs of precursor ions detected in both polarities (Hac=acetic acid, FA=formic acid, TFA=trifluoroacetic acid).

Positive precursor ion	Negative precursor ion	Δ*m/z*	ΔNeutral Mass
M+H	M–H_2O–H	20.025666	18.011114
M+H	M–CH_3	16.030376	14.015824
M+H	M–H	2.014552	0
M+H	M+Cl	−33.962126	−35.976678
M+H	M+FA–H	−43.990925	−46.005477
M+H	M+Hac–H	−58.006575	−60.021127
M+H	M+TFA–H	−111.97831	−113.992862
M+H–H_2O	M–CH_3	−1.980738	−3.99529
M+H–H_2O	M–H	−15.996562	−18.011114
M+H–H_2O	M+Cl	−51.97324	−53.987792
M+H–H_2O	M+FA–H	−62.002039	−64.016591
M+H–H_2O	M+Hac–H	−76.017689	−78.032241
M+H–H_2O	M+TFA–H	−129.989424	−132.003976
M+K	M–H_2O–H	57.981548	55.966996
M+K	M–CH_3	53.986258	51.971706
M+K	M–H	39.970434	37.955882
M+K	M+Cl	3.993756	1.979204
M+K	M+FA–H	−6.035043	−8.049595
M+K	M+Hac–H	−20.050693	−22.065245
M+K	M+TFA–H	−74.022428	−76.03698
M+Na	M–H_2O–H	42.007608	39.993056
M+Na	M–CH_3	38.012318	35.997766
M+Na	M–H	23.996494	21.981942
M+Na	M+Cl	−11.980184	−13.994736
M+Na	M+FA–H	−22.008983	−24.023535
M+Na	M+Hac–H	−36.024633	−38.039185
M+Na	M+TFA–H	−89.996368	−92.01092
M+NH_4	M–H_2O–H	37.052213	35.037661
M+NH_4	M–CH_3	33.056923	31.042371
M+NH_4	M–H	19.041099	17.026547
M+NH_4	M+Cl	−18.950131	−18.950131
M+NH_4	M+FA–H	−26.964378	−28.97893
M+NH_4	M+Hac–H	−40.980028	−42.99458
M+NH_4	M+TFA–H	−94.951763	−96.966315

In the second step (*RemoveMismatch* function), the previous data set is examined to avoid incongruities such as assignation of two different ions (*e.g.*, $[M+H]^+$ and $[M+Na]^+$) in one ionization polarity to a single molecular feature, in

which case only the most repeated assignation is considered. On the other hand, if assignations are detected with the same frequency, then only the most common ion is considered (*viz.*, $[M+H]^+$ in the positive ionization mode, and $[M-H]^-$ in the negative mode). A new table and file are thus produced.

The third step (*StudyRTdiff* and *FilterbyRT* functions) is intended to avoid mismatching due to the RT tolerance filter. Because both analyses are carried out with the same chromatographic method, the RT values for molecular features detected in both ionization modes but associated to the same metabolites should fit a correlation model. For this reason, a regression model is built to compare the two RT lists. The *StudyRTdiff* function exports the RTpositive *vs* RTnegative plot, as well as the Q–Q plot for standardized residuals and the residuals *vs* predicted values plot. A histogram of residuals is also exported. With these graphs, a cutoff limit can be set with the *FilterbyRT* function to eliminate the entities with a residual value falling above or below a preset range. The result is a new data set named "CommonEntitiesFiltered" that is ready for use in the next step.

The last step (*CombinePolarities* function) uses the previous filtered data set to identify the specific entities that should be removed from the original (positive and negative ionization) data sets. Specifically, only the molecular feature exhibiting the strongest quantitative response among those in each group is retained in the data set. The final file combines the two lists of molecular entities into a single data set in the form of a *.csv* file that is saved in the working directory and ready for statistical analysis. The package is available at the Comprehensive R Archive Network (CRAN) Repository (https://cran.r-pro ject.org/index.html).

2.7. Statistical analysis

Final data sets were processed by using the software Mass Profiler Professional (MPP 13.0, Agilent Technologies) for statistical analysis. The area from urine metabolites was normalized by the MS total useful signals (MSTUS) procedure to correct for urine dilution. A *t*-test was used to identify entities significantly differing between classes, using a *p* value threshold of 0.01. A fold-change algorithm was used as a filter to retain the molecular entities undergoing a

change in relative concentration of 2 between pairs of classes. Mass profiler professional software also afforded multivariate analysis by principal component analysis (PCA) using data scaling by mean centering as a pretreatment.

The METLIN, Human Metabolome Database (HMDB) and Massbank databases were used for metabolite identification from both MS and MSMS information.

3. Results and discussion

3.1. Theoretical performance

As stated above, most untargeted metabolomic studies based on LC–MS analysis require two injections of each sample (one in the positive ionization mode and the other in the negative mode). Although a significant fraction of metabolites is preferentially detected in either mode, many compounds can be detected in both. Because the conventional approach produces two data sets —one per ionization mode—, two statistical analyses are required to identify the specific metabolites responsible for a given biological effect. Combining the two data sets would simplify statistical analysis provided duplicate inputs representing the same tentative metabolite are eliminated. Specific software such as Qualitative Analysis from Agilent, used for extracting molecular features or entities, estimates the theoretical neutral mass by assuming the presence of potential adducts or neutral losses in addition to the common $[M+H]^+$ or $[M-H]^-$ ions. However, if the protonated or deprotonated ion is not observed or extracted, the ions detected by adduct formation or neutral losses may appear as new inputs in the data set and be spuriously assigned to $[M+H]^+$ or $[M-H]^-$ ions. For this reason, merely comparing the two lists of neutral masses provided by the two ionization modes does not allow one to unambiguously assign molecular entities to the same tentative metabolite. The proposed package uses a database containing differences of neutral mass or *m/z* values in high resolution for pairs of ions that can be detected in the positive and negative ionization modes. This strategy enables the MSCombine package to

be used when the datasets include *m/z* for each entity, as is the case with universal packages such as XCMS, and with datasets containing theoretical neutral masses as those generated with Agilent pipeline.

Processing with the software starts after the LC–MS software has exported the data sets containing the list of molecular features detected by analysis in the positive and negative ionization modes. The two data sets are input as *.csv* files into R–studio together with the database that includes the difference of neutral mass or *m/z* value calculated for combinations of ions or adducts that can be detected in positive and negative polarity. Then, the first step (*FindCommon* function) is executed with RT tolerance set in minutes and *m/z* tolerance in dalton. These values should be chosen in accordance with the quality of the analyses and the resolution of the mass spectrometer used. The RT tolerance parameter can be set at a moderate 0.2 min if the quality of the chromatographic analyses is acceptable. Concerning mass accuracy, continuous calibration of the dual ESI source and TOF resolution (*R*=30 000) afford setting a mass accuracy cutoff at 0.002 Da. Application of this algorithm produces a new object named "*CommonEntities*" in the workspace that is exported as a *.csv* file in the working directory. The file contains a list of the molecular features which, based on the preset criteria, could pertain to metabolites detected in the two polarities.

The "*CommonEntities*" file is subjected to two filtering steps in order to confirm the molecular features associated with the same metabolite. The first step (*RemoveMismatch* function) removes mismatched entities resulting from assignation to two or more precursor ions. The algorithm identifies the most common assignation by detection frequency in the set of samples of a certain experiment and removes all others. If several assignations with the same frequency are detected, then the algorithm retains the most common (*viz.,* $[M+H]^+$ ion in the positive ionization mode, and $[M-H]^-$ in the negative mode). This algorithm produces a new object in the workspace that can be exported to the working directory for inspection.

The second algorithm deals with the entities which have been mismatched according to RT tolerance (*viz.*, those the retention times of which are not correlated according to the RTpositive *vs* RTnegative plot). This algorithm uses two different functions. One, named *StudyRTdiff*, builds the regression curve that associates RT values in the two polarities; also, it exports the corresponding histogram and residuals curve, which can be used to set a cutoff value for residuals in order to avoid mismatching of precursor ions. For example, ensuring a normal distribution of residuals in the case illustrated by Fig. 2 would require setting the residual limit at ±0.2 min. This limit is used by the second function, *FilterbyRT*, to remove mismatched entities and confirm the entities assigned to the same metabolite. This step is crucial to avoid mismatching peaks to isomers eluted at close RTs.

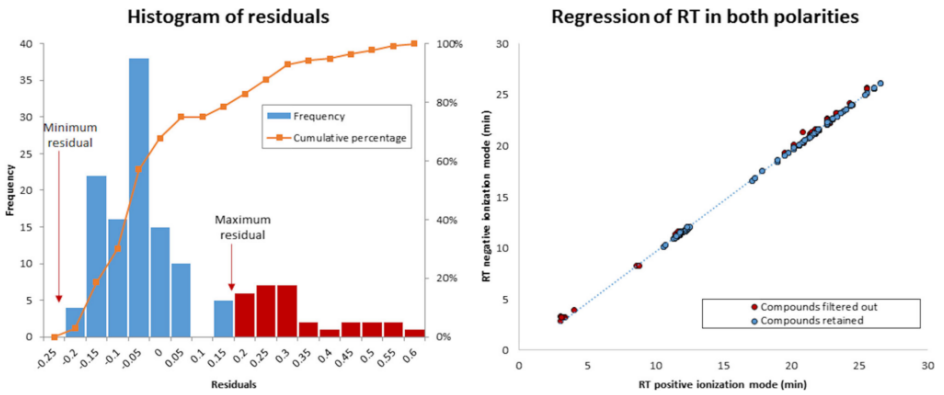


Fig. 2. Histogram of residuals for associated entities and regression curve obtained by comparing retention times (RT) for associated entities in both ionization modes.

Finally, the *CombinePolarities* function compares the mean area of the molecular entities detected in both polarities which have been associated to the same metabolite. The algorithm removes the features detected with the lowest sensitivity but also allows deleting other entities associated to the same metabolite detected in the same ionization mode. For example, if a metabolite is detected as two molecular entities (*e.g.*, $[M+H]^+$ and $[M+Na]^+$) in the positive ionization mode, but only one ($[M-H]^-$) in the negative mode, then only the entity exhibiting

the highest peak area will be included in the final data set. A new file named "*Positive–Negative"* containing the data set obtained by combining information from the analysis in both polarities is thus exported into the working directory.

3.2. Practical performance

The functionality of the proposed algorithm was assessed by using real data sets obtained from the analysis of serum and urine samples. Both are commonly used in metabolomic studies and contain a wide variety of metabolites including vitamins, amino acids, acylcarnitines, bile acids, fatty acids and phospholipids. Worth special note is the presence of glycerophospholipids in serum; these compounds are very easily ionized to various adducts in mass spectrometry —some families even in both polarities. Also, glycerophospholipids occur as positional isomers eluting closely in reversed phase chromatography. These characteristics allowed testing the ability of MSCombine to deal with different real-life situations such as the presence of metabolites giving no protonated or deprotonated ion upon ionization. As stated above, the resulting theoretical neutral mass is miscalculated as a result of the strongest signal being assigned to a tentative protonated or deprotonated precursor ion. Such is the case with glycerophosphatidylcholines, a common family of compounds widely detected in serum. In the presence of formic acid (FA), these compounds tend to form $[M+FA–H]^-$ adducts instead of $[M-H]^-$ ions in the negative ionization mode, and $[M+H]^+$ ions in the positive mode (Pulfer and Murphy, 2013). For this reason, the neutral mass calculated by any software will invariably differ between ionization modes. Supplementary Table S1 lists the lysophospholipids matched with MSCombine. As can be seen, the package accurately matched not only the formic acid adduct to the protonated ion, but also the common fragment of phosphatidylcholines $[M-CH_3]^-$, which is frequently detected in MS scans as a result of the cleavage of a methyl group in the source. If present, the $[M+Cl]^-$ adduct was also associated. In this example, the demethylated forms of lysoPC(14:0), lysoPC(16:0), lysoPC(18:0), lysoPC(18:1) and lysoPC(18:2) were pinpointed as redundant information, together with the $[M+FA–H]^-$ adduct in the negative mode corresponding to the protonated ion in the positive mode, which was detected with a higher sensitivity.

Most lysophospholipids have isomers differing in the location of the aliphatic chain (position sn-1 or sn-2). These isomers can be easily separated by reversed phase LC and eluted at close retention times. To avoid mismatching isomers, the algorithm uses a filter to build a regression curve between RTs in the two modes and removing the associated entities with a residual RT value outside the confidence interval. The residuals cutoff has to be specified after the regression curve and residuals histogram are produced. The histogram of residuals for the serum samples is shown in Fig. 2 together with the regression curve of RTs for the entities associated in both polarities. As can be seen, the residuals cutoff was set at ±0.2 min, which is an acceptable, non-stringent value in conventional chromatography. In fact, this is a common value used as an RT tolerance threshold to align molecular entities in LC–MS analysis. As can be seen in Supplementary Table S1, the filter removed all lysophospholipid isomers mismatched in the two polarities.

The number of entities removed by applying MSCombine to the serum and urine samples was proportionally similar. The sum of potential entities detected in the positive and negative ionization modes for serum and urine were 647 and 1451, respectively, from which the package removed 106 and 206 entities, respectively, corresponding to redundant information (Table 2). Therefore, 27 and 41% of the potential entities detected in the positive and negative ionization mode, respectively, corresponded to redundant information in serum. With urine, 26 and 31% of the potential entities detected in the positive and negative ionization mode, respectively, were also detected in the other mode. The 206 potential entities removed from the urine data sets corresponded to 165 tentative metabolites, for which different adducts can be detected. Similarly, the 106 potential entities removed from the serum data sets corresponded to 62 tentative metabolites.

Supplementary Table S2 includes some of the common entities detected in the positive and negative ionization modes in urine that were tentatively identified. As can be seen, this table contains a wide variety of compounds including amino acids, alkaloids and carboxylic acids, among others. Tentative identification of the 85 molecular entities that were detected in both serum sets revealed the massive presence of glycerophosphatidylethanolamines and glycerophosphatidylcholines,

Table 2. Number of molecular features detected in the positive and negative ionization modes, and those resulting after application of MSCombine tool in serum and urine

Sample	Ionization mode		Sum	Deleted entities	Total after MSCombine
	Positive	Negative			
Serum	387	260	647	106	541
Urine	787	664	1451	206	1245

which were removed in the positive and negative ionization mode, respectively, and fitted the less sensitive polarities for each family of phospholipids. Polar metabolites such as amino acids (tyrosine, phenylacetylglutamine) or lipids such as bile acids (glycocholic acid and glycodeoxycholic acid) were also efficiently removed from the data set for the least sensitive ionization mode.

The proportion of entities removed from the whole serum and urine matrices in the positive and negative ionization modes ranged from 14 to 16%. Although some compounds are known to be detected in both ionization modes, the proportion of urine metabolites detected in both ionization modes had never been accurately established. In fact, previous studies estimated the proportion of entities detected in opposite polarities in serum to be 2.8%, but this figure includes sodium adducts only; therefore, it is unrealistic (Sana *et al.,* 2008). In fact, based on our results, at least 27% of entities detected in one mode can also be detected in the other.

3.3. Applicability of MSCombine relative to the conventional data treatment

Once the proposed software was configured, it was tested by using it for statistical analysis of the resulting data sets. The results were compared with those obtained by independent analysis of data sets exported from raw data for the positive and negative ionization modes, as well as with the matrix built by merging the data sets for the two polarities without filtering. The positive and negative ionization data sets obtained by analysis of the urine cohort consisted of 787 and 664 molecular features, respectively. Combining the two matrices led to one consisting of 1451 entities, and application of the package reduced the data set to 1245 variables. A principal component analysis (PCA) was used to calculate the

coefficient of variation explained in each case. The combined matrix from the two ionization modes (with and without application of the software) led to a three-dimensional PCA matrix that explained a similar amount of variability: around 26% (see Fig. 3). Such a low proportion can be ascribed to the only major difference between the two sets of samples being treatment with ciprofloxacin or no treatment (control individuals). The similarity in explained variability can be ascribed to the information removed by the package being redundant. In both cases, PCA discriminated between the two classes according to metabolomic profile, namely: individuals taking ciprofloxacin and control individuals. Independent analysis of the data sets provided by the two ionization polarities led to similar results; thus, the proportion of explained variability did not exceed 30% and the two groups were well–differentiated. Judging by the entities detected in both polarities, some might have contributed to separation in the independent analyses. A deeper statistical analysis was conducted by filtering all experiments through a Volcano plot (p <0.01, fold change >2.0) in order to identify the significant entities explaining the observed variability between the two groups (control individuals and individuals on ciprofloxacin). A total of 41 features were significantly different between groups in the data set provided by MSCombine, and 54 based on the data set obtained by combining the positive and negative ionization matrices. The 13 molecular features absent from the MSCombine data set correspond to entities removed as redundant entities. In fact, tentative identification of those 13 entities revealed that they were associated to 10 metabolites already included in the MSCombine matrix. Those 10 metabolites were tentatively identified as hippuric acid, dopaquinone, pseudouridine, acetamido-butanoate, ascorbic acid, citric acid, isocitric acid, pyroglutamic acid, uric acid and α–*N*–phenylacetyl–L–glutamine. Redundancy in the entities associated with the same tentative metabolite was examined in terms of p value, means and 95% least significance difference (LSD) interval. It is worth emphasizing that the 13 deleted entities presented very similar statistical behavior (in terms of p value) to the 10 entities selected by MSCombine. Fig. 4 shows the mean and LSD interval for 4 of

the 10 tentative metabolites. As can be seen, all adducts and ions detected for each metabolite behaved identically.

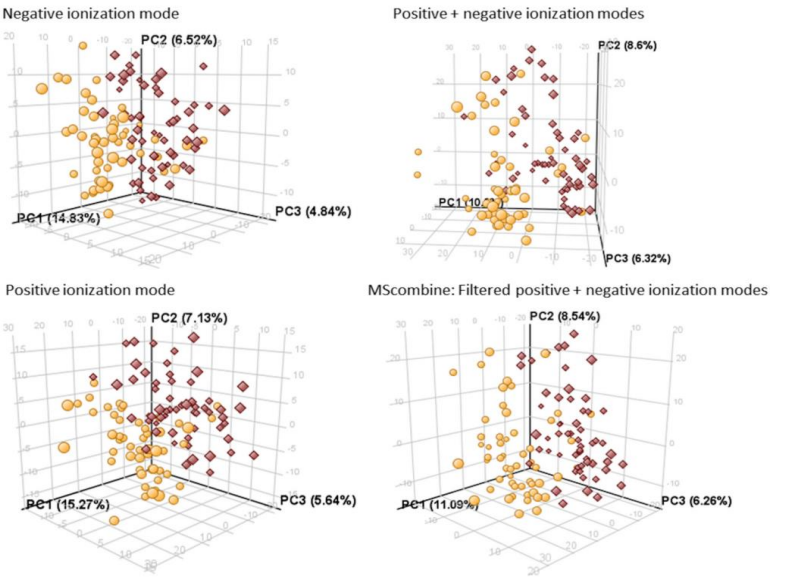


Fig. 3. PCA for urine samples from individuals on ciprofloxacin (filled diamond) and control individuals (filled circle). The data sets for the positive and negative ionization modes are shown separately and merged by application of MSCombine and without this tool.

Tentative identification of the remaining significant metabolites that were only detected in either polarity revealed the presence of conventional metabolites such as creatinine, urea, phosphoric acid, and sulfate derivatives such as *p*-cresol sulfate, in addition to ciprofloxacin, oxoprofloxacin, desethyleneciprofloxacin together with their sulfate and glucosylated derivatives.

4. Conclusions

The high chemical diversity of the metabolome in biological samples entails analyzing samples in the positive and negative ionization modes in untargeted analyses by LC–MS. This approach requires separate processing of data obtained in the two polarities, which is rather time-consuming. The proposed package

combines data from the positive and negative ionization modes in LC–MS to overcome this limitation. For this purpose, those entities introducing redundant

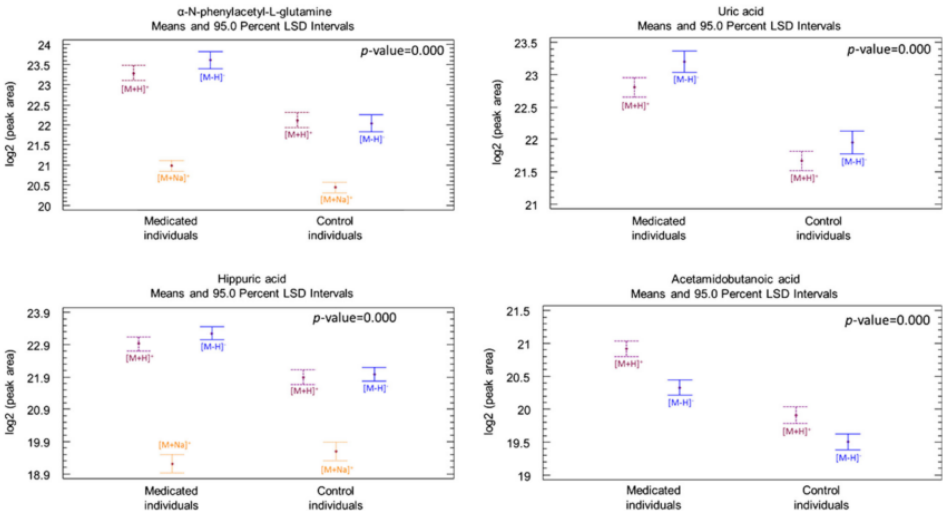


Fig. 4. Mean and 95% LSD interval for the different ions or adducts detected for α-N-phenylacetyl-L-glutamine, uric acid, hippuric acid and acetamidobutanoic acid.

information are removed and only the molecular features with the highest sensitivity for each tentative metabolite are retained. The package is an effective alternative to time-consuming analysis of data from each polarity set separately as it allows one to use a single matrix that is more representative of the target biological sample.

Acknowledgments

The Spanish Ministry of Economy and Competitiveness (MINECO), and the FEDER Program, are gratefully acknowledged for financial support (Project CTQ2012–37428). The Maimónides Biomedical Research Institute (Córdoba, Spain) is also grateful to the donors of the urine and serum samples.

References

Álvarez-Sánchez, B., Priego-Capote, F., & de Luque Castro, M. D. (2010). Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation. *TrAC–Trends in Analytical Chemistry*, 29(2), 111–119.

Arbulu, M., Sampedro, M. C., Gómez-Caballero, A., Goicolea, M. A., & Barrio, R. J. (2015). Untargeted metabolomic analysis using liquid chromatography quadrupole time-of-flight mass spectrometry for non-volatile profiling of wines. *Analytica Chimica Acta*, 858, 32–41.

Aslan, M., Özcan, F., Aslan, I., & Yücel, G. (2013). LC-MS/MS analysis of plasma polyunsaturated fatty acids in type 2 diabetic patients after insulin analog initiation therapy. *Lipids in Health and Disease*, 12, 169.

Calderón-Santiago, M., Priego-Capote, F., Turck, N., *et al.* (2015). Human sweat metabolomics for lung cancer screening. *Analytical and Bioanalytical Chemistry*, 407, 5381–5392.

Chen, Q., Park, H.-C., Goligorsky, M. S., *et al.* (2012). Untargeted plasma metabolite profiling reveals the broad systemic consequences of xanthine oxidoreductase inactivation in mice. *PLoS One*, 7(6), e37149.

Evans, C. R., Karnovsky, A., Kovach, M., Kovach, M. A., *et al.* (2014). Untargeted LC-MS metabolomics of bronchoalveolar lavage fluid differentiates acute respiratory distress syndrome from health. *Journal of Proteome Research*, 13(2), 640–649.

Fraser, K., Lane, G. A., Otter, D. E., *et al.* (2014). Non-targeted analysis by LC-MS of major metabolite changes during the oolong tea manufacturing in New Zealand. *Food Chemistry*, 151, 394–403.

Gika, H. G., Theodoridis, G. A., Plumb, R. S., & Wilson, I. D. (2014). Current practice of liquid chromatography-mass spectrometry in metabolomics and

metabonomics. *Journal of Pharmaceutical and Biomedical Analysis*, 87, 12–25.

Gilbert-López, B., García-Reyes, J. F., Meyer, C., *et al.* (2012). Simultaneous testing of multiclass organic contaminants in food and environment by liquid chromatography/dielectric barrier discharge ionization-mass spectrometry. *The Analyst*, 137(22), 5403–5410.

Herrera-Herrera, A. V., Ravelo-Pérez, L. M., Hernández-Borges, J., *et al.* (2011). Oxidized multi-walled carbon nanotubes for the dispersive solid-phase extraction of quinolone antibiotics from water samples using capillary electrophoresis and large volume sample stacking with polarity switching. *Journal of Chromatography A*, 1218(31), 5352–5361.

Kivilompolo, M., Öhrnberg, L., Oresic, M., & Hyötyläinen, T. (2013). Rapid quantitative analysis of carnitine and acylcarnitines by ultra-high performance-hydrophilic interaction liquid chromatography-tandem mass spectrometry. *Journal of Chromatography A*, 1292, 189–194.

Pulfer, M., & Murphy, R. C. (2013). Electrospray mass spectrometry of phospholipids. *Mass Spectrometry Reviews*, 22(5), 332–364.

Sana, T. R., Waddell, K., & Fischer, S. M. (2008). A sample extraction and chromatographic strategy for increasing LC/MS detection coverage of the erythrocyte metabolome. *Journal of Chromatography B*, 871(2), 314–321.

Sommer, U., Herscovitz, H., Welty, F. K., & Costello, C. E. (2006). LC–MS-based method for the qualitative and quantitative analysis of complex lipid mixtures. *Journal of Lipid Research*, 47(4), 804–814.

Vuckovic, D. (2012). Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry*, 403(6), 1523–1548.

Xiao, J. F., Varghese, R. S., Zhou, B., *et al.* (2012a). LC-MS based serum metabolomics for identification of hepatocellular carcinoma biomarkers in Egyptian cohort. *Journal of Proteome Research*, 11(12), 5914–5923.

Xiao, J. F., Zhou, B., & Ressom, H. W. (2012b). Metabolite identification and quantitation in LC–MS/MS-based metabolomics. *TrAC–Trends in Analytical Chemistry*, 32, 1–14.

Zhang, A., Sun, H., Wang, P., Han, Y., & Wang, X. (2012). Modern analytical techniques in metabolomics analysis. *The Analyst*, 137(2), 293–300.

Zhou, B., Xiao, J. F., Tuli, L., & Ressom, H. W. (2012). LC–MS- based metabolomics. *Molecular BioSystems*, 8(2), 470–481.

Supplementary material

Supplementary Table S1. List of phospholipids detected in serum in the positive and negative ionization modes (PIM and NIM, respectively). These entities had been matched by the *FindCommon* function. Mismatched entities removed by the *FilterbyRT* function are also shown. All lysoPC pertain to the Lipid Maps subclass GP0105 and lysoPE to GP0205. LysoPC with plasmalogen or alkyl ether substituent pertain to other classes, GP0107 and GP0106, respectively. Their specific ID can not be provided since isomers are not distinguishable by MS/MS.

Compound name	Adduct		Theoretical neutral mass		Retention time (min)		Mean area		Number of observations	
	PIM	NIM	PIM	NIM	PIM	NIM	PIM	NIM	PIM	NIM
LysoPE(16:0)	M+H	M-H	453.286	453.285	22.72	22.22	2.09E+06	2.18E+06	10	10
LysoPC(14:0) (1)	M+H	M+FA-H	467.302	513.307	20.19	19.82	1.52E+06	5.05E+05	12	12
	M+H	M+Cl	467.302	503.278	20.19	19.64	1.52E+06	1.46E+05	12	7
LysoPC(14:0)*	M+H	$M-CH_3$	467.302	453.285	20.19	20.04	1.52E+06	2.68E+05	12	10
	M+H	M+Cl	467.302	503.278	20.19	20.03	1.52E+06	6.59E+05	12	8
	M+H	M+FA-H	467.302	513.307	20.19	20.08	1.52E+06	3.66E+06	12	12
LysoPC(14:0) (2)	M+H	M+FA-H	467.302	513.307	20.58	20.08	1.12E+07	3.66E+06	15	12
	M+H	$M-CH_3$	467.302	453.285	20.58	20.04	1.12E+07	2.67E+05	15	10
	M+H	M+Cl	467.302	503.278	20.58	20.03	1.12E+07	6.59E+05	15	8
LysoPE(18:2)*	M+H	M-H	477.286	477.285	21.45	21.24	1.14E+06	4.93E+06	11	13
LysoPE(18:2) (1)	M+H	M-H	477.286	477.285	21.45	20.96	1.14E+06	1.57E+06	11	10
LysoPE(18:2) (2)	M+H	M-H	477.286	477.285	21.79	21.24	4.28E+06	4.93E+06	18	13
LysoPE(18:1)	M+H	M-H	479.301	479.301	23.36	22.87	2.03E+06	2.88E+06	10	10
LysoPC(P-16:0)	M+H	M+FA-H	479.337	525.343	24.05	23.58	5.50E+06	1.83E+06	12	10
LysoPC(15:0)*	M+H	M+FA-H	481.317	527.321	21.34	21.21	1.11E+06	9.77E+05	10	10
LysoPC(15:0) (1)	M+H	M+FA-H	481.317	527.322	21.34	20.87	1.11E+06	3.56E+05	10	10
LysoPC(15:0) (2)	M+H	M+FA-H	481.318	527.321	21.77	21.21	3.96E+06	9.77E+05	10	10

Cont. Supplementary Table S1

LysoPC(18:2)	M+H	M+FA-H	519.333	565.337	22.01	21.54	1.79E+08	2.73E+07	15	16
	M+H	M-CH_3	519.333	505.316	22.01	21.54	1.79E+08	1.26E+06	15	10
	M+H	M+Cl	519.333	555.309	22.01	21.65	1.79E+08	2.11E+06	15	6
LysoPC(18:1)*	M+H	M-CH_3	521.349	507.332	23.31	23.21	4.79E+06	8.71E+05	11	10
	M+H	M+Cl	521.349	557.325	23.31	23.19	4.79E+06	3.47E+06	11	6
	M+H	M+Cl	521.349	557.325	23.31	23.19	4.79E+06	3.35E+06	11	6
	M+H	M+FA-H	521.349	567.353	23.31	23.2	4.79E+06	3.26E+07	11	13
LysoPC(18:1) (2)	M+H	M-CH_3	521.349	507.332	23.66	23.21	1.23E+08	8.71E+05	15	10
	M+H	M+Cl	521.349	557.325	23.66	23.19	1.23E+08	3.47E+06	15	6
	M+H	M+Cl	521.349	557.325	23.66	23.19	1.23E+08	3.35E+06	15	6
	M+H	M+FA-H	521.349	567.353	23.66	23.2	1.23E+08	3.26E+07	15	13
LysoPC(18:0)*	M+H	M-CH_3	523.365	509.347	25.53	25.66	1.41E+07	8.10E+05	13	10
	M+H	M+Cl	523.365	559.34	25.53	25.64	1.41E+07	3.39E+06	13	6
	M+H	M+FA-H	523.365	569.369	25.53	25.6	1.41E+07	2.47E+07	13	15
LysoPC(18:0) (1)	M+H	M+FA-H	523.365	569.369	25.53	25.12	1.41E+07	4.65E+06	13	10
LysoPC(18:0) (2)	M+H	M+FA-H	523.365	569.369	26.07	25.6	1.24E+08	2.47E+07	14	15
	M+H	M-CH_3	523.365	509.347	26.07	25.66	1.24E+08	8.10E+05	14	10
	M+H	M+Cl	523.365	559.34	26.07	25.64	1.24E+08	3.39E+06	14	6
LysoPE(22:6)	M+H	M-H	525.286	525.285	21.7	21.24	3.10E+06	2.17E+06	11	10
LysoPC(20:5)	M+H	M+FA-H	541.317	587.322	20.84	20.4	3.85E+06	1.37E+06	10	10
LysoPC(20:4)*	M+H	M+FA-H	543.333	589.337	21.76	21.6	3.57E+06	5.35E+06	11	12
LysoPC(20:4) (2)	M+H	M+FA-H	543.333	589.338	21.96	21.6	2.00E+07	5.36E+06	15	12

Cont. Supplementary Table S1

LysoPE(18:0)	M+H	M-H	481.318	481.316	25.41	25.01	2.56E+06	1.95E+06	12	12
LysoPC(16:1)	M+H	M+FA-H	493.317	539.322	21.21	20.74	2.18E+07	6.25E+06	10	10
	M+H	M+Cl	493.317	529.293	21.21	20.72	2.18E+07	8.93E+05	10	8
LysoPC(16:0)*	M+H	M-CH_3	495.333	481.316	22.61	22.63	4.50E+07	1.37E+06	16	10
	M+H	M+Cl	495.333	531.309	22.61	22.64	4.50E+07	2.01E+06	16	6
	M+H	M+FA-H	495.333	541.337	22.61	22.58	4.50E+07	1.98E+07	16	19
LysoPC(16:0) (2)	M+Na	M-CH_3	517.315	481.316	23.05	22.63	5.18E+06	1.37E+06	18	10
	M+Na	M+Cl	517.315	531.309	23.05	22.64	5.18E+06	2.01E+06	18	6
	M+Na	M+FA-H	517.315	541.337	23.05	22.58	5.18E+06	1.98E+07	18	19
	M+H	M+FA-H	495.332	541.337	23.00	22.58	1.66E+08	1.98E+07	20	19
	M+H	M-CH_3	495.332	481.316	23.00	22.63	1.66E+08	1.37E+06	20	10
	M+H	M+Cl	495.332	531.309	23.00	22.64	1.66E+08	2.01E+06	20	6
LysoPC(16:0) (1)	M+H	M+FA-H	495.333	541.338	22.61	22.12	4.50E+07	7.82E+06	16	16
	M+H	M-CH_3	495.333	481.317	22.61	22.37	4.50E+07	9.66E+05	16	9
	M+H	M+Cl	495.333	531.309	22.61	22.05	4.50E+07	1.85E+06	16	6
LysoPE(20:4)*	M+H	M-H	501.286	501.285	21.51	21.33	4.27E+05	3.06E+06	11	10
LysoPE(20:4) (1)	M+H	M-H	501.286	501.286	21.51	21.03	4.27E+05	4.98E+05	11	9
LysoPE(20:4) (2)	M+H	M-H	501.286	501.285	21.82	21.33	3.03E+06	3.06E+06	14	10
LysoPC(O-18:1)	M+H	M+FA-H	507.369	553.373	24.48	24.00	6.05E+06	1.52E+06	10	10
LysoPC(17:0)	M+H	M+FA-H	509.349	555.354	24.4	24.02	4.34E+06	2.53E+06	11	10
LysoPC(hydroxy-16:0)	M+H	M+FA-H	511.327	557.332	17.34	16.87	1.21E+05	4.30E+05	13	10
LysoPC(18:3)	M+H	M+FA-H	517.317	563.321	20.96	20.52	3.53E+06	1.02E+06	11	10
LysoPC(20:3)	M+H	M+FA-H	545.349	591.353	22.86	22.38	8.00E+06	1.68E+06	10	10
LysoPC(20:2)	M+H	M+FA-H	547.364	593.368	24.36	23.93	2.39E+06	9.37E+05	11	10
LysoPC(20:1)	M+H	M+FA-H	549.38	595.384	26.58	26.14	1.41E+06	4.31E+05	10	10

* Mismatched isomers detected and removed by filtering of the residuals of a PIM RT *vs* NIM RT plot.

Supplementary Table S2. List of identified metabolites detected in urine in the positive and negative ionization modes (PIM and NIM, respectively). These entities had been matched with the *FindCommon* function. Mismatched removed by the *FilterbyRT* function are also shown.

Compound name	Adduct		Theoretical neutral mass		Retention time (min)		Mean area		Number of observations	
	PIM	NIM	PIM	NIM	PIM	NIM	PIM	NIM	PIM	NIM
Histidine (HMDB00177)	M+H	M-H	155.069	155.069	1.03	0.99	1.82E+05	2.58E+05	336	336
Acetyl-neuraminic acid (HMDB00230)	M+H	M-H	309.106	309.107	1.21	1.15	9.98E+04	1.61E+05	336	336
Isocitric acid (HMDB00193)	M+Na	M-H	214.009	192.028	1.54	1.5	2.38E+05	6.65E+06	336	336
Ascorbic acid (HMDB00044)	M+H	M-H	176.032	176.034	1.74	1.63	5.40E+05	2.25E+06	336	336
Pyroglutamic acid (HMDB00267)	M+H	M-H	129.044	129.043	2.07	1.99	1.09E+06	7.85E+05	336	336
	M+H	M+FA-H	129.044	175.049	2.07	1.79	1.09E+06	9.10E+05	336	336
Citric acid (HMDB00094)	M+H	M-H	192.028	192.028	2.09	2.02	8.48E+05	2.74E+07	336	336
	M+Na	M-H	214.009	192.028	2.09	2.02	1.15E+06	2.74E+07	336	336
Acetyl-glutamic acid (HMDB01138)	M+H	M-H	189.064	189.064	2.44	2.37	2.21E+05	3.66E+05	336	336
Uric acid (HMDB00289)	M+H	M-H	168.028	168.029	2.54	2.49	9.00E+06	1.43E+07	336	336
Hypoxanthine (HMDB00157)	M+H	M-H	136.039	136.039	2.60	2.57	2.28E+06	1.74E+05	336	336
Tyrosine (HDMB00158)	M+H	M-H	181.074	181.075	2.92	2.83	3.67E+06	5.10E+05	336	336
Xanthine (HMDB00292)	M+H	M-H	152.033	152.034	3.26	3.19	6.65E+05	6.85E+05	336	336
Acetamidobutanoate (HMDB60265)	M+H	M-H	145.074	145.074	3.49	3.4	1.37E+06	9.89E+05	336	336
	M+Na	M-H	167.055	145.074	3.5	3.4	2.05E+05	9.89E+05	336	336
Phenylalanine (HMDB00159)	M+H	M-H	165.079	165.079	4.55	4.42	4.88E+06	3.29E+05	336	336

Cont. Supplementary Table S2

4-Pyridoxic acid (HMDB00017)	M+H	M-H	183.053	183.054	4.60	4.5	5.68E+05	9.53E+05	336	336
cAMP (HMDB00058)	M+H	M-H	329.053	329.053	4.68	4.61	2.07E+05	1.74E+05	336	336
3-Methylxanthine (HMDB01886)	M+H	M-H	166.049	166.05	5.05	5.17	1.08E+06	4.91E+05	336	318
Methylxanthine*	M+H	M-H	166.05	166.05	5.26	5.41	5.57E+05	1.14E+06	336	336
	M+H	M-H	166.05	166.05	5.26	5.17	5.57E+05	4.91E+05	336	318
Pantothenic acid (HMDB00210)	M+H	M-H	219.112	219.111	5.31	5.21	6.62E+05	5.00E+05	336	336
1-Methylxanthine (HMDB10738)	M+H	M-H	166.049	166.05	5.50	5.41	1.36E+06	1.14E+06	336	336
Aminohippuric acid (HMDB01867)	M+H	M-H	194.069	194.069	5.65	5.51	7.66E+05	3.33E+05	336	336
Aspartyl-phenylalanine (HMDB28760)	M+H	M-H	280.105	280.106	5.88	5.78	6.64E+05	6.25E+05	336	336
Acetyl-valine (HMDB11757)	M+H	M-H	159.09	159.09	6.21	6.1	2.18E+05	2.47E+05	336	336
Pyrocatechol (HMDB00957)	M+H	M-H	110.037	110.037	6.33	6.5	9.50E+04	3.06E+05	331	328
Tryptophan (HMDB00929)	M+H	M-H	204.09	204.089	6.34	6.24	4.89E+06	4.56E+05	336	336

Cont. Supplementary Table S2

Glutamine (HMDB00641)	M+H	M-H	146.069	146.07	7.32	7.25	3.80E+05	8.46E+05	336	336
Phenylacetamide (HMDB10715)	M+H	M-H	135.068	135.069	7.32	7.25	2.19E+05	8.81E+05	336	336
Phenylacetyl-glutamine (HMDB06344)	M+H	M-H	264.111	264.111	7.33	7.25	1.92E+07	2.92E+07	336	336
Hippuric acid (HMDB00714)	M+H	M-H	179.058	179.059	7.36	7.25	1.51E+07	1.85E+07	336	336
	M+H	M+Cl	179.058	215.035	7.36	7.26	1.51E+07	1.19E+05	336	150
	M+Na	M-H	201.04	179.059	7.37	7.25	9.82E+05	1.85E+07	336	336
	M+Na	M+Cl	201.04	215.035	7.37	7.26	9.82E+05	1.19E+05	336	150
Hydroxyindole acetic acid (HMDB00763)	M+H	M-H	191.058	191.059	7.47	7.38	1.52E+05	1.51E+05	336	336
Acetyl-leucine (HMDB11756)	M+H	M-H	173.105	173.106	7.64	7.56	1.13E+05	1.86E+05	333	329
Indolelactic acid (HMDB00671)	M+H	M-H	205.074	205.075	9.29	9.27	3.74E+05	5.35E+05	330	317
Traumatic acid (HMDB00933)	M+Na	M-H	250.12	228.137	11.05	10.98	1.86E+05	2.34E+05	336	318

* Mismatched isomers detected and removed by filtering of the residuals of a PIM RT *vs* NIM RT plot.

CHAPTER 7

Study of exhaled breath condensate sample preparation for metabolomics analysis by LC–MS/MS in high resolution mode

Talanta 144 (2015) 1360–1369

Talanta

144 (2015) 1360–1369

Study of exhaled breath condensate sample preparation for metabolomics analysis by LC–MS/MS in high resolution mode

M.A. Fernández-Peralbo [a,b], M. Calderón Santiago [a,b], F. Priego-Capote [a,b], M.D. Luque de Castro [a,b]

[a] *Department of Analytical Chemistry, Campus of Rabanales, University of Córdoba, Córdoba, Spain*

[b] *Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía Hospital, University of Córdoba, Córdoba, Spain*

Study of exhaled breath condensate sample preparation for metabolomics analysis by LC–MS/MS in high resolution mode

M.A. Fernández-Peralbo, M. Calderón Santiago, F. Priego-Capote, M.D. Luque de Castro

Abstract

Metabolomic analysis of exhaled breath condensate (EBC) requires an unavoidable sample preparation step because of the low concentration of its components, and potential cleanup for possible interferents. Sample preparation based on protein precipitation (PP), solid-phase extraction (SPE) by hydrophilic and lipophilic sorbents or lyophilization has demonstrated that the analytical sample from the last is largely the best because lyophilization allows reconstitution in a volume as small as required (preconcentration factors up to 80-times with respect to the original sample), thus doubling the number of detected compounds as compared with the other alternatives (47 *versus* 25). In addition, PP and/or SPE cleanup are unnecessary as no effect from the EBC components removed by these steps appears in the chromatograms. The total 49 EBC compounds tentatively identified and confirmed by MS/MS in this research include amino acids, fatty acids, fatty amides, fatty aldehydes, sphingoid bases, oxoanionic compounds, imidazoles, hydroxy acids and aliphatic acyclic acids.

Keywords Exhaled breath condensate; LC–QTOF; Lyophilization; Metabolomics; Sample preparation.

1. Introduction

Exhaled breath condensate (EBC) is one of the most accessible human biofluids because of its inexpensive and non-invasive sampling and relatively simple composition [1,2]. EBC is representative of the airway lining fluid (ALF) of the respiratory tract [3]; therefore, its analysis of EBC could lead to identify potential biomarkers with capability to reflect the underlying state of pulmonary inflammation and alteration in various respiratory diseases such as asthma, chronic obstructive pulmonary disease (COPD), bronchiectasis, infections, cystic fibrosis, primary ciliary dyskinesia or even lung cancer [4–6]. In this context, EBC could be an alternative sample to those obtained by cumbersome and invasive techniques such as bronchoscopy with bronchoalveolar lavage or sputum induction.

EBC contains a mixture of volatile organic compounds (VOCs), non-volatile compounds and water [7]. Volatile compounds are present in EBC due to partitioning between the gaseous and the aqueous phase of the exhaled breath [8]. Thus, water condensation traps volatile substances with a high water–gas partition coefficient. The concentration of each individual VOC in EBC depends on multiple factors including water solubility, gas–liquid partition coefficient and temperature of the condenser. Non-volatile compounds are released to the exhaled air as aerosolized particles from the ALF. This fraction contains a broad range of compounds of different size and nature such as inorganic ions (*e.g.,* sodium or chloride) [9], small organic metabolites (*e.g.,* isoprene or acetone [10] and lipids, including leukotrienes [11] and isoprostanes [12]) or macromolecules (*e.g.,* cytokines or nucleic acids) [13]. In addition, polar compounds constitute a significant fraction of EBC due to the aqueous nature of the matrix [14]. The number and low concentration of EBC components demand for sample preparation steps involving preconcentration in dealing with research on this biofluid. The most used approaches to achieve this goal have been lyophilization (that allows as higher preconcentration factors as needed) and SPE, with less versatile preconcentration factors but with availability to endow the step with a given degree of selectivity [15]. The nature of the SPE sorbents (C18 sorbents [16] and immu-

noaffinity sorbents [17] in most cases) establishes the interaction with the analytes, and thus, the selectivity and sensitivity of the method in which this step is involved, which can be enhanced by mixtures of different sorbents. Sample concentration via lyophilization and reconstitution has been successfully applied before measurement of cytokines, proteins, 8-isoprostane, o- tyrosine and 8-OH-2′-deoxy-guanosine in EBC [18], and isoprostanes have been preconcentrated from EBC samples by either lyophilization or SPE [12,19]. A less common technique for sample preparation of EBC is liquid–liquid extraction, used to extract eoxins by tertiary butyl methyl ether [20], which was then evaporated for subsequent reconstitution of the extract components in methanol. However, evaporation is not recommended in sample preparation of EBC as it always involves a significant loss of semivolatile compounds.

Problems related to sensitivity of the analytical techniques have nowadays been overcome by the use of LC or GC coupled to mass spectrometers (MS), essentially operating in MS^2 mode [21]. There is not agreement in the literature about the most accurate, sensitive and specific approach to identify and quantify metabolites in EBC; therefore, the analytical technique is usually chosen as a function of the compounds to be determined. In fact, most studies dealing with the analysis of this biofluid have been focused on target compounds, as previously described [9–13,18,19,22], or fingerprinting [23–25], which allowed the identifycation of metabolites pertaining to chemical families such as amino acids (alanine, leucine, lysine, proline, phenylalanine, serine, threonine, tyrosine, valine), carboxylic acids (acetic acid, propionic acid, succinic acid), keto acids (pyruvate), hydroxyacids (lactic acid), alkylamines (trimethylamine) and fatty acids (butyric acid). Montuschi *et al.* studied EBC by NMR searching for metabolic profiles able to discriminate patients with both unstable and stable cystic fibrosis, and healthy individuals [26]. In this study, a total number of 11 metabolites were identified in EBC, including amino acids (glutamic acid, glutamine), carboxylic acids (acetic acid, formic acid, propionic acid), hydroxyacids (lactic acid), alcohols (propanol, methanol, ethanol, 2-propanol) and ketones (acetone).

Several methodological aspects of NMR spectroscopy for analysis of EBC have been addressed thus showing the reliability and reproducibility of this technique for this kind of samples [27]. Concerning MS, Carraro *et al.* described a method for EBC analysis without sample preparation based on LC–MS by using an Orbitrap mass spectrometer to discriminate between severe asthma cases and healthy controls [28]. Among the metabolites identified, retinoic acid, deoxy-adenosine, ercalcitriol, 20-hydroxy-PGF2α, thromboxane B2 and 6-keto-prosta-glandin F1 were relevant to discriminate the groups. In a comprehensive study carried out by Aksenov *et al.* on untargeted analysis of EBC from dolphins [29], preconcentration by lyophilization followed by both LC–MS and GC–MS allowed detecting amino acids, peptides, steroids, phospholipids, prostaglandins, carbohydrates and small molecules such as carbonic acids, amines and pharmaceuticals.

The detection of non-volatile compounds in EBC in previous studies is the reason why LC–MS has been used in this research to obtain a complementary metabolite profile to that offered by GC–MS. To avoid potential interferents and preconcentrate metabolites of interest were the main aims in this study to achieve appropriate selectivity and sensitivity of the instrumental signal. Therefore, three different alternatives for processing EBC prior to untargeted analysis by LC–MS/MS —*viz.,* protein precipitation (PP), SPE with two complementary sorbents based on lipophilic and hydrophilic interactions (HI) and EBC lyophilization— were assayed. The profiles of the compounds tentatively identified by the different strategies were then used to discuss their applicability for the development of a method for EBC analysis.

2. Materials and methods

2.1. Reagents

Acetonitrile (ACN) (LC–MS optima grade) from Fisher Scientific (Fair Lawn, NJ, USA) and Milli-Q water (Millipore, Bedford, MA, USA) were used for sample

preparation and to prepare chromatographic mobile phases. MS-grade formic acid from Scharlab (Barcelona, Spain) was used as ionization agent in LC–MS/MS analysis. Ammonium formate from Fluka (Sigma–Aldrich, St. Louis, USA) and formic acid were used to improve the retention and elution steps in the SPE protocols.

2.2. Apparatus and instruments

EBC was collected during spontaneous breathing using the condenser ECOScreen-2 (FILT Thorax-und Lungen Diagnostik GmbH, Berlin, Germany). A commercial filter from Scharlab (Barcelona, Spain) was placed over the inlet air valve to retain undesired organic compounds and particles present in the room atmosphere.

Centrifugation was carried out by a Thermo Sorvall Legend Micro 21 R thermostated centrifuge from Thermo (Thermo Fisher Scientific, Bremen, Germany). A lyophilizer (FTS Systems, Stone Ridge, New York) was used to remove the aqueous phase of the sample.

The EBC samples were analyzed using an Agilent 1200 Series LC system (consisting of a binary pump, a vacuum degasser, an autosampler and a thermostated column compartment) hyphenated to an Agilent 6540 QTOF with dual electrospray ionization (ESI) source (Agilent Technologies, Santa Clara, CA, USA). Agilent MassHunter Workstation was the software for data acquisition.

2.3. EBC sampling and collection

All steps involved in sampling and storage were developed according to the guidelines dictated in 2004 by the World Medical Association Declaration of Helsinki. A group of four healthy subjects who did not suffer from any disease or respiratory infection within the last 30 days were encouraged on an empty stomach to perform tidal breath through a mouthpiece connected to the commercially available condenser EcoScreen-2 at (–20 °C) while wearing a nose-clip. Sampling was carried out for 20 min, in which approximately 4.5 mL of EBC was collected from each individual. An EBC pool from all collected samples was aliquoted in

Eppendorf tubes that were immediately frozen and stored at –80 °C. 2-mL aliquots of water stored for 20 min in EBC collection bags were lyophilized and used as analytical blanks to evaluate the presence or absence of contamination owing to the collectors.

2.4. Sample preparation

Four different sample preparation strategies —protein precipitation (PP), SPE with two different sorbents (C18 and HI), combination of PP and SPE, and lyophilization— were used to evaluate the influence of this step on the detection coverage.

2.4.1. Protein precipitation

500 µL of the EBC pool was deproteinized with 500 µL of ACN. The mixture was vortexed for 1 min, then cooled at 4 °C for 5 min, the resulting precipitate separated by centrifugation at 14 000 × *g* for 5 min at 4 °C, and the liquid phase isolated. The step involved a dilution factor of 2.

2.4.2. Solid-phase extraction

Aliquots of EBC were cleaned up and preconcentrated by centrifugal SPE using C18 or HI (Hydrophilic Interaction) MSCs (Micro Spin Cartridges) from Harvard Apparatus (Holliston, Massachusetts, USA) following the protocol recommended by the manufacturer for each sorbent. The protocol for C18-MSC was as follows: 150 µL of water for solvation, 150 µL of ACN for sorbent conditioning, 150 µL of 96:4 water–ACN for sorbent equilibration, 150 µL of sample, 150 µL of water to obtain the sample free from the undesirable retained compounds which were subsequently eluted by 75 µL of 96:4 ACN–water containing 0.1% formic acid in two steps. All centrifugation steps were developed at 4 °C and 1 000 × *g*, with the exception of the second elution step, which was developed at 3 000 × *g* to ensure total recovery of the eluate.

The protocol for HI-MSC was: 150 mL of water for solvation, 150 µL of ACN for sorbent conditioning, 150 µL of 90:10 ACN–water containing 1mM ammonium formate for sorbent equilibration, 75 µL of sample 1:1 diluted with ACN, 150 µL of

90:10 ACN–water containing 1mM ammonium to remove undesired no retained compounds, and 150 μL of 90:10 water–ACN containing 1mM ammonium formate to elute the retained target compounds. The centrifugation conditions were as for the C18 sorbent. In both SPE protocols the preconcentration factor was 2. The eluates thus obtained constituted in both cases the analytical samples.

2.4.3. Combination of PP and SPE

A combination of PP and SPE was also tested to complete the study. To aliquots of 150 μL of the liquid phase after PP were applied the SPE protocols; therefore, no preconcentration nor dilution factors resulted from the overall step.

2.4.4. Lyophilization

Depending on the preset preconcentration factor, 1- and 2-mL aliquots of the EBC pool were lyophilized for 12 h at −40 °C and at a pressure of 40 mT. The resulting residues were reconstituted with volumes of 1:1 ACN–water containing 0.1% formic acid to obtain preconcentration factors of 20, 40 and 80 with respect to the initial volume of the aliquots. The resulting solutions were analyzed in triplicate by LC–MS/MS.

2.5. LC–QTOF MS/MS analysis

LC separation and QTOF detection were used for analysis of the EBC. Chromatographic separation was performed by using a Mediterranea Sea C18 analytical column (50 × 4.6mm i.d., 3 mm particle size) from Teknokroma (Barcelona, Spain) which was thermostated at 25 °C. The initial mobile phase was a mixture of 98% phase A (0.1% formic acid in water) and 2% phase B (0.1% formic acid in ACN). After injection, the initial mobile phase was kept under isocratic conditions for 1 min. Then, a linear gradient of phase B from 2% to 100% was applied in 16 min. The flow rate was 0.6 mL/min during the whole chromatographic step. The total analysis time was 17 min, and 5 min were required to re-establish and equilibrate the initial conditions. The injected volume was 5 μL and the injector needle was washed 5 times with 20:80 ACN–water between injections. Also, the needle seat back was flushed with 20:80 ACN–water at 4 mL/min for 10

s to avoid cross contamination. The autosampler was kept at 4 °C to increase sample stability.

The settings of the electrospray ionization source, operated in the negative and positive ionization modes, were as follows: capillary voltage ±3.5 kV, fragmentor voltage 130 V, N_2 pressure in the nebulizer 40 psi; N_2 flow rate and temperature as drying gas 8 L/min and 325 °C, respectively. The instrument was calibrated and tuned according to procedures recommended by the manufacturer. Each sample was analyzed in triplicate by setting data acquisition in centroid mode at 2 spectra per second in MS mode and 1 spectrum per second in MS2 mode in the extended dynamic range mode (2 GHz). A gas-phase fractionation approach was used to increase the identification capability [30]. For this purpose, the acquisition range in MS scanning mode was from 60 to 1200 *m/z*. The maximum number of precursors selected per cycle for MS/MS fragmentation was set at 2 with an exclusion window of 0.25 min after two consecutive selections of the same precursor. Criteria for selection of precursor ions were established for each of the three runs per sample to obtain the maximum spectral information. Thus, in both ionization modes the mass range for precursor selection in the first run was set at 60–250 *m/z* with 10 eV as collision energy. The second and third runs involved precursor selection in the range 250–450 and 450–1200 *m/z* with collision energies of 20 and 30 eV, respectively. The acquisition range in MS/MS mode was from 30 to 1200 *m/z*. To assure the desired mass accuracy of the recorded ions, continuous internal calibration was performed during analyses by using signals at *m/z* 121.0508 (protonated purine) and *m/z* 922.0097 [protonated hexakis (1H, 1H, 3H-tetrafluoropropoxy) phosphazine or HP-921] in positive ionization mode. In negative ionization mode, ions with *m/z* 119.0363 (purine anion) and *m/z* 1033.9881 adduct of HP-921 were used.

2.6. Data processing and statistical analysis

The MassHunter Workstation software package (B.05.00 Qualitative Analysis and B.06.00 Profinder, Agilent Technologies, Santa Clara, CA, USA) was

used to process the data obtained by LC–QTOF in MS/MS mode. The recursive feature extraction algorithm in the software MassHunter Profinder was used to extract and align potential molecular features in all injections. This algorithm initially extracts all potential entities from the chromatograms and aligns these features across the selected sample files in terms of mass and retention time; then, it uses the mass and retention time of each feature for recursive targeted feature extraction. This two-step procedure reduces the number of both false negatives and false positives in feature extraction. The target parameters for feature extraction included a threshold of 1500 counts and a maximum charge state of 2. In addition, the isotopic distribution for a valid feature had to be defined by two or more ions —with a peak spacing tolerance of 0.0025 *m/z*, plus 10.0 ppm. The features were aligned by using a tolerance window of 0.30 min and a mass accuracy of 10 ppm for retention times and *m/z* values across all data files, respectively. The minimum absolute height required for feature extraction was set at 3000 counts. The molecular features extracted from the analysis of analytical blanks were removed from those extracted from EBC samples.

Once all features were extracted and aligned, the software MassHunter Qualitative was used for the targeted extraction of MS/MS information from all molecular features in the whole set of analyses. The entities were tentatively identified by using MS and MS/MS information, and searching the METLIN MS and MS/MS database (http://metlin.scripps.edu) and the Human Metabolome Database (HMDB, version 3.5).

Finally, raw data files were created in compound exchange format (*.cef* files) 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. The sum of the area of all compounds detected per sample was used for the normalization of each compound. Venn diagrams were then performed to compare the number of entities detected by LC–QTOF after each sample preparation protocol. Only entities tentatively identified by MS/MS were included in the Venn diagrams.

3. Results and discussion

As commented before, shortcomings in the analysis of EBC are the low concentration of metabolites and the lack of assays with enough sensitivity for clinical use. Different sample preparation protocols looking for sufficient preconcentration and interferents removal, if required, are compared and discussed below. The repeatability (intra-day) was satisfactory with the CV ranging from 2.6% to 15.3%.

3.1. Comparison of SPE sample preparation by C18 versus HI

The sorbents were selected as a function of their selective sorbent–analyte interaction mechanisms to allow the extraction/preconcentration of different families of metabolites present in the sample. Two types of sorbents, considered as complementary, were evaluated: C18 sorbent for lipophilic interactions, and HI sorbent mainly for hydrophilic interactions. Centrifugal SPE in a miniaturized format was selected for development of the protocol using both C18- and HI-MSCs cartridges. The analytical sample thus prepared was injected in the LC–QTOF equipment for separation by a C18 chromatographic column and analysis both in positive and negative ionization modes.

The profiles of total ion chromatograms (TICs) of the sample prepared by HI-MSC were different to those resulting from C18- MSC preparation, as Fig. 1 shows. This behavior was foreseeable taking into account the retention mechanisms of the sorbents used in both types of devices. Thus, the eluate from the HI sorbent provided, at short retention times and using the same chromatographic method, peaks higher than the eluate from C18; on the contrary, the chromatographic profile provided by the latter eluate was characterized by more intense peaks at long retention times. Therefore, both sorbents provided complementary information.

Qualitative comparison of the chromatographic profiles from the eluates obtained by both sorbents was carried out by the Venn diagrams, as shows Fig. 2.A, while tentative identification provided the 25 compounds listed in Table 1.

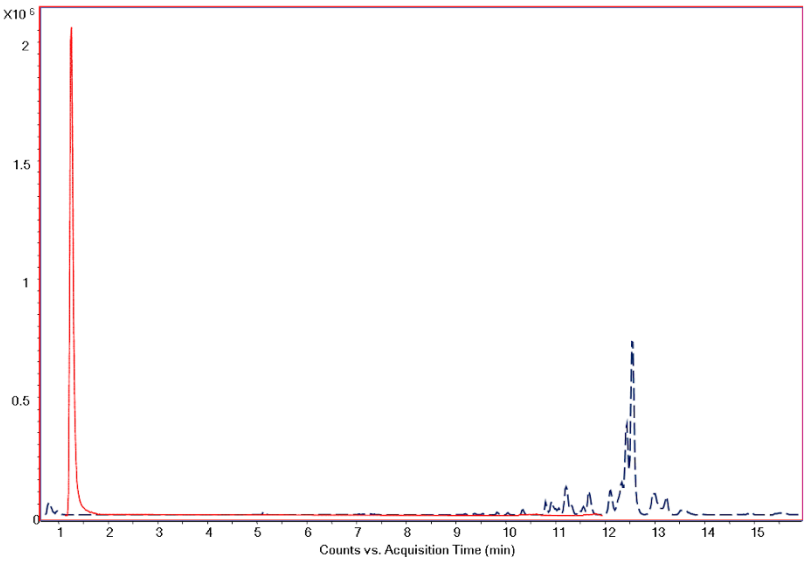


Fig. 1. Profiles of the total ion chromatograms (TICs) from samples prepared by SPE-HI (continuous red line) and by SPE-C18 (discontinuous blue line).

Three of these compounds were exclusively detected in the eluate from the HI sorbent: aspartic and phosphoric acid in both negative and positive ionization modes, and pyrophosphate only in the negative ionization mode. The C18 sorbent was not able to retain these three polar compounds owing to the non-polar nature of their interactions. Six metabolites were detected in the eluates from both sorbents due to their interaction with the two materials. Among these compounds it is worth mentioning lactic acid, able to interact with both sorbents with a similar efficiency, as Supplementary Fig. 1 shows. The other common tentatively identified compounds —*viz.,* nonanedioic acid, decanamide, dodecanamide, (d14:1) sphingosine, heptadecasphinganine— are characterized by the presence of a polar group (amide, amine or carboxylic acid) and an aliphatic chain, which are responsible for the dual retention mechanisms. Among them, nonanedioic acid, decanamide and (d14:1) sphingosine were also extracted with the same efficiency by the two sorbents, while dodecanamide, and specially heptadecasphinganine, were preferentially extracted by the C18 sorbent, as Supplementary Fig. 1 shows.

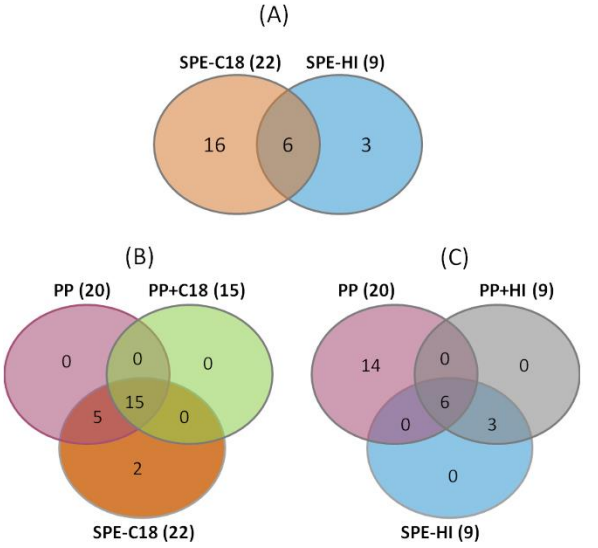


Fig. 2. Venn diagrams comparing compounds tentatively identified after, (A) SPE treatment by both C18 and HI sorbents; (B) PP, SPE-C18 and a combination of PP and SPE- C18; (C) PP, SPE-HI and a combination of PP and SPE-HI.

On the other hand, 16 exclusive metabolites were identified in the analytical sample constituted by the eluate from the C18 sorbent. This set of metabolites is characterized by a high chemical heterogeneity: three alkylamines, one amino acid derivative (nonanoylglycine), two fatty acids amino derivatives, one fatty aldehyde, two fatty amides, two sphinganines, one sphingenine (4-hydroxy-8-sphingenine), three sphingosines with long chain fatty alcohols (C18:1, C18:2 and C18:3) and one imidazol (urocanic acid).

3.2. Deproteination versus solid-phase extraction

PP should improve chromatographic retention of target analytes by removal of potential interferents such as proteins, unless compounds of interest coprecipitate with them. Owing to the very low concentration in EBC of most of the compounds, this sample preparation step, which does not involve preconcentration, was not useful when used isolated; therefore, combinations of PP with SPE were tested for sample preparation. Fig. 2B shows that PP by itself did not allow identi-

Table 1. List of metabolites identified in EBC classified by families. Information about MS/MS identification (precursor and product ions) is included.

Family	Name	RT	Precursor ion (*m/z*) [adduct]	Product ions (*m/z*)
Alkylamines	[a]Diethanolamine	1.04	106.0864 [M+H]	70.0640/88.0753/44.0491
	[b]2-Diethylaminoethanol	1.39	118.1225 [M+H]	100.1122/72.0806/43.0176
	[a]2-(Diethylamino)ethane-1,1-diol	1.66	134.1173 [M+H]	116.107/106.0863/86.0964
	[a]Trolamine	1.07	150.1121 [M+H]	89.059/104.6476/132.1023
	[b]Tris(3-hydroxypropyl)amine	1.74	192.1597 [M+H]	175.1510/156.1356/101.1165/ 98.0958/174.1483/55.0393
	[b]2,2′-(Octylimino)diethanol	8.67	218.2117 [M+H]	88.0756/106.0854/57.0697/ 200.2035
Amino acids and derivatives	[b]Valine	1.50	118.0862	72.0811/55.0544
	[a]1-Pyrroline-4-hydroxy-2-carboxylate	1.68	130.0498 [M+H]	86.0961/57.0697/41.0381/ 84.0467/56.0497
	[a]Leucine	1.69	132.102 [M+H]	86.0961
	[a]2,4-Diaminopentanoate	0.95	133.0973 [M+H]	116.0691/69.0338/66.9465
	Aspartic acid	1.13	134.0448 [M+H]	88.0392/74.0227/46.0362
	[a]Glutamic acid	1.18	148.0601 [M+H]	130.0485/84.0454
	[a]Phenylalanine	4.89	166.0879 [M+H]	120.0806/103.0541

Cont. Table 1

	[a]Phenylalanine	4.89	166.0879 [M+H]	120.0806/103.0541
	[a]Citrulline	1.21	176.1029 [M+H]	159.0758/133.0951/113.0708/ 107.0846/70.0652
	[a]Glutamic acid n-butyl ester	6.24	204.1232 [M+H]	187.9119
	[b]N-Nonanoylglycine	9.94	214.1455 [M+H]	197.1178/153.1295/118.1502/ 41.9985
	[a]N(alpha)-t-Butoxycarbonyl-L-leucine	8.23	230.1399 [M-H]	59.0139/170.1185/172.0988
	[a]N(6)-(Octanoyl)lysine	6.52	273.2178 [M+H]	200.1283/100.1128/128.1066/ 201.1310/241.1623/73.0295
Fatty acids	[a]7-Aminoheptanoic acid	5.90	146.1174 [M+H]	128.1060/83.0856/55.0545
	[a]5-(Acetylamino)pentanoate	6.07	160.097 [M+H]	142.0866/125.0596/114.0923/ 115.0943
	[b]8-Aminooctanoic acid	7.15	160.1327 [M+H]	97.1010/142.1221/43.0180/ 47.0143/125.0942/100.1136/ 143.1053
	[a]Octanedioic acid	8.76	173.0821 [M-H]	111.0814
	[b]8-Amino-7-oxo-nonanoic acid	8.43	186.1137 [M-H]	125.0971
	[c]Nonanedioic acid	9.6	187.0983 [M-H]	125.0972
	[a]Decanedioic acid	10.35	201.1139 [M-H]	139.1125/183.1023
	[a]2-Amino-9,10-epoxy-8-oxodecanoic acid	9.60	214.1091	125.09846

Cont. Table 1

	[a]12-Amino-dodecanoic acid	10.56	216.1956 [M+H]	143.1054/43.0538/71.0860/ 83.0854/170.1888/181.1604/ 67.0538
Fatty aldehydes	[b]6-Hydroxy-2,4-hexadienal	12.18	113.0597 [M+H]	69.0340/59.9905/41.0382
	[a]9,12-Tetradecadienal	12.98	209.1903 [M+H]	121.1020/193.1596/135.1158/ 79.0536/85.0645/191.1780/ 149.1341/69.0705/113.0942/ 155.1436
Fatty amides	[c]Decanamide	12.86	172.1695 [M+H]	64.7460/88.0758/43.0542
	[c]Dodecanamide	14.30	200.2006 [M+H]	88.0764/141.1266/57.0688/ 81.0678/44.0121
	[b]Heptanoylhomoserine lactone	9.72	212.1295 [M-H]	168.1394/41.9987/194.1183
	[a]2,4-Decadienoic Isobutylamide	11.27	224.2007 [M+H]	107.0848/43.0539/69.0685/ 67.0542/95.0842/71.0508/ 91.0548/121.1004
	[b]9,12-Octadecadienamide	13.97	280.2632 [M+H]	57.0693/69.0697/55.0546/ 67.0529/127.1102/245.2280/

Cont. Table 1

Sphingoid bases	[a]1-Deoxy-tetradecasphing-13-enine	15.94	228.2316 [M+H]	85.0996/102.0912/68.0574
	[c](4E,d14:1) Sphingosine	12.84	244.2273 [M+H]	98.0850/57.0693/109.1013/ 43.0176/71.0855/81.0855/ 123.0799/107.0858/57.0326
	[b]Hexadecasphinganine	10.78	274.274 [M+H]	256.2635/106.0863
	[c]Heptadecasphinganine	11.08	288.2902 [M+H]	270.2794/88.0758/106.0859
	(4E,8E,10E-d18:3) Sphingosine	14.76	296.2586 [M+H]	57.0697/81.0698/153.1269/ 123.1173/141.1272
	[b](4E,8Z,d18:2) Sphingosine	12.99	298.2744 [M+H]	57.0697/43.0544/81.0700/ 119.0860/93.0687/181.1238/ 281.2525
	[b](4OH,8Z,t18:1) Sphingosine	13.02	316.2849 [M+H]	83.0857/57.0695/135.1166/ 109.1014/111.1158
	[b]4-Hydroxy-8-sphingenine	12.49	316.2849 [M+H]	69.0701/128.1066/97.1013/ 55.0545/155.1418/149.1328/ 137.0971/81.0680
	[b]4-Hydroxysphinganine	10.83	318.3001 [M+H]	256.2631/300.2916/88.0755/ 209.1647
Aliphatic acyclic compounds	[a]Urea	1.24	61.0393 [M+H]	44.0129
Imidazoles	[b]Urocanic acid	1.66	139.0502 [M+H]	121.0393/93.0436
	[a]3-(2-Propan-2-ylimidazol-1-yl)propanimidamide	1.40	181.1443	138.1011/42.0337/111.0916/

Cont. Table 1

			[M+H]	85.0745/112.0881/139.1095/ 123.0910/69.0345
Hydroxy acids	[c]Lactic acid	1.77	89.0246 [M-H]	43.0193/71.0135
Oxoanionic compounds	[d]Phosphoric acid	1.26	96.9699 [M-H]	78.9587
	[d]Pyrophosphate	1.26	176.9353 [M-H]	158.9249/78.9589

[a] Compounds only detected in lyophilized EBC.
[b] Compounds exclusively detected in the eluate from C18-SPE in comparison with that from HI-SPE.
[c] Compounds on grey background were detected in eluates from both sorbents.
[d] Compounds exclusively detected in the eluate from HI-SPE.

fication of any compound additional to those obtained either by the SPE and by the combination of PP and SPE involving C18 as sorbent. On the contrary, when the analytical sample generated by PP was compared with the eluate from SPE involving the HI sorbent (Fig. 2C) 14 compounds were detected in the former analytical sample, which were not found in the latter. These compounds were clearly identified when the C18 sorbent was selected for SPE.

The capability of each sample preparation approach was quantitatively evaluated for the six main families of identified compounds: alkylamines, amino acids and derivatives, fatty acids, fatty aldehydes, fatty amides and sphingoid bases. For this purpose, the peak area of the compounds pertaining to the same chemical family was summed up and divided by the total peak area considering all the identified compounds. The resulting value was used as quantitative response. The comparison among sample preparation approaches is presented in Fig. 3 for the six chemical families of compounds. As the bars diagram shows, SPE with C18 sorbent was the suited sample preparation for detection of alkylamines, amino acids derivatives, fatty aldehydes and fatty amides. A PP step before SPE with the C18 sorbent was the preferred approach for detection of sphingoid bases, while PP with subsequent analysis of the supernatant phase was the best alternative for detection of fatty acids and derivatives.

3.3. Lyophilization as alternative for EBC processing

Lyophilization provided the highest number of identified compounds (a total of 47 compounds) as compared to the previous sample preparation procedures assayed. Preconcentration factors of 20, 40 and 80 were achieved using the lyophilized from 1 or 2 mL of EBC reconstituted in volumes of the 1:1 ACN–water 0.1% formic acid solution from 25 to 50 mL. Nevertheless, the increase of sensitivity ascribed to the concentration effect did not have a proportional effect on the detection coverage since the number of compounds tentatively identified did not increase with the preconcentration factor. In fact, only one extra metabolite, urea, was identified in the reconstituted lyophilized sample with the highest preconcentration factor. Urea was identified according to the MS/MS spectrum illustrated in

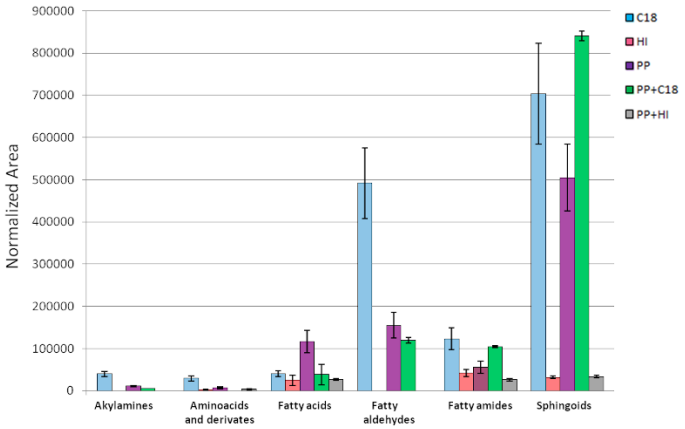


Fig. 3. Histogram plotting the sum of the areas of the compounds determined in each family after different EBC treatments (PP, SPE-C18, SPE-HI, and a combination of PP and SPE-C18 or SPE-HI).

Supplementary Fig 2. The EICs for monitoring this metabolite revealed how it could only be detected in the reconstituted sample with a preconcentration factor of 80. Taking into account the biochemical relevance of the urea cycle, this protocol was selected to compare its capability with the approaches previously evaluated.

The most important result from the tests dealing with lyophilization of EBC as processing alternative was the highest number of compounds as compared to PP and SPE with C18 and HI sorbents. The Venn diagram in Fig. 4 shows that lyophilization allowed the exclusive detection of 24 compounds present at low levels in EBC, which had not been detected in any experiment based on PP or SPE. Apart from them, the efficiency of this step allowed detecting 23 compounds previously identified with the resting approaches, which represents practically the complete list of identifications, except for phosphoric acid and pyrophosphate, only detected after SPE with the HI sorbent. Some of the 24 compounds had been previously detected by NMR, as was the case of valine, leucine and glutamic acid. Other two key amino acids detected were phenylalanine and citrulline, together with amino acid derivatives such as N(6)-(octanoyl)lysine or glutamic acid n-butyl ester. One other important family of compounds detected were fatty acids,

identified as C8-C10 dicarboxylic acids and fatty acids with amino substituent. A long-chain dialdehyde (9,12-tetradecadienal), and a sphingenine (1-deoxy-tetradecasphing-13-enine) were also identified after a high preconcentration step. Table 1 lists the compounds tentatively identified in EBC after lyophilization. The preconcentration factor provided by lyophilization as compared with the other evaluated alternatives for sample preparation is visualized in the TICs in Fig. 5.

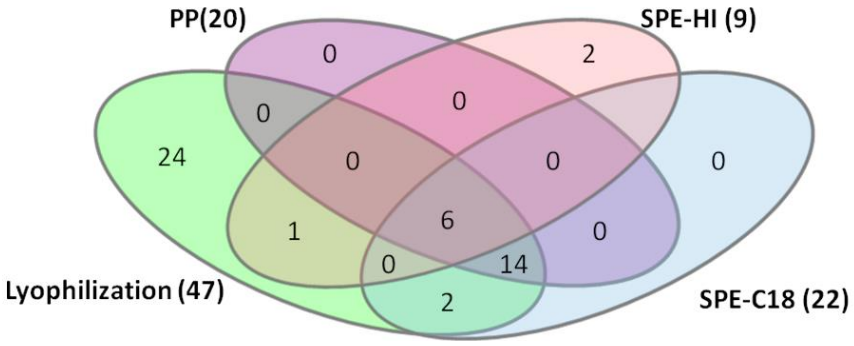


Fig. 4. Venn diagram for comparison of the compounds detected after PP, SPE-C18, SPE-HI and lyophilization.

3.4. Identification of potential entities detected by MS/MS

A total of 49 EBC compounds were tentatively identified by MS/MS in this research, including amino acids, fatty acids, fatty amides, fatty aldehydes, sphingoid bases, oxoanionic compounds imidazoles, hydroxy acids and aliphatic acyclic acids. Only compounds with high quality MS/MS spectra were considered in this step. Identification was supported on mass accuracy of precursor ions, considering mass accuracy error, expressed as ppm, from 0 to 5 ppm, and their most representative fragments by searching in the METLIN database and HMDB, mainly. Table 1 lists the tentatively identified metabolites, which are classified by families. Information on identification in terms of MS precursor ion and the most representative product ions is also included. Several of the compounds were detec-

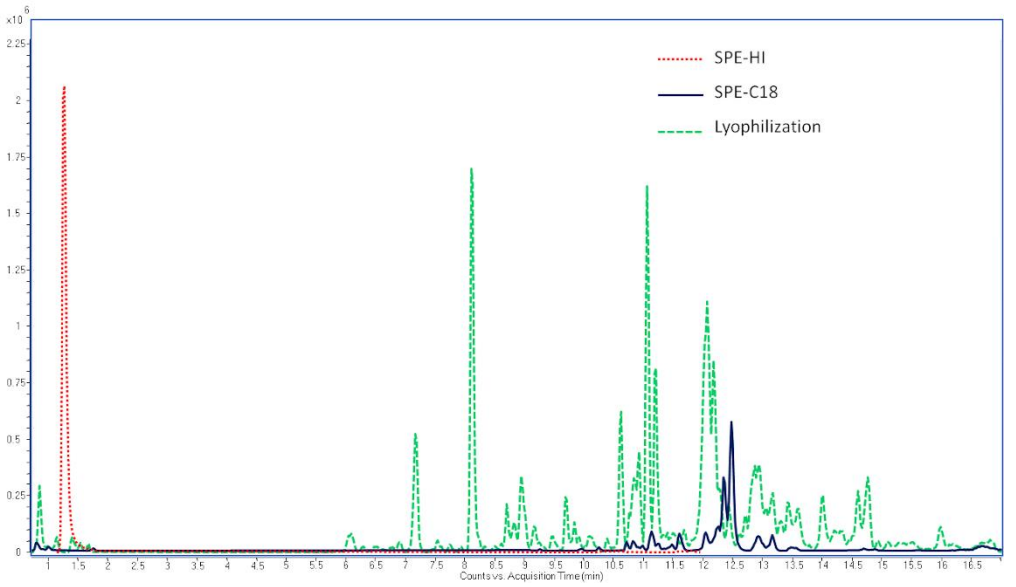


Fig. 5. TICs from samples prepared by SPE-HI (dotted red line), SPE-C18 (continuous blue line) and lyophilized EBC (discontinuous green line).

ted in both positive and negative ionization polarities, but only the ionization mode with the highest sensitivity is included.

Lyophilization provided the wider number of compounds, mainly amino acids and fatty acids, which are representative of crucial biochemical pathways. Thus, 13 amino acids and derivatives were identified in EBC, 11 of them exclusively in the lyophilized samples. Among them, 2,4-diaminopentanoate, precursor of ornithine, aspartic acid, citrulline and 1-pyrroline-4-hydroxy-2- carboxylate belong to the arginine metabolism and, therefore, to the urea cycle [31,32]. Furthermore, three essential amino acids (leucine, phenylalanine and valine) that the human body cannot synthesize at the concentration demanded for normal development and may be obtained from diet were detected in lyophilized EBC. Leucine and valine are also branched-chain amino acids (BCAA); that is, essential amino acids the carbon structure of which is characterized by a branch. These amino acids, critical to human life, are particularly involved in stress, energy and muscle metabolism [33]. Despite their structural similarities, the branched amino

acids have different metabolic pathways, with valine being metabolized only to carbohydrates, and leucine only to fats. One other amino acid detected in lyophilized EBC was glutamic acid, a key molecule in cellular metabolism [34]. It plays a central role both in transamination processes and in the synthesis of various amino acids (such as proline, hydroxyproline, ornithine and arginine) that need prior formation of this acid.

Additionally, it is worth mentioning the identification of 9 fatty acid derivatives, fatty aldehydes, fatty amides and 9 sphingoid bases with non-polar character, which demonstrate the presence of non-polar metabolites in EBC, despite the aqueous nature of this biofluid.

Three dicarboxylic acids (octanedioic acid, known as suberic acid; nonanedioic acid, also called azelaic acid, and decanedioic acid, known as sebacic acid) were detected in EBC. The levels of octanedioic acid and decanedioic acid were found elevated in the urine of patients with fatty acid oxidation disorders as medium-chain acyl-CoA dehydrogenase (MCAD) deficiency [35]. It is worth emphasizing the presence of nitrogen in most identified lipids: most of the fatty acids contain amino substituents. The same applies to sphinganines and sphingosines compounds. Furthermore, nitrogen is present through amide group in the fatty amides family. Other remarkable aspect of the identified metabolites was the presence of oxidized compounds such as, 2-amino-9,10-epoxy- 8-oxodecanoic acid, 8-amino-7-oxo-nonanoic acid, 6-hydroxy-2,4- hexadienal and 1-deoxy-tetradecasphing-13-enine. Several of the alkylamines identified in this research had already been detected in EBC at the ppb level as involved in the renal function [36].

Only the SPE-HI protocol extracted phosphoric acid and pyrophosphate. The former is a metabolic product of phospholipids, nucleic acids, phosphoproteins and phosphoglycerides, while pyrophosphate results from the hydrolysis of ATP to AMP in the cells and, hence, it is a product of a number of metabolic reactions [37].

The remaining families contain a shorter number of compounds. Nevertheless, it should be mentioned the presence of some metabolites of interest. For instance, lactic acid, known as component of EBC [38], is involved in glucose cell

metabolism, fundamental for tumor progression. In fact, many cancers obtain energy by a high rate glycolysis, resulting in lactic acid secretion even in the presence of oxygen (aerobic glycolysis) [39]. This marks a straightforward difference with normal cells, where the oxidative breakdown of pyruvate within the mitochondria is the prevalent source of energy.

Urocanic acid, a breakdown (deamination) product of histidine, had previously been identified in human saliva [40], and in the liver, where it is an intermediate in the conversion of histidine to glutamic acid. Urea, the principal end product of protein catabolism, is formed in the urea cycle, of which aspartic acid (also detected in EBC) is a constituent. A small amount of urea is also excreted in human sweat [41].

Among the 49 metabolites identified in this research using LC–QTOF, some of them (*viz.,* valine, lactic acid, leucine,12-amino-dodecanoic acid, alkylamines, urea) had already been identified in EBC, mainly by NMR [23–25,42]. In addition, some of the identified metabolites had been evaluated as potential biomarkers in other biofluids [43–45]; aspect that may be of interest to study EBC as potential biofluid for biomarker search.

4. Conclusions

The research carried out allows concluding that:

1. The matrix of EBC from healthy individuals does not cause interferences in analyses by LC–QTOF; therefore, deproteination or SPE steps for cleanup are unnecessary.

2. Despite EBC components are very diluted, 25 of the most concentrated can be determined after PP or SPE (the latter requiring to be applied in the centrifugal format because the low volumes of sample usually available).

3. Depending on the compounds to be analyzed, SPE by one of the studied sor-

bents could provide the required sample preparation in a time shorter than when lyophilization is involved.

4. Lyophilization is the most versatile of the assayed sample preparation strategies concerning preconcentration (after checking the absence of EBC matrix effects, drugs mainly, when the samples come from patients). In addition, this step can be used for sample storage; thus, the prepared sample being ready for analysis by simple thawing and reconstitution in the appropriate medium.

Acknowledgments

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 M.A. Fernández-Peralbo 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] 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.

[2] Y. Liang, S.M. Yeligar, L.A.S. Brown, Exhaled breath condensate: a promising source for biomarkers of lung disease, Sci. World J. 2012 (2012) 7, http://dx.doi.org/10.1100/2012/217518, Article ID 217518.

[3] P. Bajaj, F.T. Ishmael, Exhaled breath condensates as a source for biomarkers for characterization of inflammatory lung diseases, JASMI 3 (2013) 17–29.

[4] H.P. Chan, C. Lewis, P.S. Thomas, Exhaled breath analysis: novel approach for early detection of lung cancer, Lung Cancer 63 (2009) 164–168.

[5] H. Ahmadzai, S. Huang, R. Hettiarachchi, J.L. Lin, P.S. Thomas, Q. Zhang, Ex- haled breath condensate: a comprehensive update, Clin. Chem. Lab. Med. 51 (7) (2013) 1343–1361.

[6] P. Montuschi, D. Paris, S. Montella, D. Melck, V. Mirra, G. Santini, N. Mores, E. Montemitro, F. Majo, V. Lucidi, A. Bush, A. Motta, F. Santamaria, Nuclear magnetic resonance-based metabolomics discriminates primary ciliary dyskinesia from cystic fibrosis, Am. J. Respir. Crit. Care Med. 190 (2014) 229–233.

[7] P. Montuschi, Analysis of exhaled breath condensate in respiratory medicine: methodological aspects and potential clinical applications, Ther. Adv. Respir. Dis. 1 (1) (2007) 5–23.

[8] J. Hunt, M.D. Charlottesville, Exhaled breath condensate: an evolving tool for noninvasive evaluation of lung disease, JACI 110 (1) (2002) 28–34.

[9] A. Zacharasiewicz, N. Wilson, C. Lex, A. Li, M. Kemp, J. Donovan, J. Hooper, S. A. Kharitonov, A. Bush, Repeatability of sodium and chloride in exhaled breath condensates, Pediatr. Pulmonol. 37 (2004) 273–275.

[10] K.H. Kim, S.A. Jahan, E. Kabir, A review of breath analysis for diagnosis of human health, Trends Anal. Chem. 33 (2012) 1–8.

[11] P. Montuschi, G. Santini, S. Valente, C. Mondino, F. Macagno, P. Cattani, G. Zini, N. Mores, Liquid chromatography-mass spectrometry measurement ofleuko- trienes in asthma and other respiratory diseases, J. Chromatogr. B 964 (2014) 12–25.

[12] P. Montuschi, P.J. Barnes, G. Ciabattoni, Measurement of 8-isoprostane in exhaled breath condensate, Methods Mol. Biol. 594 (2010) 73–84.

[13] U. Sack, R. Scheibe, M. Wötzel, S. Hammerschmidt, H. Kuhn, F. Emmrich, G. Hoheisel, H. Wirtz, C. Gesnner, Multiplex analysis of cytokines in exhaled breath condensate, Cytom. A 69A (2006) 169–172.

[14] R.M. Effros, K.W. Hoagland, M. Bosbous, D. Castillo, B. Foss, M. Dunning, M. Gare, W. Lin, F. Sun, Dilution of respiratory solutes in exhaled condensates, Am. J. Respir. Crit. Care Med. 65 (2002) 663–669.

[15] M. Janicka, A. Kot-Wasik, J. Kot, J. Namieśnik, Isoprostanes-biomarkers of lipid peroxidation: their utility in evaluating oxidative stress and analysis, Int. J. Mol. Sci. 11 (2010) 4631–4659.

[16] D.C. White, R. Geyer, J. Cantu, S.C. Jo, A.D. Peacock, A.M. Saxton, S. Mani, M. Jett, O.R. Moss, Feasibility of assessment of regulatory lipids in breath condensate as potential presymptomatic harbingers of pulmonary pathobiology, J. Microbiol. Methods 62 (2005) 293–302.

[17] C.W. Hu, M.T. Wu, M.R. Chao, C.H. Pan, C.J. Wang, J.A. Swenberg, K.Y. Wu, Comparison of analyses of urinary 8-hydroxy-2′-deoxyguanosine by isotope-dilution liquid chromatography with electrospray tandem mass spectrometry and by enzyme-linked immunosorbent assay, Rapid Commun. Mass Spectrom. 18 (2004) 505–510.

[18] P. Rosias, Methodological aspects of exhaled breath condensate collection and analysis, J. Breath Res. 6 (2) (2012) 13 027102.

[19] K. Syslová, P. Kacer, M. Kuzma, P. Klusácková, Z. Fenclová, J. Pelclová, D. Lebedová, Determination of 8-iso-prostaglandin F(2alpha) in exhaled breath condensate using combination of immunoseparation and LC–ESI-MS/MS, J. Chromatogr. B 867 (2008) 8–14.

[20] C. Sachs-Olsen, M. Sanak, A.M. Lang, A. Gielicz, P. Mowinckel, K.C. Lødrup Carlsen, K.H. Carlsen, A. Szczeklik, Eoxins: a new inflammatory pathway in childhood asthma, J. Allergy Clin. Immunol. 126 (4) (2010) 859–867.

[21] S. Forcisi, F. Moritza, B. Kanawatia, D. Tziotisa, R. Lehmannc, P. Schmitt-Kopplin, Liquid chromatography–mass spectrometry in metabolomics research: mass analyzers in ultra high pressure liquid chromatography coupling, J. Chromatogr. A 1292 (2013) 51–65.

[22] P. Montuschi, LC/MS/MS analysis of leukotriene B4 and other eicosanoids in exhaled breath condensate for assessing lung inflammation, J. Chromatogr. B 877 (2009) 1272–1280.

[23] I. Bertini, C. Luchinat, M. Miniati, S. Monti, L. Tenori, Phenotyping COPD by 1H NMR metabolomics of exhaled breath condensate, Metabolomics 10 (2014) 302–311.

[24] G. De Laurentiis, D. Paris, D. Melck, M. Maniscalco, S. Marsico, G. Corso, A. Motta, M. Sofia, Metabonomic analysis of exhaled breath condensate in adults by nuclear magnetic resonance spectroscopy, Eur. Respir. J. 32 (5) (2008) 1175–1183.

[25] S. Carraro, S. Rezzi, F. Reniero, K. Héberger, G. Giordano, S. Zanconato, C. Guillou, E. Baraldi, Metabolomics applied to exhaled breath condensate in childhood asthma, Am. J. Respir. Crit. Care Med. 175 (10) (2007) 986–990.

[26] P. Montuschi, D. Paris, D. Melck, V. Lucidi, G. Ciabattoni, V. Raia, C. Calabrese, A. Bush, P.J. Barnes, A. Motta, NMR spectroscopy metabolomic profiling of exhaled breath condensate in patients with stable and unstable cystic fibrosis, Thorax 67 (3) (2012) 222–228.

[27] A. Motta, D. Paris, D. Melck, G. de Laurentiis, M. Maniscalco, M. Sofia, P. Montuschi, Nuclear magnetic resonance-based metabolomics of exhaled breath condensate: methodological aspects, Eur. Respir. J. 39 (2012) 498–500.

[28] S. Carraro, G. Giordano, F. Reniero, D. Carpi, M. Stocchero, P.J. Sterk, E. Baraldi, Asthma severity in childhood and metabolomic profiling of breath condensate, Allergy 68 (2013) 110–117.

[29] A.A. Aksenov, L. Yeates, A. Pasamontes, C. Siebe, Y. Zrodnikov, J. Simmons, M.M. McCartney, J.P. Deplanque, R.S. Wells, C.E. Davis, Metabolite content profiling of bottlenose dolphin exhaled breath, Anal. Chem. 86 (21) (2014) 10616–10624.

[30] M. Calderón-Santiago, F. Priego-Capote, M.D. Luque de Castro, Enhanced detection and identification in metabolomics by use of LC–MS/MS untargeted analysis in combination with gas-phase fractionation, Anal. Chem. 86 (15) (2014) 7558–7565.

[31] D. Mavri-Damelin, S. Eaton, L.H. Damelin, M. Rees, H.J. Hodgson, C. Selden, Ornithine transcarbamylase and arginase I deficiency are responsible for diminished urea cycle function in the human hepatoblastoma cell line HepG2, Int. J. Biochem. Cell. B 39 (2007) 555–564.

[32] J.L. Deignan, J.C. Livesay, P.K. Yoo, S.I. Goodman, W.E. O´Brien, R.K. Iyer, S. D. Cederbaum, W.W. Grody, Ornithine deficiency in the arginase double knockoutmouse, Mol. Genet. Metab. 89 (2006) 87–96.

[33] B.C. Batch, S.H. Shah, C.B. Newgard, C.B. Turer, C. Haynes, J.R. Bain, M. Muehlbauer, M.J. Patel, R.D. Stevens, L.J. Appel, L.K. Newby, L.P. Svetkey, Branched chain amino acids are novel biomarkers for discrimination of metabolic wellness, Metabolism 62 (7) (2013) 961–969.

[34] J.F. Harrison, M.L. Rinne, M.R. Kelley, N.M. Druzhyna, G.L. Wilson, S.P. Ledoux, Altering DNA base excision repair: use of nuclear and mitochondrial-targeted N-methylpurine DNA glycosylase to sensitize astroglia to chemotherapeutic agents, GLIA 55 (2007) 1416–1425.

[35] J.S. Frederick, J.S. Ronald, F.B. William, Liver Disease in Children, Lippincott Williams & Wilkins, Philadelphia, Pennsylvania, 2007.

[36] A. Mazzatenta, C. Di Giulio, M. Pokorski, Pathologies currently identified by exhaled biomarkers, Respir. Physiol. Neurobiol. 187 (2013) 128–134.

[37] D.L. Nelson, M.M. Cox, Lehninger Principles of Biochemistry, W. H. Freeman, New York, 2008.

[38] R.M. Effros, R. Casaburi, J. Porszasz, Assessment of exhaled breath condensate pH, Am. J. Respir. Crit. Care. Med. 183 (2011) 952–953.

[39] M.G. Vander Heiden, L.C. Cantley, C.B. Thompson, Understanding the Warburg effect: the metabolic requirements of cell proliferation, Science 324 (2009) 1029–1033.

[40] M. Calderón-Santiago, F. Priego-Capote, 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.

[41] C.R. Esther, G. Boysen, B.M. Olsen, L.B. Collins, A.J. Ghio, J.W. Swenberg, R. C. Boucher, Mass spectrometric analysis of biomarkers and dilution markers in exhaled breath condensate reveals elevated purines in asthma and cystic fibrosis, Am. J. Physiol. Lung Cell. Mol. Physiol. 296 (2009) 987–993.

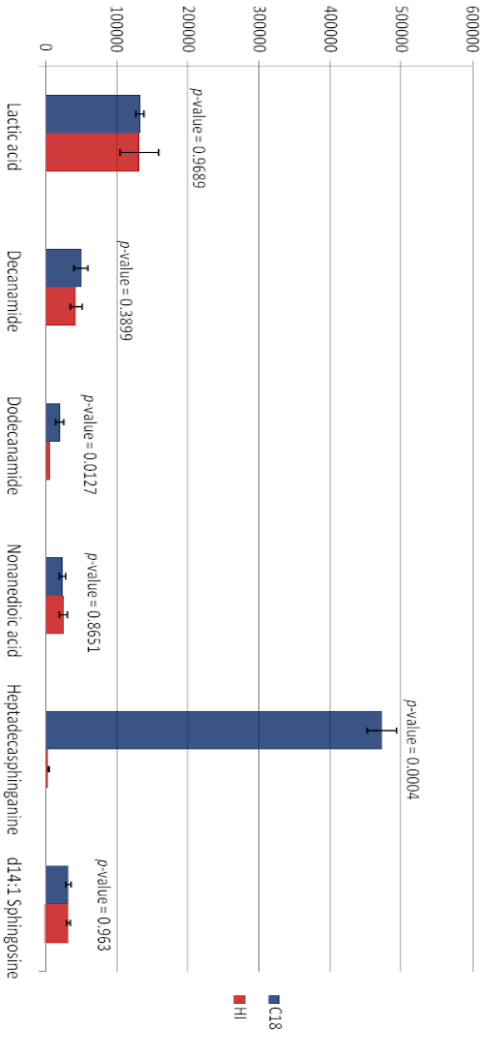
[42] C. Berchtold, L. Meier, R. Steinhoff, R. Zenobi, A new strategy based on real-time secondary electrospray ionization and high-resolution mass spectrometry to discriminate endogenous and exogenous compounds in exhaled breath, Metabolomics 10 (2014) 291–301.

[43] R. Madsen, T. Lundstedt, J. Trygg, Chemometrics in metabolomics —a review in human disease diagnosis, Anal. Chim. Acta 659 (2010) 23–33.

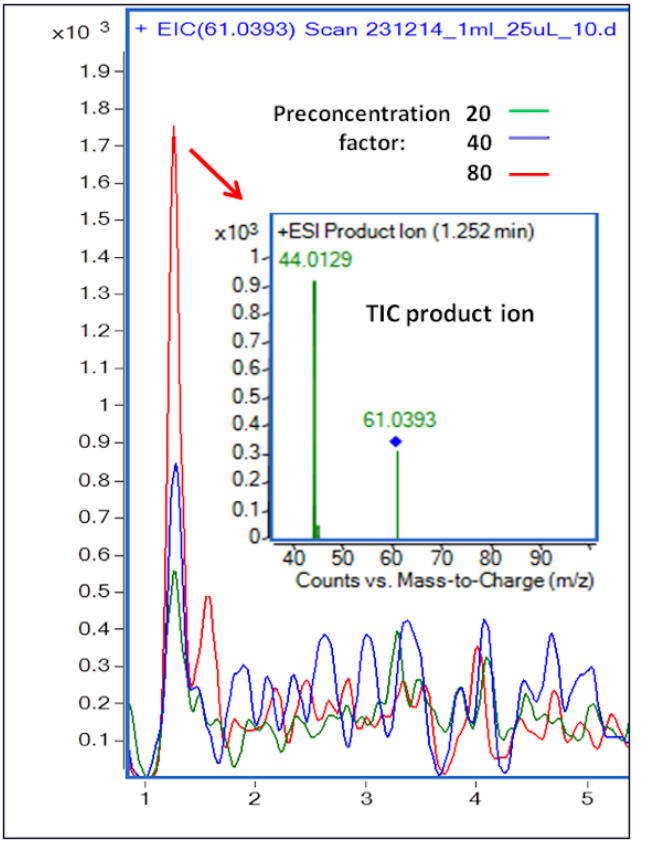
[44] Y. Li, A.G. Tang, S. Mu, HPLC–FLD determination of serum aromatic amino acids: application in chronic kidney disease patients, Clin. Chim. Acta 412 (11– 12) (2011) 1032–1035.

[45] Q. Wang, P. Gao, X. Wang, Y. Duan, The early diagnosis and monitoring of squamous cell carcinoma via saliva metabolomics, Sci. Rep. 4 (6802) (2014) 1–9.

Supplementary material



Supplementary Fig. 1. Histogram plotting the average area of the compounds common in the eluates from both sorbents by SPE.



Supplementary Fig. 2. EICs of urea from the different preconcentration factors provided by lyophilized EBC. TIC product ion from urea in 80-times preconcentrated EBC is also shown.

CHAPTER 8

Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC-QTOF: upstream information on other omics

Scientific reports 6:38243 (2016) 1–12

Scientific reports

6:38243 (2016) 1–12

Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC-QTOF: upstream information on other omics

M. A. Fernández-Peralbo[1,2], E. Gómez-Gómez[2,3], M. Calderón-Santiago[1,2], J. Carrasco-Valiente[2,3], J. Ruiz-García[2,3], M. J. Requena-Tapia[2,3], M. D. Luque de Castro[1,2], F. Priego-Capote[1,2]

[1]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain.*

[2]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía University Hospital, University of Córdoba, E-14004, Córdoba, Spain.*

[3]*Department of Urology (IMIBIC), Reina Sofia University Hospital, University of Córdoba, E-14004, Córdoba, Spain.*

Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC-QTOF: upstream information on other omics

M. A. Fernández-Peralbo, E. Gómez-Gómez, M. Calderón-Santiago, J. Carrasco-Valiente, J. Ruiz-García, M. J. Requena-Tapia, M. D. Luque de Castro, F. Priego-Capote

Abstract

The existing clinical biomarkers for prostate cancer (PCa) diagnosis are far from ideal (*e.g.*, the prostate specific antigen (PSA) serum level suffers from lack of specificity, providing frequent false positives leading to over-diagnosis). A key step in the search for minimum invasive tests to complement or replace PSA should be supported on the changes experienced by the biochemical pathways in PCa patients as compared to negative biopsy control individuals. In this research a comprehensive global analysis by LC–QTOF was applied to urine from 62 patients with a clinically significant PCa and 42 healthy individuals, both groups confirmed by biopsy. An unpaired *t*-test (p-value<0.05) provided 28 significant metabolites tentatively identified in urine, used to develop a partial least squares discriminant analysis (PLS-DA) model characterized by 88.4 and 92.9% of sensitivity and specificity, respectively. Among the 28 significant metabolites 27 were present at lower concentrations in PCa patients than in control individuals, while only one reported higher concentrations in PCa patients. The connection among the biochemical pathways in which they are involved (DNA methylation, epigenetic marks on histones and RNA cap methylation) could explain the concentration changes with PCa and supports, once again, the role of metabolomics in upstream processes.

1. Introduction

Despite prostate cancer (PCa) is a slow growing tumor mainly affecting old men, it is the second most frequently occurring cancer and the third leading cause of cancer death in men [1]. The main limiting factor in the diagnosis of PCa is the absence of symptoms at early stages. Despite the pathogenesis of PCa has not been completely elucidated, major risk factors associated to its development include age, race, and occurrence of PCa in the family line [2]. Likewise, lifestyle, diet or environmental factors can induce genetic modifications and initiate carcinogenesis processes [3].

To date few biomarkers have reached clinical use since the proposed candidates need to be specific for PCa to cause no alteration or expression in other human tissues or diseases. Urinary biomarkers can be broadly classified into DNA, RNA, protein and, recently, metabolite-based markers. Currently, prostate specific antigen (PSA) is the marker of choice for PCa screening. However, it is still not able to distinguish clinically relevant tumors from indolent ones and PSA use in screening is being downgrading. Various PSA-related derivatives, such as PSA velocity, PSA density, and free-to-total PSA ratio, have provided only limited improvements in terms of specificity [4]. RNA urine markers, the most clinically developed to date, include prostate cancer antigen 3 (PCA3), and a fusion of the 50-untranslated region of androgen-regulated transmembrane-serine protease gene (TMPRSS2) with virus E26 oncogene (ERG) (TMPRSS2:ERG) among the most used [5]. Recently, the Food and Drug Administration has approved PCA3 for its ability to predict PCa in patients with elevated PSA and a negative biopsy, but they are still considered at risk and may require a repeated biopsy. Furthermore, other studies have concluded that PCA3 has limited value in predicting aggressive cancers [6]. Similarly, although the TMPRSS2:ERG fusion was useful as a screening tool, its prognostic utility is controversial [7].

As stated above, genetic alterations caused by epigenetic factors can also play a significant role on PCa occurrence [8]. Increased methylation of cytosine at CpG islands of the glutathione S-transferase Pi 1 (GSTP1) promoter has shown to be a

very common epigenetic alteration in PCa, which has been well characterized in tissues and seems to be very promising for urine biomarker development [9]. Furthemore, aberrations in post-translational modifications (PTMs) of histones have shown to occur in PCa cells, in particular acetylation and methylation of specific lysine and arginine residues are being investigated as predictor of the risk of PCa recurrence [10]. Regarding to metabolite markers, sarcosine (*N*-methyl-glycine) is one of the main candidates found to be associated with PCa progression to metastasis, with a significant predictive value for PCa detection in urine samples [11]. However, several studies reveal that the role of sarcosine as urinary biomarker is still controversial and further validation studies are warranted [12]. Some metabolites related to genetic alterations in PCa have been found altered in biological samples from PCa patients. For instance, Jung *et al.* found 7-methyl-guanine up-regulated in malignant tissue from PCa individuals *versus* non-malignant tissue [13]; while both methylguanosine and dimethylguanosine appeared down-regulated in urine from PCa patients [14].

So far, the lack of conclusive studies makes urgently needed alternative screening tools based on no or minimum invasive tests applied to biological samples such as blood or urine; tools that allow discrimination of individuals with clinically significant PCa from those with no disease. Recent advances in analytical instrumentation have facilitated the present development of metabolomics. Since metabolite profiles closely reflect the total cellular state, metabolomics has been developed into a more integrative approach in comparison with other omics approaches.

Urine was chosen as biological material for development of the present research due to its non-invasiveness, easy collection and transportation, stability during short and long-term storage and, as the most important, the presence of end-products from cancer metabolism released either into the urine or carried within prostatic cells and directly released into the urethra through prostatic ducts. All these aspects suggest urine as an ideal biofluid for PCa studies; therefore, comprehensive analysis of urine metabolome by LC–MS/MS in high-resolution mode was applied to search for discrimination between patients with clinically

significant PCa and healthy individuals, both groups confirmed after biopsy. The set of significant metabolites tentatively identified was used to study the metabolic pathways potentially involved in PCa.

2. Methods

2.1. Ethical statement

The selected individuals were part of a cohort of patients scheduled for prostate biopsy according to the European Association of Urology (EAU) guidelines [15]. All members of the cohort signed the informed consent before involvement in the project. The experiment was planned following the guidelines dictated by the World Medical Association Declaration of Helsinki (2004), and was supervised by the Reina Sofia Hospital Biomedical Research Ethical Committee and University of Cordoba Biomedical Experimentation Ethical Committee that approved the experiments.

2.2. Subjects

Morning urine samples were collected in the Urology Department of the Reina Sofia Hospital from 104 patients on the same day of prostate biopsy before the intervention. First morning urine was discarded for development of this research. All individuals were treated with the same dose of ciprofloxacin the day before sampling, which was carried out after overnight fasting of at least for 12 h. The samples were aliquoted in 1.5 mL tubes and stored at −80 °C until analysis. The individuals were grouped in two categories, PCa patients and negative biopsy controls, as Table 1 shows. The PCa category included 62 patients with clinically significant PCa confirmed by prostate biopsy. Clinically significant PCa had to meet at least one of these criteria: clinical stage>T1c and/or Gleason pattern >6 and/or more than 3 positive cores and/or >50% cancer per core. On the other hand, no cancer was detected by histological analysis of the control group, which included 42 individuals. 36 out of 62 (58%) PCa patients showed a positive digital rectal examination (DRE) in contrast to negative biopsy controls who gave negative

DRE. The mean age of patients with PCa was 71 *versus* 62 years of the control group. Additionally, no significant differences were found in terms of prostate volume between the two groups, as well as in the number of biopsies.

Two PSA measurements were carried out to each patient: one before the biopsy, as a decision tool to undergone prostate biopsy, and a second one on the same day of the biopsy. Logically, as shown in Table 1, PCa patients provided higher PSA in both measures. In the first test, all PCa individuals gave values above 3 ng/mL, with 27 cases between 3 and 10 ng/mL and 35 patients above 10 ng/mL. These values were quite similar in the second PSA test.

On the other hand, the negative biopsy control included 39 patients (92.9%) with serum PSA above 3 ng/mL. However, this group was clearly dominated by individuals with PSA levels between 3 and 10 ng/mL. A clear change was observed in the second PSA test since most of the control group individuals gave serum PSA levels below 3 ng/mL in the measurement of the day of biopsy.

Table 1. Characteristics of the cohort under study. PSA (1): Serum PSA before the biopsy; PSA (2): Serum PSA on the day of biopsy.

Variable	PCa (n = 62)	Negative biopsy (n = 42)	*p*-value
Age (years)	71 (64–80)	62.5 (57.8–69)	<0.05
Positive DRE	36 (58.1%)	0	<0.05
*Serum PSA (1) (ng/mL)	11 (6.9–29.75)	4.1 (3.6–5.8)	<0.05
<3	0	3 (7.1%)	
3–10	27 (43.5%)	35 (83.3%)	
>10	35 (56.5%)	4 (9.6%)	
Prostate volumen (mL)	30 (22–40)	27.5 (24–47)	>0.05
N° biopsy (1°)	52 (83.9%)	38 (90.5%)	>0.05
*Serum PSA (2) (ng/mL)	11 (6.9–40.7)	2.3 (1.8–2.7)	<0.05
<3	0	41 (97.6%)	
3–10	28 (45.2%)	1 (2.4%)	
>10	34 (54.8%)	0	
Number of cylinders	12 (12–13)	12 (12–12)	>0.05
Gleason pattern			
6	4 (6.5%)		
7	26 (41.9%)		
≥8	32 (51.6%)		

*Data are expressed as median value (range of variability) (Q1-Q3).

2.3. Reagents

Acetonitrile (ACN) (LC–MS optimum grade) from Fisher Scientific (Fair Lawn, NJ, USA) and deionized water from a Millipore Milli-Q water purification system (Millipore, Bedford, MA, USA) were used to prepare the chromatographic mobile phases. LC–MS grade formic acid from Scharlab (Barcelona, Spain) was used as ionization agent in LC–MS/MS analysis. Ammonium formate from Fluka (Sigma–Aldrich, St. Louis, USA) was used to adjust the pH of urine.

2.4. Apparatus and instruments

Centrifugation was carried out by a Thermo Sorvall Legend Micro 21R thermostated centrifuge (Thermo Fisher Scientific, Bremen, Germany). The urine samples were analyzed by an Agilent 1200 Series LC system hyphenated to an Agilent 6540 QTOF with a dual electrospray ionization (ESI) source (Agilent Technologies, Santa Clara, CA, USA). Agilent MassHunter Workstation (version B05.01) was the software for data acquisition.

2.5. Sample preparation

After thawing at room temperature, urine samples were vortex-mixed for 1 min and centrifuged at 21000 × *g* for 5 min. Then, 50 µ L of the supernatant were 1:2 (v/v) diluted with 5 mM ammonium formate in water (pH 5.5–7.5) prior to LC–MS/MS analysis.

2.6. LC–QTOF analysis

Chromatographic separation was performed by using a Mediterranea Sea C18 analytical column (50 × 4.6 mm i.d., 3 µm particle size) from Teknokroma (Barcelona, Spain), thermostated at 25 °C. The initial mobile phase was a mixture of 98% phase A (0.1% formic acid in water) and 2% phase B (0.1% formic acid in ACN). After injection, the initial mobile phase was kept under isocratic conditions for 1 min; then, a linear gradient of phase B from 2% to 100% was applied within 16 min. The flow rate was 0.6 mL/min during the chromatographic step. The total analysis time was 17 min, and 5 min were required to re-establish the initial conditions. The injected volume was 5 µL in each ionization mode and the samples

were randomly run. The injector needle was washed 5 times with 20:80 ACN–water between injections and the needle seat back was flushed with 20:80 ACN–water at 4 mL/min for 10s to avoid cross contamination. The autosampler was kept at 4 °C to increase sample stability. Urine samples from several healthy volunteers were collected, mixed to obtain a pool, aliquoted and stored as quality control (QC) samples to be daily analyzed at the beginning of the sequence and after every 8 injections to ascertain that the mass spectrometer performance was stable during the analysis of the samples set. Thus, 2 QCs were run per day following the same protocol as that used for samples.

The settings of the electrospray ionization source, operated in the negative and positive ionization modes, were as follows: capillary voltage ± 3.5 kV, fragmentor voltage 130 V, N_2 pressure in the nebulizer 40psi, N_2 flow rate and temperature as drying gas 12 L/min and 325 °C, respectively. The instrument was calibrated and tuned according to procedures recommended by the manufacturer. Each sample was analyzed in triplicate by setting data acquisition in centroid mode at 2 spectra per second in MS mode and 1 spectrum per second in MS^2 mode in the extended dynamic range mode (2 GHz). A gas-phase fractionation approach was used to increase the identification capability [16]. For this purpose, each of the three replicates from each sample was acquired with different ranges of precursor selection for MS/MS detection. In all cases, the acquisition range in MS scanning mode was from 60 to 1200 *m/z*. The maximum number of precursors selected per cycle for MS/MS fragmentation was set at 2, with an exclusion window of 0.25 min after two consecutive selections of the same precursor ion. Criteria for selection of precursor ions were established for each of the three runs per sample to obtain the maximum spectral information. Thus, in both ionization modes the mass range for precursor selection in the first run was set at 60–250 *m/z* with 10 eV as collision energy. The second and third runs involved precursor selection in the range 250–450 and 450–1200 *m/z* with collision energies of 20 and 30 eV, respectively. The instrument gave typical resolution 15000 FWHM (Full Width at Half Maximum) at *m/z* 118.0863 and 30000 FWHM at *m/z* 922.0098. To assure the desired mass accuracy of the recorded ions, continuous internal calibration was performed

during analyses by using signals at *m/z* 121.0508 (protonated purine) and *m/z* 922.0097 [protonated hexakis (1 H, 1 H, 3H-tetrafluoropropoxy) phosphazine or HP-921] in positive ionization mode. In negative ionization mode, ions with m/z 119.0363 (purine anion) and *m/z* 1033.9881 (adduct of HP-921) were used.

2.7. Data processing and statistical analysis

Raw data files were converted to *mzData* files using Mass Hunter Workstation software (version B.07.00 Qualitative Analysis, Agilent Technologies, Santa Clara, CA, USA); then, the data from each polarity were separately processed in R statistical language (version 3.1.3, http:// www.r-project.org/) using the open-free XCMS (version 1.44.0)[17] and CAMERA (version1.24.1) R-packages. The XCMS package was used for processing mzData files by extracting potential molecular features (MFs) considering only ions exceeding 1500 counts peak height with a peak width between 10 and 60s, a signal-to-noise threshold of 10 and 10 ppm of error in mass accuracy. Only molecular entities detected in at least 75% of the samples belonging to at least one of the groups under study (PCa and control individuals) were considered. Then, alignment of retention times across samples, grouping and integration were executed using the obiwarp method from the "retcor" function. In the next step, the CAMERA package was used to correlate potential precursor ions as [M−H]−, [M + Cl]− and [M + HCOO]− in negative ionization mode, and as [M + H]+ and [M + Na]+ in positive ionization mode, as well as isotopic peaks from the same molecular entity. Ions with identical elution profiles and related *m/z* values (representing different adducts or isotopes of the same compound) were grouped as a unique feature to remove redundant information. This process resulted in a data set for each polarity mode containing the peak area values in the apex of chromatographic peaks for all molecular entities characterized by accurate mass and retention time (RT). As the samples were injected in triplicate, the average area of the three injections was calculated for each potential molecular feature.

Normalization of the data set in each polarity was based on the mass spectrometry total useful signal (MSTUS) method that attempts to limit the

contributions of xenobiotics and endogenous substances to the normalization factor by including only peaks present in all samples [18]. Thus, this normalization takes into account the differences in volume of urine sampled per individual since metabolites would be more concentrated and give higher signal when the volume of collected urine is lower.

The data matrix from each polarity was introduced as *.csv* format into the Mass Profiler Professional (MPP) software (version 13.1, Agilent Technologies, Santa Clara, CA, USA) for statistical analysis. A supervised analysis of the data set was done by partial least squares discriminant analysis (PLS-DA). Additionally, an unpaired *t*-test with Benjamini-Hochberg false discovery rate (FDR) was executed to find metabolites significantly different between the two groups of patients, considering a *p*-value threshold of 0.05. Identification of the most relevant entities was supported on MS and MS/MS information by searching in the METLIN MS and MS/MS database (http://metlin.scripps.edu) and the Human Metabolome Database (HMDB, v. 3.6). An accuracy error of 5 ppm was set both in MS and MS/MS data to confirm the tentative identification of metabolites.

Normalized data were uploaded at MetaboAnalyst 3.0, a web-tool based on analytical high-throughput metabolomics studies (http://www.metaboanalyst.ca) for metabolomic data analysis and interpretation.

3. Results and discussion

3.1. Data pretreatment and exploratory analysis

Normalization of urine is a critical step in metabolomics analysis owing to variability of volume excretion among individuals. Under normal conditions, urinary creatinine excretion is relatively constant and easily measurable; therefore, it is a common practice to normalize urinary metabolite levels to the response of this endogenous metabolite. However, creatinine production may vary by an external stressor such as kidney impairment [19]. For this reason, alternatives such as the so-called “MSTUS” normalization method are gaining popularity. This

method has been previously compared to other common normalization strategies such as urine volume, osmolarity and creatinine concentration, concluding that it reduces the external variability among biological replicates [18]. For this reason, MSTUS was selected for normalization of urine samples. Thus, the peak area of each detected compound was divided by the sum of peak areas of all MFs detected in each sample. The MSTUS approach was preliminarily applied to a data set generated by analysis of QC samples injected along the batch for estimation of the methodological variability in the detection of MFs. As a result, the within-laboratory reproducibility, expressed as relative standard deviation (RSD), ranged from 3.15 to 16.9%. These values indicated that the method was reproducible enough for metabolomics research and no QCs correction within days was required.

The number of MFs extracted by the XCMS R-package from the data obtained after analysis of the cohort in negative and positive ionization modes was 1127 and 781, respectively. These results were obtained after removing redundant information from adducts and isotopes provided by the CAMERA R-package, and considering only MFs present in at least 75% of the samples pertaining to one of the groups (healthy individuals or PCa patients).

3.2. Tentative identification of significant entities detected by MS/MS

An unpaired *t*-test was applied to identify statistically significant differences ($p< 0.05$), in terms of concentration, between the two groups under comparison, using the Benjamini-Hochberg FDR to avoid false positive. The statistical test reported 111 and 75 statistically significant entities in the data sets obtained in positive and negative ionization modes, respectively. Among them, 28 metabolites were tentatively identified by MS/MS in this research according to the requirements previously established. The rest of significant MFs were not identified owing to either bad quality of the MS/MS spectra or lack of information included in most databases; therefore, only compounds with high quality MS/MS spectra were considered in this step. Table 2 lists the tentatively identified metabolites, with special emphasis on amino acids and derivatives, which represents 78.6% of all

them. The list also included sulfonic acids (10.7%) and nitrogenous bases derivatives (10.7%). Moreover, it should be mentioned that 28.6% of the metabolites were acetylated compounds such as acetylcitrulline, acetylarginine, acetylputrescine or acetyllysine, while 25% of them represented methylated amino acids or nitrogenous bases as methylguanosine, methyllysine, methylhistidine or dimethylarginine. Information on identification in terms of MS precursor ion, the most representative product ions, RT, formula, *p*-value, corrected *p*-value and regulation in PCa compared to control individuals is also included. Several of these metabolites were detected in both positive and negative ionization polarities, but only the ionization mode providing the highest sensitivity was considered.

Supplementary Table 1 shows the reproducibility values for each identified metabolite, expressed as RSD. The values have been calculated prior to and after MSTUS normalization in order to evaluate the correction effect. As can be seen, RSD was reduced in all cases after application of the MSTUS normalization strategy.

3.3. Partial Least Squares Discriminant Analysis (PLS–DA)

As commented above, one of the main limitations concerning PCa diagnosis is the stage at which the disease is typically detected and the characteristics of the existing available tests. Therefore, new tests with low false negative and positive rates are demanded for preliminary PCa screening. In this study, a discrimination model based on PLS-DA was built using a new data set exclusively constituted by the tentatively identified significant metabolites. The resultant PLS-DA model showed a clear separation between patients with PCa and control individuals, as the scores plot in Fig. 1 illustrates.

The combination of principal component 1 (PC1) and PC2 provides the greatest discrimination between the two classes: PCa and healthy individuals, which are separated through the diagonal between components in Fig. 1. The values of sensitivity and specificity obtained in the training model considering 70% of the samples randomly selected were 88.4% and 92.9%, respectively. Then, the stability of the model was assessed by external validation using 30% of the samples

excluded in the training step. Validation reported sensitivity and specificity of 63.2% and 78.6%, respectively. Table 3 summarizes this information and includes the percentage of positive and negative predictive values.

Since the combination of PC1 and PC2 led to the best discrimination between the two groups, these principal components were selected to evaluate the loadings plot corresponding to the PLS-DA, shown in Fig. 2.

Thus, through PC1, which explains 28.1% of the variability, several metabolites such as urea and 7-methylguanine contributed to the clustering of the PCa patients; while amino acids such as tyrosine, citrulline and histidine together with acetylated and methylated amino acids such as acetyl-lysine, acetylhistidine, dimethyl-lysine and trimethyl-lysine were relevant to discriminate the control group of individuals with negative biopsy. On the other hand, and through PC2, imidazole lactate, acetylputrescine and dimethylarginine contributed to characterize the healthy group, while a heterogeneous group of metabolites including 5-methyldeoxycytidine-5-phosphate, 7-methylguanosine, acetylcitrulline, acetylaspartatylglutamic acid and acetyltaurine supported the differentiation of the PCa patients group.

3.4. Biological interpretation according to the obtained results

After identification of metabolites found at significantly different concentrations between control and PCa individuals, they were studied and classified according to their connection across the biochemical pathways in which they are involved, and grouping them according to their involvement in the pathways. As can be seen in Table 2, 27 significant metabolites were present at lower concentration in PCa patients than in control individuals, while only one reported higher concentration in PCa patients. All these metabolites have been grouped in the table according to the metabolic pathway in which they are preferentially involved. As can be seen, the metabolism of amino acids was predominantly altered by the occurrence of prostate tumor. In fact, the concentration of 3 amino acids and 19 related metabolites, with special emphasis on acetylated and methylated amino acids, were altered in PCa individuals. The main pathways involving

Table 2. Metabolites tentatively identified by MS/MS with significant different levels in urine from PCa versus that of negative biopsy patients. The compounds are grouped by the metabolism pathways in which they are preferentially involved and sorted by their fold change. The fold change value ranges are indicated on the table by the following legend: ***fold change values from 1.5 to 1.75. **fold change values from 1.25 to 1.5. *fold change values from 1.1 to 1.25.

Pathway	Compound name	RT (min)	*m/z*	Adduct	Formula	Fragments	*p*-Value	*p*-Value (corrected)	Fold change (Controls vs. PCa)	Regulation in cancer
Lysine degradation	Dimethyllysine***	1.05	175.14409	M+H	C8H18N2O2	84.0815/130.0878	1.01E-06	1.70E-04	1.74	↓
	5-acetamidovalerate***	1.73	160.0968	M+H	C7H13NO3	142.0862/70.0655/98.0623	1.19E-03	2.56E-02	1.51	↓
	Acetyllysine**	1.48	189.1230	M+H	C8H16N2O3	84.0811/126.0914	6.19E-05	2.53E-03	1.33	↓
	Trimethyllysine**	1.05	189.16005	M+H	C9H20N2O2	130.0861/84.0805/60.0806	8.22E-04	1.31E-02	1.28	↓
Histidine metabolism	Imidazole lactate**	1.18	157.0607	M+H	C6H8N2O3	111.0550/83.0610	5.64E-06	3.89E-04	1.46	↓
	Histidine**	1.05	156.07684	M+H	C6H9N3O2	110.0718/83.0625/95.0600	1.69E-06	6.02E-03	1.43	↓
	Methylhistidine**	1.06	170.09247	M+H	C7H11N3O2	124.0864/109.0756/96.0679/ 83.0600/68.0493	1.64E-03	2.33E-02	1.40	↓
	Acetylhistidine**	1.28	198.0874	M+H	C8H11N3O3	152.0817/110.0717/180.3059	5.84E-03	4.24E-02	1.26	↓
Arginine metabolism	Urea**	1.24	61.0399	M+H	CH4N2O	44.0133	8.39E-03	1.89E-02	1.37	↓
	Acetylarginine**	1.67	215.11621	M-H	C8H16N4O3	173.1032/129.1063	2.35E-03	2.58E-02	1.34	↓
	Acetylcitrulline**	2.52	216.0990	M-H	C8H15N3O4	173.0931/131.0821	3.35E-02	3.75E-02	1.29	↓
	Acetylputrescine**	1.25	131.1178	M+H	C6H14N2O	72.0809/114.0907	5.28E-03	4.19E-02	1.30	↓
	Dimethylarginine**	1.17	203.15036	M+H	C8H18N4O2	70.0653/116.0698/158.1285 /88.0874	3.89E-03	3.58E-02	1.25	↓
	Citrulline*	1.34	176.10316	M+H	C6H13N3O3	159.0764/130.0975/70.0651	8.09E-03	4.90E-02	1.23	↓
Tyrosine metabolism	Tyrosine**	2.96	180.0660	M-H	C9H11NO3	119.0503/163.0404/136.075 9/93.0318	2.44E-03	2.58E-02	1.34	↓
Tryptophan metabolism	8-methoxykynurenate**	7.79	220.0610	M+H	C11H9NO4	174.0554/116.9734/202.0484	1.39E-02	2.27E-02	1.31	↓
	Kynurenic acid**	7.00	188.0352	M-H	C10H7NO3	144.0454	6.21E-04	1.31E-02	1.30	↓
	Xanthurenic acid*	6.80	206.0450	M+H	C10H7NO4	188.0345/160.0386/132.0459	9.79E-03	1.92E-02	1.22	↓

Cont. Table 2

Taurine metabolism	Sulfoacetate**	1.25	138.9708	M-H	C2H4O5S	94.9810/79.9575	1.56E-03	2.26E-02	1.28	↓
	Isethionate*	1.23	124.9916	M-H	C2H6O4S	79.9572/94.985/106.9797	6.71E-03	4.24E-02	1.19	↓
	Acetyltaurine*	1.56	166.0194	M-H	C4H9NO4S	79.9579/124.0078	1.63E-02	2.27E-02	1.12	↓
Alanine, aspartate and glutamate metabolism	Acetylaspartylglutamic acid*	2.81	303.0833	M-H	C11H16N2O8	96.0077/128.0350	5.21E-03	4.19E-02	1.24	↓
	Acetylaspartate*	1.86	174.0410	M-H	C6H9NO5	88.0403/130.0507/115.0037 /58.0296	5.96E-03	4.19E-02	1.19	↓
	2-oxoglutaramate*	1.49	144.03145	M-H	C5H7NO4	126.0202/100.0409/82.0305/ 72.0454/41.9983	9.85E-04	1.31E-02	1.18	↓
Glutamine and glutamate metabolism	2-pyrrolidone-5-carboxylate**	2.10	128.03542	M-H	C5H7NO3	85.0295/41.0399/110.8772	8.21E-03	4.90E-02	1.25	↓
Purine and pyrimidine metabolism	5-methyldeoxycytidine-5′ -phosphate*	2.81	320.0620	M-H	C11H15NO10	110.0245/240.0523	1.88E-02	2.37E-02	1.21	↓
	7-methylguanosine*	5.31	296.100	M-H	C11H15N5O5	164.0574	1.97E-02	2.37E-02	1.15	↓
	7-methylguanine**	2.10	164.0580	M + H	C6H7N5O	149.0497/124.0501	2.64E-02	3.19E-02	−1.30	↑

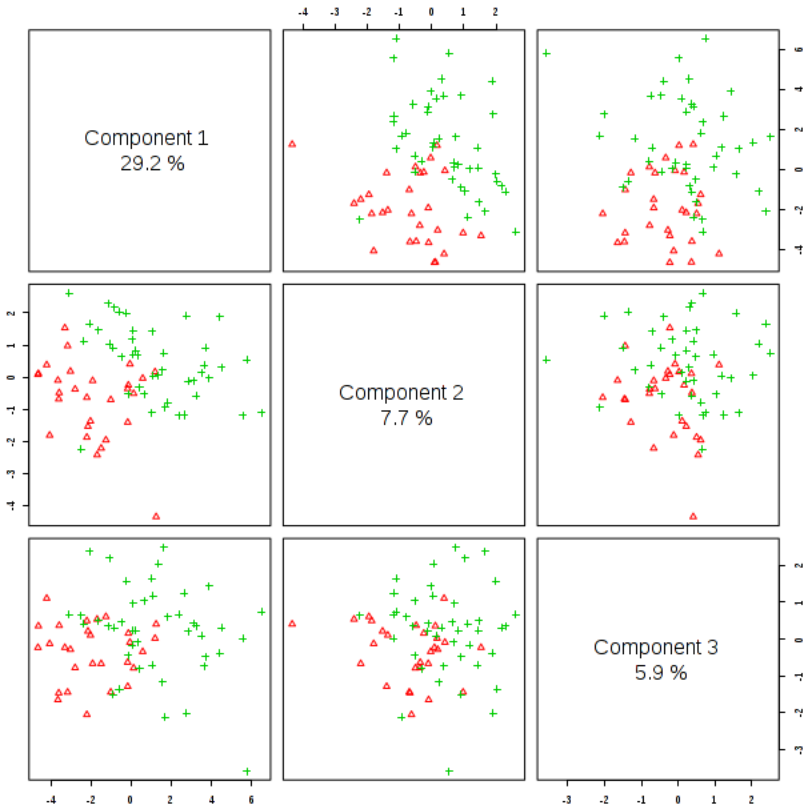


Fig. 1. PLS-DA scores plot for discrimination of PCa and negative biopsy patients by urine metabolomics analysis. PLS models were built with 70% of samples randomly selected from each group.

Table 3. Discrimination capability for the target PLS-DA model.

Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
Training with 70% of the samples			
88.4	92.9	95.0	83.9
External validation with 30% of the samples			
63.2	78.6	80	61.1

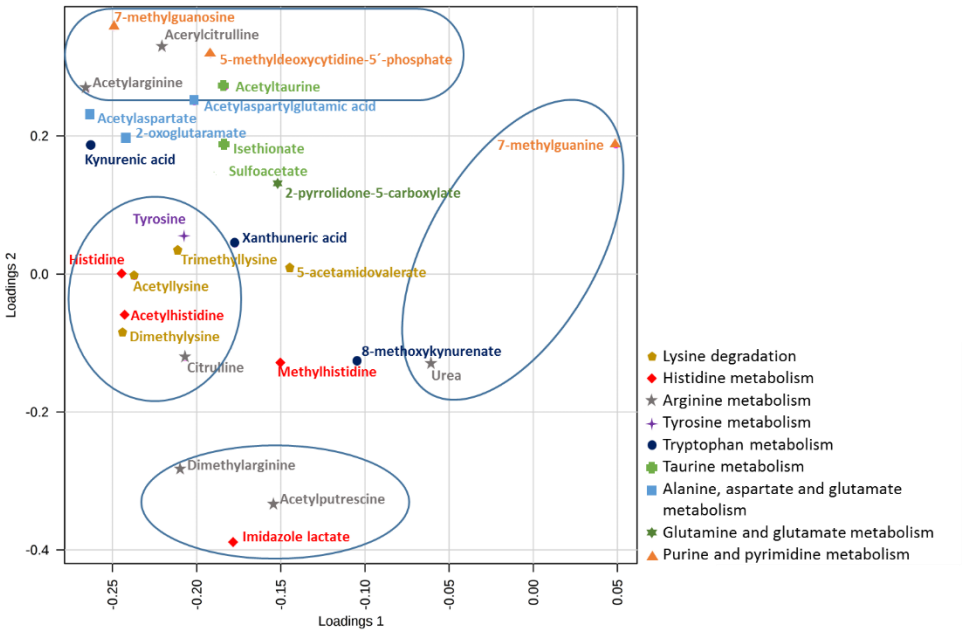


Fig. 2. PLS-DA loadings plot to discriminate patterns of PCa and negative biopsy individuals.

metabolism of amino acids altered by prostate tumor occurrence affected lysine, histidine, arginine, alanine, aspartate and glutamate, as well as aromatic amino acids, particularly tryptophan and tyrosine.

On the other hand, 3 purines and pyrimidines derivatives were also altered in cancer. Concretely, one of them (7-methylguanine) was the exclusive significant metabolite down regulated in PCa, as Table 2 shows. According to the list of significant metabolites, three biological processes can be mainly emphasized to be discussed in this research: DNA methylation, epigenetic marks on histones and RNA cap methylation. The three biological processes are independently discussed to explain the potential reason for the metabolic discrimination detected in human urine from PCa patients and negative biopsy individuals.

Methylated and acetylated amino acids related to histone protein modifications. Histone modifications together with DNA methylation are well-recognized epigenetic mechanisms of gene transcriptional regulation and play essential roles in PCa tumor initiation and progression. The N-terminal tails of histones, peripherally positioned around the nucleosome core, are subjected to various covalent post-translational modifications (PTMs) such as acetylation, methylation, phosphorylation, citrullination and ubiquitination, among others, by specific chromatin-modifying enzymes. The pattern of these modifications has been referred to as the "histone code" and determines whether the chromatin will adopt a transcriptionally active or inactive state [20]. Acetylation and methylation are widely considered the two most important PTMs occurring in histones. Acetylation induces an open chromatin conformation to regulate the accessibility of transcription factors by blocking the normal electrostatic interaction between positively charged N-terminal basic residues such as lysine, arginine and histidine located at tails of the histone, and negatively charged DNA phosphate groups. Acetylation has shown to alter the structural core of nucleosomes and chromatin, enhancing gene expression as well as DNA reparation and cytokine-activated signal transduction [21,22]. The level of acetylation is balanced by two classes of enzymes: histone acetyl transferases (HATs) and histone deacetylases (HDACs). Aberrant activities of these enzymes such as excessive recruitment of histone deacetylases and mutation in acetyltransferases have been implicated in silencing of key oncosuppressor genes identified as an early event of oncogene activation to cancer pathways [23].

Methylation is based on the same biochemical mechanism and, therefore, also contribute to modify chromatin conformation. Histone methylation is characterized by a dynamic mark in health and inheritance. By analogy to acetylation, the enzymes involved in these reversible marks include histone methyltransferases (HMTs) and histone demethylases (HDMs). However, in contrast with acetylation, methylation of specific lysine or arginine residues can result in transcriptional activation or repression for gene silencing. Additionally, combined methylation marks can present different roles to the same marks appearing in isolation.

Inappropriate targeting of histone modifying enzymes is often responsible for aberrant histone modifications. Thus, abnormal modifications in the profile of histone acetylation and methylation have been found in PCa patients, which has opened a new group of promising biomarkers with diagnostic potential to predict disease severity. As previously emphasized, these epigenetic alterations are interestingly reversible by enzymatic action. Thus, identifying key epigenetic pathways in cancer cells might pave the way to innovative therapeutic approaches [24].

In the present research, the most interesting result came from the levels of acetylated and methylated basic amino acid residues by application of a metabolomics workflow for urine analysis. Particularly, acetylated and methylated lysine, arginine and histidine residues were among the most statistically significant metabolites contributing to explain the differences in urine composition between PCa patients and healthy individuals. Acetyl-lysine, acetylarginine and acetyl-histidine were found at high-significantly lower concentration in urine from PCa patients. Fig. 3 shows the box-and-whisker plots for these three acetylated basic residues in the two evaluated groups.

The fold change values considering control individuals versus PCa patients were 1.33, 1.34 and 1.26 for acetylated lysine, arginine and histidine, respectively. This result strongly agrees with expression levels of histone acetylases (HACs), HDACs and HATs in PCa patients. Thus, HDACs are often overexpressed in prostate cancer associated to reduced histone acetylation [25], responsible for tumor suppressor gene silencing and malignant transformation. Sato *et al.* proved that the inhibition of histone deacetylase efficiently suppressed cell growth of three prostate cancer lines together with down-regulation of the androgen receptor, regardless of their hormone sensitivity, based on miRNA-mediated suppression [26]. Citrullination is an additional PTM occurring in histones, which involves the enzymatic conversion (by peptidylarginine deiminase 4, PADI4) of peptidyl arginine to citrulline to neutralize the positive charge of the former. Different studies have pointed out a transcriptional repression role to citrullination through connection to deacetylation [27], although this PTM also seems to facilitate gene

expression in early embryos by creating a platform for HAT assembly, thus leading to the enhancement of histone acetylation [28]. At the metabolite level, the present study reports that both citrulline and acetylcitrulline were significantly at higher concentrations in healthy individuals as compared to PCa patients. Therefore, both

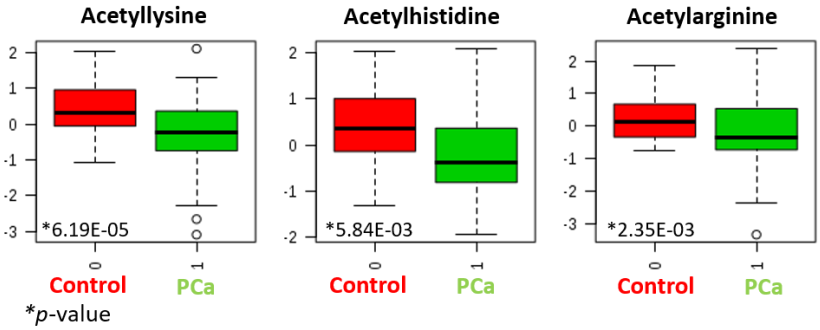


Fig. 3. Box and whisker plots from the three acetylated basic residues –lysine, arginine and histidine– in the two evaluated groups.

citrullination and acetylation of citrulline reported a behavior similar to that experienced by acetylated arginine. Citrullination could also contribute to explain the metabolic differences in urine from PCa patients as compared with that from negative biopsy individuals. Supplementary Fig. 1 shows the box and whisker plots comparing the concentration levels of citrulline and acetylcitrulline found in urine from PCa patients *versus* that from healthy individuals.

Similar to acetylation, several evidences of histone methylation marks were detected as significant metabolites by comparing PCa and healthy individuals. Thus, dimethyl and trimethyllysine, dimethylarginine and methylhistidine were in the list of significant metabolites. Among them, it is worth mentioning the role of dimethyllysine, the most significant compound in terms of *p*-value, that also experienced the highest concentration change when comparing controls and PCa patients (in fact, the fold change for this metabolite was 1.74). Fig. 4 shows the box and whisker plots for methylated residues that could aid to interpret the role of histone methylation in PCa.

Concerning methylation, the lower levels found for methylated residues in PCa patients would reveal a higher demethylation in this group of individuals leading to a higher gene silencing effect. Furthermore, several studies have sugges-

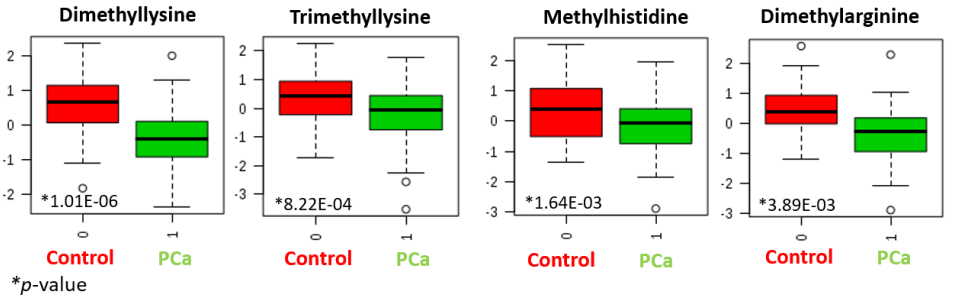


Fig. 4. Box and whisker plots for methylated residues to help in interpreting the role of histone methylation in PCa.

ted that decreased histone acetylation and methylation may be a consequence of cancer cell metabolism, since a rapid proliferation of cells and a higher activation of macromolecular biosynthesis may deplete the levels of acetyl and methyl donors such as acetyl coenzyme A and S-adenosyl methionine (SAM) [29].

On the other hand, three metabolic products of lysine, arginine and histidine significantly detected in this study were 5-acetamidovalerate, acetylputrescine and imidizol lactate, respectively. All these metabolites were present at higher concentrations in negative biopsy patients as compared to PCa cases (see Supplementary Fig. 2), thus also indicating a higher activation of basic amino acids metabolism.

DNA methylation. DNA methylation occurs when a methyl group is attached to the fifth carbon of cytosine nucleotide in a process catalyzed by DNA methyltransferases with SAM as methyl donor. The two most common mechanisms of aberrant methylation are global hypomethylation and site specific hypermethylation. Both processes impact gene expression, genome stability, genetic imprinting and cellular differentiation. DNA methylation can be enzymatically removed by several mechanisms including base excision repair, nucleotide excision repair and hydrolysis [30,31].

In mammalians, 5-methylcytosine (5mC) is sequentially oxidized into 5-hydroxymethylcytosine (5hmC), further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) [32,33]. These three oxidized methylcytosines point out the occurrence of both passive and active DNA demethylation and also serve as stable epigenetic marks [34]. Enzymatic mechanisms for removal of DNA methylation may release repaired products into the bloodstream and consequently they can appear in the urine. In fact, the deoxynucleosides of 5mC and 5hmC –namely, 5-methyl-2-deoxycytidine (5mdC) and 5-hydroxymethyl-2-deoxycytidine (5hmdC)– have recently been detected and quantified in urine [35,36], and several biological processes such as apoptosis can degrade DNA into single deoxynucleosides [37]. According to these studies, urinary 5mdC and 5hmdC contents may offer a view of the metabolic and functional status of tissues or organs in the human body.

In this research 5mdC phosphate has been significantly detected at different concentration in PCa patients as compared to negative biopsy individuals. Particularly, its concentration was lower in PCa cases, as Fig. 5A shows.

This result could be interpreted by a lack of efficiency in the machinery for repairing DNA methylation. As a result, this 5mdC epigenetic mark would be less metabolized and, therefore, detected at lower concentration level in urine from PCa patients. This finding is consistent with that reported by Morey *et al.*, who found genomic DNA hypomethylation and loss of 5-methyl-2-deoxycytidine in both early and late stages of mouse PCa adenocarcinoma [38].

Increased evidences also connect DNA methylation and methylated DNA lesions. Previous studies have reported that SAM was able to methylate DNA (O-methylguanine, 7-methylguanine and 3-methyladenine) without enzymatic involvement to produce promutagenic and procarcinogenic lesions [39,40]. This assumption could also explain the higher levels of 7-methylguanine in urine from PCa patients as compared to negative biopsy controls. In fact, previous studies have found elevated levels of methylguanine in serum and urine from cancer patients. For instance, Jung *et al.* found a significant increase of 7-methylguanine in malignant tissue from PCa patients as compared to non-malignant tissue [13].

RNA cap methylation. Eukaryotic mRNA is modified by the insertion of 7-methylguanosine 'cap' to the first transcribed nucleotide. This modification is required for efficient gene expression to produce proteins and cell viability as the

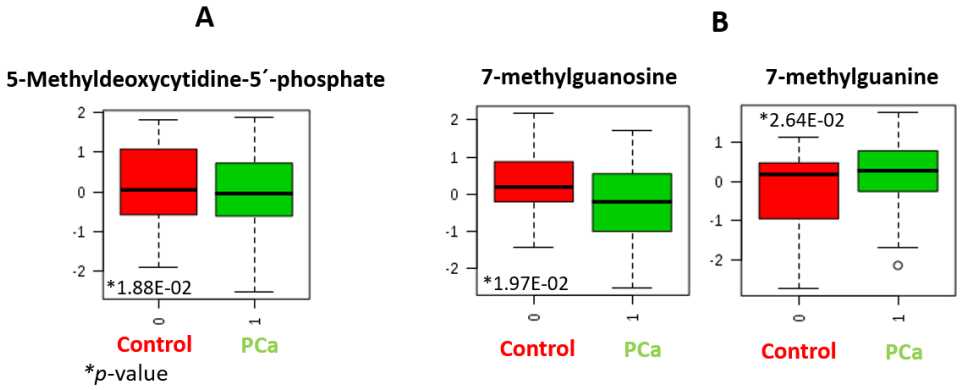


Fig. 5. Box and whisker plots comparing the concentration of (A) 5-methyldeoxycytidine-5′ phosphate and (B) 7-methylguanosine and 7-methylguanine, detected in urine from PCa patients and negative biopsy individuals.

7-methylguanosine cap allows translation of the majority of mRNAs, stabilizes mRNA against exonucleases and promotes transcription, splicing, polyadenylation and nuclear export of mRNA.

In the present research, 7-methylguanosine was found at significantly different concentration in PCa patients *versus* negative biopsy individuals. Particularly, this metabolite was found at lower concentration in PCa patients, which suggests that the translation process would be impeded in patients with the tumor.

The metabolite resulting from potential depurination of 7-methylguanosine (7-methylguanine) was also significantly different in the comparison of PCa patients and healthy individuals. Particularly, the concentration profile of 7-methylguanine was inverse to that described for 7-methylguanosine. Thus, 7-methylguanine was detected at higher concentration in PCa individuals than in healthy cases. Fig. 5B shows the box-and-whisker plots comparing the concentration of 7-methylguanosine and 7-methylguanine detected in urine from PCa patients and negative biopsy individuals. Depurination could explain the reduced

concentration of 7-methylguanosine in PCa patients. This statement is supported on the concept of "decapping", the major mechanism by which the mRNA cap methylation is reversed by removing the entire 7-methylguanosine cap. This process would cause serious errors during genes translation, with cell apoptosis as a result.

Miscellaneous. Tryptophan metabolism has been widely related both to cancer and immune system. In fact, kynurenic acid has been proposed as a biomarker in urine sediment collected from subjects following a DRE to discriminate PCa patients from healthy subjects. In the present study, kynurenic acid, xanthunerate and 8-methoxykynurenate, involved in tryptophan metabolism, presented lower levels in PCa individuals as compared to controls.

Other group of metabolites significantly decreased in urine from PCa patients in comparison to the control group included acetyltaurine, isethionate and sulfoacetate. These three metabolites, excreted in urine, are end-products from the metabolism of taurine. Chatzakos *et al.* showed that N-acyltaurines are anti-proliferative in PCa cells [41]. Also, Tang *et al.* investigated the anti-prostate cancer metastasis effect of taurine, and they proved that taurine attenuated PSA and several metastasis-related genes in human PCa cells, such as LNCaP and PC-3. In addition, taurine inhibited migration of LNCaP and PC-3 [42]; therefore, it seems that taurine metabolism could be crucial in PCa regulation.

Conclusions

In this research, a comprehensive global analysis by LC–QTOF of urine from PCa patients and negative biopsy control individuals has allowed discrimination between both groups. An unpaired *t*-test (*p*-value<0.05) provided 28 significant metabolites involved in biological processes such as the regulation of epigenetic marks, essentially DNA methylation and modifications of histones, as well as RNA cap methylation. A PLS-DA model allowed discriminating both groups of individuals with sensitivity and specificity values of 88.4% and 92.9%, respectively.

This study revealed changes at metabolomics level that can be associated with epigenetic marks widely studied as promising markers for diagnostic of PCa due to their involvement in gene expression or silencing. An additional process such as RNA cap methylation also reported significant differences between PCa patients and controls, which emphasizes the relevance of RNA transcription. Therefore, the analysis of urine revealed metabolite changes that can be interpreted according to alterations in the expression and transcription of genes, thus showing once again the capability of metabolomics to provide information on upstream processes.

Acknowledgements

The Spanish Ministerio de Economía y Competitividad (MINECO), Junta de Andalucía and FEDER program are gratefully acknowledged for financial support (Projects "Development of methods for early cancer detection", FQM-1602 and CTQ2015-68813-R). M. A. Fernández-Peralbo and M. Calderón-Santiago also thank to these organisms their research contracts. The biological samples repository of Reina Sofía Hospital (Córdoba, Spain) is also gratefully acknowledged for coordination tasks in the selection of urine samples.

References

[1] Desantis, C. E. *et al.* Cancer treatment and survivorship statistics 2014. CA Cancer J Clin 64(4), 252–271 (2014).

[2] Carter, H. B. *et al.* Early detection of prostate cancer: AUA guideline. J Urol 190(2), 419–426 (2014).

[3] Key, T. J. Prostate cancer prevention (eds. Cuzick, J. & Thorat, M. A.) 39–46 (Springer, 2014).

[4] Shariat, S. F., Karam, J. A., Margulis, V. & Karakiewicz, P. I. New blood-based biomarkers for the diagnosis, staging and prognosis of prostate cancer. BJU Int 101, 675–683 (2007).

[5] Truong, M., Yang, B. & Jarrard, D. Towards the detection of prostate cancer in urine: A critical analysis. J Urol 189(2), 422–429 (2012).

[6] Auprich, M. et al. Contemporary role of prostate cancer antigen 3 in the management of prostate cancer. Eur Urol 60, 1045–1054 (2011).

[7] Winnes, M., Lissbrant, E., Damber, J. & Stenman, G. Molecular genetic analyses of the TMPRSS2-ERG and TMPRSS2-ETV1 gene fusions in 50 cases of prostate cancer. Oncology Reports 17, 1033–1036 (2007).

[8] Jerónimo, C. *et al*. Epigenetics in prostate cancer: Biologic and clinical relevance. Eur Urol 60, 753–766 (2011).

[9] Roobol, M. J., Haese, A. & Bjartell, A. Tumour markers in prostate cancer III: Biomarkers in urine. Acta Oncologica 50(1), 85–89 (2011).

[10] Seligson, D. B. *et al*. Global histone modification patterns predict risk of prostate cancer recurrence. Nature 435, 1262–1266 (2005).

[11] Sreekumar, A. *et al*. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature 457, 910–4 (2009).

[12] Jentzmik, F. *et al*. Sarcosine in urine after digital rectal examination fails as a marker in prostate cancer detection and identification of aggressive tumours. Eur Urol 58, 12–18 (2010).

[13] Jung, K. *et al*. Tissue metabolite profiling identifies differentiating and prognostic biomarkers for prostate carcinoma. Int J Cancer 133, 2914–2924 (2013).

[14] Struck-Lewicka, W. *et al*. Urine metabolic fingerprinting using LC–MS and GC–MS reveals metabolite changes in prostate cancer: A pilot study. J Pharm Biomed Anal 111, 351–361 (2015).

[15] Heidenreich, A. *et al*. EAU Guidelines on prostate cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent — Update 2013. Eur Urol 65, 124–137 (2014).

[16] Calderón-Santiago, M., Priego-Capote, F. & Luque de Castro, M. D. Enhanced detection and identification in metabolomics by use of LC–MS/MS untargeted analysis in combination with gas-phase fractionation. Anal Chem 86, 7558–7565 (2014).

[17] Tautenhahn, R., Böttcher, C. & Neumann, S. Highly sensitive feature detection for high resolution LC/MS. Bio Med Central 16, 1–16 (2008).

[18] Warrack, B. M. *et al*. Normalization strategies for metabonomic analysis of urine samples. J Chromatogr B 877, 547–52 (2009).

[19] Miller, R. C. *et al*. Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations. Clin Chem 932, 924–932 (2004).

[20] Vogelmann, J., Valeri, A., Guillou, E. & Cuvier, O. Roles of chromatin insulator proteins in higher-order chromatin organization and transcription regulation. Nucleus 2(5), 358–369 (2013).

[21] Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science 272, 408–411 (1996).

[22] Kim, S. C. *et al*. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Mol Cell 23, 607–618 (2006).

[23] Baylin, S. B. & Ohm, J. E. Epigenetic gene silencing in cancer–a mechanism for early oncogenic pathway addiction? Nature Rev 6, 107–116 (2006).

[24] Kelly, T. K., De Carvalho, D. D. & Jones, P. A. Nat Biotechnol 28(8), 1069–1078 (2012).

[25] Halkidou, K. *et al*. Upregulation and nuclear recruitment of HDAC1in hormone refractory prostate cancer. Prostate 59, 177–189 (2004).

[26] Sato, S. *et al*. Histone deacetylase inhibition in prostate cancer triggers miR-320-mediated suppression of the androgen receptor. Cancer Res, doi: 10.1158/0008-5472.CAN-15-3339 (2016).

[27] Deplus, R., Putmans, P. & Yamada, M. Functional connection between deimination and deacetylation. Mol Cell Biol 29, 4982–4993 (2009).

[28] Kan, R. *et al.* Potential role for PADI-mediated histone citrullination in preimplantation development. BMC Dev Biol 12, 1 (2012).

[29] Kurdistani, S. K. Epigenetics and Disease. (eds. Gasser, S. M. & Li, E.) 91–106 (Springer, New York, 2011).

[30] Barreto, G. *et al.* Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature 445, 671–675 (2007).

[31] Zhu J. K. Active DNA demethylation mediated by DNA glycosylases. Ann Rev Genet 43, 143–166 (2009).

[32] Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466, 1129–1133 (2010).

[33] Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324(5929), 930–935 (2009).

[34] Ko, M., An, J. & Rao, A. ScienceDirect DNA methylation and hydroxymethylation in hematologic differentiation and transformation. Curr Opin Cell Biol 37, 91–101 (2015).

[35] Hu, C., Liu, H., Li, Y. & Chao, M. Direct analysis of 5-methylcytosine and 5-methyl-2′-deoxycytidine in human urine by isotope dilution LC-MS/MS: correlations with N-methylated purines and oxidized DNA lesions. Chem Res Toxicol 25, 462–470 (2012).

[36] Yin, R., Mo, J., Lu, M. & Wang, H. Detection of human urinary 5hydroxymethylcytosine by stable isotope dilution HPLC-MS/MS analysis. Anal Chem 87, 1846–1852 (2015).

[37] Kawane, K., Motani, K. & Nagata, S. DNA degradation and its defects. Cold Spring Harb Perspect Biol 6, 1–14 (2014).

[38] Kinney, S. R. M. *et al.* Stage-specific alterations of DNA methyltransferase expression, DNA hypermethylation, and DNA hypomethylation during prostate cancer progression in the transgenic adenocarcinoma of mouse prostate model. Mol Cancer Res 6(8), 1365–1374 (2008).

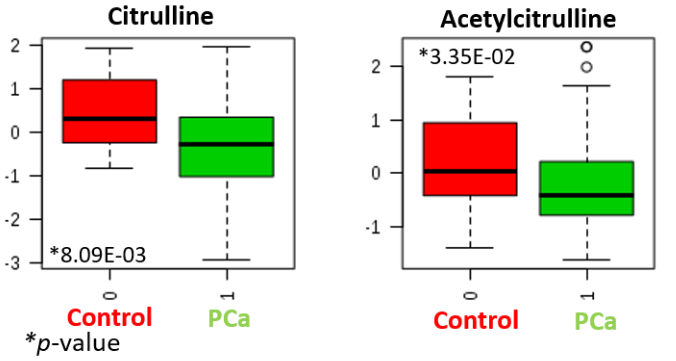
[39] Barrows, L. R. & Magee, P. N. Nonenzymatic methylation of DNA by S-adenosylmethionine in vitro. Carcinogenesis 3, 349–351 (1982).

[40] Rydberg, B. & Lindahl, T. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. EMBO 1, 211–216 (1982).

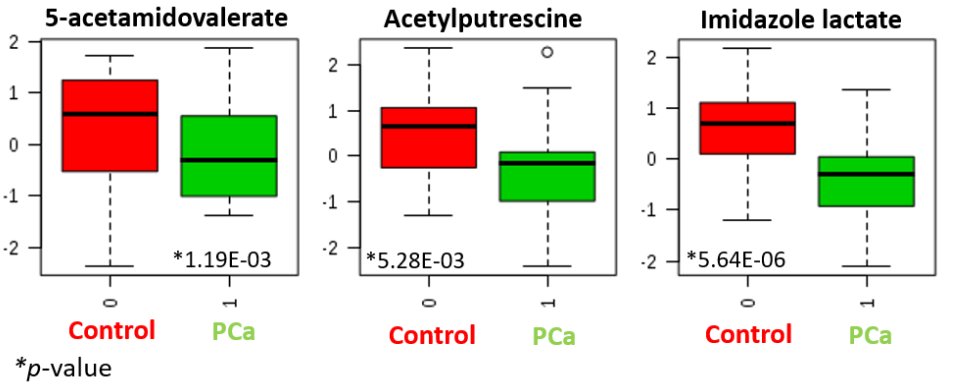
[41] Hunt, M. C. N-Acyl taurines are anti-proliferative in prostate cancer cells. Lipids 47, 355–361 (2012).

[42] Tang *et al.* Taurine 9 (eds. Marcinkiewicz, J. & Schaff, S. W.) 203–216 (Springer, New York, 2015).

Supplementary material



Supplementary Fig. 1. Box and whisker plots from citrulline and acetylcitrulline in the two evaluated groups.



Supplementary Fig. 2. Box and whisker plots from 5-acetamidovalerate, acetylputrescine and imidazole lactate in the two evaluated groups.

CHAPTER 9

Urinary metabolite panels for prostate cancer detection

Sent to Cancer Letters

Sent to Cancer letters

Urinary metabolite panels for prostate cancer detection

M. A. Fernández-Peralbo[1,2], E. Gómez-Gómez[2,3], M. Calderón-Santiago[1,2], J. Carrasco-Valiente[2,3], J. Ruiz-García[2,3], M. J. Requena-Tapia[2,3], M. D. Luque de Castro[1,2], F. Priego-Capote[1,2]

[1]*Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain.*

[2]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía University Hospital, University of Córdoba, E-14004, Córdoba, Spain.*

[3]*Department of Urology (IMIBIC), Reina Sofia University Hospital, University of Córdoba, E-14004, Córdoba, Spain.*

Urinary metabolite panels for prostate cancer detection

M. A. Fernández-Peralbo, E. Gómez-Gómez, M. Calderón-Santiago, J. Carrasco-Valiente, J. Ruiz-García, M. J. Requena-Tapia, M. D. Luque de Castro, F. Priego-Capote

Abstract

The lack of specificity of the prostate specific antigen (PSA) serum level as screening marker for diagnostic of prostate cancer (PCa) drives the search for new biomarkers, preferably non-invasive biomarkers. In the present research, a comprehensive untargeted metabolomics analysis by LC–QTOF of urine from 62 patients with a clinically significant PCa and 42 negative biopsy controls, both groups confirmed by biopsy, was applied to search for potential biomarkers that may help to understand pathological processes related to PCa. The significant identified metabolites were alternatively combined with PSA measurements in 5-variable panels (5 metabolites or 4 metabolites+PSA) for diagnostic of PCa. The panels were prepared to maximize sensitivity, specificity or accuracy (sensitivity+specificity). The main goal of the panels was to propose a screening tool to complement other diagnostic tests to reduce the false positive cases, which means to reduce the rate of individuals subjected to invasive biopsy. The panels were also characterized by high sensitivity outcomes, which means that they can also detect true positive cases. In fact, the best results achieved in terms of sensitivity and specificity were 97.7% and 100%, respectively, which were obtained by combination of dimethyllysine, dimethylarginine, acetylputrescine, citrulline and PSA. Thus, our diagnostic models suggest that urine metabolomic profiling may have potential for clinical diagnosis of PCa.

Keywords Biomarkers; LC–QTOF; Metabolomics; Panelomix; Prostate cancer; Urine

1. Introduction

Prostate cancer (PCa) is the second most frequently diagnosed cancer in men and the fifth leading cause of cancer death among men worldwide [1]. In 2017 an estimated 16.136 new cases will be diagnosed with this type of cancer and 26.730 men will die in the United States [2]. Outcomes in PCa are favorable if diagnosed early, when the disease has not achieved the metastasis phase. Current PCa screening tests consist of quantitative measurement of serum prostate specific antigen (PSA) levels and digital rectal examination (DRE) as compared to decisive diagnostic that involves prostate biopsy. Additionally, the lack of specificity of PSA as PCa marker combined with the asymptomatic and slow-growing nature of prostate tumors has resulted in a high number of over-diagnosis cases subjected to surgery or treatments with adverse side effects. Therefore, the development of discrimination models targeted at reducing false positives would be useful as screening strategy to minimize the number of patients subjected to invasive confirmatory tests.

Nowadays there are some new biomarkers already approved by the FDA organization such as the Prostate Health Index (PHI), based on the combination of PSA, free PSA and p2PSA; or the PCA 3 (Progensa) based on post digital rectal examination and subsequent analysis of PCa gene 3 in urine. There are other laboratory tests such as 4K score, Mi-Prostate score, and Prostarix (based on a metabolomics panel) that have provided promising results; nevertheless, they are not still implemented worldwide [3]. Furthermore, some emerging areas of study including miRNAs, long noncoding RNA, metabolites, exosomes, or microbiota are being explored at present [4].

Cell proliferation during PCa pathogenesis may lead to alterations in metabolite profiles, since such biomolecules represent the end products of physiological processes. Metabolomics as the youngest of the great omics has recently experienced a significant development supported on the advances in analytical instrumentation. This omics discipline allows analyzing the entire metabolic profile of a biological system [5] and can be used to discover cancer bio-

Table 1. Characteristics of the cohort under study.

Variable	Negative biopsy (n= 42)	PCa (n= 62)	*p*-value
Age (years)	62.5 (57.8-69)	71 (64-80)	<0.05
Positive DRE	0	36 (58.1%)	<0.05
***Serum PSA(1) (ng/mL)**	4.1 (3.6-5.8)	11 (6.9-29.75)	
≤ 10	38 (90.4%)	27 (43.5%)	<0.05
>10	4 (9.6%)	35 (56.5%)	
Prostate volumen (mL)			
> 30cc	23 (54.7)	27 (43.5)	>0.05
***Serum PSA(2) (ng/ml)**	2.3 (1.8-2.7)	11 (6.9-40.7)	
<3	41 (97.6%)	0	<0.05
3-10	1 (2.4%)	28 (45.2%)	
>10	0	34 (54.8%)	
Gleason pattern			
6		4 (6.5%)	
≥7		58 (93.5%)	

PSA (1): Serum PSA before the biopsy; PSA (2): Serum PSA the day of biopsy

*Data are expressed as median value (range of variability) (Q1-Q3)

markers useful for an early diagnosis of the disease, as well as assessment of prognosis and treatments response. Despite the lack of biomarkers reaching clinical practice, metabolomics proved to be a powerful tool in the discovery of new biomarkers for PCa detection, with sarcosine being one of the most promising biomarkers identified to date [6]. However, the use of sarcosine as a PCa biomarker remains a controversial issue within the scientific community [7]. Besides, metabolic discrimination based on metabolite panels by monitoring a group of molecules may lead to more accurate diagnosis.

Since metabolite profiles closely reflect the total cellular situation, metabolomics has been developed into a more integrative approach as compared to other omics approaches. In the last decade, advances in nuclear magnetic resonance spectroscopy (NMR) [8] and mass spectrometry (MS) —the latter coupled to separation techniques, mostly to liquid chromatography (LC–MS) [9], and in a lesser proportion to capillary electrophoresis (CE–MS) [10] or gas

chromatography (GC–MS) [11]— have been applied to identify metabolic alterations in PCa that may lead to clinically useful biomarkers. LC–MS has been mainly used for the study of PCa in serum, plasma and urine [12–14], being NMR more widely used in tissue samples [15], but also in seminal [16] and prostatic fluids [17]. These studies highlighted the alterations produced in the synthesis and degradation of glycine [6], in the amino acids metabolism [12], kynurenine and pyrimidine metabolism [18], membrane phospholipidic metabolism [19], fatty acids [13], energetic metabolism [17] and polyamines synthesis [20], among others, when PCa patients are compared to healthy controls.

Panels of metabolites to discriminate PCa in urine are still scarce in the literature. On the contrary, more studies have established panels of proteins [21] and genes [22]. However, many of them do not carry out external validation and the values of sensitivity and specificity are not good enough. Thus, no invasive or minimum invasive tests applied to biological samples such as blood or urine, and based on more sensitive biomarkers, are urgently needed to separate individuals with clinically significant PCa from those with indolent disease.

In this work, comprehensive analysis of urine metabolome was applied to search for potential biomarkers that may help to understand pathological processes related to PCa. Urine has some advantages for metabolomics studies since it has a high abundance of metabolites, and low content of proteins; it is collected non-invasively and requires minimal sample preparation prior to analysis. For our purpose, urine from patients with aggressive PCa and healthy individuals, both groups confirmed after biopsy, was analyzed by LC–MS/MS in high-resolution mode with the objective of discovering alterations in the metabolic phenotype of PCa patients that help in the search for new diagnosis biomarkers.

2. Materials and methods

2.1. Ethical statement

A cohort of patients scheduled for prostate biopsy according to the European

Association of Urology (EAU) guidelines were selected for this study. Informed consent before enrolment in the project was signed by all members of the cohort. The experiment was designed following the guidelines dictated by the World Medical Association Declaration of Helsinki (2004), and was supervised by the Reina Sofia Hospital Biomedical Research Ethical Committee and University of Cordoba Biomedical Experimentation Ethical Committee that approved the experiments.

2.2. Subjects and sampling

Morning urine samples were collected in the Urology Department of the Reina Sofia Hospital from 104 patients the day of prostate biopsy and before the intervention. First morning urine was discarded for development of this research. The same dose of ciprofloxacin was administered to all the patients the day before sampling after overnight fasting. The samples were aliquoted in 1.5 mL tubes and stored at −80 ºC until analysis. The patients were divided into two groups, PCa patients and negative biopsy controls with the characteristics defined in Table 1. The PCa group included 62 patients with clinically significant PCa confirmed by prostate biopsy. Clinically significant PCa cases matched at least one of the following conditions: clinical stage > T1c, Gleason pattern > 6, more than 3 positive cores and/or > 50% cancer per core. On the other hand, the control group, which included 42 individuals, was confirmed by histological analysis. Negative biopsy controls gave negative DRE in contrast to PCa patients who reported positive DRE in 58% of cases (36 out of 62). The mean age of patients with PCa was 71 *versus* 62 years of the control group, and no significant differences in terms of prostate volume were found between the two groups.

An initial PSA test was applied to each patient as a decision tool to undergone prostate biopsy (following EAU guidelines recommendation at least two PSA levels were needed to indicate a prostate biopsy in clinical practice). Moreover, a further PSA test (second PSA) was carried out the day of the biopsy. As foreseeable, PCa patients provided higher PSA in both measures, as shows Table 1. In the first test, all PCa individuals gave values above 3 ng/mL, with 27 cases between 3 and 10

ng/mL and 35 patients above 10 ng/mL. The values were quite similar in the second PSA test. Concerning the negative biopsy group, it included 39 patients (92.9%) with serum PSA above 3 ng/mL after the first analysis; however, this group was clearly dominated by individuals with PSA levels between 3 and 10 ng/mL. A clear change was observed in the second PSA test since most of the control group individuals gave serum PSA levels below 3 ng/mL the day of biopsy.

2.3. Untargeted metabolomics analysis of urine

Untargeted metabolomics analysis of urine samples was addressed by application of the protocol described by Fernández-Peralbo et al. [23]. This protocol, which includes sample preparation, analysis by LC–MS/MS and data pretreatment, is detailed in Supplementary Information.

2.4. Statistical analysis

The predictive capability of each metabolite by Random Forest Analysis was determined by MetaboAnalyst 3.0, a web-tool based on analytical high-throughput metabolomics studies (http://www.metaboanalyst.ca) for data analysis, visualization and interpretation. Finally, the PanelomiX toolbox [24] was used to configure panels of metabolites by their iterative combination in sets of 3–5 compounds according to selected thresholds to provide optimal classification performance in terms of sensitivity and specificity. The PanelomiX methodology was applied to 70% of patients for development of the models whereas the rest 30% was used as external validation set.

3. Results and discussion

3.1. Tentative identification of metabolites significantly altered by prostate cancer

A preliminary data mining step was applied to filter out MFs corresponding to redundant information from adducts and isotopes. This pretreatment step also removed MFs not detected in at least 75% of the samples pertaining to one of the

groups (negative biopsy or PCa patients). After this operation, an unpaired *t*-test was applied to target the identification step at those statistically significant MFs. No multiple testing correction was considered in this analysis to increase the list of significant metabolites prior to evaluation of their biomarker capability. Forty-two significant metabolites were tentatively identified by MS/MS. Despite most of the metabolites were detected in both positive and negative ionization polarities, only the ionization mode providing the highest sensitivity was considered. The tentatively identified metabolites are listed in Table 2 that includes information of the precursor ion, the most representative product ions supporting identification, retention time, chemical formula, *p*-value, fold change ratio, and regulation in PCa cases as compared to control individuals. Among them, amino acids and derivatives constituted 38.1% of significant metabolites, nitrogenous bases derivatives 12%, sulfonic acids 7.1%, quinolones 7.1%. The rest is formed by a miscellaneous group that includes carnitines, ketoacids and acetamides. It is remarkable that 21.4% of the metabolites were acetylated compounds (acetylcitrulline, acetylarginine, acetylputrescine or acetyllysine among them), while a similar percentage corresponded to methylated amino acids or nitrogenous bases (methylguanosine, methyllysine, methylhistidine or dimethylarginine among them).

3.2. Prediction capability of metabolites discriminating prostate cancer patients

An adequate selection of the patients forming each discriminated group is extremely advisable to obtain quality results. Most previous studies conducted with PCa used a supposedly healthy control cohort, which were not confirmed by a biopsy, probably due to normal PSA serum values. The attractiveness of the control group in this study lies in the fact that all patients were confirmed as negative biopsy cases. Moreover, a ciprofloxacin doses was administered to all patients the day of sampling, as a mandatory treatment before biopsy. The research carried out by Struck-Lewick *et al.* to search for metabolite changes in urine from PCa pointed out that the group of healthy men did not undergo medication at the time of sample collection but the authors did not detail the medication of PCa patients [25]. Sreekumar et al. found five metabolites (sarcosine, proline, kynuren-

Table 2. Metabolites tentatively identified by MS/MS with significant different levels in urine from PCa versus those of negative biopsy patients. (Grouping of the compounds was based on the metabolic pathways in which they are preferentially involved and sorted by their fold change).

Pathway	Compound name	RT	Precursor ions *m/z*	Adduct	Formula	Product ions *m/z*	*p*-value	Fold change (controls vs. PCa)	Regulation in cancer
Lysine degradation	Dimethyllysine	1.05	175.14409	M+H	C8H18N2O2	84.0815/130.0878	1.01E-06	1.74	↓
	5-acetamidovalerate	1.73	160.0968	M+H	C7H13NO3	142.0862/70.0655/98.0623	1.19E-03	1.51	↓
	Acetyl-lysine	1.48	189.1230	M+H	C8H16N2O3	84.0811/126.0914	6.19E-05	1.33	↓
	Trimethyl-lysine	1.05	189.16005	M+H	C9H20N2O2	130.0861/84.0805/60.0806	8.22E-04	1.28	↓
Histidine metabolism	Imidazole lactate	1.18	157.0607	M+H	C6H8N2O3	111.0550/83.0610	5.64E-06	1.46	↓
	Histidine	1.05	156.07684	M+H	C6H9N3O2	110.0718/83.0625/95.0600	1.69E-06	1.43	↓
	Methylhistidine	1.06	170.09247	M+H	C7H11N3O2	124.0864/109.075/96.0679/83.0600/68.0493	1.64E-03	1.40	↓
	Acetylhistidine	1.28	198.0874	M+H	C8H11N3O3	152.0817/110.0717/180.3059	5.84E-03	1.26	↓
Arginine metabolism	Urea	1.24	61.0399	M+H	CH4N2O	44.0133	8.39E-03	1.37	↓
	Acetylarginine	1.67	215.11621	M-H	C8H16N4O3	173.1032/129.1063	2.35E-03	1.34	↓
	Acetylcitrulline	2.52	216.0990	M-H	C8H15N3O4	173.0931/131.0821	3.35E-02	1.29	↓
	Acetylputrescine	1.25	131.1178	M+H	C6H14N2O	72.0809/114.0907	5.28E-03	1.30	↓
	Dimethylarginine	1.17	203.15036	M+H	C8H18N4O2	70.0653/116.0698/158.1285/88.0874	3.89E-03	1.25	↓
	Citrulline	1.34	176.10316	M+H	C6H13N3O3	159.0764/130.0975/70.0651	8.09E-03	1.23	↓
Steroid hormone biosynthesis	Estrone	8.97	269.1504	M-H	C18H22O2	145.061	3.60E-02	1.34	↓

Color scale values of fold change
From 1.5 to 1.75
From 1.25 to 1.5
From 1.1 to 1.25
From -1.1 to -1.25
From -1.25 to -1.5

Cont. Table 2

Glycine, serine and threonine metabolism	Threonine	1.36	120.06586	M+H	C4H9NO3	58.0655/74.0605/ 103.0632	1.17E-02	1.31	↓
Oxidative phosphorylation	Phosphoric acid	1.13	78.95933	M-H2O-H	H3O4P	62.9641	1.40E-02	1.28	↓
Nicotinate and nicotinamide metabolism	Nicotinic acid	1.98	122.0246	M-H	C6H5NO2	78.0349	1.82E-02	1.26	↓
Acetylcarnitine metabolism	Dimethylheptanoyl carnitine	9.38	302.2330	M+H	C16H31NO4	85.0289/60.0809/ 243.1592	1.65E-02	1.37	↓
	Tiglylcarnitine	6.04	244.1550	M+H	C12H21NO4	85.0289/60.0805/ 185.0792/ 124.0868	1.43E-02	1.22	↓
Tyrosine metabolism	Tyrosine	2.96	180.0660	M-H	C9H11NO3	119.0503/163.0404 / 136.0759/ 93.0318	2.44E-03	1.34	↓
	Vanillylmandelic acid	4.61	197.0460	M-H	C9H10O5	137.0231/138.0259	1.87E-02	1.21	↓
Tryptophan metabolism	8-methoxykynurenate	7.79	220.0610	M+H	C15H17NO8	174.0554/116.9734/ 202.0484	1.39E-02	1.31	↓
	Kynurenic acid	7.00	188.0352	M-H	C10H7NO3	144.0454	6.21E-04	1.30	↓
	4-hydroxy-4-methyl-2-oxoadipate	2.60	189.0404	M-H	C7H10O6	129.0191/174.0162/ 111.0086/ 99.0088/	2.77E-02	1.23	↓
	Xanthurenic acid	6.80	206.0450	M+H	C10H7NO4	188.0345/160.0386 /132.0459	9.79E-03	1.22	↓
Taurine metabolism	Sulfoacetate	1.25	138.9708	M-H	C2H4O5S	94.9810/79.9575	1.56E-03	1.28	↓
	Isethionate	1.23	124.9916	M-H	C2H6O4S	79.9572/94.985/ 106.9797	6.71E-03	1.19	↓
	Acetyltaurine	1.56	166.0194	M-H	C4H9NO4S	79.9579/124.0078	1.63E-02	1.12	↓

Cont. Table 2

Alanine, aspartate and glutamate metabolism	Acetylaspartylglutamic acid	2.81	303.0833	M-H	C11H16N2O8	96.0077/128.0350	5.21E-03	1.24	↓
	Acetylaspartate	1.86	174.0410	M-H	C6H9NO5	88.0403/130.0507/ 115.0037/ 58.0296	5.96E-03	1.19	↓
	2-Oxoglutaramate	1.49	144.03145	M-H	C5H6O5	126.0202/100.0409 /82.0305/ 72.0454/41.9983	9.85E-04	1.18	↓
Glutamine and glutamate metabolism	2-pyrrolidone-5-carboxylate	2.10	128.03542	M-H	C5H7NO3	85.0295/41.0399/ 110.8772	8.21E-03	1.25	↓
Phenylalanine metabolism	Phenylalanine	4.59	164.0720	M-H	C9H11NO2	147.0437/103.0545	1.51E-02	1.21	↓
	4-aminohippuric acid	5.57	193.0618	M-H	C9H10N2O3	92.0503/149.0717	1.22E-02	1.19	↓
Sialic acid	Sialyl-N-acetyl-lactosamine	1.36	675.2460	M+H	C25H42N2O19	204.0870/222.1008/274.0919/ 138.0516/256.0787	1.77E-02	1.16	↓
Purine and pyrimidine metabolism	5-methyldeoxycytidine 5´-phosphate	2.81	320.0620	M-H	C11H15NO10	110.0245/240.0523	1.88E-02	1.21	↓
	7-methylguanosine	5.31	296.100	M-H	C11H15N5O5	164.0574	1.97E-02	1.15	↓
	5,6-dihydrouridine	1.71	245.0778	M-H	C9H14N2O6	129.0650/155.047/ 41.9991	3.87E-02	1.12	↓
	Uric acid	2.35	169.0360	M+H	C5H4N4O3	141.0405/152.0091/ 126.0295/ 98.0341	3.67E-02	-1.20	↑
	7-methylguanine	2.10	164.0580	M+H	C6H7N5O	149.0497/124.0501	2.64E-02	-1.30	↑
Folate biosynthesis	7,8-dihydroneopterin	1.71	256.1040	M+H	C9H13N5O4	165.0658/179.079/ 208.0820/ 238.0931/220.0838 /86.0345	1.99E-02	-1.29	↑

ine, uracil and glycerol-3-phosphate) significantly increased with disease progression, from benign to PCa, and to metastatic PCa [6]. Taking into account these results Gamagedara et al. analyzed the levels of proline, kynurenine, uracil and glycerol-3-phosphate in 63 PCa patients using an LC–MS validated method and compared the results with those from 68 individuals with no evidence of malignancy volunteers [26]. The statistical results showed that these biomarkers cannot differentiate PCa from no malignancy evidence. The strong correlation between these metabolites and urinary creatinine concentrations implies that their occurrence is mainly due to renal excretion. A recent study by Liang et al. with a larger number of samples, 236 PCa patients and 233 controls, has resulted in the discovery of top 5 ranking biochemical markers for PCa in a VIP plot including glycocholic acid, hippurate, 5-hydroxy-L-tryptophan, taurocholic acid, and chenodeoxycholic acid in the PCa subjects from the control cases [27]. A ROC analysis revealed glycocholic acid, hippurate, 5-hydroxy-L-tryptophan to be potent discriminators between PCa and control groups (AUC > 0.95). However, this study does not include these metabolites in a biomarkers panel. Criteria for selection of controls were no history of PCa, denial of clinical manifestations of PCa, and a low level of urine PSA. Before collection of urine samples, patients had not undergone therapeutic interventions.

The predictive capability of significant metabolites in the present study was assessed by Random Forest analysis intended to discriminate PCa cases and negative biopsy individuals. Metabolites are ranked by a variable importance function that considers interactions among them. Fig. 1 shows the top-15 metabolites ranked by their contribution to classification accuracy, which means prediction capability.

All these metabolites except 7,8-dihydroneopterin presented higher relative concentrations in negative biopsy cases than in PCa patients. As can be seen, dimethyl-lysine, acetylputrescine and imidazole lactate presented the highest prediction capability and, potentially, the highest probability to appear in a panel for discrimination of PCa patients and control individuals. Dimethyl-lysine is highlighted as it matched the most significant *p*-value and fold change, followed by im-

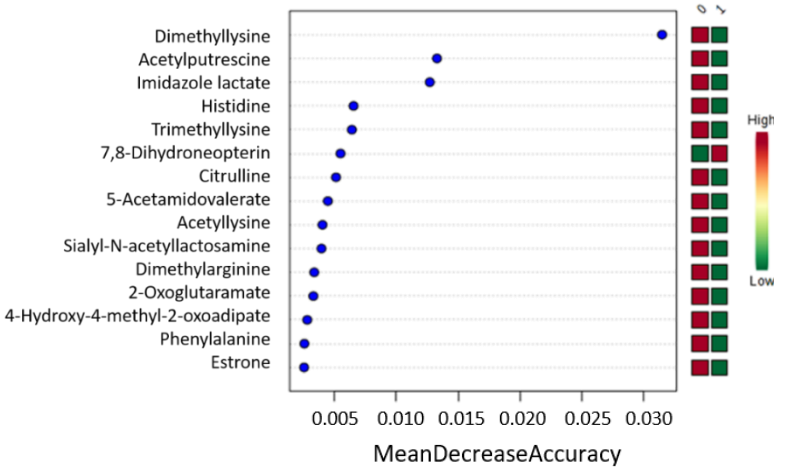


Fig. 1. Random Forest Analysis of significant metabolites found in human urine to discriminate PCa patients from negative biopsy individuals.

idazole lactate and histidine, as Table 2 shows. Also remarkable is the presence of compounds involved in the metabolism of amino acids, making up the 67% of the ranked metabolites.

These results agree with the predictive capability of the metabolites provided by the individual ROC curves. The three compounds with the best predictive capability according to the random forest analysis (viz., dimethyllysine, acetyl-putrescine and imidazole lactate) also provided the best behavior in terms of sensitivity and selectivity for discrimination of PCa patients. The sensitivity of dimethyllysine, acetylputrescine and imidazole lactate were 85.7%, 71.4 and 78.6%, respectively, while the selectivity values were 72.1%, 81.4% and 74.4%.

3.3. Configuration of biomarkers panels for detection of prostate cancer

Combination of metabolites is a suited practice to build panels of biomarkers with enhanced prediction capability as compared to individual compounds. The combination process carried out in an iterative manner is a better strategy than the selected combination of metabolites according to individual prediction capability

to take benefits from interactions among metabolites with additive effect on the discrimination performance. Metabolite panels were configured in this research by using a computational tool called Panelomix with the criterium of maximizing independently sensitivity and specificity, or global accuracy (sensitivity+specificity).

The list of all significant identified metabolites (Table 2) was considered as initial set of variables. A prefilter step was applied to identify the most frequent biomarkers and thresholds that potentially offer the most interesting combinations, and the resulting metabolites were iteratively combined in panels of five compounds with criteria described above. Table 3 lists the information associated to the three panels by combination of five metabolites. All the panels provided a positive response when 3 of the 5 metabolites gave a quantitative response below or above the corresponding threshold. It is worth mentioning that cut-off values set for metabolites of the panel corresponded to normalized concentration according to MSTUS for each polarity.

Panel A, which included acetylputrescine, dimethyllysine, imadazole lactate, tyrosine and citrulline, was targeted at maximizing accuracy. This panel provided 96.4% of specificity and 95.3% of sensitivity. Validation with the external set gave 64.3% and 79.0% of specificity and sensitivity, respectively. Fig. 2 shows the ROC curves obtained from these panels in comparison with those provided independently by the metabolites offering the best individual performance in each case. The response of panel A was compared to that of dimethyllysine, the marker with the highest predictive capability in terms of accuracy (85.7 % of specificity and 72.1 % of sensitivity). Both specificity and sensitivity were clearly improved when the panel of five compounds was considered. It is worth emphasizing that dimethyllysine, imidazole lactate and acetylputrescine, included in this panel, were the best compounds in the random forest test according to accuracy. To visualize the contribution of each compound in the panels, the data set were converted into percentage to plot Fig. 3 that shows the individual ROC and box-and-whisker plots

Table 3. Panels of five metabolites with the best predictive capability to discriminate PCa patients from control individuals. A) Panel with the best accuracy. B) Panel with optimal sensitivity. C) Panel with optimal specificity. D) Panel with the best accuracy including the PSA value.

Metabolites	A	B	C	D
	Without the serum PSA value			Including the serum PSA value
	Accuracy	*Optimal sensitivity*	*Optimal specificity*	*Accuracy*
	3+	3+	3+	3+
Acetylputrescine	< 0.0012			< 0.0012
Dimethyllysine	< 0.00066	< 0.00066	< 0.00066	< 0.00057
Imidazole lactate	< 0.001	< 0.001	< 0.0015	
Tirosine	< 0.0021	< 0.0039	< 0.0059	
Citrulline	< 0.0058		< 0.0029	< 0.0028
Dimethylarginine				< 0.004
Sialyl-N-acetyllactosamine		< 0.002		
Tiglylcarnitine			< 0.0035	
Urea		< 0.00029		
PSA				> 5
	MODEL WITH 70% OF SAMPLES			
% Sensitivity (95% CI)	95.3 (88.4–100.0)	95.3 (88.4–100.0)	88.4(79.1–97.7)	97.7 (93.0–100.0)
% Specificity (95% CI)	96.4 (89.3–100.0)	92.9 (82.1–100.0)	100 (100.0–100.0)	100 (100.0–100.0)
% AUC (95% CI)	95.3 (89.3–100.0)	* 2.4 (0.7–4.8)	*4.4 (4.0–4.9)	98.3 (94.9–100.0)
	EXTERNAL VALIDATION WITH 30% OF SAMPLES			
% Sensitivity	78.94	73.68	68.42	73.68
% Specificity	64.28	71.42	78.57	85.71
% Positive predictive value	75	77.77	81.25	87.5
% Negative predictive value	69.23	66.66	64.7	70.58
	APPLICATION TO 100% OF THE SAMPLES			
% Sensitivity	90.32	88.71	82.25	93.54
% Specificity	85.71	85.71	95.23	83.33
% Positive predictive value	90.32	90.16	96.22	89.29
% Negative predictive value	85.71	83.72	78.43	89.74

AUC: Area under the curve
* pAUC: Partial area under the curve

by considering the thresholds obtained by PanelomiX. The dashed lines pinpoint the quantitative response imposed to the panel for a positive response. As can be seen in Fig. 3A, the threshold established for tyrosine allowed classifying correctly 100% of negative biopsy cases. On the other hand, citrulline gave the best performance for PCa patients, as it classified correctly 93.3% of them.

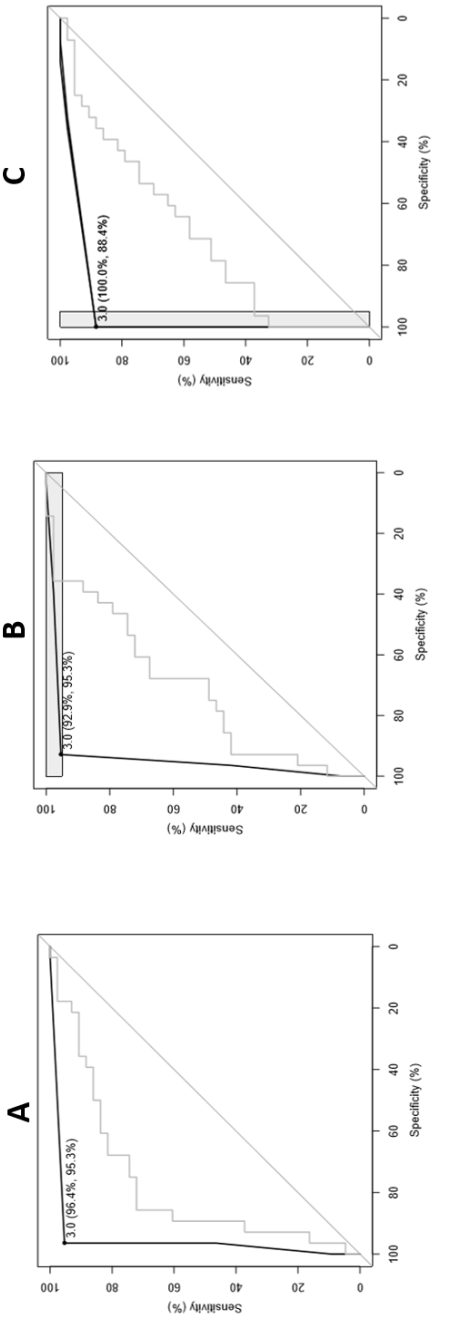


Figure 2. ROC curves of panels with five metabolites showing the best predictive capability to discriminate PCa patients from control individuals. A) Panel with the best accuracy, B) panel with optimal sensitivity, C) panel with optimal specificity.

Panel B, designed to maximize sensitivity, included dimethyllysine, imidazole lactate, tyrosine, sialyl-N-acetyllactosamine and urea. This panel provided 92.9%

of specificity and 95.3% of sensitivity, whereas the validation with the external set gave 71.4% and 73.7% of specificity and sensitivity, respectively. Despite panel B was designed to maximize sensitivity, this effect was not achieved by comparison to panel A; therefore, panel A is recommended when sensitivity is demanded. The parameters provided by histidine, the marker with the highest predictive capability in terms of sensitivity in the panel B, were 35.7% of specificity and 97.7% of sensitivity. Clearly, the panel improved the predictive accuracy as compared to histidine as individual marker, as Fig. 2 shows. Curiously, histidine was not included in this panel; however, imidazole lactate, a final product in the histidine metabolism, does. Fig. 3B illustrates the ROC curves and box-and-whisker plots of the five metabolites that constituted the panel with at least 95% sensitivity and the cut-off values imposed by the panel. As can be seen, sialyl-N-acetyllactosamine classified correctly 83.7% of PCa patients, followed by imidazole lactate and dimethyllysine with 74.4% and 72.2% of good performance for PCa patients, respectively. Again, the best specificity was provided by dimethyl-lysine.

Panel C, designed to maximize specificity, included dimethyl-lysine, imidazole lactate, tyrosine, citrulline and tiglylcarnitine. This panel provided 100% of specificity while sensitivity decreased to 88.4%. Validation with the external set gave 78.6% and 68.4% of specificity and sensitivity, respectively. As compared to panel A specificity is clearly enhanced from 64.3 to 78.6%. The parameters provided by citrulline, the compound with the highest predictive capability in this panel, in terms of specificity, were 96.4% of specificity and 37.2% of sensitivity. Therefore, the utility of the panel is proved. As Fig. 3C shows, citrulline classified correctly 96.4% of PCa patients, followed by tiglylcarnitine and dimethyl-lysine with a 92.9% and 95.7%, respectively. Tiglylcarnitine, a derivative of carnitine, is a new compound in this panel. Carnitine is the final product of lysine degradation, where dimethyllysine is an intermediate in the pathway.

It is worth mentioning that three metabolites (*viz.*, dimethyl-lysine, tyrosine and imidazole lactate) were present in the three panels, which reveals the incidence of the amino acids metabolism in the discrimination of PCa patients to improve

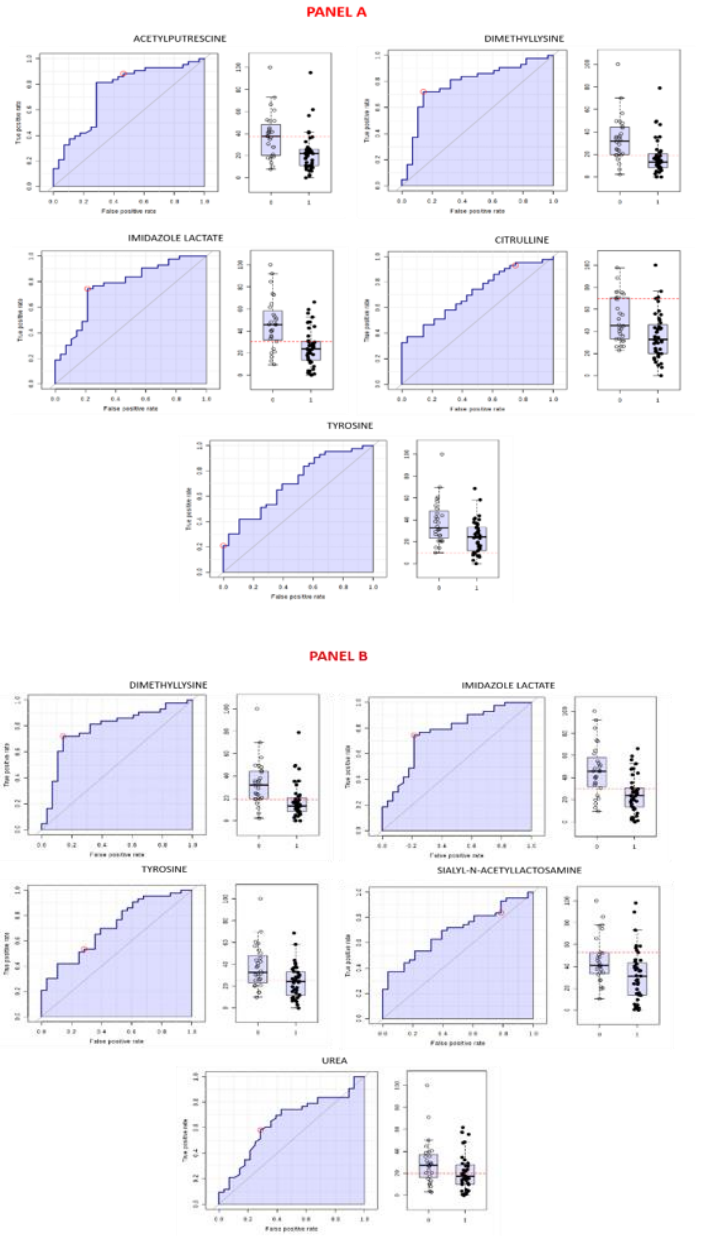


Fig. 3 It continues in page 348.

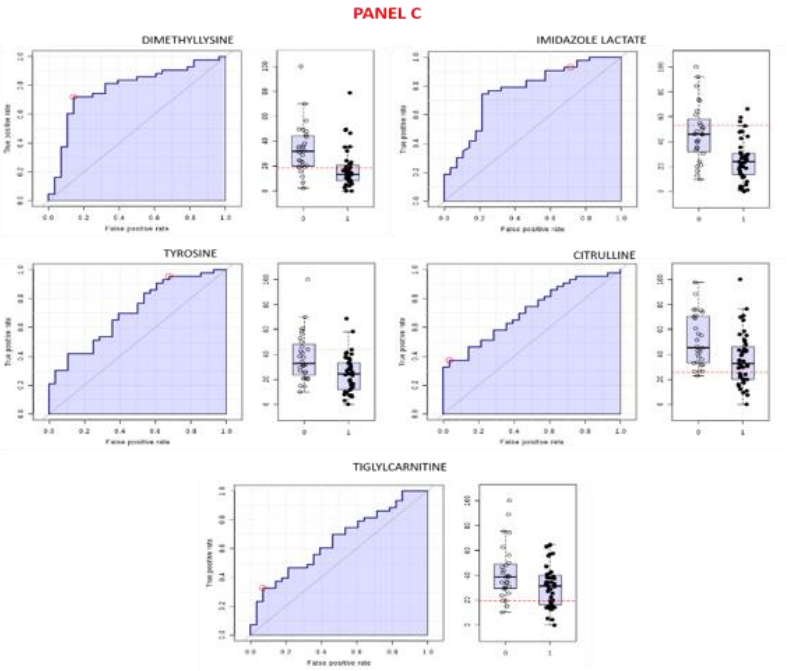


Fig. 3. The ROC curves and box-and-whisker plots of the five metabolites that constitute the panel with the best accuracy (A), with at least 95% specificity (B) and with at least 95% specificity (C) with distribution of quantitative responses for all individuals. The *y* axis in the box-and-whisker plots represents the normalized concentration of the metabolites expressed as percentage according to the samples with the highest and the lowest quantitative response. The dashed lines pinpoint the quantitative response of each marker to provide a positive response. (0=Controls, 1=PCa).

performance. Additionally, one of the most promising results is the specificity achieved by these panels. These results supported the combination of metabolites with PSA, a powerful sensitive marker, to configure a panel with high predictive capability to classify PCa patients and negative biopsy cases.

3.4. Configuration of biomarkers panels with PSA for detection of prostate cancer

The lack of specificity of PSA as a marker for PCa supported the configuration of a panel combining PSA and metabolites to complement the high sensitivity of PSA. For this purpose, panel D was designed to maximize the global accuracy by iterative combination of PSA serum levels with significant metabolites. As Table 3

shows panel D included PSA, acetylputrescine, dimethyllysine, citrulline and dimethylarginine to provide 100% of specificity and 97.7% of sensitivity. Validation with the external set gave 85.7 and 73.7% of specificity and sensitivity, respectively. The sensitivity and specificity provided by the PSA were 88.4% and 85.7%, respectively, as shows the corresponding ROC curve in Fig. 4A. However, these values were significantly improved by including in the panel the other 4 metabolites.

Among the three compounds included in the three previous panels, dimethyllysine, imidazole lactate and tyrosine, only dimethyllysine remained in this panel. Remarkable is the presence of citrulline, which improved the specificity in this study. As Fig. 4B shows, citrulline classified correctly 100% of negative biopsy controls, followed by dimethyllysine with an 89.3%. The high specificity of these metabolites allowed complementing the high sensitivity of PSA as screening marker for confirmatory analysis by biopsy. On the other hand, dimethylarginine, related to citrulline through the urea cycle, classified properly 97.7% of PCa patients.

Dimethyllysine and dimethylarginine are related to epigenetic marks on histone proteins. Histone methylation, mainly occurring at lysine and arginine residues, plays a key role in the regulation of chromatin conformation and, therefore, in its function [28]. Lysine methylation is dynamically regulated by protein lysine methyltransferases and demethylases. As dimethyllysine is down-regulated in the urine from PCa population, aberrant activities of histone demethylases or mutation into methyltransferases could promote the silencing of key oncosuppressor genes. Previous research has shown that methylation of H3 at lysine 4 is associated to inactive transcription of the PSA gene in the prostate cancer cell line LNCaP [29]. On the contrary, the transcription of the PSA gene mediated by the androgen receptor (AR) is accompanied by rapid decrease in di- and trimethylated H3 at lysine 4 [30].

Aberrant levels of histone modified arginine have also been linked to the development of several diseases including cancer [31]. In contrast to lysine methyl-

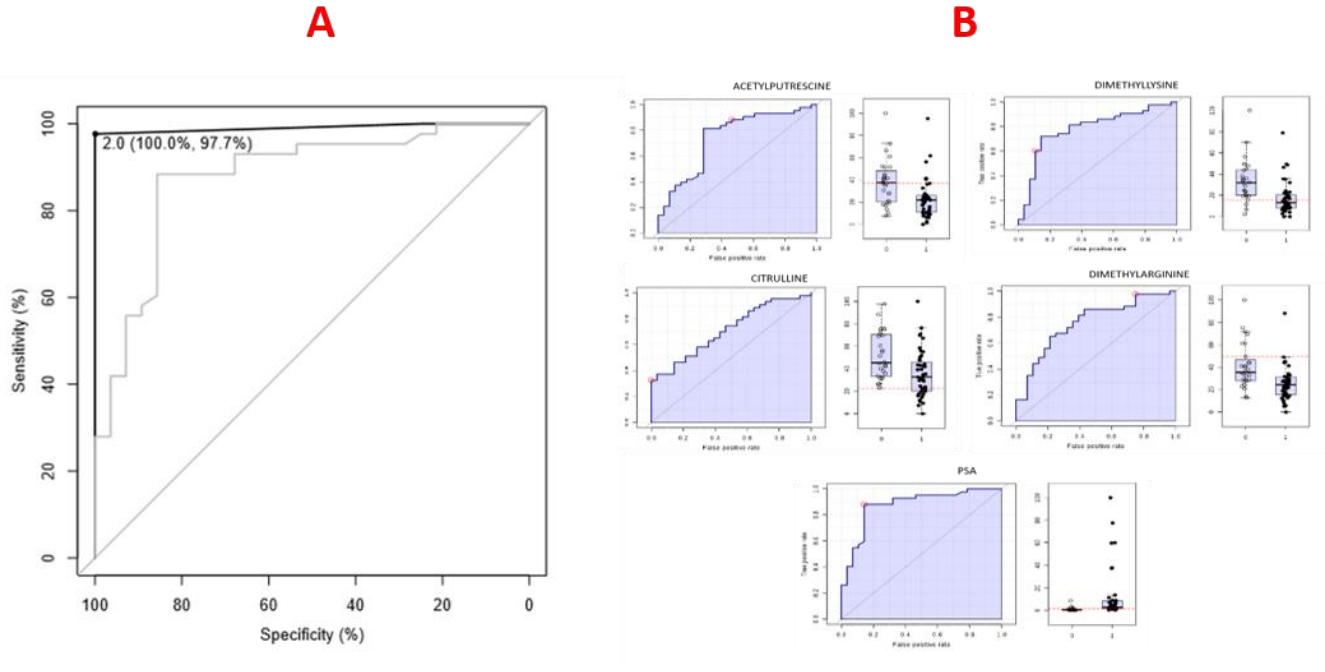


Fig. 4. (A) ROC curve of panel with five metabolites with the best accuracy including the PSA serum value. (B) ROC curves and box-and-whisker plots of the five metabolites that constitute the panel with the best accuracy including the serum PSA values. The *y* axis in the box-and-whisker plots represents the normalized concentration of the metabolites expressed as percentage according to the samples with the highest and the lowest quantitative response. The dashed lines pinpoint the quantitative response of each marker to provide a positive response. (0=Controls, 1=PCa).

ation, the extent to which arginine methylation is dynamic is much less clear. However, the reversibility of arginine methylation has recently been established through the discovery of histone arginine demethylases uncovering a new level of histone plasticity. This is the case with Jumonji domain–containing 6 protein (JMJD6) that demethylates histone H3 at arginine 2 (H3R2) and histone H4 at arginine 3 (H4R3) [32]. Since the relative concentration of dimethylarginine in the PCa group is lower than in the negative biopsy control group, histone arginine demethylases might be overexpressed in PCa individuals.

On the other hand, the observed existence of dimethylarginine dimethylaminohydrolase 1 (DDAH-1) in prostasomes derived from prostate cancer metastasis, as well as in urinary exoxomes could also be related to our results. DDAH-1 is a zinc-containing enzyme that, through hydrolysis of side-chain methylated L-arginine, regulates the activity of nitric-oxide (NO) synthase [33] and, accordingly, DDAH-1 plays a key role in NO metabolism.

Concerning citrulline, Dereziński *et al.* have recently determined the profile of amino acids in urine and serum in the search for PCa biomarkers [34]. They found significant differences in citrulline between the urine of patients with PCa and controls (p-value = 0.008), with the lowest concentration of citrulline in the cancer group, which is consistent with our results. The variation of citrulline in urine may be related to citrullination process. Citrullination is an additional post-translational modification (PTM) occurring in histones that involves the enzymatic deimination of peptidyl arginine residues yielding citrulline, mediated by a family of five enzymes known as protein arginine deiminases (PADs). For instance, PAD4 regulates histone arginine methylation levels via demethylimination, converting a methylarginine residue to citrulline at specific sites on the tails of histones H3 and H4 [31]. This activity is linked to repression of the estrogen-regulated pS2 promoter.

The dynamic appearance and disappearance of citrullination marks in histones suggests the existence of enzymes that might reverse the modifications. In fact, different studies have pointed out a transcriptional repression role to

citrullination through connection to deacetylation [35], as is the case with PAD4 that also associates with histone deacetylase 1 (HDAC1). The present study reports that both citrulline and acetylcitrulline were at significantly higher concentrations in healthy individuals than in PCa patients. A similar behavior was experienced by acetylarginine, as shows Table 2.

Acetylputrescine, the acetylated derivative of diamine putrescine, is a metabolic product of arginine that, together with the citrulline, participates in the urea cycle. This compound was found statistically significant, with lower levels in the PCa cohort. Shuster et al. published in 2013 a patent revealing higher levels of polyamines in PCa individuals, including acetylputrescine, which indicated that the level and/or activity of the enzyme ornithine decarboxylase was increased [36]. However, putrescine has been found at lower concentrations in prostate cancer tissue than in healthy tissue [37].

4. Conclusions

Using the urine of PCa patients and negative biopsy controls, we assessed several novel metabolite panels that showed good PCa detection capability. Based on the results from this study the panel combining dimethyllysine, dimethyl-arginine, acetylputrescine, citrulline and serum PSA value provided the best results in terms of sensitivity (97.7%) and specificity (100%), improving the prediction capacity of the PSA applied to the same cohort. The results of our diagnostic model suggest that urine metabolic profiling may have potential for clinical diagnosis of PCa. These preliminary results emphasize the necessity of a large-scale study to validate the proposed panels and obtain improve robustness. The use of these panels would tentatively reduce the proportion of individuals to be subjected to a biopsy and also to detect PCa in less advanced stages.

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", FQM-1602 and CTQ2015-68813-R). M. A. Fernández Peralbo and M. Calderón-Santiago also thank to these organisms their research contracts. E. Gómez Gómez thanks to The Carlos III Health Institute (ISCIII) and European Social Founds (ESF) his Rio Hortega research grant contract. The biological samples repository of Reina Sofía Hospital (Córdoba, Spain) is also gratefully acknowledged for coordination tasks in the selection of urine samples. Xavier Robin *et al.* are also acknowledged for providing the access to the PanelomiX toolbox.

References

[1] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman, F. Bray, Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, Int. J. Cancer 136 (2015) E359–386.

[2] R.L. Siegel, K.D. Miller, A. Jemal, Cancer Statistics, 2017, CA Cancer J. Clin. 67 (2017) 7–30.

[3] S. Saini, V. Affairs, C. Street, S. Francisco, PSA and beyond: alternative prostate cancer biomarkers, Cell Oncol. 39 (2016) 97–106.

[4] J.J. Tosoian, A.E. Ross, L.J. Sokoll, A.W. Partin, C.P. Pavlovich, Urinary biomarkers for prostate cancer, Urol. Clin. N. Am. 43 (2016) 17–38.

[5] R.S. Kelly, M.G. Vander Heiden, E. Giovannucci, L.A. Mucci, Metabolomic biomarkers of prostate cancer : prediction, diagnosis, progression, prognosis, and recurrence, Cancer Epidemiol. Biomarkers Prev. 25(6) (2016) 887–906.

[6] A. Sreekumar, L.M. Poisson, T.M. Rajendiran, A.P. Khan, Q. Cao, J. Yu, B. Laxman, R. Mehra, R.J. Lonigro, Y. Li, M.K. Nyati, A. Ahsan, S. Kalyana-

Sundaram, B. Han, X. Cao, J. Byun, G.S. Omenn, D. Ghosh, S. Pennathur, D.C. Alexander, A. Berger, J.R. Shuster, J.T. Wei, S. Varambally, C. Beecher, A.M. Chinnaiyan, Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression, Nature 457 (2009) 910–914.

[7] D.L. Cao, D.W. Ye, Y. Zhu, H.L. Zhang, Y.X. Wang, X.D. Yao, Efforts to resolve the contradictions in early diagnosis of prostate cancer: a comparison of different algorithms of sarcosine in urine, Prostate Cancer Prostatic Dis. 14 (2011) 166–172.

[8] V. Kumar, D.K. Dwivedi, N.R. Jagannathan, High-resolution NMR spectroscopy of human body fluids and tissues in relation to prostate cancer, NMR Biomed. 27(1) (2013) 80–89.

[9] T. Zhang, D.G. Watson, L. Wang, M. Abbas, L. Murdoch, L. Bashford, I. Ahmad, N.-Y. Lam, A.C.F. Ng, H.Y. Leung, Application of holistic liquid chromatography–high resolution mass spectrometry based urinary metabolomics for prostate cancer detection and biomarker discovery, PloS One 8 (2013) e65880.

[10] D. Theodorescu, E. Schiffer, H.W. Bauer, F. Douwes, R. Polley, T. Schmidt, W. Schöfer, P. Zürbig, M. Good, J.J. Coon, H. Mischak, Discovery and validation of urinary biomarkers for prostate cancer, Proteomics Clin Appl. 2 (2009) 556–570.

[11] H. Wu, T. Liu, C. Ma, R. Xue, GC/MS-based metabolomic approach to validate the role of urinary sarcosine and target biomarkers for human prostate cancer by microwave-assisted derivatization, Anal. Bioanal. Chem. 401 (2011) 635–646.

[12] Y. Miyagi, M. Higashiyama, A. Gochi, M. Akaike, T. Ishikawa, T. Miura, N. Saruki, E. Bando, H. Kimura, F. Imamura, M. Moriyama, Plasma free amino acid profiling of five types of cancer patients and its application for early detection, PLoS One 6 (2011) e24143.

[13] X. Zang, C.M. Jones, T.Q. Long, M.E. Monge, M. Zhou, D. Walker, R. Mezencev, A. Gray, J.F. Mcdonald, F.M. Fernandez, Feasibility of detecting

prostate cancer by ultra performance liquid chromatography–mass spectrometry serum metabolomics, Proteome Res. 13 (2014) 3444–3454.

[14] D. Cao, D. Ye, H. Zhang, Y. Zhu, Y. Wang, X. Yao, A multiplex model of combining gene-based, protein-based, and metabolite-based with positive and negative markers in urine for the early diagnosis of prostate cancer, Prostate 71 (2011) 700–710.

[15] M. Tessem, M.G. Swanson, K.R. Keshari, M.J. Albers, Z.L. Tabatabai, J.P. Simko, K. Shinohara, S.J. Nelson, B. Vigneron, I.S. Gribbestad, J. Kurhanewicz, Evaluation of lactate and alanine as metabolic biomarkers of prostate cancer using 1H HR-MAS spectroscopy of biopsy tissues, Magn. Reson. Med. 60 (2009) 510–516.

[16] E.E. Kline, E.G. Treat, T.A. Averna, M.S. Davis, A.Y. Smith, L.O. Sillerud, Citrate concentrations in human seminal fluid and expressed prostatic fluid determined via 1H nuclear magnetic resonance spectroscopy outperform prostate specific antigen in prostate cancer detection, J. Urol. 176 (2006) 2274–2279.

[17] N.J. Serkova, E.J. Gamito, R.H. Jones, C. O'Donnell, J.L. Brown, S. Green, H. Sullivan, T. Hedlund, E.D. Crawford, The metabolites citrate, myo-inositol, and spermine are potential age-independent markers of prostate cancer in human expressed prostatic secretions, Prostate 68 (2008) 620–628.

[18] J.E. McDunn, Z. Li, K.P. Adam, B.P. Neri, R.L. Wolfert, M. V Milburn, Y. Lotan, T.M. Wheeler, Metabolomic signatures of aggressive prostate cancer, Prostate 73 (2013) 1547–60.

[19] G.F. Giskeødegård, A.F. Hansen, H. Bertilsson, S.V. Gonzalez, K.A. Kristiansen, P. Bruheim, S.A. Mjøs, A. Angelsen, T.F. Bathen, M.B. Tessem, Metabolic markers in blood can separate prostate cancer from benign prostatic hyperplasia, Br. J. Cancer. 113 (2015) 1–8.

[20] M.L. Lynch, J.K. Nicholson, Proton MRS of human prostatic fluid: correlations between citrate, spermine, and myo-inositol levels and changes with disease, Prostate 255 (1997) 248–255.

[21] G. Sardana, E.P. Diamandis, Biomarkers for the diagnosis of new and recurrent prostate cancer, Biomarkers Med. 6 (2012) 587–596.

[22] M.J. Roobol, A. Haese, A. Bjartell, M.J. Roobol, A. Haese, A. Bjartell, Tumour markers in prostate cancer III: biomarkers in urine, Acta Oncologica 50 (2011) 85–89.

[23] M. A. Fernández-Peralbo, E. Gómez-Gómez, M. Calderón-Santiago, J. Carrasco-Valiente, J. Ruiz-García, M. J. Requena-Tapia, M. D. Luque de Castro, F. Priego-Capote, Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC–QTOF: upstream information on other omics, Scientific Reports 6 (2016) 6:38243.

[24] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J. Sanchez, M. Müller, PanelomiX: a threshold-based algorithm to create panels of biomarkers, Translant. Proteomics 1 (2013) 57–64.

[25] W. Struck-Lewicka, M. Kordalewska, R. Bujak, A. Yumba Mpanga, M. Markuszewski, J. Jacyna, M. Matuszewski, R. Kaliszan, M.J. Markuszewski, Urine metabolic fingerprinting using LC–MS and GC–MS reveals metabolite changes in prostate cancer: a pilot study, J. Phar. Biomed. Anal. 111 (2015) 351–361.

[26] S. Gamagedara, A.T. Kaczmarek, Y. Jiang, Validation study of urinary metabolites as potential biomarkers for prostate cancer detection, Validation study of urinary metabolites as potential biomarkers for prostate cancer detection, Bioanalysis 4(10) (2012) 1175–1183.

[27] Q. Liang, H. Liu, L. Xie, X. Li, A. Zhang, High-throughput metabolomics enables biomarker discovery in prostate cancer, RSC Adv. 7 (2016) 2587–2593.

[28] E.L. Greer, Y. Shi, Histone methylation: a dynamic mark in health, disease and inheritance, Nat. Rev. 13 (2012) 343–357.

[29] L.C. Li, P.R. Carroll, R. Dahiya, Epigenetic changes in prostate cancer: implication for diagnosis and treatment, JNCI 97 (2005) 103–115.

[30] J. Kim, L. Jia, W.D. Tilley, G. Coetzee, Dynamic methylation of histone H3 at lysine 4 in transcriptional regulation by the androgen receptor, Nucleic Acids Res. 31 (2003) 6741–6747.

[31] J. Fuhrmann, P.R. Thompson, Protein arginine methylation and citrullination in epigenetic regulation, ACS Chem. Biol. 11 (2016) 654–668.

[32] R.K.B. Bingsheng Chang, Yue Chen, Yingming Zhao, JMJD6 is a histone arginine demethylase, Science 318 (2007) 444–448.

[33] T. Lu, N. Kanzelmeyer, M.J. Kemper, D. Tsikas, A.M. Das, Developmental changes in the L-arginine/nitric oxide pathway from infancy to adulthood: plasma asymmetric dimethylarginine levels decrease with age, Clin. Chem. Lab. Med. 45(11) (2007) 1525–1530.

[34] P. Dereziński, A. Klupczynska, W. Sawicki, J.A. Pałka, Z.J. Kokot, Amino acid profiles of serum and urine in search for prostate cancer biomarkers: a pilot Study, Int. J. Med. Sci. 14 (2017) 9–13.

[35] R. Deplus, P. Putmans, M. Yamada, functional connection between deimination and deacetylation of histones, Mol. Cell. Biol. 29 (2009) 4982–4993.

[36] J.R. Shuster, M.W. Mitchell, J.E. McDunn, B.Neri, Biomarkers for prostate cancer and methods using the same, Pub. No.: US 2013/0217647 A1. (2013)

[37] G.F. Giskeødegård, H. Bertilsson, K.M. Selnæs, A.J. Wright, T.F. Bathen, T. Viset, J. Halgunset, A. Angelsen, I.S. Gribbestad, M.B. Tessem, Spermine and citrate as metabolic biomarkers for assessing prostate cancer aggressiveness, Plos One. 8 (2013) 1–9.

Supplementary information

Reagents

LC–MS grade acetonitrile (ACN) from Fisher Scientific (Fair Lawn, NJ, USA) and deionized water from a Millipore Milli-Q water purification system (Millipore, Bedford, MA, USA) were used to prepare the chromatographic mobile phases. LC–MS grade formic acid from Scharlab (Barcelona, Spain) was used as ionization agent in LC–MS/MS analysis. The pH of the samples was adjusted with ammonium formate from Fluka (Sigma–Aldrich, St. Louis, USA).

Apparatus and instruments

Centrifugation was carried out by a Thermo Sorvall Legend Micro 21 R thermostated centrifuge (Thermo Fisher Scientific, Bremen, Germany). The urine samples were analyzed by an Agilent 1200 Series LC system hyphenated to an Agilent 6540 QTOF with a dual electrospray ionization (ESI) source (Agilent Technologies, Santa Clara, CA, USA). Agilent MassHunter Workstation (version B05.01) was the software for data acquisition.

Sample preparation

The urine samples were thawed at room temperature, vortex-mixed for 1 min and centrifuged at 21000 × g for 5 min. Then, 50 µL of the supernatant was 1:2 (v/v) diluted with 5 mM ammonium formate in water (pH 5.5–7.5) prior to LC–MS/MS analysis.

Analysis by LC–QTOF

Chromatographic separation was performed by using a Mediterranea Sea C18 analytical column (50 × 4.6 mm i.d., 3 µm particle size) from Teknokroma

(Barcelona, Spain), thermostated at 25 °C. The initial mobile phase was a mixture of 98% phase A (0.1% formic acid in water) and 2% phase B (0.1% formic acid in ACN). After injection, the initial mobile phase was kept under isocratic conditions for 1 min; then, a linear gradient of phase B from 2% to 100% was applied within 16 min. The flow rate was 0.6 mL/min during the chromatographic step. The total analysis time was 17 min, and 5 min were required to re-establish the initial conditions. The sample volume injected in each ionization mode was 5 µL and the samples were randomly run. The injector needle was washed 5 times with 20:80 ACN–water between injections and the needle seat back was flushed with 20:80 ACN–water at 4 mL/min for 10 s to avoid cross contamination. The autosampler was kept at 4 °C to provide stability to the samples. Urine samples from several healthy volunteers were collected, mixed to obtain a pool, aliquoted and stored as quality control (QC) samples to be daily analyzed at the beginning of the sequence and after every 8 injections to ascertain that the mass spectrometer performance was stable during the analysis of the samples batch. Thus, 2 QCs were run per day following the protocol used for samples.

The settings of the electrospray ionization source, operated in the negative and positive ionization modes, were as follows: capillary voltage ±3.5 kV, fragmentor voltage 130 V, N_2 pressure in the nebulizer 40 psi, N_2 flow rate and temperature as drying gas 12 L/min and 325 °C, respectively. The procedures recommended by the manufacturer were used for calibration and tune of the instrument. The samples were analyzed in triplicate by setting data acquisition in centroid mode at 2 spectra per second in MS mode and 1 spectrum per second in MS^2 mode in the extended dynamic range mode (2 GHz). A gas-phase fractionation approach was used to increase the identification capability [1]. For this purpose, each of the three replicates from each sample was acquired with different ranges of precursor selection for MS/MS detection. The acquisition range in MS scanning mode was from 60 to 1200 *m/z* in all instances. The maximum number of precursors selected per cycle for MS/MS fragmentation was set at 2 with an exclusion window of 0.25 min after two consecutive selections of the same precursor ion. Criteria for selection of precursor ions were established for each of

the three runs per sample to obtain the maximum spectral information. Thus, in both ionization modes the mass range for precursor selection in the first run was set at 60–250 *m/z* with 10 eV as collision energy. The second and third runs involved precursor selection in the range 250–450 and 450–1200 *m/z* with collision energies of 20 and 30 eV, respectively. The instrument gave typical resolution 15000 FWHM (Full Width at Half Maximum) at *m/z* 118.0863 and 30000 FWHM at *m/z* 922.0098. To assure the desired mass accuracy of the recorded ions, continuous internal calibration was performed during analyses by using signals at *m/z* 121.0508 (protonated purine) and *m/z* 922.0097 [protonated hexakis (1H, 1H, 3H-tetrafluoropropoxy) phosphazine or HP-921] in positive ionization mode. In negative ionization mode, ions with *m/z* 119.0363 (purine anion) and *m/z* 1033.9881 (adduct of HP-921) were used.

Data mining

Raw data files were converted into mzData files using Mass Hunter Workstation software (version B.07.00 Qualitative Analysis, Agilent Technologies, Santa Clara, CA, USA); then, the data from each polarity were separately processed in R statistical language (version 3.1.3, http://www.r-project.org/) using the open-free XCMS (version 1.44.0) [2] and CAMERA (version1.24.1) R-packages. Processing of mzData files was performed by XCMS package to extract potential molecular features (MFs) considering only ions exceeding 1500 counts peak height with a peak width between 10 and 60 s, a signal-to-noise threshold of 10, and 10 ppm of error in mass accuracy. The adopted criterion was to consider only molecular entities detected in at least 75% of the samples belonging to at least one of the groups under study (PCa and control individuals). Then, the obiwarp method from “retcor” function was used to execute alignment of retention times across samples, grouping and integration. In the next step, the CAMERA package was used to correlate potential precursor ions as ([M–H]–, [M+Cl]– and [M+HCOO]– in negative ionization mode and as [M+H]+ and [M+Na]+ in positive ionization mode, as well as isotopic peaks from the same molecular entity.

Ions with identical elution profiles and related m/z values (representing different adducts or isotopes of the same compound) were grouped as a unique feature to remove redundant information. This process resulted in a data set for each polarity mode containing the peak area values in the apex of the chromatographic peaks for all molecular entities characterized by accurate mass and retention time (RT). The average area from the triplicate injection of each sample was calculated for each potential molecular feature.

The mass spectrometry total useful signal (MSTUS) method —that attempts to limit the contributions of xenobiotics and endogenous substances to the normalization factor by including only peaks that are present in all samples— was used for normalization of the data set from each polarity [3]. This normalization strategy considers the differences in volume of urine sampled per individual since metabolites would be more concentrated and give higher signal when the volume of collected urine is lower.

A *.csv* format was used to introduce the data matrix from each polarity into the Mass Profiler Professional (MPP) software (version 13.1, Agilent Technologies, Santa Clara, CA, USA) for statistical analysis. An unpaired *t*-test was executed to find metabolites significantly different between the two groups of patients, considering a *p*-value threshold of 0.05. Identification of the most relevant entities was supported on MS and MS/MS information by searching in the METLIN MS and MS/MS database (http://metlin.scripps.edu) and the Human Metabolome Database (HMDB, v. 3.6). An accuracy error of 5 ppm was set both in MS and MS/MS data to confirm the tentative identification of metabolites.

References

[1] M. Caldero´n-Santiago, F. Priego-Capote, M.D. Luque de Castro, Enhanced detection and identification in metabolomics by use of LC–MS/MS untargeted analysis in combination with gas-phase fractionation, Anal. Chem. 86 (2014) 7558–7565.

[2] R. Tautenhahn, C. Böttcher, S. Neumann, Highly sensitive feature detection for high resolution LC/MS, BMC Bioinformatic 9(504) (2008) 1–16.

[3] B.M. Warrack, S. Hnatyshyn, K.H. Ott, M.D. Reily, M. Sanders, H. Zhang, D.M. Drexler, Normalization strategies for metabonomic analysis of urine samples, J. Chromatogr. B 877 (2009) 547–52.

Discusión de los resultados

La normativa vigente en 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, establece que ha de incluirse una discusión conjunta de los resultados.

La investigación recogida en esta Memoria tiene como denominador común la metabolómica en el área clínica y se ha dividido en tres secciones en función del objetivo perseguido. En la Sección I, que comprende los Capítulos 1 y 2, se discuten las revisiones bibliográficas dedicadas a la preparación de la muestra de orina y a métodos de detección de componentes del condensado de aire exhalado (EBC), dos de los biofluidos investigados en esta Tesis.

La investigación experimental, que se recoge en las Secciones II y III, está dedicada a las dos estrategias básicas en metabolómica: análisis orientado ("*targeted analysis*") y análisis global ("*untargeted analysis*"), respectivamente. La Sección II engloba la investigación sobre la determinación de compuestos de interés clínico tales como los eicosanoides derivados de los ácidos omega-6 (Capítulo 3), implicados en la respuesta inflamatoria. La mejora en la cuantificación de una clase de ellos, los ácidos hidroxieicosatetraenoicos (HETEs), mediante el uso del estándar interno deuterado 15-HETEd8, constituye el Capítulo 4, que complementa al 3. Por otro lado, el desarrollo de un método para la cuantificación del fármaco quimioterapéutico paclitaxel y sus principales metabolitos permite el seguimiento de su efectividad y/o toxicidad, tal como recoge el Capítulo 5. La Sección III se dedica a la investigación sobre perfiles metabólicos ("*profiling analysis*") realizada con dos objetivos: superar retos quimiométricos como la reducción del tiempo que requiere el tratamiento de datos en espectrometría de masas (Capítulo 6) y abarcar la mayor parte posible del metaboloma de los biofluidos tratados en la Sección I. Para conseguir este último objetivo se ha optimizado la preparación de la muestra de EBC (Capítulo

7) y se ha llevado a cabo la búsqueda de posibles biomarcadores del cáncer de próstata (PCa) en orina (Capítulos 8 y 9). Los resultados obtenidos en cada una de las tres secciones se discuten después de comentar algunos aspectos instrumentales.

El uso de la espectrometría de masas como técnica de detección es común a las secciones experimentales. Las propiedades analíticas de esta técnica, en términos de sensibilidad, selectividad, exactitud, precisión y resolución, la convierten en la herramienta más útil en análisis clínico metabolómico, tanto orientado como global. Las dos estrategias típicas en metabolómica se benefician de esta técnica en sus diferentes modos, que son función del objetivo perseguido: análisis cualitativo o cuantitativo. Entre los diferentes tipos de espectrómetros de masas los más utilizados para análisis global son el QTOF y el Q-Orbitrap gracias a su alta resolución (exactitud en la medida de la relación *m/z*), selectividad y sensibilidad. Por estas razones son los dos detectores más adecuados para análisis cualitativo y también para cuantificación relativa. En esta investigación se ha utilizado un QTOF que, como ventaja adicional, presenta una alta velocidad de barrido. Por otra parte, el QqQ es tremendamente operativo en análisis orientado por su sensibilidad y selectividad para análisis confirmatorio, que lo convierte en el detector de masas óptimo con fines cuantitativos. Teniendo en cuenta la complejidad de las muestras clínicas, los espectrómetros de masas son muy adecuados en metabolómica clínica, especialmente cuando se acoplan a sistemas de separación cromatográficos. En este caso la cromatografía de líquidos en su modalidad de fase reversa ha sido también común a todas las secciones experimentales.

Sección I: Revisiones bibliográficas

La investigación sobre la documentación bibliográfica realizada por la doctoranda durante los primeros años de desarrollo de su Tesis alertó sobre el escaso aprovechamiento de la orina y el aire exhalado como muestras en investigación clínica, a pesar de las ventajas que ofrecen, ya que, además de un muestreo fácil y no invasivo, hacen previsible la obtención de información

interesante sobre patologías con las que están relacionadas. Es el caso de dos de los cánceres con mayor incidencia en la población, el de próstata en hombres y el de pulmón, que han sido objeto de estudio en esta Tesis Doctoral. El limitado uso de estas muestras se debe a la escasa información sobre el tratamiento adecuado; por lo que se ha llevado a cabo una exhaustiva revisión bibliográfica del estado actual de la preparación de las muestras de orina para su estudio metabolómico, así como de los sistemas disponibles para la detección del cáncer de pulmón basados en EBC. Los Capítulos 1 y 2, que constituyen la Sección I de la Tesis, abarcan estos estudios bibliográficos.

En el Capítulo 1, publicado con el título "Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis" se discuten las etapas de pretratamiento y preparación como tal de la muestra de orina previas a la separación individual y/o detección, poniendo de manifiesto la atención poco rigurosa que han merecido estas etapas hasta la fecha. Se hace hincapié en el diferente uso de estas etapas en función de la estrategia metabolómica elegida, aconsejando que la muestra permanezca lo más intacta posible empleando el mínimo número de subetapas en análisis globales del metaboloma, ya que no se conoce el comportamiento de los compuestos que comprende, mientras que en análisis dirigido estas etapas deberán ser lo más selectivas posible para los compuestos de interés.

Las etapas que comprende el pretratamiento de la orina, desde el muestreo hasta su almacenamiento, pasando por el tipo de dispositivo para la recolección, inhibición del metabolismo, adición de conservantes, normalización del volumen, ajuste de pH y desproteinización, se han discutido exhaustivamente analizando los pros y los contras de cada etapa sin perder de vista en ningún momento el objetivo del análisis.

La preparación de la muestra propiamente dicha en el caso de la orina constituye la segunda parte de la revisión. A pesar de que esta etapa engloba todas las subetapas durante las cuales los analitos se adecuan para la detección, el hecho de que los detectores utilizados en metabolómica generalmente estén

conectados en línea con un equipo de separación (cromatógrafo o, menos comúnmente, equipo de electroforesis capilar), hace que se suela considerar terminada la etapa de preparación de la muestra cuando ésta (o lo que de ella queda tras su preparación) llega al detector o al sistema de separación, por lo que recibe el nombre de "muestra analítica". Las etapas que puede comprender la preparación de la muestra de orina son: hidrólisis enzimática, limpieza de la muestra y preconcentración de analitos —mediante extracción en fase sólida o líquida principalmente—, evaporación o liofilización de la muestra y adición de estándar interno.

Las diferentes opciones que existen para llevar a cabo cada etapa se han discutido y la idoneidad de una(s) frente a otra(s) se justifica a través de los estudios de optimización. Las subetapas que son esenciales, aconsejables o innecesarias también se plantean en esta publicación sugiriendo algunas propuestas emergentes como la hidrólisis asistida por ultrasonidos para acelerar el proceso. Los errores a que da lugar un mal desarrollo de estas etapas se enfatizan en la publicación, donde también se discuten las formas de realizarlas para evitar o minimizar los errores. Especial hincapié se hace en las subetapas que pueden provocar supresión de la ionización como la adición de surfactantes para la conservación de la orina y la evaporación o liofilización, aconsejando el uso de estándares internos marcados con isótopos estables para compensar estos efectos indeseados.

La segunda parte de la Sección I, el Capítulo 2, contiene la revisión publicada con el título "Analytical methods based on exhaled breath for early detection of lung cancer". Se pone de manifiesto en ella la alta mortalidad causada por este tipo de cáncer como consecuencia de la ausencia de síntomas en estadios tempranos y la limitada capacidad de las técnicas de diagnóstico actuales, que dan lugar a un elevado número de falsos positivos que derivan en biopsias innecesarias.

Se describen en esta publicación las características y la composición del EBC, comparándola con los componentes orgánicos volátiles (VOCs). Se enfatiza

el hecho de que a pesar de que su uso en métodos de cribado de enfermedades respiratorias ha aumentado exponencialmente, existen aún pocos estudios sobre el cáncer de pulmón utilizando EBC como muestra.

Los métodos de detección del cáncer de pulmón mediante perros entrenados aprovechando su desarrollado olfato, así como los métodos basados en otras grandes ómicas como la genómica y la proteómica y en el uso de sensores, se discuten en esta publicación. Se sugiere, asimismo, la aportación que supondría la investigación basada en el binomio químicos analíticos–metabolómica. Se razona que si los perros detectan el cáncer de pulmón en el aire exhalado condensado incluso en estadios tempranos, es decir, cuando existen muy bajas concentraciones de los biomarcadores de la enfermedad, la instrumentación analítica de última generación podría aportar la suficiente sensibilidad para el estudio y análisis de estos biomarcadores. Este logro eliminaría las limitaciones del uso de perros para la detección del cáncer de pulmón, tales como la detección de compuestos no relacionados con cáncer (consecuencia de inflamación, necrosis, tabaquismo, etc.), o por no ser selectivos para el tipo de cáncer en concreto, por lo que es fundamental la identificación de estos compuestos. La mayoría de los componentes del EBC son metabolitos, de forma que se hace hincapié en el uso de la metabolómica para obtener una información más racional, fiable y amplia que permita el establecimiento de bases que faciliten la detección de la enfermedad. La tecnología analítica que actualmente se utiliza en metabolómica constituye una excelente herramienta para tratar de alcanzar los límites de detección atribuibles al olfato de los perros. Asímismo, un adecuado tratamiento de la muestra ayudaría a conseguir más información de su contenido; lo que constituye la investigación que recoge el Capítulo 7.

Se concluye que un adecuado tratamiento de la muestra y la elección acertada del sistema de detección permiten una correcta interpretación biológica de los resultados. Por ello, esta Sección I trata de poner de manifiesto qué se ha conseguido hasta ahora y qué retos hay que lograr para consolidar el uso de la orina y del EBC en investigación clínica para la detección y seguimiento de enfermedades, lo que resultó clave para el desarrollo de los siguientes capítulos

de la Tesis, además de ofrecer una visión que puede ser de utilidad para bioquímicos, clínicos u otros investigadores no familiarizados con la temática.

Sección II: Contribuciones al análisis metabolómico dirigido

Se define el análisis orientado en metabolómica como aquél que se aplica a una muestra para la cuantificación de un grupo o familia de compuestos, generalmente los precursores y los metabolitos de unas determinadas rutas metabólicas. A pesar de que sólo se aplica a compuestos predefinidos esta estrategia posee una serie de ventajas respecto al análisis metabolómico global. La principal de ellas es que tanto la preparación de la muestra como la detección de los analitos de interés se optimizan para reducir el número de interferentes y maximizar la selectividad y la sensibilidad mediante la adecuada elección de la plataforma analítica.

La preparación de la muestra depende de la naturaleza de los analitos, siendo la extracción líquido–líquido (LLE) y la extracción en fase sólida (SPE) las dos alternativas más usadas para eliminar interferentes y preconcentrar los metabolitos problema. En concreto la investigación recogida en los Capítulos 3 y 4 se ha beneficiado de las ventajas de la SPE automatizada y acoplada en línea con etapas posteriores del proceso analítico, utilizándose la LLE en la investigación que ha dado lugar al Capítulo 5.

Cuando el número de compuestos es suficientemente pequeño para optimizar la preparación de la muestra y la detección de cada uno de ellos y se puede obtener su recta de calibrado individual —porque la existencia de patrones lo permite— se consigue la cuantificación de todos los compuestos problema. La inclusión de un estándar interno adecuado mejora sustancialmente la exactitud en la cuantificación; lo que se justifica en el Capítulo 3.

Los paneles de compuestos seleccionados en esta investigación para análisis orientado estuvieron formados tanto por metabolitos implicados en rutas metabólicas conocidas y que juegan un papel relevante en ciertas patologías, como en

el caso de un fármaco utilizado en la terapia del cáncer y sus principales metabolitos cuya determinación es clave para el seguimiento de su eficiencia y/o toxicidad.

El desarrollo de un método siempre ha de completarse con su validación y aplicación a muestras reales que pongan de manifiesto su utilidad. Los métodos desarrollados en esta sección han sido adecuadamente validados a través de la evaluación de sus características analíticas y aplicados a muestras reales procedentes de cohortes con alguna patología como la formada por pacientes diagnosticados con diversas afecciones cardiacas en el Capítulo 4 y mujeres con carcinomatosis peritoneal ovárica en el Capítulo 5.

Capítulos 3 y 4

El Capítulo 3 se dedica al estudio de una familia de eicosanoides, los HETEs, para evaluar la inclusión de un estándar interno estable marcado isotópicamente, el 15HETEd8 (también usado en el Capítulo 4), dando como resultado la publicación "Stable isotopic internal standard correction for quantitative analysis of hydroxyeicosatetraenoic acids (HETEs) in serum by on-line SPE–LC–MS/MS in selected reaction monitoring mode". Por otro lado, el desarrollo de un método totalmente automatizado para la determinación de eicosanoides derivados de ácidos grasos omega-6 (ácido araquidónico y linoleico), en concreto prostaglandinas (PGs), ácidos hidroxioctadecadienoicos (HODEs), (HETEs) y tromboxano (TX) en suero para el estudio de enfermedades arteriales coronarias fue el objetivo del Capítulo 4, publicado con el título "Targeted analysis of omega-6-derived eicosanoids in human serum by SPE–LC–MS/MS for evaluation of coronary artery disease". Ya que se utilizó la misma plataforma analítica y es común la cuantificación de 8,11,12, y 15HETE con la ayuda de 15HETE-d8, se realiza la discusión conjunta de los resultados de ambos capítulos.

Los HETEs desempeñan un papel regulador en la respuesta inflamatoria estando asociados a diferentes patologías como enfermedades cardiovasculares, cáncer y osteoporosis, por lo que es extremadamente útil un método que permita

determinarlos lo más exactamente posible en un rango de concentraciones que abarque los valores extremadamente bajos a los que se encuentran en fluidos humanos. Además, la complejidad de las muestras biológicas demanda adecuados protocolos para el tratamiento de muestra que eliminen potenciales interferentes antes de la separación cromatográfica. Con este propósito, se adoptó la plataforma SPE–LC–MS/MS, donde la SPE totalmente automatizada permite eliminar las atapas de precipitación de proteínas y de evaporación, disminuyendo la pérdida y degradación de analitos a la vez que se reduce el tiempo de análisis.

El número de metabolitos objetivo permitió la optimización completa de las variables que afectan tanto a la detección para cada compuesto (voltaje, energía de colisión e iones producto) como a la extracción (tipo de sorbente, disolventes, volumen y flujo de la carga, lavado y elución), así como la caracterización del método desarrollado: curvas de calibrado, rango dinámico, límites de detección y de cuantificación, repetitividad, reproducibilidad y exactitud, que fueron notablemente mejoradas con la inclusión de un estándar interno deuterado. Asimismo, la comparación con métodos anteriores basados en LLE y SPE fuera de línea para los mismos analitos reveló una mejora de la sensibilidad usando la configuración en línea.

El estudio del uso de un HETE deuterado como estándar interno se consideró de interés para asegurar la calidad de los resultados respecto a los compuestos de este grupo. Para ello, y en las condiciones de trabajo óptimas, se desarrollaron 4 modelos de calibración, en un pool de suero, con disoluciones del estándar y aplicando o no la corrección con el estándar interno. Se observó un considerable efecto matriz en los modelos de calibración preparados en suero al compararlos con los obtenidos con disoluciones estándar, consiguiéndose un notable efecto de corrección con el uso del estándar interno deuterado. Esta corrección fue mayor para el 15HETE y el 12HETE que para el 8HETE y el 11HETE, lo que se explica por la semejanza química de los dos primeros con el 15HETE-d8, ya que el grupo hidroxilo está más alejado de la posición 15 en el 11HETE y en el 8HETE. La validación del modelo derivó en varios estudios, como la evaluación de la

eficiencia en la retención de los analitos en el cartucho mediante un sistema de doble cartucho, obteniéndose una excelente retención, por encima del 90%, para todos los analitos. La estimación de la exactitud, crucial para evaluar el efecto matriz, se calculó mediante el factor de recuperación a tres niveles de concentración en suero con y sin corrección mediante el estándar interno. La cuantificación del 8HETE, el 11HETE y el 15HETE se apoyó adecuadamente en los parámetros analíticos a concentraciones medias de ellos (35 y 55 ng/mL), mientras que el análisis del 12HETE proporcionó un bajo factor de recuperación a los tres niveles de concentración estudiados. Por otra parte, la cuantificación a baja concentración de eicosanoides se caracterizó por una disminución de los factores de recuperación por debajo del 78.5%. Por lo tanto, la implementación de factores de corrección es necesaria para la cuantificación exacta en rangos de concentración específicos.

Una vez optimizado y validado el método, se aplicó al análisis de muestras de suero procedentes de 45 pacientes ateroscleróticos (15 con angina estable, 15 con infarto de miocardio sin elevación del segmento ST —NSTEMI— y 15 que habían sufrido un infarto agudo de miocardio) y 16 controles. Aunque el método permitió un análisis cuantitativo de 12 metabolitos, sólo 5 eicosanoides estuvieron por encima del límite de cuantificación en un número representativo de muestras.

Un análisis univariante ANOVA puso de manifiesto la significancia del tromboxano (TX) y el 11HETE (producidos por la acción catalítica de cicloxigenasas —COXs) en los individuos control en comparación con los pacientes con lesión coronaria. Además, la concentración relativa de ambos compuestos se evaluó mediante su magnitud de cambio, obteniéndose que el TX y el 11HETE estaban 14.66 y 1.18 veces aumentados en el grupo control, por lo que en realidad el TX fue el único que experimentó un cambio muy significativo entre ambos grupos. Aunque el TX actúa como un potente agente de agregación plaquetaria, este resultado puede explicarse porque la mayoría de los pacientes tomaron aspirina, que inhibe la actividad enzimática de las COXs, reduciendo los niveles de TX en sangre con un efecto plaquetario antiagregante, incluso a dosis bajas. A pesar de

la inhibición de las COXs por la ingestión de aspirina en los pacientes lesionados, los HETEs y los HODEs se generan por acción de otras dos enzimas (lipoxigenasas —LOXs— y citocromo P450 —CYP) que actúan sobre el ácido araquidónico y el ácido linoleico.

Los niveles séricos de los eicosanoides en los pacientes diagnosticados con enfermedad coronaria también se examinaron estadísticamente según la gravedad de la lesión coronaria estratificada como angina estable, NSTEMI e infarto de miocardio agudo (AMI). El 12HETE fue el único metabolito significativo para explicar las diferencias entre los individuos que sufrieron un episodio de AMI y los diagnosticados con angina estable con una significación del 95%. Sintetizado por la acción enzimática de una LOX, la concentración de este compuesto no se ve afectada por el tratamiento con aspirina, que se encontró 2.5 veces más concentrado en pacientes que habían sufrido AMI; lo que corrobora estudios previos que reportan la formación de 12HETE durante la isquemia. La significancia del 12HETE en AMI frente a angina estable fue además evaluada mediante curvas ROC estimando el área parcial bajo la curva restringida entre un 90 y 100% de especificidad, dando como resultado un 90%.

Merece destacarse de estos estudios que los niveles de 12HETE presentaron diferencias significativas en los pacientes que habían sufrido AMI en comparación con los pacientes con angina estable. Estos resultados pueden ayudar a diagnosticar y prevenir un posible infarto agudo de miocardio en pacientes ya lesionados. Para su cuantificación, una plataforma que conecte en línea la SPE con el equipo LC–MS/MS se presenta como la elección óptima siendo necesario el uso de un adecuado estándar interno para conseguir una buena exactitud.

Capítulo 5

Otra aplicación de la metabolómica en el área clínica se desarrolló en la investigación que dio lugar a la publicación “LC–MS/MS quantitative analysis of paclitaxel and its major metabolites in serum, plasma and tissue from women with ovarian cancer after intraperitoneal chemotherapy”.

Los efectos secundarios derivados del uso de un metabolito exógeno como el antineoplásico paclitaxel requirieron su cuantificación junto con la de sus principales metabolitos: el 6α-hidroxipaclitaxel (6α-OHP) y el *p*-3´-hidroxipaclitaxel (C3´-OHP). El enfoque farmacometabolómico permite, en este caso, estudiar la farmacocinética del placlitaxel con el objetivo de optimizar dosis y tiempos de exposición de las terapias y evitar de esta forma la toxicidad, ayudando a personalizar la medicina.

El presente estudio farmacológico se apoyó en la plataforma LC–MS/MS con modo de seguimiento selectivo de reacciones (SRM) para cuantificar simultáneamente paclitaxel, 6α-OHP, C3´-OHP, utilizando docetaxel como estándar interno. La optimización de los parámetros del detector y de la fuente de ionización ESI para los metabolitos y para el estándar interno condujeron a utilizar ácido fórmico al 0.1% como agente ionizante y a la selección de aductos $[M+Na]^+$ como iones precursores tras observar que los principales iones productos derivaban ellos.

La etapa de preparación de las muestras (en este caso suero y tejido) se optimizó utilizando un pool de cada una de ellas tomado antes de la quimioterapia, tras asegurarse de que no existían valores detectables de los analitos objetivo. Al observar una baja sensibilidad usando una etapa de precipitación de proteínas, se eligió una simple LLE con metil *ter*butil éter (MTBE) para eliminar interferentes y preconcentrar los analitos objetivo mediante evaporación y reconstitución en metanol. Se evaluaron las características analíticas del método, consiguiendose límites de detección y cuantificación menores en suero que en tejido para todos los analitos y siendo ambos límites mayores para el 6α-OHP que para el paclitaxel y el C3´-OHP. La precisión y exactitud del método para el tejido y el suero estuvieron en unos valores aceptables, corroborando así la capacidad del método para la cuantificación de los compuestos en estudio.

El método se aplicó a muestras de suero y tejido procedentes de 13 mujeres diagnosticadas con carcinomatosis peritoneal por cáncer epitelial de ovario tratadas con quimioterapia intraperitoneal hipertérmica (HIPEC), así como al

plasma de una de ellas. La toma de muestra se llevó a cabo a diferentes tiempos: antes, durante y tras el tratamiento con HIPEC para conocer la evolución del fármaco en la aplicación postquimioterapia. Así, el suero y el plasma se tomaron inmediatamente después del tratamiento con HIPEC y una hora después, mientras que la biopsia del tejido se realizó antes y 1 h después del tratamiento con el fármaco.

El paclitaxel y sus metabolitos no se detectaron en las muestras extraídas antes del tratamiento de quimioterapia ni sus metabolitos en el tejido después de la quimioterapia, un hecho esperado ya que el paclitaxel apenas se metaboliza en el tejido peritoneal. Sin embargo, la concentración de este fármaco en el tejido fue, en todos los casos, significativamente mayor que en el suero. En la mayoría de los casos, la concentración de paclitaxel en suero había disminuido una hora después del tratamiento, metabolizado preferentemente a 6α-OHP puesto que la concentración de este metabolito aumentó, siendo significativamente superior a la de C3´-OHP una hora después del tratamiento con HIPEC.

Cabe destacar que la idoneidad del método para la determinación de los analitos en la muestra de plasma de una de las pacientes fue patente ya que las diferencias en la concentración de paclitaxel en ambas muestras de sangre (suero y plasma) y a los dos tiempos de muestreo (inmediatamente después del tratamiento y 1 h más tarde) estuvieron dentro del error experimental del método.

Sección III: Contribuciones al análisis metabolómico global

La obtención de información lo más completa posible sobre la naturaleza de los metabolitos existentes en la muestra constituye el objetivo principal del análisis metabolómico global. Para su desarrollo se requieren analizadores de masas de alta resolución que permitan conseguir una buena exactitud en la medida de la relación *m/z* que haga posible la identificación del máximo número de metabolitos detectados en el biofluido en estudio. Equipos QTOF o Q–Orbitrap son los preferidos en estos casos puesto que permiten obtener espectros

MS/MS en alta resolución de los iones precursores detectados que, junto con una técnica de separación previa, normalmente cromatográfica, proporcionan la separación adecuada para la identificación inequívoca de los metabolitos. Las herramientas quimiométricas y bioinformáticas son claves en este tipo de análisis no dirigidos para extraer información del a veces inmenso volumen de datos, de una forma lo más rápida posible y representativa del problema en estudio, además de para sacar conclusiones fiables de los resultados obtenidos. La información biológica disponible en la HMDB y las rutas biosintéticas de la base de datos 'Kyoto Encyclopedia of Genes and Genomes' (KEGG) han sido de gran utilidad para la interpretación de los resultados obtenidos. En este bloque de la parte experimental se utilizó un equipo QTOF que existe en el laboratorio del grupo de investigación. Hubiera sido muy interesante comparar los resultados obtenidos con un equipo Q–Orbitrap, pero no se pudo concretar una colaboración para tal fin.

Se persiguieron 4 objetivos mediante esta estrategia metabolómica, que constituyen los 4 capítulos que conforman esta sección: mejorar y acelerar el tratamiento de los datos generados mediante LC–QTOF, estudiar el tratamiento de una muestra biológica poco estudiada como es el EBC y conseguir información adecuada de las muestras de orina de enfermos de PCa para finalmente desarrollar paneles de posibles biomarcadores. Los resultados obtenidos se discuten a continuación.

Capítulo 6

Aunque en la literatura se encuentra un gran número de estrategias de análisis de datos, ninguna de ellas puede elegirse como la opción óptima en todos los casos, lo que hace del análisis de datos un área activa de investigación en bioinformática. El artículo "MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes" que constituye este Capítulo, es un ejemplo de ello.

LC–QTOF en modo MS/MS es una de las mejores plataformas para obtener información completa de una amplia gama de metabolitos presentes en una muestra ya que combina dos modos de ionización, positivo y negativo, cuya elección en cada caso viene dictada por la estructura de los metabolitos. La decisión, sin embargo, se adopta tras un procesamiento por separado de los datos obtenidos en cada polaridad, lo que requiere un tiempo considerable. Aunque una fracción significativa de metabolitos se detecta preferentemente en un modo de ionización, muchos compuestos pueden detectarse en ambos, generando información redundante. Para superar estas limitaciones se utilizó el software de análisis de datos R para desarrollar los algoritmos que conforman el paquete denominado MSCombine. Esta herramienta combina las matrices de datos generados en ambos modos de ionización tras eliminar las entidades moleculares comunes y conservar las detectadas con mayor sensibilidad, resultando así una única matriz más manejable y representativa de la muestra biológica en estudio.

Se utilizaron diferentes algoritmos para conseguir este propósito, como la correlación a través de la función *Findcommon* de ambas matrices con una tolerancia del valor de masa de 0.002 Da y 0.2 min en el tiempo de retención, generando el archivo *CommonEntities*. A través de la función *RemoveMismatch* se eliminan iones asociados a la misma identidad molecular, decantándose por el ion más común que en caso de duda o frecuencia parecida será $[M+H]^+$ o $[M-H]^-$ en ionización positiva o negativa, respectivamente. En el tercer paso la función *StudyRTdiff* construye un modelo de regresión que compara las dos listas de tiempos de retención de ambas polaridades para establecer un límite de corte —en este caso ± 0.2 min— filtrando la matriz con la función *FilterbyRT*, que elimina entidades no coincidentes y confirma las entidades asignadas al mismo metabolito. Este paso es crucial para obviar la falta de coincidencia de los picos con isómeros eluidos a tiempos de retención cercanos. El último paso (función *CombinePolarities*) compara el área media de las entidades moleculares detectadas en ambas polaridades que se han asociado al mismo metabolito, eliminando las entidades detectadas con menor sensibilidad, además de suprimir otras

entidades asociadas al mismo metabolito detectado en el mismo modo de ionización.

La funcionalidad del paquete MSCombine se evaluó utilizando conjuntos de datos reales obtenidos del análisis de muestras de suero y orina para hacer frente a diferentes situaciones de la vida real (compuestos que solo tienen aductos, isómeros a tiempos de retención similares). El 27 y 41% de las entidades potenciales detectadas en el modo de ionización positivo y negativo, respectivamente, correspondieron a información redundante en suero, mientras que en orina el 26 y el 31% de las entidades potenciales detectadas en el modo positivo y negativo, respectivamente, también se detectaron en el otro modo.

MSCombine se validó también mediante el análisis estadístico de los datos obtenidos del análisis de la cohorte de orina que consistieron en 787 y 664 entidades moleculares en los modos positivo y negativo, respectivamente. La combinación de las dos matrices condujo a otra que constaba de 1451 entidades, y la aplicación del paquete la redujo a 1245. El análisis estadístico a través de un diagrama de Volcano ($p < 0.01$, cambio de magnitud > 2) proporcionó 41 entidades significativamente diferentes entre los grupos (controles e individuos que tomaron ciprofloxacin) en el conjunto de datos proporcionados por MSCombine, y 54 basado en el conjunto de datos obtenidos mediante la combinación de las matrices de los modos de ionización positiva y negativa. Las 13 características moleculares ausentes del conjunto de datos MSCombine correspondieron a entidades eliminadas como entidades redundantes. De hecho, la identificación tentativa de esas 13 entidades reveló que estaban asociadas a 10 metabolitos ya incluidos en la matriz MSCombine, confirmándose la utilidad del paquete estadístico creado.

Capítulo 7

"Study of exhaled breath condensate sample preparation for metabolomics analysis by LC–MS/MS in high resolution mode", es el artículo que ha derivado en el Capítulo 7. Como ya se comentó en el Capítulo 2, el EBC es un biofluido de

muestreo fácil en condiciones normales, de bajo coste y no invasivo, que al ser representativo del líquido que recubre las vías respiratorias hace que su análisis pueda conducir a identificar posibles biomarcadores de diversas enfermedades respiratorias, siendo una excelente alternativa a actuales técnicas incómodas e invasivas. Sin embargo, el mayor porcentaje de EBC está constituido por agua, por lo que la búsqueda de tratamientos adecuados que preconcentren sus componentes y los limpien de posibles interferentes es el reto que se ha considerado en este estudio.

La mayoría de los estudios sobre análisis de este biofluido se han centrado en compuestos concretos. Un enfoque no dirigido permite en este caso detectar los metabolitos resultantes tras aplicar diferentes tratamientos de la muestra y dilucidar de esta manera si uno de ellos destaca sobre los demás en función del número de compuestos detectados, o cual es el más idóneo para detectar/cuantificar algunos metabolitos en concreto. Para obtener un perfil de metabolitos complementario al que puede ofrecer la plataforma GC–MS, se ha vuelto a usar la plataforma LC–QTOF para el estudio de posibles compuestos no volátiles y semivolátiles en EBC.

Se evaluaron tres alternativas para el procesamiento de un pool de EBC procedente de 4 voluntarios sanos antes del análisis global, que consistieron en: (i) precipitación de proteínas (PP); (ii) SPE adaptada a un formato miniaturizado con centrifugación (debido a los pequeños volúmenes de muestra habitualmente disponibles) con dos sorbentes complementarios con interacciones lipofílicas (C18) e hidrofílicas (HI); (iii) liofilización del EBC.

Un estudio cualitativo de los resultados mediante diagramas de Venn considerando sólo las entidades tentativamente identificadas por MS/MS permitió comparar el número de entidades detectadas tras cada tratamiento de la muestra, proporcionando la liofilización el mayor número de compuestos identificados (un total de 47 compuestos frente a los 49 identificados en total) de los que 24 se encontraron exclusivamente en las muestras liofilizadas. Aunque la reconstitución de la muestra liofilizada se llevó a cabo utilizando tres factores de

preconcentración (20, 60, y 80), el número de compuestos no aumentó con el aumento del factor, identificándose sólo un metabolito adicional, la urea, en la muestra liofilizada y reconstituida con el factor de preconcentración más alto.

Por otro lado, el eluato del sorbente HI proporcionó picos considerablemente mayores que los del sorbente C18 a tiempos de retención cortos, mientras que el perfil cromatográfico proporcionado por el eluato de la SPE-C18 se caracterizó por picos más intensos a tiempos de retención más largos; por tanto, estos sorbentes proporcionan información complementaria. Tres metabolitos se detectaron exclusivamente en el eluido de SPE-HI debido a su naturaleza polar: los ácidos láctico y fosfórico y el pirofosfato, siendo estos dos últimos los dos metabolitos no detectados en las muestras liofilizadas. Por otra parte, se identificaron 16 metabolitos exclusivos de la muestra analítica constituida por el eluato del sorbente C18, y 6 en los eluatos de ambos sorbentes debido a su interacción con los dos materiales. La concentración en EBC de la mayoría de los compuestos es tan baja que una etapa de PP, que no implica preconcentración, no fue útil cuando se usó aislada al no identificarse ningún metabolito exclusivo tras ella. Lo mismo ocurrió al combinar la PP con la SPE con C18 como sorbente. Por el contrario, se detectaron 14 compuestos tras la PP+SPE-HI, que no se consiguieron sin la etapa PP. La capacidad de cada método de preparación del EBC fue además cuantificado de forma relativa usando el área normalizada en cada una de las principales familias de compuestos.

Además de la detección de aminoácidos y derivados, alquilaminas, oxoaniones, imidazoles, hidroxiácidos y ácidos acíclicos alifáticos, es de destacar la presencia en EBC de metabolitos no polares —el 51% del total de metabolitos identificados— a pesar de la naturaleza acuosa de este biofluido. Concretamente se identificaron 9 derivados de ácidos grasos, siendo la PP la mejor opción para su detección más sensible, 9 bases esfingoideas donde la SPE-C18 tras previa PP fue el enfoque más adecuado, y 5 amidas lipídicas que junto con 2 aldehídos de este tipo se cuantificaron más sensiblemente tras SPE-C18. Además, la evaluación de algunos de los metabolitos identificados como biomarcadores potenciales

en otros biofluidos constituye un aspecto interesante para estudiar el EBC como biofluido para la búsqueda de biomarcadores.

Capítulos 8 y 9

Dos de los principales objetivos en metabolómica clínica son estudiar e interpretar rutas metabólicas que se perturban o modifican por la presencia de una enfermedad determinada, así como encontrar y caracterizar nuevos biomarcadores para el seguimiento de su desarrollo, de modo que puedan utilizarse para prevenir la enfermedad en sí, para detectarla en los estadios iniciales, e incluso para estudiar la evolución durante su tratamiento. Las dos primeras etapas en la búsqueda de biomarcadores son la identificación del biomarcador candidato y su validación analítica. La primera generalmente implica utilizar una estrategia de análisis no orientado o de huella dactilar con el fin de identificar moléculas en una concentración significativamente diferente en grupos de individuos. El siguiente paso es la evaluación del potencial de estas moléculas como biomarcadores utilizando algoritmos adecuados, tales como el análisis de curvas ROC. En general, la combinación de biomarcadores en paneles potencia la capacidad discriminante respecto al uso de moléculas individuales.

El estudio de estas dos primeras etapas necesarias para descubrir biomarcadores de uso en análisis clínico ha dado lugar a la redacción de dos artículos que conforman los Capítulos 8 y 9 de esta Memoria de Tesis. Así, el artículo "Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC–QTOF: upstream information on other omics" que constituye el Capítulo 8, pone de manifiesto diferentes familias de compuestos en orina que son de interés clínico por su implicación en rutas biológicas cruciales que se alteran o desregulan como consecuencia del PCa.

La presencia de productos finales del metabolismo del cáncer liberados ya sea en la orina o transportados al interior de las células prostáticas y directamente liberados en la uretra a través de conductos prostáticos, junto con todas las ventajas que este biofluido presenta y que se han expuesto en el Capítulo 1,

fueron los motivos de su elección para desarrollar esta investigación. Otros aspectos que han impulsado este estudio han sido la falta de especificidad del antígeno prostático específico (PSA), el actual marcador de uso clínico del PCa, junto con la ausencia de síntomas en etapas tempranas del tumor, que resultan en un elevado número de hombres sobrediagnosticados sometidos a cirugía o a tratamientos con efectos secundarios adversos.

El análisis global mediante LC–QTOF de la orina de 62 pacientes con PCa clínicamente significativo y 42 individuos sanos, ambos confirmados por biopsia, proporcionó 28 metabolitos significativos tentativamente identificados tras un test de la *t* no pareado (valor de $p < 0.05$) que se utilizaron para desarrollar un modelo de análisis discriminante de mínimos cuadrados parciales (PLS-DA). La clara separación entre los pacientes con PCa y los controles mediante el PLS-DA, considerando el 70% de las muestras seleccionadas al azar, se caracterizó por el 88.4 y el 92.9% de sensibilidad y especificidad, respectivamente. La estabilidad del modelo se evaluó mediante validación externa utilizando el 30% de las muestras excluidas en el entrenamiento, reportando una sensibilidad y especificidad del 63.2% y el 78.6%, respectivamente.

Entre los compuestos significativos es destacable la presencia de aminoácidos y sus derivados, representando el 78.6% del total. Además, la presencia de aminoácidos acetilados como la acetilcitrulina, la acetilarginina, la acetiltriputrescina o la acetil-lisina y aminoácidos metilados, así como la de bases nitrogenadas como la metilguanosina, la metil-lisina, la metilhistidina o la dimetilarginina representan el 25 y 28.6% de todos los metabolitos. Concretamente, los aminoácidos lisina, arginina e histidina acetilados y metilados fueron los metabolitos más significativos del estudio, encontrándose a una concentración más baja en la orina de pacientes con PCa. Cabe destacar el papel de la dimetil-isina, el compuesto más significativo en términos del valor de *p*, que también experimentó la mayor magnitud de cambio en su concentración: 1.74.

La metilación del ADN, las marcas epigenéticas en residuos de aminoácidos en histonas y la metilación de la cápsula del ARNm constituyen tres de los princi-

pales procesos biológicos relacionados con varios de los metabolitos significativos encontrados, como la 7-metilguanina, la 7-metilguanosina, el 5-metildeoxicotidina-5′-fosfato y los ya mencionados aminoácidos metilados y acetilados. Estos procesos están bien reconocidos en la regulación de la transcripción de genes y desempeñan papeles claves en la iniciación y progresión del PCa.

La evaluación de los metabolitos significativos considerados potenciales biomarcadores del PCa utilizando algoritmos como el análisis de curvas ROC fue el siguiente paso del estudio que constituye el Capítulo 9 y se recoge en el artículo "Urinary metabolite panels for prostate cancer detection".

La capacidad predictiva de metabolitos significativos se evaluó previamente mediante un análisis de Random Forest para discriminar los casos de PCa y los individuos control. La dimetil-lisina, la acetilputrescina y el lactato de imidazol presentaron la mayor capacidad de predicción, teniendo por tanto mayor probabilidad de aparecer en un panel de biomarcadores. Estos resultados están de acuerdo con la capacidad predictiva de los metabolitos proporcionados por sus curvas ROC individuales, siendo la sensibilidad de la dimetil-lisina, la acetilputrescina y el lactato de imidazol del 85.7%, el 71.4% y el 78.6%, respectivamente, y la selectividad del 72.1%, el 81.4% y el 74.4%.

En general, la combinación de biomarcadores en paneles potencia la capacidad discriminante respecto al uso de biomarcadores aislados. Así, en esta investigación se configuraron paneles de metabolitos usando una herramienta computacional llamada Panelomix con diferentes criterios: maximizar independientemente la sensibilidad y especificidad, o considerar conjuntamente ambos parámetros (sensibilidad+especificidad). Los metabolitos significativos identificados se combinaron alternativamente con los valores de PSA en paneles de 5 variables (5 metabolitos o 4 metabolitos + PSA) para el diagnóstico del PCa. La metodología PanelomiX se aplicó al 70% de los pacientes para el desarrollo de los modelos, mientras que el 30% restante se utilizó como conjunto de validación externa. Todos los paneles proporcionaron una respuesta positiva cuando 3 de los 5 metabolitos dieron una respuesta cuantitativa por debajo o por encima del

umbral correspondiente. La dimetil-lisina, la tirosina y el lactato de imidazol estuvieron presentes en todos los paneles en los que no se incluyó el valor del PSA, revelando la importancia de los aminoácidos para mejorar la discriminación de los pacientes con PCa.

La falta de especificidad de PSA como marcador de PCa apoyó la configuración de un panel que combina PSA y 4 metabolitos para complementar la alta sensibilidad de PSA. Este panel se diseñó para maximizar la exactitud global e incluyó, además del PSA, la acetilputrescina, la dimetil-lisina, la citrulina y la dimetilarginina, proporcionando un 100% de especificidad y un 97.7% de sensibilidad. La validación con el 30% de las muestras ofreció un 85.7 y un 73.7% de especificidad y sensibilidad, respectivamente.

Con estos estudios se pone de manifiesto el papel clave del metabolismo de los aminoácidos y sus derivados, especialmente de los derivados de la arginina y la lisina, en la discriminación del PCa.

Discussion of the results

The present rules in the University of Córdoba concerning the model of Thesis-Book in the format in which the articles as such (either published or close to be published) are in the book, make mandatory to include a joint discussion of the obtained results.

The research in this Thesis-Book, with the common denominator metabolomics and clinical area, has been divided into three sections as a function of the pursued objectives. Thus, Section I, constituted by Chapters 1 and 2, discusses bibliographical reviews devoted to sample preparation of urine, and to methods for determination of exhaled breath condensate (EBC) components, two of the biofluids investigated in this Thesis.

The experimental research that constitutes Sections II and III is devoted to the two basic strategies in metabolomics: targeted and untargeted analysis. Section II encompasses the research on determination of compounds of clinical interest such as eicosanoid derivatives from omega-6 (Chapter 3), involved in the inflammatory response. The improvement in the quantification of a class of these derivatives, hydroxyeicosatetraenoic acids (HETEs), by using the deuterated internal standard 15HETE-d8, constitutes Chapter 4, a complement of Chapter 3. On the other hand, the development of a method for quantification of the chemotherapeutic drug paclitaxel and its main metabolites allows monitoring their effectiveness y/or toxicity, as shown in Chapter 5. Section III is devoted to the research on metabolomics profiles (profiling analysis) carried out with two objectives: to achieve chemometric challenges as the shortening of the time required for treatment of the data provided by mass spectrometric detectors (Chapter 6) and encompass as much as possible of the metabolome of the biofluids that are the subjects of Section I. To fulfil this last objective the sample preparation step of EBC has been optimized (Chapter 7) and a search for potential biomarkers of prostate cancer has been carried out (Chapters 8 and 9). The results obtained in

each of the three sections are discussed after an overview of some experimental aspects.

The use of mass spectrometry (MS) for detection is common to both experimental sections. The analytical properties of this technique, in terms of sensitivity, selectivity, accuracy, precision and resolution, make MS the most useful tool for metabolomics analysis (both targeted and untargeted) in the clinical field. The two typical strategies in metabolomics take advantage of MS in its different modes as a function of the pursued objective: qualitative or quantitative analysis. Among the different types of mass spectrometers the most used for untargeted analysis are QTOF and Q-Orbitrap thanks to their high resolution (accuracy in measurement of the *m/z* ratio), selectivity and sensitivity. These reasons make them the most appropriate detectors for qualitative analysis, and also for relative quantification. In this research QTOF equipment has been used, which shows as additional advantage a high scan rate. On the other hand, the QqQ instrument is very useful for targeted analysis thanks to its sensitivity and selectivity for confirmatory analysis, which make it the mass detector of choice for quantitative purposes. Taking into account the complexity of clinical samples, mass spectrometers are very appropriate for use in clinical metabolomics, particularly when coupled to chromatographs. Liquid chromatography in the reverse phase mode has also been common to both experimental sections.

Section I: Bibliographical studies

The research on bibliography developed by the PhD student during the first years of her Thesis alerted about the scan exploitation of urine and exhaled breath as samples for clinical research, despite the advantages of these samples. In addition to their easy and non-invasive sampling, they foreseeably can provide interesting information on the pathologies with which they are related. This is the case with the two types of cancer with higher incidence in human population: that of prostate in men and that of lung, both subjects of this Doctoral Thesis. The limited use of these samples is due to the scant information on the appropriate treatment; therefore, an exhaustive bibliographic study on the present situation of

both preparation of urine samples for metabolomics analysis, and the available systems for lung cancer detection based on EBC. Chapters 1 and 2 of this Thesis Book encompass these bibliographic studies.

Chapter 1, content published with the tittle "Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis", discusses the steps of sample pretreatment and sample preparation of urine prior to individual separation and/or detection, and shows the scantly rigorous attention devoted to these steps so far. Special attention is given to these steps as a function of the selected metabolomics strategy, with clear emphasis on keeping the samples as intact as possible by developing the minimum number of stages in dealing with global analyses of metabolome. All due to the fact that the behavior of the compounds in it are not known; while in targeted analysis these stages must be as selective as possible for the target compounds.

The stages involved in urine pretreatment, from sampling to storage, through the collection device, metabolism inhibition, preservatives addition, volume normalization, pH adjustment and deproteination, have been exhaustively discussed, and the advantages and disadvantages of them considered in the light of the analysis objective.

Sample preparation of urine as such is the subject of the second part of the review. This step encompasses all the sub-steps to adequate the analytes to detection; nevertheless, as most of the detectors used in metabolomics are on-line connected to equipment for previous separation (a chromatograph or, less common, capillary electrophoresis equipment), the sample preparation step is considered finished when the sample (or what of it rests after preparation; thus is, the "analytical sample") is ready for being inserted into the high resolution separation equipment. The sub-steps involved in sample preparation of urine are enzymatic hydrolysis, sample cleanup and preconcentration of the target analytes mainly by SPE or LLE, evaporation or lyophilization and addition of the internal standard.

The different options to carry out each sub-step have been discussed and the suitability of some of them as compared with others are justified through the optimization studies. The sub-steps that are essential, advisable or unnecessary are also discussed in this publication, which also suggests some emergent proposals as the use of ultrasound to accelerate the hydrolysis step. The errors associated to an incorrect development of these sub-steps are also emphasized, as well as the ways in which they must be developed to avoid or minimize errors. Especial emphasis is given to sub-steps that can promote ionization suppression such as the addition of surfactants for urine preservation, evaporation or lyophilization. The use of internal standards labelled with stable isotopes to compensate for undesirable effects is advised.

The second part of Section I, Chapter 2, is the review "Analytical methods based on exhaled breath for early detection of lung cancer". It shows the high mortality caused for this type of cancer as a consequence of both the absence of symptoms during the first steps of the disease and the limited possibilities of the present techniques for diagnostic, which give place to a high number of false positives, and unnecessary biopsies as a result.

The characteristics and composition of EBC are described in this publication, and the latter is compared with that of volatile organic compounds (VOCs). The fact that the increased use of EBC for screening of respiratory diseases has not involved a similar use for lung cancer studies has been emphasized.

The methods for detection of lung cancer based on trained dogs taking advantage of their sensitive smell, as well as methods based on other great omics such as genomics and proteomics, and the use of sensors are also discussed in this publication. The potential contribution of research based on the analytical chemists–metabolomics binomial is outlined, based on the fact that an earlier detection of this cancer by dogs (that is, when the biomarkers of this disease must be in EBC at a very low concentration) could be surpassed by using cutting-edge analytical instrumentation endowed with enough sensitivity for the study and analysis of these biomarkers. The achievement would avoid the limitations in-

herent to the use of dogs for detection of lung cancer. For example, detection of compounds non-related with cancer formed as a consequence of inflammation, necrosis, tabaquism, etc., or those that are not selective for this type of cancer; thus making mandatory identification of these compounds. The majority of EBC components are metabolites; therefore, emphasis is made on the necessity of metabolomics to obtain more rational, reliable, and wide information to establish the basis for an easy and accurate detection of this disease. The present analytical technology endowed with high sensitivity for metabolomics studies is a promising tool to reach the detection limits provided by dogs. In addition, an appropriate sample preparation would help to enhance the information it contains, as shows the research that constitutes Chapter 7.

The conclusion from Section I is that an appropriate sample preparation and the proper selection of the detector provide a correct biological interpretation of the results. In short, Section I shows what has been done and what challenges must be attained to consolidate the use of urine and EBC in clinical research for the detection and monitoring of given diseases. This state-of-the art was the key for development of the following Thesis Chapters, in addition to offering to biochemists, clinicians and researchers non familiar with the subject a discussed overview of it.

Section II: Contributions to targeted metabolomics analysis

Targeted metabolomics analysis is that applied to a sample for quantification of a family or group of compounds, usually to precursors or to metabolites belonging to given metabolic pathways. Despite this strategy is only applied to predefined compounds, it is endowed with a number of advantages as compared with metabolomics untargeted analysis. The main advantage is that sample preparation and analytes detection are optimized both to reduce the number of interferents and maximize sensitivity and selectivity by selection of the appropriate analytical platform.

The sample preparation step depends on the analytes nature, being liquid–liquid extraction (LLE) and solid-phase extraction (SPE) the two more used alternatives to remove interferents and preconcentrate the target metabolites. In short, the research in Chapters 3 and 4 has took advantage from automated SPE on-line coupled to subsequent steps of the analytical process; while LLE has been used in the research in Chapter 5.

When the number of compounds to be determined is small enough to optimize sample preparation and detection for each of them, and individual calibration can be carried out —providing standards of them exist— quantification of all target compounds can be achieved. Inclusion of an appropriate internal standard substantially improves accuracy of quantification, as justified in Chapter 3.

The panels of the compounds selected in this research for targeted analysis were formed by metabolites involved in known metabolic pathways and with a relevant role in given pathologies. Such is the case with a drug used for cancer therapy and its main metabolites, the determination of which is the key to monitor their efficiency and/or toxicity.

The development of a method must always finish with its validation and application to real samples to show its usefulness. The methods in this section have been properly assessed by evaluation of their analytical features and then applied to real samples from cohorts with some pathology as those formed by patients diagnosed with different cardiac affections (Chaper 4), and women with ovarian peritoneal carcinomatosis (Chapter 5).

Chapters 3 and 4

Chapter 3 is devoted to the study of an eicosanoid family, hydroxyeicosatetranoic acids or HETEs to evaluate the inclusion of an internal standard labelled with stable isotopes, 15HETE-d8 (also used in Chapter 4), which resulted in the publication “Stable isotopic internal standard correction for quantitative analysis

of hydroxyeicosatetraenoic acids (HETEs) in serum by on-line SPE–LC–MS/MS in selected reaction monitoring mode". On the other hand, the development of a fully automated method for the determination of derivatives from eicosanoids of omega-6 fatty acids (arachidonic and linoleic acids); in fact prostaglandins (PGs), hydroxyoctadecadienoic acids (HODEs), HETEs and thromboxane (TX) in serum for the study of coronary artery diseases was the subject in Chapter 4. It gave place to the publication "Targeted analysis of omega-6-derived eicosanoids in human serum by SPE–LC–MS/MS for evaluation of coronary artery disease". As the same analytical platform was used for development of the research in both chapters and also quantification of 8,11,12 and 15HETE with the help of 15HETE-d8 is common to them, the discussion of the results from both is developed jointly.

HETEs play a regulatory role in the inflammatory response, and are associated to different pathologies as cardiovascular diseases, cancer and osteoporosis. Therefore, a method for their determination within a concentration range that encompasses the extremely low values at which they are in human biofluids is very useful. In addition, the complexity of the biological samples demands for appropriate protocols for sample preparation that remove potential interferents prior to the chromatographic separation. With this purpose, an SPE–LC–MS/MS arrangement was used, in which the fully automated SPE avoids the steps of proteins precipitation and evaporation, thus decreasing losses and degradation of analytes, and shortening the time for analysis as a result.

The number of target metabolites allowed complete optimization of the variables that affect both the detection of each compound (voltage collision energy and product ions) and extraction (type of sorbent, solvents, volume and flow-rate of charge, wash and elution), as well as characterization of the developed method: calibration curves, dynamic ranges, detection and quantification limits, repeatability, reproducibility, and accuracy, which were markedly improved by including the deuterated internal standard. Also, the comparison with previous methods for the same analytes based on of-line LLE and SPE revealed an improved sensitivity thanks to the on-line arrangement.

The use of a deuterated HETE as internal standard as a way of improving the quality of the results for the compounds belonging to this group was considered of interest. For this study, four calibration models were developed under the optimum working conditions using either a serum pool or standard solutions and applying or not internal standard correction. A significant matrix effect was observed in the calibration models prepared in serum as compared to those obtained by standard solutions, thus achieving an also significant correction effect with the use of the deuterated internal standard. The correction effect was higher for 15HETE and 12HETE than for 8HETE and 11HETE, justified by the chemical similarity of the two former with 15HETE-d8, as the hydroxyl group is further from the position 15 in 11HETE and 8HETE. The validation of the model was complemented by other studies such as evaluation of the efficiency of cartridge retention by a dual cartridge system, which showed an excellent retention, above 90%, for all metabolites. The accuracy of the method, crucial to evaluate the matrix effect, was calculated by the recovery factor at three concentration levels in serum, with and without internal standard correction. The quantification of 8HETE, 11HETE and 15HETE was properly supported on the analytical parameters and averaged concentrations (35 and 55 ng/mL), while the analysis of 12HETE provided a low recovery factor at the three concentration levels. On the other hand, quantification of low concentrated eicosanoids was characterized by a decrease of the recovery factors below 78.5%. Therefore, implementation of correction factors is mandatory for accurate quantification within given ranges.

Once the method was optimized and validated, it was applied to the analysis of serum samples from 45 aterosclerotic patients (15 with stable angina, 15 with myocardial infarction non-ST-elevated elevation acute coronary syndrome —NSTEMI—, 15 who suffered acute myocardial infarction —AMI—, and 16 controls. Despite the method allowed the quantitative analysis of 12 eicosanoids, only 5 of them were at concentrations above their quantification limit in a representative number of samples.

An ANOVA univariate analysis showed the significance of TX and 11HETE (products of the catalytic action of cyclooxygenases) in control individuals as

compared with patients with coronary disease. The relative concentration of both compounds was evaluated by their fold-change, and TX and 11HETE were found 14.66 and 1.18, respectively, increased in the group; therefore, only TX experienced a significant change between both groups. Although TX acts as a strong agent for platelet aggregation the found result can be explained because the majority of the patients had been treated with aspirin that inhibits the enzymatic activity of COXs, thus reducing the levels of TX in blood with an antiaggregation platelet effect, even at low doses. Despite the inhibitory effect of aspirin on the activity of COXs caused on patients with lesion, HETEs and HODEs are generated through other two enzymes (lipooxygenases —LOXs— and cytochrome P450 —CYP), which act on arachidonic and linoleic acids.

The levels of eicosanoids in serum of patients diagnosed with coronary artery disease were also statistically examined depending on the severity of the coronary lesion as stable angina, NSTEMI or AMI. 12HETE was the only significant metabolite to explain the differences between individuals who suffered an AMI episode and those diagnosed with stable angina at a 95% significance level. Synthesized under the enzymatic action of an LOX, the concentration of this compound is not affected by aspirin treatment; therefore, it was 2.5 times more concentrated in patients affected by AMI, in agreement with previous studies that reported 12HETE formation during ischemia. The significance of 12HETE in AMI *versus* stable angina was also evaluated by ROC curves with estimation of the partial area under the curve restricted between 90 and 100% of specificity, yielding a result of 90%.

The significant differences in the 12HETE levels shown by patients affected by AMI as compared with stable angina patients must be emphasized. These results can help to the diagnostic and prevention of a new AMI in patients previously affected. For quantification of this potential biomarker, an analytical platform with on-line connection of an SPE device with LC–MS/MS equipment constitutes an optimum choice, without forgetting the use of the appropriate internal standard to achieve the required accuracy.

Chapter 5

Another analytical application in the clinical area was the research published under the title "LC–MS/MS quantitative analysis of paclitaxel and its major metabolites in serum, plasma and tissue from women with ovarian cancer after intraperitoneal chemotherapy".

The secondary effects from the use of an exogenous metabolite as the antineoplasic paclitaxel made necessary its quantification together with that of its main metabolites: 6α-hydroxypaclitaxel (6α-OHP) y el *p*-3´-hydroxypaclitaxel (C3´-OHP). The pharmacometabolomics approach allows, in this case, studying the pharmacokinetics of paclitaxcel with the aim of optimizing the doses and exposition times for therapies, avoiding in this way the toxic effects and giving a new step towards personalized medicine.

This pharmacologic study was supported on the LC–MS/MS platform with selected reaction monitoring (SRM) for the simultaneous quantification of paclitaxel, 6α-OHP, and C3´-OHP, with docetaxel as internal standard. Optimization of the detector parameters and those of the ESI ionization source for the analytes and the internal standard led to the use of 0.1% formic acid as ionizing agent, and selection of the [M+Na]+ adducts as precursor ions as the main product ions were produced from them.

The sample preparation step (serum and tissue as samples) was optimized using a pool of each sample that was taken before chemotherapy (after checking the absence of the target analytes). After protein precipitation the sensitivity of the method was very low; therefore, this step was avoided and, instead, LLE with methyl t*ert*butyl ether (MTBE) and evaporation plus reconstitution in methanol were used to remove interferents and preconcentrate the analytes, respectively. The characterization of the method provided limits of detection and quantification lower in serum than in tissue for all analytes, being both limits higher for 6α-OHP than for paclitaxel and C3´-OHP. The precision and accuracy of the method for

tissue and serum samples were acceptable, thus corroborating the capability of the method for quantification of the target compounds.

The method was applied to serum and tissue from 13 women diagnosed with ovarian peritoneal carcinomatosis after hyperthermic intraperitoneal intra-operative chemotherapy (HIPEC), and also to plasma from one of them. Samples were taken at different times: before HIPEC treatment, immediately after and 1 h after the treatment for serum and plasma sampling, while the tissue biopsy was performed before treatment and after 1 h application.

Paclitaxel and metabolites were not detected in the samples obtained before treatment with the drug, as did the metabolites in the tissue after chemotherapy. This behavior was foreseeable as paclitaxel is hardly metabolized in peritoneal tissue. Nevertheless, the concentration of the drug was significantly higher in tissue than in serum. In most cases, the concentration of paclitaxel in serum decreased 1 h after HIPEC treatment, being mainly metabolized to 6α-OHP, as this metabolite increased significantly its concentration as compared to that of C3´-OHP.

It must be pointed out the suitability of the method for the determination of the target analytes in plasma from one of the patients, as the differences in concentration as compared with that in serum at the two sampling times (immediately after treatment and 1 h later) were within the experimental error of the method.

Section III: Contributions to global metabolomics analysis

The main objective of untargeted metabolomics analysis is to obtain as much as possible information on the nature of the metabolites in the sample under study. A key requirement for its development is the use of high resolution mass analyzers that allow good resolution in the measurement of the *m/z* ratio, thus making possible identification in the target biofluid of a number of metabolites as high as possible. QTOF or Q–Orbitrap equipment are the preferred options in these cases, as they allow obtainment of MS/MS spectra from the detected precursor ions that

—together with a device for appropriate separation of the target compounds, usually a chromatograph— make possible unequivocal identification of the metabolites. Both chemometric and bioinformatics tools are the key in untargeted analysis to mine useful information from the most times vast volume of data in a way as representative and fast as possible, and with confident conclusions. The biological information available in HMDB and the biosynthetic pathways of the Kyoto Encyclopedia of Genes and Genomes (KEGG) have shown to be very useful for understanding the results and interpreting them. QTOF equipment existing in the laboratory of the research group was used in this study. It would be very interesting comparison of the obtained results with those from Q–Orbitrap equipment, but a collaboration with this purpose was not be possible. materialized.

The 4 objectives pursued by this metabolomics strategy, that constitute the 4 chapters of this section, were as follows: to improve and accelerate treatment of the results generated by LC–TOF, to study the treatment of a scarcely studied biological sample as EBC and achieve appropriate information from samples of urine from PCa patients to finally develop panels of potential biomarkers. The obtained results are discussed below.

Chapter 6

Despite a great number of strategies for data analysis appear in the literature, none of them can be selected as optima in all cases; therefore, this is an area of active research, particularly in dealing with bioinformatics. The article "MSCombine: a tool for merging untargeted metabolomic data from high-resolution mass spectrometry in the positive and negative ionization modes", subject of this chapter, is a clear example of the activity in this area.

LC–QTOF in MS/MS mode is one of the best platforms to obtain very comprehensive information on a wide number of the metabolites in a sample, as it combines two ionization modes —positive and negative— the choice of one of them depends on the metabolites structure. Nevertheless, the mode is selected after separate processing of the results obtained by each polarity; a time consuming

step. Despite a significant fraction of the metabolites are preferentially detected by one of the ionization modes, a number of compounds can be detected by both; thus generating redundant information. To overcome these limitations a tool to combine the data matrices from the positive and negative ionization modes in MS was proposed. It was based on an open source data analysis software R through its graphical user interface R-studio. This tool combines the two matrices with the data generated by both ionization modes after removing common molecular entities and keeping those detected with the highest sensitivity; thus resulting a single matrix easier to hand and representative of the target biological sample.

The different algorithms used to achieve the pursued aim were: the *Findcommon* function, executed specifying a mass and retention time (RT) tolerance of 0.002 Da and 0.2 min, respectively, thus generating the file *CommonEntities*. Then, and through the function *RemoveMismatch*, ions associated to the same molecular feature were removed, selecting only the most common ion (*viz.*, $[M+H]^+$ in the positive ionization mode, and $[M-H]^-$ in the negative mode). In the third step the function *StudyRTdiff* built a regression model that compared the two RT lists to establish a threshold —in this case ± 0.2 min— by filtering the matrix using the function *FilterbyRT* that eliminated no coincident entities and confirmed entities associated to the same metabolite. This is a crucial step to avoid mismatching due to the RT tolerance filter, resulting in a new data set named "CommonEntitiesFiltered" that is ready for use in the next step. The last step (*CombinePolarities* function) compared the average area of the molecular entities detected by both polarities and associated to the same metabolite, and eliminated the less sensitive entities. In addition, this function suppressed other entities associated to the same metabolite associated to the same ionization mode.

The functionality of the proposed algorithm was assessed by using real data sets obtained from the analysis of serum and urine samples. The 27 and 41% of the potential entities detected in the positive and negative ionization modes, respectively, corresponded to redundant information in serum, while in urine percentages 26 and 31 of the potential entities detected in the positive and negative modes, respectively, were also detected in the other mode.

MSCombine was also validated by statistical analysis of the data from the urine cohort, which consisted of 787 and 664 molecular entities in the ionization modes positive and negative, respectively. Combination of the two matrices provided a matrix with 1451 entities, reduced to 1245 by application of the proposed method. Statistical analysis by a Volcano plot ($p < 0.01$, fold change > 2) provided 41 significant features that were different between the two groups (controls and individuals who ingested ciprofloxacin) in the data set provided by MSCombine, and 54 in the data set that combined the positive and negative ionization matrices. The 13 molecular features absent from the MSCombine data set correspond to entities removed as redundant entities. In fact, tentative identification of those 13 entities revealed that they were associated to 10 metabolites already included in the MSCombine matrix.

Chapter 7

"Study of exhaled breath condensate sample preparation for metabolomics analysis by LC–MS/MS in high resolution mode" is the article that gave place to Chapter 7. As commented in Chapter 2, EBC is a sample of easy and non-invasive sampling, and low cost. As this sample is representative of the liquid that covers the respiratory tract, it is foreseeable that its analysis can lead to identifying biomarkers of different respiratory diseases; thus being an excellent alternative to present cumbersome and invasive techniques. Nevertheless, a high percentage of EBC is water, thus making mandatory the search for appropriate sample preparation steps that preconcentrate EBC components and remove potential interferents. This was the challenge faced by this study.

Most of the studies on EBC have been focused on given compounds. A non-focused approach allowed in this case to detect the metabolites in the sample after application of different treatments, and elucidate in this way if one of them provides better results expressed either as the number of detected compounds or what is the best to detect/quantify given metabolites. With the aim of obtaining a metabolites profile complementary to that provided by a GC–MS platform, LC–

QTOF was used, once again, to study potential volatile and semi-volatile compounds in EBC.

Three alternatives were evaluated before global analysis to process an EBC pool from 4 healthy volunteers: (i) precipitation of proteins (PP); (ii) SPE in a miniaturized format assisted by centrifugation (because the low sample volume available in most cases), and with 2 complementary sorbents with lipophilic interactions (C18) and hydrophilic interactions (IH); (iii) lyophilization of EBC.

The qualitative study of the results by Venn plots by using only the entities tentatively identified by MS/MS allowed comparing the number of entities detected after each sample treatment. Lyophilization provided the highest number of identified compounds (47 identified compounds *versus* 49 as the total number of compounds identified by using the analytical samples resulting from the 3 alternative treatments. Despite reconstitution of the lyophilized sample was carried out using 3 concentration factors (20, 60 and 80, as compared with the volume of the original sample) the number of detected compounds did not increase, but only a new metabolite (urea) was found in the lyophilized sample reconstituted at the highest preconcentration factor.

Concerning sorbents, the eluate from sorbent HI provided, at short retention times, much higher chromatographic peaks than sorbent C18; and the opposite happened at long retention times; therefore, the information from both sorbents was complementary. Three metabolites were exclusively detected in the eluate from SPE-HI because its polar character: lactic and phosphoric acids and pyrophosphate, being the latter two the only metabolites non-detected in the lyophilized samples. On the other hand, 16 metabolites were exclusive in the analytical sample from the eluate from sorbent C18, and 6 the eluated from both sorbents due to its interaction with the two materials. The concentration of the majority of compounds in EBC is so low that a step as PP, that does not involve preconcentration, was not useful when used isolated, as the analytical sample thus obtained did not provide any exclusive metabolite. Similar behavior was obtained by combination of PP and SPE-C18. On the contrary, 14 compounds non-detected by only SPE-HI were detected by combining PP and SPE-HI. The ability of each

method for preparation of EBC was also relatively quantified using the normalized area for each family of compounds.

It must be pointed out that in addition to the detection of amino acids and derivatives, alkilamines, oxoanionic compounds, imidazoles, hydroxyl acids and aliphatic acyclic acids, non-polar metabolites —51% of the identified compounds— were identified, despite the polar nature of EBC. The identified compounds were 9 fatty acid derivatives —for which PP provided the best analytical sample—, 9 sphingoid bases —best analytical sample from PP+SPE-C18—, and 5 lipidic amides+2 lipidic aldehydes provided a more sensitive relative quantification when the analytical sample was from the single SPE-C18 step. Finally, some of the identified metabolites had previously been identified as biomarkers in other biofluids, thus openning a door to the study of EBC as biofluid to search for biomarkers.

Chapters 8 and 9

Two of the main objectives of clinical metabolomics are to find and study metabolic pathways which are disturbed of modified by the presence of a given disease, and also to search for and characterize new biomarkers to monitor its development. In this way, biomarkers can be used to prevent the disease, to detect it at the initial stages, even to study its evolution during a given treatment.

The first two steps in the search for biomarkers are identification of the candidate to biomarker and its analytical validation. The first step involves the use of a metabolomics strategy of either untargeted analysis or fingerprint to identify molecules at significant different concentration in groups of individuals. The next step is to evaluate the potential of these molecules as biomarkers by using the appropriate algorithms, such as analysis of ROC curves. In general, the combination of biomarkers into panels potentiates the discriminant capacity as compared to the use of individual biomarkers.

The study of the two steps mandatory to find clinical biomarkers has given place to two articles that constitute Chapters 8 and 9 of this Thesis-Book. Thus, the article "Prostate cancer patients–negative biopsy controls discrimination by untargeted metabolomics analysis of urine by LC-QTOF: upstream information on other omics", that constitutes Chapter 8, shows the different families of compounds in urine that are of clinical interest because their involvement in crucial biological pathways, which can be altered or deregulated as a consequence of prostate cancer (PCa).

The main reasons for selection of urine as sample biofluid were: (i) the presence of final products of cancer metabolism that are released into urine or transported into prostatic cells; (ii) the fact that these last products are then released in the urethra through the prostate ducts; (iii) the advantages of urine as clinical sample, as widely discussed in Chapter 1. Other reasons that have promoted this study have been: (i) the lack of specificity of the prostate specific antigen (PSA), the clinical biomarker used presently for PCa, and; (ii) the lack of symptoms in the first steps of this cancer, that result in a high number of over-diagnosed men, then subjected to either surgery or treatments that produce adverse side effects.

Untargeted metabolomics analysis by LC–QTOF of urine from 62 patients with clinically significant PCa, and 42 healthy individuals, both confirmed by biopsy, provided 28 significant metabolites tentatively identified after an unpaired *t*-test (p value < 0.05). These metabolites were used to develop a model of partial least squares discriminant analysis (PLS-DA). The clear separation between patients with PCa and controls provided by PLS-DA, taking 70% of the samples randomly selected, was characterized by 88.4% and 92.9% sensitivity and specificity, respectively. The stability of the model was assessed by external validation using the remaining 30% of the samples excluded from the training set, reporting sensitivity and specificity of 63.2% and 78.6%, respectively.

Among the significant compounds it is remarkable the presence of amino acids and derivatives that constitute 78.6% of the total. In addition, the presence

of acetylated amino acids such as acetylcitrulline, acetylputrescine or acetillysine, methylated amino acids as well as nitrogenous bases derivatives such as methylguanosine, methyllisine, methylhistidine and dimethylarginine constitute 25.0 and 28.6%, respectively of all metabolites. In short, acetylated and methylated lysine, arginine and histidine were the most significant metabolites in this study, which were at lower concentration in urine from PCa patients than in that from healthy individuals. It is remarkable the role of dimethyllisine, the most significant compounds in terms of *p* value, which also experienced the highest change in concentration: 1.74.

DNA methylation, epigenetic marks in residual amino acids in histones and RNA cap methylation constitute three of the main biological processes related to several significant found metabolites, such as 7-methylguanine, 7-methylguanosine and 5- metildeoxycotidine-5′- phosphate and the above commented methylated and acetylated amino acids. These processes are well known in the regulation of gene transcription, and they play key roles in the initiation and progression of PCa.

Evaluation of significant metabolites considered as potential biomarkers for PCa by using algorithms such as ROC curves was the next step of the study that constitutes Chapter 9 and gave place to the article “Urinary metabolite panels for prostate cancer detection”.

The prediction capability of significant metabolites was evaluated by Random Forest analysis to discriminate PCa cases and individuals control. Dimethyllisine, acetylputrescine and imidazole lactate provided the highest prediction capability; therefore, they have also the highest possibility to be in biomarkers panels. These results are in agreement with the prediction capability provided by individual ROC curves, in which the sensitivity for dimethyllisine, acetylputrescine and imidazole lactate was 85.7%, 71.4% and 78.6%, respectively; while selectivity was 72.1%, 81.4% and 74.4%.

In general, combination of several biomarkers potentiates the discrimination capability as compared with individual biomarkers. Thus, in this research metab-

olites panels were configured using a chemometric tool known as PanelomiX attending to different criteria: maximize sensitivity or specificity independendly, or use both parameters jointly (sensitivity+specificity). The identified significant metabolites were alternatively combined with the PSA values in 5-variable panels (5 metabolites or 4 metabolites+PSA) to diagnose PCa. For development of the panels the PanelomiX tool was applied to 70% of the patients, and the rest (30%) was used for external validation. All panels provided positive response when 3 out the 5 metabolites provided quantitative response below or above the preset threshold. Dimethyllisine, acetylputrescine and imidazole lactate were always present in the panels involving PSA, thus supporting the importance of amino acids to improve discrimination of PCa patients.

The lack of specificity of PSA as PCa biomarker supports the panel that combines PSA and 4 metabolites to complement the high sensitivity of PSA. This panel was designed to maximize the overall accuracy and included, in addition to PSA, acetylputrescine, dimethyllisine, citruline and dimethylarginine, providing 100% specificity and 97.9% sensitivity. Validation of the panel using 30% of the samples provided 85.7 and 73.7% of specificity and sensitivity, respectively.

These studies show the key role of the metabolism of amino acids and derivatives, particularly arginine and lysine derivatives, in the discrimination of PCa.

CONCLUSIONES

Una búsqueda bibliografía en profundidad ha permitido evaluar críticamente el estado actual del tratamiento de muestra de orina en metabolómica y de los sistemas existentes para la detección del cáncer de pulmón basados en aire exhalado, mostrando sus deficiencias o carencias y proporcionando pautas para explotar estos dos biofluidos en investigación clínica a través de la metabolómica.

La investigación experimental que constituye esta Tesis se ha centrado en el desarrollo de aplicaciones basadas en análisis metabolómico global y orientado en el área clínica, destacando la utilidad de la cromatografía de líquidos acoplada a la espectrometría de masas para ambas estrategias. La mejora del tratamiento de la muestra y de los datos constituye otro de los aspectos que caracterizan la investigación desarrollada.

Las conclusiones más destacadas de la investigación dedicada a análisis dirigido pueden resumirse en los siguientes puntos:

- La aplicabilidad de un método para la determinación de eicosanoides en suero utilizando la plataforma SPE–LC–MS/MS se ha demostrado mediante el análisis de muestras de pacientes ateroscleróticos diagnosticados con angina estable, infarto de miocardio sin elevación del segmento ST (NSTEMI) o infarto de miocardio agudo (AMI). Los resultados ponen de manifiesto niveles significativos del ácido 12-hidroxieicosanoide (12HETE) en pacientes que han sufrido un AMI en comparación con individuos con angina de pecho estable, lo que podría mejorar el diagnóstico y prevenir la reincidencia de AMI en pacientes ya lesionados.
- El uso del estándar interno deuterado 15HETE-d8 es crucial para el análisis cuantitativo de isómeros HETEs por SPE–LC–MS/MS, requiriéndose la aplicación de un factor de corrección de recuperación para

compensar los efectos de la matriz que afectan a la retención de los analitos en el sorbente y disminuyen la ionización.

- Una simple extracción liquido–liquido de paclitaxel y sus metabolitos en tejido peritoneal, plasma y suero, previa a su análisis por LC–MS/MS, ha permitido su cuantificación de forma sensible y selectiva. La aplicación del método a muestras de 13 mujeres afectadas por carcinomatosis peritoneal ovárica y sometidas a un tratamiento con HIPEC puso de manifiesto su potencial para el seguimiento de la evolución de este fármaco. Por lo tanto, la investigación futura sobre la toxicidad del paclitaxel y sus metabolitos, la dosis a aplicar y la cinética de su metabolismo puede apoyarse en este método.

La investigación soportada en análisis metabolómico global permitió extraer las siguientes conclusiones:

- El paquete estadístico "MSCombine" combina datos procedentes de los modos de ionización positivo y negativo en LC–MS de alta resolución, eliminando entidades que introducen información redundante y conservando las entidades moleculares con mayor sensibilidad para cada metabolito tentativo. Este paquete es una alternativa eficaz al análisis de datos de cada polaridad por separado ya que permite utilizar una única matriz, más representativa de la muestra biológica en estudio, y agilizar el proceso.

- El estudio de la preparación de la muestra de aire exhalado condensado (EBC) mediante precipitación de proteínas (PP), extracción en fase sólida (SPE) por sorbentes hidrófilos y lipófilos y liofilización ha demostrado que la muestra analítica resultante de la liofilización es la más versátil, ya que permite la reconstitución en un volumen tan pequeño como se requiera, lo que permite duplicar el número de compuestos detectados en comparación con las otras alternativas (47 compuestos frente a 25). Además, la muestra analítica resultante de esta preparación puede usarse para el almacenamiento, estando lista para el análisis tras reconstitución en el medio apropiado. Por otro lado, la limpieza mediante PP y/o SPE es

innecesaria, ya que en los cromatogramas no aparece ningún efecto de los componentes del EBC eliminados por estas etapas.

- Un modelo PLS-DA con los metabolitos significativos del cáncer de próstata (PCa) encontrados en orina ha permitido discriminar pacientes con PCa de individuos sanos, ambos grupos confirmados por biopsia, con sensibilidad y especificidad del 88.4% y el 92.9%, respectivamente. Este estudio reveló cambios a nivel metabolómico que pueden estar asociados con marcas epigenéticas, ampliamente estudiadas como prometedores marcadores para el diagnóstico de PCa debido a su participación en la expresión o el silenciamiento de genes. Se demuestra así una vez más la capacidad de la metabolómica para proporcionar información de procesos cuyo estudio parecía restringido a la genómica u otras ómicas más próximas a ella que la metabolómica.

- La capacidad de la orina para la detección del PCa, evaluada mediante la combinación de los metabolitos significativos en paneles de 5 variables — 5 metabolitos o 4 metabolitos + los valores en suero del antígeno prostático específico (PSA)— hace recomendable un panel que combine la dimetil-lisina, la acetilputrescina, la citrulina y la dimetilarginina con el valor del PSA sérico, como primera opción para detección del PCa, ya que este panel proporcionó los mejores resultados en cuanto a sensibilidad (97.7%) y especificidad (100%). Los resultados de este modelo de diagnóstico sugieren que el perfil metabólico de la orina puede tener aplicación para el diagnóstico clínico del PCa. El uso de estos paneles permitirá reducir los casos de falsos positivos y, por tanto, la proporción de individuos sometidos a pruebas de confirmación invasivas, así como la detección del PCa en etapas menos avanzadas, e incluso incipientes.

CONCLUSIONS

An in-depth bibliographic search has allowed a critical evaluation of both the state-of-the-art of the treatment of urine sample in metabolomics studies and the existing systems for lung cancer detection in exhaled air samples, showing their deficiencies and/or lacks and providing guidelines to take profit from these bio-fluids in clinical research through metabolomics.

The experimental research in this Thesis was focused on the development of methods/applications based on targeted and untargeted metabolomics analysis in the clinical area, being remarkable the usefulness of liquid chromatography coupled to mass spectrometry in both strategies. The improvement of both sample preparation and data treatment constitutes other of the characteristic aspects of the developed research.

The most outstanding conclusions from the research devoted to targeted analysis can be summarized in the following points:

- The applicability of a method for the determination of eicosanoids in serum based on an SPE–LC–MS/MS platform has been demonstrated by analysis of serum samples from atherosclerotic patients diagnosed with stable angina, non-ST-elevation acute coronary (NSTEMI), and acute myocardial infarction (AMI). The obtained results showed significant levels of the 12-hydroxyeicosanoid acid (12HETE) in patients with AMI as compared with those with stable angina; information that will allow diagnosing and preventing reincidence of AMI patients.
- The use of a deuterated internal standard as 15HETE-d8 is crucial for quantitative analysis of HETE isomers by SPE–LC–MS/MS, which requires a recovery correction factor to compensate for the matrix effects that modify analytes retention on the sorbent and suppress ionization.

- A simple liquid–liquid extraction of paclitaxel and its metabolites from peritoneal tissue, plasma and serum, prior to analysis by LC–MS/MS, has allowed a sensitive and selective quantification of the target compounds. Application of the method to samples from 13 women with ovaric peritoneal carcinomatosis and subjected to HIPEC treatment showed the potential of the method to monitor the drug. Therefore, future research on the toxicity of paclitaxel and its metabolites, doses to be applied and kinetics of paclitaxel metabolism can be supported on the developed method.

The research based on untargeted metabolomics analysis allows the following conclusions:

- The statistical tool "MSCombine" combines data from the positive and negative ionization modes in high resolution LC–MS, and eliminates entities with redundant information by keeping the molecular entities with the highest sensitivity for each tentative metabolite. The proposed package uses a single matrix, more representative of the biological sample under study, and expedites the process as a result.
- The study on sample preparation of exhaled breath condensate (EBC) by proteins precipitation (PP), solid-phase extraction (SPE) by hydrophilic and lipophilic sorbents, and lyophilization has demonstrated that the analytical sample resulting from lyophilization is more versatile, as it allows reconstitution using a volume of liquid as small as required; thus allowing almost duplication of the number of detected compounds when compared with the other alternatives (47 compounds versus 25). In addition, the analytical sample from this step can be used for sample storage, and it is ready for analysis after reconstitution by the appropriate agent. Also, clean-up by PP and/or SPE is unnecessary as any of the sample components removed by these steps do not appear in the chromatograms from lyophilized EBC.

- A PLS-DA model involving the metabolites significant in prostate cancer (PCa) found in urine have allowed discriminating PCa patients and healthy individuals, both groups confirmed by biopsy, with sensitivity and specificity of 88.4% and 92.9%, respectively. This study revealed changes at the metabolomics level that can be associated to epigenetic marks, widely studied as promising biomarkers for PCa diagnostic because their involvement in the gene expression and silencing. Once again it is thus demonstrated the ability of metabolomics to provide information on processes the study of which seemed to be restricted to genomics or other omics closer to it than metabolomics.
- The ability of urine for detection of PCa, evaluated by 5-variable panels — 5 metabolites or 4 metabolites + the values in of prostatic specific antigen (PSA)— makes recommendable a panel including dimethyllisine, acetylputrescine, citrulline and dimethylarginine with the serum PSA value as the first option for PCa detection. This panel provided the best results dealing with sensitivity (97.7%) and specificity (100%). The results of this diagnostic model suggests that the metabolomics profile of urine can be used for the clinical diagnosis of PCa. The use of these panels will allows decreasing false positives; thus resulting in less individuals subjected to invasive confirmation tests, as well as in the detection of PCa in earliest, even incipient, steps of the disease.

Anexos

annexes

ANNEX I

Other publications co-authored by the PhD student

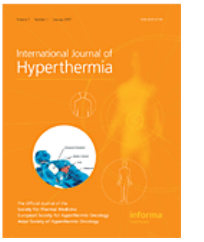
1. Role of Intraperitoneal Intraoperative Chemotherapy with Paclitaxel in the Surgical Treatment of Peritoneal Carcinomatosis from Ovarian Cancer. Hyperthermia versus Normothermia. A randomized controlled trial.

   A. Casado-Adam, S. Rufian-Peña, A. Arjona-Sánchez, F. C. Muñoz-Casares, T. Caro-Cuenca, R. Ortega-Salas, ***M.A. Fernández-Peralbo****, M. D. Luque de Castro, J. Briceño-Delgado*

   Sent to International Journal of Hyperthermia

Taylor & Francis
Taylor & Francis Group

Sent to International
Journal of Hyperthermia



Role of Intraperitoneal Intraoperative Chemotherapy with Paclitaxel in the Surgical Treatment of Peritoneal Carcinomatosis from Ovarian Cancer. Hyperthermia versus Normothermia. A randomized controlled trial

A. Casado-Adam[1,5], S. Rufian-Peña[1,5], A. Arjona-Sánchez[1,5], F. C. Muñoz-Casares,[1,5], T. Caro-Cuenca[2,5], R. Ortega-Salas[2,5], M. A. Fernández-Peralbo[3,5], M. D. Luque de Castro[3,5], J.M. Sánchez-Hidalgo[1,5], C. Hervas-Martinez[4,5], J. Briceño-Delgado[1,5]

[1]*Oncologic and Pancreatic Surgery Unit, University Hospital Reina Sofia, Córdoba, Spain*

[2]*Department of Pathology, University Hospital Reina Sofía, Córdoba, Spain.*

[3]*Department of Analytical Chemistry. University of Córdoba,*

Campus of Rabanales, Annex Marie Curie Building, 14071, Córdoba. Spain.

[4]*Department of Computer Science and Numerical Analysis, University of Córdoba, Campus of Rabanales, Albert Einstein Building, 3rd Floor, 14071, Córdoba, Spain.*

[5]*Institute of Biomedical Research Maimónides (IMIBIC), Reina Sofía University Hospital, University of Córdoba, Córdoba, Spain.*

Abstract

Background

The treatment of ovarian carcinomatosis with cytoreductive surgery and HIPEC is still controversial. The effect and pharmacokinetics of the chemotherapeutics used (specially taxanes) are under consideration at present.

Methods

A phase II, simple blind and randomized controlled trial (NTC02739698) was performed between July 2012 and November 2014. Thirty-two patients with primary or recurrent ovarian carcinomatosis undergoing cytoreductive surgery (CRS) and intraoperative intraperitoneal chemotherapy with paclitaxel (PTX) were included; 16 in hyperthermic and 16 in normothermic conditions. Tissue, serum and plasma samples were taken in every patient before and after intraperitoneal chemotherapy to measure the concentration of PTX. To analyze the inmunohistochemical profile of p53, p27, p21, ki67, PCNA, caspasa-3 and the pathological response a scale of intensity and percentage of expression and a grouped Miller and Payne system were used, respectively. Perioperative characteristics and morbi-mortality were also analyzed.

Results

Main characteristics of patients, surgical morbidity, haematoxicity and nephrotoxicity were similar in both groups. The concentration of paclitaxel in the tissue was higher than that observed in plasma and serum, although no statistically significant differences were found between the two groups. No statistically significant association regarding pathological response and apoptosis (caspasa-3) between both groups was proved. Intraperitoneal PTX reduced the expression of p53, p27, p21, ki67 and PCNA more in hyperthermia group, but not significantly.

Conclusion

The use of intraperitoneal PTX has proven an adequate pharmacokinetics with reduction of cell cycle and proliferation markers globally without finding

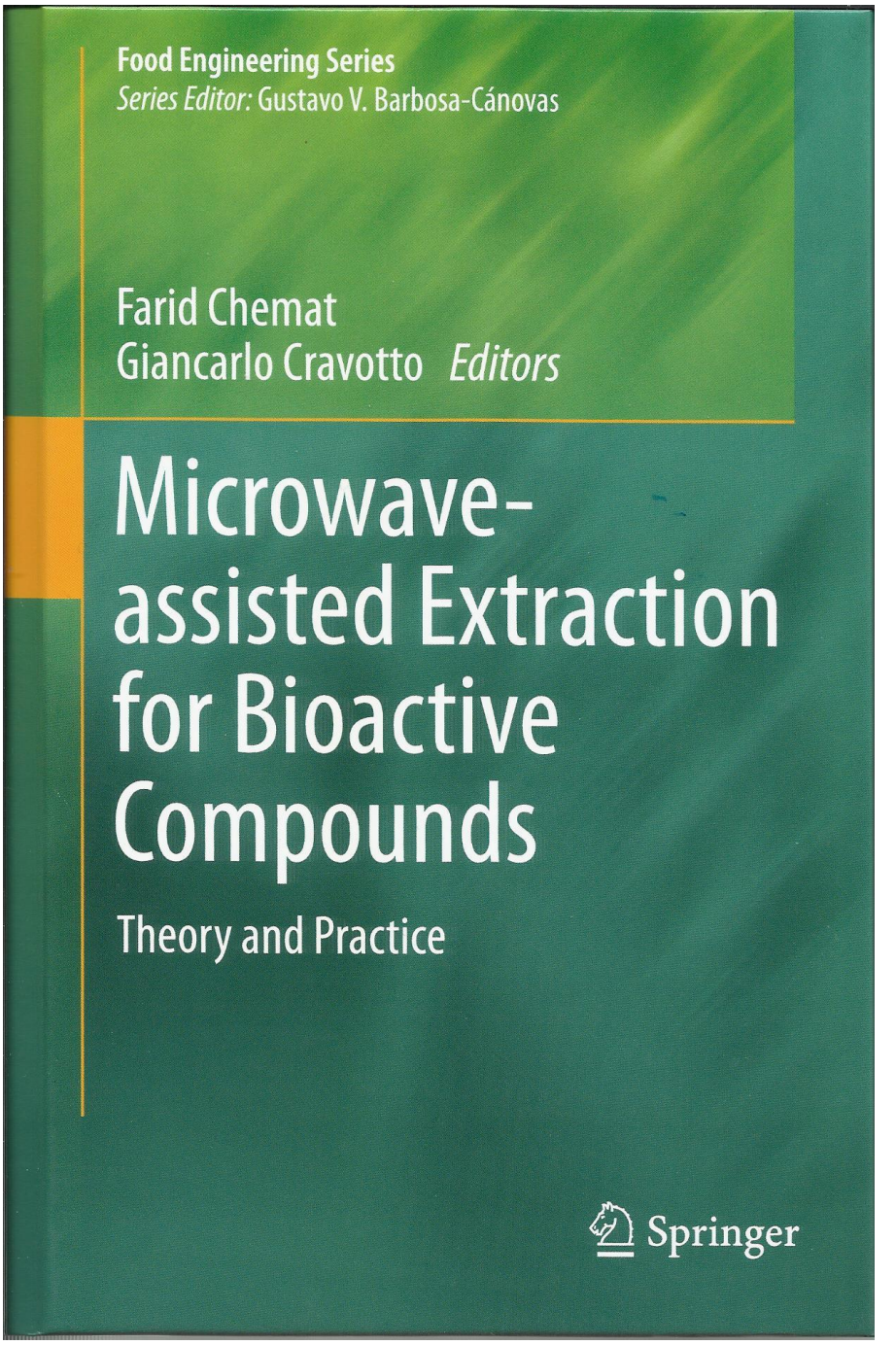
differences between its administration in hyperthermia versus normothermia conditions.

ANNEX II

Book chapters co-authored by the PhD student

1. M.D. Luque de Castro, **M.A. Fernández-Peralbo**, B. Linares-Zea, J. Linares, Chapter 4: The role of microwaves in the extraction of fats and oils, in "Microwave-assisted extraction for bioactive compounds", edited by Farid Chemad and Giancarlo Cravotto, Springer 2012, pp. 69–101.

2. M.D. Luque de Castro, **M.A. Fernández-Peralbo**, Chapter 6: The role of microwaves in omics disciplines, in "Microwave-assisted extraction for bioactive compounds", edited by Farid Chemad and Giancarlo Cravotto, Springer 2012, pp. 127–210.


Food Engineering Series
Series Editor: Gustavo V. Barbosa-Cánovas
Farid Chemat
Giancarlo Cravotto Editors
Microwave-assisted Extraction for Bioactive Compounds
Theory and Practice
Springer

Chapter 4
The Role of Microwaves in the Extraction of Fats and Oils

M.D. Luque de Castro, M.A. Fernández-Peralbo, B. Linares-Zea, and J. Linares

4.1 Introduction

4.1.1 Properties of Fats and Oils

Fats is the general name given to a category of lipids usually referred to as acylglycerides, which are esters in which two or three fatty acids are bonded to a glycerol molecule forming monoglycerides, diglycerides, or triglycerides, respectively. The most common fats are triglycerides; these are triesters of glycerol and fatty acids, and can be solid or liquid at room temperature depending on their particular structure and composition. Although the names "oils," "fats," and "lipids" are widely used to refer to fats, the words oils and fats usually apply to lipids that are liquid and solid, respectively, at ambient temperature.

The types of fatty acids constituting a given ester (particularly, their degree of unsaturation) dictate its physical state at ambient temperature. Thus, fats are formed mainly from fatty acids with a saturated long chain (more than eight carbon atoms) including lauric, myristic, and palmitic acids. Fatty acids are present in bacon, cocoa, and various other foods and are believed to raise the levels of low-density

M.D. Luque de Castro (✉) • M.A. Fernández-Peralbo
Department of Analytical Chemistry, Maimónides Institute of Biomedical Research (IMIBIC), Reina Sofía Hospital, University of Córdoba, Marie Curie Annex Building, Campus of Rabanales, Córdoba E-14071, Spain
e-mail: qa1lucam@uco.es; q32fepem@uco.es

B. Linares-Zea • J. Linares
DEOLEO, S.A., Carretera de Arjona, 4, Andújar E-23740, Spain
e-mail: belinares16@hotmail.com; jolinfer@gmail.com

F. Chemat and G. Cravotto (eds.), *Microwave-assisted Extraction for Bioactive Compounds: Theory and Practice*, Food Engineering Series 4,
DOI 10.1007/978-1-4614-4830-3_4,

Chapter 6
The Role of Microwaves in Omics Disciplines

M.D. Luque de Castro and M.A. Fernández-Peralbo

6.1 Introduction

6.1.1 Definition of Genomics, Proteomics, and Metabolomics

Genomics addresses the study of all genes and their interrelationships to identify their combined influence on the growth and development of an organism. Similarly, proteomics is defined as the study of the expression, localizations, functions, and interactions of all proteins expressed by the genetic material of an organism. Finally, metabolomics is concerned with the quantitative measurement of all low molecular weight metabolites (sugars, amino acids, organic acids, fatty acids, and others) in an organism's cells at a specified time under specific environmental/biological conditions.

Oliver et al. [1] coined the term "metabolomics" in their systematic functional analysis of the yeast genome and proposed the challenge of discovering what each gene product does and how genes in a living yeast cell interact to shape molecular and systems biology. Based on evidence gathered over the past few decades [2], the flow of information from genes to function is linear and translated through transcripts, proteins, and, finally, metabolites.

Microwaves have been used to a dissimilar extent to facilitate work on analytical platforms in the different omics. Thus, metabolomics has for several decades taken advantage of the large number of methods developed under the umbrella of the reductionist theory in molecular biology by using MWs to improve sample preparation steps. By contrast, the other omics have adopted MWs mainly to accelerate

M.D. Luque de Castro (✉) • M.A. Fernández-Peralbo
Department of Analytical Chemistry, Maimónides Institute of Biomedical Research (IMIBIC), Reina Sofía Hospital, University of Córdoba, Marie Curie Annex Building, Campus of Rabanales, Córdoba E-14071, Spain
e-mail: qa1lucam@uco.es; q32fepem@uco.es

F. Chemat and G. Cravotto (eds.), *Microwave-assisted Extraction for Bioactive Compounds: Theory and Practice*, Food Engineering Series 4,
DOI 10.1007/978-1-4614-4830-3_6,

ANNEX III

Patent co-authored by the PhD student

No Patente: P201631044

Biomarcadores para el cribado y diagnóstico de Cáncer de Próstata

M. D. Luque de Castro, F. Priego Capote, M. Calderón Santiago, ***M. A. Fernández Peralbo,*** *M. J. Requena Tapia, E. Gómez Gómez, J. Carrasco Valiente, J. Ruiz García.*

Los autores de la presente invención han analizado la concentración de los distintos metabolitos de la orina en individuos que no padecen cáncer de próstata y en individuos que padecen cáncer de próstata. Han encontrado una serie de marcadores para el diagnóstico de los individuos con cáncer de próstata, diferenciando los sujetos con cáncer de próstata 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 próstata.

ANNEX IV

Communications to congresses and meetings

1. Discriminación entre pacientes con cáncer de próstata y controles con biopsia negativa mediante análisis metabolómico en orina por LC–QTOF

M.A. Fernández Peralbo, M. Calderón Santiago, F. Priego Capote, M.D. Luque de Castro

V CONGRESO CIENTÍFICO DE INVESTIGADORES EN FORMACIÓN DE LA UNIVERSIDAD DE CÓRDOBA (CREANDO REDES)

Córdoba, 2016

Tipo de evento: Comunicación oral en Congreso **Ámbito:** Nacional

---

2. Search for prostate cancer potential biomarkers in urine by LC–QTOF

M.A. Fernández Peralbo, E. Gómez Gómez, M. Calderón Santiago, J. Carrasco Valiente, M. J. Requena Tapia, J. Ruiz García, F. Priego Capote, M.D. Luque de Castro

7th YOUNG INVESTIGATORS MEETING

Córdoba, 2016

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

---

3. Role of the temperature on the pharmacokinetics of paclitaxel in intraperitoneal intraoperative administration in the surgical treatment of peritoneal carcinomatosis from ovarian cancer. Hyperthermia versus normothermia

A. Casado Adam, F.C. Muñoz Casares; A. Arjona Sánchez; M.A. Fernández Peralbo, M.D. Luque de Castro, M.C. Muñoz Villanueva, J.M. Sánchez Hidalgo, T. Caro cuenca, E. Ortega Salas, S. Rufián Peña

7th YOUNG INVESTIGATORS MEETING

Córdoba, 2016

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

---

4. Metabolómica y cáncer de próstata: biomarcadores en orina

E. Gómez Gómez, M.A. Fernández Peralbo, J. Carrasco Valiente, M. Calderón Santiago, J. Ruiz García, F. Priego Capote, M.D. Luque de Castro, M. J. Requena Tapia

LXXXI CONGRESO NACIONAL DE UROLOGÍA

Toledo, 2016

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

---

5. Metabolomics in prostate cancer diagnosis

E. Gómez Gómez, M.A. Fernández Peralbo, J. Carrasco Valiente, M. Calderón Santiago, J. Ruiz García, F. Priego Capote, M.D. Luque de Castro, M. J. Requena Tapia

8th EUROPEAN MULTIDISCIPLINARY MEETING ON UROLOGICAL CANCERS

Milan, 2016

Tipo de evento: Póster en Congreso **Ámbito:** Internacional

---

6. Metabolomics in prostate cancer

E. Gómez Gómez, J. Carrasco Valiente, M.A. Fernández Peralbo, M. Calderón Santiago, J. Ruiz García, F. Priego Capote, M.D. Luque de Castro, M. J. Requena Tapia

XXIX CONGRESO DE LA ASOCIACIÓN ANDALUZA DE UROLOGÍA

Córdoba, 2016

Tipo de evento: Comunicación oral en Congreso **Ámbito:** Nacional

---

7. Resultados preliminares en ensayo clínico. Papel de la quimioterapia intraoperatoria intraperitoneal con paclitaxel en el tratamiento quirúrgico radical de la carcinomatosis peritoneal de origen ovárico: hipertermia *vs*. normotermia.

A. Casado Adam, F.C. Muñoz Casares; A. Arjona Sánchez; M.A. Fernández Peralbo, M.D. Luque de Castro, M.C. Muñoz Villanueva, J.M. Sánchez Hidalgo, T. Caro cuenca, E. Ortega Salas, S. Rufián Peña

XVI REUNIÓN DEL GRUPO ESPAÑOL DE CIRUGÍA ONCOLÓGICA PERITONEAL

Madrid, 2015

Tipo de evento: Comunicación oral en Congreso **Ámbito:** Nacional

---

8. Targeted analysis of paclitaxel and its major metabolites in serum, plasma and tissue from woman with ovarian cancer by SPE–LCMS/MS

M.A. Fernández Peralbo, F. Priego Capote, M.D. Luque de Castro

XVIII REUNIÓN DE LA SOCIEDAD ESPAÑOLA DE QUÍMICA ANALÍTICA

Úbeda, 2013

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

---

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, 2013

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

---

10. Mass spectrometry as a fingerprinting metabolomic tool applied to the study of cardiovascular diseases

M.A. Fernández Peralbo, F. Priego Capote, M.D. Luque de Castro

XXIII REUNIÓN NACIONAL DE ESPECTROSCOPÍA. VII CONGRESSO IBÉRICO DE ESPECTROSCOPÍA

Córdoba, 2012

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

---

11. Plataformas para análisis cualitativo/cuantitativo en metabolómica clínica y nutricional

F. Priego Capote, M.D. Luque de Castro, C. Ferreiro Vera, M. Calderón Santiago, M. Orozco Solano, M. A. Fernández Peralbo

XII REUNIÓN DEL GRUPO REGIONAL ANDALUZ DE LA SOCIEDAD ESPAÑOLA DE QUÍMICA ANALÍTICA

Málaga, 2012

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

---

12. Solid–liquid and liquid-liquid extraction methods for oil enrichment

M.A. Fernández Peralbo, M.D. Luque de Castro

THE INTERNATIONAL CONFERENCE ON NATURAL PRODUCTS.

Castrex Cedex (Francia) 2011

Tipo de evento: Póster en Congreso **Ámbito:** Internacional

---

13. Métodos de lixiviación y de extracción líquido–líquido para el enriquecimiento de aceites

M.A. Fernández Peralbo, M.D. Luque de Castro

XII REUNIÓN DEL GRUPO REGIONAL ANDALUZ DE LA SOCIEDAD ESPAÑOLA DE QUÍMICA ANALÍTICA

Córdoba, 2010

Tipo de evento: Póster en Congreso **Ámbito:** Nacional

V Congreso Científico de Investigadores en Formación de la Universidad de Córdoba,
CREANDO REDES

UNIVERSIDAD D CÓRDOBA

educo
Escuela de Doctorado de la Universidad de Córdoba

eidA3
escuela internacional de doctorado en agroalimentación
SEDE CÓRDOBA

Córdoba, 30 de noviembre, 1 y 2 de diciembre de 2016

En el Rectorado de la Universidad de Córdoba

Avda. Medina Azahara 5

Discriminación entre pacientes con cáncer de próstata y controles con biopsia negativa mediante análisis metabolómico en orina por LC–QTOF

M.A. Fernández-Peralbo, M. Calderón Santiago, F. Priego-Capote, M.D. Luque de Castro

Universidad de Córdoba. Facultad de Ciencias. Departamento de Química Analítica. E-mail: q32fepem@uco.es

Summary

The existing clinical biomarkers for prostate cancer (PCa) diagnosis are far from ideal (e.g., the prostate specific antigen serum –PSA– level suffers from lack of specificity, providing frequent false positives leading to over-diagnosis). A key step in the search for minimum invasive tests to complement or replace PSA should be supported on the changes experienced by the biochemical pathways in PCa patients as compared to negative biopsy control individuals. In this research a comprehensive global analysis by LC–QTOF was applied to urine from 62 patients with a clinically significant PCa and 42 healthy individuals, both groups confirmed by biopsy. An unpaired *t*-test (*p*-value<0.05) provided 42 significant metabolites tentatively identified in urine, used to develop a partial least squares–discriminant analysis (PLS–DA) model characterized by 86.1% and 92.9% of sensitivity and specificity, respectively. Among the 42 significant metabolites 39 were present at lower concentrations in PCa patients than in control individuals, while only three reported higher concentrations in PCa patients. The connection among the biochemical pathways in which they are involved (DNA methylation, epigenetic marks on histones and RNA cap methylation) could explain the concentration changes with PCa and supports, once again, the role of metabolomics in upstream processes.

Resumen

Los biomarcadores para diagnosticar el cáncer de próstata (CaP) no son ideales. Por ejemplo, el marcador PSA carece de suficiente especificidad, proporcionando falsos positivos que implican a un exceso de tratamiento. Un paso clave en la búsqueda de pruebas mínimamente invasivas para complementar o sustituir el PSA debe estar soportado en los cambios experimentados por las rutas bioquímicas en pacientes con CaP en comparación con individuos control. En esta investigación se llevó a cabo un análisis global por LC–QTOF de la orina de 62 pacientes con CaP y 42 individuos sanos, ambos grupos confirmados por biopsia. Un test de la *t* (*p* <0,05) proporcionó 42 metabolitos significativos en la orina, que se utilizaron para desarrollar un modelo PLS-DA, caracterizado con 86,1% de sensibilidad y 92,9% de especificidad. Entre los metabolitos, 39 están presentes en concentraciones más bajas en pacientes con CaP que en los individuos control, y sólo tres están más concentrados en pacientes con CaP. La conexión entre las rutas bioquímicas en las que están involucrados (la metilación del ADN, marcas epigenéticas en histonas y la metilación del ARN-cap) podría explicar los cambios en la concentración con CaP y soportar el papel de la metabolómica en los procesos anteriores.

Introducción

El cáncer de próstata (CaP) es el segundo cáncer más frecuente y la tercera causa principal de muerte por cáncer en hombres [1]. En la actualidad, el PSA es el marcador usado para la detección del CaP, que no es capaz de distinguir los tumores indolentes clínicamente relevantes [2]. Las alteraciones genéticas causadas por factores epigenéticos pueden desempeñar un papel significativo en el CaP [3]. Se ha demostrado que existen aberraciones en las modificaciones post-traduccionales (MPTs) de las histonas en células de CaP, en particular, la acetilación y la metilación de residuos de lisina y arginina específicos están siendo investigadas como predictores del riesgo de CaP [4]. Dado que el perfil de metabolitos refleja con exactitud el total del estado celular, la metabolómica se ha convertido en un enfoque más integrador en comparación con otras ómicas. La orina se presenta como el biofluido ideal en este estudio debido a su carácter no invasivo, recogida fácil y a la presencia de productos finales del metabolismo del cáncer liberados en la orina o dentro de las células prostáticas liberadas en la uretra a través de los conductos prostáticos.

Metodología

La recogida de la orina de 104 pacientes en ayunas se realizó en el departamento de urología del Hospital Reina Sofía. Estos individuos se agruparon en dos categorías, como muestra la Tabla 1: los pacientes con CaP y los controles con biopsia negativa. Después de su descongelación a temperatura ambiente, las muestras de orina se agitaron durante 1 min y se centrifugaron a 21.000 × *g* durante 5 min. A continuación, 50 µL del sobrenadante se diluyeron 1:2 (v/v) con formiato de amonio 5 mM en agua antes del análisis por LC–MS/MS. MetaboAnalyst 3.0, una herramienta web para estudios de metabolómica, se utilizó para el análisis e interpretación de los datos.

Tabla 1. Características de la población.

Variable	CaP (n= 62)	Biopsia negativa (n= 42)	Valor de p
Edad (años)	71 (64-80)	62.5 (57.8-69)	<0.05
Tacto rectal patológico	36 (58.1%)	0	<0.05
PSA(1) (ng/mL)	11 (6.9-29.75)	4.1 (3.6-5.8)	
<3	0	3 (7.1%)	<0.05
3-10	27 (43.5%)	35 (83.3%)	
>10	35 (56.5%)	4 (9.5%)	
Volumen prostático (mL)	30 (22-40)	27.5 (24-47)	>0.05
N° Biopsia (1°)	52 (83.9%)	38 (90.5%)	>0.05
PSA(2) (ng/mL)	11 (6.9-40.7)	2.3 (1.8-2.7)	
<3	0	41 (97.6%)	<0.05
3-10	28 (45.2%)	1 (2.4%)	
>10	34 (54.8%)	0	
Número de cilindros	12 (12-13)	12 (12-12)	>0.05
Gleason			
6	4 (6.5%)		
7	26 (41.9%)		
≥8	32 (51.6%)		

PSA (1): PSA indicación de biopsia; PSA (2): PSA el día de la biopsia.

Resultados

Tras la extracción y normalización de los compuestos detectados en la orina, 42 compuestos tentativamente identificados fueron significativos entre los dos grupos en estudio. De ellos el 38.1% son aminoácidos y derivados y el 12% derivados de bases nitrogenadas. Además, cabe mencionar que el 21,4% de los metabolitos son compuestos acetilados tales como la cetilacitrullina, la acetilarginina, la acetilputrescina o la acetillisina, y el mismo porcentaje está representado por aminoácidos metilados o bases nitrogenadas como la metilguanosina, la metilisina, la metilhistidina o la dimetilarginina.

Se construyó un modelo basado en la discriminación PLS-DA usando los metabolitos significativos identificados tentativamente; modelo que mostró una clara separación entre los pacientes con CaP y los individuos control, como muestra la Figura 1.

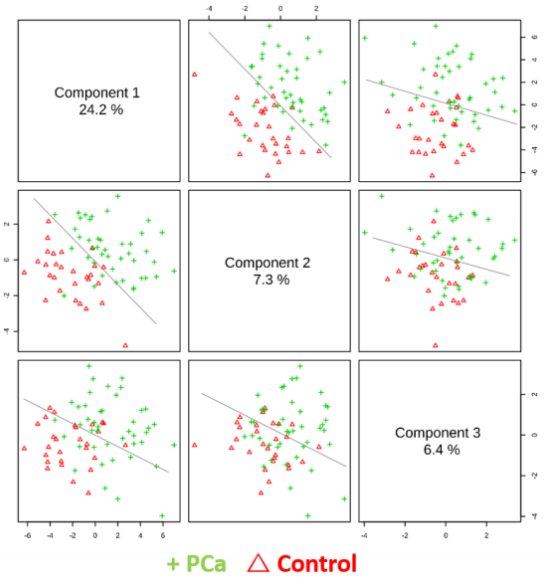


Figura 1. PLS para la discriminación entre pacientes con CaP e individuos control.

Los valores de sensibilidad y especificidad obtenidos en el modelo de predicción teniendo en cuenta el 70% de las muestras seleccionadas al azar fueron 86,1% y 92,9%, respectivamente. La estabilidad del modelo se evaluó mediante la validación externa utilizando el 30% de las muestras restantes. La validación proporcionó una sensibilidad y especificidad del 73,7% y 78,6%, respectivamente. El metabolismo de los aminoácidos está predominantemente alterado por la presencia de CaP. De hecho, la concentración de 5 aminoácidos y 25 metabolitos relacionados, con especial énfasis en aminoácidos acetilados y metilados, está alterada en individuos con CaP. Las principales rutas metabólicas alteradas por la aparición del CaP son la de la lisina, histidina, arginina, taurina, alanina, aspartato y glutamato, así como la de áminoácidos aromáticos, particularmente el triptófano, la fenilalanina y la tirosina.

Por otro lado, 5 derivados de purinas y pirimidinas también se alteran en los casos de CaP. En concreto dos de ellos, el ácido úrico y la 7-metilguanina, son objeto de regulación en el CaP.

De acuerdo con la lista de los metabolitos significativos, tres procesos biológicos pueden estar implicados en la presencia del CaP: la metilación del ADN, las marcas epigenéticas en las proteínas histonas y la metilación del casquete del RNA.

Conclusiones

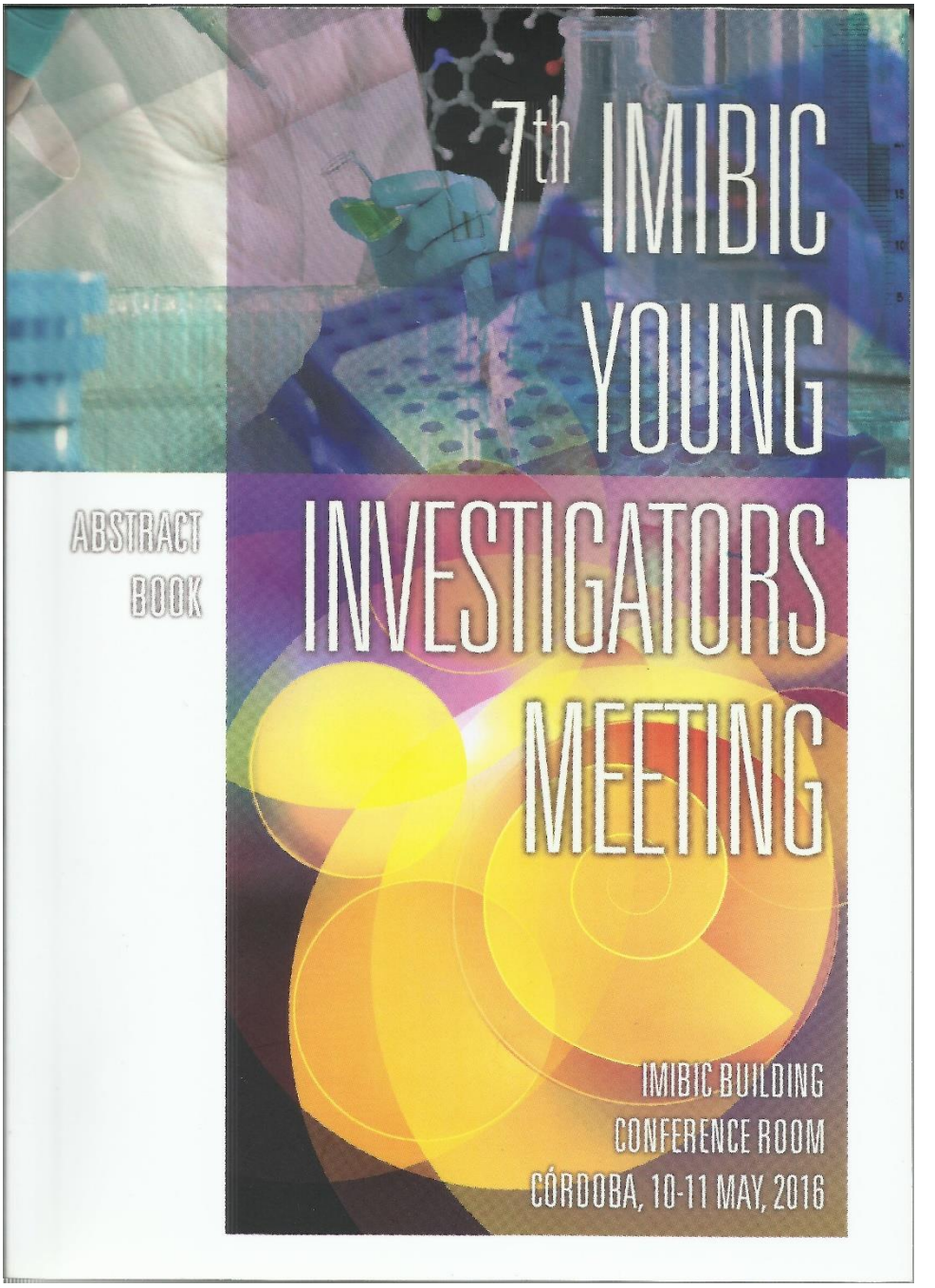
En esta investigación, un análisis global por LC–QTOF de la orina de pacientes con CaP e individuos control con biopsia negativa ha permitido la discriminación entre ambos grupos. Un test de la *t* no apareado (p <0,05) proporcionó 42 metabolitos significativos que intervienen en procesos biológicos, tales como la regulación de las marcas epigenéticas, principalmente la metilación del ADN y las modificaciones de las histonas, así como la metilación del casquete del ARN. Un modelo PLS-DA permitió discriminar ambos grupos de individuos con valores de sensibilidad y especificidad del 86,1% y 92,9%, respectivamente. Este estudio reveló cambios a nivel metabolómico que se pueden asociar con marcas epigenéticas y que han sido ampliamente estudiados como marcadores prometedores para el diagnóstico de CaP debido a su participación en la expresión o silenciamiento de genes. El proceso de metilación del casquete de ARN también proporcionó diferencias significativas entre los pacientes con CaP y los controles, lo que incide en la importancia de la transcripción del ARN. Por lo tanto, el análisis de orina reveló cambios en metabolitos que pueden ser interpretados de acuerdo con las alteraciones en la expresión y la transcripción de los genes, lo que demuestra una vez más la capacidad de la metabolómica para proporcionar información sobre los procesos anteriores.

Bibliografía

[1] C. E. Desantis et al. Cancer treatment and survivorship statistics 2014. *Cancer J Clin* 64(4) (2014) 252-271.
[2] H. B Carter et al. Early detection of prostate cancer: AUA guideline. *J Urol* 190(2) (2014) 419-426.
[3] C. Jerónimo et al. Epigenetics in prostate cancer: Biologic and clinical relevance. *Eur Urol* 60 (2011) 753-766.
[4] D. B. Seligson et al. Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* 435 (2005) 1262-1266.

Agradecimientos

Agradecemos la ayuda financiera soportada por el Ministerio Español de Economía y Competitividad (MINECO), la Junta de Andalucía y el programa FEDER en los proyectos: “Development of methods for early cancer detection”, FQM-1602 and CTQ2015-68813-R.


7th IMIBIC
YOUNG
INVESTIGATORS
MEETING
ABSTRACT
BOOK
IMIBIC BUILDING
CONFERENCE ROOM
CÓRDOBA, 10-11 MAY, 2016

P9. Search for prostate cancer potential biomarkers in urine by LC–QTOF.

Authors: *M.A. Fernández-Peralbo, E.Gómez-Gómez, M. Calderón Santiago, J. Carrasco-Valiente, M.J. Requena-Tapia, J. Ruiz-García, F. Priego-Capote, M.D. Luque de Castro.*

Group: GC21 Metabolomics Identification/Quantification of Bioactive Components.

The accurate detection of prostate cancer (PCa) is urgently needed to reduce overdiagnosis and over-treatment, while maintaining a reduction in mortality. In this sense, the existing clinical biomarkers for PCa diagnosis are far from ideal, as is the case with prostate specific antigen serum (PSA) level that suffers from lack of specificity, thus giving place to frequent false positives. For this reason, minimum invasive tests based on markers from biological samples such as blood or urine that could complement or replace PSA represent a goal in PCa research. Recent advances in analytical instrumentation have facilitated the development of metabolomics, providing an insight into the metabolic state and biochemical processes of the organism. In this research, a comprehensive global analysis by LC–QTOF have been applied to urine from 62 patients with a clinically significant PCa and 42 healthy individuals, both groups confirmed by biopsy. An unpaired t-test (p-value<0.05) revealed 42 significant metabolites tentatively identified in urine, that were considered to develop a partial least squares–discriminant analysis (PLS–DA) model, characterized by 86.05 and 92.85% of sensitivity and specificity, respectively. Then, an external validation using the 30% of the samples reported a sensitivity and specificity of 73.68% and 78.57%, respectively. These 42 urinary metabolites are involved in key biochemical pathways like amino acids metabolism (e.g. lysine degradation, taurine, and tryptophan metabolism), the urea cycle and purine and pyrimidine metabolism, among others. These results indicate that deregulation of amino acids metabolism may be specific for the PCa metabolic phenotype, which can be also associated to abnormal cell growth and intensive cell proliferation. These preliminary results emphasize the necessity for a large scale study to validate the proposed metabolites, as well as to include urine from less advanced PCa stages.

P19. Role of temperature on the pharmacokinetics of paclitaxel in intraperitoneal intraoperative administration in the surgical treatment of peritoneal carcinomatosis from ovarian cancer. Hyperthermia versus Normothermia.

Authors: *Casado-Adam A, Muñoz-Casares F.C, Arjona-Sánchez A, Fernández Peralbo M.A, Luque de Castro M.D, Muñoz Villanueva M.C, Sánchez Hidalgo J.M, Caro Cuenca T, Ortega Salas R, Rufián Peña S.*

Groups: GC18 Translational research in surgery of solid organ transplants.

Background: The treatment of ovarian carcinomatosis with cytoreductive surgery and HIPEC (hyperthermic intraperitoneal chemotherapy) is still controversial. The effect and pharmacokinetics of drugs used are under consideration at present. Our objetive is to analyze the effect of temperature on palitaxel´s pharmacokinetics (tissue, plasma and serum concentration) in his intraperitoneal intraoperative administration. Patients and method: Randomised, single centre, simple blind, prospective clinical trial was performed between July 2012 and November 2014. 32 patients diagnosed with primary or recurrent ovarian peritoneal carcinomatosis where incluyed and underwent cytoreductive surgery and intraoperative intraperitoneal chemotherapy with paclitaxel (16 in hypertermic and 16 in normothermic conditions). In every patient, tissue, serum and plasma samples were taken before and after intraperitoneal chemotherapy to measure the concentration of paclitaxel. Also demographic, surgical and morbidity characteristics were analyzed. Results: Degree of cytoreduction achieved, peritoneal carcinomatosis index (PCI), type of peritonectomy procedure and other demographic data were similar between the two groups except BMI (body mass index) , which was higher in the normothermia group against hyperthermia (29(5) against 25(3), respectively; p=0,007). Surgical morbidity (Clavien >III), haematotoxicity and nephrotoxicity according CTCAE v4.0 were similar in both groups. Paclitaxel levels in serum and plasma immediately after administration were slightly superior in the hyperthermia group, while levels in serum, plasma and tissue 1 hour after administration were higher in the normothermia group. These differences were not statistically significant (p=0,372; p=0,856; and p=0,525; p=0,119 and p=0,252, respectively). Conclusion: Paclitaxel has proven adequate pharmacokinetics to treat peritoneal carcinomatosis of ovarian origin, reaching optimal concentrations in tissue and minimal in serum and plasma when administered in the peritoneal cavity during cytoreductive surgery. However there were no differences observed between the group treated with normothermia and the group treated with hyperthermia.


LXXXI
CONGRESO
NACIONAL
DE UROLOGIA
Del 15 al 18
Junio 2016
TOLEDO

Resumen nº: C-1

Título: **Metabolómica y cáncer de próstata: Biomarcadores en orina**

Autores: Gómez-Gómez E; Fernández-Peralbo M.A*; Carrasco-Valiente J; Calderón Santiago M. *; Ruiz-García J; Priego-Capote F*; Luque de Castro M.D*; Requena-Tapia M.J

Institución: *Departmento de Urologia (IMIBIC), Hospital Universitario Reina Sofía, Universidad de Córdoba, Córdoba, España * Departmento de Química analítica (IMIBIC), Campus de Rabanales, Universidad de Córdoba, Córdoba, España*

Introducción

El cribado del cáncer de próstata (CaP) es un tema controvertido debido al problema causado por el sobrediagnóstico y sobretratamiento. En este sentido, los biomarcadores existentes en la actualidad, como el PSA, no son ideales debido a su falta de especificidad. Por esta razón, la identificación de nuevos biomarcadores con mayor sensibilidad y especificidad obtenidos a través de pruebas mínimamente invasivas en muestras biológicas tales como sangre u orina que complementen o sustituyan al PSA, representa una meta en la investigación del CaP. Recientes avances en la instrumentación analítica han favorecido el desarrollo de la metabolómica, proporcionando mayor información sobre la expresión de metabolitos y sus procesos bioquímicos.

Materiales y métodos

Se llevó acabo el análisis global mediante LC-QTOF de orina procedente de 62 pacientes con CaP clínicamente significativo y 42 controles (biopsia negativa y PSA posterior < 3 ng/dL).

El análisis estadístico de los datos proporcionó los metabolitos significativamente diferentes entre ambas poblaciones. Se desarrolló un modelo discriminante mediante análisis parcial de mínimos cuadrados (PLS-DA).

Resultados

Ambas poblaciones diferían en edad, siendo mayores los pacientes con cáncer (71 vs 63), además de presentar un mayor valor de PSA (11 ng/ml vs 4,1 ng/ml). Casi la mitad de pacientes con cáncer (43,5%) presentaban un PSA < 10 ng/ml en la indicación de biopsia frente a un 83% de sanos.

Cuarenta y dos metabolitos significativos fueron tentativamente identificados. El PLS-DA en el 70% de la población proporcionó una sensibilidad y especificidad del 86,05% y

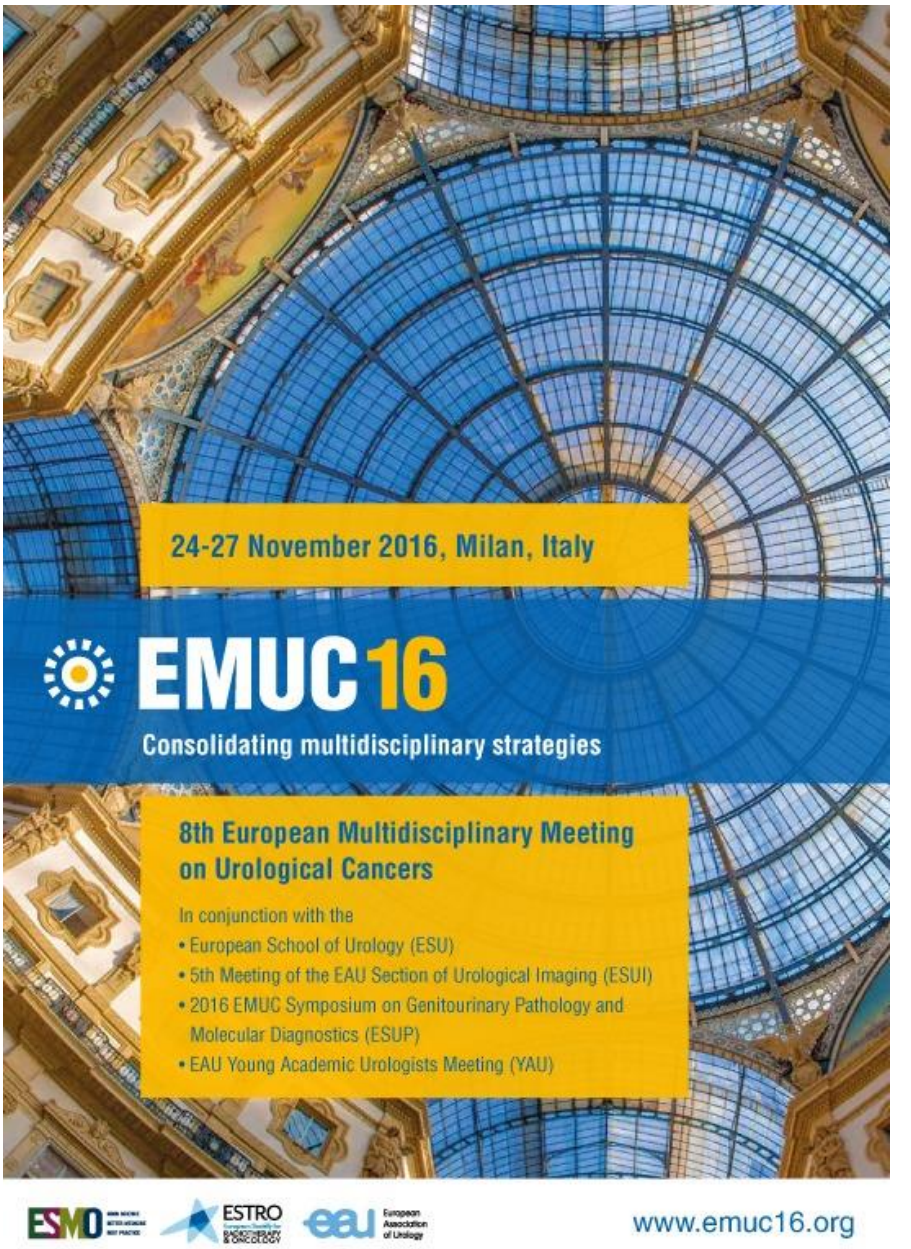
92,85% respectivamente. Posteriormente, se realizó una validación externa con el 30% de la población restante obteniéndose una sensibilidad del 73,68% y una especificidad del 78,57%.

Conclusiones

Estos resultados muestran que la desregulación del metabolismo de aminoácidos puede ser específico del fenotípico metabólico del CaP y que podría utilizarse un panel de metabolitos como diagnóstico de CaP en orina. Se están desarrollando paneles de biomarcadores de entre 3 y 5 metabolitos en una población de práctica clínica habitual.

Agradecimientos

Se agradece la ayuda financiera soportada por el Ministerio de Economía y Competitividad (MINECO) y el programa FEDER en los proyectos "Development of methods for early cancer detection" y "Optimización y aplicación de plataformas metabolómicas de análisis de biofluidos no invasivos para la búsqueda de biomarcadores de diagnóstico precoz de cáncer de próstata


24-27 November 2016, Milan, Italy
EMUC16
Consolidating multidisciplinary strategies
8th European Multidisciplinary Meeting on Urological Cancers
In conjunction with the
• European School of Urology (ESU)
• 5th Meeting of the EAU Section of Urological Imaging (ESUI)
• 2016 EMUC Symposium on Genitourinary Pathology and Molecular Diagnostics (ESUP)
• EAU Young Academic Urologists Meeting (YAU)
ESMO
ESTRO
eau European Association of Urology
www.emuc16.org

P066

Metabolomics in prostate cancer diagnosis

By:

Gomez Gomez E.[1], Fernández-Peralbo M.A.[2], Carrasco-Valiente J.[1], Calderón-Santiago M.[2], Ruiz-García J.[1], Priego-Capote F.[2], Luque De Castro M.D.[2], Requena-Tapia M.J.[1]

Institutes:

[1]Reina Sofia University Hospital/IMIBIC, Dept. of Urology, Cordoba, Spain, [2]IMIBIC/ Reina Sofía University Hospital, Dept. of Analytical Chemistry, Cordoba, Spain

Introduction & Objectives

The accurate detection of Prostate Cancer (PCa) is urgently needed to reduce overdiagnosis and over-treatment, while maintaining a reduction in mortality. In this sense, the existing clinical biomarkers for PCa diagnosis are far from ideal such as the Prostate Specific Antigen serum (PSA) level that suffers from lack of specificity, providing frequent false positives. For this reason, minimum invasive tests in biological samples such blood or urine based on more sensitive and specific biomarkers that can complement or replace PSA represents a goal in PCa research. Recent advancements in analytical instrumentation have made possible the development of metabolomics, providing an insight into the metabolic state and the biochemical processes of the organism.

Material & Methods

In this research we applied a comprehensive global analysis by liquid-chromatography – quadrupole time of flight (LC–QTOF) of urine from 62 patients

with a clinically significant PCa and 42 healthy individuals, both groups confirmed by biopsy. Morning urine samples without prostate stimulation were collected from patients before undergoing prostate biopsy.

Results

A t-test unpaired (p-value<0.05) provided 42 significant metabolites tentatively identified in urine, that were considered to develop a Partial Least Squares–Discriminant Analysis (PLS–DA) model, characterized by 86.05 and 92.85% of sensitivity and specificity, respectively. Then, an external validation using the 30% of the samples reported a sensitivity and specificity of 73.68% and 78.57%, respectively. These 42 urinary metabolites are involved in biochemical pathways like amino acids metabolism such as lysine degradation, taurine, and tryptophan metabolism, as well as urea cycle and purine and pyrimidine metabolism, among others. These results indicate that deregulation of amino acids metabolism may be specific for PCa metabolic phenotype, which can be also associated with abnormal cell growth and intensive cell proliferation.

Conclusions

These results indicate that deregulation of amino acids metabolism may be specific for prostate cancer metabolic phenotype. Also, they emphasize the necessity of a large scale study to validate the proposed metabolites.


XXIX
Congreso de
la Asociación
Andaluza de
Urología
Córdoba
13 y 14 de Octubre de 2016
Parador de la Arruzafa

Metabolomics in prostate cancer

Gómez Gómez E.[1], Carrasco-Valiente J.[1], Fernández-Peralbo M.A.[2], Calderón-Santiago M.[2], Ruiz-García J.[1], Priego-Capote F.[2], Luque de Castro M.D.[2], Requena-Tapia M.J.[1]

[1]Reina Sofia University Hospital/IMIBIC, Dept. of Urology, Cordoba, Spain, [2]IMIBIC/ Reina Sofía University Hospital, Dept. of Analytical Chemistry, Cordoba, Spain

Introduction & Objectives

The accurate detection of Prostate Cancer (PCa) is urgently needed to reduce overdiagnosis and over-treatment, while maintaining a reduction in mortality. In this sense, the existing clinical biomarkers for PCa diagnosis are far from ideal such as the Prostate Specific Antigen serum (PSA) level that suffers from lack of specificity, providing frequent false positives. For this reason, minimum invasive tests in biological samples such blood or urine based on more sensitive and specific biomarkers that can complement or replace PSA represents a goal in PCa research. Recent advancements in analytical instrumentation have made possible the development of metabolomics, providing an insight into the metabolic state and the biochemical processes of the organism.

Material & Methods

In this research we applied a comprehensive global analysis by liquid-chromatography – quadrupole time of flight (LC–QTOF) of urine from 62 patients with a clinically significant PCa and 42 healthy individuals, both groups confirmed by biopsy. Morning urine samples without prostate stimulation were collected from patients before undergoing prostate biopsy.

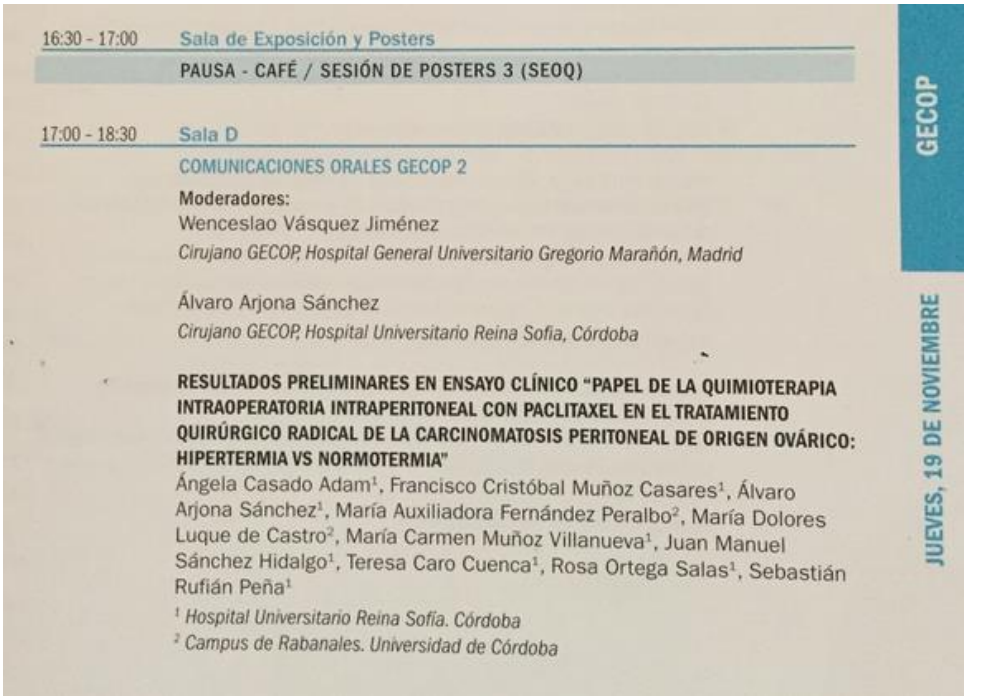
Results

A t-test unpaired (p-value<0.05) provided 42 significant metabolites tentatively identified in urine, that were considered to develop a Partial Least Squares–Discriminant Analysis (PLS–DA) model, characterized by 86.05 and 92.85% of sensitivity and specificity, respectively. Then, an external validation using the 30% of the samples reported a sensitivity and specificity of 73.68% and 78.57%, respectively. These 42 urinary metabolites are involved in biochemical pathways like amino acids metabolism such as lysine degradation, taurine, and tryptophan metabolism, as well as urea cycle and purine and pyrimidine metabolism, among others. These results indicate that deregulation of amino acids metabolism may be specific for PCa metabolic phenotype, which can be also associated with abnormal cell growth and intensive cell proliferation.

Conclusions

These results indicate that deregulation of amino acids metabolism may be specific for prostate cancer metabolic phenotype. Also, they emphasize the necessity of a large scale study to validate the proposed metabolites.


REGE CAROLO
ANNO
MDCCLXXVIII
MADRID
IV Congreso Nacional
SEOQ
Sociedad Española de Oncología Quirúrgica
VI Reunión
Grupo Español de Cirugía Oncológica Peritoneal
GECOP
19 y 20 de Noviembre de 2015
PROGRAMA FINAL



16:30 - 17:00 Sala de Exposición y Posters

PAUSA - CAFÉ / SESIÓN DE POSTERS 3 (SEOQ)

17:00 - 18:30 Sala D

COMUNICACIONES ORALES GECOP 2

Moderadores:

Wenceslao Vásquez Jiménez

Cirujano GECOP, Hospital General Universitario Gregorio Marañón, Madrid

Álvaro Arjona Sánchez

Cirujano GECOP, Hospital Universitario Reina Sofía, Córdoba

RESULTADOS PRELIMINARES EN ENSAYO CLÍNICO "PAPEL DE LA QUIMIOTERAPIA INTRAOPERATORIA INTRAPERITONEAL CON PACLITAXEL EN EL TRATAMIENTO QUIRÚRGICO RADICAL DE LA CARCINOMATOSIS PERITONEAL DE ORIGEN OVÁRICO: HIPERTERMIA VS NORMOTERMIA"

Ángela Casado Adam[1], Francisco Cristóbal Muñoz Casares[1], Álvaro Arjona Sánchez[1], María Auxiliadora Fernández Peralbo[2], María Dolores Luque de Castro[2], María Carmen Muñoz Villanueva[1], Juan Manuel Sánchez Hidalgo[1], Teresa Caro Cuenca[1], Rosa Ortega Salas[1], Sebastián Rufián Peña[1]

[1] *Hospital Universitario Reina Sofía. Córdoba*

[2] *Campus de Rabanales. Universidad de Córdoba*

GECOP

JUEVES, 19 DE NOVIEMBRE

RESULTADOS PRELIMINARES EN ENSAYO CLÍNICO "PAPEL DE LA QUIMIOTERAPIA INTRAPERITONEAL INTRAOPERATORIA CON PACLITAXEL EN EL TRATAMIENTO QUIRÚRGICO RADICAL DE LA CARCINOMATOSIS PERITONEAL DE ORIGEN OVÁRICO: HIPERTERMIA VS NORMOTERMIA".

Nº EudraCT: 2011-004373-89.

Autores: A. Casado Adam[1], F.C. Muñoz Casares[1], A. Arjona Sánchez[1], M.A. Fernández Peralbo[2], M.D. Luque de Castro[2], M.C Muñoz Villanueva[3], J.M. Sánchez Hidalgo[1], T. Caro Cuenca[4], R Ortega Salas[4], S Rufián Peña[1].

[1]. Unidad de Cirugía Oncológica Peritoneal. Hospital Universitario Reina Sofía. Córdoba.

[2]. Departamento de Química Analítica. Campus de Rabanales. Universidad de Córdoba.

[3]. Unidad de Metodología y Bioestadística del IMIBIC. Hospital Universitario Reina Sofía. Córdoba.

[4]. Servicio de Anatomía Patológica. Hospital Universitario Reina Sofía. Córdoba.

Objetivos: El tratamiento de la carcinomatosis de ovario con cirugía citorreductora y HIPEC es aún controvertido. El efecto de los quimioterápicos empleados así como su farmacocinética están en estudio en la actualidad. Nuestro objetivo es analizar el efecto que ejerce en la farmacocinética (concentración tisular, plasmática y sérica) la administración por vía intraperitoneal intraoperatoria del quimioterápico paclitaxel en condiciones de hipertermia versus normotermia.

Material y Métodos: Ensayo clínico unicéntrico, aleatorizado, prospectivo, simple ciego, realizado entre julio de 2012 y noviembre de 2014. Se incluyeron 32 pacientes con carcinomatosis peritoneal de ovario estadio IIIc de FIGO, 16 recibieron paclitaxel intraperitoneal intraoperatorio en hipertermia versus 16 en normotermia, tras cirugía citorreductora en ambos grupos. Variable principal: Concentración de paclitaxel en plasma, suero y tejido pre-y post tratamiento.

Resultados: En las 32 pacientes analizadas, el nivel de cirugía citorreductora obtenida, PCI, extensión de peritonectomía, tiempo operatorio, y demás datos demográficos fueron similares entre ambos grupos, excepto en el IMC, que fue superior en el grupo de normotermia frente a hipertermia (29±5 frente a 25±3, respectivamente; p=0,007). La morbilidad quirúrgica (Clavien >III), hematotoxicidad y nefrotoxicidad según CTCAE v4.0 fueron similares en ambos grupos.

Los niveles de placlitaxel en suero y plasma inmediatamente tras su administración fueron ligeramente superiores en el grupo de hipertermia frente a los niveles en suero y plasma 1 hora post-administración y en tejido, que fueron ligeramente superiores en el grupo de normotermia, aunque estas diferencias no fueron estadísticamente significativas (p=0,372, p=0,856, y p=0,525, p=0,119 y p=0,252, respectivamente).

Conclusión: A la vista de estos resultados preliminares, podemos avanzar que el paclitaxel ha demostrado tener una adecuada farmacocinética para alcanzar concentraciones en suero, plasma y tejido durante su administración intraperitoneal intraoperatoria tras cirugía citorreductora. Sin embargo, no se demostraron diferencias significativas entre su administración en condiciones de hipertermia frente a normotermia.


XVIII REUNIÓN DE LA SOCIEDAD ESPAÑOLA DE QUÍMICA ANALÍTICA
SEQ 2013
2013 SEQA
VI REUNIÓN DE LA SOCIEDAD ESPAÑOLA DE ESPECTROMETRÍA DE MASAS
Úbeda, 16-19 junio 2013
SEQA
S.E.E.M

DIM-18

TARGETED ANALYSIS OF PACLITAXEL AND THEIR MAJOR METABOLITES IN SERUM, PLASMA AND TISSUE FROM WOMAN WITH OVARIAN CANCER BY SPE–LC–MS/MS.

M.A. Fernández-Peralbo, F. Priego Capote, M.D. Luque de Castro

Department of Analytical Chemistry, Annex C-3 Building, Campus of Rabanales, University of Cordoba, E-14071 Córdoba, qa1lucam@uco.es.

Paclitaxel is a drug used to treat multiple types of tumor including lung, stomach, breast, ovary and endometrial cancer [1–3]. This drug is metabolized in the human liver by cytochrome P450 (CYP) enzymes. CYP2C8 metabolizes paclitaxel to 6a-hydroxypaclitaxel and CYP3A4 metabolizes paclitaxel to *p*-3´-hydroxypaclitaxel [4, 5]. The concentration of the active principle and its metabolites in blood and ovary tissue provides information on the pharmacokinetics of paclitaxel.

This study was aimed at determining the concentration of paclitaxel and its metabolites in different biological samples (serum, plasma and tissue). Samples from 13 patients were taken before and after washing the affected area of the ovaries with a solution of paclitaxel at different temperatures and one hour after chemotherapy in case of blood.

The analytical platform to develop the corresponding method consists of a fully automated system for solid-phase extraction (SPE) on-line connected to a liquid chromatograph (LC) and a triple quad mass detector (MS/MS).

The variables that affect each analytical step were optimized:
(i) SPE variables: type of cartridge, solvent, volume and flow for both sample loading and washing, and elution time.
(ii) LC variables: ionization agent.
(iii) MS/MS variables: fragmentor, collision energy, gas temperature, gas flow, nebulizer pressure, sheath gas temperature, sheath gas flow and capillary voltage.

The method thus optimized was applied to the three types of samples and the results compared with the calibration curves for the three target compounds. Differences on both absorption of the drug and its metabolism as a function of the temperature were significant, both in blood and in tissue. Not significant differentes were found between the use of plasma or serum. A wider study using a more representative number of patients is mandatory to corroborate the results found in the present research.

[1] A. Younes, J.P. Ayoub, A. Sarris, L. North, O. Pate, P. McLaughlin, M.A. Rodriguez, J. Romaguera, F. Hagemeister, C. Bachier, A. Preti, F. Cabanillas, Ann. Oncol. 8 (Suppl. 1) (1997) 129.
[2] W.B. Derry, L. Wilson, M.A. Jordan, Cancer Res. 58 (1998) 1177.
[3] M.A. Jordan, K. Wendell, S. Gardiner, W.B. Derry, H. Copp, L. Wilson, Cancer Res. 56 (1996) 816.
[4] R. Vaclavikova, P. Soucek, L. Svobodova, P. Anzenbacher, P. Simek, F.P. Guengerich, I. Gut, Drug Metab. Dispos. 32 (2004) 666.
[5] E.K. Rowinsky, M. Wright, B. Monsarrat, R.C. Donehower, Ann. Oncol. 5 (Suppl. 6) (1994) S7.


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:

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

1

Mass spectrometry as a fingerprinting metabolomic tool for the study of cardiovascular diseases through urine samples

M.A. Fernandez-Peralbo[1, 2], F. Priego-Capote[1, 2], M.D. Luque de Castro[1, 2]

[1]Department of Analytical Chemistry, Annex Marie Curie Building, Campus of Rabanales, University of Córdoba, E-14071, Córdoba, Spain

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

Cardiovascular diseases (CVDs) are a major cause of premature death worldwide [1]. Atherosclerosis is produced by a coronary injury that implies an irreversible deposition of lipids, particularly cholesterol, into blood decreasing flow transported by arteries to the myocardium [2].

Omics —particularly proteomics and genomics— have so far been employed to study CVDs and to search for ideal biomarkers [3–5]. Unfortunately, the majority of the identified biomarkers are not sensitive and/or specific enough to be used for population screening. Metabolomics can provide interesting tools for CVD studies and accurate information about the clinical situation of patients by using non-invasive analysis. Small biomolecules are the end result of all regulatory complexity present in cells, tissues, or organisms, including transcriptional and translational regulations, and post-translational modifications [6]. Metabolic changes are the closest reporters of alterations in the body in response to a disease process or drug therapy. Based on these aspects, metabolomics is a complementary approach to other omics in clinical diagnosis.

Different pathologies have been studied by metabolomics analysis of urine, but this biofluid has not been the preferred option to correlate metabolite profiles with cardiovascular diseases. For this, the potential of LC–QqTOF MS/MS) for metabolomic profiling of urine from patients with coronary artery diseases (CAD) was examined. Urine samples from 69 patients with different cardiac events and 23 controls from healthy persons

were analyzed. The studies were focused on discrimination among individuals who had suffered acute myocardial infarction (AMI) from individuals who had undergone one of the previous events to the situation of AMI, such as both stable and unstable angina.

Stepwise simplification of data dimensionality was conducted by fold change analysis to keep only those molecular features that experienced a change in relative concentration between pathological states of 2.0. After this filtering step, statistical analysis was carried out by PCA and PLS–DA. Figure 1 shows a discrimination between individuals with unstable angina and those who have suffered an AMI case, where PC1 and PC2 explain 40% of the observed variability. In contrast to PCA, the PLS–DA algorithm reported the graph shown in Figure 2, with a clear discrimination between individuals affected by an unstable angina and those with an episode of AMI. The accuracy of the method was estimated by cross-validation, where 84% was achieved in this case, as shows Table 1. The most significant molecular features were identified by ANOVA (p-value = 0.01) for each of the three studies (stable angina versus unstable angina, stable angina versus AMI, unstable angina versus AMI). As can be seen there were not potential metabolites in common among the three studies. In fact, most of the features were specific for each study.

The integration of metabolomics in the systems biology scenario should provide a more complete view of CVD. Therefore, future research is demanded to attain prognostic/diagnostic tools with appropriate selectivity and specificity considering the current relevance of CVD.

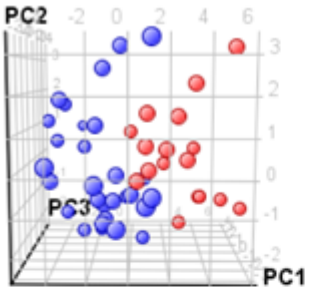


Figure 1: PCA scores plot provides an almost separation between unstable angina patients (🔴) and those who have suffered an AMI (🔵).

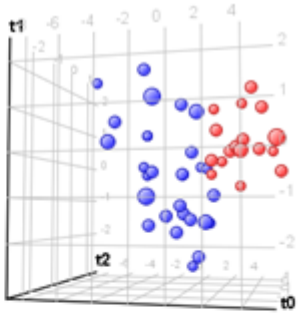


Figure 2: PLS-DA scores plot provides complete separation between patients with unstable angina (🔴) and AMI event (🔵).

Unstable angina (U.A) vs AMI	(U.A)	AMI	Accuracy (%)
Model Validation			
(U.A)	16	1	94.118
AMI	6	23	79.310
Recognition ability (%)			**84.783**
Model Training			
(U.A)	16	1	94.118
AMI	0	29	100
Predictive ability (%)			**97.826**

Table 1: Classification of the results using the PLS–DA model

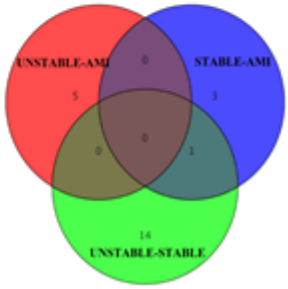


Figure 3: Venn diagram of three different studies.

References

1. Petersen S, Peto V, Rayner M, Leal J, Luengo-Fernández R, Gray A. European Cardiovascular Disease Statistics: 2005 edition. London: British Heart Foundation; 2005.
2. Libby P. Nature, 420, 868–874, (2002).
3. C.X. Kim, K.R. Bailey, G.G. Klee, A.A. Ellington, G. Liu, T.H. Mosley Jr, H. Rehman, I.J. Kullo, PLoS One 5(2) (2010) e9065. doi:10.1371/journal.pone.0009065.
4. C. López-Pedrera, N. Barbarroja, J.M. Villalba, Proteomics Clin. Appl. 3(2), 213–225, (2009).
5. H.J. Um, J. Min, H.J. Lee, Y.H. Kim, J. Biotechnol. 131(2), S55–S57, (2007).
6. G.D. Lewis, A. Asnani, R.E. Gerszten, J. Amer. College Cardiol. 52(2), 117–123, (2008).


XIII
GRASEQA 2012
Reunión del Grupo Regional Andaluz de la Sociedad Española de Química Analítica
MÁLAGA,
7 y 8 de junio de 2012
PROGRAMA CIENTÍFICO

Plataformas para análisis cualitativo/cuantitativo en metabolómica clínica y nutricional

F. Priego Capote, M.D. Luque de Castro, C. Ferreiro Vera, M. Calderón Santiago, M. Orozco Solano, A. Fernández Peralbo
Departamento de Química Analítica, Universidad de Córdoba, Campus de Rabanales, CP14071, Córdoba, Tel. 957218615, q72prcaf@uco.es

Resumen:

La metabolómica es actualmente una de las disciplinas ómicas más utilizadas en estudios clínicos y nutricionales debido a la interacción directa entre metaboloma y fenotipo. Así, el desarrollo de una determinada patología o el efecto nutricional de una dieta pueden alterar de manera significativa las rutas bioquímicas implicadas en el metabolismo de un individuo. El análisis metabolómico debe aportar una visión complementaria a la ofrecida por las otras disciplinas ómicas y, por tanto, debe contribuir al carácter integrador de la biología de sistemas. El papel del químico analítico resulta clave en la optimización de los métodos a desarrollar con el fin de conseguir el mayor nivel de información química posible para su conversión en información biológica. La complejidad del metaboloma compuesto por metabolitos con diferentes propiedades químicas y presentes en un amplio rango de concentración aumenta la dificultad del análisis metabolómico que conlleva:

- Estudio pormenorizado del tipo de muestra a estudiar y cuáles son las condiciones de almacenamiento para asegurar su estabilidad durante el estudio.
- Comparación entre alternativas de preparación de muestra destinadas al tratamiento de biofluidos humanos con el fin analizarlas en las mejores condiciones posibles.
- Integración de técnicas analíticas complementarias con el fin de maximizar el número de metabolitos detectados e identificados en cada muestra biológica.

- Validación de los métodos existentes desde un punto de vista analítico para asegurar la calidad de los resultados obtenidos.

Basados en este esquema, el Grupo PAIDI FQM-227 ha desarrollado plataformas metabolómicas para los siguientes estudios tanto en análisis orientado como global:

- La evaluación de la ingesta de aceites vegetales (con contenido variado en antioxidantes y sometidos a fritura) en el perfil de metabolitos de ácidos grasos poliinsaturados implicados en el mecanismo de la inflamación.
- El efecto en el perfil metabólico de individuos sanos de diferentes dietas de intervención en condiciones controladas.
- El desarrollo de métodos discriminantes de pacientes con lesión coronaria y diferente diagnóstico clínico a partir del análisis de suero y de orina.
- La aplicación de métodos de análisis cuantitativo de ácidos tricarboxílicos, aminoácidos, eicosanoides, fenoles y sus metabolitos, etc. a estudios con el fin de orientarlos a un panel representativo de metabolitos.

Los resultados obtenidos avalan la adecuación de la metabolómica para proporcionar una visión adicional en análisis clínico y nutricional, que hasta hace poco tiempo se sustentaba exclusivamente en genómico, transcriptómica y proteómica.

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-O

Solid–liquid and liquid–liquid extraction methods for oil enrichment

A. Fernández Peralbo and 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 enormous influence of food on health has given place to a new aspect of the present culture within which words such as "functional food", "supplemented food", "nutraceutical", "fat-free", "cholesterol-free", "low-sugar content", etc. are common place. This mouvement leads to an annual expenditure in US of more than $25 billion devoted to products labeled with the previous words; products that are consumed by more than 160 million of individuals.

Edible oils constitute one of the foods more frequently enriched with different compounds (particularly antioxidants) and with different objectives (*e.g.* to increase resistance to oxidation during storage and frying, to improve their organoleptic properties). Taking into account the proved healthy characteristics of extra-virgin olive-oil (EVOO), the most abundant research en this area is devoted to the enrichment of other oils with the compounds that endow EVOO with its unique features (in general phenol compounds, and especially secoiridoids), using as raw materials to obtain these compounds olive leaves or alperujo (the waste from oil production by the two-phase system).

An exhaustive study has shown that the two most effective ways to transfer olive phenols to any type of oil (the transfer process being always a function of the characteristics of the most abundant fatty acids in the oil) is by solid–liquid extraction (when the raw material is olive leaves), or by liquid–liquid extraction when the enrichment is based on an extract, preferably aqueous and very concentrated, obtained from either olive leaves or alperujo. In any case, application of ultrasound dramatically increases the efficiency of the process and shortens the time necessary for mass transfer. The enrichment degree is usually established by the total content of phenol compounds (use of the Folin–Ciocalteu method) or by the concentration of the antioxidant compounds considered more beneficial for health, such as oleuropein, hydroxytyrosol, luteolin, apigenin, etc. In this case, the determination step is carried out by liquid chromatography with detection by absorption spectrometry (diode array detector) or by mass spectrometry.

Enrichments up to 1 000 μg/g have been reported, which can be tailor-made, as it has been assessed that depending on the olive-tree cultivar, the collection time and the raw material acting and donor phase, the relative concentration of the different phenols can vary widely.

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 D CORDOBA

GRASEQA 2010

SC-43

Métodos de lixiviación y de extracción líquido–líquido para el enriquecimiento de aceites

M.A. Fernández Peralbo y M.D. Luque de Castro

Departamento de Química Analítica, Anexo al edificio Marie Curie, Campus de Rabanales Universidad de Córdoba, 14071 Córdoba, España

La constatación de la enorme influencia de la alimentación en la salud ha dado lugar a un nuevo aspecto de la cultura del que forman parte palabras como "alimento funcional", "alimento suplementado", "nutracéutico", o el de alimento bajo en grasa, en colesterol, en azúcares, etc. La extensión de este movimiento la pone de manifiesto el hecho de que en Estados Unidos se gastan más de 25 billones de dólares anuales en la compra de alimentos funcionales, nutracéuticos y suplementos dietéticos, que los consumen más de 160 millones de individuos.

Los aceites comestibles constituyen uno de los alimentos que con más frecuencia ha sido objeto de enriquecimiento con diferentes compuestos (especialmente antioxidantes) y con diferentes objetivos (*e.g.* aumentar su resistencia a la oxidación durante su almacenaje o en la fritura y prevención del desarrollo de enfermedades cardiovasculares y de ciertos tipos de cáncer). Dadas las características probadamente saludables del aceite de oliva virgen extra, la investigación más abundante en esta área se dedica a enriquecer otros aceites con los compuestos que le dan al de oliva virgen sus especiales características (en general compuestos fenólicos, pero de forma especial secoiridoides), utilizando como materia prima para su obtención hojas de olivo o alperujo.

El estudio desarrollado ha puesto de manifiesto que las dos formas más efectivas de producir la transferencia de los fenoles del olivo a cualquier tipo de aceite (el proceso siempre es función de las características de los ácidos grasos mayoritarios en el aceite) es mediante extracción sólido–líquido (cuando la materia prima está constituida por hojas de olivo), y la extracción líquido–líquido cuando se parte de un extracto, preferiblemente acuoso y muy concentrado, obtenido de las hojas de olivo o del alperujo. En ambos casos, la aplicación de ultrasonidos aumenta enormemente la eficiencia del proceso y acorta de forma drástica el tiempo necesario para la transferencia de masa. El grado de enriquecimiento se establece por el contenido total de compuestos fenólicos (método de Folin–Ciocalteu) o por la concentración de aquellos antioxidantes que se consideran más beneficiosos para la salud, tales como oleuropeína, hidroxitirosol, luteolina, apigenina, etc. En este último caso, la determinación se lleva a cabo por cromatografía de líquidos con detección mediante espectrometría de diodos en fila o con espectrómetro de masas.

Se han conseguido mediante estos procedimientos enriquecimientos de hasta 1.000 µg/g; con los que se pueden preparar aceites enriquecidos "a la carta", tanto por la concentración total de antioxidantes, como por el tipo de ellos, ya que se ha comprobado que, en función de la variedad de olivo, de la época de recolección y de la materia prima que actúa como fase donadora, la concentración relativa de los diferentes fenoles puede variar enormemente.

Abreviaturas

abbreviations

5mC, 5-methylcytosine

5hmC, 5-hydroxymethylcytosine

5fC, 5-formylcytosine

5caC, 5-carboxyl-cytosine

5mdC, 5-methyl-2-deoxycytidine

5hmdC, 5-hydroxymethyl-2-deoxycytidine

AA, arachidonic acid

ACE, automated cartridge exchange

ACN, acetonitrile

ACS, acute coronary syndrome

ALF, airway lining fluid

AMI, acute myocardial infarction

AMP, adenosine monophosphate

ANOVA, analysis of variance

APCI, atmospheric pressure chemical ionization

APPI, atmospheric pressure photoionization

AR, androgen receptor

AS, autosampler

ATP, adenosine triphosphate

AUC, area under the curve

BCAAs, branched chain amino acids

BMI, body mass index

CAD, coronary artery disease

CE, capillary electrophoresis

CI, confidence interval

CID, collision-induced dissociation

COPD, chronic obstructive pulmonary disease

COX, cyclooxygenase

CT, computed tomography

CV, coefficient of variation

CVD, cardiovascular disease

CYP, cytochrome P450

Da, dalton

DAPPI, desorption atmospheric pressure photoionization

DART, direct analysis in real time

DDAH-1, dimethylarginine dimethylaminohydrolase 1

DESI, desorption electrospray ionization

DIMS, direct infusion mass spectrometry

DRE, digital rectal examination

EAU, European association of urology

EBC, exhaled breath condensate

EDP, desorption electrospray ionization

EDTA, ethylenediaminetetraacetic acid

EETs, epoxyeicosatrienoic acids

EI, electron impact ionization

EIC, extracted ion chromatogram

ERB, ethical review board

ESI, electrospray ionization

FA, formic acid

FDA, food and drugs administration

FDR, false discovery rate

FC, fold change

FTICR, Fourier transform ion cyclotron resonance

FWHM, full width at half maximum

GC, gas chromatography

GSTP1, glutathione S-transferase Pi 1

H3K4, histone 3 at lysine 4

H3R2, histone H3 at arginine 2

H4R3, histone H4 at arginine 3

HATs, histone acetyl transferases

HAc, acetic acid

HACs, histone acetylases

HDACs, histone deacetylases

HDMs, histone demethylases

HETEs, hydroxyeicosatetraenoic acids

HI, hydrophilic interactions

HILIC, hydrophilic interaction liquid chromatography

HIPEC, hyperthermic intraperitoneal intraoperative chemotherapy

HLA, human leukocyte antigen

HLB, hydrophilic–lipophilic balance

HMDB, Human Metabolome Database

HMTs, histone methyltransferases

HMUFA, high fat content diet with monounsaturated fatty acids

HODE, hydroxyoctadecadienoic acid

HP-921, hexakis (1H, 1H, 3H-tetrafluoropropoxy) phosphazine

HPD, high pressure dispenser

IL-6, interleukin 6

IMS, ion-mobility spectrometry

IP, intraperitoneal chemotherapy

IS, internal standard

JARID1B, jumonji AT-rich interactive domain 1B

JMJD6, jumonji domain–containing 6 protein

KEGG, Kyoto Encyclopedia of Genes and Genomes

LC, liquid chromatography

LC–SSI–ITMS, liquid chromatography–ion trap mass spectrometry with a sonic spray ionization interface

LLE, liquid–liquid extraction

LOD, limit of detection

LOQ, limit of quantitation

LOXs, lipoxygenases

LysoPC, lysophosphatidylcholine

LSD, least significance difference

LSD1, lysine-specific demethylase 1

LT, leukotriene

MALDI, matrix-assisted laser desorption ionization

MCAD, medium-chain acyl-CoA dehydrogenase

METLIN, Metabolites and Tandem MS Database

MF, molecular feature

MHC, major histocompatibility complex

MIDAS, autosampler

MPP, mass profiler professional

MS, mass spectrometry

MSC, micro spin cartridges

MSTUS, mass spectrometry total useful signal

MTBE, metil-*ter*butil-éter

NMR, nuclear magnetic resonance

MPP, mass profiler professional

MRM, multiple reaction monitoring

NO, nitric-oxide

NSTE-ACS, non-ST-elevation acute coronary syndrome

NSTEMI, non-ST elevation myocardial infarction

OH-PAHs, monohydroxylated polycyclic aromatic hydrocarbons

OPLS-DA, orthogonal partial least squares discriminant analysis

OT, orbitrap

PAD, protein arginine deiminase

pAUC, partial area under the curve

PCa, prostate cancer

PCA, principal component analysis

PCA3, prostate cancer antigen 3

PGs, prostaglandins

PHI, prostate health index

PLS, partial least squares

PLS-DA, partial least squares discriminant analysis

PP, protein precipitation

PSA, prostate specific antigen

PTMs, post-translational modifications

PTR, proton-transfer reaction

QC, quality control

QIT, quadrupole ion trap

QqQ, triple quadrupole

QTOF, quadrupole–time of flight

rcf, relative centrifugal force

RF, random forest

ROC, receiver-operating characteristic

RSD, relative standard deviation

RT, retention time

SAM, standard addition method

SCs, synthetic cannabinoids

SD, standard deviation

SF, supercritical fluid

SPE, solid-phase extraction

SPME, solid-phase microextraction

SIL-IS, stable isotopic labeled internal standard

S/N, signal/noise

SOPs, standard operating procedures

SRM, selected reaction monitoring

TFA, trifluoroacetic acid

TICs, total ion chromatograms

TMPRSS2, androgen-regulated transmembrane-serine protease gene

TNF, tumor necrosis factor

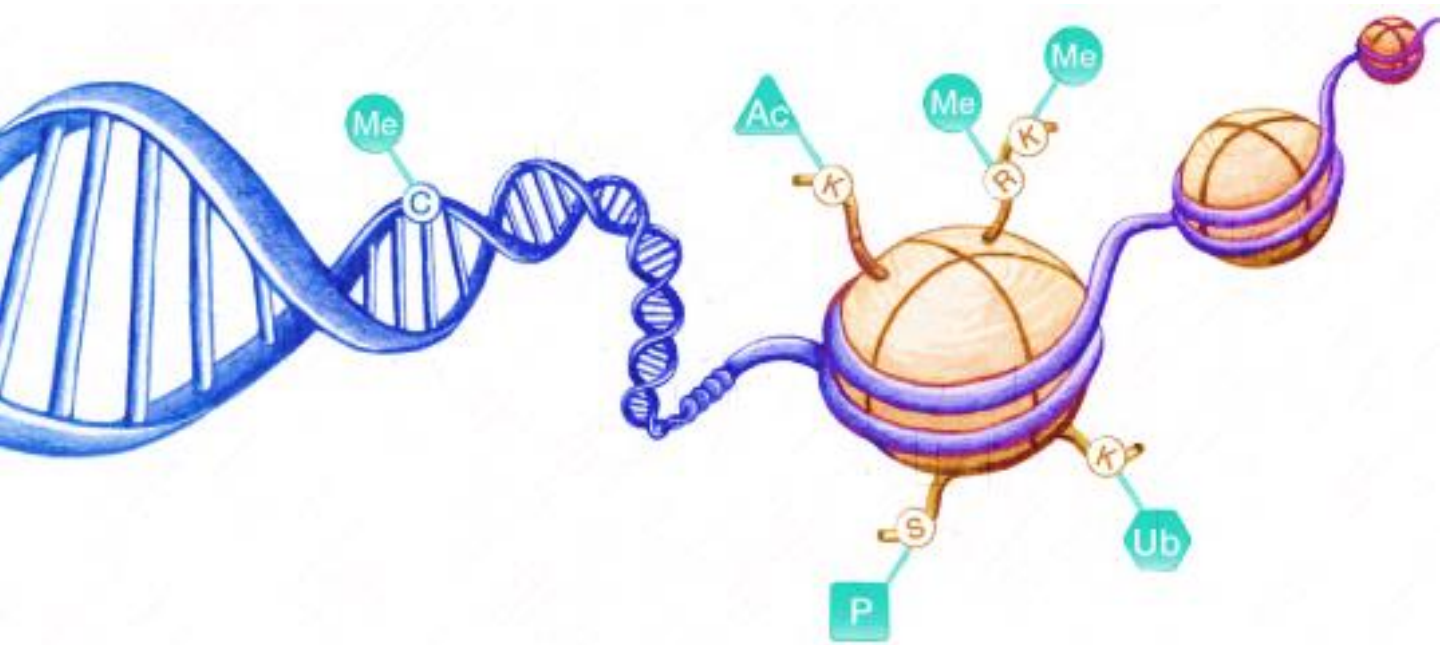
TrAC, trends in analytical chemistry

TX, thromboxane

UPLC, ultra-performance liquid chromatography

VEGF, vascular endothelial growth factor

VOC, volatile organic compound



IMIBIC
INSTITUTO MAIMÓNIDES DE INVESTIGACIÓN BIOMÉDICA DE CÓRDOBA

UNIVERSIDAD DE CÓRDOBA

ceiA3
campus de excelencia internacional agroalimentario