UNIVERSITAT DE BARCELONA

Efecto de un patrón de alimentación mediterránea sobre los perfiles metabolómicos asociados a salud metabólica y microbiota intestinal. Estudios de biomarcadores mediante una aproximación metabolómica no dirigida por resonancia magnética nuclear

Enrique Almanza Aguilera

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) i a través del Dipòsit Digital de la UB (**diposit.ub.edu**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) y a través del Repositorio Digital de la UB (**diposit.ub.edu**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service and by the UB Digital Repository (**diposit.ub.edu**) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

UNIVERSIDAD DE BARCELONA

FACULTAD DE FARMACIA Y CIENCIAS DE LA ALIMENTACIÓN

DEPARTAMENTO DE NUTRICIÓN, CIENCIAS DE LA ALIMENTACIÓN Y GASTRONOMÍA

EFECTO DE UN PATRÓN DE ALIMENTACIÓN MEDITERRÁNEA SOBRE LOS PERFILES METABOLÓMICOS ASOCIADOS A SALUD METABÓLICA Y MICROBIOTA INTESTINAL. ESTUDIOS DE BIOMARCADORES MEDIANTE UNA APROXIMACIÓN METABOLÓMICA NO DIRIGIDA POR RESONANCIA MAGNÉTICA NUCLEAR

ENRIQUE ALMANZA AGUILERA

2017

Universidad de Barcelona

Facultad de Farmacia y Ciencias de la Alimentación

Departamento de Nutrición, Ciencias de la Alimentación y Gastronomía

Programa de Doctorado

Alimentación y Nutrición

2013-2017

EFECTO DE UN PATRÓN DE ALIMENTACIÓN MEDITERRÁNEA SOBRE LOS PERFILES METABOLÓMICOS ASOCIADOS A SALUD METABÓLICA Y MICROBIOTA INTESTINAL. ESTUDIOS DE BIOMARCADORES MEDIANTE UNA APROXIMACIÓN METABOLÓMICA NO DIRIGIDA POR RESONANCIA MAGNÉTICA NUCLEAR

Memoria presentada por Enrique Almanza Aguilera para optar al título de Doctor por la Universidad de Barcelona, dirigida por:

Dr. Cristina Andrés Lacueva

(Director/Tutor)

Dr. Mireia Urpí Sardà

(Director)

Enrique Almanza Aguilera

2017

La presente tesis doctoral ha sido financiada por:

Consejo Nacional de Ciencia y Tecnología

Beca de formación doctoral

Registro/Solicitud 231248/359380

JPI HDHL FOODBALL

Project PCIN-2014-133

Ministerio de Economía y Competitividad y Fondo Europeo Europeo de Desarrollo Regional

Project AGL 2009-13906-C02-01

Agencia de Gestión de Ayudas Universitarias y de Investigación (AGAUR)-Generalitat de Catalunya

Ayuda de apoyo para potenciar los grupos de investigación consolidados

2014-SGR-1566

Associació Catalana de Diabetes

Project ACD2015

LISTA DE ABREVIATURAS Y ACRÓNIMOS

A-ADM	Alta adherencia a la dieta Mediterránea
ANOVA	Análisis de Varianza (del inglés, *Analisis of Variance*)
AUC	Área bajo la curva (del inglés, *Area Under the Curve*)
B-ADM	Baja adherencia a la dieta Mediterránea
DM2	Diabetes Mellitus Tipo 2
ECV	Enfermedad cardiovascular
FDR	Tasa de descubrimiento falsos (del inglés, *False Discovery Rate*)
FFQ	Cuestionario de Frecuencia de Consumo de Alimentos (del inglés, *Food Frequency Questionnaire*)
HDL	Colesterol unido a lipoproteínas de alta densidad (del inglés, *High Density Lipoprotein cholesterol*)
HMDB	Base de datos del Metaboloma Humano (del inglés, *Human Metabolome Database*)
IMC	Índice de Masa Corporal
KEGG	Enciclopedia "Kyoto" de Genes y Genomas (del inglés, *Kyoto Encyclpedia of Genes and Genomes*)
LC-MS	Cromatografía líquida acoplada a espectrometría de masas (del inglés, *Liquid Chromatography coupled to Mass Spectrometry*)
LDL	Colesterol unido a lipoproteínas de baja densidad (del inglés, *Low Density Lipoprotein cholesterol*)
DietMed	Dieta Mediterránea
NOESY	Espectroscopia de efecto nuclear *Overhauser* (del nglés, *Nuclear Overhauser Effect Spectroscopy*)
OSC	Corrección de señales ortogonales (del inglés, *Orthogonal Signal Correction*)
OSC-PLS-DA	Analisis discriminante por mínimos cuadrados parciales con corrección de señales orthogonales (del inglés, *Partial Least Square Discriminant Analysis with Orthogonal Signal Correction*)
PCA	Análisis de componentes principales (del inglés, *Principal Component Analysis*)
PLS	Regresión de mínimos cuadrados parciales (del inglés, *Partial Least Square regression*)
PLS-DA	Análisis discriminante por mínimos cuadrados parciales (del inglés, *Partial Least Square Discriminant Analysis*)
PREDIMED	Prevención con Dieta Mediterránea

rdCV	Doble validación cruzada repetida (del inglés, *repeated double Cross Validation*)
rdCV-PLS	Regresión de mínimos cuadrados parciales con doble validación cruzada repetida (del inglés, *Partial Least Square regression with repeated double Cross Validation*)
rdCV-RF	Bosques Aletaorios con doble validación cruzada repetida (del inglés, *Random Forest with repeated double Cross Validation*)
RF	Bosques Aletaorios (del inglés, *Random Forest*)
RMN	Resonancia Magnética Nuclear
ROC	Característica Operativa del Receptor (del inglés, *Receiver Operating Characteristic*)
STOCSY	Correlación estadística total de espectroscopía (del inglés, *Statistical Total Correlation Spectroscopy*)
VIP	Importancia de las variables en la proyección (del inglés, *Variable Importance Projection*)

ÍNDICE

RESUMEN/*ABSTRACT* 1

INTERÉS/THESIS RATIONALE 7

II. INTRODUCCIÓN 17

2.1 METABOLÓMICA 19

2.1.1 Introducción 19

2.1.2 Flujo de trabajo de la metabolómica 21

2.2 RIESGO CARDIOVASCULAR 24

2.2.1 Generalidades 24

2.2.1 Factores de riesgo 24

2.2.2 Obesidad 25

2.2.3 Obesidad Metabólicamente Sana 25

2.3.4 Diabetes Mellitus 28

2.3 DIETA MEDITERRANEA 33

2.3.1 Definición 33

2.3.2 Dieta Mediterránea y enfermedad cardiovascular 34

2.3.3 Dieta Mediterránea, obesidad, diabetes y síndrome metabólico 37

2.3.4 Componentes bioactivos 38

2.3.5 Dieta Mediterránea, microbiota y salud metabólica 40

2.3.6 Evaluación de la adherencia a la dieta Mediterránea 41

2.4 APLICACIONES DE LA METABOLÓMICA EN EL ESTUDIO DE BIOMARCADORES DE RIESGO CARDIOVASCULAR Y ESTILOS DE VIDA CARDIOSALUDABLE 47

2.4.1 Biomarcadores de Diabetes Mellitus 2 47

2.4.2 Biomarcadores de patrones alimentarios e intervenciones dietéticas 58

2.4.3 Biomarcadores de dieta Mediterránea 60

III. OBJETIVOS 63

IV. METODOLOGIA 67

4.1 ESTUDIOS CLÍNICOS 69

4.1.1.1 Características de los participantes 70

4.1.1.2 Mediciones, cuestionarios y recogida de muestras biológicas 71

4.1.1.3 Cuestionario de adherencia a la dieta Mediterránea 71

4.1.1.4 Estratificación de la población para los estudios de biomarcadores de diabetes y de adherencia a la dieta Mediterránea 72

4.1.2 Diseño del estudio INTERVENCIÓN EN EL ESTILO DE VIDA EN MUJERES OBESAS METABÓLICAMENTE SANAS 74

4.1.2.1 Características de los participantes 74

4.1.2.3 Mediciones, cuestionarios y recogida de muestras biológicas 75

4.2 ANÁLISIS METABOLÓMICO NO DIRIGIDO 76

4.2.1 Recolección y preparación de las muestras biológicas 76

4.2.2 Adquisición de los datos y procesamiento de los espectros 77

4.2.3 Análisis de los datos 77

4.2.4 Identificación de metabolitos 94

4.2.5 Interpretación biológica 95

V. RESULTADOS 97

5.1 BIOMARCADORES METABOLOMICOS Y METABOTIPOS DE DIABETES MELLITUS TIPO 2: ANÁLISIS TRANSVERSAL EN EL ESTUDIO PREDIMED 99

5.2 BIOMARCADORES METABOLÓMICOS DE ADHERENCIA A LA DIETA MEDITERRÁNEA 117

5.3 BIOMARCADORES DEL EFECTO DE UNA INTERVENCIÓN EN EL ESTILO DE VIDA PARA LA PÉRDA DE PESO EN MUJERES OBESAS METABÓLICAMENTE SANAS 127

VI. DISCUSIÓN 139

VII. CONCLUSIONES/*CONCLUSIONS* 151

VIII. REFERENCIAS BIBLIOGRÁFICAS 161

IX. ANEXOS 183

9.1 PUBLICACIONES DERIVADAS DE LA TESIS DOCTORAL 185

9.2 COMUNICACIONES EN CONGRESOS 267

RESUMEN

Las enfermedades cardiovasculares (ECV) ocupan el primer lugar a nivel mundial en causas de mortalidad. Los cambios pertinentes en el estilo de vida como son la dieta y la actividad física proporcionan beneficios a nivel metabólico frente a factores de riesgo cardiovascular intermedio como la obesidad y la diabetes. La aplicación de la metabolómica en estudios observacionales y de intervención en un estilo de vida saludable permite profundizar en el conocimiento de los mecanismos moleculares subyacentes, así como en la identificación de biomarcadores que describan estos mecanismos asociados de una forma más global.

El objetivo principal de la presente tesis doctoral fue identificar biomarcadores relacionados con factores de riesgo cardiovascular y el estudio de los perfiles metabolómicos asociados a un patrón de alimentación mediterránea mediante el empleo una aproximación metabolómica no dirigida basada en resonancia magnética nuclear (RMN). Para la consecución de este objetivo se han analizado muestras biológicas y datos metabolómicos de dos estudios clínicos. El primero de estos estudios corresponde al ensayo clínico aleatorizado y multicéntrico de intervención "Prevención con Dieta Mediterránea" (PREDIMED), del cual se han incluido muestras de orina al inicio del estudio. El segundo es un estudio clínico de intervención aleatorizado, paralelo y simple ciego, basado en dieta y actividad física en un grupo de mujeres obesas metabólicamente sanas, con el objetivo de evaluar la pérdida de peso. De este último estudio se han obtenido muestras de plasma en el momento inicial y después de 3 y 12 meses de intervención. La plataforma y enfoque analítico en común para el análisis de las muestras biológicas en ambos estudios fue una aproximación metabolómica por RMN.

Para el estudio de los biomarcadores relacionados con un factor de riesgo cardiovascular, se caracterizó la huella metabólica asociada a la diabetes mellitus tipo 2 (DM2) mediante la comparación de la composición del metaboloma urinario de individuos diabéticos y no diabéticos dentro de la población del estudio PREDIMED. Con este análisis se obtuvo un total de 33 metabolitos discriminantes entre los dos grupos de individuos. Posteriormente, a partir de esta caracterización se determinaron los metabolitos con mayor capacidad de discriminación de los grupos a través del cálculo de un marcador multimetabolito que incluyó altos niveles de ácido metilsuccínico, alanina, dimetilglicina, ácido guanidinoacético, y niveles bajos de glutamina, metilguanidina, ácido 3-hidroximandélico y ácido hipúrico. Este marcador mostró una alta capacidad discriminante de la diabetes (área bajo la curva [AUC] 96,4%), incluso con una mayor capacidad para discriminar que la glucosa (AUC 89,8%). Adicionalmente, a partir de la caracterización metabólica inicial se logró clasificar a la población diabética en dos fenotipos metabólicos (metabotipos). En concreto, los individuos diabéticos con niveles más altos de

fenilalanina, fenilacetilglutamina, *p*-cresol y ácido acetoacético se caracterizaron por presentar también los niveles más altos de glucosa plasmática.

Referente a los perfiles metabolómicos asociados a un patrón de alimentación mediterránea, se han realizado dos aproximaciones. La primera aproximación ha sido a partir del estudio PREDIMED, estratificando a la población disponible según su nivel de adherencia a la dieta Mediterránea (DietMed), y la segunda aproximación a través de una intervención en el estilo de vida para la pérdida de peso.

En el estudio PREDIMED, se ha determinado un biomarcador de alta adherencia a la DietMed. Se estratificó a la población en alta (A-ADM) y baja (B-ADM) adherencia a la DietMed en función de la puntuación individual obtenida a través de un cuestionario de adherencia compuesto por 14 ítems (escala total: 0-14 puntos). Individuos con puntuaciones de $\leq$7 se incluyeron al grupo B-ADM e individuos con puntuaciones $\geq$10 fueron asignados al grupo A-ADM. Posteriormente, se determinaron las diferencias en la composición del metaboloma urinario entre estos dos grupos, obteniendo así un total de 34 metabolitos discriminantes. Entre los metabolitos asociados a la A-ADM algunos se correlacionaron con la ingesta de alimentos característicos de la DietMed. Finalmente, el biomarcador de A-ADM estuvo compuesto por tres metabolitos derivados del metabolismo de la microbiota: fenilacetilglutamina, *p*-cresol y ácido 4-hidroxifenilacético. Este biomarcador demostró tener una alta capacidad discriminante entre B-ADM y A-ADM (AUC, 97,0%; especificidad, 94,7%; y sensibilidad, 95,2%). Por otra parte, el biomarcador de A-ADM se correlacionó positiva y significativamente con la ingesta de frutas, verduras, legumbres y pescado.

La identificación de biomarcadores de efecto y asociados a una intervención en el estilo de vida saludable se logró mediante el análisis comparativo de los cambios en la composición del metaboloma plasmático de dos grupos de mujeres obesas y definidas como metabólicamente sanas, asignadas a un grupo control (recomendaciones generales de dieta saludable y actividad física) o tratamiento (DietMed hipocalórica y actividad física regular) para la pérdida de peso. Después de tres meses de estudio y comparadas con el grupo control, las participantes en el grupo tratamiento tuvieron una mayor pérdida de peso (kg) (3,04 ± 3,62 *vs* 9,10 ± 3,50 $p<0{,}001$). Los cambios en la composición del metaboloma en el grupo intervención incluyeron niveles más altos de fosfocreatina, ácido 3-hidroxibutírico, metilguanidina, mio-inositol y ácido fórmico, así como niveles más bajos de trimetilamina, prolina y fragmentos de colesterol LDL/VLDL. La pérdida de peso en los dos grupos se asoció positivamente con niveles más altos también de fosfocreatina, ácido fórmico, asparagina y ácido hipúrico; así como con niveles más bajos de creatina, trimetilamina, fragmentos de colesterol LDL/VLDL y ácido 2-hidroxibutírico.

Los resultados derivados de los presentes estudios que componen esta Tesis Doctoral refuerzan la importancia de la metabolómica para la identificación y desarrollo de biomarcadores metabólicos que permitan profundizar en los mecanismos moleculares subyacentes a enfermedades crónicas y estilos de vida saludable, así como en la relación de estos biomarcadores con la prevención, prevalencia y mejora de estas enfermedades.

ABSTRACT

Cardiovascular diseases (CVD) rank first in the world in causes of mortality. The timely changes in lifestyle such as diet and physical activity provide benefits at metabolic level against intermediate cardiovascular risk factors including obesity and diabetes. The application of metabolomics in observational and interventional studies through a healthy lifestyle allows a deeper understanding of the underlying molecular mechanisms, as well as the identification of biomarkers that describe these mechanisms in a more global way.

The main objective of this Doctoral Thesis was to identify biomarkers related to cardiovascular risk factors and to study the metabolic profiles associated with a Mediterranean dietary pattern by using an untargeted metabolomic approach based on nuclear magnetic resonance (NMR). To achieve this objective, biological samples and metabolomic data from two clinical studies have been used. The first of these studies corresponds to the randomized and multicenter clinical trial of "Prevención con Dieta Mediterránea" (PREDIMED), from which urine samples were used at the baseline of the study. The study belongs to a randomized, parallel and single blind intervention in lifestyle (consistent in diet and physical activity) for weight loss in metabolically healthy obese women (MHO). From this last study, plasma samples were collected at baseline and after 3 and 12 months of intervention. The common analytical platform and approach for analyzing the biological samples from both studies was a metabolomic approach by NMR.

For the study of biomarkers related to a cardiovascular risk factor, the metabolic footprint associated with type 2 diabetes mellitus (DM2) was characterized by comparing the composition of the urinary metabolome of diabetic and non-diabetic individuals within the population of PREDIMED study. From this analysis, a total of 33 discriminant metabolites were obtained between the two groups (DM2 and non-DM2). Subsequently, from this characterization, the metabolites with the greatest discrimination capacity of the groups were determined by calculating a multimetabolite biomarker, which included high urine levels of methylsuccinic acid, alanine, dimethylglycine, guanidinoacetic acid, and low levels of glutamine, methylguanidine, 3-hydroxymandelic acid and hippuric acid. This marker showed a high discriminant capacity for diabetes (area under the curve [AUC] 96.4%), even with a greater ability to discriminate than glucose (AUC 89.8%). Additionally, from the initial metabolic characterization, the diabetic population was classified into two metabolic phenotypes (metabotypes). Specifically, diabetic subjects with higher levels of phenylalanine, phenylacetylglutamine, *p*-cresol and acetoacetic acid were also characterized by higher plasma glucose levels.

With respect to the metabolic profiles associated with a Mediterranean dietary pattern, two approaches were conducted. The first approach was made from the PREDIMED study,

stratifying the available population according to their level of adherence to the Mediterranean diet (MedDiet). Meanwhile, the second approach was based on the intervention study of lifestyle for weight loss in MHO women.

In the PREDIMED study, a biomarker of high adherence to MedDiet has been determined. The population was stratified in high (H-MDA) and low (L-MDA) adherence to MedDiet according to the individual score obtained through an adherence questionnaire composed of 14 items (total scale: 0-14 points). Individuals with scores of ≤ 7 were included in the L-MDA group and individuals with scores ≥ 10 were assigned to the H-MDA group. Subsequently, the differences in the composition of the urinary metabolome between these two groups were determined, obtaining a total of 34 discriminant metabolites. Among the metabolites associated with H-MDA some were correlated with intake of characteristic foods of MedDiet. Finally, the biomarker of H-MDA was composed of three metabolites derived from the metabolism of the microbiota: phenylacetylglutamine, *p*-cresol and 4-hydroxyphenylacetic acid. This biomarker demonstrated a high discriminant capacity between L-MDA and H-MDA (AUC, 97.0%, specificity, 94.7%, and sensitivity, 95.2%). On the other hand, the biomarker of H-MDA correlated positively and significantly with the intake of fruits, vegetables, legumes and fish.

The identification of biomarkers of effect and associated with a healthy lifestyle intervention was achieved through the comparative analysis of changes in the composition of the plasma metabolome of two groups of MHO women assigned to either a control group (General dietary and physical activity recommendations) or treatment group (MedDiet hypocaloric and regular physical activity). After three months of study and compared with the control group, participants in the treatment group had greater weight loss (kg) (3.04 ± 3.62 vs 9.10 ± 3.50 $p < 0.001$). Changes in metabolome composition in the intervention group included higher levels of phosphocreatine, 3-hydroxybutyric acid, methylguanidine, myo-inositol and formic acid, as well as lower levels of trimethylamine, proline and LDL/VLDL cholesterol signals. Weight loss in the two groups was positively associated with higher levels of phosphocreatine, formic acid, asparagine and hippuric acid; as well as with lower levels of creatine, trimethylamine, LDL/VLDL cholesterol signals and 2-hydroxybutyric acid.

The findings derived from the present studies that compose this Doctoral Thesis reinforce the importance of the metabolomics for the identification and development of metabolic biomarkers that allow to deep in the molecular mechanisms underlying chronic diseases and healthy lifestyles, as well as in the relation of these biomarkers with the prevention, prevalence and improvement of these diseases.

I. INTERÉS

La obesidad y la diabetes son dos de los principales factores de riesgo cardiovascular con altas tasas de morbilidad y mortalidad a nivel mundial (Organización Mundial de la Salud (OMS) 2017). En su prevención y tratamiento, se reconoce que la adopción de estilos de vida que incluyan patrones alimentarios saludables o con baja ingesta calórica, así como una regular actividad física, contribuye significativamente en una disminución del riesgo cardiovascular (Calder et al. 2011; Verheggen et al. 2016). En este sentido, se ha demostrado que el seguimiento de un patrón alimentario de tipo mediterráneo es un componente del estilo de vida altamente recomendable debido a su relación inversa con la incidencia y mortalidad por ECV (Estruch et al. 2013; Trichopoulou et al. 2003; Martinez-Gonzalez & Martin-Calvo 2016; Grosso et al. 2017). Algunos de los principales mecanismos que explican esta relación incluyen una reducción en la incidencia y prevalencia del síndrome metabólico (MetS) y sus componentes (Perez-Martinez et al. 2017), la mejora de los índices antropométricos (Schroder et al. 2004; Shai et al. 2008; Esposito et al. 2003) y una disminución del riesgo de desarrollar diabetes mellitus (Jannasch et al. 2017), entre otros. Sumado a estos mecanismos, en los últimos años se demostrado que la pérdida de peso a través de intervenciones en el estilo de vida (basadas por ejemplo en patrones alimentarios con restricción calórica y/o actividad física) impacta positivamente en distintos marcadores clínicos, lo que se traduce en un menor riesgo cardiovascular (Abbenhardt et al. 2013; Petelin et al. 2014). Ante tales evidencias, sin embargo, existe todavía un desconocimiento en torno a los procesos bioquímicos que subyacen a las alteraciones producidas por los factores de riesgo cardiovascular y su relación con el seguimiento o adopción de estilos de vida cardiosaludable. Desentrañar estos procesos a través de la identificación de moléculas y rutas metabólicas asociadas, resulta de especial importancia en áreas como la epidemiología nutricional, no sólo para una mejor comprensión del funcionamiento y evolución de las patologías, sino también para el establecimiento de estrategias terapéuticas o de intervención en el estilo de vida necesarias para su control. En los últimos años con el desarrollo de las llamadas ciencias *ómicas,* y entre estas particularmente la metabolómica, se han logrado numerosos avances en la identificación de moléculas características de procesos biológicos normales o patológicos.

La metabolómica se define como una disciplina que estudia globalmente las moléculas de bajo peso molecular (<1500 Da), conocidas también como metabolitos, dentro de un sistema biológico (p.e., célula, tejido u organismo) en un momento determinado (Johnson et al. 2016; Monteiro et al. 2013; Gibbons et al. 2015). Al conjunto de metabolitos que constituyen el sistema biológico se les conoce como metaboloma y su estudio a través de la metabolómica permite conocer de manera directa las relaciones bioquímicas y los procesos biológicos que en el sistema se llevan a cabo. El análisis del metaboloma requiere de la aplicación de un flujo de

trabajo multidisciplinario en el que se incluye la obtención de los datos a través de técnicas analíticas avanzadas, como la espectroscopia por resonancia magnética nuclear (RMN), y su análisis mediante técnicas estadísticas multivariantes. El uso más frecuente de la RMN en la metabolómica es en estudios con aproximaciones no dirigidas o *untargeted*, que consisten en el análisis de los cambios en metaboloma por efecto de condiciones específicas, pero sin previo conocimiento de los metabolitos que se han de determinar. Estas aproximaciones *untargeted* además tienen la potencialidad de generar un marcado aumento en la identificación de potenciales biomarcadores de procesos biológicos en condiciones normales o patológicas (Bujak et al. 2015).

De este modo, una de las aplicaciones más comunes de la metabolómica es la identificación de biomarcadores de enfermedades, como la DM2. Concretamente, en la investigación sobre la DM2 la metabolómica se ha aplicado exitosamente en el descubrimiento de biomarcadores de diagnóstico, pronóstico, alteraciones metabólicas y de mecanismos de acción de medicamentos (Urpi-Sarda et al. 2015). La mayoría de estos biomarcadores, sin embargo, suelen estar compuestos por un solo metabolito, lo que dificulta a veces su replicación en diferentes estudios. Algunos modelos, en cambio, emplean la combinación de metabolitos y la información de otras variables como la edad, género y estilos de vida, para la construcción de biomarcadores, principalmente de predicción de DM2 (Herder et al. 2014). Es importante tomar en cuenta que la mayoría de estos modelos se han construido y probado a partir de muestras de plasma sanguíneo, lo que en cierto modo dificulta su aplicación en estudios observacionales o clínicos de gran escala. En este sentido, el desarrollo de biomarcadores a partir de muestras biológicas menos invasivas, como la orina, constituye un campo de oportunidad en la investigación de las alteraciones metabólicas producidas por la DM2. Por otra parte, la aplicación de la metabolómica en el estudio de la DM2 resulta esencial hoy en día para entender mejor la intervariabilidad metabólica que existe dentro de esta enfermedad. Por lo tanto, el estudio de fenotipos metabólicos (metabotipos) relacionados con la DM2 a través de la metabolómica representa un punto de partida en la nueva “medicina estratificada” (Pearson 2016) para la identificación de individuos con mayor riesgo o el establecimiento de programas de intervención en el estilo de vida focalizados en cada uno de estos metabotipos (Heinzmann et al. 2012).

Por otra parte, y como ya se ha mencionado anteriormente, los estilos de vida que incluyan el seguimiento o adopción de patrones alimentarios saludables como la DietMed constituyen un componente esencial en la prevención y disminución del riesgo de ECV. La valoración del cumplimiento o adherencia a este patrón, sin embargo, supone un reto en estudios observacionales o de intervención, en los que hasta la fecha se utilizan principalmente

cuestionarios autoinformados de frecuencia de consumo de alimentos para tal valoración (Hu 2002). Algunos de estos cuestionarios se presentan en forma de índices o *scores* de adherencia, y son construidos en base a recomendaciones dietéticas como resultado de un consenso científico o como propuesta de investigadores que emplean un enfoque basado en la evidencia científica (Scientific Report of the 2015 Dietary Guidelines Advisory Committee (DGAC) 2015). Hasta la fecha, la adherencia a la DietMed se ha evaluado mediante el uso de distintos *scores* descritos en la bibliografía, entre los que destacan los cuestionarios de 9 ítems, propuesto por Tricopoluus (Trichopoulou et al. 2003), y el de 14 ítems, propuesto por el estudio PREDIMED (Schroder et al. 2011). Estos *scores* se han empleado además en distintos estudios para estudiar la asociación de la adherencia a la MeDiet con la composición de la microbiota intestinal (Gutierrez-Diaz et al. 2017), factores de riesgo cardiovascular (Martinez-González et al. 2012), enfermedades relacionadas con el envejecimiento (Talegawkar et al. 2012) y la mortalidad total, entre otros (Trichopoulou et al. 2003). Sin embargo, debido a la complejidad de la DietMed, en términos de la cantidad y frecuencia con la que se consumen los alimentos que lo componen, la evaluación de su adherencia a través de los métodos tradicionales puede resultar inexacta, por lo que se condisera que el empleo de biomarcadores de exposición a los alimentos (*food metabolome*) podría ser de gran complementariedad (Andersen et al. 2014). En este sentido durante los últimos años la metabolómica ha jugado un papel importante en el descubrimiento de nuevos biomarcadores dietéticos, ya sean de consumo de alimentos individuales (Garcia-Aloy et al. 2014; Heinzmann et al. 2010) o de patrones alimentarios (Gibbons et al. 2017; Playdon et al. 2017). Ejemplos de patrones alimentarios descritos a través de biomarcadores metabolómicos incluyen la nueva dieta nórdica (Lankinen et al. 2016; Andersen et al. 2014) y la dieta occidental (Bouchard-Mercier et al. 2013). Sin embargo, biomarcadores de adherencia a la DietMed han sido poco estudiados (Gutierrez-Diaz et al. 2017). Se reconoce por lo tanto que el uso de biomarcadores metabolómicos podría no solo complementar la información de los *scores* tradicionales de adherencia a la DietMed, sino mejorar la precisión en su evaluación y establecer relaciones de su efecto en el organismo (Bondia-Pons et al. 2015). De esta manera, la caracterización de la adherencia a un patrón alimentario mediterráneo mediante un enfoque metabolómico resulta de especial interés en la investigación nutricional ya que permitiría comprender mejor la relación entre la dieta, teniendo en cuenta la biodisponibilidad de los compuestos bioactivos de la DietMed, y el riesgo de enfermedades crónicas asociadas a la ECV.

Por otra parte, en estudios de intervención en el estilo de vida dirigidos a disminuir el riesgo de ECV o de factores de riesgo intermedios, la información de la evaluación del seguimiento de un patrón alimentario saludable sería insuficiente si no se consideran también los efectos

beneficiosos de tales intervenciones. Con respecto a esto último, existen numerosas evidencias en torno a los múltiples beneficios que proporcionan intervenciones en el estilo de vida, basadas por ejemplo en dietas hipocalóricas y actividad física regular, dirigidas a la pérdida de peso en poblaciones obesas (Petelin et al. 2014; Abbenhardt et al. 2013). Sin embargo, la mayoría de los estudios de intervención conducidos hasta ahora con esta aproximación han focalizado su interés en la obesidad en general, descartando, por ejemplo, la variabilidad metabólica que existe dentro de este fenotipo. En este sentido, estudios recientes han demostrado la existencia de un fenotipo discordante dentro de los individuos obesos caracterizado por presentar un menor riesgo de ECV o mortalidad total (Calori et al. 2011; Hamer & Stamatakis 2012). A los individuos con este fenotipo se les conoce como obesos metabólicamente sanos (del inglés, *metabolically healthy obese*, MHO) y recientemente han comenzado a ganar interés en la epidemiología nutricional. No obstante, aunque la dieta y la actividad física son factores de riesgo cardiovascular modificables bien conocidos, su potencial impacto beneficioso en condiciones controladas sobre el fenotipo la MHO recién está ganando interés (Dalzill et al. 2014; Arsenault et al. 2009). La metabolómica aplicada en estudios de intervención para la pérdida de peso en el fenotipo MHO ofrece por tanto la oportunidad de profundizar en el conocimiento de los mecanismos moleculares a través de los cuales la intervención y la pérdida de peso modulan, por separado o conjuntamente, los beneficios metabólicos que se traducen en una mejora del riesgo cardiovascular. Por lo tanto, el estudio de biomarcadores metabolómicos relacionados con intervenciones en el estilo de vida que combinan patrones alimentarios saludables y actividad física, resulta de especial interés en el fenotipo MHO, no por estos biomarcadores contribuyen a confirmar los efectos beneficiosos de dichas intervenciones sino también porque permiten entender de qué manera este fenotipo se acerca hacia un estado más saludable o de menor riesgo cardiovascular.

Por todo lo anteriormente expuesto, se presenta una Tesis Doctoral que integra la aplicación de la metabolómica mediante una aproximación no dirigida por RMN en el estudio de biomarcadores y patrones metabólicos asociados a un factor de riesgo cardiovascular, como es la DM2, y de un patrón alimentario saludable, compuesto por una DietMed y su efecto tras una intervención para la pérdida de peso en el fenotipo MHO.

THESIS RATIONALE

Obesity and diabetes are two of the major cardiovascular risk factors with high worldwide morbidity and mortality rates (WHO 2017). In their prevention and treatment, it is recognized that the adoption of lifestyles that include healthy dietary patterns or dietary patterns with low caloric intake, as well as practice of regular physical activity, contributes significantly to decrease cardiovascular risk (Calder et al., 2011; Verheggen et al. Al. 2016). In this sense, it has been demonstrated that the adherence to a Mediterranean-type food pattern is a highly recommended lifestyle component due to its inverse relationship with CVD incidence and mortality (Estruch et al., 2013; Trichopoulou et al., 2003; Martinez-Gonzalez & Martin-Calvo 2016; Grosso et al, 2017). Some of the main mechanisms explaining this relationship include a reduction in the incidence and prevalence of Metabolic Syndrome (MetS) and its components (Perez-Martinez et al., 2017), improvement of anthropometric indices (Schroder et al., 2004; and a reduced risk of developing diabetes mellitus (Jannasch et al., 2017), among others. In addition to these mechanisms, in recent years it has been demonstrated that weight loss through lifestyle interventions (based, for example, on calorie restricted dietary patterns and /or physical activity) has a positive impact on different clinical markers that results in a lower cardiovascular risk (Abbenhardt et al., 2013, Petelin et al., 2014). In the face of such evidence, however, there is still a lack of knowledge about the biochemical processes that underlie the alterations produced by cardiovascular risk factors and their relation with the compliance or adoption of healthy lifestyles. To unravel these processes through the identification of molecules and associated metabolic pathways, is of particular importance in areas such as nutritional epidemiology, not only because this could provide a better understanding of the functioning and evolution of pathologies, but also for the establishment of therapeutic or intervention strategies for their control. In recent years, with the development of the called *omics*, and among them, metabolomics, many advances have been made in the identification of molecules characteristic of normal and pathological biological processes.

Metabolomics is defined as the discipline that studies globally low molecular weight molecules (<1500 Da), also known as metabolites, within a biological system (e.g., cells, tissues or organisms) at a given time (Johnson et al. 2016; Monteiro et al., 2013; Gibbons et al., 2015). To the set of metabolites that make up the biological system are known as metabolome and its study through metabolomics allows directly to know the biochemical relationships and biological processes that are carried out in the system. Metabolomic analysis requires the application of a multidisciplinary workflow that includes obtaining the data through advanced analytical techniques such as nuclear magnetic resonance (NMR) spectroscopy and its analysis

by multivariate statistical techniques. The most frequent use of NMR in metabolomics is in studies with untargeted approaches, which consist in the analysis of changes in metabolome by effect of specific conditions, but without previous knowledge of the metabolites to be determined. These untargeted approaches also have the potential to generate a marked increase in the identification of potential biomarkers of biological processes under normal or pathological conditions (Bujak et al., 2015).

One of the most common applications of metabolomics is the identification of disease biomarkers, such as DM2. Specifically, in DM2 research metabolomics has successfully been applied in the discovery of diagnostic, prognosis, metabolic alterations biomarkers and drug action mechanisms (Urpi-Sarda et al., 2015). Most of these biomarkers, however, are usually composed of a single metabolite, which sometimes makes it difficult to replicate in different studies. Some models, however, use the combination of metabolites and information from other variables such as age, gender and lifestyles, for the construction of biomarkers, mainly for predicting DM2 (Herder et al., 2014). It is important to take into account that most of these models have been constructed and tested from blood plasma samples, which in some ways makes it difficult to apply them in observational or large-scale clinical studies. In this sense, the development of biomarkers from less invasive biological samples, such as urine, constitutes a field of opportunity in the investigation of the metabolic alterations produced by DM2. On the other hand, the application of metabolomics in the study of DM2 is essential today to better understand the metabolic intervariability that exists within this disease. Therefore, the study of metabolic phenotypes (metabotypes) related to DM2 through metabolomics represents a starting point in the new "stratified medicine" (Pearson 2016) for the identification of individuals at greater risk and to establish programs of intervention in lifestyle focused on each of these metabotypes (Heinzmann et al., 2012).

On the other hand, and as mentioned above, a lifestyle that include compliance or adoption to healthy dietary patterns such as MedDiet is an essential component in the prevention and reduction of CVD risk. The evaluation of compliance or adherence to a dietary pattern, however, is a challenge in observational or intervention studies, on which to date self-reported food consumption frequency questionnaires are mainly used for their assessment (Hu 2002). Some of these questionnaires are presented in the form of indexes or scores of adherence, and are constructed based on dietary recommendations as a result of a scientific consensus or as proposed by researchers using a scientific evidence approach (Scientific Report of the 2015 Dietary Guidelines Advisory Committee [DGAC] 2015). To date, adherence to MedDiet has been evaluated using different scores described in the literature, including the 9-item questionnaires, proposed by Tricopoulus (Trichopoulou et al., 2003), and the 14-item

questionnaire, proposed by the PREDIMED study (Schroder et al., 2011). These scores have also been used in different studies to study the association of MeDiet adherence with the composition of the intestinal microbiota (Gutierrez-Diaz et al., 2017), cardiovascular risk factors (Martinez-González et al. Age-related diseases (Talegawkar et al., 2012) and total mortality, among others (Trichopoulou et al., 2003). However, because of the complexity of MedDiet, in terms of the amount and frequency with which the foods that comprise it are consumed, evaluation of its adherence through traditional methods may be inaccurate. Therefore, the use of biomarkers of food exposure could be of great complementarity (Andersen et al., 2014). In recent years, metabolomics has played an important role in the discovery of new dietary biomarkers, either of individual food consumption (Garcia-Aloy et al., 2010) or dietary patterns (Gibbons Et al., 2017; Playdon et al., 2017). Examples of dietary patterns described through metabolomic biomarkers include the new Nordic diet (Lankinen et al., 2016; Andersen et al., 2014) and the western diet (Bouchard-Mercier et al., 2013). However, biomarkers of adherence to MedDiet have been poorly studied (Gutierrez-Diaz et al., 2017). Therefore, it is recognized that the use of metabolomic biomarkers could not only supplement the information of the traditional MedDiet adherence scores, but also improve the accuracy of its evaluation and help to establish relationships of its effect on the organism (Bondia-Pons et al. 2015). Thus, the characterization of adherence to a Mediterranean dietary pattern through a metabolic approach is of special interest in nutritional research since it would allow a better understanding of the relationship between diet, taking into account the bioavailability of the bioactive compounds of MedDiet and the risk of chronic diseases associated with CVD.

On the other hand, in lifestyle intervention studies aimed to reduce the risk of CVD or intermediate risk factors, the information from the evaluation of the compliance of a healthy dietary pattern would be incomplete if the beneficial effects of such interventions are not also measured. With respect to the latter, there is a great deal of evidence regarding the multiple benefits of lifestyle interventions based, for example, on low-calorie diets and regular physical activity, aimed to loss weight in obese populations (Petelin et al., 2004 Abbenhardt et al., 2013). However, most of the intervention studies conducted so far with this approach have focused their interest in obesity in general, ruling out, for example, the metabolic variability that exists within this phenotype. In this sense, recent studies have demonstrated the existence of a discordant phenotype within obese individuals characterized by a lower risk of CVD or total mortality (Calori et al., 2011; Hamer & Stamatakis 2012). Individuals with this phenotype are referred to as MHO and have recently begun to gain interest in nutritional epidemiology. However, although diet and physical activity are well-known modifiable cardiovascular risk factors, their potential beneficial impact under controlled conditions on the MHO phenotype is

only gaining interest (Dalzill et al., 2004; Arsenault et al., 2009). Metabolomics applied in interventional studies for weight loss in the MHO phenotype thus offers the opportunity to deep in the knowledge of the molecular mechanisms through which intervention and weight loss modulate, separately or together, the metabolic benefits that can be translated into improved cardiovascular risk. Therefore, the study of metabolomic biomarkers related to lifestyle interventions that combine healthy dietary patterns and physical activity, is of particular interest for the study of the MHO phenotype, not only because these biomarkers contribute to confirm the beneficial effects of such interventions but also because they allow us to understand how this phenotype changes towards a healthier state or lower cardiovascular risk.

For all of the above, it is presented a Doctoral Thesis that integrates the application of the metabolomics by an untargeted approach with NMR in the study of biomarkers and metabolic profiles of one cardiovascular risk factor, such as DM2, and a healthy dietary pattern composed of a Mediterranean diet and its effect on the MHO phenotype after an intervention for weight loss.

II. INTRODUCCIÓN

2.1 METABOLÓMICA

2.1.1 Introducción

De acuerdo a la *Metabolomics Society*, el término metabolómica se refiere a la caracterización global de moléculas de bajo peso molecular (<1500 Da), llamadas metabolitos, dentro de un sistema biológico (p.e., célula, tejido u organismo) en un momento determinado. El conjunto de metabolitos que componen un sistema biológico recibe el nombre de metaboloma y su estudio a través de la metabolómica permite conocer más acerca de las vías metabólicas o procesos biológicos que en ese sistema se llevan a cabo.

Actualmente puede considerarse a la metabolómica como el último eslabón en la cadena o cascada de las ciencias "ómicas". En esta cascada suelen producirse numerosas interacciones entre los distintos niveles conocidos como genoma, transcriptoma, proteoma y metaboloma (Figura 2.1). De este modo los cambios observados en el metaboloma suelen representar las perturbaciones dinámicas de los distintos niveles *ómicos* (Bujak et al. 2015). A diferencia de los genes y proteínas, cuyas funciones están sujetas a regulación epigenética y modificaciones postraduccionales, respectivamente, los metabolitos están directamente relacionados con la actividad bioquímica y por lo tanto son más fáciles de correlacionar con el fenotipo (Patti et al. 2012). En otras palabras, podría considerarse al metaboloma como una firma de la actividad bioquímica de un sistema biológico en momento específico y bajo unas condiciones fisiológicas y/o perturbaciones determinadas (Patti et al. 2012; Kell et al. 2005).

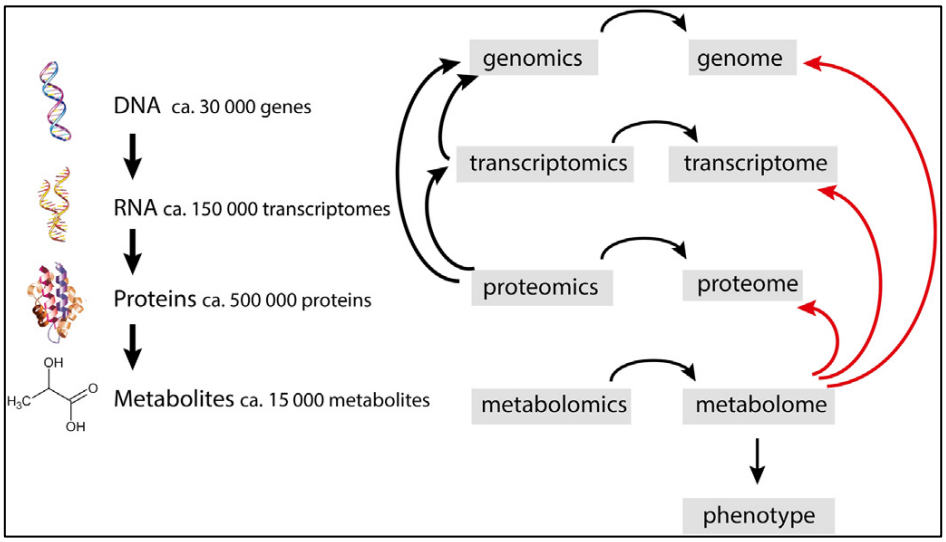


Figura 2.1. La cascada de las ómicas en un enfoque de biología de sistemas. Bujak *et al*., 2015.

En este sentido, la valiosa información que proporciona el estudio del metaboloma justifica en parte las ventajas que ofrece el empleo de la metabolómica en la biología de sistemas y en particular en la investigación de áreas como la biomedicina, toxicología, farmacología y de la alimentación y nutrición.

De acuerdo a Scalbert *et al*, el metaboloma humano se ve influenciado por otros dos subtipos de metaboloma: i) metaboloma endógeno, que incluye todos los metabolitos intrínsecos relacionados con el metabolismo primario e intermedio del organismo, y ii) metaboloma exógeno, compuesto por todos los metabolitos derivados de factores extrínsecos como la medicación (*drug metabolome*), contaminación (*pollulatn metabolome*) y la dieta (*food metabolome*) (Figura 2.2) (Scalbert et al. 2014). Específicamente, el *food metabolome* consiste en el conjunto de metabolitos derivados de la digestión de los compuestos nutricionales esenciales y no esenciales, del metabolismo derivado del individuo (o huésped) y de su microbiota intestinal (Scalbert et al. 2014). Debido a que el metaboloma endógeno y el *food metabolome* se encuentran fuertemente influenciados por la dieta, ambos recibirán especial atención a lo largo de la presente tesis doctoral.

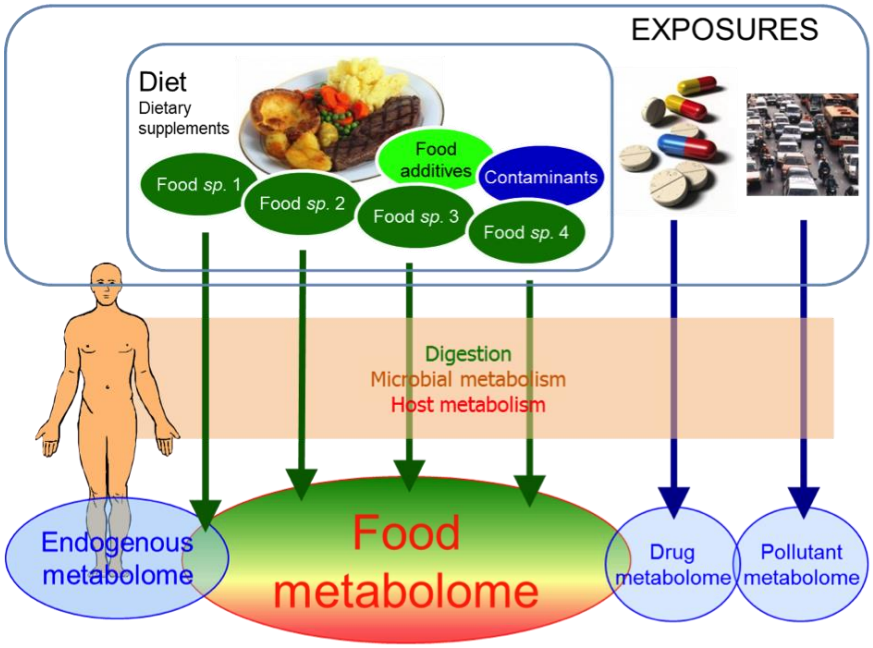


Figure 2.2. The human metabolomes. *Sp*., species. Scalbert *et al*., 2014.

El análisis del metaboloma requiere de la aplicación de un flujo de trabajo multi-disciplinario en el que se incluye la obtención de los datos a través de técnicas analíticas avanzadas. Una de las técnicas analíticas más comúnmente empleadas en la metabolómica es la espectroscopia por

resonancia magnética nuclear (**RMN**) (Beckonert et al. 2007; Beisken et al. 2015; Brennan 2014). Entre las ventajas que ofrece esta técnica se encuentran la de ser altamente reproducible, no destructiva, requiere poco o ningún tratamiento previo de la muestra, permite la detección de metabolitos presentes en concentraciones relativamente altas (μM-mM) en varias muestras biológicas (p.e., orina, sangre, saliva, heces y tejidos), además de proveer información estructural para la identificación de metabolitos no identificados (*unknowns*) (Beckonert et al. 2007). El uso más frecuente de la RMN en la metabolómica es en estudios con **aproximaciones no dirigidas o *untargeted***, que consisten en el análisis de los cambios en metaboloma por efecto de condiciones específicas, pero sin previo conocimiento de los metabolitos que se han de determinar. Estas aproximaciones *untargeted* además tienen la potencialidad de generar un marcado aumento en la identificación de potenciales biomarcadores de procesos biológicos en condiciones normales o patológicos (Bujak et al. 2015).

Un **biomarcador se** define como aquella sustancia medida objetivamente y evaluada como indicador de un proceso biológico normal, estado patogénico o de respuesta a un tratamiento farmacológico (Strimbu & Tavel 2010). En metabolómica se reconocen tres tipos de biomarcadores: i) biomarcadores de exposición, definidos como aquellos compuestos exógenos, o alguno de sus metabolitos, que pueden ser medidos en alguna muestra biológica del organismo; ii) biomarcadores de efecto, aquellos relacionados con alguna alteración bioquímica o fisiológica en el organismo que, dependiendo de la magnitud, se pueden asociar con el deterioro en la salud o enfermedad; y iii) biomarcadores de susceptibilidad, aquellas sustancias indicadoras de la capacidad del organismo para responder a una determinada exposición (Schulte 2005).

2.1.2 Flujo de trabajo de la metabolómica

Con un previo planteamiento de la pregunta de investigación, la hipótesis y el diseño experimental, los estudios de metabolómica se llevan a cabo en cinco etapas que se describen a continuación.

1) Recolección y preparación de las muestras biológicas.

La recolección incluye un control riguroso de la toma, almacenamiento y distribución de las muestras. Mientras que la preparación incluye una serie de operaciones (opcionalmente) como la homogenización, extracción, purificación, concentración, fraccionamiento, etc., de las muestras antes de ser analizadas. Debido a la practicidad en su colección (poco o nada invasiva) y manejo, así como a la información que de ellas puede obtenerse, el plasma, suero y orina

suelen ser las muestras biológicas más comúnmente empleadas en estudios metabolómica (Álvarez-Sánchez et al. 2010; Claus & Swann 2013). El plasma, por ejemplo, proporciona una "lectura instantánea" de los procesos anabólicos y catabólicos (normales o anormales) que ocurren en el organismo entero en el momento del muestreo (Álvarez-Sánchez et al. 2010). Mientras tanto, la orina al contener productos metabólicos finales de reacciones bioquímicas, provee al investigador una visión global de lo que ha ocurrido en el organismo por efecto de la interacción entre los metabolomas endógeno y exógeno (p.e., dieta y microbiota) (Claus & Swann 2013).

2) Adquisición y procesamiento de los datos.

La adquisición de los datos consiste en la generación de los datos a partir del análisis de las muestras biológicas a través de una plataforma analítica. En el caso de la RMN, los datos obtenidos son compuestos por espectros que a su vez incluyen las señales de resonancia y desplazamiento químico de los distintos compuestos encontrados en las muestras. Mientras tanto, el procesamiento de los datos incluye una serie de acciones que contribuyen a reducir la variabilidad de tipo "ruido" dentro y entre las muestras para la obtención de variables (señales, picos, etc.) biológicamente relevantes al momento de realizar los análisis estadísticos.

3) Análisis de los datos.

El análisis de los datos en metabolómica suele ser un proceso "multipaso" que comienza con un pretratamiento de los mismos y que a su vez incluye operaciones de normalización, transformación y escalado (Beisken et al. 2015). Estas operaciones, al igual que los posteriores análisis estadísticos, son llevadas a cabo mediante el empleo de métodos quimiométricos y bioinformáticos en la matriz de datos que contiene por un lado las muestras (o individuos, en filas) y por otro las variables (p.e., señales de RMN) obtenidas de la adquisición de los datos. Las principales técnicas de análisis estadístico empleadas en metabolómica suelen incluir análisis de regresión, correlación, clasificación y agrupamiento de variables, entre otros. La finalidad de estas técnicas es la de reducir significativamente el volumen de datos procedentes de la etapa anterior, seleccionando las variables con la información más relevante o relacionada con los fenómenos estudiados.

4) Identificación de metabolitos.

La identificación de metabolitos es sin duda una de los mayores retos en los experimentos de metabolómica, principalmente en aquellos que se realizan con una aproximación no dirigida. Esta etapa consiste en asignar la identidad (nomenclatura y estructura química) de las señales/variables que fueron significativamente relevantes del análisis estadístico previo. Para la identificación de metabolitos se suele recurrir a la comparación de los espectros en bases de

datos de uso comercial o público, a la búsqueda minuciosa en literatura científica y en algunos casos a la realización de experimentos adicionales para confirmar su identidad.

5) Interpretación biológica.

Mediante la interpretación biológica se busca dar sentido a los resultados obtenidos y relacionarlos con el objetivo principal del estudio, dando a su vez respuesta a la pregunta de investigación inicialmente planteada. Para lograr esto, se recurre al uso de bases de datos que contienen información acerca de los metabolitos identificados, así como a la búsqueda bibliográfica de investigaciones similares o relacionadas con el tema. La construcción de redes metabólicas o el situar a los metabolitos dentro de rutas metabólicas conocidas, son herramientas otro tipo de herramientas que permiten conocer o profundizar en los mecanismos o procesos biológicos que subyacen al fenómeno estudiado (p.e., enfermedad y dieta).

2.2 RIESGO CARDIOVASCULAR

2.2.1 Generalidades

Las enfermedades cardiovasculares (ECV) son un conjunto de trastornos del corazón y los vasos sanguíneos, entre los que se incluyen las cardiopatías coronarias, las enfermedades cerebrovasculares, el aumento de la tensión arterial, las vasculopatías periféricas, las cardiopatías reumáticas y congénitas, y la insuficiencia cardiaca.

Durante los últimos 15 las ECV han ocupado el primer lugar en las causas de mortalidad a nivel mundial, afectando principalmente a los países de ingresos bajos y medios: más del 80% de las defunciones por esta causa se producen en esos países y afectan casi por igual a hombres y mujeres. De aquí a 2030, se estime que 23,6 millones de personas morirán por alguna enfermedad cardiovascular, principalmente por cardiopatías y accidentes cerebrovasculares. Se prevé que estas enfermedades sigan siendo la principal causa de muerte (Organización Mundial de la Salud (OMS) 2017).

2.2.1 Factores de riesgo

La mayoría de las ECV son causadas por factores de riesgo que pueden ser controlados, tratados o modificables. Los principales factores de riesgo se resumen en la tabla 2.1

Tabla 2.1 Principales factores de riesgo de las ECV

Conductuales (modificables)	Metabólicos (controlables)	Otros (no modificables)
Consumo de tabaco	Hipertensión	Edad
Sedentarismo	Hiperglucemia y diabetes	Género
Dieta (malsana)	Hiperlipidemia	Antecedentes familiares de ECV
Consumo nocivo de alcohol	Sobrepeso y obesidad	

Fuente: OMS, 2017

En términos de muertes atribuibles, el principal factor de riesgo cardiovascular a nivel mundial es la hipertensión (a la que se atribuye el 13% de las muertes mundiales), seguido por el

consumo de tabaco (9%), elevada glucosa en sangre (6%), y el sobrepeso y obesidad (5%) (Organization 2011).

Debido a que la presente tesis doctoral centra su interés en factores de riesgo cardiovascular como la obesidad, la diabetes y la dieta, los siguientes apartados se focalizarán en tratar estos.

2.2.2 Obesidad

El sobrepeso y la obesidad se definen como una acumulación anormal o excesiva de grasa que puede ser perjudicial para la salud. La causa fundamental del sobrepeso y la obesidad es un balance energético positivo crónico, lo que significa que las calorías consumidas son mayores a las calorías gastadas (Lavie et al. 2016).

Además del peso, el índice de masa corporal (IMC), la circunferencia de la cintura y la relación cintura-cadera son algunos de los parámetros empleados para medir la obesidad (Mandviwala et al. 2016). El IMC es un indicador simple de la relación entre el peso y la talla que se utiliza con mayor frecuencia para identificar el sobrepeso y la obesidad en los adultos(Organización Mundial de la Salud (OMS) 2016). Se calcula dividiendo el peso de una persona en kilos por el cuadrado de su talla en metros (kg/m^2). En términos generales, se considera que el peso de un adulto es normal (normopeso) cuando su IMC encuentra entre 18.5 y 25 kg/m^2, con sobrepeso si el IMC está entre 25.0 y 29.9 kg/m^2, y obeso si su IMC es $\geq$30 kg/m^2 (Mandviwala et al. 2016).

La obesidad es uno de los principales factores de riesgo cardiovascular debido su relación con numerosas comorbilidades como: la hipertensión y la diabetes mellitus tipo 2 (DM2), además de una relación directa con las ECV, principalmente a través del desarrollo de cardiopatías y accidentes cerebrovasculares (Mandviwala et al. 2016). Los mecanismos intermedios a través de los cuales estas comorbilidades se desarrollan en el marco de la obesidad incluyen: un incremento en la resistencia a la insulina, dislipidemia, trastornos hemodinámicos, daños en la función y estructura cardiaca, inflamación, y desregulaciones neurohumorales y del ciclo celular (Lavie et al. 2016). Otras comorbilidades relacionadas con la obesidad incluyen ciertos tipos de cáncer, trastornos locomotores y apnea/trastornos del sueño, entre otros (Poirier et al. 2006). La obesidad, por lo tanto, se asocia con un mayor riesgo de morbilidad y mortalidad, así como con una reducción de la esperanza de vida (Poirier et al. 2006).

2.2.3 Obesidad Metabólicamente Sana

A pesar de los efectos adversos conocidos en relación a la obesidad, numerosos estudios recientemente han evidenciado la existencia de un fenotipo obeso caracterizado por un bajo

riesgo cardiovascular. A este fenotipo se le conoce como Obesos Metabólicamente Sanos (del inglés, *metabolically healthy obese*, MHO).

Aunque al día de hoy no existe un criterio unificado para su definición, la mayoría de los estudios concuerdan en que la MHO se caracteriza por la ausencia de enfermedades metabólicas como DM2, dislipidemia e hipertensión. La figura 2.3 ilustra las principales diferencias metabólicas entre los MHO y su contraparte, los obesos metabólicamente enfermos.

Figura 2.3. Factores que podrían distinguir entre individuos obesos metabólicamente sanos y metabólicamente enfermos a pesar de poseer una similar cantidad de masa grasa. Basado en Primeau et al., 2011.

Existe una creciente evidencia que sugiere que la capacidad de expansión del tejido adiposo y la inflamación subclínica son los principales mecanismos que determinan si un individuo es MHO o no-MHO (Muñoz-Garach et al. 2016). El primer mecanismo se basa en la "hipótesis de la expansibilidad del tejido adiposo", la cual sostiene que una vez que los adipocitos alcanzan el umbral de su capacidad de almacenamiento (p.e., debido a una sobrecarga calórica crónica), estos comienzan a aumentar en número (hiperplasia) o en tamaño (hipertrofia) (Badoud et al.

2015) (Figura 2.5). Eventualmente, la falta de capacidad de expansión del tejido adiposo conlleva a un estado de disfunción caracterizado por una disminución de la capacidad lipogénica, y aumentos en la producción de productos lipotóxicos y secreción de proteínas proinflamatorias, éste último con el consecuente aumento de infiltración de células inmunes (Badoud et al. 2015; Bell & Hamer 2016). La inflamación subclínica, por su parte, se caracteriza por un aumento en los niveles plasmáticos de moléculas proinflamatorias como la proteína C-reactiva (del inglés, *C*-reactive *protein*, PCR), interleucina-6, (del inglés, *interleukine-6*, IL-6), factor de necrosis tumoral alfa (del inglés, *tumoral necrosis factor alpha* TNF-α), ácidos grasos libres (del inglés, *free fatty acids*, FFA) y adipoquinas (p.e., leptina y resistina), entre otras (Teixeira et al. 2015; Muñoz-Garach et al. 2016; Badoud et al. 2015). Finalmente, todos estos mecanismos se encuentran asociados con la resistencia a la insulina, la cual, como es sabido, es uno de los principales componentes de la enfermedad metabólica. Con todo esto, los individuos MHO se caracterizan por tener una mejor capacidad de expansión de su tejido adiposo, un menor grado de inflamación local y sistémica, y una mayor sensibilidad a la insulina en comparación con individuos no-MHO (Blüher 2014). Estas condiciones metabólicas en el fenotipo MHO se ven además influenciadas por una distribución específica de la grasa corporal que incluye una menor acumulación de grasa visceral y ectópica en comparación con fenotipos no-MHO (Blüher 2014; Muñoz-Garach et al. 2016) (Figura 2.4).

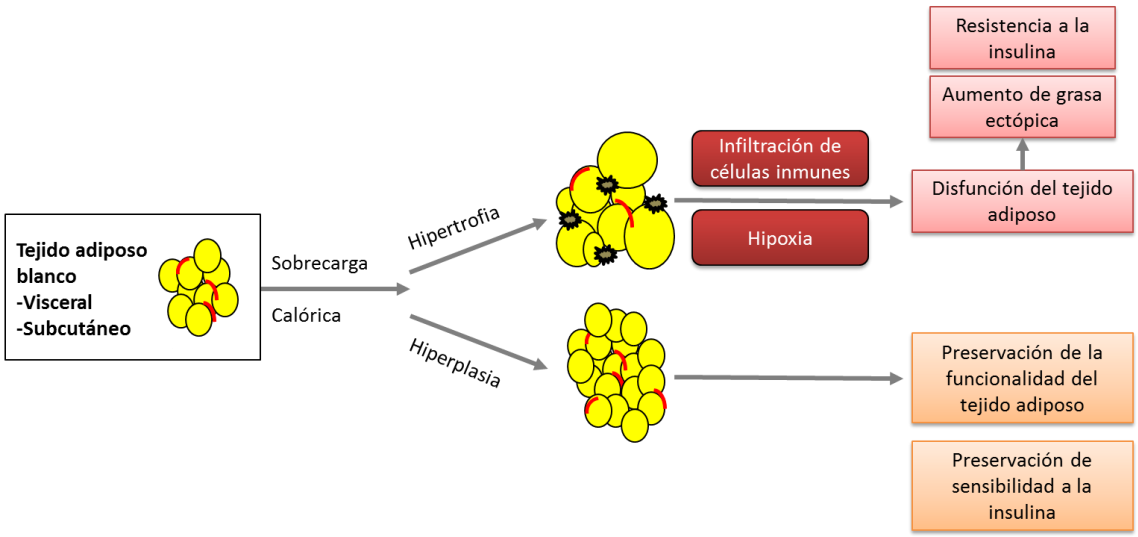


Figura 2.4 Interacción entre la remodelación del tejido adiposo, inflamación y resistencia a la insulina. Adaptado de Badoud et al. (2015).

Se estima que la prevalencia de este fenotipo varía entre el 10 y 34%, dependiendo los criterios empleados para su definición; es más común en las mujeres que en los hombres y parece disminuir con la edad en ambos sexos (Phillips 2017). Aunque existen discrepancias en torno a si la MHO refleja en realidad un estado saludable, es decir un estado permanente de salud, la mayoría de los estudios coinciden en que la MHO es en realidad un estado de transición hacia la obesidad metabólicamente enferma (del inglés, *metabolically unhealthy obese*, MUO) (Bell & Hamer 2016; Blüher 2014; Muñoz-Garach et al. 2016). Diferentes estudios, por ejemplo, han evidenciado que aún en ausencia de factores de riesgo metabólico, los individuos con MHO pueden llegar a tener una mayor inflamación vascular subclínica en comparación con individuos normopeso y metabólicamente sanos (Muñoz-Garach et al. 2016). Otros estudios de larga duración han demostrado no sólo que un número considerable de individuos con MHO llegan a desarrollar MUO, sino que los riesgos de mortalidad por enfermedad cardiovascular y mortalidad total llegan a ser hasta 1,4 y 1,2 veces mayor en este fenotipo en comparación con individuos normopeso sanos, respectivamente (Bell & Hamer 2016)). Ante tales evidencias, por lo tanto, se ha sugerido que la disfunción metabólica y la obesidad son condiciones inseparables, y que la obesidad saludable en realidad debería ser vista como un estado de relativa salud, en lugar de una salud absoluta (Bell & Hamer 2016)).

Por otra parte, debido a que la obesidad en sí misma constituye un factor de riesgo de ECV y ante la falta de evidencias definitivas acerca de una predisposición genética a la MHO, se reconoce que la pérdida de peso, a través de intervenciones quirúrgicas o modificaciones del estilo de vida, podría ser la estrategia con mayores efectos beneficiosos (Muñoz-Garach et al. 2016; Bell & Hamer 2016). De manera general, aunque las intervenciones quirúrgicas contribuyen con una mayor pérdida de peso y beneficios a la salud, también suponen mayores riesgos, complicaciones y costos económicos con respecto a las intervenciones en el estilo de vida (Beaulac & Sandre 2017). En este sentido se ha demostrado que dentro de las intervenciones conductuales y de estilo de vida, la mayor pérdida de peso en individuos obesos se alcanza en aquellas que combinan la ingesta de patrones alimentarios con restricción calórica y la realización de ejercicio (Hassan et al. 2016). Por otro lado, sin embargo, no se tiene completamente claro si el fenotipo MHO podría beneficiarse igualmente de este tipo de intervenciones (Muñoz-Garach et al. 2016).

2.3.4 Diabetes Mellitus

La diabetes es un conjunto de enfermedades metabólicas caracterizadas por elevados niveles de glucosa en sangre (hiperglicemia), como resultado de defectos en la secreción o en la acción de

la insulina (o ambos) (American Diabetes Asociation (ADA) 2017). Los procesos patógenos que están involucrados en el desarrollo de la diabetes van desde la destrucción autoinmune de las células β-pancreáticas, con la consiguiente deficiencia de insulina, hasta anomalías que dan lugar a la resistencia a la acción de la insulina. Frecuentemente, tanto el deterioro de la secreción de la insulina como la resistencia a la misma coexisten en un mismo paciente, lo que dificulta determinar cuál de estas dos es la causa primaria de la hiperglucemia.

La hiperglucemia crónica de la diabetes se asocia con daños, disfunción e insuficiencia de diferentes órganos a largo plazo. Algunas de las principales complicaciones derivadas de la diabetes incluyen: i) retinopatía, con pérdida potencial de la visión; ii) nefropatía, que conduce a insuficiencia renal; iii) neuropatía periférica, con riesgo de úlceras en los pies, artropatía neuropática y amputaciones; iv) neuropatía autónoma, que causa síntomas gastrointestinales, genitourinarios y cardiovasculares y disfunción sexual (American Diabetes Asociation (ADA) 2017). El deterioro del crecimiento y la susceptibilidad a ciertas infecciones también pueden acompañar a la hiperglucemia crónica. Otras consecuencias agudas y mortales de la diabetes no controlada son la hiperglucemia con cetoacidosis y el síndrome hiperosmolar no cetósico (American Diabetes Asociation (ADA) 2017).

Los pacientes con diabetes además tienen una mayor incidencia de enfermedades cardiovasculares, arteriales periféricas y cerebrovasculares ateroscleróticas (Beckman et al. 2013). Uno de los principales mecanismos propuestos para explicar la relación entre la hiperglicemia y la resistencia a la insulina, con los eventos cardiovasculares, comienza señalando que, debido a los niveles fluctuantes de glucosa en sangre, se produce un desbalance en el óxido nítrico (del inglés, *nitric oxide*, NO) endotelial disponible y la acumulación de especies reactivas de oxígeno (del inglés, *reactive oxygen species*, ROS). La generación de ROS inactiva rápidamente el NO para formar peroxinitrito ($ONOO^-$), un potente oxidante que desencadena la nitrosilación de proteínas y disfunción de enzimas implicadas con la homeostasis endotelial (Paneni et al. 2014). Como consecuencia también de la acumulación de ROS se produce la activación de importantes rutas metabólicas que incluyen: i) un aumento en la formación de productos de glicosilación avanzada; ii) activación de las isoformas de la proteína cinasa C; iii) aumento en el flujo de polioles; iv) aumento en el flujo de hexosamina; y iv) aumento en la expresión el factor nuclear kB (NF-kB).

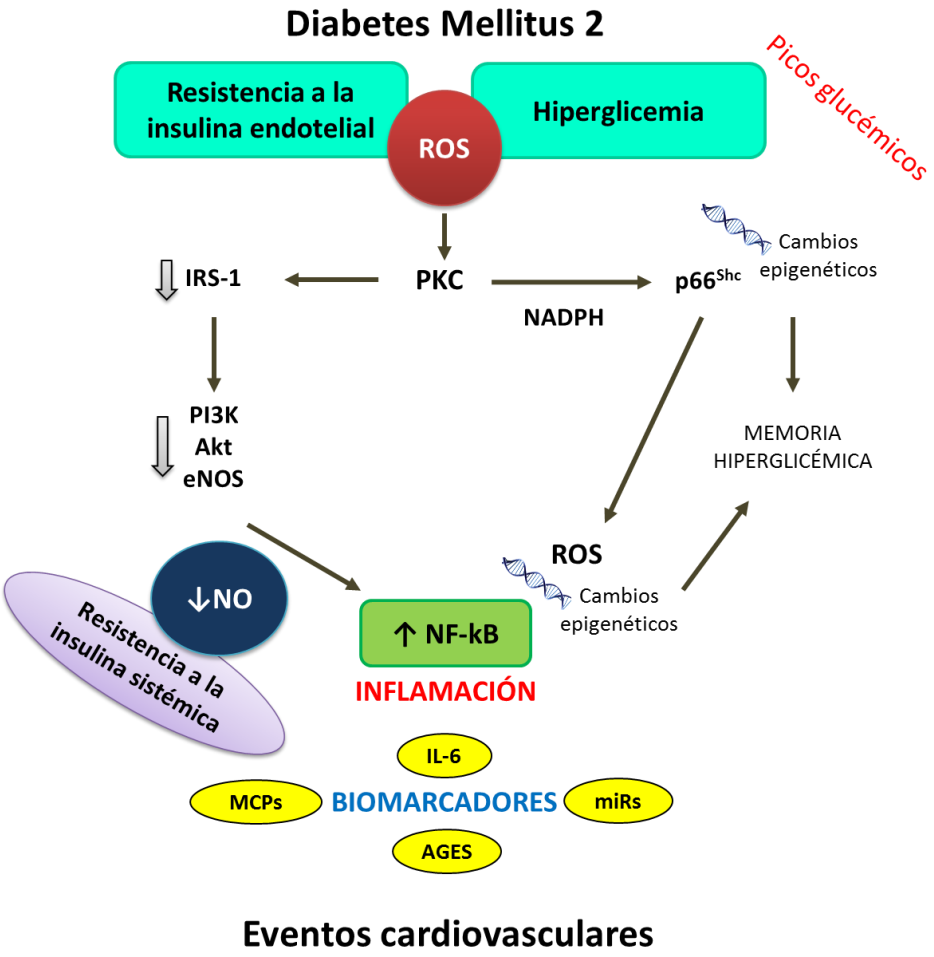


Figura 2.5 Esquema que representa la relación molecular de la diabetes y los eventos cardiovasculares. La inhibición fosforilada del IRS-1 por la PKC perjudica la expresión de PI3K y Akt, lo que conduce a una disfunción de la eNOS y un síntesis reducida de NO. Esta cadena de eventos bloquea el reclutamiento capilar mediado por el NO y perjudica la administración de insulina en órganos sensibles a esta hormona, lo que conduce a un resistencia a la insulina sistémica. Por otro lado, la hiperglucemia provoca la activación de la NADPH oxidasa, dependiente de la PKC, y el adaptador mitocondrial $p66^{Shc}$, que conduce a la generación de ROS, la activación de NF-kB y la regulación positiva de las moléculas inflamatorias. Los picos hiperglucémicos transitorios, así como los ROS desencadenan cambios epigenéticos, que son responsables de la disfunción vascular persistente a pesar de la restauración de la normoglucemia. Tal medio oxidativo e inflamatorio desencadena importantes precursores del daño vascular incluyendo citoquinas circulantes, microRNAs, micropartículas y AGEs, que pueden servir como importantes biomarcadores de ECV en individuos con obesidad y diabetes. Abreviaciones: ECV, enfermedades cardiovasculares; AGEs, productos de glicosilación avanzada (del inglés, *advanced glycosylated end products*); Akt, proteína quinasa B; eNOS, enzima óxido nítrico sintasa endotelial (del inglés, *endotelial nitric oxide synthase*); IL-6 interleuquina (del inglés, *interleukine-6*); IRS-1, receptor de insulina sustrato-1 (del inglés *insulin receptor substrate-1*); NADPH, nicotinamida adenina dinucleótido fosfato (del inglés, *nicotinamide adenine dinucleotide phosphate*); NF-kB, factor nuclear kB (del inglés, *nuclear factor* kB); NO óxido nítrico (del inglés, *nitric oxide*); PI3K, fosfoinositol 3-quinasa (del inglés, *phosphatidylinositide 3-kinase*); PKC proteína quinasa C (del inglés, *Protein kinase C*); ROS especies reactivas de oxígeno (del inglés, *reactive oxygen species*). Tomado de Paneni et al., 2014.

El incremento de la actividad de estas vías, por lo tanto, conlleva a un estado de disfunción endotelial e inflamación que desembocan en un daño vascular y la aparición de eventos cardiovasculares (Paneni et al. 2013; Paneni et al. 2014). En la figura 2.5 se representa de manera esquemática las relaciones moleculares que se desencadenan a partir de la hiperglucemia y la resistencia a la insulina hasta la aparición de eventos cardiovasculares.

La diabetes es una de las mayores emergencias mundiales de salud del siglo XXI. La prevalencia de la diabetes ha aumentado de manera constante en los últimos tres decenios, en paralelo al incremento de la prevalencia de la obesidad y el sobrepeso. En particular, la prevalencia de la diabetes aumenta con mayor rapidez en los países de ingresos bajos y medios, donde el acceso a las tecnologías y medicamentos necesarios es limitado (Organización Mundial de la Salud 2016). Se estima, por ejemplo, que tan sólo en el año 2015, 415 millones de adultos padecían esta enfermedad y que otros 318 millones tenían intolerancia a la glucosa, lo que les situaba ya con un alto riesgo de desarrollar la enfermedad en el futuro (International Diabetes Dederation 2015).

Existen tres tipos principales de diabetes: i) diabetes tipo 1 (DM1), ii) DM2 y iii) diabetes gestacional. La DM1, o diabetes insulinodependiente, es causada por una reacción autoinmune, en la que el sistema de defensa del cuerpo ataca a las células-beta productoras de insulina en el páncreas. Esta forma de diabetes, que es más frecuente en niños y jóvenes adultos, representa del 5 al 10% de los diabéticos. En la DM2, o diabetes no insulinodependiente, el cuerpo es capaz de producir insulina pero se vuelve resistente a ella, de modo que la insulina es ineficaz e incluso, con el tiempo, insuficiente. Este tipo de diabetes es la más común de todas, al representar del 90-95% de los diabéticos, ocurre generalmente en adultos, aunque se observa cada vez más en niños y adolescentes. La mayoría de los pacientes con DM2 suelen ser obesos, cuya condición por sí misma provoca ya un cierto grado de resistencia a la insulina (American Diabetes Asociation (ADA) 2017). Frecuentemente este tipo de diabetes no es diagnosticada sino hasta después de varios años, debido a que la hiperglucemia se desarrolla gradualmente y en etapas tempranas que a menudo no es lo suficientemente grave como para que el paciente note cualquiera de los síntomas clásicos de la diabetes (International Diabetes Dederation 2015).

Por otra parte, se reconoce que la implementación de una serie de intervenciones costoeficientes y oportunas pueden ayudar a las personas en riesgo o con diagnóstico de DM2. Entre las intervenciones cabe mencionar la reducción del peso corporal y el control de la glucemia mediante modificaciones en el estilo de vida y, en caso necesario, medicación; el control de la tensión arterial y el perfil lipídico para reducir el riesgo de ECV y otras complicaciones; además de revisiones médicas periódicas que permitan detectar a tiempo posibles complicaciones de la DM2 para facilitar la introducción oportuna de un tratamiento (Organización Mundial de la

Salud 2016). Específicamente, dentro de los cambios en el estilo de vida, se ha demostrado que la adopción de patrones alimentarios ricos en cereales integrales, frutas, verduras, legumbres y frutos secos; con moderado consumo de alcohol; y menor consumo de granos refinados, carnes rojas o procesadas y bebidas azucaradas, reducen el riesgo de diabetes y mejoran el control glucémico y los lípidos en sangre en pacientes con diabetes (American Diabetes Association (ADA) 2017).

2.3 DIETA MEDITERRANEA

2.3.1 Definición

Declarada patrimonio cultural inmaterial de la humanidad, la dieta mediterránea (DietMed) es un patrón alimentario característico de las zonas oleicas de la región del Mediterráneo y descrito a partir de la década de los años 60 (Trichopoulou et al. 2014). Aunque existen variantes en su descripción de acuerdo al lugar donde se consume, se reconoce que la tradicional DietMed se caracteriza por i) una alta ingesta de aceite de oliva virgen extra, verduras (incluyendo las de hoja verde), frutas, cereales integrales, legumbres y frutos secos; ii) un consumo moderado de pescado y aves de corral; iii) un bajo consumo de productos lácteos, carnes rojas, productos cárnicos procesados, y dulces; y iv) un consumo moderado de alcohol (principalmente en forma de vino) entre las comidas (Davis et al. 2015).

Recientemente, en una adaptación del esquema de la pirámide de la DietMed se han incorporado elementos cualitativos y cuantitativos en la selección de alimentos, tomando en cuenta aspectos de carácter cultural y social íntimamente ligados al estilo de vida mediterráneo (Bach-Faig et al. 2011). Dichos aspectos incluyen entre otros: a) moderación, haciendo hincapié en que el tamaño de las porciones de los alimentos representados en la parte inferior de la pirámide debe ser mayor y consumirse frecuentemente, ya que proporcionan saciedad junto con cantidades moderadas de energía, y los alimentos en los niveles superiores deben ser consumidos en cantidades menores y menos frecuentemente debido a su alto contenido de grasa y azúcares simples; b) socialización, más allá de aspectos nutricionales, este aspecto se refiere a factores relacionados con la alimentación (entendida como un hecho social), como las actividades culinarias, el conocimiento transmitido de generación en generación, el tiempo dedicado a comidas, compartir la comida compañía de familiares y amigos; c) la práctica regular de actividad física (al menos 30 minutos a lo largo del día), como un complemento básico a la dieta equilibrando la ingesta de energía, manteniendo el peso corporal saludable y proporcionando muchos otros beneficios para la salud; d) descanso, durante el día, en forma de siesta, como en la noche; e) estacionalidad, refiriéndose a la preferencia por alimentos estacionales, frescos y mínimamente procesados, contribuyendo con esto a maximizar el contenido de nutrientes y compuestos bioactivos en la dieta; y f) consumo de productos tradicionales, locales y *eco-friendly* (Bach-Faig et al. 2011). A continuación, en la figura 2.6 se muestra la pirámide de la DietMed en la que se incluyen los aspectos señalados anteriormente.

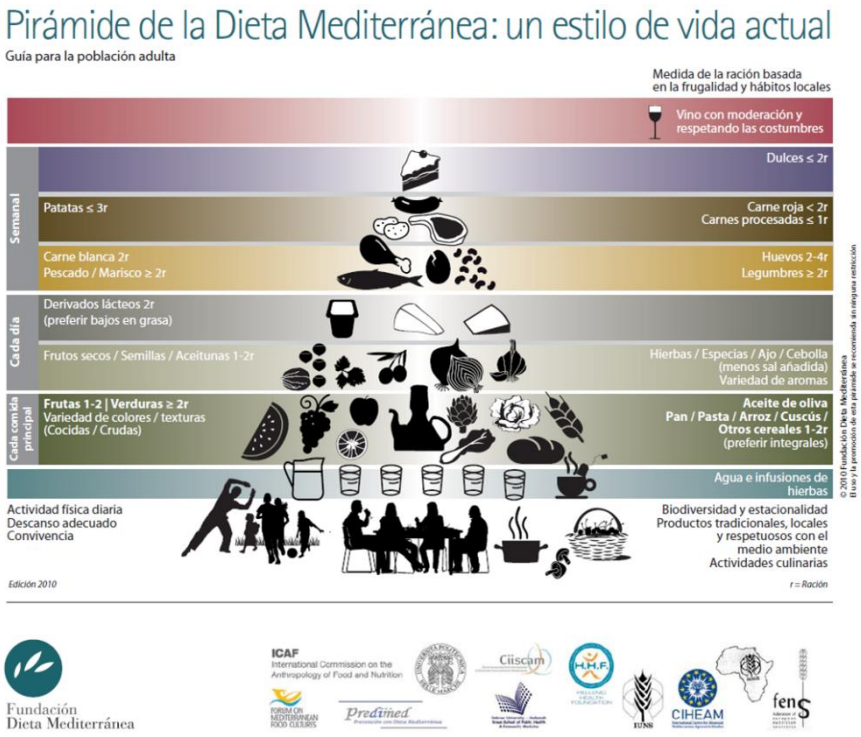


Figura 2.6 La nueva pirámide de la Dieta Mediterránea: Un estilo de vida actual. Fuente Bach-Faig et al (2011).

2.3.2 Dieta Mediterránea y enfermedad cardiovascular

Existe una numerosa y cada vez más creciente evidencia científica acerca de los efectos beneficiosos de la DietMed en la salud (Figura 2.7). La mayoría de dichas evidencias documentadas hasta el día de hoy giran en torno al efecto preventivo y protector de la DietMed frente a las ECV, diabetes, síndrome metabólico (SM), cáncer y, más recientemente, enfermedades neurodegenerativas (Dinu et al. 2017; Martinez-Gonzalez & Martin-Calvo 2016). Los mecanismos a través de los cuales se producen estos efectos beneficiosos de la DietMed también se han documentado y abarcan desde la modulación en la expresión de genes y moléculas consideradas de riesgo (p.e., moléculas de inflamación), hasta la modulación positiva de parám etros clínicos (p.e., perfil lipídico) e índices antropométricos (p.e., peso e IMC). No es de extrañar, por lo tanto, que debido a esta consistente y amplia variedad de evidencias científicas el departamento de agricultura de los Estados Unidos (USDA, por sus siglas en inglés) decidiera recientemente incluir a la DietMed dentro de su lista de patrones alimentarios saludables y recomendados para la población estadounidense, junto al *U.S.-style Pattern* y al

Healthy Vegetarian Pattern (Scientific Report of the 2015 Dietary Guidelines Advisory Committee [DGAC] 2015).

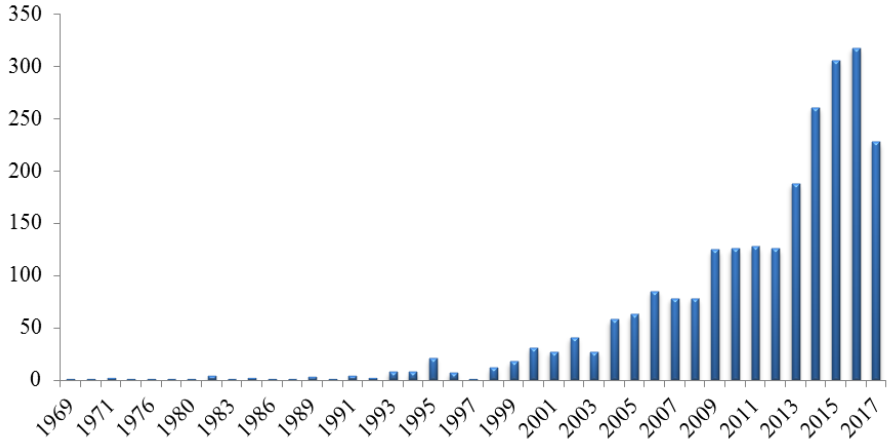


Figura 2.7 Número de publicaciones proporcionadas por el motor de búsqueda PubMed mediante la búsqueda con las palabras "Mediterranean diet AND health".

El mayor número de evidencias científicas generadas en relación a los efectos beneficiosos de la DietMed se centra en la relación inversa que existe entre el seguimiento de este patrón alimentario y el riesgo de desarrollar (y morir por causas relacionadas con) ECV. Cronológicamente, algunos de los estudios que han contribuido sustancialmente con estas evidencias incluyen el *Seven countries: a multivariate analysis of death and coronary heart disease*, conducido por Ancel Keys y colaboradores durante la década de los años 50 (Keys et al. 1980); el *Lyon Diet Heart Study*, dirigido por de Lorgeril y colaboradores en la década de los 90 (de Lorgeril et al. 1994; de Lorgeril et al. 1999); el estudio conducido por Trichopoulou y colaboradores, también durante la década de los 90 (Trichopoulou et al. 2003); y el estudio "Prevención con Dieta Mediterránea" (PREDIMED), recientemente publicado por Estruch y colaboradores (Estruch et al. 2013). Estos y otros muchos estudios realizados en las últimas dos décadas han demostrado que estilos de vida con una alta adherencia a la DietMed se relacionan no sólo con una menor incidencia de eventos cardiovasculares fatales y no fatales (incluyendo la enfermedad coronaria, insuficiencia cardiaca y accidente cerebrovascular) en individuos con riesgo (lo que se conoce como prevención primaria), sino también con un retraso en la progresión de las ECV y una disminución en la aparición futuros eventos cardiovasculares en individuos con ECV preexistente (prevención secundaria) (Martinez-Gonzalez & Martin-Calvo 2016; Shen et al. 2015; Sofi et al. 2013). Por otro lado, tal ha sido la importancia y el interés en que ha generado este patrón alimentario en la salud pública que en los últimos años se han replicado con éxito varios estudios observacionales y de intervención con DietMed en

poblaciones no mediterráneas (Martinez-Gonzalez & Martin-Calvo 2016). Sumado a todo lo anterior, numerosas revisiones sistemáticas y meta-análisis han concluido que la DietMed es con toda seguridad uno de los factores en el estilo de vida que mejor protege frente al desarrollo y muerte por ECV. Por ejemplo, uno de los más recientes meta-análisis conducidos por Grosso y colaboradores (2017) concluyó que la alta adherencia a la DietMed se asocia hasta con un 19% de disminución en el riesgo de incidencia y mortalidad por ECV (Grosso et al. 2017).

Como se mencionó anteriormente, los efectos cardioprotectores de la DietMed se encuentran íntimamente relacionados, de manera directa o indirecta, con la modulación y efecto que tiene este patrón alimentario sobre diversos factores de riesgo cardiovascular como (Martinez-Gonzalez et al. 2009; Shen et al. 2015):

- Sobrepeso y obesidad
- Dislipidemia
- Presión arterial
- Desregulación en el metabolismo de la glucosa
- Resistencia a la insulina
- Inflamación
- Disfunción endotelial
- Estrés oxidativo
- Expresión de variantes genéticas (polimorfismos) proaterogénicas

A su vez, de acuerdo a la literatura, la modulación positiva de la mayoría de estos factores de riesgo cardiovascular se ve influenciada en mayor o menor medida por la ingesta de distintos componentes de la DietMed como:

a) Fibra, especialmente proveniente de frutas, verduras, legumbres y granos integrales
b) Una favorable relación de ácidos grasos, con especial abundancia de ácidos grasos poliinsaturados omega-3 (como el ácido α-linoleico, de origen vegetal)
c) Ácidos eicosapentanoico y docosahexanoico, provenientes del pescado
d) Vitaminas, minerales y compuestos fenólicos, provenientes de frutas y verduras, con propiedades antioxidantes y antiinflamatorias

No obstante, pese a que los efectos beneficiosos de estos componentes y los alimentos que los incluyen (p.e., aceite de oliva, frutos secos, pescado, vino) sobre los citados factores de riesgo cardiovascular se encuentran documentados, se reconoce ampliamente también que el consumo combinado de estos componentes a través de la DietMed proporciona los mayores beneficios sobre la salud cardiovascular (efecto sinérgico) (Grosso et al. 2017; Martinez-Gonzalez & Martin-Calvo 2016).

2.3.3 Dieta Mediterránea, obesidad, diabetes y síndrome metabólico

La obesidad, diabetes y SM son facores de riesgo cardiovascular que suelen estar íntimamente relacionados y que a su vez se ven disminuidos en el riesgo de incidencia y progresión por el consumo de una DietMed.

El papel de la DietMed en el desarrollo y/o modificación del sobrepeso y la obesidad siempre ha sido de gran interés para la investigación clínica. Hasta la fecha, por ejemplo, se desconoce algún tipo de evidencia que sugiera que la adherencia a la DietMed se encuentre relacionada con el aumento de peso en personas con normopeso y/o proporcione protección contra el riesgo de experimentar un estado de sobrepeso y/u obesidad (Dinu et al. 2017). No obstante, existe un gran número de evidencias acerca de los efectos beneficiosos del seguimiento de una DietMed sobre el control del peso en individuos obesos. Específicamente, se ha demostrado que la alta adherencia a la DietMed en poblaciones con sobrepeso y obesidad se encuentra directamente asociada con un mantenimiento y pérdida de peso, y reducciones en la circunferencia de la cintura y el índice de masa corporal, así como con una reducción en la grasa abdominal (Estruch et al. 2016; Funtikova et al. 2014; Eguaras et al. 2015; Bertoli et al. 2015; Park et al. 2017). Distintos meta-análisis han soportado estas evidencias y han concluido, además, que la DietMed es una herramienta útil para la pérdida de peso en individuos obesos, y que estos beneficios se acentúan más cuando la DietMed incluye restricciones calóricas y/o la práctica regular de actividad física (Esposito et al. 2011; Sofi et al. 2013; Romieu et al. 2017).

En cuanto a la diabetes mellitus, existen publicaciones de numerosos estudios transversales, prospectivos y de ensayos clínicos aleatorizados que evidencian el papel protector de la DietMed frente al riesgo de desarrollar esta enfermedad. Por ejemplo, en un reciente meta-análisis de estudios prospectivos de cohorte publicados entre 2007 y 2014, e incluyendo 122.810 individuos, se concluyó que una mayor adherencia a la DietMed se asociaba con hasta un 19% de menor riesgo de desarrollar diabetes (Schwingshackl et al. 2015). En otro meta-análisis de estudios prospectivos, similar pero independiente al primero, e incluyendo hasta 136.846 individuos, se concluyó que el seguimiento de una alta adherencia a la DietMed se encontraba relacionado con un 23% de menor riesgo de desarrollar DM2 (Koloverou et al. 2014). Por su parte, aunque con menor número de publicaciones, ha habido estudios clínicos de intervención que han investigado la relación entre la DietMed y la diabetes. Uno de los más recientes, conducido por Salas-Salvadó y colaboradores, y desarrollado en el marco del estudio PREDIMED, logró demostrar que los individuos asignados a los grupos de intervención con DietMed tuvieron hasta un 30% de reducción en el riesgo de desarrollar DM2 en comparación con el grupo control (Salas-Salvadó et al. 2014). Interesantemente las evidencias acerca de

efecto protector de la DietMed frente al desarrollo de la diabetes no se limitan sólo a la DM2. Estudios recientes, por ejemplo, han demostrado que la exposición a DietMed disminuye también significativamente el riesgo de desarrollar de diabetes gestacional en mujeres embarazadas (Salas-Salvadó et al. 2016). Por otra parte, la DietMed no sólo es efectiva en la prevención de la diabetes, sino también en su control. Esto último se ha demostrado a través de varios estudios clínicos y resumido en un meta-análisis que reunió las evidencias de 20 ensayos clínicos con 3.074 individuos (Ajala et al. 2013). En este meta-análisis de estudios de intervención, con periodos iguales o mayores a seis meses, se encontró que los individuos diabéticos asignados al grupo de intervención con DietMed mostraban mayores mejoras en el control glucémico y sensibilidad a la insulina en comparación con los individuos asignados a un grupo control (Ajala et al. 2013). Todas estas evidencias en relación a la diabetes mellitus, han sumado para considerar a la DietMed como una de las opciones no farmacológicas válidas y más efectivas en su prevención y tratamiento (Sofi et al. 2013).

Por otra parte, varios estudios transversales y prospectivos han sugerido que la DietMed tiene efectos protectores sobre diferentes componentes de SM. Algunas de estas evidencias han quedado resumidas en un meta-análisis sobre estudios epidemiológicos y ensayos clínicos, dirigido por Kastorini y colaboradores (2011). En este trabajo, los autores concluyeron, mediante la combinación de los dos tipos de estudios, no sólo que la adherencia a la DietMed se asociaba con un reducido riesgo de SM (hasta un 31% de menor riesgo), sino que también las intervenciones con DietMed, en comparación con dietas control, se encontraban asociadas con efectos beneficiosos en todos los componentes del SM, incluyendo la circunferencia de cintura, colesterol HDL, triglicéridos, presión arterial, glucemia en ayunas y resistencia a la insulina (Kastorini et al. 2011). Uno de los últimos estudios clínicos en evidenciar los efectos beneficiosos de la DietMed sobre el SM fue el llevado a cabo por Babio y colaboradores (2014) en el marco del estudio PREDIMED. Este estudio evidenció que la intervención con DietMed se asoció con una reversión del SM, particularmente por una reducción de la hiperglicemia y la circunferencia de cintura, pero no con la prevención de su incidencia (Babio et al. 2014). Tales han sido las evidencias en torno a los efectos beneficiosos de la DietMed en la prevención del SM y la mejora de sus componentes, que recientemente un panel internacional de expertos en el área decidió incluir a este patrón alimentario dentro de sus recomendaciones en el estilo de vida para prevenir y tratar el SM (Perez-Martinez et al. 2017).

2.3.4 Componentes bioactivos

La DietMed es un patrón alimentario rico en alimentos de origen vegetal, los cuales a su vez constituyen una excelente fuente de componentes nutricionales y bioactivos que en lo individual

o de manera sinérgica son los principales causantes de los efectos beneficiosos sobre la salud metabólica. Los principales componentes bioactivos conocidos en la DietMed incluyen los ácidos grasos insaturados, carbohidratos complejos y fibra, proteínas vegetales, minerales no sódicos, fitoesteroles y polifenoles. Estos componentes y su relación directa beneficiosa sobre distintos marcadores de riesgo de diabetes y SM se ilustran a continuación en la figura 2.8.

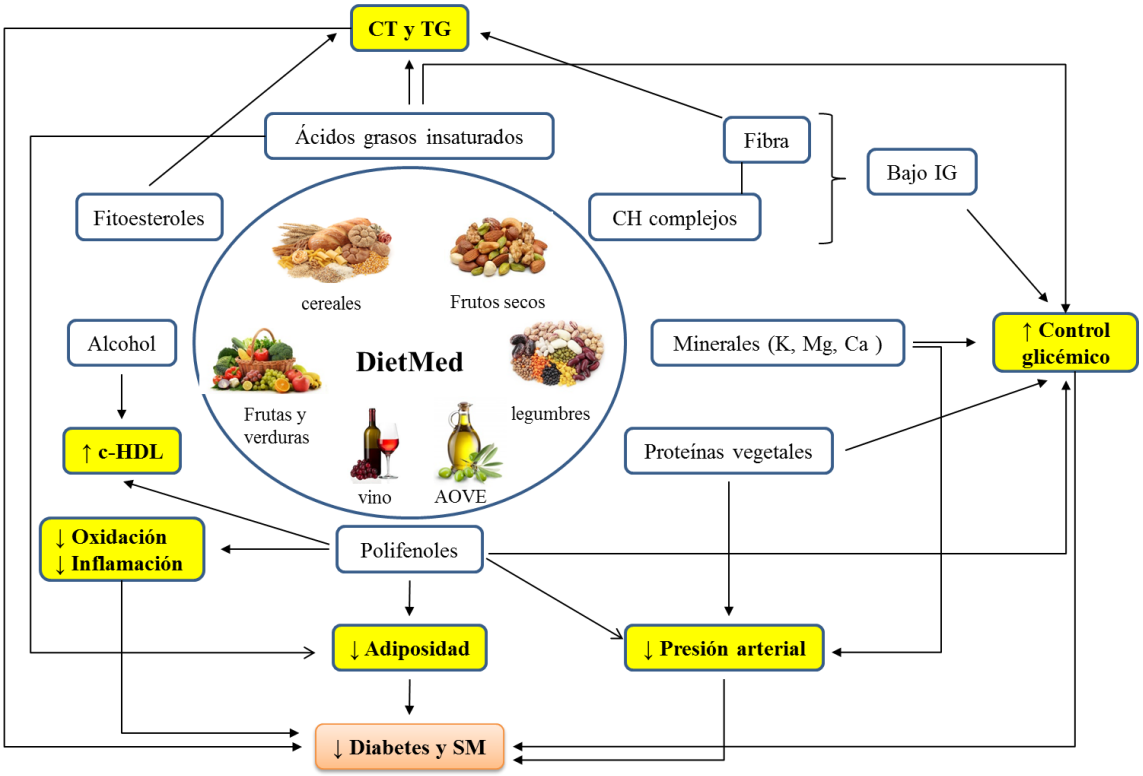


Figura 2.8 Mecanismos propuestos que relacionan la dieta Mediterránea con prevención y tratamiento de la diabetes y síndrome metabólico. Basado en Salas-Salvadó et al. 2016. Abreviaciones y acrónimos: AOVE, aceite de oliva virgen extra; c-HDL, colesterol de lipoproteínas de alta densidad; CT, colesterol total; DietMed, dieta Mediterránea; IG, índice glicémico; SM, síndrome metabólico; TG, triglicéridos

Interesantemente, a pesar de que la DietMed es un patrón alimentario rico en grasas, estas son de origen vegetal, obtenidas principalmente por la ingesta de aceite de oliva y frutos secos, los cuales a su vez son también fuente de micronutrientes y fitoquímicos altamente bioactivos. Este último hecho, por lo tanto, hace de la DietMed un patrón alimentario altamente saludable, incluso superior con respecto a otros patrones alimentarios con bajo contenido de grasa, los cuales también son recomendados desde el punto de vista clínico como beneficiosos para la salud metabólica (Salas-Salvadó et al. 2016).

2.3.5 Dieta Mediterránea, microbiota y salud metabólica

Actualmente la microbiota intestinal es considerada como un órgano más del cuerpo humano ya contribuye a regular numerosos procesos fisiológicos y diferentes funciones del huésped (Woting & Blaut 2016). Se cree que una parte sustancial de la fisiología intestinal, incluyendo la modulación de la motilidad intestinal, la homeostasis de la barrera intestinal, la absorción de nutrientes y la distribución de la grasa, están influenciada por la relación mutua entre la microbiota intestinal y su huésped humano (Duranti et al. 2017). Recientemente, además, se ha comprobado el papel que juega la microbiota intestinal como factor ambiental crucial en el desarrollo de enfermedades metabólicas como la obesidad y la DM2, debido principalmente a su contribución en la modulación de varios procesos que incluyen el metabolismo energético del huésped, permeabilidad epitelial intestinal, secreción de hormonas peptídicas intestinales, y el estado inflamatorio del huésped (Palau-Rodriguez et al. 2015). No es de extrañar por lo tanto que exista una relación directa entre la composición de la microbiota, su actividad y la ingesta de alimentos. En este sentido, durante las últimas décadas se ha logrado evidenciar el papel que juegan diversos alimentos por separado o a través de patrones alimentarios en la modulación de la microbiota intestinal y su relación con la incidencia de enfermedades crónicas como la obesidad y la DM2 (Sonnenburg & Bäckhed 2016). Por ejemplo, se ha observado que patrones alimentarios con un alto contenido energético, principalmente por consumo de grasas de origen animal, y alto consumo de carnes rojas, se encuentran asociados con un aumento en la abundancia de bacterias del género *Bacteroides*; mientras que patrones alimentarios ricos en fibra y compuestos fenólicos favorecen la abundancia de bacterias intestinales del género *Prevotella* (Eid et al. 2017; Arora & Bäckhed 2016; Sonnenburg & Bäckhed 2016). En consonancia con este último tipo de patrones alimentarios, durante los últimos años el interés por los efectos beneficiosos de la DietMed se ha extendido al estudio de su capacidad para modular la composición y actividad de la microbiota intestinal.

Por ejemplo, en dos estudios independientes conducidos por Haro y colaboradores (2016) se observó que la intervención con DietMed no solo favoreció la abundancia de microbiota intestinal relacionada de manera inversa con la incidencia de diabetes en individuos obesos, sino que también contribuyó con la restauración de microbiota intestinal afectada por el SM (Haro, Montes-Borrego, et al. 2016; Haro, Garcia-Carpintero, et al. 2016). Resultados similares se observaron en otro estudio dirigido por De Filippis y Colaboradores (2016) en una población sana. En este último estudio, interesantemente se observó que la composición de la microbiota intestinal contribuyó por sí misma a distinguir entre una alta y baja adherencia a la DietMed en los participantes (De Filippis et al. 2015). Estas y otras evidencias ponen de manifiesto una vía más a través de la cual la DietMed proporciona múltiples beneficios a la salud, tanto en la

prevención como en el control de enfermedades crónicas y que son factores de riesgo cardiovascular.

2.3.6 Evaluación de la adherencia a la dieta Mediterránea

La evaluación del seguimiento o adherencia a la DietMed es un elemento imprescindible en los estudios clínicos y epidemiológicos que buscan establecer asociaciones entre este patrón alimentario y los efectos beneficiosos a la salud. Hasta el día de hoy, una de las forma más comunes empleadas para medir la adherencia a un patrón alimentario es a través del uso de cuestionarios que se presentan en forma de índices o *scores* (Hu 2002). Estos índices de adherencia se construyen se construye en base a recomendaciones dietéticas, las cuales son el resultado de un consenso científico o la propuesta de un grupo de investigadores que emplean un enfoque basado en las evidencias científica (Scientific Report of the 2015 Dietary Guidelines Advisory Committee [DGAC] 2015).

En este sentido, y respecto a la DietMed, en un interesante trabajo publicado por Hernández-Ruiz y colaboradores (2015) se reportó la existencia de hasta 22 índices que cuantifican la adherencia a este patrón alimentario. Una lista de estos índices, sus componentes y sus rangos de puntuación, basada en el trabajo de estos autores, se muestra en la tabla 2.2 (Hernandez-Ruiz et al. 2015). De acuerdo a esta revisión, Trichopoulou y colaboradores (1995) fueron los primeros en definir un índice de adherencia a la DietMed, para evaluar la asociación entre la adherencia a la misma y el riesgo de mortalidad en una población griega de edad avanzada (Trichopoulou et al. 1995). Este primer índice estuvo compuesto por 8 componentes, de los cuales 7 correspondían a alimentos/grupos de alimentos (ingesta de frutas, verduras, frutos secos, legumbres, granos, carne y productos cárnicos, leche y productos lácteos, y vino tinto) y 1 *ratio* para considerar la calidad de la grasa (ácidos grasos monoinsaturados/ácidos grasos saturados). Posteriormente este índice fue modificado por los mismos autores mediante la inclusión de la ingesta de pescado, obteniéndose después que, de igual modo, una alta adherencia a la DietMed se asociaba con un bajo riesgo de mortalidad (Trichopoulou et al. 2003).

A partir de estos primeros trabajos, se han publicado varios índices más para evaluar la adhesión a la DietMed (Tabla 2.2). Algunos de estos, por ejemplo, se basan en ligeras modificaciones de la escala inicialmente propuesta por Trichopolou y colaboradores (1995), mientras que otros no comparten aspectos importantes en su definición.

Tabla 2.2 Descripción de los componentes de los índices de adherencia a la dieta Mediterránea[1]

#	Referencia	Índice	Componentes	Puntuación	Rango
1	(Trichopoulou et al. 1995)	*Mediterranean Diet Scale* 1995 (MDS-95)	8 (7 grupos de alimentos/alimentos individuales, 1 *ratio*)	0-1	0-8
2	(Trichopoulou et al. 2003)	*Mediterranean Diet Scale* 1995 (MDS-03)	9 (7 grupos de alimentos/alimentos, 1 nutriente, 1 *ratio*)	0-1	0-9
3	(Asghari et al. 2013)	*Mediterranean Diet Scale* 2013 (MDS-13)	10 (8 grupos de alimentos/alimentos, 2 relaciones)	0-1	0-10
4	(Trichopoulou et al. 2005)	*Modified Mediterranean Diet* (MMD)	9 (8 grupos de alimentos/alimentos, 1 nutriente)	0-1	0-9
5	(Martinez-Gonzalez et al. 2002)	*Mediterranean Dietary Pattern* 2002 (MDP-02)	8 (6 grupos de alimentos/alimentos, 2 nutrientes)	0-5	5-40
6	(Ciccarone et al. 2003)	*Mediterranean Dietary Pattern* 2003 (MDP-03)	18 (17 grupos de alimentos/alimentos, 1 nutriente)	0-1	0-14
7	(Sanchez-Villegas et al. 2006)	*Mediterranean Dietary Pattern* 2006 (MDP-06)	10 (9 grupos de alimentos/alimentos, 1 nutriente)	1-3	10-30
8	(Sanchez-Villegas et al. 2006)	*Mediterranean-Style Dietary Pattern Score*(MSDPS)	13 (13 grupos de alimentos/alimentos)	0-10	0-100
9	(Sanchez-Villegas et al. 2006)	*Mediterranean Diet Quality Index* (MDQI)	7 (5 grupos de alimentos/alimentos, 2 nutrientes)	2-0	14-0
10	(Sanchez-Villegas et al. 2006)	*Mediterranean Adequacy Index* (MAI)	16 (16 grupos de alimentos/alimentos)	-	0-100

Tabla 2.2 Continuación

10	(Sanchez-Villegas et al. 2006)	*Mediterranean Adequacy Index* (MAI)	16 (16 grupos de alimentos/alimentos)	-	0-100
11	(Fung et al. 2005)	*Alternate Mediterranean Diet Index* (aMED)	9 (8 grupos de alimentos/alimentos, 1 nutriente)	0-1	0-9
12	(Agnoli et al. 2011)	*Italian Mediterranean Index* (ITALIAN-MED)	11 (11 grupos de alimentos/alimentos)	0-1	0-11
13	(Buckland et al. 2009)	*Relative Mediterranean Diet Score* (rMED)	9 (8 grupos de alimentos/alimentos, 1 nutriente)	0-2	0-18
14	(Woo et al. 2001)	*Mediterranean Diet Score 2001* (MDS-01)	8 (6 grupos de alimentos/alimentos, 1 nutriente, 1 *ratio*)	0-1	0-8
15	(Schroder et al. 2004)	*Mediterranean Diet Score 2004* (MDS-04)	9 (9 grupos de alimentos/alimentos)	1-3	9-27
16	(Pitsavos et al. 2005)	*Mediterranean Diet Score 2005* (MDS-05)	11 (11 grupos de alimentos/alimentos)	0-5	0-55
17	(Panagiotakos et al. 2007)	*Mediterranean Diet Score 2007* (MDS-07)	11 (11 grupos de alimentos/alimentos)	0-5	0-55
18	(Yang et al. 2014)	*Modified Mediterranean Diet Score* (mMDS)	10 (8 grupos de alimentos/alimentos, 2 factores de estilo de vida)	0-2 0-4 0-8	0-42
19	(Martinez-Gonzalez et al. 2002)	*Cardioprotective Mediterranean Diet Index* (Cardio)	9 (9 grupos de alimentos/alimentos)	0-1	0-9

Tabla 2.2 Continuación

20	(Martinez-González et al. 2012)	*Mediterranean food pattern PREDIMED Study* (MeDiet-PREDIMED/ MEDAS)	14 (14 grupos de alimentos/alimentos)	0-1	0-14
21	(Sofi et al. 2014)	*Mediterranean diet based on the literature* (MED-LITE)	9 (8 grupos de alimentos/alimentos, 1 nutriente)	0-2	0-18
22	(Sotos-Prieto et al. 2015)	*Mediterranean Lifestyle* (MEDLIFE index)	28 (21 grupos de alimentos/alimentos, 1 nutriente, 6 factores de estilo de vida)	0-1	0-28

[1]Basada en Hernández-Ruiz et al 2015

Algunos de estos índices además han sido desarrollados para adaptar la evaluación de la adherencia a la DietMed a diferentes poblaciones, países y grupos de edad. Por otra lado, las diferencias entre los distintos índices reportados se relacionan principalmente con los componentes y el sistema de puntuación empleado en cada uno, dando lugar a una gran diversidad de índices respecto al número y tipos de componentes (alimentos, grupos de alimentos, nutrientes y/o factores de estilo de vida), los criterios de puntuación y los puntos de corte utilizados para la puntuación (principalmente medianas o tertiles) (Hernandez-Ruiz et al. 2015). La mayoría de estos índices se han empleado para asociar la DietMed con múltiples efectos beneficios para la salud y prevención de enfermedades crónicas (Tabla 2.3). Por otro lado, sin embargo, la mayoría de los estudios difieren en método del establecimiento de los niveles de adherencia a la DietMed. El método más común, por ejemplo, se basa en la comparación entre terciles o quintiles (más alto vs más bajo) en base a la puntuación continua total obtenida por cada individuo en un índice determinado. Sin embargo, sumada a la variedad de índices que existen, los métodos de estratificación del nivel de adherencia a la DietMed siguen siendo objeto de debate entre la comunidad científica debido a las complicaciones que supone trasladar o replicar estas aproximaciones en distintas poblaciones. Además, debido a la complejidad de la DietMed, en términos de la cantidad y frecuencia con la que se consumen los alimentos que la componen, la evaluación de su adherencia a través de cualquiera de los índices tradicionales puede resultar inexacta.

Tabla 2.3 Relación del uso de índices de adherencia a la dieta Mediterránea en el estudio del riesgo y prevención de la mortalidad y enfermedades crónicas[1]

Índice de adherencia a la DietMed[2]	Efectos beneficiosos en la salud	Referencias
MDS-1995	Prevención de obesidad	(Schroder et al. 2006)
	Prevención de cáncer	(Praud et al. 2014)
MDS-2003	Mejorara del perfil cardiometabólico	(Cuenca-Garcia et al. 2014)
	Mejora de la función hepática y renal	(Alkerwi et al. 2015)
	Mejora de la función cognitiva	(Ye et al. 2013)
MMD	Disminución del riesgo de ECV	(Hoevenaar-Blom et al. 2012)
	Disminución del riesgo de cáncer	(Couto et al. 2011)
	Disminución de la hipertensión	(Psaltopoulou et al. 2004)
MDP-2006	Prevención de obesidad	(Schroder et al. 2007)
MDQI	Prevención de ECV	(Najafi & Sheikhvatan 2013)
	Prevención de diabetes	(Mosharraf et al. 2013)
MAI	Reducción de la presión arterial	(di Giuseppe et al. 2008)
aMED	Bajo riesgo de mortalidad	(Harmon et al. 2015)
	Bajo riesgo de cáncer	(Reedy et al. 2014)
	Bajo riesgo de diabetes	(Jacobs et al. 2015)
	Bajo riesgo de fracturas de cadera	(Zeng et al. 2014)
ITALIAN-MED	Prevención de cáncer de colon	(Agnoli et al. 2013)
rMED	Disminución del riesgo de ECV	(Buckland et al. 2009)
	Disminución del riesgo de mortalidad global y mortalidad por ECV	(Buckland et al. 2011)
	Bajo riesgo de cáncer gástrico	(Buckland et al. 2010),
	Bajo riesgo de obesidad	(Romaguera et al. 2010)
	Bajo riesgo de diabetes	(Romaguera et al. 2011)

Tabla 2.3 Continuación

MeDiet-PREDIMED	Bajo riesgo de hiperuricemia	(Guasch-Ferre et al. 2013)
	Bajo riesgo de psoriasis	(Barrea et al. 2015)
	Bajo riesgo de obesidad	(Martinez-González et al. 2012)

[1]Basada en Hernández-Ruiz et al 2015; 2la definición de los acrónimos de los índices se mencionan la tabla 2.2

Esto hace necesario, por lo tanto, el empleo de otros indicadores complementarios o más precisos que la información obtenida a través de un índice de adherencia (Maj-Britt S Andersen et al. 2014). En este sentido, los biomarcadores obtenidos a través de aproximaciones metabolómicas han emergido como una opción novedosa, complementaria y más precisa en la evaluación del seguimiento de patrones alimentarios como la DietMed (Maj-Britt S Andersen et al. 2014; Playdon et al. 2017). Este tipo de biomarcadores, por lo tanto, podría aportar mayor precisión en la evaluación de la adherencia a la DietMed, reduciendo por un lado la impresición de la información obtenida a través de los índices empleados hasta ahora, y posibilitando su extrapolación a otras poblaciones. Finalmente, los biomarcadores de adherencia a la DietMed podrían ser aprovechados en el estudio de asociaciones entre el consumo de este patrón dietético y la prevención y/o tratamiento de enfermedades crónicas, incluidos los factores de riesgo cardiovascular.

2.4 APLICACIONES DE LA METABOLÓMICA EN EL ESTUDIO DE BIOMARCADORES DE RIESGO CARDIOVASCULAR Y ESTILOS DE VIDA CARDIOSALUDABLE

2.4.1 Biomarcadores de Diabetes Mellitus 2

La DM2 es una enfermedad causada por un conjunto complejo de interacciones entre factores genéticos y ambientales, y cuya prevalencia está aumentando en todo el mundo (Bain et al. 2009). Esto hace que la identificación de biomarcadores para su detección y estadificación, así como la implementación de estrategias eficaces para su prevención sean objetivos científicos y médicos especialmente críticos (Gonzalez-Franquesa et al. 2016). En este sentido, aunque los niveles altos de glucosa en sangre siguen siendo uno de los primeros biomarcadores de DM2, las alteraciones metabólicas asociadas a la DM2 no se limitan sólo al metabolismo de los carbohidratos, sino también al de los lípidos, proteínas e incluso la microbiota. Esta complejidad presenta desafíos para una comprensión completa de los procesos moleculares que contribuyen al desarrollo de esta grave enfermedad (Klein & Shearer 2016). Afortunadamente, con los avances en técnicas analíticas de alto rendimiento y la aplicación de enfoques bioinformáticos y de modelación estadística, a través de la metabolómica, en los últimos años se ha podido describir mejor el metaboloma asociados a la DM2, y a diferentes estadios que pueden ser previos o posteriores al diagnóstico (Guasch-Ferre et al. 2016; Gonzalez-Franquesa et al. 2016).

La información respecto a la aplicación de la metabolómica en el estudio de biomarcadores de DM2 en esta Tesis Doctoral se presenta en formato de revisión bibliográfica, la cual ha sido publicada en una revista científica internacional.

Mireia Urpi-Sarda, Enrique Almanza-Aguilera, Sara Tulipani, Francisco J. Tinahones, Jordi Salas-Salvadó and Cristina Andres-Lacueva. Metabolomics for Biomarkers of Type 2 Diabetes Mellitus: Advances and Nutritional Intervention Trends. *Current Cardiovascular Risk Reports*. 2015; 9:12. DOI 10.1007/s12170-015-0440-y.

En la presente revisión se resumen distintas clases de metabolitos que han sido propuestos, a través de numerosos estudios metabolómicos, como potenciales biomarcadores de DM2 y de estados previos como la intolerancia a la glucosa y resistencia a la insulina. Se proporciona además una visión global de las tendencias en las intervenciones dietéticas dirigidas a reducir el riesgo y mejorar el control de la DM2.

Curr Cardiovasc Risk Rep (2015) 9: 12
DOI 10.1007/s12170-015-0440-y

LIPIDS (L PARNELL AND J ORDOVAS, SECTION EDITORS)

Metabolomics for Biomarkers of Type 2 Diabetes Mellitus: Advances and Nutritional Intervention Trends

Mireia Urpi-Sarda · Enrique Almanza-Aguilera · Sara Tulipani · Francisco J. Tinahones · Jordi Salas-Salvadó · Cristina Andres-Lacueva

Published online: 17 February 2015
© Springer Science+Business Media New York 2015

Abstract Metabolic characterization of type 2 diabetes mellitus (T2DM) is crucial for the identification of individuals at risk for developing diabetes and T2DM-related vascular complications as well as for monitoring disease progression. The application of metabolomics to diabetes research may lead to the identification and discovery of diagnostic and prognostic T2DM biomarkers, in addition to elucidating disease pathways. In the present review, we summarize the distinct classes of metabolites that have been proposed as potential biomarkers for progressing stages of T2DM by metabolomic approaches. Several studies have demonstrated that the metabolism of carbohydrates, lipids, and amino acids is considerably altered in prediabetes and continue to vary over the course of T2DM progression. The identification of intermediate metabolites involved in glycolysis, gluconeogenesis, the tricarboxylic acid cycle, lipolysis, and proteolysis have provided evidence of these metabolic dysfunctions. Finally, given the increasing worldwide incidence of T2DM and its related complications, research should focus on the impact of lifestyle factors, particularly diet, at the metabolomic level for better understanding and improved healthcare strategies.

Keywords Type 2 diabetes mellitus · Metabolomics · Diabetes research · Biomarkers · Pathways · Progression states · Lifestyle factors · Diet

This article is part of the Topical Collection on *Lipids*

Electronic supplementary material The online version of this article (doi:10.1007/s12170-015-0440-y) contains supplementary material, which is available to authorized users.

M. Urpi-Sarda (✉) · E. Almanza-Aguilera · S. Tulipani · C. Andres-Lacueva
Biomarkers and Nutritional and Food Metabolomics Research Group, Nutrition and Food Science Department, XaRTA, INSA, Pharmacy Faculty, University of Barcelona, Avenida Joan XXIII s/n, 08028 Barcelona, Spain
e-mail: murpi@ub.edu

S. Tulipani
Biomedical Research Institute (IBIMA), Service of Endocrinology and Nutrition, Hospital Complex (Virgen de la Victoria), Campus de Teatinos s/n, University of Málaga, 29010 Málaga, Spain

F. J. Tinahones · J. Salas-Salvadó
CIBER Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Instituto de Salud Carlos III, Madrid, Spain

F. J. Tinahones
Endocrinology and Nutrition Service, Virgen de la Victoria Clinical Hospital, Málaga, Spain

J. Salas-Salvadó
Human Nutrition Unit, Hospital Universitari de Sant Joan de Reus, Institut d'Investigació Sanitària Pere Virgili (IISPV), Universitat Rovira i Virgili, Reus, Spain

Introduction

In addition to obesity, T2DM is a major risk factor for cardiovascular disease (CVD). Hyperglycemia and insulin resistance (IR) are powerful predictors of adverse cardiovascular events, and these two risk factors in combination exert a detrimental synergistic effect [1]. However, despite a number of recognized risk factors including family history of diabetes, age, sex, and numerous anthropometric, biochemical, socioeconomic, and lifestyle variables, the identification of individuals with increased risk of T2DM and/or CVD remains a laborious task [2]. Typical biomarkers in T2DM prediction models include elevated concentrations of fasting plasma glucose and insulin, glycated hemoglobin (HbA_{1c}), serum ferritin, C-reactive protein (CRP), and interleukin-2 receptor alpha (IL2RA); these models also predict a decrease in effective serum insulin and adiponectin concentrations [3, 4]. Recent

advances in "omics" such as genomics and metabolomic technologies are generating an increasing number of potential biomarkers to identify crucial stages in the pathogenesis and progression of a determined pathological state, including T2DM. In addition, they are also gaining insights in the underlying molecular disease-causing mechanisms through the discovery of associations between some genetic variants and key small molecules. These relationships highlight the power of integrating multiomic approaches (Systeomics) to better understand the causal mechanisms [5]. For instance, recently, Wang et al. [6] delineated the role of fatty acid desaturases (FADs) in regulating human liver lipid composition through a targeted lipidomic analysis and associated them with FADS single-nucleotide polymorphisms (SNPs) from genome-wide association studies (GWASs). They suggested that FADS1 and its polymorphisms were related with long-chain fatty acid accumulation in human liver [6]. Metabolomics is defined as a comprehensive characterization of endogenous or exogenous metabolites representing the metabolome [7]; the use of this technology enables the detection of physiological or pathological changes in cells, tissues, or body fluids and represents a useful tool for biomarker detection [8, 9]. In diabetes research, metabolomics has been successfully applied to diagnostic and prognostic biomarker discovery, elucidation of disease pathways, identification of drug side effects, and discovery of functional biomarkers for drug activity [10]. The present review aims to summarize the distinct family of metabolites that have been proposed as potential biomarkers of different stages of T2DM by metabolomic approaches. Additionally, the impact of diet, as an important lifestyle factor, on classical and metabolomic biomarkers will be reviewed for better understanding the pathophysiology of diabetes, aiming to implement healthcare strategies in the future.

Metabolomic Biomarkers in Prediabetes

The term prediabetes refers to the impaired glucose tolerance (IGT) and/or impaired fasting glucose (IFG) of subjects with a relatively high risk of developing diabetes. IFG is characterized by fasting plasma glucose levels between 100 mg/dl (5.6 mmol/l) and 126 mg/dl (7.0 mmol/l). IGT is defined by 2-h plasma glucose values after an oral glucose tolerance test (OGTT); values between 140 mg/dl (7.8 mmol/l) and 200 mg/dl (11.1 mmol/l) are considered indicative of IGT [11]. In addition to IFG and IGT, IR is also considered a crucial metabolic status because it can precede the dysglycemic states of prediabetes and T2DM [12]. Because prediabetic stages are asymptomatic, extended time periods may elapse before diagnosis of T2DM, hampering early detection. In addition to the most common test used to assess impaired insulin sensitivity (IS), homeostasis model assessment of insulin resistance (HOMA-IR), OGTTs are also indirectly used to assess insulin resistance.

Metabolomics may be extremely helpful in the identification of novel biomarkers of prediabetes and metabolic disturbances that precede the new onset of T2DM. In this regard, a targeted metabolomic approach has shown that IR emerges in insulin-dependent processes, such as proteolysis, lipolysis, ketogenesis, and glycolysis, in addition to the reduction in glucose uptake and suppression of gluconeogenesis, thus, reflecting a broad switch from catabolism to anabolism [13].

The understanding of the role of dyslipidemia in prediabetes has progressed significantly with the implementation of lipidomics, a new branch of metabolomics [14–16]. The current and advanced analytical techniques used in lipidomics such as chromatography coupled to mass spectrometry (MS) or nuclear magnetic resonance (NMR) as well as other spectroscopic approaches are powerful techniques used in lipidomics for lipid detection and characterization [17]. They allow detection and characterization up to several hundreds of lipids belonging to major lipid classes (i.e., fatty acyls, phospholipids, glycerolipids, glycerophospholipids, sphingolipids, sterol lipids) [17]. As an example of the high throughput obtained by these technologies, a recent lipidomic study reported over 500 different lipid molecular species among the main lipid classes in plasma of individuals [18]. The importance of lipoprotein fatty acid composition and its role in IS has been emphasized in a study using lipidomic techniques, including ultra-performance liquid chromatography coupled to mass spectrometry (UPLC/MS). In this study, the degree of fatty acid saturation in triacylglycerols (TAG) within the VLDL-IDL-LDL axis and HDL were differentially related to IR [15]. Specifically, serum TAG molecules, such as 16:0/16:0/18:1 and 16:0/18:1/18:0, correlated positively with HOMA-IR; however, TAG containing essential fatty acids, such as 18:1/18:2/18:2, correlated negatively. The findings of this study reinforce the role that fatty acids may have in the pathogenesis of IR; therefore, the serum fatty acid composition may be considered a more precise marker of insulin resistance than total serum TAG concentrations [15]. Consistent with this hypothesis, in the Framingham Heart Study (FHS), serum TAG characterized by relatively low carbon number and double bond content (i.e., C46:1, 48:1) were positively associated with HOMA-IR; conversely, TAG with increased carbon number and double bond content (i.e., C56:9, C58:10) were not correlated with HOMA-IR [16]. These results were consistent even after the participants were grouped in quartiles according to HOMA-IR [16].

Zhao et al. [19] also observed a characteristic lipid profile for individuals with IGT as a prediabetic condition. Applying untargeted metabolomics using UPLC-qTOF-MS, the authors reported increased plasma levels of free fatty acids (FFA) (i.e., C16:0, C18:0, C18:1) and glycochenodeoxycholic acid as well as decreased concentrations of lysophosphatidylcholines

derived from OGTT. In the follow-up of both studies, α-HB was a positive predictor and L-GPC a negative predictor of dysglycemia (RISC study) or T2DM (Botnia Study), independently of the family history of diabetes, sex, age, BMI, and fasting plasma glucose [27]. Recently, a novel branched-chain ketoacid derivative of isoleucine, called 3-methyl-2-oxovalerate, was found to be significantly associated to IFG, both in the plasma and urine of individuals with prediabetes relative to control individuals with diabetes [29]. The observation that 3-methyl-2-oxovalerate was the second strongest predictive biomarker for IFG, after glucose, was first reported in a cohort of 2204 females from Twins, UK, and was subsequently replicated in individuals with IFG (n=536) in the follow-up study KORA F4 [29].

Other metabolites that are not directly linked with major metabolic pathways have been significantly associated with prediabetes. For example, decreased urinary levels of gut microbiota-associated metabolite biomarkers, including hippuric acid, methylxanthine, methyluric acid, and 3-hydroxyhippuric acid, have been linked to IGT [19].

Metabolomics Biomarkers and Pathways Altered in T2DM

Because T2DM triggers multiple metabolic disorders, efforts have been made to elucidate the mechanisms causing these disorders and systemic complications. Metabolomics has been successfully applied in T2DM research to elucidate novel metabolic pathways as well as to define relationships between significant metabolites in these pathways. Nowadays, the knowledge of interactions between affected T2DM pathways is improving through building biological pathways and network analysis techniques which integrate data from the different "omics" [30, 31]. To obtain a more comprehensive analysis of the metabolic processes negatively regulated by T2DM, a summary of metabolomic studies and the resulting metabolites significantly altered in diabetics is shown in Table 1. Furthermore, Fig. 1 illustrates a summary of metabolic networks presumably affected by T2DM. To this end, we used the MetaCore™ software (Thomson Reuters) from GeneGo, uploading the complete list of T2DM biomarkers reported in literature (Table 1).

Metabolomic studies evaluating individuals with T2DM confirm that metabolic networks of primary biomolecules, such as carbohydrates, lipids, and amino acids, are altered as a consequence of the diabetic stage (Fig. 1). Hyperglycemia and glycosuria are the major biomarkers of uncontrolled T2DM [29, 54]. However, abnormal levels of other metabolites reflect dysregulation of carbohydrate metabolism (i.e., fructose, mannose) (Fig. 1). Impairment of glycolysis and gluconeogenesis has been demonstrated by metabolomic approaches through the identification of metabolites included in these pathways: glycerol-3-phosphate, phosphoenolpyruvate, pyruvate, and lactate. Additionally, downstream tricarboxylic acid cycle (TCA) metabolites, such as citrate, 2-oxoglutarate, succinate, fumarate, and malate, are deregulated in diabetes (Table 1). Controversial results suggest that levels of circulating and urinary glucogenic amino acids [29, 33, 36, 39, 41, 45, 46, 52, 55] in diabetic subjects indicate deregulation of glucose biosynthesis (Table 1). Furthermore, significant increases in three ketone bodies, acetone, acetoacetate, and β-hydroxybutyrate, in plasma [32, 41] and urine [34], reflect a reduction in glucose uptake and the onset of ketosis in T2DM [56].

A considerable number of metabolomic studies have reported a positive association between abnormal circulating concentrations of lipid derivatives and T2DM progression. Although not necessarily consistent with prediabetes with respect to saturation, higher concentrations of long-chain (i.e., oleic, palmitic) and lower concentrations of medium-chain FFAs (i.e., caproate, pelargonate, 10-undecenoate) are a characteristic lipid signature among individuals with T2DM [41, 43, 45, 57]. Using GC-MS analysis, Han et al. demonstrated that primarily esterified fatty acids (EFA) are decreased in patients with T2DM, while non-esterified fatty acids (NEFA) are increased. This occurs even when including the variability of groups with different stages of diabetic nephropathy, suggesting a combination of lipotoxicity and toxicology repair mechanism [47]. Applying a lipidomic approach, Ståhlman et al. characterized the lipid composition of ApoB-containing lipoproteins isolated from control, normolipidemic, and dyslipidemic individuals with T2DM [58]. Significant increases in PC 16:0-20:3 (in VLDL and LDL) and PC 18:0-20:3 (in LDL) were detected in normolipidemic T2DM compared with control individuals. These alterations were more pronounced in the dyslipidemic T2DM group, which also had a relatively increased amount of PC 16:0-16:1. Similarly, significant increases in CE 16:1 (in VLDL and LDL) and CE 20:3 (in LDL) were detected in lipoproteins from dyslipidemic T2DM participants. Furthermore, levels of palmitic acid (C16:0) in VLDL and LDL TAG correlated positively with IR [58]. Abnormal circulating levels of distinct subclasses of phospholipids, including PC, lysoPC, phosphatidylinositol (PI), PE, lysoPE, SM, PG, and sphingosine-1-phosphate (Table 1), have also been identified by metabolomic approaches and were considered potential biomarkers of the diabetic dyslipidemia [48, 49, 53]. Ceramides, another important class of bioactive lipids, have recently garnered attention due their pathophysiological relevance in the development of IR and impaired glycemic control [59]; therefore, ceramide concentrations are significantly deregulated in T2DM [37].

A research focused on acylcarnitines (AcylCN) and their byproducts has generated insights into the dysregulation of fatty acid oxidation associated with T2DM. Variations in the levels of AcylCN, mostly from short to medium chains, have been detected by applying targeted metabolomic analyses to

Table 1 Metabolomic studies showing metabolites significantly associated to T2DM

Study design	Analytic technique	Biofluid	Metabolites significantly affected by the disease*	Reference
T2DM subjects (*n*=33) Healthy subjects (*n*=20)	^{1}H NMR	Urine	(↑): Lactate, alanine, citrate, DMA, TMAO, hippurate, glycine, creatine, acetate, betaine, acetone, acetoacetate, β-hydroxybutyrate	[32]
T2DM subjects (*n*=11) Healthy subjects (*n*=16)	^{1}H NMR	Plasma	(↑): Lactate (↓): Leucine, isoleucine, valine, β-hydroxybutyrate, alanine, glutamine, citrate, tyrosine, formate	[33]
		Urine	(↑): Alanine, citrate, phenylalanine, tyrosine, hippurate, phospho(enol)pyruvate (↓): Glutamate, glutamine, *N*-methylnicotinamide, uridine	[33]
T2DM subjects (*n*=30) Healthy subjects (*n*=12)	^{1}H NMR	Urine	(↑): TMA, DMA, DMG, betaine, TMAO, citrate, acetate, acetoacetate, butyrate, 2-hydroxybutyrate, β-hydroxybutyrate, alanine, glutamine, ornithine, taurine, *N*-methylnicotinamide, *N*-methyl-2-pyridone-5-carboxamide (↓): Creatine, creatinine, malate, fumarate, succinate, 2-oxoglutarate, leucine, isoleucine, histidine, tryptophan, allantoin, *N*-methylnicotinate	[34]
T2DM subjects (*n*=28) Healthy subjects (*n*=26)	GC-MS	Urine	(↑): 4-Aminobenzoic acid (↓): Maleic acid dimethyl ester, oxyl acetic acid, 2,5-bisoxy-benzeneacetic acid	[35]
T2DM subjects (*n*=82) Healthy subjects (*n*=36)	GC-MS	Serum	(↑): Butyrate, valine, glutamate, palmitate (C16:0), urate, oleate (C18:1), stearate (C18:0), arachidonate, maltose, octadecanoate (↓): Lactate, lysine, glucuronolactone	[36]
Obese subjects with T2DM (*n*=13) Healthy subjects (*n*=14)	GC-MS	Plasma	(↑): Ceramides: C18 (*N*-stearoylsphingosine), C20 (*N*-eicosanoylsphingosine), C24:1 Ceramide	[37]
T2DM subjects (*n*=48) Healthy subjects (*n*=31)	GC×GC-TOF-MS	Plasma	(↑): Glucose, 2-hydroxyisobutyric acid	[38]
T2DM subjects (*n*=74) NGT subjects (*n*=80)	^{1}H NMR	Serum	(↑): Glucose (↓): Isoleucine, leucine, valine, alanine, methionine, glutamine, citrate, lysine, choline, lactate, tyrosine, phenylalanine, histidine	[39]
T2DM subjects (*n*=33) Healthy subjects (*n*=25)	UPLC-oaTOF-MS	Serum	(↓): Phytosphingosine, dihydrosphingosine, leucine	[40]
T2DM subjects (*n*=26) Healthy subjects (*n*=26)	GC-MS	Plasma	(↑): Lactate, alanine, 2-hydroxyisobutyric acid, β-hydroxybutyric acid, phosphate, leucine, isoleucine, serine, pyroglutamic acid, palmitic acid, oleic acid, stearic acid, arachidonic acid, 1-monopalmitin, 1-monostearin (↓): 2-ketoisocaproic acid	[41]
Obese subjects with T2DM (*n*=44) Obese subjects without T2DM (*n*=12)	HPLC-MS	Plasma	(↑): AcylCN: total-free, C2 (acetylcarnitine), C6 (hexanoyl carnitine), C8 (octanoyl carnitine), *cis*-3, 4-methylene-nonanoyl carnitine, C14 (myristoyl carnitine), C18:1 (oleoyl carnitine), C8-dicarb (suberoyl carnitine), summed C10-C14 AcylCN, total AcylCN (↓): AcylCN: C3 (propionyl carnitine)	[42]
Obese subjects with T2DM (*n*=44) Obese subjects without T2DM (*n*=12)	GC-TOF-MS	Plasma	(↑): β-hydroxybutyrate, oleic acid, gluconic acid, fructose, palmitoleic acid, 3,6-anhydrogalactose, glucuronic acid, glucose, heptadecanoic acid, inulobiose, leucine, 2-hydroxybutyrate, 2-deoxyerythritol, palmitic acid, 2-ketoisocaproic acid, uridine, cysteine, xylose, histidine, stearic acid (↓): Benzylalcohol, benzoic acid, lysine, ethanolamine, arachidonic acid, glycine, glycerol-3-phosphate	[43]
T2DM subjects (*n*=10) Lean subjects without T2DM (*n*=12) Obese subjects without T2DM (*n*=14)	HPLC-ESI-MS/MS	Plasma	(↑): AcylCN: C3 (propionyl carnitine), C5 (isovaleryl carnitine), C8 (octanoyl carnitine), C4-OH (3-hydroxy-butyryl carnitine), C5-OH (3-hydroxy-isovaleryl carnitine), C6-OH (hydroxyhexanoyl carnitine)	[44]
T2DM subjects (*n*=40) Healthy subjects (*n*=60)	UHPLC-MS/MS2 GC-MS NMR	Serum and plasma	(↑): Desoxyhexose, glucose, glycolipids (H3-HNAc2-NANA, HNAC, HNAc-H2-dH), uronic acid, dihexose, mannose, creatinine, glutamylvaline, gamma-glutamylisoleucine, β-hydroxybutyrate, PAGN, phenylalanine,	[45]

Table 1 (continued)

Study design	Analytic technique	Biofluid	Metabolites significantly affected by the disease*	Reference
			3-indoxyl sulfate, kynurenine, homocitrulline, myristate, palmitate, 2-hydroxypalmitate, margarate, 10-heptadecenoate, stearate, 2-hydroxystearate, oleate, linoleate, linoleamide, linolenate, eicosenoate, dihomo-alpha-linolenate, adrenate, isoleucine, leucine, gamma-glutamylleucine, valine (↓): 1,5-anhydroglucitol, caproate, heptanoate, pelargonate, glycerophosphorylcholine, PC a C20:4, PC aa (OH, COOH) C28:4, PC aa C34:4, SM C14:0, SM C22:2, 10-undecenoate, arachidonate	
T2DM subjects (n=18) Healthy subjects (n=19)	^{1}H NMR	Serum	(↑): Fumarate, methylguanidine, pyruvate, glucose, DMA, methylamine, mannose, TMAO, uridine (↓): Taurine, pyroglutamate, threonine, phenylalanine, serine, glycine	[46]
T2DM subjects (n=30) Healthy subjects (n=30)	GC-MS	Plasma	(↑): NEFA (C16:0, C18:2, C18:1, C18:0, C20:4, C20:3, C20:2, C20:0, C22:6) (↓): NEFA (C10:0) EFA (C10:0, C14:0, C16:0, C16:1n-9, C18:2, C18:1, C18:0, C20:4, C20:5, C20:3, C20:2, C22:6)	[47]
T2DM subjects (n=30) Healthy subjects (n=30)	LC-TOF-MS	Plasma	(↑): LysoPC (C18:2) (↓): PC (C18:0/20:4), PI (C18:0/20:4)	[48]
T2DM subjects (n=26) Healthy subjects (n=27)	UPLC/Q-TOF-MS	Plasma	(↑): Dodecanoic acid, myristic acid, leucine, lysine, phenylalanine, propionyl carnitine, octanoyl carnitine, decanoyl carnitine, dodecanoyl carnitine, palmityl carnitine, heptadecanoyl carnitine, linoleyl carnitine, vaccenyl carnitine, lysoPC (14:0, 16:1, 18:1, 18:3, 20:5, 22:6), lysoPE (18:2, 22:6) (↓): serine, lysoPE (18:1)	[49]
T2DM subjects (n=9) Lean subjects without T2DM (n=39) Obese subjects without T2DM (n=64)	HPLC-ESI-MS/MS	Plasma	T2DM vs other groups: (↑): 3-hydroxyoleoylcarnitine (↓): AcylCN (C2, C6) leucine, isoleucine, valine, phenylalanine, methionine, alanine	[50]
T2DM subjects (n=60) Healthy subjects (n=25)	HPLC	Plasma	(↑): Myristic acid, palmitic acid, palmitoleic acid, stearic acid, oleic acid, linoleic acid, linolenic acid, γ-linolenic acid, eicosadienoic acid, arachidonic acid, eicosapentaenoic acid, docosahexaenoic acid	[51]
T2DM subjects (obese n=31, lean=95) Healthy subjects (obese n=80, lean n=20)	LC-MS/MS	Plasma	(↑): Hydroxyproline, glutamine, ethanolamine, citrulline, sarcosine, β-alanine, glutamate, 3-methyl histidine, γ-aminobutyric acid, β-aminoisobutyric acid, proline (↓): Phosphoserine, phosphoric acid ethanolamine, taurine, serine, aspartate, histidine, 1-methylhistidine, arginosuccinic acid, carnosine, anserine, α-aminoadipic acid, δ-hydroxylysine, lysine, homocysteine, leucine, tryptophan	[52]
T2DM subjects (n=105) Healthy subjects (n=77)	UPLC-ESI-QTOF-MS	Plasma	(↑): Itaconic acid, leucine, PC (18:0/0:0), sphingosine-1-phosphate, PG (18:0/18:1) (↓): Inosine, uric acid, 3-hydroxymethylglutaric acid, succinic acid, taurine, PE (P-16:0/22:6)	[53]
		Urine	(↑): *N*-acetyl-D-phenylalanine,serotonin (↓): 2-ketoglutaric acid, 2-ketobutyric acid, 1-methylhistidine, kynurenic acid, xanthurenic acid, pyruvic acid	
T2DM subjects (n=115) Healthy subjects (n=1897)	UPLC-MS/MS	Plasma	(↑): 2-hydroxybutyrate, proline, 3-methyl-2-oxobutyrate, 3-methyl-2-oxovalerate, 4-methyl-2-oxopentanoate, isoleucine, leucine, valine, fructose, mannose, glucose, lactate, arabinose, malate, erythrytol (↓): *N*-acetylglycine, citrulline, dimethylarginine (SDMA + ADMA), 1,5-anhydroglucitol, octanoylcarnitine, 15-methylpalmitate, 10-heptadecenoate, myristate, myristoleate, palmitoleate, pentadecanoate, 5-dodecenoate, heptanoate, pelargonate, palmitoyl sphingomyelin, cholesterol	[29]
T2DM subjects (n=81)	UPLC-QTOF-HDMS	Urine	(↑): Acylcarnitines, 3-indoxylsulfate, glucose, glycine	[54]

Table 1 (continued)

Study design	Analytic technique	Biofluid	Metabolites significantly affected by the disease*	Reference
Healthy subjects (n=42) T2DM subjects (n=35) Healthy subjects (n=35)	^{1}H NMR	Plasma	(↓): Citric acid, kynurenic acid, urate, glucuronolactone, lysine, phosphate (↑): α-glucose, β-glucose (↓): Isoleucine, leucine, valine, lactate, alanine, glutamate, creatine, creatinine, myo-inositol, scyllo-inositol, choline, tyrosine, phenylalanine, 1-methylhistidine	[55]

*Increase (↑) or decrease (↓) of metabolites in the same line

^{1} *H NMR* proton nuclear magnetic resonance, *AcylCN* acyl carnitine, *ADMA* asymmetric dimethylarginine, *DMA* dimethylamine, *DMG* dimethylglycine, *EFA* esterified fatty acids, *GC× GC-TOF-MS* two-dimensional gas chromatography-time-of-flight mass spectrometry, GC-MS gas chromatography-mass spectrometry, *GC-TOF-MS* gas chromatography-time-of-flight mass spectrometry, *HPLC* high-performance liquid chromatography, *HPLC-ESI-MS/MS* high-performance liquid chromatography-electrospray ionization with tandem mass spectrometry, *LC-MS/MS* liquid chromatography with tandem mass spectrometry, *LC-TOF-MS* liquid chromatography-time-of-flight mass spectrometry, *lysoPC* lysophosphatidylcholine, *lysoPE* lysophosphatidylethanolamine, *NEFA* non-esterified fatty acids, *NGT* normal glucose tolerance, *PC* phosphatidylcholine, *PG* phosphatidylglycerol, *PI* phosphatidylinositol, *SDMA* symmetric dimethylarginine, *SM* sphingomyelin, *T2DM* type 2 diabetes mellitus, *TMA* trimethylamine, *TMAO* trimethylamine *N*-oxide, *UHPLC/MS/MS2* ultra-high-performance liquid chromatography/tandem mass spectrometry, *UPLC/Q-TOF-MS* ultra-performance liquid chromatography-quadrupole/time-of-flight-mass spectrometry, *UPLC-ESI-QTOF-MS* ultra-performance liquid chromatography-electrospray ionization-quadrupole-time-of-flight-mass spectrometry, *UPLC-MS/MS* ultra-performance liquid chromatography with tandem mass spectrometry, *UPLC-oaTOF-MS* ultra-performance liquid chromatography coupled with orthogonal acceleration time-of-flight mass spectrometry, *UPLC-QTOF-HDMS* ultra-performance liquid chromatography-quadrupole time-of-flight high-definition mass spectrometry

the plasma of diabetic individuals [42, 44, 50]. In contrast to previous studies carried out with adult subjects, a targeted metabolomic study conducted by Mihalik et al. showed that the plasma levels of C2 and C6, AcylCN metabolites were lower in 17 adolescents (~13–15 years) with T2DM compared to those with normal weight (n=39) or obese (n=64). Additionally, short- and medium-chain AcylCN metabolites (C3-, C4-, C5-, C10:1) and free carnitine concentrations were lower in diabetic adolescents compared to normal weight subjects [50]. Unlike adults, obese adolescents with or without T2DM do not demonstrate impaired fatty acid or amino acid metabolism. These contrasting results between adolescents and adults could be reflective of the duration of obesity; chronic obesity may result in a gradual progression in the failure of mitochondrial adaptive mechanisms [50].

Amino acids and their derived metabolites have become the third group of endogenous biomolecules heavily studied in metabolomic analyses (Table 1 and Fig. 1). Many amino acids are involved in other metabolic process impaired by T2DM, such as gluconeogenesis. Several intermediary molecules involved in amino acid metabolism have been identified and proposed as T2DM biomarkers, along with other metabolites. The list of amino acid derivatives includes precursors for biosynthesis (i.e., TCA intermediates) and catabolism products (i.e., 2-hydroxybutyrate, 3-methyl-2-oxovalerate, 2-oxobutyrate), serving later as precursors of other molecules. Furthermore, protein catabolism, possibly leading to muscle atrophy, has been evidenced by the identification of reduced circulation and/or excretion of compounds such as creatine, creatinine, 3-methylhistidine, methylguanidine, and sarcosine among others (Table 1).

Untargeted metabolomic approaches increasingly allow for the discovery of small molecules beyond those naturally endogenous metabolites that are associated to T2DM risk. Among these uncommon metabolites are choline byproducts (i.e., DMA, TMA, DMG, TMAO, betaine), sugar acids (i.e., uronic acid), purine metabolites (i.e., allantoin, inosine), nucleotide sugars (i.e., arabinose), vitamin derivatives (i.e., *N*-methylnicotinamide), as well as products of microbial metabolism (i.e., phenylacetylglutamine, 3,6-anhydrogalactose) (Table 1). However, further research and confirmation through several metabolomic studies (considering a targeted approach) will be required to validate these metabolites as T2DM biomarkers.

Metabolomic Biomarkers as Predictors of T2DM Development

Biomarkers that predict future T2DM are useful to identify those individuals at high risk of developing diabetes, and these allow to focus public health preventive strategies in this population. Given the metabolic heterogeneity associated with

Fig. 1 Summary of metabolic networks affected in T2DM according to the presence of metabolites in urine (*top orange bar*) and serum/plasma (*lower blue bar*)*. *Networks listed were obtained by an Enrichment Analysis in MetaCore™ (Genego, St. Joseph, MI) and ordered according to the major metabolic pathways involved. Figure includes metabolites listed in Table 1 separated by urine and serum/plasma and their direction. A figure with the full list of metabolic networks resulting from MetaCore™ is available in the supplementary material

T2DM before its onset, there is great demand for reliable, predictive biomarkers. Targeted metabolomic studies have increasingly aided in the development of novel biomarkers in large prospective studies [16, 60••, 61, 62]. Consistent with previous observations in individuals with IR, Rhee et al. observed a characteristic association between carbon number and bond content that was predictive of developing T2DM. Specifically, TAG with a lower carbon number and double bond content were associated with an increased odds ratio (OR) for diabetic subjects, while TAG with a higher carbon number and double bond content were associated with an OR of less than one. Moreover, the inverse relationship between diabetes risk, carbon number, and double bond content persisted after multivariable adjustment for lysoPC, PC, and possibly lysophosphatidylethanolamines (lysoPE), but not for cholesterol esters (CE) [16]. In another but complementary study, Wang et al. carried out two parallel and independent studies based on the same sample population. These authors discovered two novel metabolic signatures for the prediction of T2DM [60••, 62]. In the first study, higher levels in a panel of five amino acids (isoleucine, leucine, phenylalanine, tyrosine, and valine) showed a strong association with future development of diabetes. Moreover, a combination of three amino acids (isoleucine, tyrosine, and phenylalanine) was shown to be a better predictor of future diabetes than all five amino acids; individuals in the top quartile of this 3-amino acid score had a five- to sevenfold higher risk of developing new-onset diabetes compared with individuals in the lowest quartile [60••]. These results, with the exception of a non-significance of isoleucine, were replicated by the same authors in the Malmö Diet and Cancer (MDC) study [60••]. Later, Wang et al. reported that the odds of developing T2DM were increased fourfold for individuals in the higher quartile of plasma 2-aminoadipic acid (2-AAA) concentrations over the

current diabetes epidemic, future research should consider the relevance of novel biomarkers for the prediction and diagnosis of T2DM and the elucidation of disease pathways implicated in this disease. Metabolomic approaches have identified distinct classes of metabolites as potential biomarkers for different stages of T2DM. Several studies have demonstrated that the metabolism of carbohydrates, lipids, and amino acids are considerably altered in the prediabetic state and at different stages of T2DM progression. The identification of intermediate metabolites included in glycolysis, gluconeogenesis, the tricarboxylic acid cycle, lipolysis, and proteolysis have provided evidence for this metabolic dysfunction. Due to the scarcity of information on the effects of lifestyle changes on metabolomic biomarkers, more effort should be directed in expanding our knowledge to the metabolic modulations caused by dietary patterns in T2DM patients. Lifestyle interventions have a significant impact on diabetes prevention and control through modeling peripheral classical biomarkers of T2DM; therefore, future studies should aim to develop novel biomarkers that are sensitive to food challenge. This could establish appropriate dietary strategies to help improve the life course of T2DM patients.

Acknowledgments This study was supported by CICYT AGL2009-13906-C02-01 from the Spanish Ministerio de Economía y Competitividad (MINECO) and PI13/01172 Project, (Plan N de I + D + i 2013-2016) co-funded by ISCII-Subdirección General de Evaluación y Fomento de la Investigación and Fondo Europeo de Desarrollo Regional (FEDER). We also thank the award of 2014SGR1566 from the Generalitat de Catalunya's Agency AGAUR and the EU Joint Programming Initiative A Healthy Diet for a Healthy Life on Biomarkers BioNH-FOODBALL (PCIN-2014-133-MINECO-Spain). M.U.-S. would like to thank the "Ramón y Cajal" program (RYC-2011-09677) from MINE CO and the Fondo Social Europeo. EAA would like to thank to CONACYT (México) for the Ph.D. fellowship. ST acknowledges the Juan de la Cierva program (MINECO).

Compliance with Ethics Guidelines

Conflicts of Interest Cristina Andres-Lacueva, Francisco J Tinahones, Jordi Salas-Salvadó, Mireia Urpi-Sarda, Sara Tulipani, and Enrique Almanza-Aguilera have no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance,
•• Of major importance

1. Paneni F, Costantino S, Cosentino F. Insulin resistance, diabetes, and cardiovascular risk. Curr Atheroscler Rep. 2014;16:419.
2. Herder C, Karakas M, Koenig W. Biomarkers for the prediction of type 2 diabetes and cardiovascular disease. Clin Pharmacol Ther. 2011;90:52–66.
3. Kolberg JA, Jørgensen T, Gerwien RW, Hamren S, McKenna MP, Moler E, et al. Development of a type 2 diabetes risk model from a panel of serum biomarkers from the Inter99 cohort. Diabetes Care. 2009;32:1207–12.
4. Zhang X, Gregg EW, Williamson DF, Barker LE, Thomas W, Bullard KM, et al. A1C level and future risk of diabetes: a systematic review. Diabetes Care. 2010;33:1665–73.
5. Gieger C, Geistlinger L, Altmaier E, Hrabé de Angelis M, Kronenberg F, Meitinger T, et al. Genetics meets metabolomics: a genome-wide association study of metabolite profiles in human serum. PLoS Genet. 2008;4:e1000282.
6. Wang L, Athinarayanan S, Jiang G, Chalasani NZhang M, Liu W. Fatty acid desaturase 1 (FADS1) gene polymorphisms control human hepatic lipid composition. Hepatology. 2015;61:119–28.
7. Preet A, Karve TM, Rizk N, Cheema AK. Metabolomics: approaches and applications to diabetes research. J Diabetes Metab. 2012; S6:001.
8. Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. Metabolomics applied to diabetes research: moving from information to knowledge. Diabetes. 2009;58:2429–43.
9. Friedrich N. Metabolomics in diabetes research. J Endocrinol. 2012;215:29–42.
10. Suhre K. Metabolic profiling in diabetes. J Endocrinol. 2014;221: R75–85.
11. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care. 2014;37 Suppl 1:S81–90.
12. Cobb J, Gall W, Adam KP, Nakhle P, Button E, Hathorn J, et al. A novel fasting blood test for insulin resistance and prediabetes. J Diabetes Sci Technol. 2013;7:100–10.
13. Shaham O, Wei R, Wang TJ, Ricciardi C, Lewis GD, Vasan RS, et al. Metabolic profiling of the human response to a glucose challenge reveals distinct axes of insulin sensitivity. Mol Syst Biol. 2008;4:214.
14. Lam SM, Shui G. Lipidomics as a principal tool for advancing biomedical research. J Genet Genomics. 2013;40:375–90.
15. Kotronen A, Velagapudi VR, Yetukuri L, Westerbacka J, Bergholm R, Ekroos K, et al. Serum saturated fatty acids containing triacylglycerols are better markers of insulin resistance than total serum triacylglycerol concentrations. Diabetologia. 2009;52:684–90.
16. Rhee EP, Cheng S, Larson MG, Walford GA, Lewis GD, Mccabe E, et al. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. J Clin Invest. 2011;121:1402-11.
17. Li M, Yang L, Bai Y, Liu H. Analytical methods in lipidomics and their applications. Anal Chem. 2014;86:161–75.
18. Quehenberger O, Armando AM, Brown AH, Milne SB, Myers DS, Merrill AH, et al. Lipidomics reveals a remarkable diversity of lipids in human plasma. J Lipid Res. 2010;51:3299–305.
19. Zhao X, Fritsche J, Wang J, Chen J, Rittig K, Schmitt-Kopplin P, et al. Metabonomic fingerprints of fasting plasma and spot urine reveal human pre-diabetic metabolic traits. Metabolomics. 2010;6:362–74.
20. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab. 2009;9:311–26.
21. Huffman KM, Shah SH, Stevens RD, Bain JR, Muehlbauer M, Slentz CA, et al. Relationships between circulating metabolic intermediates and insulin action in overweight to obese, inactive men and women. Diabetes Care. 2009;32:1678–83.
22. Würtz P, Mäkinen V-P, Soininen P, Kangas AJ, Tukiainen T, Kettunen J, et al. Metabolic signatures of insulin resistance in 7, 098 young adults. Diabetes. 2012;61:1372–80.

23. Palmer ND, Stevens RD, Antinozzi PA, Anderson A, Bergman RN, Wagenknecht LE, et al. A metabolomic profile associated with insulin resistance and conversion to diabetes in the Insulin Resistance Atherosclerosis Study. J Clin Endocrinol Metab. 2014;jc20142357.
24. Klimentidis YC, Divers J, Casazza K, Beasley T, Allison DB, Fernandez JR. Ancestry-informative markers on chromosomes 2, 8 and 15 are associated with insulin-related traits in a racially diverse sample of children. Hum Genomics. 2011;5:79–89.
25. Tai ES, Tan ML, Stevens RD, Low YL, Muehlbauer MJ, Goh DL, et al. Insulin resistance is associated with a metabolic profile of altered protein metabolism in Chinese and Asian-Indian men. Diabetologia. 2010;53:757–67.
26.• Wang-Sattler R, Yu Z, Herder C, Messias AC, Floegel A, He Y, et al. Novel biomarkers for pre-diabetes identified by metabolomics. Mol Syst Biol. 2012;8:615. *Authors identified three metabolites (glycine, lysophosphatidylcholine (LPC) (18:2) and acetylcarnitine) that had significantly altered levels in IGT individuals as compared to those with NGT. These biomarkers were first identified in the KORA cohort and further validated in EPIC-Postdam cohort.*
27. Gall WE, Beebe K, Lawton KA, Adam KP, Mitchell MW, Nakhle PJ, et al. Alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. PLoS ONE. 2010;5:e10883.
28.• Ferrannini E, Natali A, Camastra S, Nannipieri M, Mari A, Adam KP, et al. Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. Diabetes. 2013;62:1730–7. *In this papers α-HB and L-GPC are proposed as novel and independent potent predictors for incident dysglycemia after their confirmation in two observational cohorts: RISC study and Botnia study.*
29. Menni C, Fauman E, Erte I, Perry JR, Kastenmüller G, Shin SY, et al. Biomarkers for type 2 diabetes and impaired fasting glucose using a nontargeted metabolomics approach. Diabetes. 2013;62:4270–6.
30. Kelder T, Eijssen L, Kleemann R, van Erk M, Kooistra T, Evelo C. Exploring pathway interactions in insulin resistant mouse liver. BMC Syst Biol. 2011;5:127.
31. Kutmon M, Evelo CT, Coort SL. A network biology workflow to study transcriptomics data of the diabetic liver. BMC Genomics. 2014;15:971.
32. Messana I, Forni F, Ferrari F, Rossi C, Giardina B, Zuppi C. Proton nuclear magnetic resonance spectral profiles of urine in type II diabetic patients. Clin Chem. 1998;44:1529–34.
33. Van Doorn M, Vogels J, Tas A, van Hoogdalem EJ, Burggraaf J, Cohen A, et al. Evaluation of metabolite profiles as biomarkers for the pharmacological effects of thiazolidinediones in Type 2 diabetes mellitus patients and healthy volunteers. Br J Clin Pharmacol. 2007;63:562–74.
34. Salek RM, Maguire ML, Bentley E, Rubtsov D V, Hough T, Cheeseman M, et al. A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. Physiol Genomics. 2007;29:99–108.
35. Yuan K, Kong H, Guan Y, Yang J, Xu G. A GC-based metabonomics investigation of type 2 diabetes by organic acids metabolic profile. J Chromatogr B Analyt Technol Biomed Life Sci. 2007;850:236–40.
36. Bao Y, Zhao T, Wang X, Qiu Y, Su M, Jia W, Jia W. Metabonomic variations in the drug-treated type 2 diabetes mellitus patients and healthy volunteers. J Proteome Res. 2009;8:1623–30.
37. Haus JM, Kashyap SR, Kasumov T, Zhang R, Kelly KR, Defronzo RA, et al. Plasma ceramides are elevated in obese subjects with type 2 diabetes and correlate with the severity of insulin resistance. Diabetes. 2009;58:337–43.
38. Li X, Xu Z, Lu X, Yang X, Yin P, Kong H, et al. Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: biomarker discovery for diabetes mellitus. Anal Chim Acta. 2009;633:257–62.
39. Zhang X, Wang Y, Hao F, Zhou X, Han X, Tang H, et al. Human serum metabonomic analysis reveals progression axes for glucose intolerance and insulin resistance statuses research articles. J Proteome Res. 2009;5188–95.
40. Zhang J, Yan L, Chen W, Lin L, Song X, Yan X, et al. Metabonomics research of diabetic nephropathy and type 2 diabetes mellitus based on UPLC-oaTOF-MS system. Anal Chim Acta. 2009;650:16–22.
41. Zeng M, Che Z, Liang Y, Wang B, Chen X, Li H, et al. GC–MS based plasma metabolic profiling of type 2 diabetes mellitus. Chromatographia. 2009;69:941–8.
42. Adams SH, Hoppel CL, Lok KH, Zhao L, Wong SW, Minkler PE, et al. Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid b-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. J Nutr. 2009;139:1073–81.
43. Fiehn O, Garvey WT, Newman JW, Lok KH, Hoppel CL, Adams SH. Plasma metabolomic profiles reflective of glucose homeostasis in non-diabetic and type 2 diabetic obese African-American women. PLoS One. 2010;5:e15234.
44. Mihalik SJ, Goodpaster BH, Kelley DE, Chace DH, Vockley J, Toledo FG, et al. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipotoxicity. Obesity (Silver Spring). 2010;18:1695–700.
45. Suhre K, Meisinger C, Döring A, Altmaier E, Belcredi P, Gieger C, et al. Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. PLoS One. 2010;5:e13953.
46. Jung J, Jang Z, Hwang G-S. Metabolomics approach for discovering disease biomarkers and understanding metabolic pathway. J Anal Sci Technol. 2011;2:A189–93.
47. Han LD, Xia JF, Liang QL, Wang Y, Wang YM, Hu P, et al. Plasma esterified and non-esterified fatty acids metabolic profiling using gas chromatography-mass spectrometry and its application in the study of diabetic mellitus and diabetic nephropathy. Anal Chim Acta. 2011;689:85–91.
48. Zhu C, Liang QL, Hu P, Wang YM, Luo G. Phospholipidomic identification of potential plasma biomarkers associated with type 2 diabetes mellitus and diabetic nephropathy. Talanta. 2011;85:1711–20.
49. Ha CY, Kim JY, Paik JK, Kim OY, Paik YH, Lee EJ, et al. The association of specific metabolites of lipid metabolism with markers of oxidative stress, inflammation and arterial stiffness in men with newly diagnosed type 2 diabetes. Clin Endocrinol (Oxf). 2012;76:674–82.
50. Mihalik SJ, Michaliszyn SF, de las Heras J, Bacha F, Lee S, Chace DH, et al. Metabolomic profiling of fatty acid and amino acid metabolism in youth with obesity and type 2 diabetes: evidence for enhanced mitochondrial oxidation. Diabetes Care. 2012;35:605–11.
51. Xu W, Zhang L, Huang Y, Yang Q, Xiao H, Zhang D. Discrimination of type 2 diabetes mellitus corresponding to different traditional Chinese medicine syndromes based on plasma fatty acid profiles and chemometric methods. J Ethnopharmacol. 2012;143:463–8.
52. Zhou Y, Qiu L, Xiao Q, Wang Y, Meng X, Xu R, et al. Obesity and diabetes related plasma amino acid alterations. Clin Biochem. 2013;46:1447–52.
53. Kaur P, Rizk N, Ibrahim S, Luo Y, Younes N, Perry B, et al. Quantitative metabolomic and lipidomic profiling reveals aberrant amino acid metabolism in type 2 diabetes. Mol Biosyst. 2013;9:307–17.
54. Zhang AH, Sun H, Yan GL, Yuan Y, Han Y, Wang XJ. Metabolomics study of type 2 diabetes using ultra-performance LC-ESI/quadrupole-TOF high-definition MS coupled with pattern recognition methods. J Physiol Biochem. 2014;70:117–28.
55. Mamtimin B, Hizbulla M, Kurbantay N, You L, Yan X, Upur H. An magnetic resonance-based plasma metabonomic investigation on

abnormal Savda in different complicated diseases. J Tradit Chin Med. 2014;34:166–72.
56. Lu H, Hu F, Zeng Y, Zou L, Luo S, Sun Y, et al. Ketosis onset type 2 diabetes had better islet β-cell function and more serious insulin resistance. J Diabetes Res. 2014;2014:510643.
57. Xu J, Cai S, Li X, Dong J, Ding J, Chen Z. Statistical two-dimensional correlation spectroscopy of urine and serum from metabolomics data. Chemom Intell Lab Syst. 2012;112:33–40.
58. Ståhlman M, Pham HT, Adiels M, Mitchell TW, Blanksby SJ, Fagerberg B, et al. Clinical dyslipidaemia is associated with changes in the lipid composition and inflammatory properties of apolipoprotein-B-containing lipoproteins from women with type 2 diabetes. Diabetologia. 2012;55:1156–66.
59. Larsen PJ, Tennagels N. On ceramides, other sphingolipids and impaired glucose homeostasis. Mol Metab. 2014;3:252–60.
60.•• Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, et al. Metabolite profiles and the risk of developing diabetes. Nat Med. 2011;17:448–53. *Five branched-chain and aromatic amino acids had highly significant associations with future diabetes: isoleucine, leucine, valine, tyrosine and phenylalanine. A combination of three amino acids predicted future diabetes (with a more than five-fold higher risk for individuals in top quartile). The results were replicated in an independent, prospective cohort. These findings underscore the potential key role of amino acid metabolism early in the pathogenesis of diabetes and suggest that amino acid profiles could aid in diabetes risk assessment.*
61. Floegel A, Stefan N, Yu Z, Mühlenbruch K, Drogan D, Joost HG, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. Diabetes. 2013;62:639–48.
62. Wang TJ, Ngo D, Psychogios N, Dejam A, Larson MG, Vasan RS, et al. 2-Aminoadipic acid is a biomarker for diabetes risk. J Clin Invest. 2013;123:4309–17.
63. Lorber D. Importance of cardiovascular disease risk management in patients with type 2 diabetes mellitus. Diabetes Metab Syndr Obes. 2014;7:169–83.
64. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, et al. The food metabolome: a window over dietary exposure. Am J Clin Nutr. 2014;99:1286–308.
65.•• Ley SH, Hamdy O, Mohan V, Hu FB. Prevention and management of type 2 diabetes: dietary components and nutritional strategies. Lancet. 2014;383:1999–2007. *This review highlight the role of dietary components and nutritional strategies for prevention and management of T2DM. Several dietary patterns such as Mediterranean, low glycaemic index, moderately low carbohydrate, and vegetarian diets are proposed as adecuate dietary options, independently of personal and cultural food preferences.*
66. de Koning L, Chiuve SE, Fung TT, Willett WC, Rimm EB, Hu FB. Diet-quality scores and the risk of type 2 diabetes in men. Diabetes Care. 2011;34:1150–6.
67. InterAct Consortium. Adherence to predefined dietary patterns and incident type 2 diabetes in European populations: EPIC-InterAct Study. Diabetologia. 2014;57:321–33.
68. Ajala O, English P, Pinkney J. Systematic review and meta-analysis of different dietary approaches to the management of type 2 diabetes. Am J Clin Nutr. 2013;97:505–16.
69.•• Estruch R, Ros E, Salas-Salvadó J, Covas MI, Corella D, Arós F, et al. Primary prevention of cardiovascular disease with a Mediterranean diet. N Engl J Med. 2013;368:1279–90. *In this primary prevention trial authors observed that among persons at high cardiovascular risk (including cases of T2DM), a Mediterranean diet supplemented with extra-virgin olive oil or nuts reduced significantly the incidence of major cardiovascular events.*
70. Viscogliosi G, Cipriani E, Liguori ML, Marigliano B, Saliola M, Ettorre E, et al. Mediterranean dietary pattern adherence: associations with prediabetes, metabolic syndrome, and related microinflammation. Metab Syndr Relat Disord. 2013;11:210–6.
71. Koloverou E, Esposito K, Giugliano D, Panagiotakos D. The effect of Mediterranean diet on the development of type 2 diabetes mellitus: a meta-analysis of 10 prospective studies and 136,846 participants. Metabolism. 2014;63:903–11.
72. Salas-Salvadó J, Bulló M, Babio N, Martínez-González MÁ, Ibarrola-Jurado N, Basora J, et al. Reduction in the incidence of type 2 diabetes with the Mediterranean diet: results of the PREDIMED-Reus nutrition intervention randomized trial. Diabetes Care. 2011;34:14–9.
73. Salas-Salvadó J, Bulló M, Estruch R, Ros E, Covas MI, Ibarrola-Jurado N, et al. Prevention of diabetes with Mediterranean diets: a subgroup analysis of a randomized trial. Ann Intern Med. 2014;160:1–10.
74. Magnusdottir OK, Landberg R, Gunnarsdottir I, Cloetens L, Åkesson B, Önning G, et al. Plasma alkylresorcinols reflect important whole-grain components of a healthy Nordic diet. J Nutr. 2013;143:1383–90.
75. Magnusdottir OK, Landberg R, Gunnarsdottir I, Cloetens L, Akesson B, Landin-Olsson M, et al. Plasma alkylresorcinols C17:0/C21:0 ratio, a biomarker of relative whole-grain rye intake, is associated to insulin sensitivity: a randomized study. Eur J Clin Nutr. 2014;68:453–8.
76. Akesson A, Andersen LF, Kristjánsdóttir AG, Roos E, Trolle E, Voutilainen E, et al. Health effects associated with foods characteristic of the Nordic diet : a systematic literature review. Food Nutr Res. 2013;57. doi:10.3402/fnr.v57i0.22790.
77. Kanerva N, Rissanen H, Knekt P, Havulinna AS, Eriksson JG, Männistö S. The healthy Nordic diet and incidence of type 2 diabetes—10-year follow-up. Diabetes Res Clin Pract. 2014;106:e34–7.
78. Stella C, Beckwith-Hall B, Cloarec O, Holmes E, Lindon JC, Powell J, et al. Susceptibility of human metabolic phenotypes to dietary modulation. J Proteome Res. 2006;5:2780–8.
79. Walsh MC, Brennan L, Pujos-guillot E, Sébédio J, Scalbert A, Fagan A, et al. Influence of acute phytochemical intake on human urinary metabolomic profiles. Am J Clin Nutr. 2007;86:1687–93.
80. Rasmussen LG, Winning H, Savorani F, Toft H, Larsen TM, Dragsted LO, et al. Assessment of the effect of high or low protein diet on the human urine metabolome as measured by NMR. Nutrients. 2012;4:112–31.
81. Heinzmann SS, Merrifield CA, Rezzi S, Kochhar S, Lindon JC, Holmes E, et al. Stability and robustness of human metabolic phenotypes in response to sequential food challenges. J Proteome Res. 2012;11:643–55.
82. May DH, Navarro SL, Ruczinski I, Hogan J, Ogata Y, Schwarz Y, et al. Metabolomic profiling of urine: response to a randomised, controlled feeding study of select fruits and vegetables, and application to an observational study. Br J Nutr. 2013;110:1760–70.
83. Floegel A, von Ruesten A, Drogan D, Schulze MB, Prehn C, Adamski J, et al. Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. Eur J Clin Nutr. 2013;67:1100–8.
84. Moazzami AA, Zhang J-X, Kamal-Eldin A, Aman P, Hallmans G, Johansson JE, et al. Nuclear magnetic resonance-based metabolomics enable detection of the effects of a whole grain rye and rye bran diet on the metabolic profile of plasma in prostate cancer patients. J Nutr. 2011;141:2126–32.
85. Vázquez-Fresno R, Llorach R, Marinic J, Tulipani S, Garcia-Aloy M, Espinosa-Martos I, et al. Urinary metabolomic fingerprinting after consumption of a probiotic strain in women with mastitis. Pharmacol Res. 2014;87:160-5.
86. Llorach R, Urpi-Sarda M, Tulipani S, Garcia-Aloy M, Monagas M, Andres-Lacueva C. Metabolomic fingerprint in patients at high risk of cardiovascular disease by cocoa intervention. Mol Nutr Food Res. 2013;57:962–73.

2.4.2 Biomarcadores de patrones alimentarios e intervenciones dietéticas

Una limitación importante de la investigación nutricional en poblaciones de vida libre es la evaluación objetiva de la ingesta dietética, la cual hasta la fecha se sigue realizando mediante el uso de cuestionarios de frecuencia de alimentos, recordatorios y diarios dietéticos (Thomson et al. 2003). Sin embargo, la prevalencia de declaraciones erróneas a través de estos instrumentos puede variar enormemente, desde un 30 hasta un 88% (Garcia-Perez et al. 2017). Estas declaraciones erróneas suelen incluir, por ejemplo, una subestimación de la ingesta de alimentos poco saludables y/o, en contraparte, una sobrevaloración de la ingesta de alimentos saludables (Garcia-Perez et al. 2017). Afortunadamente, durante los últimos años las aplicaciones de la metabolómica en la investigación nutricional se han incrementado significativamente, particularmente en la identificación de tres tipos de biomarcadores:

i) Biomarcadores dietéticos (Corella & Ordovas 2015; O'Gorman & Brennan 2015): Estos biomarcadores proporcionan información acerca de la ingesta de ciertos alimentos, y por esto son empleados cada vez con mayor frecuencia para evaluar la ingesta dietética (Odriozola & Corrales 2015). En términos generales se reconocen dos tipos de biomarcadores dietéticos: los biomarcadores de ingesta de alimentos específicos (*food* metabolome) y los biomarcadores de ingesta de patrones alimentarios (O'Gorman & Brennan 2015). Actualmente la mayoría de los biomarcadores dietéticos descritos en la literatura se corresponden a la ingesta de alimentos específicos como por ejemplo: cítricos, carne, pescado, café, pan, frutos secos, vino, etc. No obstante, existe un creciente interés por el estudio de biomarcadores que describan, más bien, la ingesta de varios alimentos a través de los patrones alimentarios, especialmente aquellos de carácter saludable (Hu 2002; O'Gorman & Brennan 2015).

Un patrón alimentario puede definirse como la ingesta habitual de diferentes alimentos y bebidas, variando en proporción, variedad, combinaciones y frecuencia con la que estos se consumen (Scientific Report of the 2015 Dietary Guidelines Advisory Committee [DGAC] 2015). Teniendo en cuenta esta definición, algunos ejemplos de patrones alimentarios descritos a través de biomarcadores metabolómicos incluyen: la ingesta habitual de distintos grupos de alimentos y nutrientes (Altmaier et al. 2011; M.-B.S. Andersen et al. 2014; Floegel et al. 2013; Guertin et al. 2014; Lloyd et al. 2013; Menni et al. 2013; O'Gorman et al. 2014; O'Sullivan et al. 2011), permitiendo incluso en ocasiones clasificar a estos patrones como saludables o poco saludables (Barton et al. 2015; Floegel et al. 2013; Garcia-Perez et al. 2017; Gibbons et al. 2017). En cuanto a patrones alimentarios más definidos por una región geográfica se ha logrado, por ejemplo, describir y diferenciar entre una dieta Nórdica y una Danesa (Maj-Britt S Andersen et al. 2014; Khakimov et al. 2016; Marklund et al. 2014), y la dieta occidental (Bouchard-

Mercier et al. 2013). Sin embargo, biomarcadores de adherencia a otros patrones alimentarios saludables como la DietMed han sido poco estudiados (Gutierrez-Diaz et al. 2017).

ii) Biomarcadores de enfermedades metabólicas (y riesgo) relacionadas con la dieta: Los biomarcadores de enfermedades relacionados con la dieta son aquellos que describen diferentes fenotipos intermedios de alguna enfermedad o incluso de la severidad de la misma. Un gran número de biomarcadores en esta categoría incluyen enfermedades como la obesidad (Elliott et al. 2015), diabetes (Urpi-Sarda et al. 2015) y el SM (Pujos-Guillot et al. 2017), entre otros. Sobre este tipo de biomarcadores, específicamente relacionados con la diabetes, se ha hablado ya en el apartado 2.4.1.

En este punto es interesante e importante mencionar que los estudios metabolómicos más recientes en torno a la identificación de biomarcadores dietéticos, y de enfermedades metabólicas, están reconociendo la necesidad de optar por integrar o combinar distintos metabolitos en un solo modelo que permita una mejor y más amplia descripción del proceso estudiado (Cobb et al. 2015; Garcia-Aloy et al. 2015; Madrid-Gambin et al. 2017; Pujos-Guillot et al. 2017). En el caso de los patrones alimentarios, estos llamados biomarcadores multimetabolito o paneles de biomarcadores proporcionan medidas más precisas acerca de la exposición a los alimentos, en comparación con biomarcadores simples, por separado, o de un solo metabolito (Marklund et al. 2014). Mientras tanto, en el caso de las enfermedades metabólicas, este tipo de biomarcadores son de gran utilidad tanto en la predicción y diagnóstico, como en la evolución y estratificación de las poblaciones en función de la gravedad o riesgo (Carter et al. 2016; Pujos-Guillot et al. 2017). Lo anterior es de especial interés en este apartado, además, debido a que el enfoque de biomarcador multimetabolito se empleó en dos estudios que forman parte de esta Tesis Doctoral (biomarcadores de diabetes y de adherencia a la DietMed).

iii) Biomarcadores de intervenciones nutricionales: Con respecto a los biomarcadores de intervenciones nutricionales, estos suelen obtenerse a partir de estudios de intervención con alimentos o patrones alimentarios para medir el efecto que se produce sobre ciertas moléculas diana, incluido el metaboloma. Más concretamente, este tipo de biomarcadores reflejan no sólo la ingesta del alimento (nutriente o patrón), sino también su metabolismo y los posibles efectos sobre una enfermedad (Corella & Ordovas 2015). Ejemplos de este tipo de biomarcadores incluyen aquellos que describen los efectos de intervenciones dietéticas combinadas con otras acciones sobre el estilo de vida (p.e., actividad física) para la disminución del riesgo cardiovascular (Brennan et al. 2015; Gibbons et al. 2015; Bondia-Pons, José Alfredo Martinez, et al. 2015; Piccolo et al. 2015). Otros ejemplos más concretos en cuanto a la aplicación de la identificación de biomarcadores a través de aproximaciones metabolómicas en estudios de

intervención con patrones alimentarios incluyen: los efectos de la nueva dieta Nórdica y la DietMed (con o sin restricción calórica) sobre componentes del SM y la pérdida de peso (Lankinen et al. 2016; Khakimov et al. 2016; Marklund et al. 2014; Bondia-Pons, José Alfredo Martinez, et al. 2015);

2.4.3 Biomarcadores de dieta Mediterránea

Como se ha mencionado anteriormente, la DietMed es un patrón alimentario saludable que recibe cada vez mayor interés debido a los múltiples beneficios a la salud, los cuales a su vez están mediados por la ingesta, tanto en lo individual como con efecto sinérgico, de sus componentes principales como las frutas, verduras, legumbres, frutos secos, aceite de oliva, vino tinto, pescado, etc. Esta amplia variedad de alimentos y efectos beneficiosos hacen que exista un creciente interés por describir la DietMed a través de marcadores de ingesta, también conocidos como biomarcadores metabolómicos del *food metabolome* (Scalbert et al. 2014) (Scalbert et al. 2014). En este sentido, al día de hoy se encuentran reportados en la literatura una variedad de metabolitos o biomarcadores del *food metabolome*, descritos a través de estudios metabolómicos. Cabe mencionar que la mayoría de estos han podido ser identificados en muestras de suero/plasma y orina, y a través de plataformas analíticas basadas en cromatografía líquida acoplada a espectrometría de masas (del inglés, *liquid chromatography coupled to mass spectrometry*, LC-MS), mientras que sólo unos cuantos por RMN. Una lista con algunos de los biomarcadores de ingesta de alimentos característicos de la DietMed, identificados mediante RMN, y frecuentemente reportados en la bibliografía se muestra en la tabla 2.4.

Como se observa en esta tabla, no sólo el número de biomarcadores específicos de alimentos característicos de la DietMed e identificados mediante RMN es limitado, sino también el número y variedad de alimentos descritos mediante esta técnica analítica. Caso contrario sucede con los biomarcadores identificados mediante técnicas basadas en LC-MS, al haber mayor número de estos reportados y de alimentos (incluidos y no en la DietMed) en la literatura (Scalbert et al., 2014; Brennan et al. 2015). Esto se debe principalmente a que pese a la mayor robustez que ofrece la RMN frente a las técnicas de LC-MS, estas últimas proporcionan mayor sensibilidad en la identificación de metabolitos derivados de la ingesta de alimentos, los cuales suelen encontrarse en concentraciones tan pequeñas que en ocasiones a través de la RMN no puede detectar (Guertin et al. 2014).

Tabla 2.4 Ejemplos de metabolitos propuestos en la literatura como biomarcadores de la ingesta de alimentos característicos de la dieta Mediterránea[1]

Grupo de alimentos/alimento	Metabolito	Muestra biológica	Referencias
(↑) Frutas	Prolina betaína Ácido hipúrico Ácido tartárico	Orina	(Heinzmann et al. 2010) (Heinzmann et al. 2012)
(↑) Verduras	Fenilacetilglutamina	Orina	(Stella et al. 2006) (O'Sullivan et al. 2011)
(↑) Verduras: crucíferas	SMCS	Orina	(Edmands et al. 2011)
(↑) Legumbres	Dimetilamina 3-metilhistidina	Orina	(Madrid-Gambin et al. 2017)
(↑) Fibra dietética	Ácido hipúrico Ácido tartárico		(Rasmussen et al. 2012)
(-) Vino tinto	Ácido 4-OHFA Manitol Etanol	Orina	(Vazquez-Fresno et al. 2012)
(↓) Carne	*O*-Acetilcarnitinas 1 y 3 metilhistidina Creatina Carnitina TMAO	Orina	(Stella et al. 2006) (O'Sullivan et al. 2011) (Cross et al. 2011)

[1]Identidicados a través de una plataforma analítica de RMN. Abreviaciones y simbología: 4-OHFA, ácido 4-hidroxifenilacético, TMAO, trimetilamina, N-óxido; RMN, resonancia magnética nuclear; S-metil-cisteina sulfóxido; ↑, alto consumo; ↓, bajo consumo; -, consumo moderado.

De este modo, mediante técnicas de LC-MS no sólo se ha logrado confirmar y extender la lista de metabolitos mencionados en la tabla 2.3, sino que se han descrito otros asociados al consumo de otros alimentos característicos de la DietMed como los frutos secos (Garcia-Aloy et al. 2014), vino (Boto-Ordonez et al. 2013), pescado (Schmidt et al. 2015). No obstante, a pesar de la amplia variedad de biomarcadores reportados en la literatura, son pocos los estudios que han evaluado o descrito la ingesta de un patrón alimentario en específico a través de un enfoque metabolómico. En este sentido, existen menos estudios todavía en relación a la descripción de la DietMed y/o su adherencia mediante biomarcadores metabolómicos.

En un trabajo reciente publicado por Playdon y colaboradores (2017) se identificaron biomarcadores metabolómicos correlacionados con 4 índices de patrones alimentarios: el *Healthy Eating Index* (HEI) 2010, el *alternative Medieterranean Diet Score* (aMED), el *WHO Healthy Diet indicator* (HDI), y el *Baltic Sea Diet*. En este estudio, se encontró que el aMED se asoció con los niveles plasmáticos de 46 metabolitos, de los cuales sólo 21 pudieron ser identificados. La mayoría de estos metabolitos, sin embargo, no se encontraron asociados exclusivamente al aMED (Playdon et al. 2017).

Es importante reconocer además que tras la ingesta de alimentos, ya sea individualmente o a través de patrones alimentarios, se producen interacciones entre los compuestos derivados de estos y el metabolismo endógeno, así como con la microbiota intestinal, lo que da lugar a otra

variedad de metabolitos relacionados sí con la ingesta de alimentos pero no derivados directamente de su consumo. Este tipo de metabolitos, por lo tanto, a través de las huellas metabolómicas asociadas a la ingesta de patrones alimentarios pueden contribuir tanto en la descripción de estos últimos como en sus efectos beneficiosos a la salud. En este sentido, existen también pocos estudios que a través de la metabolómica hayan evidenciado no sólo la ingesta de la DietMed (Vazquez-Fresno et al. 2015) sino también los efectos beneficiosos de esta en la salud metabólica (Bondia-Pons et al. 2015). Se reconoce, por lo tanto, que la evaluación de la adherencia a la DietMed y el estudio de los mecanismos específicos de sus efectos beneficiosos en la salud, en el contexto de un ensayo de intervención u observacional, y mediante el uso de aproximaciones metabolómicas, pueden contribuir a proporcionar mensajes más sólidos de salud pública y ayudar a identificar objetivos moleculares clave para una prevención y manejo más eficaces de las ECV (Martinez-Gonzalez et al. 2016).

En resumen, describir en su totalidad a los componentes de la DietMed a través de la metabolómica resulta un desafío a superar en el campo de la investigación nutricional y epidemiológica. Además, aún con la identificación de biomarcadores de ingesta de todos estos componentes dietéticos, se ha de considerar que debido al intervariabilidad metabólica de las personas, estos no siempre podrán ser encontrados en las mismas proporciones en las muestras biológicas que normalmente se emplean en estudios de metabolómica. En este sentido, el concepto de biomarcadores de adherencia a la DietMed puede resultar de mayor interés y aplicabilidad, ya que estos, a diferencia de los biomarcadores de ingesta y de los cuestionarios e índices de cumplimiento, reflejarían de manera más precisa el nivel de seguimiento de la DietMed y su modulación en el metabolismo endógeno y de la microbiota intestinal.

III. OBJETIVOS

3.1 OBJETIVOS

El **objetivo principal** de esta tesis doctoral es identificar biomarcadores relacionados con factores de riesgo cardiovascular y el estudio de los perfiles metabolómicos asociados a un patrón de alimentación mediterránea mediante el empleo una aproximación metabolómica no dirigida basada en resonancia magnética nuclear.

Para la consecución de este objetivo se plantearon los siguientes **objetivos específicos**:

1. Caracterizar la huella metabólica urinaria asociada a la diabetes mellitus tipo 2 (marcadores de enfermedad) en una población de elevado riesgo cardiovascular.

2. Calcular un biomarcador metabolómico con una alta capacidad predictiva (elevada sensibilidad y especificidad) de diabetes prevalente, a partir de la huella metabólica asociada a la misma.

3. Definir fenotipos metabólicos dentro de la población diabética a partir de la huella metabólica asociada a la diabetes.

4. Identificar la huella metabólica urinaria que permita diferenciar entre individuos con una elevada y baja adherencia a un patrón de dieta Mediterránea.

5. Calcular un biomarcador metabolómico de alta adherencia a la dieta Mediterránea con una elevada sensibilidad y especificidad.

6. Determinar los cambios producidos en el metaboloma plasmático por efecto de una intervención basada en dieta Mediterránea hipocalórica y actividad física en individuos obesos metabólicamente sanos.

7. Estudiar la asociación entre los cambios producidos en el metaboloma y la pérdida de peso tras la intervención con dieta y actividad física en individuos obesos metabólicamente sanos.

IV. METODOLOGÍA

4.1 ESTUDIOS CLÍNICOS

Los objetivos de la presente Tesis Doctoral están enmarcados en dos estudios clínicos que a continuación se describen.

4.1.1 Diseño del estudio PREDIMED

El estudio PREDIMED es un ensayo clínico multicéntrico, paralelo, aleatorizado y controlado, realizado en España desde el año 2003 hasta el 2010 con el objetivo de evaluar el efecto de una intervención con DietMed en la prevención primaria de la enfermedad cardiovascular en individuos con alto riesgo (Estruch et al. 2013). La intervención consistió en el seguimiento de una DietMed con suplementación de aceite de oliva virgen extra o una mezcla de frutos secos, comparadas ambas con una dieta control baja en grasa (Martínez-González et al. 2012; Estruch et al. 2013), y tuvo una mediana de seguimiento de 4,8 años (Figura 4.1).

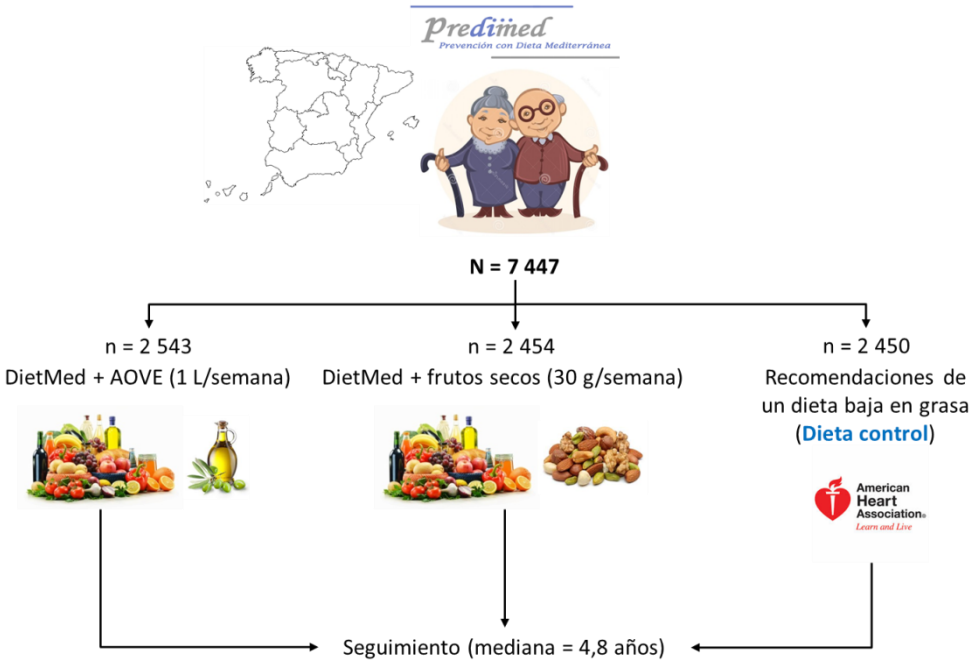


Figura 4.1 Esquema del diseño del estudio PREDIMED llevado a cabo en España entre los años 2003 y 2010. Abreviaciones: DietMed, dieta Mediterránea; AOVE, aceite de oliva virgen extra.

Este ensayo clínico se registró en el *Current Controlled Trials* de Londres, con el ID ISRCTN35739639, y su protocolo se ajustó a la Declaración de Helsinki, siendo aprobado además por los comités de ética de cada uno de los centros participantes (Martínez-González et al. 2012). Para el desarrollo de los dos estudios metabolómicos conducidos en la presente Tesis Doctoral, se contó con muestras de orina, y datos clínicos, sociodemográficos y de ingesta de alimentos de 210 individuos provenientes de dos centros clínico del estudio PREDIMED (Hospitales clínicos de Valencia y Barcelona) en el tiempo inicial, es decir antes de la intervención.

Por otra parte, los diseños de estos dos estudios metabolómicos corresponden al proyecto AGL2009-13906-C01/02: "Desarrollo de biomarcadores de consumo y efecto de un patrón de alimentación Mediterránea en la prevención de la enfermedad cardiovascular: una aproximación metabolómica", coordinado en colaboración entre la Universidad de Barcelona y el Hospital Clínico de Barcelona.

4.1.1.1 Características de los participantes

El estudio PREDIMED incluyó hombres con edades comprendidas entre 50 y 80 años, y mujeres de entre 60 y 80 años de edad. Al momento de su inclusión ninguno de los participantes presentaba ECV diagnosticada pero cumplían con alguno de los dos siguientes criterios de inclusión:

i) DM2, definida por una glucemia basal $\geq$126 mg/dL; glucemia casual $\geq$200 mg/dL acompañada de síntomas de diabetes; glucemia después de un test de tolerancia a la glucosa $\geq$200 mg/dL en dos determinaciones; o tratamiento con insulina o medicamentos hipoglucemiantes orales.

ii) Presentar tres o más de los siguientes factores de riesgo cardiovascular:

- Tabaquismo (>1 cigarrillo/día, durante el mes anterior a la inclusión)
- Hipertensión arterial (presión arterial sistólica $\geq$140 mmHg o presión arterial diastólica $\geq$90 mmHg, o tratamiento con medicamentos antihipertensivos).
- Hipercolesterolemia (colesterol LDL $\geq$160 mg/dL o tratamiento con medicamentos hipolipemiantes).
- Bajos niveles de colesterol HDL (HDL $\leq$ 40 mg/dL en hombres o $\leq$ 50 mg/dL en mujeres).
- Sobrepeso u obesidad (IMC $\geq$ 25 kg/m^2).
- Historia familiar de ECV prematura (familiares de primer orden hombres < 55 años o mujeres < 65 años).

Por otro lado, se excluyeron aquellos individuos que presentaban alguno de los siguientes criterios:

- Historia previa de ECV (enfermedad coronaria, accidente cerebrovascular o ateriopatía periférica).
- Enfermedad crónica grave.
- Inmunodeficiencia o infección por el VIH.
- Antecedentes de hipersensibilidad o alergia al aceite de oliva o frutos secos.
- Baja predisposición a cambiar los hábitos dietéticos, según los estadios de disposición al cambio del modelo de Prochaska y DiClemente (Nigg et al. 1999).

4.1.1.2 Mediciones, cuestionarios y recogida de muestras biológicas

Al momento de su inclusión, los participantes acudieron en ayunas a los centros cínicos. En esta primera visita, a los participantes se les tomaron medidas antropométricas (peso, altura e IMC) y de presión arterial, y se les proporcionó los siguientes cuestionarios:

i) Cuestionario general. Compuesto por 47 ítems que incluían información sobre las variables socio-demográficas, medicación actual y los factores de riesgo.

ii) Cuestionario de frecuencia de consumo de alimentos. Semicuantitativo y compuesto por 137 ítems (Fernandez-Ballart et al. 2010). A partir de este cuestionario pudo calcularse tanto la composición de nutrientes como la ingesta energética de acuerdo a tablas españolas de composición de alimentos (Mataix 2003).

iii) Cuestionario de actividad física. (Elosua et al. 1994).

iv) Cuestionario de 14 puntos para medir la adherencia a la DietMed. Explicado con mayor detalle en el siguiente apartado de este capítulo (4.1.1.3).

En cuanto a las muestras biológicas, por cada participante se recogió orina de primera hora de la mañana y en condiciones de ayuno. Esta muestra se alicuotó, codificó y almacenó a -80 ºC hasta el momento de sus análisis.

4.1.1.3 Cuestionario de adherencia a la dieta Mediterránea

Con la finalidad de definir el nivel de adherencia a la DietMed entre los participantes, se empleó un cuestionario validado previamente en el estudio PREDIMED compuesto por 14 preguntas (Schroder et al. 2011). Brevemente, este cuestionario está compuesto por 12 preguntas acerca de la frecuencia del consumo de alimentos y otras 2 preguntas sobre los hábitos de la ingesta de alimentos considerados característicos de la dieta Mediterránea española (Tabla 4.1). La respuesta a cada pregunta se puntúa como 0 o 1, de acuerdo a la respuesta (Tabla 4.1). La

puntuación o *score* total posible a este cuestionario va de 0 a 14 puntos, teniendo en cuenta que a mayor puntuación, mayor adherencia a la DietMed. De este modo, el *score* de adherencia a la DietMed fue calculado para cada uno de los participantes en este estudio y posteriormente empleado para su estratificación.

4.1.1.4 Estratificación de la población para los estudios de biomarcadores de diabetes y de adherencia a la dieta Mediterránea

Para el estudio de biomarcadores de riesgo cardiovascular, se incluyeron datos y muestras biológicas 154 participantes, de los cuales 85 eran diagnosticados como diabéticos tipo 2 (DM2) y 69 libres de DM2 (Figura 4.2). Mientras que para el estudio de biomarcadores de alta adherencia a la DietMed, 122 participantes fueron incluidos y clasificados en grupos de baja (n=63) y alta (n=56) adherencia a la DietMed, según su *score* total obtenido del cuestionario de 14 preguntas (Figura 4.2). Específicamente, se consideraron los *scores* $\leq$ 7 y $\geq$ 10 como los puntos de corte para la clasificación de baja y alta adherencia a la DietMed, respectivamente (Martinez-González et al. 2012).

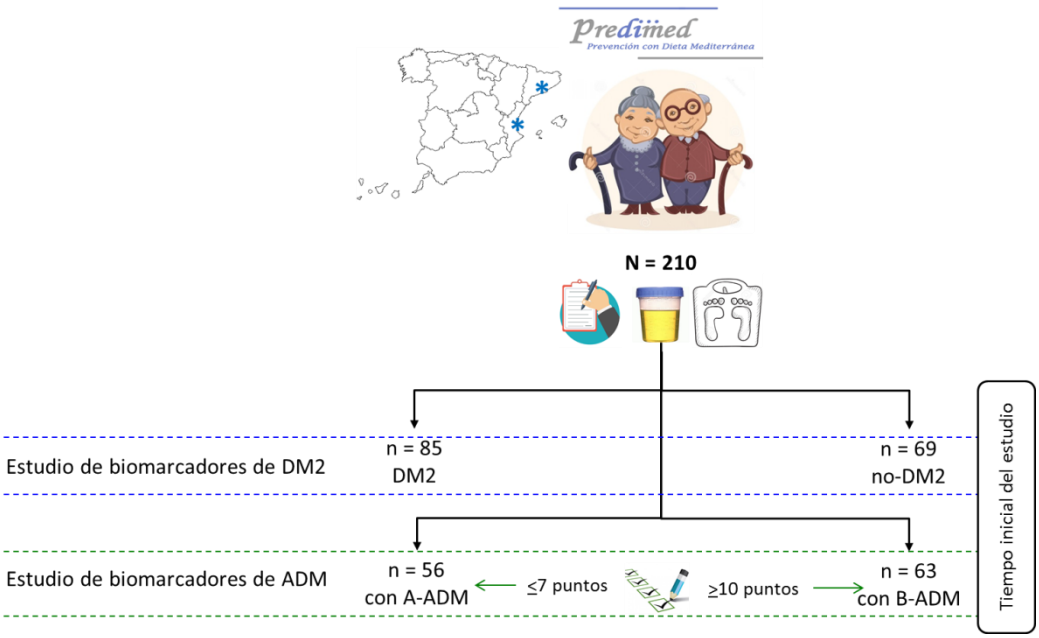


Figura 4.2 Esquema de los diseños de estudios de biomarcadores de diabetes y adherencia a la dieta Mediterránea, a partir de las muestras biológicas, medidas antropométricas y datos alimentarios y sociodemográficos de 210 participantes del estudio PREDIMED (centros Valencia y Barcelona). Abreviaturas: ADM, adherencia a la dieta Mediterránea; A-ADM, alta adherencia a la dieta Mediterránea; B-ADM, baja adherencia a la dieta Mediterrána; DM2, diabetes mellitus tipo 2.

Tabla 4.1 Cuestionario para evaluar la adherencia a la dieta Mediterránea en la población del estudio PREDIMED

Pregunta		Criterio de puntuación
1	¿Usa usted el aceite de oliva como principal grasa para cocinar?	Sí = 1 punto
2	¿Cuánto aceite de oliva consume en total al día (incluyendo el usado para freír, comidas fuera de casa, ensaladas, etc.)?	4 o más cucharadas = 1 punto
3	¿Cuantas raciones de verdura u hortalizas consume al día? (las guarniciones o acompañamientos = 1/2 ración) 1 ración = 200g.	2 o más (al menos una de ellas en ensalada o crudas) = 1 punto
4	¿Cuantas piezas de fruta (incluyendo zumo natural) consume al día?	3 o más al día = 1 punto
5	¿Cuantas raciones de carnes rojas, hamburguesas, salchichas o embutidos consume al día? (ración: 100 - 150 g)	menos de 1 al día = 1 punto
6	¿Cuantas raciones de mantequilla, margarina o nata consume al día? (porción individual: 12 g)	menos de 1 al día = 1 punto
7	¿Cuantas bebidas carbonatadas y/o azucaradas (refrescos, colas, tónicas, bitter) consume al día?	menos de 1 al día = 1 punto
8	¿Bebe usted vino? ¿Cuánto consume a la semana?	7 o más vasos a la semana = 1 punto
9	¿Cuantas raciones de legumbres consume a la semana? (1 plato o ración de 150 g)	3 o más a la semana = 1 punto
10	¿Cuantas raciones de pescado-mariscos consume a la semana? (1 plato pieza o ración: 100 - 150 de pescado o 4-5 piezas o 200 g de marisco)	3 o más a la semana = 1 punto
11	¿Cuantas veces consume repostería comercial (no casera) como galletas, flanes, dulce o pasteles a la semana?	menos de 2 a la semana = 1 punto
12	¿Cuantas veces consume frutos secos a la semana? (ración 30 g)	3 o más a la semana = 1 punto
13	Consume usted preferentemente carne de pollo, pavo o conejo en vez de ternera, cerdo, hamburguesas o salchichas? (carne de pollo: 1 pieza o ración de 100 - 150 g)	Sí = 1 punto
14	¿Cuantas veces a la semana consume los vegetales cocinados, la pasta, arroz u otros platos aderezados con salsa de tomate, ajo, cebolla o puerro elaborada a fuego lento con aceite de oliva (sofrito)?	2 o más a la semana = 1 punto

4.1.2 Diseño del estudio INTERVENCIÓN EN EL ESTILO DE VIDA EN MUJERES OBESAS METABÓLICAMENTE SANAS

El presente estudio fue un ensayo clínico, aleatorizado, paralelo y simple ciego, conducido por el hospital regional universitario de Málaga y el hospital Virgen de la Victoria, en Málaga, España, con el objetivo de evaluar el efecto de una intervención en el estilo de vida sobre la pérdida de peso y otros marcadores cardiometabólicos en mujeres obesas metabólicamente sanas. La intervención consistió en un grupo tratamiento y un grupo control. Los participantes del grupo tratamiento recibieron un programa basado en una DietMed hipocalórica y la realización de actividad física de manera regular. La DietMed hipocalórica incluyó la ingesta de aceite de oliva virgen extra y frutos secos, pero con una restricción calórica total de 600 kcal (aproximadamente un 30% de los requerimientos de energía estimados). La distribución de la ingesta calórica diaria en este grupo fue: 35-40% proveniente de las grasas (de las cuales 8-10% provenientes de ácidos grasos saturados), 40-45% proveniente de carbohidratos de bajo índice glicémico, y un 20% proveniente de proteínas. En cuanto a la actividad física, esta consistió en la práctica de al menos 150 min/semana de caminata. Por otra parte, los participantes del grupo control recibieron recomendaciones generales acerca de una dieta saludable y actividad física. Este estudio se condujo de acuerdo a las directrices establecidas en la Declaración de Helsinki, y todos los protocolos fueron aprobados por un comité de Ética (Comité Coordinador de Ética de la Investigación Biomédica de Andalucía). Este estudio además se encuentra registrado en el *Current Controlled Trials* con el ID ISRCTN88315555.

Para el desarrollo del estudio metabolómico conducido en esta Tesis Doctoral, a partir del presente estudio clínico, se contó con muestras de plasma, y datos clínicos y sociodemográficos de 57 participantes en el tiempo inicial y después de 3 y 12 meses de intervención.

4.1.2.1 Características de los participantes

Los participantes del presente estudio fueron mujeres, con edades entre 35 y 55 años, obesas (IMC $\geq$ 30kg/m^2) y definidas como metabólicamente sanas, provenientes de cuatro centros de salud del distrito de Málaga (Andalucía, España). El diagnóstico de salud metabólica se realizó en base a los criterios propuestos por la Federación Internacional de Diabetes. En concreto, las participantes debían tener $\leq$ 1 de los siguientes factores de riesgo cardiovascular:

i) Elevados niveles de glucosa en sangre. Glucosa plasmática $\geq$ 100 mg/dL.

ii) Hipertensión. Presión arterial sistólica $\geq$ 135 mmHg y/o presión arterial diastólica $\geq$ 85 mmHg; o uso de medicamentos antihipertensivos.

iii) Hipertrigliceridemia. Triglicéridos ≥ 150 mg/dL o estar bajo tratamiento con medicamentos hipolipemiantes, o

iv) Bajos niveles de colesterol HDL. HDL < 50 mg/dL.

Las participantes fueron excluidas si presentaban alguno de los siguientes criterios:

- Diagnóstico previo de diabetes.
- Embarazo o planes de estarlo durante el estudio.
- Enfermedad cardiovascular o cualquier enfermedad crónica.
- Abuso de alcohol o drogas.
- Estar o haber estado en un programa para la pérdida de peso durante los últimos tres meses, que incluyera una intervención en la actividad física o la dieta.

4.1.2.3 Mediciones, cuestionarios y recogida de muestras biológicas

Se tomaron medidas de tipo antropométricas, incluyendo peso, altura, circunferencia de la cintura, e IMC, además de tomas de la presión arterial.

Las muestras biológicas empleadas para este estudio fueron muestras de sangre extraídas en condiciones de ayunas en tubos que contenían EDTA. Estas muestras se colectaron el día de la inscripción (tiempo inicial) y después de 3 y 12 meses de estudio. A partir de estas muestras se determinaron:

- Glucemia (glucosa plasmática después de un test de tolerancia a la glucosa de 75 g a 2 horas, y hemoglobina glicosilada [HbA1c]).
- Insulina (incluyendo el modelo homeostático de evaluación de resistencia a la insulina [HOMA-IR].
- Perfil lipídico. Triglicéridos; colesterol total y colesterol tipos LDL y HDL.

Estas determinaciones se realizaron dentro de las 12 horas posteriores a la recolección de las muestras y siguiendo los métodos estandarizados que se emplean en los centros de salud de adscripción de los participantes.

Por su parte, para el estudio metabolómico, las muestras de sangre se centrifugaron y el plasma obtenido se recogió en los mismos puntos de tiempo (inicial, 3 y 12 meses). Estas muestras de plasma entonces se alícuotaron, etiquetaron y almacenaron inmediatamente a -80 °C hasta su análisis.

4.2 ANÁLISIS METABOLÓMICO NO DIRIGIDO

Los estudios metabolómicos se realizaron siguiendo el esquema general del flujo de trabajo de la metabolómica, mencionadas anteriormente en el capítulo de antecedentes bibliográficos, y que se describen con mayor detalle en las siguientes secciones (Figura 4.3).

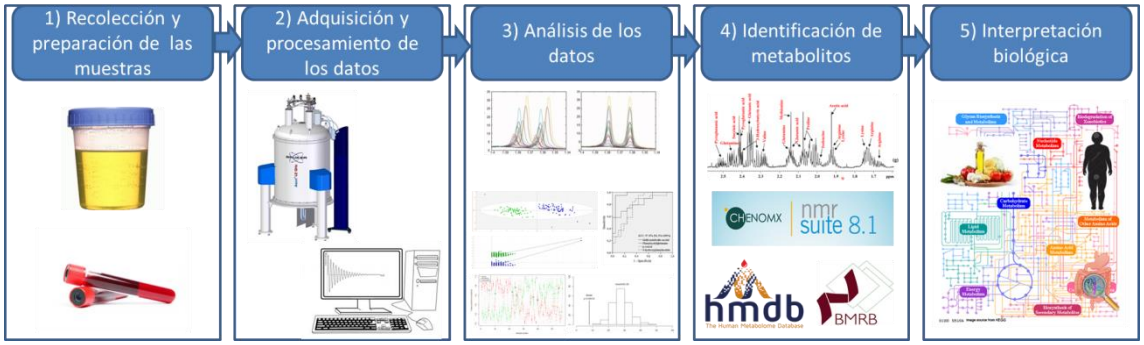


Figura 4.3 Diagrama de flujo de trabajo general aplicado en los estudios metabolómicos con aproximación no dirigida por RMN, y que componen la presente Tesis Doctoral.

4.2.1 Recolección y preparación de las muestras biológicas.

El análisis de muestras biológicas en estudios de metabolómica con una aproximación no dirigida requiere de una serie de operaciones previas en su preparación, de modo que se minimicen los errores de tipo analítico y se detecte el mayor número posible de metabolitos (Álvarez-Sánchez et al. 2010). De manera general, estas operaciones son lo más simples posibles con el fin de preservar en la medida de lo posible la muestra original. Estas operaciones previas a los experimentos de RMN varían además en función del tipo de muestra biológica, siendo la centrifugación, adición de soluciones tampón o *buffer* y agitado, las que se llevan a cabo comúnmente para muestras de orina y plasma (Beckonert et al. 2007).

Considerando estos aspectos, a continuación se describen las operaciones específicas en la preparación de las muestras analizadas en la presente Tesis Doctoral.

Orina. Es importante indicar en este punto que el protocolo seguido para la preparación de las muestras de orina y su posterior análisis del metaboloma se basó en una metodología desarrollada y validada por el grupo de investigación (Vazquez-Fresno et al. 2016). En concreto, las muestras fueron descongeladas a 4 °C dentro de un frigorífico, agitadas mediante *vortex* y posteriormente centrifugadas a 13.200 revoluciones por minuto (rpm) durante 5 minutos y una

temperatura constante de 4 °C. Trescientos microlitros de sobrenadante se vertieron en un tubo Eppendorf y se mezclaron con 120 µL de solución *buffer* (1,5 M de fosfato potásico monobásico [KH_2PO_4], 2 mM de azida sódica [NaN_3] y 0,1% de *3-(trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt* [TSP] en agua deuterada [D_2O]98% y ajustado a pH 7) y 200 µL de agua desionizada/agua deuterada en proporción 0,8/0,2. El volumen final de esta mezcla (620 µL) fue vertido en un tubo de cristal para RMN de 5 mm de diámetro y almacenado a 4 °C hasta el momento de su análisis.

Plasma. La preparación de las muestras de plasma y su posterior análisis del metaboloma se realizaron siguiendo una metodología desarrollada y validada por el el *Swedish NMR centre* (Radjursoga et al. 2017). En específico, las muestras fueron descongeladas en frigorífico a 4 °C y posteriormente agitadas mediante *vortex*. 150 µL de cada muestra fueron añadidos en una placa de 96 pocillos y diluidos con 150 µL de agua MilliQ y 600 µL de metanol puro a -20 °C. Las placas se agitaron mediante *vortex*, se incubaron (primero a 800 rpm por 10 min y a 12 °C, y luego por 30 minutos a -20 °C) y centrifugaron (a 2.250 g, por 60 min a 4 °C) para la precipitación de proteínas. 600 µL de sobrenadante fueron añadidos a una nueva placa de 96 pocillos y liofilizados a -4 °C durante 16 h. Posteriormente, las muestras secas o *pellets* se disolvieron en 50 µL de metanol deuterado (MeOD) y de nuevo liofilizadas hasta remover el exceso de MeOD. Los nuevos *pellets* fueron reconstituidos en 200 µL solución *buffer* (37,5 mM de fosfato de sodio, pH 6,95, 100% D_2O, 0.02% NaN_3, 0.25 mM de *4,4-dimethyl-4-silapentane-1-sulfonic acid* deuterado [DSS-d_6] y 1 mM de imidazol) y agitados a 800 rpm, a 22 °C, durante 30 minutos. Finalmente, 180 µL de esta mezcla fueron vertidos a un tubo de cristal para RMN de 3 mm de diámetro y puestos en cola a temperatura constante de 4 °C para su posterior análisis.

4.2.2 Adquisición de los datos y procesamiento de los espectros

4.2.2.1 Adquisición de los datos

Una vez preparadas para su análisis, las muestras (contenidas en los tubos de cristal) son colocadas en el espectrómetro de RMN, y una serie de parámetros son fijados en el software de uso del equipo para optimizar la adquisición de los datos de cada muestra. A continuación, en la tabla 4.1, se detallan los equipos y los parámetros empleados para adquisición de los datos en cada una de las muestras biológicas estudiadas en esta Tesis Doctoral.

4.2.2.2 Procesamiento de los espectros

Posterior a su adquisición, los espectros fueron procesados mediante la realización de tres operaciones conocidas como i) faseado (corrección de la fase), ii) referenciado (de los desplazamientos químicos en cada espectro, con respecto al desplazamiento químico de un estándar externo [δ 0.0] como el TSP [para las muestras de orina] o el DSS-d_6 [para las muestras de plasma] y iii) corrección de la línea base. El procesado de todos los espectros se realizó haciendo uso del programa TopSpin 3.0 (Bruker Biospin Ltd, Billerica, MA, USA).

Tabla 4.1 Parámetros de adquisición de los espectros de 1D-RMN a partir de las muestras biológicas empleadas para los estudios metabolómicos.

	Orina	Plasma
Equipo	Varian Inova, 500 MHz, con intercambiador de muestras automático	Bruker Oxford, 800 MHz, con sonsola Avance III HD y criosonda TCI de 3 mm
Experimento	NOESYPRESAT[1]	PRESAT[1]
Temperatura (K)	298	298
Número de barridos	128	128
Anchura espectral (ppm)	14	20
Puntos de datos	32	65
Tiempo de adquisición (s)	2,0	2,05
Tiempo de relajación (s)	5	3

[1]El término PRESAT hace referencia al experimento de pulso con supresión del agua

4.2.3 Análisis de los datos

4.2.3.1 Pre-tratamiento de los datos

Las señales o picos obtenidos en los espectros de RMN, especialmente los de ^{1}H, suelen sufrir desalineaciones debidas a diversos factores como variaciones de pH, fuerza iónica y temperatura, entre otros. Para corregir estas desalineaciones existen dos métodos que se emplean comúnmente en estudios de metabolómica. Estos métodos son el *bucketing* y el *peak alignment*. Mediante la aplicación de estos métodos (en conjunto o por separado) se robustecen

los posteriores análisis estadísticos y la consecuente interpretación de los resultados (Sousa et al. 2013).

Convencionalmente, el método de *bucketing* (también llamado *binning*) consiste en la segmentación de un espectro en pequeñas áreas llamadas *buckets* o *bins* en rangos variables que van de los 0,01 a los 0,05 ppm. Preferentemente, el tamaño de un *bucket* debería ser lo suficientemente grande como para que un pico determinado quede dentro de un mismo *bucket*, y al mismo tiempo no tan grande como para incluir picos de otros compuestos dentro de un mismo *bucket* (Vu & Laukens 2013). Las intensidades dentro de cada *bucket* se suman, de manera que el área bajo cada región espectral se utiliza en lugar de las intensidades individuales. Por lo tanto, un nuevo conjunto de variables (cada una como resultado de la suma de intensidades) es creado y utilizado posteriormente en los análisis estadísticos (Figura 4.4A).

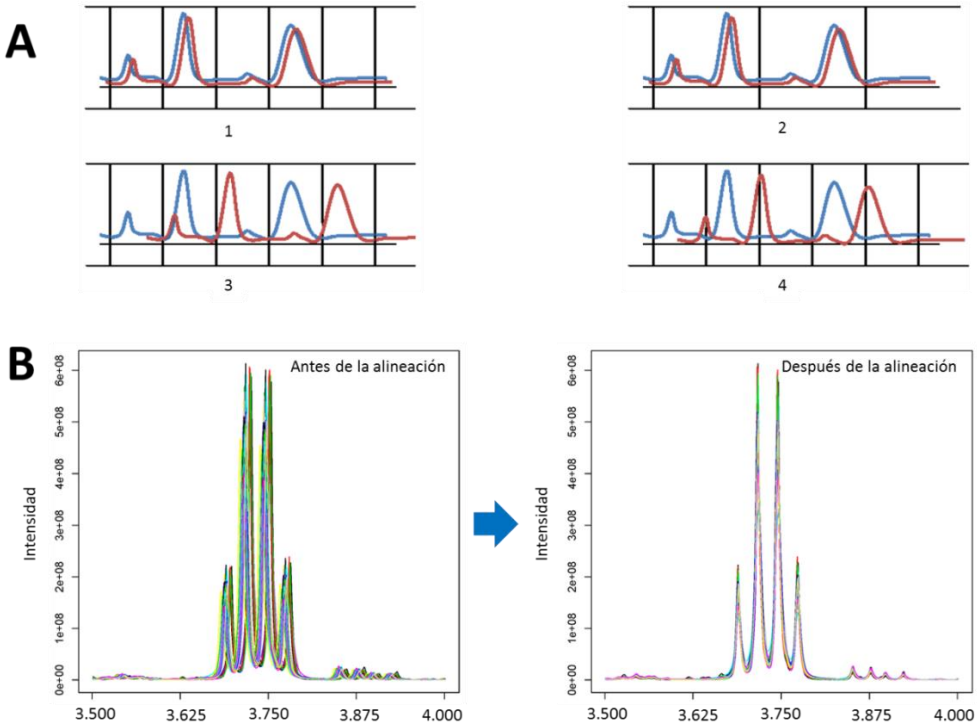


Figura 4.4 Ejemplos de cómo el *bucketing* (A) y el *peak aligment* (B) abordan la desalineación de distintas señales de la región espectral de RMN. Idealmente, mediante el *bucketing* cada pico debería quedar dentro un mismo *bucket* (A1); sin embargo, dependiendo de la anchura del *bucket* (A2) o la desalineación de las señales (A3 y A4), los picos podrían quedar mal distribuidos en los *buckets*. Adaptado de Vu & Laukens, 2013.

Por otro lado, mediante el *peak alignment*, la alineación de los picos se logra tomando cada segmento del espectro individualmente, desplazándolo lateralmente e interpolándolo linealmente para estirar o contraerse hasta obtener la mejor correlación con un segmento de espectro de referencia correspondiente (Forshed et al. 2003) (Figura 4.4B).

En la presente Tesis Doctoral los espectros obtenidos de las muestras de orina se segmentaron siguiendo el método de *bucketing*, fijando regiones de 0,005 ppm de anchura en el rango espectral de 0-10 ppm. Esta segmentación de los espectros y la integración de los picos se realizaron haciendo uso del programa ACD/NMR *Processor* 12.0 (*Advanced Chemistry Development*, Toronto, Canadá). En cuanto a los espectros de las muestras de plasma, se aplicó un algoritmo de *peak alignment*, según la función *reference picking* del paquete *speaq* (version 1.2.1) en el programa R 3.3.1.

4.2.3.2 Tratamiento de los datos

Los datos obtenidos a través de experimentos de metabolómica con una aproximación no dirigida suelen tener diferentes formas de variación que pueden ser atribuibles a los procesos biológicos estudiados o a factores del procedimiento analítico empleado (p.e., interferencias experimentales e instrumentales generadas durante la preparación de las muestras y la posterior adquisición de datos) (van den Berg et al. 2006). Por lo tanto, el objetivo del tratamiento de los datos es focalizar los posteriores análisis estadísticos en la información biológicamente relevante, minimizando al máximo las otras fuentes de variación, además de procurar obtener una distribución de los datos más adaptada a la curva de normalidad para su posterior comparación (Garcia-Aloy 2014; van den Berg et al. 2006).

Las siguientes operaciones (incluyendo métodos específicos) son algunas de las más comúnmente empleadas en el tratamiento de datos obtenidos por RMN en estudios de metabolómica, y particularmente aplicadas en el tratamiento de los datos obtenidos del análisis de las muestras de orina en la presente Tesis Doctoral. Dichas operaciones además se aplicaron en orden consecutivo y mediante el empleo de la plataforma web MetaboAnalyst 3.0 (Xia et al. 2015; Xia & Wishart 2016).

- Normalización por el método suma de filas (*row-wise*). La finalidad de la normalización es la de eliminar la variación no deseada entre las muestras con el fin de hacerlas comparables entre sí. Usualmente esta operación se emplea para minimizar o eliminar los efectos de la dilución, como es el caso de la orina (Craig et al. 2006). Específicamente, el método *row-wise* consiste en dividir la intensidad de cada variable (pico o *bin*) por la suma de las intensidades totales del espectro.

- Transformación por el método de logaritmo. Incluye una serie de conversiones para linealizar o cambiar la escala de los datos con el fin de corregir su heterocedasticidad (Goodacre et al. 2007). Dicho de otro modo, la transformación de los datos se utiliza para normalizar las variables y que éstas sigan una distribución normal.
- Escalado por el método Pareto. El objetivo del escalado es hacer las variables comparables entre sí, especialmente cuando en la matriz de datos existen variables que son muy diferentes en el orden de magnitud. Concretamente, el escalado consiste en dividir cada variable de cada metabolito de la matriz de datos por un factor escalar, el cual es diferente para cada variable y normalmente es una medida de la dispersión para esa variable. En el escalado por el método Pareto se divide cada variable por la raíz cuadrada de su desviación estándar, de este modo los grandes cambios son menos predominantes (van den Berg et al. 2006).

4.2.3.3 Análisis estadísticos

La multidimensionalidad de los datos es una de las principales características en un estudio metabolómico con aproximación no dirigida. Esta multidimensionalidad viene dada por una matriz de datos cuyo número de variables (columnas) suelen ser muy superior con respecto al número de muestras (filas), lo que la hace compleja y difícil de interpretar. En este sentido, la aplicación de métodos quimiométricos a través de análisis estadísticos univariantes y multivariantes resulta esencial en la reducción de la dimensionalidad y la extracción de información biológica relevante de los datos (Saccenti et al. 2014).

Las técnicas de análisis estadístico (así como los programas informáticos para su conducción) empleadas en los diferentes estudios que componen esta Tesis Doctoral se describen a continuación. Para todas estas pruebas el nivel de significancia fue fijado en $p < 0,05$.

4.2.3.3.1 Análisis univariantes

Las técnicas de análisis univariante incluyeron pruebas *t*-student y U de Mann-Whitney para comparar la media y mediana, respectivamente, de variables cuantitativas entre dos grupos. Específicamente, la prueba *t*-student se empleó cuando los datos cumplían con el supuesto de distribución normal (prueba paramétrica), y la prueba U de Mann-Whitney cuando este supuesto no se cumplía (prueba no paramétrica). De forma similar, pero para comparaciones entre más de dos grupos, se emplearon las pruebas ANOVA (paramétrica) y Kruskal-Wallis (no paramétrica).

Mientras tanto, la comparación de variables categóricas entre dos o más grupos fue realizada mediante pruebas *chi*-cuadrado (X^2).

Más concretamente, estas pruebas estadísticas se aplicaron en la comparación de: las intensidades de señales de los metabolitos identificados, los parámetros clínicos y variables sociodemográficas entre los grupos de participantes formados en cada estudio. Todas estas pruebas univariantes se llevaron a cabo con el programa estadístico SPSS® 21 (IBM Corp., Armonk, NY, USA).

Correcciones para comparaciones múltiples.

Las pruebas univariantes de datos metabolómicos con dos o más grupos suelen llevar el riesgo de arrojar valores de significancia estadística que en realidad sean falsos positivos (error de tipo 1, lo que significa no aceptar la hipótesis nula cuando esta es verdadera). Es decir, cuanto mayor sea el número de comparaciones, mayor la probabilidad de encontrar diferencias significativas que en realidad se deban al azar y no al efecto del fenómeno biológico estudiado. Por lo tanto, para reducir el número de falsos positivos en estas comparaciones se recurre a métodos de corrección o ajuste en la significancia. Específicamente, en las pruebas univariantes realizadas con las intensidades de los metabolitos identificados en esta Tesis Doctoral se empleó el método conocido como "tasa de descubrimiento falsos" (del inglés, *false discovery rate*, FDR). La aplicación de este método se realizó en el entorno del programa estadístico R 3.3.1.

4.2.3.3.2 Análisis multivariantes

A diferencia del análisis univariante, el análisis multivariante comprende la observación y análisis estadístico simultáneo de más de dos variables. Las técnicas multivariantes aplicadas en el análisis de los datos metabolómicos de la presente Tesis se describen a continuación.

- **Análisis de componentes principales (del inglés, *Principal Component Analysis*, PCA).**

El PCA es una técnica no supervisada que tiene como objetivo reducir la dimensionalidad de los datos mediante la realización de combinaciones lineales de las variables originales (correlacionadas entre sí) en nuevos conjuntos de variables más pequeños (no correlacionados entre sí) (Trygg et al. 2007). Las nuevas combinaciones se denominan componentes principales (PC) y se ordenan de forma decreciente según el porcentaje de la varianza total de los datos que explican. Así, por ejemplo, en el PC1 se describe la máxima varianza mientras que, en los componentes siguientes (PC2, PC3, etc.), el porcentaje de varianza explicada decrece (Yamamoto et al. 2014). Los resultados obtenidos de un PCA se visualizan mediante dos

gráficos de dispersión conocidos como *score plot* y *loading plot* (Figura 4.5). En el *score plot* se expresan los datos originales de las muestras (individuos) distribuidos en el nuevo espacio de coordenadas (PC), de modo que la variación sistemática puede ser interpretada en términos de similitudes o diferencias entre las muestras. Mientras tanto, en el *loading plot* se grafican los coeficientes de las transformaciones lineales de las variables originales que dan lugar a los nuevos PC y, por tanto, corresponden a los metabolitos analizados (o señales espectrales de RMN) en el proceso de proyección. En este sentido, la dirección en la cual se proyectan los puntos en el *loading plot* corresponde a la dirección de proyección de los puntos en el *score plot*, haciendo posible determinar qué variables son responsables de la agrupación de las muestras en el PCA (Trygg et al. 2007). Por último, un PCA suele ser sensible a los valores extremos o atípicos (del inglés, *outliers*), por ello muchas veces se utiliza este análisis como proceso exploratorio para detectarlos.

La técnica de PCA fue empleada en la presente Tesis para la visualización preliminar de los datos después de su tratamiento (sección 4.2.3.2) y la detección de posibles *outliers* en cada estudio. Los programas estadísticos empleados alternativamente para realización de esta técnica fueron SPSS® 21, y R 3.3.1.

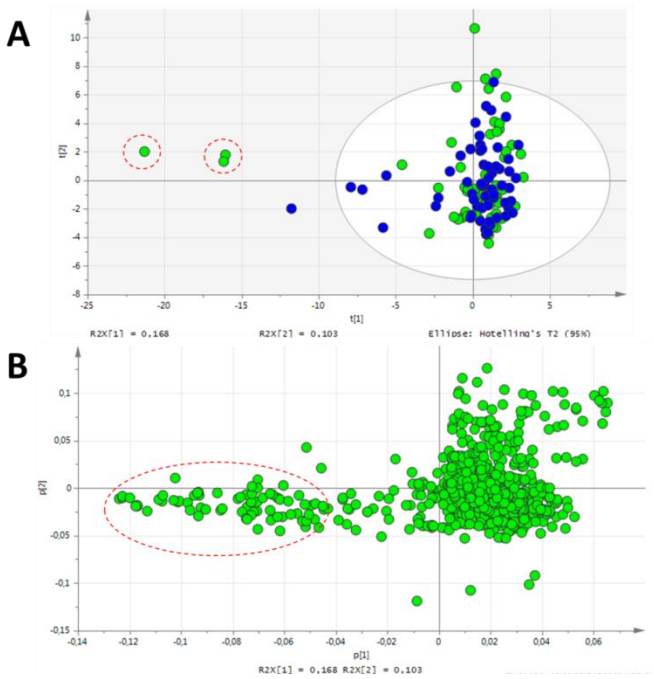


Figura 4.5 *Score plot* (a), *loading plot* (b) y detección de posibles *outliers* (círculos punteados en rojo) de un típico PCA. Los ejes X e Y en cada gráfico representan los componentes 1 y 2, respectivamente. Los dos colores en el *score plot* ejemplifican dos grupos de comparación en el PCA.

- **Regresión de mínimos cuadrados parciales (del inglés, *Partial Least Squares regression*, PLS) y análisis discriminante por mínimos cuadrados parciales (del inglés, *Partial Least Square Discriminant Analysis*, PLS-DA)**

La regresión por PLS es una técnica que reduce los predictores (variables independientes X o datos metabolómicos) a un conjunto más pequeño de componentes no correlacionados y realiza una regresión de mínimos cuadrados con estos componentes, en lugar de hacerlo con los datos originales (Westerhuis et al. 2008). De este modo, los nuevos componentes (también llamados variables latentes) son empleados para predecir la variable Y (variable continua) que, a diferencia del PCA, no maximizan la varianza explicada del conjunto de datos, sino la covarianza entre las variables X y la correspondiente variable dependiente Y (variable continua) (Alonso et al. 2015). Una variante de la regresión por PLS es la PLS-DA, la cual a diferencia del PLS se emplea para discriminar entre dos o más grupos en la variable de interés Y. Usualmente en el caso de un PLS-DA de dos clases, la variable Y se codifica en 0 y 1, donde cada número se corresponde a una de las clases. Tanto el análisis por PLS como por PLS-DA son técnicas de análisis supervisado ampliamente empleadas en estudios de metabolómica pues resultan potencialmente útiles cuando las variables predictoras están muy correlacionadas (colinealidad) o cuando el número de predictores es superior al número de casos (Gromski et al. 2015). Típicamente, los modelos obtenidos por PLS y PLS-DA generan gráficos de dispersión tipo *score plot* y *loading plot*, con las mismas propiedades que se describen para el PCA. Adicionalmente, mediante los modelos PLS y PLS-DA generan diferentes estadísticos que permiten por un lado determinar la calidad del modelo y por otro obtener las variables más importantes o de mayor peso en la predicción de la variable de interés. Dos de estos estadísticos son los coeficientes de determinación (R^2) y de predicción (Q^2). El coeficiente R^2 es un criterio que permite valorar la robustez del modelo ya que estima la bondad de ajuste o el porcentaje de variación de los resultados que puede explicarse por el modelo. Por su parte, el coeficiente Q^2 indica la fracción del total de la variación de X e Y que se puede predecir a partir de los nuevos componentes determinados. Los valores de R^2 y Q^2 se encuentran en el rango de 0 a 1, donde mientras más cercanos sean estos al 1, mayor será la capacidad discriminante y predictiva del modelo, respectivamente.

Los modelos PLS y PLS-DA suelen verse afectados negativamente por la variación sistemática presente en la matriz X que no está relacionada con la variable Y. Este hecho puede comportar algunos errores en la interpretación y tener importantes implicaciones en la selección de los marcadores discriminantes (Trygg et al. 2007). Considerando que en estudios observacionales la mayoría de las veces existen muchas fuentes de variación diferentes a la que se está estudiando (p.e., variabilidad interindividual debida a la variación fisiológica innata como la edad, el

género y la dieta, entre otros), resulta necesario aplicar filtros multivariantes de pre-procesamiento de los datos antes de proceder con la construcción de los modelos (Trygg et al. 2007). En este sentido, el filtro de corrección ortogonal de las señales (del inglés, *ortogonal signal correction*, OSC), por ejemplo, elimina de la matriz X toda la información no relacionada (ortogonal) con la variable Y (Dona et al. 2016).

En el desarrollo de los objetivos de la presente Tesis Doctoral, el análisis de regresión por PLS se empleó para estudiar la asociación de la pérdida de peso con los cambios producidos en el metaboloma como resultado de una intervención en el estilo de vida; mientras que el análisis por OSC-PLS-DA se empleó en los estudios de biomarcadores de DM2 y adherencia a la DietMed. En ambos estudios un OSC-PLS-DA se empleó para determinar inicialmente el perfil de metabolitos discriminantes entre individuos con DM2 y no-DM2, o entre individuos con A-ADM y B-ADM. En ambos tipos de modelo, PLS y OSC-PLS-DA, los estadísticos R^2 y Q^2 se emplearon como parámetros iniciales de calidad de los modelos, mientras que la validez de los mismos se evaluó mediante la aplicación de los algoritmos descritos en la sección "estrategias de validación de modelos". La conducción de los modelos PLS y OSC-PLS-DA se realizó con los programas estadísticos R 3.3.1, y SIMCA-P® 13.0 (Umetrics, Umeå, Suecia), respectivamente. En particular, los modelos PLS se construyeron con la función MUVR (del inglés, *Multivariate modelling with Unbiased Variable selection*) proporcionado por el paquete con el mismo nombre (MUVR 0.0.950)

- **Análisis de bosques aleatorios (del inglés, *Random Forest*, RF)**

La técnica RF es un algoritmo que consiste en la creación de diversos árboles de clasificación para determinar y/o predecir la pertenencia de las observaciones (individuos) a los distintos grupos o clases de interés (Breiman 2001). El algoritmo general de un análisis por RF inicia con la partición de los datos originales en numerosos subgrupos mediante una técnica de re-muestreo o *bootstrapping* con reemplazamiento, la cual consiste en la formación de diferentes grupos de datos elegidos de forma aleatoria. Este tipo de muestreo permite la construcción de dos tipos de grupos llamados grupos de entrenamiento (del inglés, *training sets*; que en conjunto incluyen el 63,2% de todas las muestras) y grupos de prueba (del inglés, *test sets*; que incluyen el 36,8% restante de todas las muestras) (Etapa 1). Los *training sets* son usados para construir los árboles, mientras que los *test sets* son usados para estimar la precisión de la clasificación. En los *training set* cada punto de división del árbol o nodo emplea un subconjunto pequeño de variables elegidas al azar para determinar aquellas que permitan la mejor clasificación de los datos (Etapa 2). En esta etapa se generan tantos árboles de decisión como *training sets* se hayan

formado, y de cada árbol se obtienen las probabilidades de clasificación en cada grupo según el subconjunto de datos empleado. Una vez finalizado el algoritmo en los *training set* se realiza la evaluación de los árboles generados mediante la introducción de nuevos casos a través de los *test sets* (Etapa 3). De este modo la categoría final de la clase (clasificación) es asignada por el voto mayoritario del conjunto árboles en el *test set* (Etapa 4) (Gromski et al. 2015). La figura 4.6 ilustra la implementación del algoritmo general de RF en un conjunto de datos tan grande como los obtenidos en estudios de metabolómica.

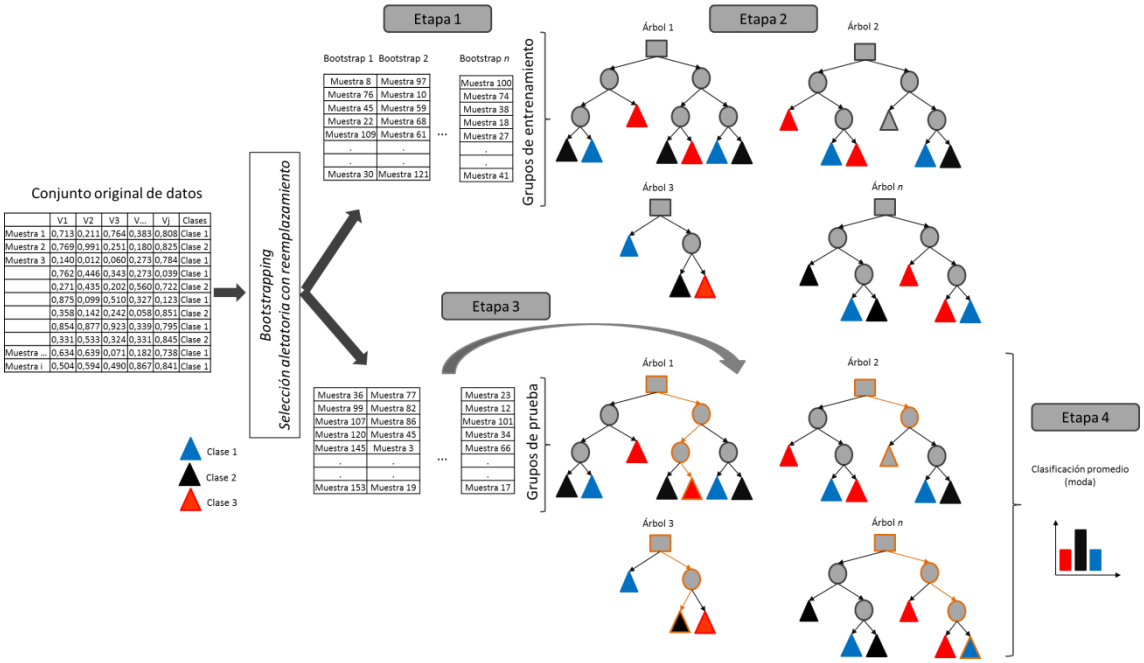


Figura 4.6 Esquema general del algoritmo *random forest*. En la primera etapa el conjunto original de datos es dividido en grupos de entrenamiento y prueba usando la técnica de *bootstrapping* (selección aleatoria con reemplazamiento). En la segunda etapa los grupos de entrenamiento son usados para construir un gran número de árboles de decisión no correlacionados que serán usados más tarde en la etapa 3. Esta misma etapa comienza con los nodos raíz (cuadrados), en los cuales diferentes subgrupos de variables de entrada son seleccionados aleatoriamente (usualmente de acuerdo a la raíz cuadrada del número total de variables). Este proceso permite optimizar las divisiones en los árboles dentro del "bosque". Los círculos corresponden a los nodos internos donde las muestras son divididas en base a los diferentes valores de las variables. Los triángulos corresponden a las hojas de los nodos, en las cuales las clases o grupos son reconocidos. En la etapa 3 cada uno de los árboles generados en los grupos de entrenamiento son probados en los grupos de prueba (líneas en naranja) y la frecuencia de clasificación en cada una de las clases o grupos es determinada. Finalmente, en la etapa 4, el algoritmo RF genera una distribución de probabilidad de las distintas clasificaciones obtenidas y determina por la máxima probabilidad la clasificación final. Fuente (Gromski et al. 2015).

El algoritmo de RF presenta las ventajas de tratar con grandes conjuntos de datos sin eliminación de variables, ser capaz de manejar valores perdidos, no requiere el escalado previo de los datos y suele ser robusto a los *ouliers* (Breiman 2001; Gromski et al. 2015).

Los modelos de clasificación por RF pueden ser visualizados a través de gráficos de dispersión como el que se detalla en la figura 4.7. Además, de forma similar a los modelos anteriormente explicados, la calidad de los modelos por RF suele evaluarse a través de diferentes parámetros estadísticos que incluyen entre otros el número de clasificaciones erróneas (del inglés, *misclassifications*). El número de clasificaciones erróneas se expresa en porcentaje y representa la proporción de muestras (individuos) clasificadas de manera errónea de acuerdo a la clase o grupo de pertenencia original. Por lo tanto, cuanto menor sea el número de clasificaciones erróneas (idealmente < 30%), mejor la calidad del modelo RF.

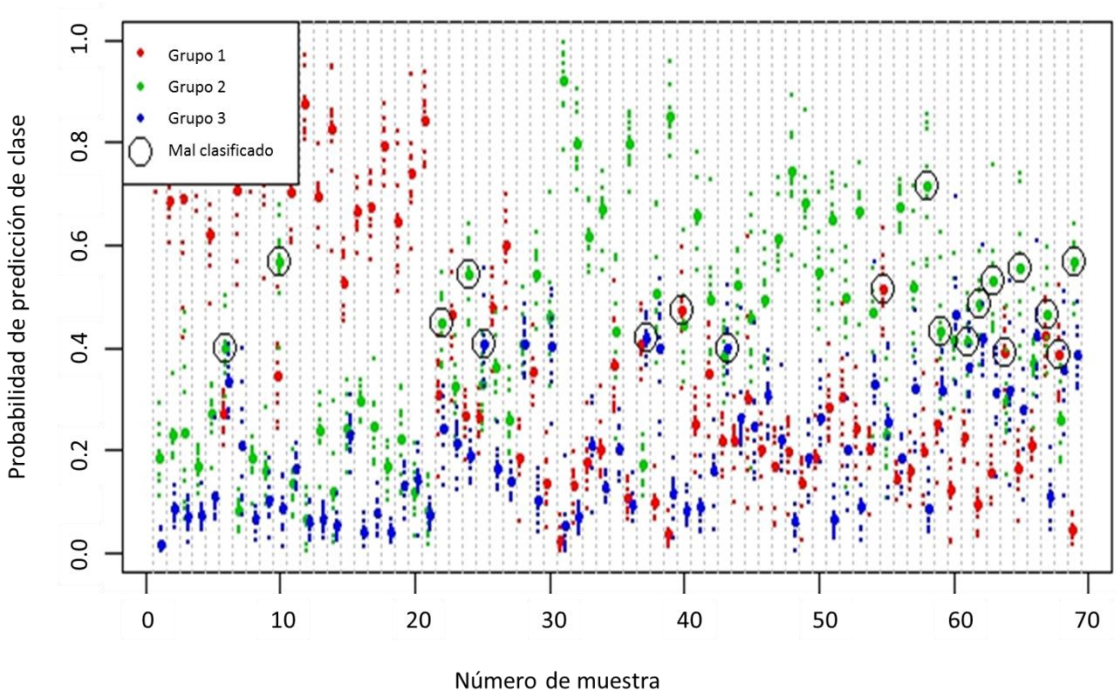


Figura 4.7. Visualización de la probabilidad de clasificación de las muestras en cada uno de los grupos de interés, a través del algoritmo *random forest* (RF). Cada punto (indistintamente del tamaño) en cada columna representa la probabilidad asignada por cada modelo RF a cada muestra de ser clasificada en cada uno de los grupos, según el color asignado. Los puntos de mayor tamaño indican la probabilidad promedio de pertenecer a cada grupo (según el color). De este modo, la probabilidad más alta de cada muestra de pertenecer a un grupo según el modelo es contrastada con su clasificación original, determinando así la capacidad del modelo para clasificar de manera correcta o incorrecta (puntos rodeados por círculos) las muestras.

La aplicación del algoritmo RF en la presente Tesis se realizó en el marco de la determinación de los cambios producidos en el metaboloma plasmático como respuesta a una intervención en el estilo de vida para la pérdida de peso. Más específicamente, se calcularon modelos RF para la clasificación de los individuos en este estudio de acuerdo al grupo de intervención (control *vs* tratamiento) en dos periodos de tiempo (3 y 12 meses). La conducción de estos modelos se llevó a cabo en el programa estadístico R 3.3.1, mediante la función MUVR, y usando el número de clasificaciones erróneas como parámetro de calidad clasificatoria.

- **Análisis de conglomerados (del inglés, *cluster analysis* o *clustering*)**

Esta técnica estadística multivariante incluye numerosos algoritmos enfocados a la división de las observaciones de un conjunto de datos en grupos o clases (del inglés, *clusters*) de modo tal que las distancias entre las observaciones dentro un mismo grupo sean menores con respecto a las distancias entre observaciones de grupos diferentes (Blekherman et al. 2011). Estas distancias son calculadas a través de una función de distancia como medida de la disimilitud entre las observaciones dentro y entre los grupos, siendo la distancia euclidiana la más comúnmente empleada. Típicamente, los algoritmos de *clustering* son no supervisados y pueden clasificarse en no jerárquicos y jerárquicos, dependiendo de si el número de grupos a formar se ha predefinido o no, respectivamente.

De este modo, con el objetivo de clasificar a un grupo de individuos diabéticos en fenotipos metabólicos relacionados con la diabetes, dentro de la presente Tesis Doctoral, se llevó a cabo un análisis de *clustering* mediante el algoritmo conocido como k-medias (del inglés, *k-means*). Brevemente, este algoritmo comienza especificando el número de grupos (*k clusters*) que se han de formar. A continuación, y de forma aleatoria, los individuos son asignados a cada uno de estos grupos, calculando a su vez la distancia entre estos y los llamados centroides de cada grupo. Posteriormente los individuos son reasignados al grupo con el centroide más cercano. Por último, los dos últimos pasos se repiten hasta que no se produzca ninguna reasignación más, es decir, hasta que los individuos se estabilizan en alguno de los grupos.

La ejecución del algoritmo de agrupación por *k-means* se realizó en la plataforma web MetaboAnalyst 3.0, introduciendo los metabolitos discriminantes que definieron el estado diabético y fijando k=2 centroides en cada grupo de diabéticos y no diabéticos. Este número se fijó debido al número limitado de individuos por grupo, ya que al reducir significativamente su tamaño podría perderse variabilidad y potencia estadística. El resultado de este análisis arrojó la clasificación de dos fenotipos metabólicos en cada grupo en base a la diferencia de abundancia

de los metabolitos en cada uno. La visualización de los cuatro fenotipos metabólicos resultantes se logró a través de un gráfico conocido como mapa de calor (del inglés, *heatmap*). La construcción de este gráfico se realizó también en la plataforma MetaboAnalyst 3.0.

- **Estrategias de validación de modelos**

La conducción de modelos predictivos conlleva el riesgo de que se produzca un sobreajuste (del inglés, *overfitting*) de los mismos. Esto es, el modelo se ajusta muy bien a los datos existentes pero tiene un pobre rendimiento para predecir nuevos resultados. Por lo tanto, la validez de un modelo requiere la realización de diferentes pruebas como la validación cruzada y de permutaciones (Westerhuis et al. 2008), y de este modo confirmar o garantizar la robustez de un modelo. Las siguientes pruebas y sus variantes se emplearon para validar los modelos predictivos de los diferentes estudios que componen esta Tesis Doctoral.

En las pruebas de validación cruzada, la matriz de datos se divide en un grupo de entrenamiento (*training set*) y un grupo de validación (del inglés, *validation set*). Mientras que en el *validation set* el modelo es creado, en el *validation set* éste es validado. Este procedimiento se repite de manera iterativa hasta que todas las muestras han sido incluidas en el *validation set* en una única ocasión, por lo que entonces se considera que el error de predicción calculado será representativo para nuevas muestras (Westerhuis et al. 2008). Este procedimiento de validación es el que se sigue de manera interna en los modelos para generar el estadístico Q^2. Una variante más robusta de este tipo de pruebas (especialmente para pequeños grupos de datos) es la doble validación cruzada repetida (del inglés, *repeated double cross validation*, rdCV). Inicialmente, en la rdCV el conjunto total de los datos es divididos en dos grupos llamados grupo de calibración (del inglés, *calibration set*) y grupo de prueba (*test set)*. Luego el *test set* es dividido en dos nuevos grupos: *training set* y *validation set*, entre los cuales se lleva a cabo una validación cruzada simple, como se describió anteriormente. Esta primera validación cruzada generará un número óptimo de componentes que servirán para construir un modelo final de clasificación que posteriormente será probado en *test set* (segunda validación cruzada). Este procedimiento entonces se repite hasta que todos los individuos han pasado por el *test set* una única vez (Filzmoser et al. 2009; Westerhuis et al. 2008). La selección de las muestras de validación se ha de realizar de forma aleatoria de modo que se aseguren diferentes combinaciones de *validation sets* y *training sets* para cada *test set*. De esta manera, el modelo final es construido en ausencia total de las muestras contenidas en el *test set* y así la predicción es independiente de la optimización del modelo.

El algoritmo descrito para las pruebas de rdCV fue implementado en el desarrollo de los modelos PLS y RF como primera prueba de su robustez mediante la función MUVR en el programa estadístico R 3.3.1.

En las pruebas de permutación, la hipótesis nula a ser probada o refutada es que el modelo óptimo encontrado durante el proceso de clasificación/predicción también podría ser encontrado si cada muestra es reasignada a otra categoría escogida al azar en la misma proporción que la clasificación original (Westerhuis et al. 2008). De este modo, durante una prueba de permutación, la estructura del modelo original y el subconjunto de variables que lo componen se mantienen fijas mientras que múltiples modelos "permutados al azar" son evaluados. Como resultado, se genera una distribución de referencia de la hipótesis nula y el rendimiento (en términos de R^2, Q^2, o número de clasificaciones erróneas) del modelo original se compara estadísticamente con esta distribución, fijando un valor de significancia *p* para rechazar o aceptar dicha hipótesis (Xia et al. 2013).

Las pruebas de permutación fueron empleadas para validar los modelos PLS, OSC-PLS-DA y RF descritos en los apartados anteriores. Valores de R^2<0,04 y Q^2 <0,05, a partir de los modelos permutados, se consideraron adecuados para validar los modelos originales (van der Voet 1994). Adicionalmente, para los modelos PLS y RF, la distribución de frecuencias de los estadísticos Q^2 y número de clasificaciones erróneas de los modelos permutados se contrastaron con los estadísticos de los modelos originales mediante pruebas *t*-student. El número de permutaciones para todos los modelos se fijó en 200 y los programas para su realización fueron SIMCA-P® (modelos OSC-PLS-DA) y R 3.3.1 (modelos PLS y RF).

- **Selección de variables discriminantes**

Después de comprobar la validez del modelo, el siguiente paso en el análisis multivariante de los datos metabolómicos es la selección y obtención de las variables relevantes o discriminantes. En este sentido, la selección de variables de acuerdo a sus valores VIP (del inglés, *Variable Importance In Projection*) es uno de los criterios más comúnmente empleados en modelos multivariantes de estudios de metabolómica. Un valor VIP es una medida de la importancia o la contribución de una variable al modelo. Concretamente es una suma ponderada de los cuadrados del *loading* que tiene en cuenta la cantidad de varianza explicada de cada componente (Xia et al. 2013; Xia & Wishart 2016). Así, cuanto mayor sea el valor VIP de una variable, mayor su contribución en modelo y/o mayor su capacidad discriminante. En la práctica, la mayoría de los estudios de metabolómica aceptan como discriminantes a las variables con valores VIP iguales o superiores a 1,0 (Feng et al. 2016).

De este modo, la obtención de las variables relevantes o discriminantes a partir de los distintos modelos generados en los estudios de esta Tesis Doctoral se realizó mediante la selección de las mismas en base a sus valores VIP. En el caso de los modelos OSC-PLS-DA, el criterio específico de selección fue variables dentro del componente no ortogonal con valores VIP$\geq$1,0, mientras que para los modelos PLS y RF esta selección se realizó considerando el 80% de las variables más significativas o con valores VIP más altos en cada modelo. Este último criterio se deriva del argumento *varRatio* de la función MUVR, mediante el cual tras cada iteración de rdCV de los modelos PLS y RF se selecciona el 80% de las variables más significativas, acortando así la lista de variables en cada iteración.

- **Regresión logística binaria y curvas ROC (del inglés, *Receiver Operating Characteristic*)**

Este tipo de regresión resulta útil para los casos en los que se desea predecir la presencia o ausencia de una característica o resultado según los valores de un conjunto de variables predictoras. Aunque es similar a un modelo de regresión lineal, la regresión logística binaria está adaptada para modelos en los que la variable dependiente es dicotómica. Los coeficientes de regresión logística pueden utilizarse para estimar la *odds ratio* de cada variable independiente del modelo. Para la construcción del modelo las variables son introducidas una a una o todas a la vez siguiendo distintos métodos. Uno de estos métodos es el de selección por pasos hacia adelante (del inglés, *forward selection stepwise*). Este método consiste en la inclusión de una variable a la vez en el modelo y que previamente ha cumplido con una serie de criterios de ajuste prefijados. De este modo, el proceso de inclusión de variables en el modelo continúa sucesivamente hasta que no existen más variables que cumplan con el criterio de ajuste fijados. La lógica subyacente de este método consiste en conservar las variables independientes que contienen información relevante y a la vez prescindir de aquellas que resulten redundantes respecto de las que quedaron en el modelo. Se trata de un procedimiento de índole exclusivamente estadístico, que discurre según algoritmos programables en los que, una vez elegido el conjunto inicial de variables, no intervienen los juicios teóricos del investigador. La regresión logística binaria suele emplearse conjuntamente con el análisis de curvas ROC en estudios de metabolómica para el desarrollo y evaluación de biomarcadores compuestos por dos o más metabolitos, respectivamente.

Los gráficos de curvas ROC muestran el poder predictivo y de discriminación del modelo generado en la regresión logística, y por lo tanto, son útiles para visualizar, organizar y seleccionar los clasificadores. La forma de cuantificar el poder discriminatorio del modelo es a través de la determinación de los siguientes parámetros: sensibilidad, especificidad y área bajo

la curva (del inglés, *Area Under the Curve*, AUC). La sensibilidad es la capacidad que posee el modelo para identificar correctamente los casos positivos (en el modelo Y=1). Es decir, a la probabilidad de que un sujeto en este caso consumidor sea clasificado correctamente como tal, mientras que la especificidad se refiere a la capacidad que posee el modelo para clasificar correctamente a los casos negativos (en el modelo Y=0). La relación de la sensibilidad y especificidad se describe mejor a través de las siguientes funciones, usando como ejemplo la clasificación de individuos enfermos y sanos:

$$\text{Sensibilidad} = VP/(VP + FN) \quad \text{Especificidad} = VN/(VN + FP)$$

Dónde:

VP = el número de individuos sujetos enfermos que se identifican correctamente como enfermos (resultado positivo y prueba positiva).

FN = el número de individuos sanos que se identifican correctamente como sanos (resultado negativo y prueba negativa).

FP = el número de individuos sanos que se identifican incorrectamente como enfermos (resultado negativo y prueba positiva).

VN = El número de individuos enfermos que se identifican incorrectamente como sanos (resultado positivo y prueba negativa).

Se entiende por lo tanto que cuanta mayor sea la sensibilidad y especificidad, mayor será la capacidad del modelo para discriminar entre los niveles de la variable dependiente. Mientras tanto, los valores de la AUC suelen oscilar entre 0 y 1, teniendo en cuenta la siguiente escala para determinar la calidad del modelo analizado: 0,9-1,0, excelente; 0,8-0,9, buena; 0,7-0,8, regular; 0,6-0,7, pobre; y 0,5-0,6, mala (Xia et al. 2013).En la figura 4.8 se muestra un ejemplo del típico gráfico generado por el análisis de curvas ROC y las diferentes capacidades predictivas.

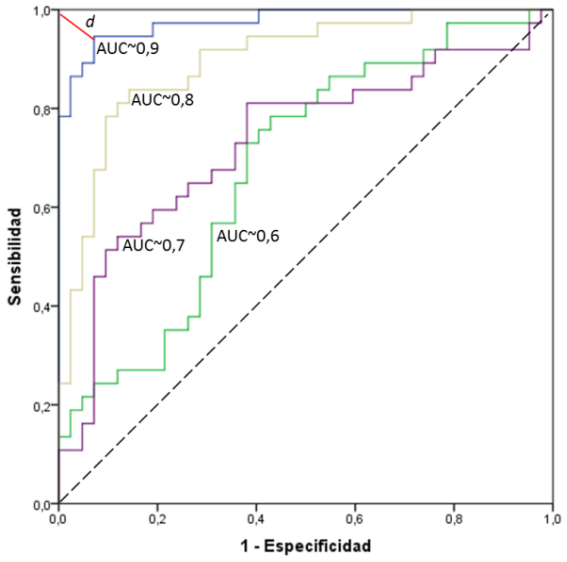


Figura 4.8 Ejemplo de curvas ROC con diferente capacidad predictiva (según se indica en el texto). En el eje de las abscisas se indica el valor resultante del cálculo 1-especificidad (lo que da una idea de los falsos positivos), mientras que en el eje de las ordenadas se indica la sensibilidad (la proporción de positivos reales). Así, la curva ROC se construye considerando diferentes puntos de corte que corresponden a las diferentes parejas de valores 1-especificidad y sensibilidad.

En el marco de los estudios de biomarcadores de diabetes y adherencia a la dieta Mediterránea, el empleo de la regresión logística binaria por el método de anteriormente descrito y el análisis de curvas ROC se realizó de la siguiente manera: Como paso inicial y para asegurar la robustez de los resultados, los conjuntos de datos se dividieron en *training set* y *validation set* de modo similar a como se ha mencionado en los apartados anteriores. Cada *training set* estuvo compuesto por dos terceras partes del conjunto total de los datos (~66,6%) y cada *validation set* por el resto (~33,3). La regresión logística se utilizó entonces para la construcción de un modelo de metabolitos combinados o multimetabolito con capacidad discriminante del fenómeno estudiado a partir del *training set*. Las variables de partida para cada modelo fueron los metabolitos identificados a partir de variables seleccionadas como discriminantes y significativas de los respectivos modelos OSC-PLS-DA. Posteriormente, el modelo multimetabolito y los metabolitos de forma individual fueron sometidos al análisis de curvas ROC para la obtención y comparación de los parámetros AUC, sensibilidad y especificidad. Éste último análisis se realizó tanto en el *training set* como en el *validation set*, aplicando en éste último el modelo formulado en el *training set*. El cálculo de la sensibilidad y especificidad se realizó considerando la distancia mínima (*d*) de la curva respecto al ángulo superior izquierdo (Figura 4.8). La conducción de la regresión logística binaria y los análisis de curvas ROC se realizó con el programa paquete SPSS® 21.

- **Análisis de correlaciones**

La finalidad de la correlación es examinar la dirección y la magnitud de la asociación entre dos variables, es decir, cómo varían conjuntamente. Por lo tanto, las correlaciones se consideran positivas cuando ambas variables se expresan en la misma dirección (p.e., ambas aumentan o ambas disminuyen) y negativas cuando la expresión de una de las variables es opuesta a la de la otra (p.e., una aumenta mientras la otra disminuye). Los dos coeficientes de correlación (r) con mayor frecuencia de aplicación en la estadística de estudios de metabolómica son: el de Pearson (paramétrico, basado en la media) y el de Spearman (no paramétrico, basado en la suma de rangos o mediana). Indistintamente del tipo de correlación, los coeficientes varían de -1,0 a 1,0, teniendo en cuenta que para: $r<0$, la correlación negativa; $r=0$, la correlación no existe; y para $r>0$ correlación positiva. Además, el grado de correlación puede considerarse como débil si $|r|<0{,}3$; moderada si $0{,}3 \leq |r| \leq 0{,}7$; o fuerte si $|r| > 0{,}7$ (Scalbert et al. 2014).

De acuerdo a los diferentes estudios que componen la presente Tesis, el análisis de correlación por los métodos citados se aplicaron para: i) la identificación de metabolitos (como se describe en la sección 4.2.4), ii) establecer asociaciones entre metabolitos e ingesta de alimentos, y iii) determinar la asociación entre metabolitos y la pérdida de peso. La realización de estas correlaciones se realizó mediante el uso de los programas SPSS® 21, (incisos i y ii) y R 3.3.1 (inciso iii).

4.2.4 Identificación de metabolitos

La identificación de metabolitos en todos los estudios que componen la presente Tesis se realizó mediante la aplicación y combinación de las siguientes estrategias:

i. Análisis de correlación entre las señales de RMN significativas de los modelos para la determinación de la multiplicidad de un mismo metabolito. Específicamente, en esta estrategia para la identificación de metabolitos a partir de las muestras de plasma se empleó la técnica STOCSY (del inglés, *Statistical Total Correlation Spectroscopy*), aplicando la función statNMR, del paquete NMRtools, versión 0.0.912 en el programa R 3.3.1. Mientras que para la identificación de metabolitos en las muestras de orina se calcularon los coeficientes de correlación de Pearson y Spearman mediante el programa SPSS® 21, usando la matriz de datos normalizados en la plataforma MetaboAnalyst 3.0.

ii. Uso del programa de identificación Chenomx NMR Suite 8.2 profiler (Chenomx Edmonton, Canadá). Chenomx es un programa de *spiking* o fortificación virtual que permite la identificación de compuestos en espectros de una dimensión (1D). Contiene una biblioteca de referencia espectral interna con más de 260 compuestos teniendo en cuenta el pH y la frecuencia de trabajo del equipo utilizado.

iii. Búsqueda de coincidencia de los desplazamientos químicos, multiplicidad y constantes de acoplamiento en bases de datos espectrales de RMN de libre uso como: *Human Metabolome Database* (HMDB, www.hmdb.ca), *Biological Magnetic Resonance Data Bank* (BMRB, www.bmrb.wisc.edu) y *Madison Metabolomics Consortium Database* (MMCD, mmcd.nmrfam.wisc.edu).

iv. Búsqueda de coincidencia de los desplazamientos químicos, multiplicidad y constantes de acoplamiento en la literatura científica, especialmente en publicaciones relacionadas con los estudios de la presente Tesis.

4.2.5 Interpretación biológica

Esta es la última etapa del *workflow* de la metabolómica. Con la interpretación biológica se complementan las respuestas a la pregunta de investigación en cada estudio, bien sea por la descripción biológica de los fenómenos estudiados o por la generación de nuevas hipótesis. A continuación, se describen las estrategias y herramientas seguidas para la interpretación biológica de los estudios aquí presentados.

i) Consulta de literatura científica, concretamente publicaciones relacionadas con el tema o los resultados obtenidos de cada estudio, para profundizar en el conocimiento de las distintas rutas y procesos bioquímicos en los que los metabolitos se ven involucrados.

ii) Consulta del papel biológico de los metabolitos en bases de datos como: HMDB, FooDB, (del inglés, *food* database) y KEGG (del inglés, *Kyoto Encyclopedia of Genes and Genomes*).

iii) Construcción de redes, listas y gráficos de enriquecimiento metabólico a través de plataformas informáticas de uso libre y comercial. Estas plataformas fueron MetaboAnalyst 3.0 (http://www.metaboanalyst.ca) y MetaCore® (GeneGo, Thomson

Reuters, St. Joseph, MI, USA). Dentro de la plataforma MetaboAnalyst se usó el módulo *Pathway Analysis* y la librería *Homo sapiens* para la construcción de un gráfico y lista de enriquecimiento metabólico. Mientras que para la plataforma MetaCore®, se usaron los módulos *one-click analysis/Metabolic networks (endogenous)* y *build network/for your experimental data* para la construcción de redes y listas de enriquecimiento metabólico. El uso de estas plataformas fue especialmente útil en los estudios de biomarcadores de diabetes y adherencia a la dieta Mediterránea, y el *input* de las presentes plataformas fueron las listas de metabolitos discriminantes de cada estudio.

V. RESULTADOS

5.1 BIOMARCADORES METABOLOMICOS Y METABOTIPOS DE DIABETES MELLITUS TIPO 2: ANÁLISIS TRANSVERSAL EN EL ESTUDIO PREDIMED

Los resultados del presente estudio se encuentran en vías de publicación en una de las revistas del área.

Mireia Urpi-Sarda[#], Enrique Almanza Aguilera[#], Rafael Llorach, Rosa Vázquez-Fresno, Ramon Estruch, Dolores Corella, Jose Vicente Sorli, Francesc Carmona, Alex Sanchez-Pla, Jordi Salas-Salvadó, and Cristina Andres-Lacueva. Nontargeted metabolomics biomarkers and metabotypes of type 2 diabetes prevalence: a cross-sectional study in the Predimed trial. [*In process of publication*] [#]*Equal contribution*.

Resumen

Los objetivos del presente estudio fueron: i) caracterizar la huella metabólica urinaria asociada a la diabetes mellitus tipo 2 en una población de elevado riesgo cardiovascular; ii) calcular un biomarcador metabolómico con una alta capacidad predictiva de diabetes prevalente, a partir de la huella metabólica asociada a la misma; y iii) definir fenotipos metabólicos dentro de la población diabética a partir de la huella metabólica asociada a la diabetes. El estudio metabolómico con aproximación no dirigida mediante RMN consistió en el análisis de muestras de orina de 85 individuos diagnosticados con DM2 y 69 individuos sin DM2 (no-DM2) del estudio PREDIMED. Los análisis multivariantes incluyeron la formulación de modelos OSC-PLS-DA y de regresión logística, curvas ROC y agrupación por *k-means*, según se describe en el apartado de metodología. La huella metabolómica asociada a la DM2 estuvo compuesta por 33 metabolitos. El biomarcador de DM2 estuvo compuesto por altos niveles de ácido metilsuccínico, alanina, dimetilglicina y ácido guanidinoacético, y bajos niveles de glutamina, metilguanidina, y ácidos 3-hidroximandelico e hipúrico. Este biomarcador tuvo AUC de 96,4%, la cual fue superior a la de los metabolitos individuales incluidos en el biomarcador y a la de la glucosa (no incluida). Se identificaron dos metabotipos asociados a la DM2, uno de ellos caracterizado por tener niveles plasmáticos más altos de glucosa, y de fenilalanina, fenilacetilglutamina, *p*-cresol y ácido acetoacético en orina. La huella metabolómica y el biomarcador de DM2 reflejan que las principales alteraciones metabólicas asociadas a la DM2 incluyen al metabolismo de aminoácidos, carbohidratos y microbiota. El metabotipo asociado a DM2 con niveles más altos de glucosa plasmática y de otros metabolitos urinarios sugiere que este podría estar relacionado con mayor riesgo cardiovascular a largo plazo. Este estudio refuerza el uso de metabolómica y metabotipos como herramientas para comprender la complejidad metabólica y la variación interindividual de la DM2, lo que podría ser útil para desarrollar una nutrición y medicina estratificadas y precisas.

5.1.2 Resultados

5.1.2.1 Características de los participantes

Las características de los inviduos que participaron en este estudio se muestran en la Tabla 5.1. De manera general los participantes tuvieron una edad promedio (± desviación estándar) de 67 ± 6 años; un tercio de la población fueron hombres; alrededor del 55% fueron diabéticos y un 47% presententaban obesidad. Al comparar las características demográficas y otras variables clínicas de riesgo cardiovascular entre los dos grupos, se observó que sólo la circunferencia de cintura, niveles de glucosa plasmática y la ingesta de medicamentos antidiabéticos fueron superiores en el grupo de inviduos diabéticos. En cuanto al resto de las variables, incluido el *score* de adherencia a la DietMed, los grupos fueron similares entre sí, no observándose más diferencias que las mencionadas anteriormente.

5.1.2.2 Validación del modelo discriminante de DM2

El análisis OSC-PLS-DA que fue conducido para caracterizar la huella metabólica discriminantes entre los grupos DM2 y no-DM2 generó un modelo compuesto por un componente latente y estadísticos R^2Y (cum) y Q^2Y (cum) con valores de 0,815 y 0,737, respectivamente; indicando de este modo una buena capacidad para clasificar a los individuos de acuerdo su estatus relacionado con la DM2 (Figura 5.1). Adicionalmente, mediante un test de permutaciones (n=200) se obtuvieron nuevos estadísticos R^2Y y Q^2Y con valores de 0,226 y -0,139, respectivamente, lo que permitió validar el modelo anterior (Figura 5.2).

5.1.2.3 Identificación de metabolitos

Tomando en cuenta las variables con valores VIP$\geq$1.0, a partir del modelo OSC-PLS-DA, y llevando a cabo una prueba de *t*-student de estas variables entre los grupos DM2 y no-DM2, se logró identificar un total de 33 metabolitos (Tabla 5.2). Diecisiete de estos metabolitos, incluida la glucosa, fueron encontrados en niveles más altos en los individuos con DM2, mientras que el resto de los metabolitos se observaron en niveles inferiores en comparación con los individuos del grupo no-DM2.

Tabla 5.1 Características de los participantes de acuerdo a su condidicón relacionada con la DM2

	Total (n=154)	Individuos DM2 (n=85)	Individuos no-DM2 (n=69)	*P*-valor
Edad (años)	67 ± 5,70	67 ± 6,00	67 ± 5,50	0,67
Hombre, n (%)	48 (31)	30 (35)	18 (22)	0,15
Fumadores actuales, n (%)	20 (13)	11 (13)	9 (13)	0,58
IMC, kg/m^2	30,2 ± 4,90	30,4 ± 4,60	30,0 ± 3,92	0,50
Obesidad (IMC≥25 kg/m^2), n (%)	71 (47)	37 (44)	34 (49)	0,29
Peso, kg	76,0 ± 12,7	77,7 ± 13,3	73,9 ± 11,6	0,06
CC, cm	101 ± 11,4	104 ± 11,3	98,3 ± 10,8	<0,01
Actividad física, MET·min/d	261 ± 268	241 ± 285	286 ± 245	0,30
PAS, mmHg	144 ± 19,7	141 ± 18,0	148 ± 21,2	0,08
PAD, mmHg	81,9 ± 10,2	81,1 ± 10,5	83,0 ± 9,95	0,38
Score de adherencia a la DietMed	8,50 ± 1,80	8,38 ± 1,80	8,70 ± 1,80	0,19
Biomarcadores plasmáticos				
Glucose, mg/dl	128,2 ± 48,5	159 ± 45,3	89,9 ± 7,20	<0,001
Colesterol HDL, mg/dl	55,1 ± 14,4	51,5 ± 14,4	58,4 ± 13,9	0,09
Colesterol LDL, mg/dl	132 ± 28,8	123,3 ± 29,7	139 ± 26,5	0,06
Colesterol total, mg/dl	214 ± 14,4	207 ± 35,0	218,8 ± 32,9	0,21
Triglicéridos, mg/dl	150 ± 114	177 ± 153	125,8 ± 49,3	0,13
Medicación, n ± (%)				
Antihipertensivos	65 (42)	32 (38)	33 (48)	0,13
Hipolipemiantes	85 (55)	42 (49)	43 (62)	0,07
Antidiabéticos orales	65 (42)	65 (76)	0	<0,01
Insulina	18 (12)	18 (21)	0	<0,01

Los datos se muestran como la media ± desviación estándar, para variables continuas; o como el número (n) y porcentaje (%) para variables categóricas, según se indica. Abreviaciones: CC, circunferencia de cintura; HDL, lipoproteína de alta densidad; IMC, índice de masa corporal; LDL, lipoproteína de baja densidad; MET, equivalente metabólico; DietMed, dieta Mediterránea PAD, presión arterial diastólica; PAS, presión aterial sistólica.

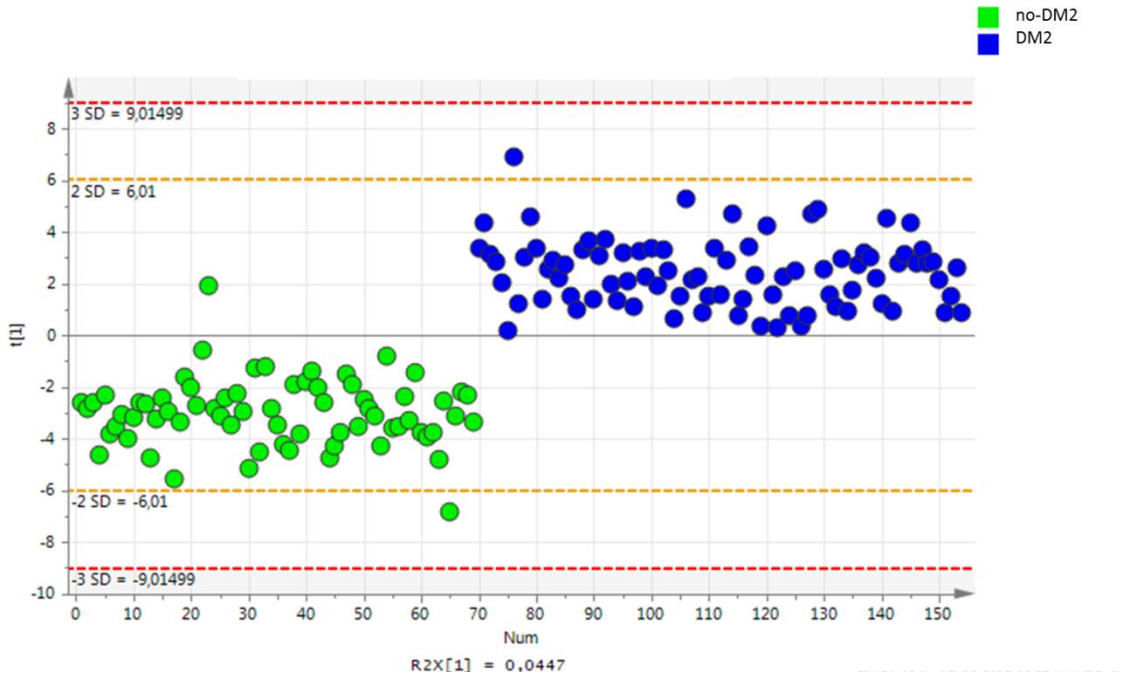


Figura 5.1 Representación gráfica del modelo OSC-PLS-DA. Las diferencias en el metaboloma urinario de los individuos DM2 (círculos en azul) y no-DM2 (círculos en verde) son representadas a través del análisis supervisado.

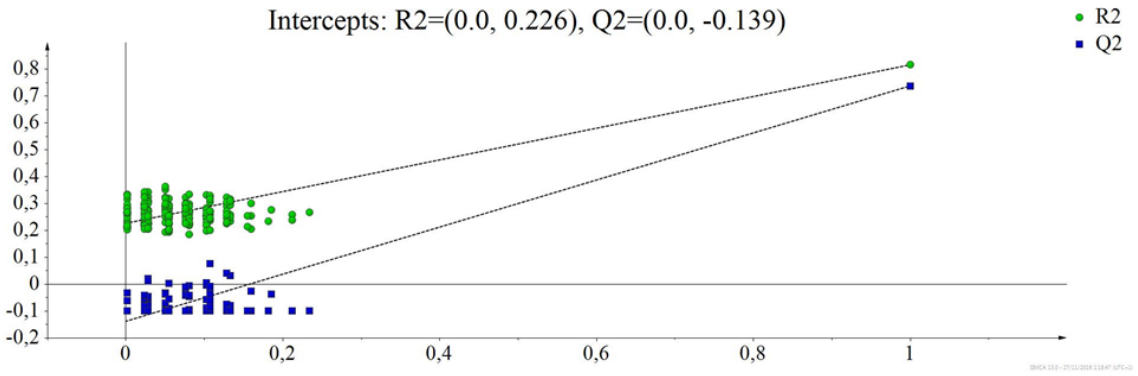


Figura 5.2 Gráfico de la prueba de permutaciones (n=200) del modelo OSC-PLS-DA para el análisis discriminante entre invididuos DM2 y no-DM2.

Tabla 5.2 Metabolitos urinarios identificados mediante RMN entre individuos DM2 y no-DM2

Metabolitos	Desplazamiento químico (δ, ppm), multiplicidad[1] y (J, Hz)	P-valor[2]	VIP[3]
Niveles más altos en DM2			
Ácido 3-hidroxihipúrico	7.12 (d, J=8.92), 7.29 (m), 7.37 (m), 7.41 (m)	0,01	1,11
Ácido 3-hidroxibutírico	1.19 (d, J= 6.24)	0,05	0,94
Ácido 3-hidroxiisovalérico	1.27 (s)	<0,01	1,26
Ácido 4-deoxitreónico	1.22 (d, J=6.47)	0,04	1,00
Alanina	1.48 (d, J=7.27)	0,01	1,45
Ácido cis-aconítico	3.13 (s), 5.75 (s)	<0,01	1,44
Ácido cítrico	2.54 (d, J=15.16), 2.68 (d, J=14.67)	0,05	1,28
Dimetilglicina	2.93 (s)	0,01	1,46
Glucosa	3.25 (m), 3.45 (m), 3.50 (m), 3.72 (m), 4.64 (d, J=7.93), 5.25 (d, J=3.71)	<0,001	3,58
Ácido guanidinoacético	3.79 (s)	<0,001	2,82
Alcohol isopropílico	1.16 (d, J=6.24)	<0,01	1,29
Ácido láctico	1.33 (d, J=6.90), 4.11 (q)	<0,001	2,03
Ácido metilsuccínico	1.07 (d, J=7.01)	<0,001	1,77
Fenilalanina	7.32 (d, J=7.15), 7.36 (m), 7.43 (t)	<0,01	1,20
Ácido pirúvico	2.38 (s)	<0,001	1,61
Ácido subérico	1.29 (m), 1.59 (t)	<0,001	1,77
Trimetilamina N-óxido /betaina[4]	3.27 (s), 3.90 (s)	<0,001	2,66

Tabla 5.2 Continuación

Niveles más bajos en DM2			
Ácido 3-hidroximandélico	6.87 (d, J=8.27), 6.93 (t), 6.99 (d, J=7.64)	<0,01	1,68
Ácido acetoacético	2.28 (s)	<0,01	1,12
Acetilcartinina	2.15 (s), 2.19 (s)	0,01	1,26
Creatinina	4.06 (s)	<0,001	2,36
Dimetilsulfona	3.16 (s)	<0,001	1,87
Glutamina	2.14 (m), 2.47 (m)	0,001	1,28
Ácido hippúrico	3.98 (d, J=5.82) 7.55 (t) 7.63 (tt) 7.83 (dd) 8.54 (bb)	<0,001	2,80
Histidina	7.08 (s)	0,01	1,35
Metilguanidina	2.82 (s)	<0,01	1,68
N-acetilglutamina	2.09 (m), 4.16 (m), 2.25 (t), 7.36 (m), 7.98 (m)	<0,01	1,09
N-acetilglicoproteinas	2.04 (s)	0,02	1,09
p-cresol	2.34 (s), 7.22 (d, J=8.37), 7.27 (d. J=8.90)	<0,01	1,70
Fenilacetilglutamina	2.27 (t), 4.17 (m), 7.42 (m), 7.26 (m), 7.98 (bb)	0,03	1,14
Scyllo-inositol	3.36 (s)	<0,001	2,13
Trigonelina	4.43 (s) 8.08 (m) 8.84 (m) 9.12 (s)	<0,01	1,44
Trimetilamina	2.86 (s)	<0,01	1,30

[1]multiplicidad (s, singlete; br. s, singlete amplio; d, doblete; dd, doble doblete; t, triplete; tt, triple triplete; m, multiplete); [2]significancia estadística obtenida a través de una prueba t-student para muestras independientes, [3]valores obtenidos a partir del modelo OSC-PLS-DA y empleados como criterio de selección de las variables de RMN (bins) para la posterior identificación de metabolitos; 4solapamiento de picos de los respectivos compuestos.

5.1.2.4 Visualización e interpretación de las alteraciones metabólicas asociadas a la DM2

Para visualizar e interpretar las alteraciones metabólicas asociadas a la DM2, la lista de de metabolitos identificados como discriminantes entre DM2 y no-DM2 fue sometida a un análisis de mapeo y de redes de rutas metabólicas mediante el empleo de las plataformas MetaboAnalyst® y MetaCore®. La visualización gráfica de los resultados de estos análisis se puede observar en las figuras 5.3 y 5.4. De acuerdo a los resultados obtenidos por el análisis en la plataforma MetaboAnalyst®, las principales alteraciones metabólicas asociadas a la DM2 se presentan en el metabolismo energético, incluyendo el ciclo de los ácidos tricarboxílicos y la glucólisis y gluconeogénesis; y en el metabolismo de aminoácidos, entre otros (Figura 5.3). A su vez las tablas 5.3 y 5.4 enlistan el total de las rutas metabólicas en las que todos los metabolitos discriminantes de DM2 podrían estar potencialmente implicados.

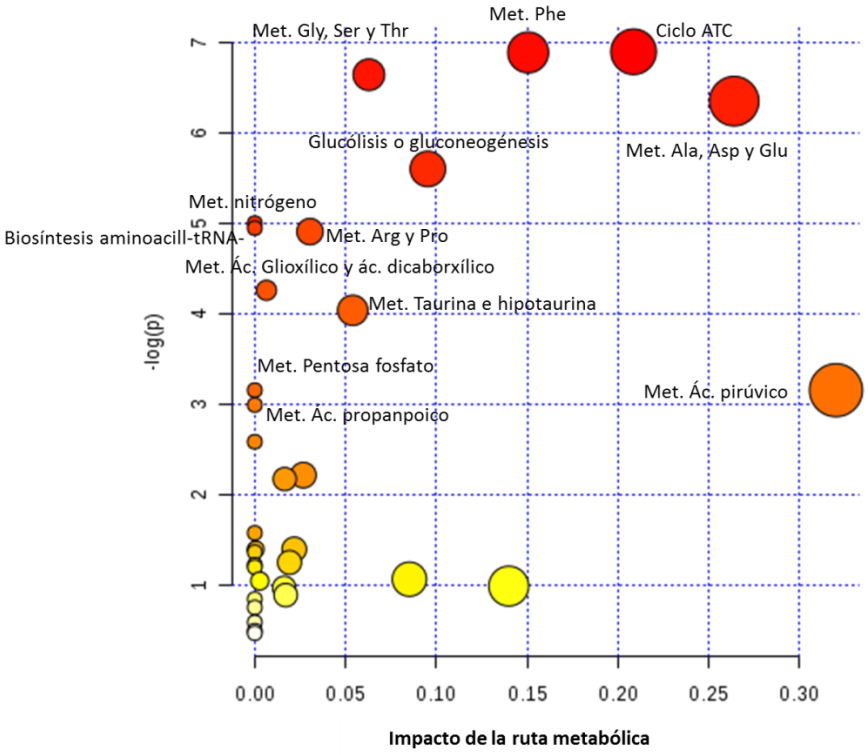


Figura 5.3 Vista de las principales rutas metabólicas afectadas por la DM2, de acuerdo al análisis hecho en la plataforma MetaboAnalyst®. El eje X y el tamaño del círculo representan el impacto de los metabolitos dentro de la correspondiente ruta metabólica, mientras que el eje Y y el color del círculo (variando de amarillo a rojo) indican la significancia de la ruta metabólica con respecto a todas las representadas en el gráfico. De este modo los círculos ubicaods hacia el fondo del gráfico con log(p)<3 y colores más claros representan rutas metabólicas con poco impacto y/o significancia en el análisis. Abreviaciones: Ác, ácido; ATC, ácidos tricarboxílicos; Met, metabolismo.

Tabla 5.3 Lista de las rutas metabólicas potencialmente alteradas en los individuos con diabetes mellitus tipo 2[1]

Nombre de la ruta metabólica	Total	Esperados	Coincidencias	p-valor	-log(p)	Ajuste Holm	FDR	Impacto
Cíclo de los ácidos tricarboxílicos	20	0,21	3	0,001	6,90	0,08	0,03	0,21
Metabolismo de fenilalanina	45	0,47	4	0,001	6,89	0,08	0,03	0,15
Metabolismo de glicina, serina y treonina	48	0,50	4	0,001	6,65	0,10	0,03	0,06
Metabolismo de alanina, ácido aspártico y ácido glutámico	24	0,25	3	0,001	6,36	0,13	0,03	0,26
Glucólisis o gluconeogénesis	31	0,32	3	<0,01	5,61	0,28	0,06	0,09
Biosíntesis de aminoacil-tRNA	75	0,78	4	<0,01	5,01	0,50	0,07	0,00
Metabolismo de nitrógeno	39	0,40	3	<0,01	4,95	0,52	0,07	0,00
Metabolismo de arginina y prolina	77	0,80	4	<0,01	4,91	0,54	0,07	0,03
Metabolismo de ácidos glioxílico y dicarboxílico	50	0,52	3	0,01	4,26	1,00	0,12	0,01
Metabolismo de taurina e hipotaurina	20	0,21	2	0,02	4,04	1,00	0,14	0,05
Ruta de la pentosa fosfato	32	0,33	2	0,04	3,16	1,00	0,28	0,00
Metabolismo de ácido pirúvico	32	0,33	2	0,04	3,16	1,00	0,28	0,32
Metabolismo de ácido propanoico	35	0,36	2	0,05	2,99	1,00	0,31	0,00
Metabolismo de ácido nicontínico y nicotinamida	44	0,46	2	0,07	2,59	1,00	0,43	0,00
Metabolismo de D-glutamina y ácido D-glutámico	11	0,11	1	0,11	2,22	1,00	0,57	0,03
Metabolismo de cisteína y metionina	56	0,58	2	0,11	2,17	1,00	0,57	0,02
Metabolismo de selenoaminoacidos	22	0,23	1	0,21	1,58	1,00	0,97	0,00
Biosíntesis de ácido pantoténico y coenzima A	27	0,28	1	0,25	1,40	1,00	0,97	0,00
Biosíntesis de fenilalanina, tirosina y triptófano	27	0,28	1	0,25	1, 40	1,00	0,97	0,00
Biosíntesis de valina, leucina e isoleucina	27	0,28	1	0,25	1, 40	1,00	0,97	0,02
Metabolismo de β-alanina	28	0,29	1	0,25	1,37	1,00	0,97	0,00
Metabolismo de vitamina B6	32	0,33	1	0,28	1,25	1,00	1,00	0,02

Tabla 5.3 Continuación

Biosíntesis de terpenoide backbone	33	0,34	1	0,29	1,23	1,00	1,00	0,00
Metabolismo de metano	34	0,35	1	0,30	1,20	1,00	1,00	0,00
Metabolismo de ácido butanoíco	40	0,41	1	0,34	1,07	1,00	1,00	0,08
Metabolismo de galactosa	41	0,42	1	0,35	1,05	1,00	1,00	<0,01
Metabolismo de histidina	44	0,46	1	0,37	0,99	1,00	1,00	0,14
Metabolismo de ácido ascórbico y aldárico	45	0,47	1	0,38	0,97	1,00	1,00	0,02
Metabolismo de almidón y sucrosa	50	0,52	1	0,41	0,89	1,00	1,00	0,02
Interconversiones de pentosa y ácido glucurónico	53	0,55	1	0,43	0,85	1,00	1,00	0,00
Metabolismo de pirimidina	60	0,62	1	0,47	0,75	1,00	1,00	0,00
Metabolismo de tirosina	76	0,79	1	0,55	0,59	1,00	1,00	0,00
Metabolismo de amino azúcares y nucleotidos azúcares	88	0,91	1	0,61	0,50	1,00	1,00	0,00
Metabolismo de purina	92	0,95	1	0,62	0,47	1,00	1,00	0,00

[1]Lista complementaria al análisis de rutas metabólicas hecho en MetaboAnalyst®

Tabla 5.4 Metabolitos urinarios encontrados significativamente diferentes entre individuos diabéticos y no diabéticos, y las rutas metabólicas específicas en las que potencialmente se encuentran involucrados[1]

Metabolito	Rutas metabólicas
Niveles más altos en DM2	
Ácido 3-hidroxivalérico	Metabolismo de leucina, isoleucina y valina
Ácido 4-deoxitreónico acid	Ciclo de ácidos tricarboxílicos
Ácido 3-hidroxibutírico	Síntesis y degradación de cuerpos cetónicos; metabolismo de ácido butanoico
Alanina	Metabolismo de alanina, ácido glutámico, cisteína, glicina, glutamina, histidina, metionina, prolina, serina y treonina
Ácido cis-aconítico	Ciclo de ácidos tricarboxílicos
Ácido cítrico	Ciclo de ácidos tricarboxílicos
Dimetilglicina	Metabolismo de glicina, serina, cisteína y treonina
Glucosa	Glucólisis y gluconeogénesis; metabolism de fructosa y galactosa; regulación del metabolism de la insulina, lípidos y ácidos grasos
Ácido guanidinoacético	Metabolismo de aminoácidos y ciclo de la urea
Alcohol isopropílico	Metabolismo de ácido pirúvico; glucólisis y gluconeogénesis
Ácido láctico	Glucólisis y gluconeogénesis
Ácido metilsuccínico	Metabolismo de isoleucina
Fenilalanina	Metabolismo de fenilalanina y tirosina
Ácido pirúvico	Metabolismo de alanina, ácido glutámico, cisteína, glicina, glutamina, histidina, metionina, prolina, serina y treonina; glucólisis y gluconeogénesis; metabolismo de ácido pirúvico; ciclo de ácidos tricarboxílicos; regulación del metabolism de la insulina, lípidos y ácidos grasos
Ácido subérico	Metabolismo de ácidos dicarboxílicos
Trimetimetilamina N-óxido /betaina‡	Metabolismo de colina

Tabla 5.4 continuación

Niveles más bajos en DM2	
Ácido 3-hidroxihipúrico	Metabolismo microbiano; metabolism de glicina
Ácido 3-hidroximandélico	Metabolismo de catecolamina; metabolism de m-octopamina, m-sinefrina (fenilefrina) y tirosina
Ácido acetoacético	Síntesis y degradación de cuerpos cetónicos; metabolism de leucina, isoleucina, valina y tirosina; metabolismo de ácido butanoico y ácido propanoico
Acetilcarnitina	Oxidación de ácidos grasos; resistencia a la insulina
Creatinina	Metabolismo de arginina y prolina
Dimetilsulfona	Metabolismo microbiano; metabolism endógeno humano de metanotiol; metabolism de azufre
Glutamina	Metabolismo de nitrógeno; metabolism de ácido glutámico, glutamina, histidina y prolina
Ácido hipúrico	Metabolismo microbiano; metabolism de fenilalanina
Histidina	Metabolismo de β-alanina; metabolism de histidina, ácido glutámico, glutamina y prolina
Metilguanidina	Catabolismo de proteínas
N-acetilglutamina	Metabolismo de amino azúcares y nucleótido azúcares
N-acetilglicoproteínas	Catabolismo de proteínas
p-cresol	Metabolismo microbiano; catabolismo de proteínas
Fenilacetilglutamina	Metabolismo de fenilalanina
Scyllo-inositol	Metabolismo de inositol fosfato
Trigonelina	Metabolismo de ácido nicotínico y nicotinamida
Trimetilamina	Metabolismo de metano; metabolism microbiano

[1]Lista complementaria al análisis de rutas metabólicas hecho en MetaboAnalyst®

La construcción de redes metabólicas a partir de la plataforma MetaCore® se realizó con el fin de visualizar con mayor detalle la relación entre los metabolitos incluidos en las dos principales rutas metabólicas afectadas por la DM2. De este modo, del total de los metabolitos encontrados identificados, seis de ellos se encontraron relacionados con el metabolismo energético (glucosa y ácidos pirúvico, láctico, hidroxibutírico, acetoacético y *cis*-aconítico) y otros 12 relacionados con el metabolismo de los aminoácidos (alanina, creatina, creatinina, glicina, glutamina, histidina, fenilalanina, fenilacetilglutamina, lisina, valina, y ácidos hipúrico y guanidinoacético) (Figura 5.4).

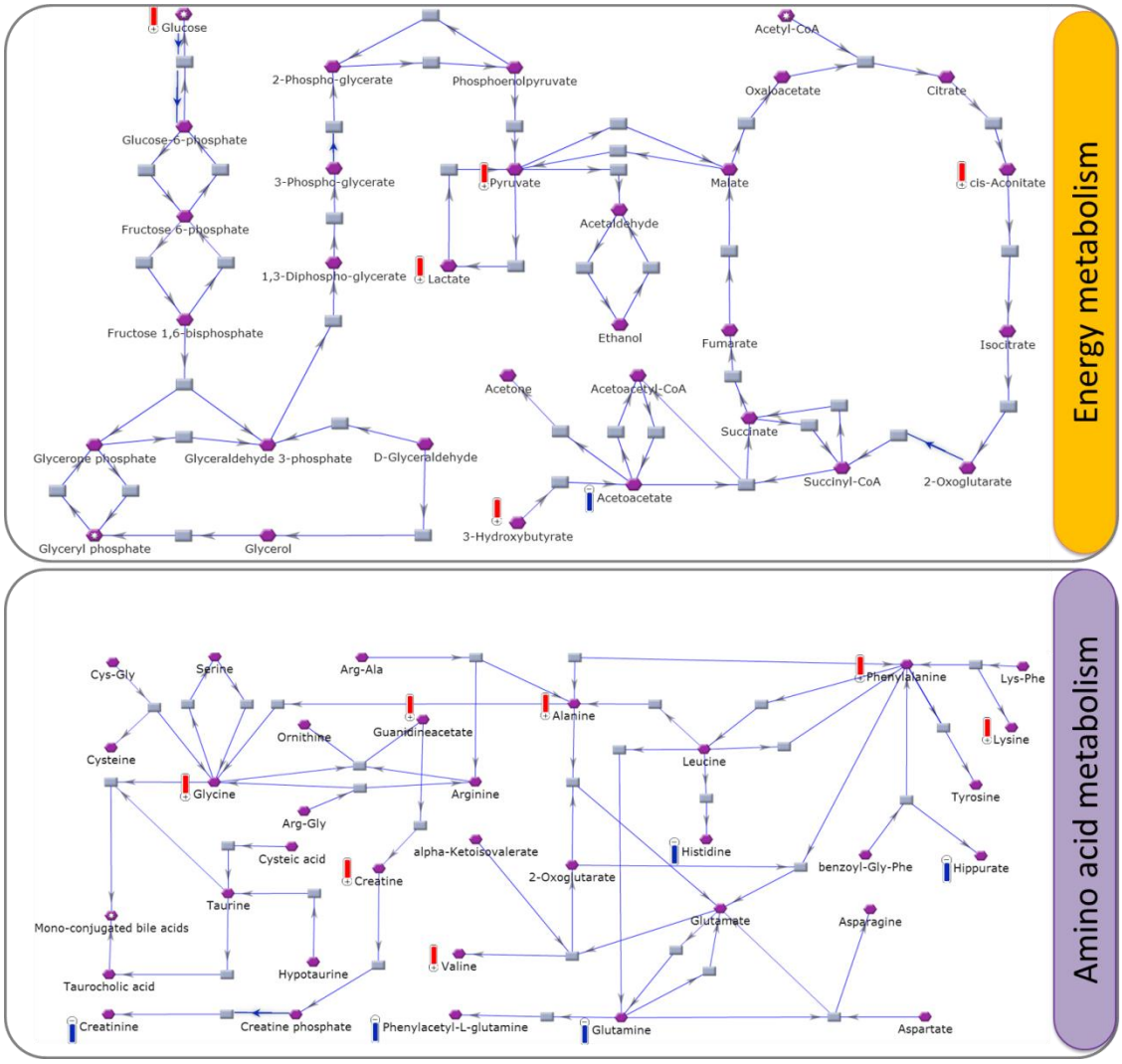


Figura 5.4 Relación entre los metabolitos urinarios identificados dentro de los dos principales tipos de metabolismo afectados por la DM2. El color de las barras y signos denotan el sentido de la excreción asociada a la DM2: rojo (+), niveles más altos; azul (-), niveles más bajos. Redes metabólicas construidas en la plataforma MetaCore®.

5.1.2.5 Cálculo del biomarcador asociado a DM2

A partir del perfil de metabolitos discriminantes obtenidos del modelo OSC-PLS-DA (Tabla 5.2), pero excluyendo a la glucosa, se construyó un modelo para definir un biomarcador multimetabolito que permitiera una mejor discriminación de la DM2 entre los participantes. La exclusión de la glucosa en la construcción del modelo se debió a que dada su alta capacidad discriminante de DM2, su inclusión en el biomarcador reduciría la inclusión de otros metabolitos que en su conjunto permitieran una similar o mejor capacidad discriminante. Para la construcción del modelo, los participantes fueron previamente asignados aleatoriamente a uno de dos grupos denominados *training set* (2/3 de la población) y *validation set* (1/3 de la población). Se construyó el modelo en el *traning set*, mediante una regresión logística binaria paso a paso para posteriormente validarlo con la población del *validation set*. Los metabolitos incluidos en el modelo y, por tanto, en el biomarcador resultante, se muestran en la tabla 5.5

Tabla 5.5 Metabolitos seleccionados por la regresión logistica binaria por pasos para la construcción de un biomarcador multimetabolito de DM2.

Metabolito	Coeficiente	Error estándar	*P*-valor
Ácido metilsuccinico	9,25	2,85	<0,01
Alanina	2,87	1,00	<0,01
Glutamina	-5,29	2,78	0,06
Metilguanidina	-5,61	2,16	0,01
Dimetilglicina	2,47	1,25	0,05
Ácido guanidinoacético	3,32	1,30	0,01
Ácido 3-hidroximandélico	-5,09	1,89	0,01
Ácido hipúrico	-10,4	2,81	<0,01

De acuerdo al modelo, el biomarcador de DM2 estuvo compuesto por niveles altos de ácido metilsuccínico, alanina, dimetilglicina y ácido guanidinoacético, así como por niveles bajos de glutamina, metilguanidina, ácido 3-hidroximandélico y ácido hipúrico. La capacidad para discriminar la DM2 del presente biomarcador se evaluó mediante análisis de curvas ROC, calculando el AUC, sensibilidad y especificidad, tanto en el *training set* como en el *validation set*. El AUC, sensibilidad y especificidad del biomarcador y de los metabolitos incluidos se

muestran en la tabla 5.6. En esta misma tabla se incluyen los valores de estos parámetros para la glucosa, que aunque fue excluída

El AUC del biomarcador en el *training set* y *validation set* fue de 96,1% y 96,4%, respectivamente, siendo estos valores superiores a los mostrados para cada metabolito incluído en el modelo. Por lo que respecta a la especificidad y sensibilidad, el biomarcador tuvo valores de 96,0% y 93,0% en el *training set*, y de 87,0% y 96,4% en el *validation set*, respectivamente. Al igual que con el AUC, los valores de especificidad y sensibilidad fueron superiores a los obtenidos para cada uno de los metabolitos. Por otra parte, al comparar el biomarcador con la glucosa en términos de capacidad discriminante, se pudo observar que los valores de AUC, especificidad y sensibilidad fueron superiores en el biomarcador, sugiriendo de este modo una alta o incluso mayor capacidad discriminante de la DM2, en comparación con la glucosa.

Tabla 5.6 Parámetros de los análisis de curvas ROC para el biomarcador de DM2 y los metabolitos incluidos en él, de acuerdo a los grupos *training* y *validation*.

Training set	AUC (95% CI)	*P*-valor	Especificidad (%)	Sensibilidad (%)
Modelo combinado	96,1 (92,3-100)	<0.001	96,0	93,0
Glucosa*	83,5 (76,0-91,1)	<0.001	91,3	66,7
Validation set				
Modelo combinado	96,4 (92,0-100)	<0.001	87,0	96,4
Ácido metilsuccinico	69,4 (54,8-84,0)	0,02	74,0	57,1
Alanina	66,3 (51,5-81,1)	0,05	78,3	53,6
Glutamina	37,6 (21,8-53,4)	0,13	34,8	67,8
Metilguanidina	31,8 (17,1-46,5)	0,03	34,8	50,0
Dimetilglicina	64,0 (48,7-79,3)	0,09	65,2	64,3
Ácido guanidinoacético	81,0 (68,8-93,4)	<0,001	74,0	82,1
Ácido 3-hidroximandélico	40,2 (24,3-56,2)	0,23	30,0	64,3
Ácido hipúrico	15,0 (3,40-36,4)	<0,001	8,69	75,0
Glucosa*	89,8 (80,2-99,3)	<0.001	82,6	93,0

*Excluída para la creación del modelo y biomarcador , pero incluída en la presente tabla con fines comparativos.

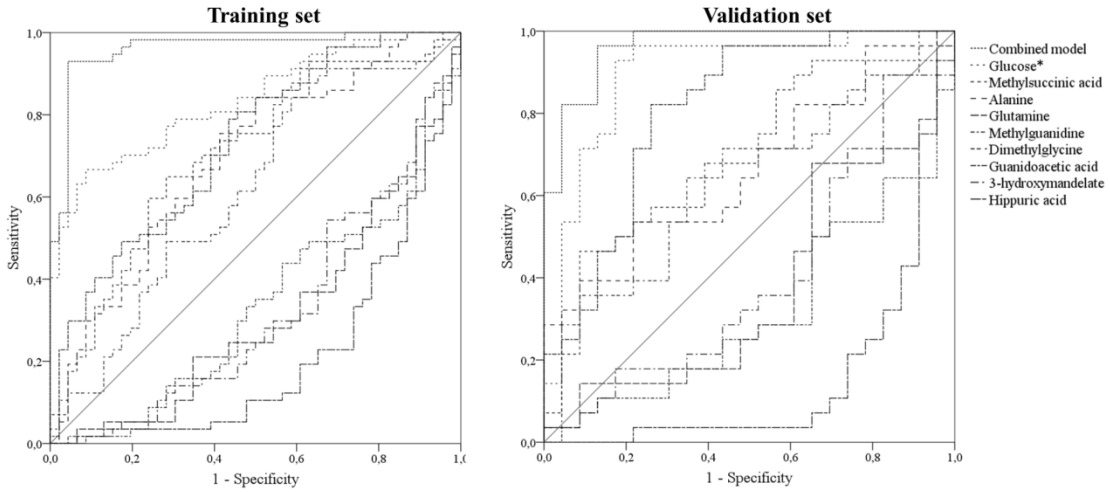


Figura 5.5 Curvas ROC del biomarcador multi-metabolito de DM2 y los metabolitos incluidos en él, de acuerdo a los grupos *training* y *validation*. *Aunque fue excluída para la construcción del biomarcador, se incluye en el presente análisis con fines de comparación con el biomarcador y el resto de los metabolitos incluidos.

5.1.2.6 Clasificación de la población estudiada en metabotipos relacionados con la DM2

La obtención de metabotipos relacionados con la DM2 se realizó a partir de un análisis de *clusters* por el método *k-means*, fijando *k clusters*=2, e incluyendo el total de los metabolitos identificados asociados a la DM2 (Tabla 5.2). De este análisis se obtuvieron cuatro metabotipos: dos en el grupo de participantes con DM2 y otros dos en el grupo no-DM2 (Figura 5.5). De este análisis también se se pudo observar que, aunque la mayoría de las diferencias observadas en los metabolitos entre los grupos DM2 y no-DM2 fueron las mismas a las obtenidas por el modelo OSC-PLS-DA, los metabotipos 1 (*cluster* 1) y 2 (*cluster* 2), dentro del grupo DM2, fueron diferentes en la excresión de cuatro metabolitos: ácido acetoacético, *p*-cresol, fenilalanina y fenilacetilglutamina. Concretamente, los niveles de excreción de estos metabolitos fueron superiores en el *cluster* 2 en comparación con el *cluster* 1. Respecto a los metabotipos 3 (*cluster* 3) y 4 (*cluster* 4), dentro del grupo no-DM2, se pudo observar un patrón similiar en los niveles de los cuatro metabolitos mencionados anteriormente. En este sentido, los niveles de excresión de ácido acetoacético, *p*-cresol, fenilalanina y fenilacetilglutamina fueron similares entre los metabotipos 2 (DM2) y 3 (no-DM) (Figura 5.7). Al comparar los datos antropometricos y clínicos entre los distintos metabotipos, se observó que los niveles de glucosa plasmática en el metabotipo 2 fueron más altos con respecto a los del metabotipo 1. Por lo que respecta a la medicación y al resto de las variables demográficas y clínicas, no se observaron diferencias significativas entre ambos metabotipos 1 y 2.

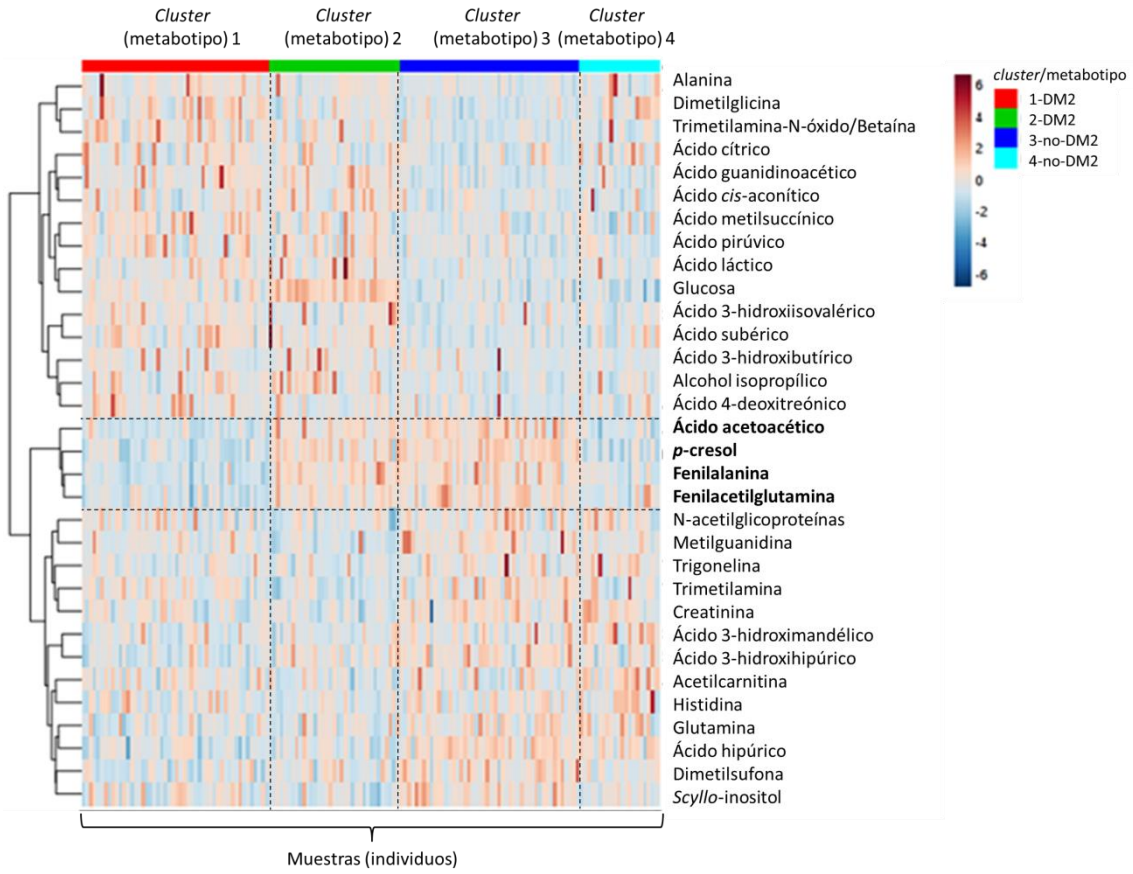


Figura 5.5 Gráfico tipo *Heatmap* de las diferencias y similitudes en la excresión urinaria de los metabolitos entre los cuatro metabotipos asociados a la DM2. Al centro, delimintados por líneas punteadas horizontales y texto en negrita, los metabolitos que distinguen entre los metabotipos 1 y 2.

Tabla 5.7 Características de los participantes de acuerdo al metabotipo asociado a la DM2

	Cluster 1 (n=50)	*Cluster* 2 (n=35)	*Cluster* 3 (n=48)	*Cluster* 4 (n=21)	*P*-valor
Edad (años)	67 ± 5	68 ± 6	67 ± 5	66 ± 6	0,42
Hombres, n (%)	20 (40)	10 (29)	12 (25)	6 (29)	0,42
Fumadores actuales, n (%)	6 (12)	5 (14)	6 (13)	3 (14)	0,87
IMC, kg/m^2	30,8 ± 3.60	29,9 ± 5,80	30,0 ± 3,90	29,7 ± 4,10	0,70
Obesidad (IMC≥25 kg/m^2), n (%)	25 (50)	12 (34)	23 (48)	11 (52)	0,45
Peso, kg	79,4 ± 11,7	75,3 ± 15,1	74,2 ± 11,7	73,3 ± 11,8	0,13
CC, cm	105 ± 9,80	102 ± 13,2 [ab]	98,0 ± 10,4	99,2 ± 11,9	0.02
Actividad física, MET·min/d	258 ± 323	217 ± 223	305 ± 257	242 ± 214	0,51
PAS, mmHg	139 ± 16,5	143 ± 19,8	144 ± 20,7	156,5 ± 20,3	0,06
PAD, mmHg	80,9 ± 10,0	81,3 ± 11,3	82,4 ± 10,3	84,2 ± 9,50	0,79
Score de adherencia a la DietMed	8,20 ± 1,70	8,50 ± 1,90	9,00 ± 1,60	8,10 ± 1.90	0,08
Marcadores plasmáticos					
Glucosa, mg/dl	148 ± 34,8	175,5 ± 53,5	89,9 ± 7,60	90,0 ± 6,40 [c]	<0.001
Colesterol total, mg/dl	197 ± 29,7	220 ± 38,7	220 ±28,5	217 ± 44,7	0,25
Triglicéridos, mg/dl	150 ± 53,3	139± 56,6	127 ± 53,3	121 ± 34,2	0,59
Medicación, n ± (%)					
Antihipertensivos	22 (44)	10 (29)	24(50)	9 (43)	0,27
Hipolipemiantes	25 (50)	17 (49)	32 (67)	11 (52)	0,29
Antidiabéticos orales	40 (80) [a]	25 (71) [a]	0 (0) [b]	0 (0) [b]	<0,001
Insulina	8 (16) [a]	10 (29) [a]	0 (0) [b]	0 (0) [b]	<0,001

Los datos se muestran como la media ± desviación estándar, para variables continuas; o como el número (n) y porcentaje (%) para variables categóricas, según se indica. Abreviaciones: CC, circunferencia de cintura; IMC, índice de mása corporal; MET, equivalente metabólico; DietMed, dieta Mediterránea; PAD, presión arterial diastólica; PAS, presión aterial sistólica. *P*-valor de significancia calculado para las diferencias entre medias mediante ANOVA. Letras diferentes en la misma fila denotan diferencias estadísticamente significativas.

5.2 BIOMARCADORES METABOLÓMICOS DE ADHERENCIA A LA DIETA MEDITERRÁNEA

Los resultados del presente estudio han sido publicados en la revista *Journal of Nutricional Biochemistry*, indexada en el *Journal Citation Reports* con un factor de impacto de 4,518 (2016), situada en el primer cuartil (Q1) de las categorías *Biochemistry and molecular biology* (62/286) y *Nutrition and dietetics* (11/81).

Enrique Almanza-Aguilera, Mireia Urpi-Sarda, Rafael Llorach, Rosa Vázquez-Fresno, Mar Garcia-Aloya, Francesc Carmona, Alex Sanchez, Francisco Madrid-Gambin, Ramon Estruch, Dolores Corella, Cristina Andres-Lacueva. Microbial metabolites are associated with a high adherence to a Mediterranean dietary pattern using a 1H-NMR-based untargeted metabolomics approach. *Journal of Nutricional Biochemistry*, 48 (2017) 36-43.

Resumen

El estudio de biomarcadores de patrones alimentarios incluyendo la DietMed es todavía escaso y podría mejorar la evaluación de estos patrones. Además, el estudio de estos biomarcadores podría proporcionar una mejor comprensión de los beneficios para la salud de los patrones alimentarios dietéticos en la epidemiología nutricional. Los objetivos del presente estudio fueron: i) identificar la huella metabólica urinaria que permita diferenciar entre individuos con una elevada y baja adherencia a un patrón de dieta Mediterránea; y ii) calcular un biomarcador metabolómico de alta adherencia a la dieta Mediterránea con una elevada sensibilidad y especificidad. En este estudio transversal, se incluyeron 56 individuos con A-ADM y 63 individuos con B-ADM, respectivamente; todos ellos pertenecientes al estudio PREDIMED. El estudio metabolómico con enfoque no dirigido consistió en el análisis de muestras de orina mediante RMN. Se realizaron análisis multivariantes para determinar las diferencias de los metabolitos entre los grupos, y para el cálculo y evaluación del biomarcador de A-ADM. Un total de 34 metabolitos fueron identificados como discriminantes entre A-ADM y B-ADM. La huella metabolómica asociada a la A-ADM incluyó una mayor excreción de prolina betaína y fenilacetilglutamina, entre otras, y una menor excreción de metabolitos relacionados con el metabolismo de la glucosa. El biomarcador de A-ADM con valores de especificidad, sensibilidad y AUC de 95%, 95%, y 97%, respectivamente, estuvo compuesto por tres metabolitos microbianos: fenilacetilglutamina, *p*-cresol y ácido 4-hidroxifenilacético. El perfil global de metabolitos identificados refleja la modulación metabólica producida por A-ADM. El biomarcador propuesto de A-ADM podría ser una mejor herramienta de apoyo en la epidemiología nutricional para establecer futuras asociaciones entre A-ADM y la prevención o mejora de enfermedades crónicas.

5.2.2 Resultados

5.2.2.1 Características de los participantes de los grupos de baja y alta adherencia

Para el presente estudio la población fue inicialmente estratificada en B-ADM o A-ADM en base a la puntuación (*score*) de la adherencia a la DietMed obtenida a través del cuestionario de de 14 puntos preguntas del estudio PREDIMED. Las características generales y la ingesta de alimentos y nutrientes de la población agrupada según su nivel de adherencia a la DietMed se recogen en las tablas 5.6 y 5.7, respectivamente. De manera general los participantes tuvieron una edad promedio de 66,7 ± 5,9 años, un IMC de 30,3 ± 4,5 kg/m^2, y un 68,9% de la población fueron mujeres. Respecto a los factores de riesgo cardiovascular no se observaron diferencias significativas entre los grupos B-ADM y A-ADM (Tabla 5.6).

Tabla 5.6 Características demográficas y de parámetros clínicos en los participantes agrupado de acuerdo al nivel de adherencia a la DietMed

	Total	B-ADM (≤ 7 DietMed *score*)	A-ADM (≥10 DietMed *score*)	*P*-valor
Participantes, *n*	119	63	56	
Edad, (años)	66,7 ± 5,90	66,8 ± 6,1	66,6 ± 5,7	0,83
Mujeres, *n* (%)	82 (68,9)	42 (66,7)	40 (71,4)	0,58
IMC, kg/m^2	30,3 ± 4,5	31,0 ± 4,10	29,5 ± 4,70	0,08
Fumadores actuales, *n* (%)	12 (10,1)	7 (11,1)	5 (8,9)	0,69
Historia familiar de ECV, *n* (%)	39 (32,8)	23 (36,5)	16 (28,6)	0,34
Diabetes, *n* (%)	63 (52,9)	36 (57,1)	26 (46,4)	0,24
Hipertensión, *n* (%)	102 (85,7)	51 (81,0)	51 (91,1)	0,12
Hipercolesterolemia, *n* (%)	85 (71,4)	44 (69,8)	41 (73,2)	0,68
Medicación				
Antidiabéticos orales, *n* (%)	46 (38,7)	26 (41,3)	20 (35,7)	0,53
Insulina, *n* (%)	18 (15,1)	10 (16,0)	8 (14,3)	0,81
Antihipertensivos, *n* (%)	56 (47,1)	27 (42,9)	29 (51,8)	0,33
Hipolipemiantes, *n* (%)	58 (48,7)	29 (46,0)	29 (51,8)	0,53

Los datos se muestran como la media ± desviación estándar o número (n) y porcentaje (%), según se indica. Las diferencias en estas variables entre la baja y alta adherencia se calcularon mediante pruebas *t*-student o *chi*-cuadrado, de acuerdo a si son variables continuas o categóricas, respectivamente. Abreviaciones: ECV, enfermedad cardiovascular; A-ADM, alta adherencia a la dieta Mediterránea; B-ADM, baja adherencia a la dieta Mediterránea; DietMed, dieta Mediterránea.

Con respecto a la ingesta de alimentos y nutrientes, y en comparación con el grupo B-ADM, los participantes en el grupo A-ADM reportaron consumir mayores cantidades de aceite de oliva, frutos secos, verduras, frutas, legumbres, pescado y fibra dietética. Mientras que la ingesta de grasas totales, carbohidratos, proteínas, y de energía total fue similar entre los dos grupos (Tabla 5.7).

Tabla 5.7 Ingesta de alimentos y nutrientes entre los participantes del estudio y de acuerdo a su nivel de adherencia a la DietMed

Ingesta de alimentos y nutrientes	Total	B-ADM (≤ 7 DietMed score)	A-ADM (≥10 DietMed score)	*P*-valor
Participantes, *n*	119	63	56	
Alimentos				
Aceite de oliva, g/d	38 ± 15	35 ± 15	41 ± 13	0,02
Frutos secos, g/d	12 ± 12	9 ± 12	15 ± 12	0,02
Verduras, g/d	402 ± 192	377 ± 215	430 ± 160	0,03
Frutas, g/d	480 ± 242	419 ± 235	550 ± 233	<0,01
Legumbres, g/d	20 ± 10	17 ± 9	22 ± 10	<0,01
Cereales, g/d	269 ± 111	275 ± 112	263 ± 111	0,56
Pescado, g/d	103 ± 50	84 ± 47	124 ± 45	<0,001
Carne, g/d	147 ± 69	152 ± 81	142 ± 52	0,41
Lácteos, g/d	358 ± 213	366 ± 207	350 ± 220	0,70
Bollería, g/d	18 ± 21	19 ± 22	16 ± 20	0,44
Vino, mL/d	56 ± 116	54 ± 4	47 ± 3	0,71
Nutrientes				
Ingesta TE, kcal/d	2397 ± 569	2399 ± 635	2394 ± 490	0,96
Carbohidratos, % del TE	43,3 ± 6,8	43,4 ± 6,6	43,2 ± 7,0	0,87
Proteínas, % del TE	16,7 ± 2,6	16,4 ± 2,8	17,1 ± 2,4	0,13
Grasa total, % del TE	38,0 ± 6,2	37,8 ± 6,5	38,2 ± 5,8	0,71
Fibra dietética, g/d	28,4 ± 8,0	26,2 ± 7,0	31,0 ± 8,3	<0,01
DietMed score	8,30 ± 2,20	6,40 ± 0,90	10,5 ± 0,60	<0,001

Los datos se muestran como la media ± desviación estándar. Las diferencias en estas variables entre la baja y alta adherencia se calcularon mediante pruebas *t*-student. Abreviaciones: A-ADM, alta adherencia a la dieta Mediterránea; B-ADM, baja adherencia a la dieta Mediterránea; DietMed, dieta Mediterránea; TE, total de la energía.

5.2.2.2 Identificación de metabolitos asociados a la alta adherencia a dieta Mediterránea

Para la obtención de la huella metabólica discriminante entre B-ADM y A-ADM, se realizó un modelo OSC-PLS-DA. Este modelo resultó en un componente latente y estadísticos R^2Y (cum) y Q^2Y (cum) con valores 0,913 and 0,764, respectivamente, indicando una buena capacidad para discriminar entre los dos grupos B-ADM y A-ADM. Posteriormente, el modelo fue validado a través de un test de permutaciones (n=200), proporcionando valores de intersección con la ordenada R^2 y Q^2 de 0,346 y -0,154, respectivamente.

Las variables con valores VIP$\geq$1,0 del modelo OSC-PLS-DA fueron sometidas a pruebas de *t*-student o U de Mann-Whitney (de acuerdo al supuesto de distribución normal) para determinar las diferencias de estas entre los grupos B-ADM y A-ADM, y la posterior identificación de metabolitos. De este modo, un total de 34 metabolitos fueron hallados como discriminantes entre B-ADM y A-ADM. En la tabla 5.8 se muestra una lista de estos metabolitos y las rutas metabólicas en las que éstos están involucrados, así como la dirección en su excreción asociada a la A-ADM.

Los participantes en el grupo A-ADM tuvieron una marcada excreción de metabolitos relacionados con el metabolismo de proteínas y aminoácidos. Concretamente este grupo de metabolitos asociados a la A-ADM estuvo caracterizado por una excreción mayor de anserina, carnosina, creatina, creatinina, ácido guanidinoacético, histidina, N-acetilglutamina, y una menor excreción de 3-metilhistidina, alanina, glicina y lisina. De forma similar, otro grupo mayoritario de metabolitos asociados a la A-ADM estuvo compuesto por metabolitos derivados del metabolismo de la microbiota. Este patrón de metabolitos se compuso por niveles más altos en orina de ácido 4-hidroxihipúrico, ácido 4-hidroxifenilacético, dimetilsulfona, ácido 3-(-3hidroxifenil)-3-hidroxipropiónico, *p*-cresol y fenilacetilglutamina, y por niveles más bajos de ácido 3-indoxilsulfhídrico, ácido hipúrico y ácido isobutírico. Por otra parte, la A-ADM se caracterizó también por una excreción más baja de metabolitos relacionados con el metabolismo energético, del ácido propiónico, purina y cafeína, así como niveles más altos en orina de metabolitos incluidos en el metabolismo de la colina (con excepción de la betaína), inositol fostato, ácido nicotínico, nicotinamida y pirimidina.

Tabla 5.8 Metabolitos urinarios discriminantes entre la alta y baja adherencia a la dieta Mediterránea

Metabolito	^{1}H δ, ppm (multiplicidad)	Excreción en A-ADM	*P*-valor[2]
Metabolismo de proteínas/aminoácidos			
3-Metilhistidina	7.95 (s)	↓	0,03
Alanina	1.49 (d)	↓	<0,01
Anserina	3.79 (s), 8.29 (s)	↑	<0,001
Carnosina	7.18 (s), 8.15 (s)	↑	0,04
Creatina	3.94 (s)	↑	<0,01
Creatinina	3.05 (s), 4.07 (s)	↑	<0,01
Glicina	3.57 (s)	↓	<0,01
Ácido guanidinoacético	3.78 (s)	↑	<0,01
Histidina	7.09 (s), 7.93 (s)	↑	<0,001
Lisina	1.72 (m)	↓	0,02
N-acetilglutamina	2.08 (m), 2.30 (m)	↑	<0,01
Prolina betaína	3.11 (s), 3.30 (s), 4.08 (m)	↑	<0,001
Metabolismo de la microbiota			
Ácido 3-indoxil sulfhídrico	7.27 (t), 7.51 (d), 7.71 (d)	↓	<0,001
Ácido 4-hidroxihipúrico	6.98 (d), 7.75 (d)	↑	<0,001
Ácido 4-hidroxifenilacético	6.88 (d), 7.16 (s)	↑	<0,001
Dimetilsulfona	3.16 (s)	↑	<0,01
Acido hipúrico	7.55 (tt), 7.82 (dd)	↓	<0,01
HPHPA	5.02 (t), 6.86 (d), 6.93 (br s), 7.30 (t)	↑	0,03
Ácido isobutírico	1.07 (d)	↓	<0,001
p-cresol	2.34 (s), 7.22 (d), 7.28 (d)	↑	<0,001
Fenilacetilglutamina	1.93 (m), 2.27 (m), 3.66 (m), 4.18 (m), 7.36 (t), 7.43 (t), 8.0 (d)	↑	<0,001
Metabolismo energético (glicólisis/gluconeogénesis, ATC, y cuerpos cetónicos)			
β-glucosa	3.39 (m), 3.44 (m), 3.49 (m), 3.90 (m), 4.66 (d)	↓	<0,001
Ácido láctico	1.35 (d)	↓	0,01

Tabla 5.8 *continuación*

Metabolito	^{1}H δ, ppm (multiplicidad)	Excreción en A-ADM	*P*-valor[2]
Ácido succínico	2.42 (s)	↓	<0,01
Metabolismo de la colina			
Dimetilamina	2.72 (s)	↑	<0,001
Betaína	3.27 (s), 3.91 (s)	↓	0,02
Trimetilamina-N-oxido	3.28 (s)	↑	<0,01
Metabolismo del inositol fosfato			
Scyllo-inositol	3.36 (s)	↑	<0,001
Metabolismo del ácido nicotínico y nicotinamida			
N-Metilnicotinamida	4.49 (s), 8.85 (m), 9.13 (s)	↑	<0,01
Metabolismo del ácido propiónico			
Alcohol isopropílico	1.15 (d)	↓	0,02
Metabolismo de la purina y cafeína			
Xantosina	7.86 (s)	↓	0,01
Metabolismo de la pirimidina			
Urea	5.82 (br.s)	↑	<0,01
Metilguanidina	2.82 (s)	↑	0,02
Ácido malónico	3.12 (s)	↑	0,03

Las flechas (↑) y (↓) indican una excreción relativamente más alta o baja, respectivamente, en la A-ADM con respecto a la B-ADM. Las diferencias en los metabolitos entre los grupos B-ADM y A-ADM se analizaron mediante pruebas *t*-student o U de Mann-Whitney, dependiendo de la distribución normal de los datos. La significancia fue corregida mediante el procedimiento de Benjamini–Hochberg.

Abreviaciones: HPHPA, ácido 3-(3-hidroxifenil)-3-hidroxipropiónico; multiplicidad (s, singlete; br. s, singlete amplio; d, doblete; dd, doble doblete; t, triplete; tt, triple triplete; m, multiplete).

5.2.2.3 Biomarcador de adherencia a la DietMed

Se realizó un modelo basado en un análisis de regresión logística binaria paso a paso, incluyendo los 34 metabolitos anteriormente identificados a partir del modelo OSC-PLS-DA. Posteriormente se empleó un análisis de curvas ROC para evaluar la capacidad discriminante del modelo resultante y de los metabolitos incluidos en este. Con este fin, el conjunto de datos de los individuos de los grupos A-ADM y B-ADM se dividió en dos subgrupos llamados *training* set y *validation* set, tal como se indica en el apartado de metodología correspondiente. De este modo, un modelo compuesto por niveles altos de fenilacetilglutamina, *p*-cresol y ácido 4-hidroxifenilacético fue asignado como el biomarcador multi-metabolito de una A-ADM (Tabla 5.9).

Tabla 5.9 Metabolitos seleccionados por el modelo de regresión logística como los más discriminantes entre una baja y alta adherencia a la dieta Mediterránea.

Metabolito	Parámetros del modelo de regresión		
	Coeficiente	Error estándar	*P*-valor
Fenilacetilglutamina	34,7	8,69	<0,001
p-cresol	9,68	3,16	<0,01
Ácido 4-hidroxifenilacético	16,0	5,62	<0,01

La especificidad y la sensibilidad del biomarcador, medidas a través del análisis de las curvas ROC, fueron superiores al 90% tanto en el *training set* como en el v*alidation set* (Tabla 5.10). La fenilacetilglutamina tuvo 85,7% y 68,4% de especificidad y sensibilidad, respectivamente en el *validation* set, mientras que el *p*-cresol y ácido 4-hidroxifenilacético mostraron valores entre 66% y 84% en estos mismos parámetros y también en el *validation set* (Tabla 5.10). Por otra parte, la calidad del biomarcador en términos de AUC fue del 97,7% en el *training set* y del 97,0% en el *validation set*, mientras que el AUC de los metabolitos individualmente se encontró entre el 59% y 86%, siendo en todo caso valores inferiores a los mostrados por el biomarcador (Tabla 5.10 y Figura 5.6).

Tabla 5.10 Parámetros del análisis de curvas ROC del biomarcador multi-metabolito de A-ADM y de los metabolitos incluidos en este.

Training set	Especificidad (%)	Sensibilidad (%)	AUC (95% CI)	*P*-valor
Biomarcador multi-metabolito	93,0	94,6	97,7 (95,0-100)	<0.001
Fenilacetilglutamina	85,7	83,8	89,7 (82,7-96,7)	<0.001
p-cresol	59,5	81,1	73,2 (61,7-84,8)	<0.001
Ácido 4-hidroxifenilacético	62,0	73,0	68,1 (56,4-79,9)	0.006

Validation set	Especificidad (%)	Sensibilidad (%)	AUC (95% CI)	*P*-valor
Biomarcador multi-metabolito	95,2	94,7	97,0 (92,3-100)	<0,001
Fenilacetilglutamina	85,7	68,4	81,7 (68,5-94,9)	0,001
p-cresol	76,2	73,7	77,2 (62,4-92,0)	0,003
Ácido 4-hidroxifenilacético	66,7	84,2	77,7 (63,1-92,3)	0,003

Acrónimos y abreviaciones: AUC, área bajo la curva; CI, intervalo de confianza; A-ADM, alta adherencia a la dieta Mediterránea; ROC, característica operativa del receptor.

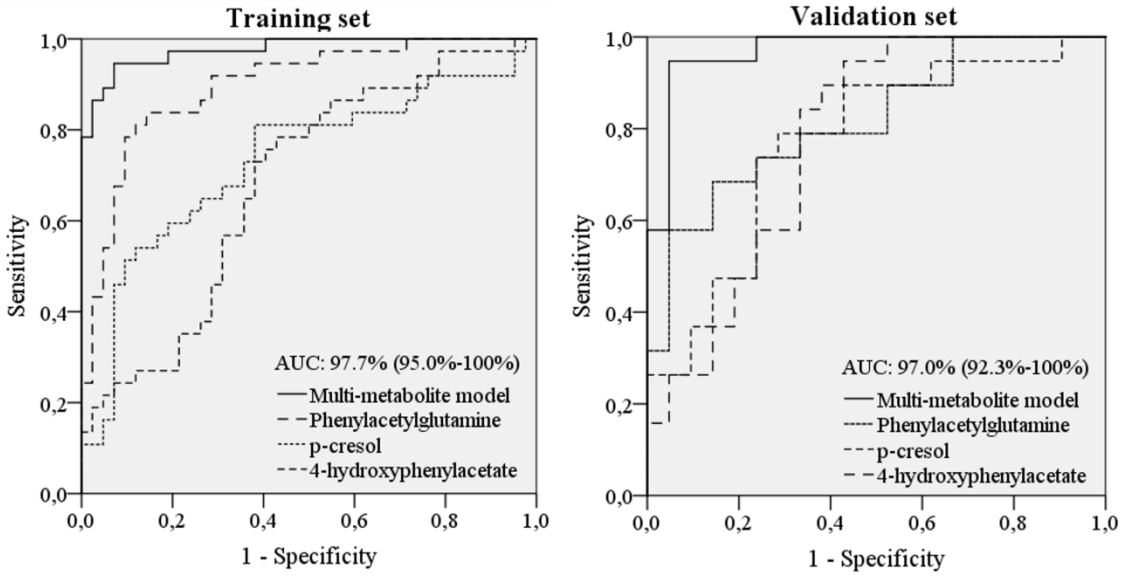


Figura 5.6 Curvas ROC del biomarcador multi-metabolito y los metabolitos individuales incluidos en este para la discriminación entre alta y baja adherencia a la dieta Mediterránea.

5.2.2.4 Relación entre la ingesta de alimentos y los metabolitos asociados de A-ADM

Se exploró la asociación entre la excreción de los metabolitos asociados al grupo de A-ADM (Tabla 5.8) y la ingesta de alimentos reportada a través de los cuestionarios de frecuencia. Tales asociaciones se establecieron mediante la determinación del coeficiente de Spearman, debido a la falta de distribución normal en el biomarcador y los metabolitos, y fijando con un nivel de significancia de $p \leq 0.05$.

Concretamente, se encontraron correlaciones significativas entre la prolina-betaína y el *Scyllo*-inositol con la ingesta de frutas, particularmente cítricos (r=0,36 y r=0,35; p<0,001, respectivamente); la trimetilamina-N-oxido y la ingesta de pescado (r=0,24, p<0,01); y entre la anserina y la ingesta de carne (r=0,231, p<0,01). En un análisis similar, se encontró que el biomarcador de A-ADM correlacionó significativamente con la ingesta de frutas, verduras, legumbres y pescado, así como con la ingesta de la fibra dietética (r = 0,2-0,3; p<0,01) (Tabla 5.11). Cabe mencionar que los metabolitos individuales incluidos en el biomarcador correlacionaron débilmente o no significativamente con los alimentos mencionados. Aunque sí que se observó una fuerte y significativa correlación entre el biomarcador de A-ADM y la puntuación del test de adherencia a la dieta Mediterránea (r = 0,7; p <0,001), mientras que las correlaciones entre los metabolitos del biomarcador y la puntuación del test de adherencia a la dieta Mediterránea fueron moderadas pero también significativas (r = 0,2-0,3; p<0,01) (Tabla 5.11).

Tabla 5.11 Correlaciones entre el biomarcador de A-ADM y los metabolitos incluidos en este con la ingesta de alimentos y la puntuación del test de adherencia a la dieta Mediterránea

Componente	Grupo de alimentos	*r*[1]	*P*-valor[2]
Biomarcador multi-metabolito	Verduras	0,27	0,01
	Frutas	0,27	0,01
	Verduras + frutas	0,34	0,001
	Legumbres	0,24	0.01
	Pescado	0,32	0,002
	DietMed *score*	0,75	<0,001
Fenilacetilglutamina	Verduras	0,29	0,004
	Frutas	0,21	0,03
	Verduras + frutas	0,31	0,002
	Legumbres	0,11	ns
	Pescado	0,19	0,05
	DietMed *score*	0,62	<0,001
p-cresol	Verduras	<0,10	ns
	Frutas	<0,10	ns
	Verduras + frutas	<0,10	ns
	Legumbres	0,20	0,04
	Pescado	0,16	ns
	DietMed *score*	0,37	<0,001
Ácido 4-hidroxifenilacético	Verduras	0,25	0.01
	Frutas	0,17	ns
	Verduras + frutas	0,23	0,02
	Legumbres	0,19	ns
	Pescado	0,30	0,003
	DietMed *score*	0,37	<0,001

[1]coeficiente de correlación de Spearman; [2]significancia corregida para comparaciones múltiples mediante el procedimiento FDR (*false discovery rate*).

Abreviaciones y acrónimos: A-ADM, alta adherencia a la dieta Mediterránea; ns, no significativo (p>0,05).

5.3 BIOMARCADORES DEL EFECTO DE UNA INTERVENCIÓN EN EL ESTILO DE VIDA PARA LA PÉRDA DE PESO EN MUJERES OBESAS METABÓLICAMENTE SANAS

Los resultados del presente estudio se encuentran en vías de publicación en una de las revistas del área.

Enrique Almanza-Aguilera, Carl Brunius, M Rosa Bernal-Lopez, Mar Garcia-Aloy, Francisco J Madrid-Gambin, Francisco J Tinahones, Ricardo Gómez-Huelgas, Rikard Landberg, Cristina Andres-Lacueva. Impact of a lifestyle intervention for weight-loss on plasma metabolome among metabolically healthy obese women. [*Under revision*].

Resumen

Las personas obesas metabólicamente sanas (del inglés, *metabolically healthy* obese, MHO) tienen menor riesgo cardiovascular que su contraparte no saludable. Sin embargo, poco se sabe acerca de los beneficios metabólicos de una intervención en el estilo de vida para la pérdida de peso en este fenotipo. Los objetivos del presente estudio fueron: i) determinar los cambios producidos en el metaboloma plasmático por efecto de una intervención basada en dieta Mediterránea hipocalórica y actividad física en individuos obesos metabólicamente sanos; y ii) estudiar la asociación entre los cambios producidos en el metaboloma y la pérdida de peso tras la intervención con dieta y actividad física en individuos obesos metabólicamente sanos. El estudio metabolómico con aproximación no dirigida mediante RMN consistió en el análisis de muestras plasma de 30 mujeres MHO asignadas a un grupo tratamiento (dieta Mediterránea hipocalórica y actividad física regular) y de 27 mujeres MHO asignadas a un grupo control (recomendaciones generales de dieta saludable y actividad física). Estos análisis fueron realizados en el tiempo basal y después de 3 y 12 meses de intervención, al igual que una serie de comparaciones de las medidas antropométricas, perfiles glucémicos y lipídicos entre los grupos. Distintas técnicas de análisis multivariante se emplearon para determinar los cambios producidos en el metaboloma y la asociación de estos con la pérdida de peso en los dos grupos. En comparación con el grupo de control, el grupo de tratamiento experimentó una mayor pérdida de peso (kg) tanto a los 3 meses (9,10 ± 3,50 vs 3,04 ± 3,62, p <0,001) como a 12 meses (9,98 ± 6,79 vs 4,62 ± 7,22, p <0,01). El análisis multivariante de los cambios en el metaboloma tras la intervención permitió clasificar correctamente a los individuos de acuerdo a su grupo (control o tratamiento) a los 3 meses (86%, *p*<0,001), pero no a los 12 meses (65%). Niveles altos de ácido fórmico y fosfocreatina, así como niveles bajos de trimetilamina y de colesterol LDL/VLDL (señales de RMN) fueron encontrados en común asociados con la intervención y la pérdida de peso. Por otro lado, en comparación con el grupo control, el grupo tratamiento presentó también niveles plasmáticos más altos de mioinositol, metilguanidina y ácido 3-hidroxibutírico, y niveles más bajo de prolina. Mientras que niveles altos de ácido hipúrico y asparagina, y niveles bajos de ácido 2-

hidroxibutírico y creatina fueron encontrados asociados con la pérdida de peso. En conjunto, una intervención en el estilo de vida (basada en una dieta Mediterránea hipocalórica y actividad física) para la pérdida de peso, y la pérdida de peso en sí, modulan positivamente el metabolismo energético y el metabolismo de aminoácidos, lipoproteínas y microbiota. A pesar de haber pérdida de peso a los 12 meses de intervención, la regulación metabólica anteriormente observada a los 3 meses fue atenuada. Se necesitan más estudios para determinar los cambios en el metabolismo y su adaptación en intervenciones de largo plazo.

5.3.1 Resultados

5.3.1.1 Características de los participantes en el tiempo basal

Las medidas antropométricas y los parámetros clínicos de los 57 participantes de acuerdo al grupo control o grupo tratamiento, tanto en el tiempo basal como a los 3 y 12 meses posteriores a la intervención se presentan en la tabla 5.12. La edad promedio de los participantes al inicio fue de 45,1 ± 3,45 años y un IMC de 35,8 ± 4,93 kg/m^2. No se observaron diferencias entre el grupo control y el grupo tratamiento en variables como el peso, circunferencia de la cintura, presión arterial u otros parámetros clínicos, incluyendo glucemia y perfil lipídico (Tabla 5.12).

5.3.1.2 Cambios en las medidas antropométricas y parámetros clínicos

Después de 3 y 12 meses de estudio, se observaron cambios significativos en las medidas antropométricas y parámetros clínicos en ambos grupos. En concreto, después de 3 y 12 meses de estudio los participantes en el grupo tratamiento tuvieron una mayor pérdida de peso comparada con la del grupo control (Tabla 5.12). Se observaron también reducciones en el IMC y la circunferencia de cintura (CC) en ambos grupos tanto a los 3 como a los 12 meses; sin embargo, estas reducciones fueron más pronunciadas en el grupo tratamiento (IMC [kg/m^2] 3 meses: 3,65 ± 1,41 *vs* 1,23 ± 1,47, p <0,001; CC [cm] 3 meses 10,1 ± 5,05 *vs* 5,43 ± 4,91, p <0,001; IMC 12 meses: 4,03 ± 2,77 *vs* 1,84 ± 2,87, p <0,001; CC 12 meses 10,8 ± 7,30 *vs* 0,72 ± 2,50, p <0,001). Comparados con el tiempo basal, a los 3 y 12 meses de intervención ambos grupos experimentaron una disminución el colesterol total y cambios en el colesterol HDL. Específicamente, a los 3 meses de estudio estos cambios fueron un incremento del colesterol HDL en el grupo control y una disminución del mismo en el grupo tratamiento. Por último, comparados con el basal, a los 3 meses, sólo los participantes en el grupo tratamiento tuvieron una disminución en sus niveles de colesterol LDL; y de presión arterial sistólica, glucosa y triglicéridos a los 12 meses. Mientras que también en este tiempo (12 meses) sólo el grupo control mostró una disminución en los niveles de colesterol LDL.

5.3.1.3 Análisis multivariante del efecto de la intervención y la pérdida de peso

Para determinar el efecto de la intervención sobre los cambios producidos en el metaboloma plasmático de los participantes, se realizó un análisis clasificatorio de tipo RF, complementado con el algoritmo rdCV, como se ha descrito en la metodología. Para este análisis se emplearon los cambios producidos en el metaboloma tras cada periodo de intervención (3 y 12 meses), como variable independiente, y la asignación del grupo (control y tratamiento), como variable dependiente. De este modo se obtuvieron dos modelos rdCV-RF, uno para cada periodo de

intervención, y usando el número de clasificaciones erróneas como estadístico para evaluar la calidad de los mismos. La visualización gráfica del número de clasificaciones predichas por los respectivos modelos rdCV-RF se muestran en la figura 5.7. Como se observa en esta figura, el número de clasificaciones erróneas producidas por el modeo rdCV-RF a los 3 meses de intervención fue de 8, lo que equivalió a un 14% del total de los individuos. Dicho de otro modo, después de 3 meses de intervención el modelo rdCV-RF clasificó de manera correcta al 86% a los individuos. Posteriormente, la validez de este modelo se pudo comprobar mediante un test de permutaciones (n=200) que se muestra de manera gráfica en la figura 5.8. En contra parte, el modelo rdCV-RF para los 12 meses de intervención mostró un número superior de clasificaciones erróneas (n=20, 35%), lo que se tradujo en una baja capacidad predictiva del modelo (65%). De esta manera, en vista de la mejor clasificación obtenida a los 3 meses de intervención, el resto de las pruebas estadísticas continuaron con la información correspondiente a este periodo.

En cuanto a la asociación entre los cambios producidos en el metaboloma tras la intervención y la pérdida de peso, esta fue analizada mediante una regresión PLS, e integrando, al igual que en los modelos anteriores, el algoritmo rdCV. Para este análisis se modelaron los cambios producidos en el metaboloma a los 3 meses de intervención, como variable independiente, y el cambio en el peso producido, como variable dependiente. Estos análisis fueron realizados tanto con los datos de ambos grupos (control y tratamiento), como con los datos del grupo tratamiento, sólo.

Tabla 5.12 Medidas antropométricas y variables clínicas de los participantes en el basal y después de 3 y 12 meses de intervención.

	Basal				3 meses			12 meses		
	Todos	Control	Tratamiento	P^1	Control	Tratamiento	P^2	Control	Tratamiento	P^2
Edad, (años)	45,1 ± 3,45	44,4 ± 3,31	45,7 ± 3,51	0,16						
Peso (kg)	90,3 ± 13,5	91,4 ± 15,6	89,3 ± 11,5	0,69	88,3 ± 13,8***	80,2 ± 10,5***	<0,001	86,7 ± 13,5**	79,3 ± 12,3***	0,03
IMC, kg/m2	35,8 ± 4,93	36,3 ± 5,74	35,4 ± 4,12	0,49	35,1 ± 4,89***	31,7 ± 3,67***	<0,001	34,5 ± 4,84**	31,3 ± 4,19***	<0,01
Circunferencia de cintura, cm	112 ± 11,2	110 ± 12,2	114 ± 10,1	0,19	105 ± 10,1***	104 ± 9,48***	<0,001	109 ± 12,5	103 ± 10,7***	<0,001
PAS, mmHg	114 ± 14,1	113 ± 15,0	115 ± 13,4	0,53	ND	ND		111 ± 12,1	110 ± 13,7*	0,35
PAD, mmHg	75,6 ± 9,29	74,7 ± 8,45	76,4 ± 10,1	0,50	ND	ND		74,9 ± 9,10	73,4 ± 10,4	0,25
Glucosa, mg/dL	88,0 ± 8,38	89,5 ± 9,30	86,4 ± 7,29	0,15	87,7 ± 6,55	80,9 ± 5,78	0,24	86,1 ± 7,12	80,4 ± 9,72**	0,28
Colesterol total, mg/mL	196 ± 28,6	195 ± 31,2	197 ± 26,5	0,72	184 ± 40,0*	175 ± 26,2***	0,16	186 ± 32,6**	189 ± 26,0*	0,93
Colesterol LDL, mg/mL	119 ± 27,5	123 ± 27,4	116 ± 27,7	0,43	135 ± 143	105,6 ± 21,8*	0,32	114 ± 27,1**	113 ± 21,6	0,98
Colesterol HDL, mg/mL	56,1 ± 12,1	52,9 ± 11,4	59,1 ± 12,2	0,05	57,0 ± 11,7**	53,2 ± 9,52***	<0,001	52,6 ± 11,0	57,8 ± 14,8	0,93
Triglicéridos, mg/mL	93,2 ± 38,7	95,7 ± 39,4	91,0 ± 38,6	0,67	95,3 ± 48,2	81,4 ± 28,1	0,29	98,9 ± 60,6	82,3 ± 29,5*	0,84

Los datos se presentan como la media ± desviación estándar. Las diferencias con respecto al basal fueron analizadas mediante pruebas pareadas de t-student (*$p \leq 0.05$; **$p \leq 0.01$;***$p \leq 0.001$). P^1, significancia de las comparaciones de los valores medios entre los grupos control y tratamiento en el momento inicial. P^2, significancia de las comparaciones de los cambios producidos (con respecto al basal) entre los grupos control y tratamiento. Todos los valores de significancia (P) fueron corregidos mediante el procedimiento FDR (false discovery rate). Abreviaciones y acrónimos: HDL, lipoproteína de alta densidad; IMC, índice de masa corporal, PAS, presión arterial diastólica; LDL, lipoproteína de baja densidad; ND, datos no disponibles, PAS, presión arterial sistólica.

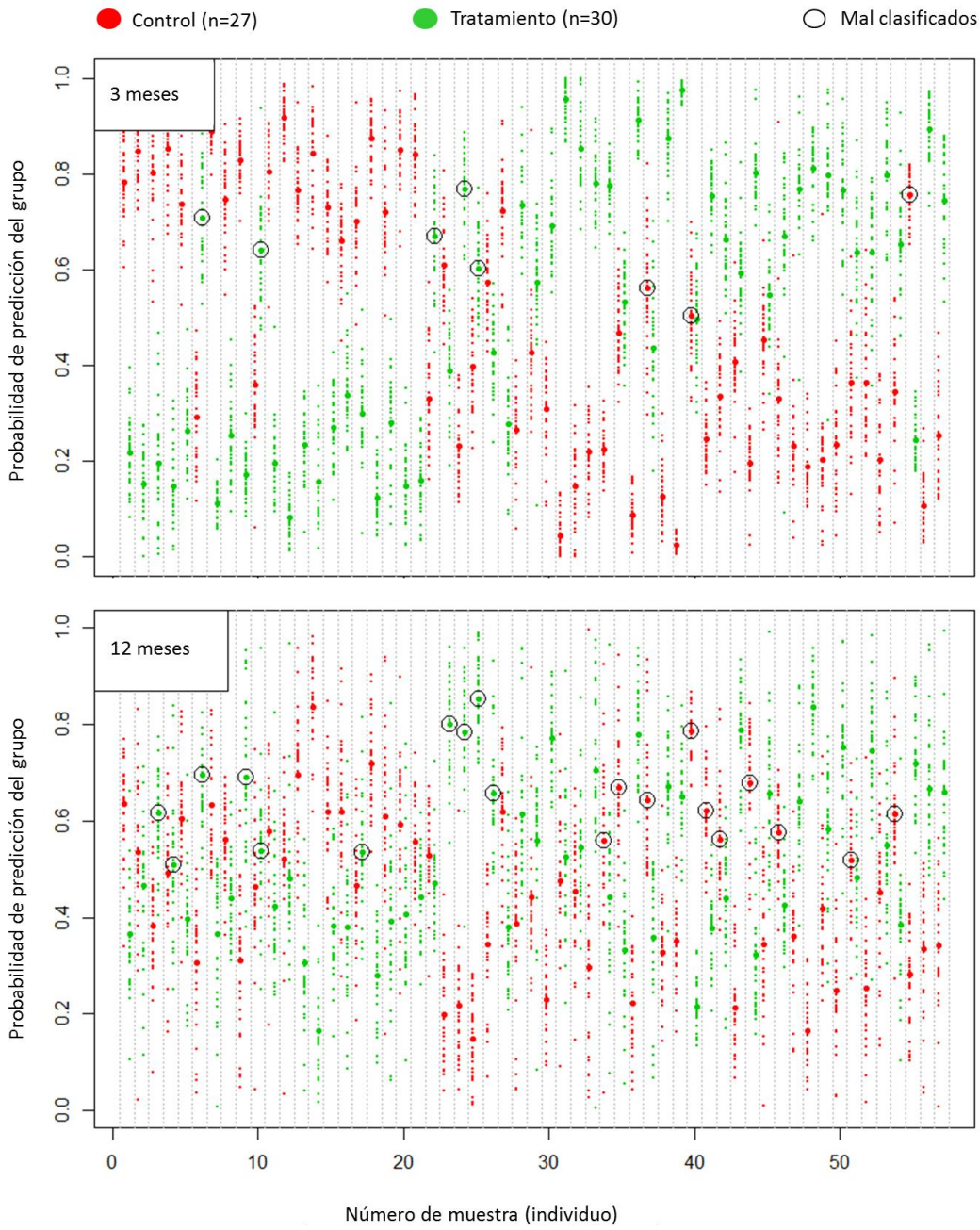


Figura 5.7 Clasificación predictiva de los participantes de acuerdo a los grupos control e intervención a los 3 y 12 meses de estudio. Los modelos *random forest* (RF) se construyeron aplicando un algoritmo de validaciones cruzadas dobles repetidas (rdCV). La probabilidad de clasificación individual de cada sub-modelo (n = 30) se muestra coloreada. La probabilidad de clasificación global por individuo según el grupo se muestra en el círculo más grande dentro de un mismo color. Las personas mal clasificadas están marcadas por un círculo negro.

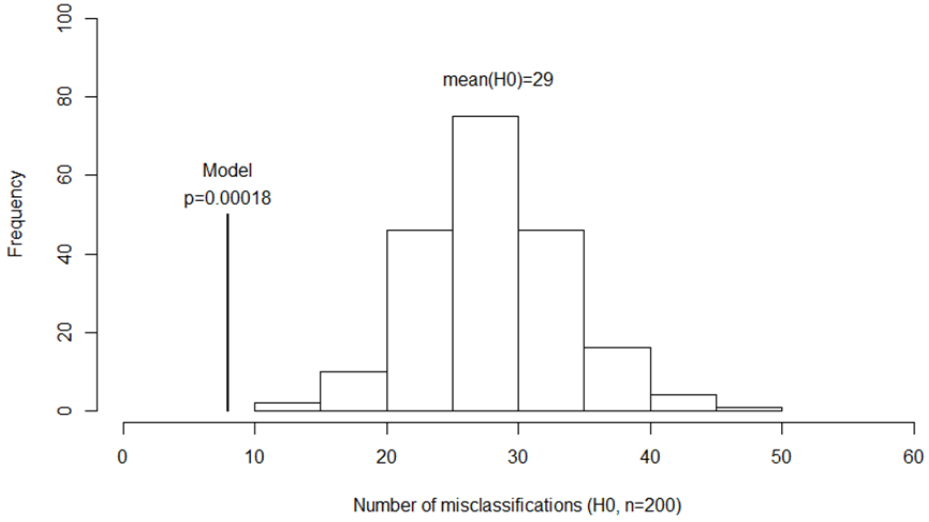


Figura 5.8 Test de permutaciones del modelo rdCV-RF para la clasificación de los participantes de acuerdo a los grupos control e intervención a los 3 meses.

Los estadísticos R^2 y Q^2 fueron empleados como parámetros iniciales de calidad para determinar la capacidad de los modelos rdCV-PLS de encontrar asociaciones entre los cambios producidos en el metaboloma y el cambio de peso. En este sentido los valores de R^2 y Q^2 para el modelo rdCV-PLS en ambos grupos fueron de 0,63 y 0,257, respectivamente; mientras que para el modelo rdCV-PLS con sólo el grupo tratamiento estos valores fueron de 0,744 y 0,298, respectivamente (Figura 5.9, A y B). Sugiriendo en este punto una moderada capacidad predictiva de los respectivos modelos rdCV-PLS. Posteriormente, ambos modelos fueron validados a través de pruebas de permutación, en las que el estadístico Q^2 de los modelos permutados (-0,130 y -0,193, para ambos grupos o sólo grupo tratamiento, respectivamente) fue comparado con el de los modelos originales. Los resultados y significancia (p<0,001) de estas comparaciones se muestran en la figura 5.9 (C y D).

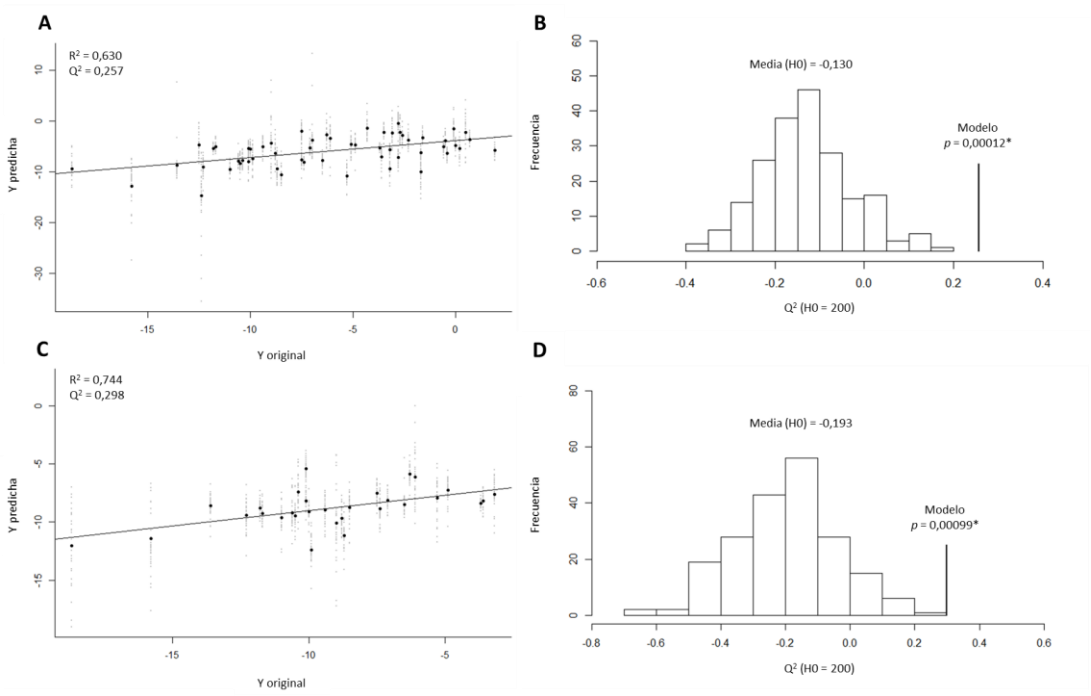


Figura 5.9 Análisis de regresión y *tests* de permutaciones correspondientes a los modelos rdCV-PLS para el cambio de peso en ambos grupos (A y B) o sólo en el grupo tratamiento (C y D). *significancia estadística de un t-test, comparando la media de la distribución de frecuencias de los Q^2 permutados y el Q^2 del modelo original.

5.3.1.4 Efecto modulador de la intervención sobre los metabolitos plasmáticos

La identificación de los metabolitos relacionados con el efecto de la intervención se realizó a partir de la selección de las variables significativas del modelo rdCV-RF correspondiente a los 3 meses de intervención, tal como se indica en el capítulo de metodología. Una vez identificados los metabolitos se calculó el *fold change* de estos en el grupo tratamiento con respecto al grupo intervención. De este modo, los cambios observados en el metaboloma después de 3 meses de intervención incluyeron niveles más altos de ácido 3-hidroxibutírico, ácido fórmico, metilguanidina, *myo*-inositol, fosfocreatina y un metabolito sin identificar (*unknown* U6.40), así como niveles más bajos de prolina, trimetilamina, señales (RMN) de LDL/VLDL y dos metabolitos *unknown* (U3.32 y U4.35) en el grupo tratamiento (Tabla 5.13). Además, los cambios observados en el ácido 3-hidroxibutírico, metilguanidina, fosfocreatina, *myo*-inositol, prolina, U4.35 y U6.40 fueron $\geq$2.0 veces mayores en el grupo tratamiento con respecto al grupo control. Los metabolitos significativos del modelo de clasificación a 3 meses fueron investigados también a los 12 meses mediante pruebas univariantes de *t*-student. De este análisis se obtuvo que sólo la fosfocreatina y U3.32 fueron hallados diferentes entre los grupos,

concretamente con niveles más altos en el grupo intervención. Los cambios en estos metabolitos, sin embargo, fueron inferiores a los observados a 3 meses (Tabla 5.13).

Tabla 5.13 Metabolitos seleccionados del modelo rdCV-RF para la clasificación de individuos según el grupo control o intervención a 3 meses, y un seguimiento de los mismos a los 12 meses.

		3 meses		12 meses	
Desplazamiento químico (δ, ppm) y multiplicidad	Metabolito[1]	*Fold change*[2]	*P*[3]	*Fold change*[2]	*P*[3]
3.32 (s), 3.61 (s)	*Unknown* (U3.32)	0.62	<0.001	0.20	0.01
3.03 (s), 3.94 (s)	Fosfocreatina	2.07	<0.001	0.12	0.01
1.27 (m), 5.32 (m)	LDL/VLDL[4]	0.55	0.01	0.26	0.32
6.40 (m)	*Unknown* (U6.40)	0.14	<0.01	0.10	0.11
2.90 (s)	Trimetilamina	0.53	<0.01	0.38	0.28
1.20 (d), 2.30 (m), 2.40 (m), 4.16 (m)	Ácido 3-hidroxibutúrico	2.21	0.04	0.66	0.69
2.82 (s)	Metilguanidina	3.17	0.10	0.98	0.97
4.35 (s)	*Unknown* (U4.35)	0.21	0.10	0.91	0.93
2.0 (m), 2.06 (m), 2.34 (m), 3.33 (m), 3.41 (m), 4.12 (dd)	Prolina	0.38	0.01	0.41	0.14
3.28 (t), 3.52 (dd), 3.61 (t), 4.05 (t)	*Myo*-inositol	12.5	0.58	0.75	0.78
8.45 (s)	Ácido fórmico	1.60	0.72	0.43	0.16

[1]Metabolitos ordenados en forma decreciente según sus valores VIP. [2]*Fold change* calculado como el cambio medio en el grupo intervención con respecto al grupo control. [3]Significancia obtenida al comparar las medias entre los grupos mediante pruebas *t*-student. Valores *P* ajustados por el procedimiento FDR *(false discovery rate*). [4]resonancias correspondientes a grupos metileno ((CH2)n) y olefinico (-CH=CH-), de los desplazamientos químicos a 1.27 ppm y 5.32 ppm, respectivamente. Abreviaciones: s, singlete; d, doblete; m, multiplete; dd, doble doblete, t, triplete.

5.3.1.5 Cambios en el metaboloma asociados con la pérdida de peso

Un total de 11 metabolitos fueron encontrados significativamente asociados al cambio de peso (y consecuentemente a la pérdida de peso) a los 3 meses en ambos grupos (Tabla 5.14). Específicamente, la pérdida de peso se asoció con niveles más bajos de ácido 2-hidroxibutírico, creatina, señales de LDL/VLDL, trimetilamina y tres metabolitos *Unknown* (carbohidratos, U2.96 y U3.32), y niveles más altos de asparagina, ácido fórmico, ácido hipúrico y fosfocreatina. Interesantemente, los cambios observados en cinco de los metabolitos proporcionados por este modelo fueron encontrados en común con el modelo anterior para evaluar el efecto de la intervención sobre los cambios en el metaboloma. Los metabolitos y las posibles regulaciones metabólicas en común y distintas entre los modelos se ilustran en la figura 5.10.

Tabla 5.14 Metabolitos seleccionados del modelo rdCV-PLS para el cambio de peso desde el basal hasta los 3 meses en los dos grupos, intervención y control.

		Coeficiente de correlación			
Desplazamiento químico, ppm (multiplicidad)	Metabolito[1]	Pearson	P^2	Spearman	P^2
3.03 (s), 3.93 (s)	Fosfocreatina	-0.35	<0.01	-0.39	<0.01
3.32 (s), 3.61 (t)	*Unknown* (U3.32)	0.44	<0.001	0.42	<0.001
2.85 (m), 2.94 (m), 4.0 (dd)	Asparagina	-0.27	0.04	-0.27	0.04
8.44 (s)	Ácido fórmico	-0.24	0.08	-0.31	0.02
0.89 (m), 1.64 (m), 1.74 (m)	Ácido 2-hidroxibutírico	0.35	<0.01	0.32	0.01
3.39 (m), 3.42 (m), 3.68 (m)	*Unknown* (carbohidrato)	0.36	<0.01	0.42	<0.001
0.90 (m), 1.25 (m)	LDL/VLDL[3]	0.32	0.01	0.4	<0.01
3.03 (s), 3.92 (s)	Creatina	0.33	0.01	0.32	0.01
2.90 (s)	Trimetilamina	0.32	0.02	0.28	0.04
2.96 (d)	*Unknown* (U2.96)	0.29	0.03	0.27	0.04
7.62 (t), 7.84 (d)	Ácido hipúrico	-0.23	0.09	-0.29	0.03

[1]Metabolitos ordenados en forma decreciente según sus valores VIP. [2]Significancia obtenida al comparar las medias entre los grupos mediante pruebas *t*-student. Valores *P* ajustados por el procedimiento FDR *(false discovery rate*). [3]resonancias correspondientes a grupos metilo (CH_3) y metileno ($(CH_2)_n$), de los desplazamientos químicos a 0.90 ppm y 1.25 ppm, respectivamente. Abreviaciones: s, singlete; d, doblete; m, multiplete; dd, doble doblete, t, triplete.

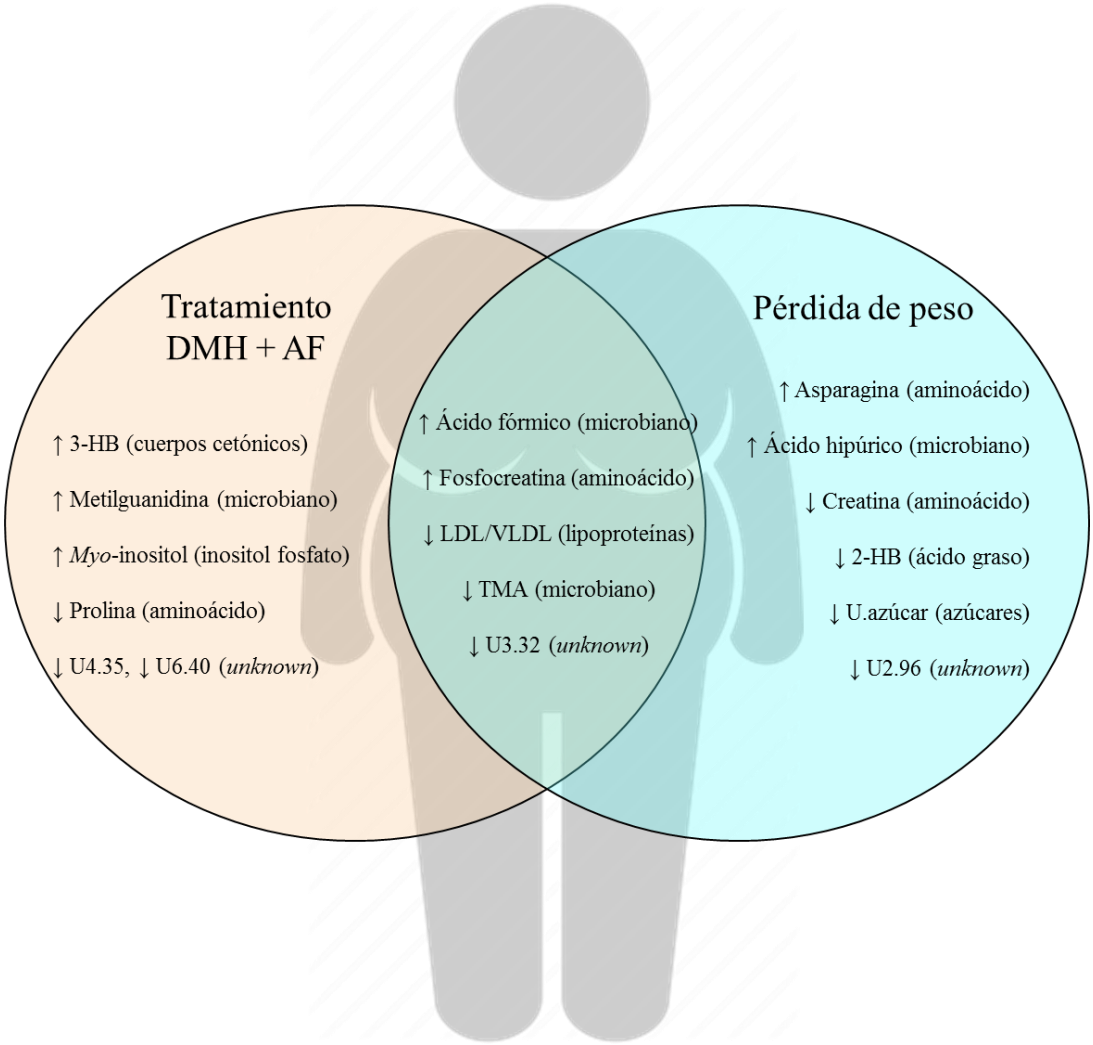


Figura 5.10 Regulaciones metabólicas diferentes y en común asociadas con la intervención y la pérdida de peso en los participantes tras 3 meses de estudio. La clase de cada metabolito se encierra entre paréntesis; la dirección de las flechas (↑) y (↓) denotan que el metabolito se encuentra en niveles más altos o más bajos, respectivamente.

Abreviaciones: 3-HB, ácido 3-hidroxibutírico; 2-HB, ácido 2-hidroxibutírico; AF, actividad física; DMH, dieta Mediterránea hipocalórica; LDL, lipoproteína de baja densidad; VLDL, lipoproteína de muy baja densidad, TMA, trimetilamina.

VI. DISCUSIÓN GLOBAL

La primera parte del trabajo de investigación de la tesis doctoral comprende la identificación de marcadores de un factor de riesgo cardiovascular como es el la DM2. Este estudio incluyó participantes del ensayo PREDIMED en el tiempo inicial del ensayo. Se identificó una huella metabólica de 33 metabolitos que fueron significativamente diferentes entre individuos con DM2 y no-DM2. Más específicamente, sólo ocho de estos metabolitos (alanina, glutamina, metilguanidina, dimetilglicina, y ácidos metilsuccínico, guanidinoacético, 3-hidroximandélico e hipúrico) fueron incluidos en un biomarcador que proporcionó la capacidad discriminante más alta de DM2 entre la población estudiada. Esta capacidad discriminante fue mayor que la mostrada individualmente por cada uno de los metabolitos incluidos en el biomarcador, y que la de la glucosa, la cual no fue incluída en la construcción del biomarcador. Además, mediante un análisis y mapeo de las rutas metabólicas implicadas, se ha podido confirmar lo anteriormente reportado por la literatura, en relación a que la DM2 altera significativamente el metabolismo energético, de aminoácidos y proteínas, así como el metabolismo de la microbiota intestinal (Urpi-Sarda et al. 2015; Gonzalez-Franquesa et al. 2016; Park et al. 2015). Por lo tanto, la huella metabólica y el biomarcador de DM2 reflejan claramente las principales alteraciones metabólicas que diferencian al individuo diabético frente al no diabético.

Los altos niveles de alanina y los bajos niveles de glutamina en orina del diabético podrían estar relacionados con un aumento de la gluconeogénesis (Salek et al. 2007; van Doorn et al. 2007). Curiosamente, en un reciente meta-análisis se observó que estos niveles se encontraban asociados con un incremento en el riesgo de incidencia de prediabetes y DM2 (Guasch-Ferre et al. 2016). También relacionados con el metabolismo de los aminoácidos, el biomarcador de DM2 incluyó niveles altos de ácidos metilsuccínico y guanidinoacético. Estos niveles de ácido metilsuccínico en orina podrían estar relacionados con un incremento del catabolismo de la isoleucina, el cual se ha reportado ampliamente como biomarcador de resistencia a la insulina, obesidad y DM2 (Giesbertz & Daniel 2016; Newgard 2012; Newgard et al. 2009). En el caso del ácido guanidinoacético, un metabolito intemedio en el ciclo de la urea, podría estar elevado en la DM2 debido a un incremento en el catabolismo de las proteínas (Ostojic et al. 2013). Por su parte, los bajos niveles en metilguanidina y los ácidos 3-hidroximandélico e ácido hipúrico, asociados a la DM2, sugieren un impacto en el metabolismo de la microbiota (Moreno-Indias et al. 2014; Ley et al. 2005; Palau-Rodriguez et al. 2015). Específicamente, los niveles bajos de metilguanidina podrían reflejar una menor concentración de su precursor, la creatinina (Wyss & Kaddurah-Daouk 2000), la cual además se encontró también con bajos niveles en los participantes del grupo DM2. Recientemente se ha publicado que una baja excreción urinaria de creatinina se encuentra asociada con diversas causas de mortalidad en individuos con diabetes y nefropatía (Sinkeler et al. 2013). El ácido hipúrico es uno de los metabolitos derivados del

metabolismo microbiano de compuestos fenólicos y fibra dietética (Heinzmann et al. 2012), por lo que sus niveles bajos en orina podrían indicar o bien una menor ingesta de alimentos ricos en polifenoles y fibra (Anderson et al. 2009; Tresserra-Rimbau et al. 2016), o una alteración de la microbiota que lo sintetiza (Lees et al. 2013). Hasta donde se tiene conocimiento, el ácido 3-hidroximandélico, un metabolito derivado del metabolismo microbiano de la tirosina, ha sido reportado previamente en estudios de biomarcadores de DM2. No obstante, en un estudio conducido por Elliot *et al*.,(2015) se observó que la concentración urinaria de este metabolito estuvo inversamente relacionada con la obesidad (Elliott et al. 2015), el cual es un factor de riesgo de DM2 bien conocido (Haslam & James 2005). Los niveles altos de dimetilglicina encontrados asociados a la DM2 en este estudio son consistentes con lo reportado en la literatura (Urpi-Sarda et al. 2015). La dimetilglicina es un metabolito derivado del metabolismo de la colina con propiedades osmorreguladoras que, al encontrarse en altas concentraciones, podría relacionarse con un efecto hiperosmótico de la glucosa, indicando además una disfunción renal (Guan et al. 2013).

En este estudio también se hallaron dos fenotipos metabólicos, o "metabotipos", dentro del grupo de participantes con DM2. En concreto, estos dos metabotipos se diferenciaron por tener diferentes niveles de fenilalanina, fenilacetilglutamina, *p*-cresol y ácido acetoacético. Interesantemente, el metabotipo 2, cuya excreción en estos metabolitos fue mayor que la del metabotipo 1, mostró los niveles de glucosa plasmática más altos. Una primera hipótesis es que los participantes en este metabotipo podrían tener un mayor descontrol o desregulación de la glucemia en sangre, lo cual en periodos más prolongados podría causar mayores complicaciones, incluyendo eventos cardiovasculares (Norhammar et al. 2016). Esta hipótesis además podría reforzarse al tener en cuenta la relación de los metabolitos en el metabotipo 2 con el riesgo cardiovascular reportado en la literatura. Por ejemplo, la fenilalanina se ha descrito como marcador de un alto riesgo de diabetes (Herder et al. 2014; Guasch-Ferre et al. 2016), además de ser incluida junto a la tirosina e isoleucina en índices de predicción a largo plazo de enfermedades y eventos cardiovasculares (Magnusson et al. 2013). Existen estudios que han destacado a la fenilacetilglutamina como un potente e independiente marcador de riesgo de enfermedades cardiovasculares y mortalidad en pacientes con enfermedades crónicas renales (Wurtz et al. 2015). De forma similar, el *p*-cresol, un soluto urémico formado por el metabolismo microbiano de la tirosina, ha sido descrito como predictor de eventos cardiovasculares, independientemente del factor de filtración glomerular en pacientes con moderada enfermedad renal (Meijers et al. 2010). El ácido acetoacético es un cuerpo cetónico que puede ser producido endógenamente a partir de la lisina o la oxidación de ácidos grasos. Diversos estudios han destacado la asociación entre niveles altos de cuerpos cetónicos en sangre

con hiperglicemia y DM2 (Mahendran et al. 2013). Por otro lado, la hipercetonemia y cetosis se han visto relacionadas con complicaciones microvasculares, del hígado y cerebro, lo cual, incrementa el riesgo de morbilidad y mortalidad a largo plazo (Kanikarla-Marie & Jain 2016).

Un segundo estudio que compone esta Tesis Doctoral, y que incluyó también participantes del ensayo clínico PREDIMED, fue el análisis metabolómico con una aproximación no dirigida para descubrir un marcador asociado a la alta adherencia a un patrón de dieta Mediterránea (DietMed). Concretamente, este biomarcador se obtuvo a partir de una huella metabólica urinaria compuesta por 34 metabolitos que discriminaron entre una alta (A-ADM) y baja (B-ADM) adherencia a la DietMed. La clasificación en A-ADM o B-ADM entre los participantes se logró mediante el uso de un cuestionario validado en el estudio PREDIMED (Schroder et al. 2011), y estableciendo los siguientes puntos de corte: $\leq$ 7 puntos, baja adherencia; y $\geq$ 10 puntos para la alta adherencia como previamente se había establecido (Martinez-González et al. 2012). Este perfil identificado de metabolitos permitió además confirmar la asociación de algunos de ellos con la ingesta de alimentos específicos, tal como se ha reportado en la literatura. Tales fueron los casos, por ejemplo, de la prolina-betaína y *Scyllo*-inositol, relacionados con la ingesta de frutas, particularmente cítricos (r=0,36 y r=0,35; p<0,001, respectivamente) (Guertin et al. 2014); o de la trimetilamina-N-óxido y anserina, relacionadas con la ingesta de pescado (r=0,24, p<0,01) y carne blanca (r=0,23, p<0,01), respectivamente. (Mitchell et al. 2008; Schmedes et al. 2016). También se observaron en la A-ADM niveles elevados de metabolitos relacionados con el metabolismo microbiano de compuestos fenólicos, incluyendo los ácidos 4-hidroxihipúrico, 3-(3-hidroxifenil)-3-hidroxipropiónico y 4-hidroxifenilacético. Los altos niveles en orina de estos metabolitos podrían, por lo tanto, reflejar una ingesta mayor de alimentos ricos en compuestos fenólicos en los participantes agrupados en la A-ADM, con respecto a los del grupo B-ADM (Scalbert et al. 2002; Manach et al. 2004; Quideau et al. 2011).

Además de estos metabolitos relacionados con el f*ood metabolome*, también se encontraron otros de origen endógeno y relacionados con el metabolismo energético, caracterizados por niveles más bajos de glucosa, ácidos láctico y succínico en el grupo de A-ADM, en comparación con el grupo de B-ADM. Esto resulta de particular interés dado que aunque algunos de los participantes incluidos en el estudio eran diabéticos, su distribución y la ingesta de medicamentos antidiabéticos fue similar entre los dos grupos. Por lo tanto, estos resultados sugieren que en comparación con los participantes con B-ADM, aquellos con una A-ADM podrían tener relativamente mejor control de su glucemia. Esta hipótesis se apoya en la evidencia generada por diversos estudios en los que se ha demostrado que una A-ADM se asocia con una mejora de los índices glucémicos en poblaciones con riesgo cardiovascular (Salas-Salvadó et al. 2014).

Hace falta resaltar que el biomarcador de A-ADM, con elevada capacidad discriminante, estuvo compuesto por tres metabolitos derivados del metabolismo de la microbiota intestinal: fenilacetilglutamina, *p*-cresol y ácido 4-hidroxifenilacético. De este modo, el papel o influencia de la alta adherencia a un patrón de dieta saludable como la DietMed en la modulación de la microbiota quedó evidenciada por los altos niveles en orina de estos tres metabolitos. Diversos estudios conducidos alrededor de la DietMed confirman esta noción. De Filippis *et al*., por ejemplo, demostraron la relación entre el nivel de adherencia a la DietMed y la composición de la microbiota intestinal, así como de los metabolitos derivados de la misma (De Filippis et al. 2015). En otro estudio, Haro *et al*., demostraron que una intervención a largo plazo con DietMed restauraba parcialmente las alteraciones producidas en la composición de la microbiota en individuos con síndrome metabólico (Haro et al. 2016).

Más recientemente, Gutierrez-Díaz *et al*., (2017) encontraron que las concentraciones fecales de ácidos benzoicos y 3-hidroxifenilacético eran mayores en aquellos individuos con una adherencia alta a la DietMed (Gutierrez-Diaz et al. 2017). A su vez, Vázquez-Fresno et al., encontraron (2015) al ácido 4-hidroxifenilacético en niveles urinarios más altos en un grupo de individuos relativamente "más saludables" en comparación con otro grupo de individuos obesos y diabéticos, después de una intervención con polifenoles de vino tinto. Estos resultados sugirieron una respuesta metabólica distinta postintervención entre los grupos posiblemente asociada con diferencias o alteraciones en el metabolismo de la microbiota intestinal. Estos últimos hallazgos, por lo tanto, resultan interesantes y consistentes con el biomarcador de A-ADM en la presente tesis, al incluir un metabolito similar, el ácido 4-hidroxifenilacético.

Hoy en día, existe un creciente interés en el estudio de la relación entre la adherencia a DietMed y las alteraciones derivadas de los alimentos en relación con los alimentos (Del Chierico et al. 2014). Los tres metabolitos incluidos en el biomarcador de A-ADM pueden provenir de distintas fuentes dietéticas relacionadas con la DietMed misma. La fenilacetilglutamina, formada por la conjugación de la glutamina con el ácido fenilacético, puede provenir de la *β*-oxidación endógena de ácidos grasos con grupos fenilos, así como del metabolismo endógeno de la fenilalanina (Andrew W. Nicholls John N. Haselden, Catherine J. Waterfield 2000), o a través de la ingesta de alimentos de origen vegetal. De hecho, en un estudio reciente se encontró una relación positiva entre las concentraciones de este metabolito en orina y la ingesta de verduras, sugiriendo que la fenilacetilglutamina podría ser un biomarcador de ingesta de este grupo de alimentos (O'Sullivan et al. 2011). El *p*-cresol, como se ha mencionado anteriormente, es un metabolito que puede provenir del metabolismo microbiano de la tirosina con el ácido 4-hidroxifenilacético como metabolito intermedio (Ramakrishna et al. 1991). Respecto al ácido 4-hidroxifenilacético, se ha encontrado que está relacionado con la ingesta de dietas vegetarianas,

y de alimentos ricos en compuestos fenólicos como el vino (Stella et al. 2006) o el chocolate negro (Martin et al. 2009). Por otro lado, la asociación del presente biomarcador con la A-ADM quedó también reflejada en las asociaciones significativas que se encontraron entre el biomarcador y la ingesta de algunos alimentos característicos de una DietMed (frutas, verduras, legumbres, pescado).

Los resultados obtenidos en el presente estudio, y en particular la determinación de un biomarcador de A-ADM, son novedosos desde el punto de vista de las potenciales aplicaciones biológicas en la epidemiología nutricional. En este campo de la nutrición, por ejemplo, el biomarcador propuesto de A-ADM podría ser empleado como una herramienta más exacta para evaluar la ingesta o seguimiento de este patrón alimentario, además de ayudar a entender la relación entre la DietMed y los beneficios a la salud.

En este punto es importante destacar que los resultados obtenidos a partir de las distintas estratificaciones empleadas (DM vs no-DM2 y B-ADM vs A-ADM) con los participantes del mismo ensayo PREDIMED en la presente Tesis Doctoral, sugieren una interesante complementariedad dada por los metabolitos encontrados en común. Por ejemplo, los niveles urinarios altos de glucosa, alanina y ácido láctico, asociados a la DM2, se observaron en niveles bajos en la A-ADM, con respecto a la B-ADM. Esto sugiere, por ejemplo, que a pesar de la desregulación en el metabolismo energético en la DM2, ésta podría verse atenuada o mejorada con una A-ADM. A su vez, otros metabolitos de origen exógeno como el scyllo-inositol, encontrado en bajos niveles en la DM2 y en altos niveles en la A-ADM, sugieren en cambio que mientras en la DM2 se asocia a una baja ingesta de alimentos saludables (p.e., frutas y verduras), en la A-ADM reflejarían o confirmarían la alta ingesta de los mismos.

La descripción metabólica de factores de riesgo cardiovascular y la evaluación del seguimiento de patrones alimentarios cardiosaludables en estudios observacionales, son componentes esenciales en la epidemiología nutricional para el establecimiento de estrategias preventivas o de mejora de riesgo cardiovascular que consistan, por ejemplo, en intervenciones oportunas en el estilo de vida. En este sentido, la determinación de biomarcadores metabolómicos de efecto de tales intervenciones, complementa a los obtenidos de las dos primeras aproximaciones (riesgo cardiovascular y patrón alimentario saludable), en el sentido de que a través de estos se puede profundizar mejor en los mecanismos moleculares que subyacen en la disminución del riesgo cardiovascular.

Por estos motivos, el tercer y último estudio que compone la presente Tesis Doctoral se basó en determinar los cambios producidos en el metaboloma plasmático por efecto de una intervención en el estilo de vida, basada en una DietMed hipocalórica y actividad física regular, en mujeres obesas y definidas como metabólicamente sanas (MHO). En este mismo contexto, se investigó la asociación de la pérdida de peso con los cambios producidos en el metaboloma.

Tal como se esperaba, comparados con los participantes del grupo control (recomendaciones generales de una dieta cardiosaludable y actividad física), los integrantes del grupo tratamiento mostraron una mayor pérdida de peso y cambios más pronunciados en algunos de los parámetros clínicos después de 3 y 12 meses de estudio. Estas pérdidas de peso, sin embargo, fueron similares entre los 3 y 12 meses para cada grupo. Estos primeros resultados confirman lo reportado en la literatura, acerca de que intervenciones para la pérdida de peso basadas en restricciones calóricas y actividad física regular, resultan efectivas en la mejora de diversos marcadores de riesgo cardiovascular en poblaciones obesas (Petelin et al. 2014; Abbenhardt et al. 2013).

Las diferencias en la composición del metaboloma plasmático entre los grupos control e intervención fueron más pronunciadas a los 3 meses de estudio. Una primera hipótesis es que la pérdida de peso entre los 3 y 12 meses fuera similar en cada grupo y/o que después de 12 meses se produjeran diferentes mecanismos compensatorios que atenuaran los cambios en el metaboloma observados a los 3 meses. Entre las distintas variables clínicas y antropométricas analizadas, se encontró que sólo la pérdida de peso se asoció con los cambios producidos en el metaboloma tras 3 meses de intervención. Curiosamente, cinco metabolitos (ácido fórmico, fosfocreatina, trimetilamina, señales de LDL/VLDL y U3.32) de los asociados a la pérdida de peso fueron hallados en común por efecto de la intervención. Así pues, tanto intervención como pérdida de peso se encontraron asociados con varios metabolitos que en su conjunto sugieren un impacto positivo en el metabolismo energético, de aminoácidos, lipoproteínas y microbiota intestinal.

Los niveles más altos en plasma de ácido 3-hidroxibutírico en el grupo intervención, comparados con el control, son consistentes con lo publicado en otros estudios de intervención para la pérdida de peso (Brehm et al. 2005; Gu et al. 2013). El ácido 3-hidroxibutírico es un cuerpo cetónico que se ve aumentado en plasma por un incremento en la lipólisis de ácidos grasos en las mitocondrias hepáticas, lo que ocurre por estados metabólicos de restricción energética, causados por condiciones de ayuno o restricción calórica en la ingesta de alimentos (Cotter et al. 2013). Por lo tanto, los altos niveles en plasma del ácido 3-hidroxibutírico podrían reflejar una homeostasis en el metabolismo energético dado por un incremento en la oxidación

lipídica (Matoulek et al. 2014). Notablemente, en este estudio se encontró también una asociación inversa entre la pérdida de peso y los niveles del ácido 2-hidroxibutírico, un conocido biomarcador temprano de alteraciones en el metabolismo de la glucosa en poblaciones no diabéticas (Gall et al. 2010). Este hallazgo por lo tanto sugiere que una disminución de este metabolito con la pérdida de peso estaría indicando un menor riesgo de desarrollar diabetes (Penn et al. 2013)

Los bajos niveles de prolina y los altos niveles de asparagina en plasma reflejan el impacto de la intervención y la pérdida de peso, respectivamente, en el metabolismo de aminoácidos. Como aminoácido glucogénico, la prolina puede verse disminuida en plasma por un incremento en la gluconeogénesis hepática, la cual suele ser característica de la restricción calórica (Browning et al. 2008) y el ejercicio físico (Donovan & Sumida 1997). Así pues, los niveles plasmáticos más bajos de prolina en el grupo de tratamiento en comparación con el grupo de control se encuentran en concordancia lo reportado en la literatura acerca de que tanto la restricción calórica (Gu et al. 2013) como un aumento de la actividad física (Morris et al. 2013) resultan en una menor concentración de este aminoácido en individuos obesos. Otros estudios también han demostrado una asociación entre la pérdida de peso a largo plazo y los bajos niveles plasmáticos de prolina (Zheng et al. 2016). Esto, sin embargo, no pudo ser corroborado en el presente estudio al no encontrarse una asociación significativa entre este metabolito y la pérdida de peso.

La asociación positiva encontrada entre la asparagina y la pérdida de peso es consistente con otros estudios que indican una relación inversa entre este aminoácido y la obesidad (Cheng et al. 2012; Takashina et al. 2016). Los niveles plasmáticos de asparagina pueden verse afectados por la ingesta de alimentos que contengan este aminoácido o por la síntesis endógena a partir del ácido oxaloacético vía el ácido aspártico. En estudios conducidos con animales se logró demostrar que la suplementación con ácido aspártico y asparagina incrementaban la incorporación de la glucosa y contenido de glucógeno en el músculo esquelético, posiblemente debido a una incorporación de transportadores de glucosa tipo 4 o vesículas al complejo del glucógeno (Lancha et al. 2009). Por lo tanto, aunque la relación de la asparagina con la pérdida de peso resulta incierta, podría sugerirse, por la evidencia citada anteriormente, que un aumento de la asparagina en plasma podría asociarse además con un balance positivo en los niveles de glucosa. Esta hipótesis, sin embargo, debería ser probada en futuras investigaciones.

Hasta este punto, y en conjunto, las asociaciones observadas entre los ácidos 2 hidroxibutírico, 3-hidroxibutíricom, prolina y asparagina con la intervención y la pérdida de peso, sugieren fuertemente un impacto positivo en la homeostasis de la glucosa en el fenotipo MHO, que a su vez podría también traducirse en una disminución de su riesgo cardiovascular.

También relacionadas con el metabolismo de aminoácidos y energético, se encontró a la fosfocreatina en niveles más altos tras la intervención y positivamente asociada con la pérdida de peso, así como una asociación inversa entre la creatina y la pérdida de peso. La creatina es producida principalmente en el hígado y músculo esquelético a partir de los aminoácidos glicina y arginina, y puede ser reversiblemente fosforilada para formar la fosfocreatina con mediación de la enzima creatina quinasa (Wyss & Kaddurah-Daouk 2000). Es probable, entonces, que la contrastante asociación de la creatina y fosfocreatina con la pérdida de peso se deba a un efecto modulador de actividad de la creatina quinasa. Diversos estudios, en este sentido, han mostrado relaciones inversas entre la obesidad (Haan et al. 2017; George et al. 2016) y la pérdida de peso con la expresión de esta enzima (Van Weyenberg et al. 2008), lo cual contribuiría a explicar la hipótesis planteada anteriormente.

La baja intensidad en las señales correspondientes a colesterol LDL/VLDL en el grupo intervención y asociada a la pérdida de peso es también consistente con lo encontrado en otros estudios similares (Dattilo & Kris-Etherton 1992), y podría estar relacionada con un incremento de la expresión del receptor de LDL y la lipoproteína lipasa (Patalay et al. 2005).

En este estudio curiosamente se identificaron cuatro metabolitos derivados del metabolismo de la microbiota: ácidos fórmico e hipúrico, metilguanidina y trimetilamina, encontrándose diferentemente afectados por la intervención, y asociados a la pérdida de peso. Este hallazgo confirma lo reportado en otros estudios que destacan el papel de las interacciones de la microbiota intestinal con el huésped en la composición del peso corporal y la pérdida de peso, ya sea por intervenciones en el estilo de vida (Remely et al. 2015) o por intervenciones quirúrgicas (Liu et al. 2017). La trimetilamina, un metabolito intermedio del metabolismo microbiano de la carnitina y colina obtenidas a través de la dieta, fue encontrada en niveles plasmáticos inferiores tras la intervención e inversamente asociada a la pérdida de peso. La trimetilamina es además precursora de la trimetilamina-N-óxido, la cual ha sido propuesta como marcador de riesgo cardiovascular debido a su actividad proaterogénica demostrada (Wang et al. 2011; Tang et al. 2013). Se sugiere, por lo tanto, que los niveles bajos de trimetilamina en plasma tras la intervención y asociados a la pérdida de peso podrían estar relacionados ya sea con una menor ingesta de sus precursores dietéticos (p.e., huevo y carne) (Tang et al. 2013; Koeth et al. 2013), o debido a un efecto modulador en el metabolismo de la colina y carnitina, y consecuentemente apuntar hacia un menor riesgo cardiovascular debido a una menor disponibilidad para sintetizar la trimetilamina-N-óxido.

Algunos de los metabolitos denominados *unknown* merecen la atención de cara a futuras investigaciones para elucidar su identidad y su papel en los procesos aquí estudiados, como es el

caso del U3.32, que fue afectado por la intervención a los 3 y 12 meses de estudio, además de encontrarse asociado a la pérdida de peso.

El alto número de clasificaciones incorrectas en el modelo para discriminar entre control y tratamiento a 12 meses sugiere una gran similitud en el metaboloma entre los grupos, probablemente debido a un incumplimiento de la intervención o a posibles adaptaciones metabólicas después de los primeros 3 meses en el grupo tratamiento. Se sabe que distintos factores, incluidos fisiológicos, conductuales y del medio ambiente, son clave en el cumplimiento o abandono de intervenciones a largo plazo para la pérdida de peso (Greenberg et al. 2009). Se ha documentado además que, aunque las intervenciones en el estilo de vida pueden ser efectivas en la pérdida de peso y mejora de marcadores cardiovasculares en plazos prolongados de tiempo, la pérdida de peso máxima suele alcanzarse dentro de los primeros seis meses seguida de un variable mantenimiento o ganancia de peso (Lean 2000). En este sentido, se pudo observar en este estudio que la pérdida de peso de los grupos fue similar entre los 3 y 12 meses, lo que apoyaría una hipotética fase de mantenimiento de peso entre estos dos periodos de tiempo. Así, sería además probable que durante el mantenimiento del peso se produjeran las adaptaciones metabólicas suficientes para no permitir la discriminación de los grupos en el largo plazo. En apoyo de esta última hipótesis se pudo comprobar que la mayoría de las diferencias observadas en los parámetros clínicos a los 3 meses disminuyeron a los 12 meses. Sin embargo, al día de hoy no existen reportes acerca de este tipo de adaptación metabólica como resultado de intervenciones de largo plazo para la pérdida de peso.

Por otra parte, aunque los participantes fueron mujeres obesas y definidas *a priori* como metabólicamente sanas, los resultados obtenidos tras a 3 meses de estudio indican que la intervención para la pérdida de peso causó cambios en un sentido hacia un fenotipo más saludable o con un menor riesgo cardiovascular. Estos beneficios, sin embargo, podrían verse atenuados, aún con una pérdida de peso, en el largo plazo. Las razones de esta atenuación no están claras y merecen ser investigadas.

En resumen, los resultados obtenidos de los estudios que constituyen la presente Tesis Doctoral refuerzan la importancia de la aplicación de la metabolómica con una aproximación no dirigida en la identificación y desarrollo de biomarcadores de distintos procesos biológicos normales, como el seguimiento, cumplimiento y/o efecto de un patrón de dieta Mediterránea; y patológicos como la DM2.

VII. CONCLUSIONES

Las siguientes conclusiones se derivan de los objetivos planteados y los resultados obtenidos en la presente Tesis Doctoral:

1. Se ha caracterizado la huella metabólica urinaria (marcadores de enfermedad) asociada a la diabetes mellitus tipo 2 (DM2) en una población de elevado riesgo cardiovascular mediante un estudio metabolómico con aproximación no dirigida por resonancia magnética nuclear (RMN).

2. A partir de esta huella se ha calculado un biomarcador de prevalencia de DM2, el cual estuvo compuesto por altos niveles de ácido metilsuccínico, alanina, dimetilglicina y ácido guanidinoacético, y bajos niveles de glutamina, metilguanidina, y ácidos 3-hidroximandelico e hipúrico. Este biomarcador mostró una alta capacidad para discriminar la DM2 (AUC=96,4%), con una elevada sensibilidad (96,4%) y especificidad (87,0%). Esta capacidad discriminante del biomarcador fue incluso superior a la mostrada por la glucosa (AUC=89,8%; sensibilidad=93,0%; especificidad=82,6%).

3. En conjunto, la huella metabólica y el biomarcador de DM2 permiten confirmar que las principales alteraciones metabólicas se encuentran relacionadas con el metabolismo energético (p.e., glucólisis, gluconeogénesis, ciclo de ácidos tricarboxílicos y cuerpos cetónicos), de aminoácidos y de microbiota intestinal.

4. Se observaron dos fenotipos metabólicos (metabotipos) relacionados con la DM2. Entre estos, el metabotipo con los niveles más altos en orina de fenilalanina, ácido acetoacético, fenilacetilglutamina y *p*-cresol se caracterizó también por tener los niveles de glucosa plasmática más altos. De acuerdo a evidencias previas, este metabotipo dentro de la población con DM2 podría estar relacionado con un mayor riesgo de futuros eventos cardiovasculares.

5. Se identificó la huella metabólica urinaria que permitió diferenciar entre una alta (H-MDA) y baja (L-MDA) adherencia a un patrón de dieta Mediterránea (MedDiet) en individuos con elevado riesgo cardiovascular.

6. Esta huella metabólica incluyó niveles urinarios bajos de algunos metabolitos de origen endógeno (p.e., glucosa) y niveles altos de otros metabolitos de origen exógeno (*food*

metabolome, p.e., prolina betaína), lo que sugiere una amplia modulación del metabolismo endógeno.

7. El biomarcador que permite discriminar entre una H-MDA y L-MDA, con una elevada sensibilidad (95%), especificidad (95%) y AUC (97%), se encuentra compuesto por niveles altos de tres metabolitos de origen microbiano: p-cresol, fenilacetilglutamina y ácido 4-hidroxifenilacético). Este biomarcador mostró además correlaciones significativas con la ingesta de alimentos (frutas, verduras, legumbres y pescado) y fibra dietaria ($r \geq 0{,}30$, $p<0.01$).

8. La huella metabólica urinaria identificada y el biomarcador asociado a la H-MDA reflejan la modulación metabólica de la adherencia a la MedDiet en la regulación del metabolismo de la microbiota intestinal. Además, el biomarcador de H-MDA tiene el potencial de ser empleado como un mejor instrumento para establecer futuras asociaciones entre la alta adherencia a la MedDiet y la prevención o mejora de enfermedades crónicas en estudios de epidemiología nutricional.

9. El efecto de una intervención en el estilo de vida, basada en una MedDiet hipocalórica y actividad física, en el metaboloma plasmático de mujeres obesas metabólicamente sanas (MHO) se pudo observar solamente después de los 3 primeros meses de intervención, pero no después de 12 meses.

10. La pérdida de peso en las mujeres MHO se asoció positivamente con los cambios en los niveles plasmáticos de ácido fórmico y fosfocreatina; y negativamente con los cambios en los niveles de trimetilamina y colesterol LDL/VLDL. Estos cambios en los niveles de estos metabolitos fueron encontrados asociados también con la intervención.

11. Estos resultados sugieren que en conjunto, una intervención basada en el seguimiento de una MedDiet hipocalórica y actividad física regular, y la pérdida de peso en sí, modulan positivamente el metabolismo energético, de aminoácidos, lipoproteínas y de la microbiota en un mujeres MHO. Estas modulaciones sugieren además que los cambios producidos en el metabolismo de este fenotipo se dirigen hacia un estado clínico “más saludable” o de menor riesgo cardiovascular.

12. La falta de cambios observables en el metaboloma tras un periodo de intervención más prolongado (12 meses) podría estar relacionada con una fase de mantenimiento en la pérdida de peso y/o a una adaptación metabólica derivada de la misma.

13. Los resultados obtenidos de los presentes estudios en general refuerzan la importancia de la aplicación de la metabolómica con una aproximación no dirigida en la identificación y desarrollo de biomarcadores de enfermedad y estilos de vida saludables. La identificación de estos biomarcadores permite no sólo profundizar en los mecanismos moleculares que subyacen a los fenómenos estudiados sino también en la relación que guardan éstos con otros procesos biológicos normales o patológicos.

14. El desarrollo de biomarcadores compuestos por más de un metabolito (multimetabolito) resulta hoy en día de gran importancia en estudios epidemiológicos observacionales y de intervención para, por ejemplo, evaluar la cumplimentación de pautas dietéticas (p.e., patrones alimentarios) y evaluar el riesgo y las asociaciones de estos biomarcadores con el desarrollo, prevalencia o mejora de enfermedades crónicas.

CONCLUSIONS

The following conclusions are derived from the objectives and results obtained in the different studies that make up the present Doctoral Thesis:

1. The urinary metabolic fingerprint (disease biomarkers) associated with type 2 diabetes mellitus (DM2) has been characterized in a population at high cardiovascular risk through a metabolomic study with an untargeted approach based on nuclear magnetic resonance (NMR).

2. From this metabolic fingerprint, a biomarker of DM2 prevalence has been calculated, which was composed of high levels of methylsuccinic acid, alanine, dimethylglycine and guanidinoacetic acid, and low levels of glutamine, methylguanidine, and 3-hydroxymandelic and hippuric acids. This biomarker showed a high capacity to discriminate DM2 (AUC = 96.4%), with high sensitivity (96.4%) and specificity (87.0%). This discriminant capacity of the biomarker was even higher than that shown by glucose (AUC = 89.8%, sensitivity = 93.0%, specificity = 82.6%).

3. Together, the metabolic fingerprint and the biomarker of DM2 confirm that the main metabolic alterations are related to energetic metabolism (e.g., glycolysis, gluconeogenesis, tricarboxylic acids cycle and ketone bodies), amino acids and gut microbiota.

4. Two metabolic phenotypes (metabotypes) related to DM2 were observed. Among these, the metabotype with the highest urine levels of phenylalanine, acetoacético acid, phenylacetylglutamine, and p-cresol was also characterized by having the highest plasma glucose levels. According to previous evidence, this metabotype within the population with DM2 could be related to an increased risk of cardiovascular events.

5. A urinary metabolic fingerprint was identified that allowed the difference between a high (H-MDA) and low (L-MDA) adherence to a Mediterranean diet pattern (MedDiet) in individuals with high cardiovascular risk.

6. This metabolic fingerprint included low urinary levels of some endogenous metabolites (e.g., glucose) and high levels of other exogenous metabolites (food metabolome, e.g., proline betaine), suggesting a broad modulation of endogenous metabolism.

7. The biomarker that discriminates between H-MDA and L-MDA, with high sensitivity (95%), specificity (95%) and AUC (97%), is composed of high levels of three metabolites of microbial origin: p-cresol, phenylacetylglutamine and 4-hydroxyphenylacetic acid). This biomarker also showed significant correlations with food intake (fruits, vegetables, legumes and fish) and dietary fiber ($r > 0.30$, $p < 0.01$).

8. The identified urinary metabolic fingerprint and the biomarker associated with H-MDA reflect the metabolic modulation of adherence to MedDiet in regulating the metabolism of the intestinal microbiota. In addition, the H-MDA biomarker has the potential to be used as a better tool to establish future associations between high adherence to MedDiet and the prevention or amelioration of chronic diseases in nutritional epidemiology studies.

9. The effect of a lifestyle intervention, based on a hypocaloric MedDiet and physical activity, on the plasma metabolome of metabolically healthy obese women (MHO) was observed only after the first 3 months of intervention, but not after 12 months.

10. Weight loss in MHO women was positively associated with changes in plasma levels of formic acid and phosphocreatine; and negatively associated with changes in levels of trimethylamine and LDL/VLDL cholesterol. These changes in the levels of these metabolites were also found to be associated with the intervention.

11. These findings suggest that together, an intervention based on a hypocaloric MedDiet and regular physical activity, and weight loss itself, positively modulate the energy, amino acids, lipoproteins and the microbiota metabolism in MHO women. These modulations also suggest that the changes produced in the metabolism of this phenotype are directed towards a "healthier" or lower cardiovascular risk.

12. The lack of observable changes in the metabolome after a longer intervention period (12 months) could be related to a maintenance phase in weight loss and/or to a metabolic adaptation derived from it.

13. In general, the findings obtained from the present studies reinforce the importance of the application of metabolomics with an untargeted approach in the identification and development of disease biomarkers and healthy lifestyles. The identification of these biomarkers allows not only to deep in the molecular mechanisms that underlie the studied phenomena but also in the relationship that they keep with other normal or pathological biological processes.

14. The development of biomarkers composed of more than one metabolite (multimetabolite) is nowadays of great importance in epidemiological observational and interventional studies, for example, to evaluate the compliance of dietary guidelines (e.g., dietary patterns) and to evaluate the risk and associations of these biomarkers with the development, prevalence or improvement of chronic diseases.

VIII. REFERENCIAS

Abbenhardt, C. et al., 2013. Effects of individual and combined dietary weight loss and exercise interventions in postmenopausal women on adiponectin and leptin levels. Journal of internal medicine, 274(2), pp.163–175.

Agnoli, C. et al., 2011. A priori-defined dietary patterns are associated with reduced risk of stroke in a large Italian cohort. The Journal of nutrition, 141(8), pp.1552–1558.

Agnoli, C. et al., 2013. Italian Mediterranean Index and risk of colorectal cancer in the Italian section of the EPIC cohort. International journal of cancer, 132(6), pp.1404–1411.

Ajala, O., English, P. & Pinkney, J., 2013. Systematic review and meta-analysis of different dietary approaches to the management of type 2 diabetes. The American journal of clinical nutrition, 97(3), pp.505–16.

Alkerwi, A. et al., 2015. Cross-comparison of diet quality indices for predicting chronic disease risk: findings from the Observation of Cardiovascular Risk Factors in Luxembourg (ORISCAV-LUX) study. The British journal of nutrition, 113(2), pp.259–269.

Alonso, A., Marsal, S. & Julia, A., 2015. Analytical methods in untargeted metabolomics: state of the art in 2015. Frontiers in bioengineering and biotechnology, 3, p.23.

Altmaier, E. et al., 2011. Questionnaire-based self-reported nutrition habits associate with serum metabolism as revealed by quantitative targeted metabolomics. European Journal of Epidemiology, 26(2), pp.145–156.

Álvarez-Sánchez, B., Priego-Capote, F. & de Castro, M.D.L., 2010. Metabolomics analysis I. Selection of biological samples and practical aspects preceding sample preparation. TrAC Trends in Analytical Chemistry, 29(2), pp.111–119.

Álvarez-Sánchez, B., Priego-Capote, F. & de Castro, M.D.L., 2010. Metabolomics analysis II. Preparation of biological samples prior to detection. TrAC Trends in Analytical Chemistry, 29(2), pp.120–127.

American Diabetes Asociation (ADA), 2017. 2. Classification and Diagnosis of Diabetes. Diabetes care, 40(Suppl 1), pp.S11–S24.

American Diabetes Association (ADA), 2017. Standard of medical care in diabetes - 2017. Diabetes Care, 40 (sup 1)(January), pp.s4–s128.

Andersen, M.-B.S. et al., 2014. Discovery and validation of urinary exposure markers for different plant foods by untargeted metabolomics. Analytical and Bioanalytical Chemistry, 406(7), pp.1829–1844.

Andersen, M.-B.S. et al., 2014. Untargeted metabolomics as a screening tool for estimating compliance to a dietary pattern. Journal of proteome research, 13(3), pp.1405–1418.

Anderson, J.W. et al., 2009. Health benefits of dietary fiber. Nutrition reviews, 67(4), pp.188–205.

Andrew W. Nicholls John N. Haselden, Catherine J. Waterfield, J.K.N., 2000. A metabonomic approach to the investigation of drug-induced phospholipidosis: an NMR spectroscopy and pattern recognition study. Biomarkers, 5(6), pp.410–423.

Arora, T. & Bäckhed, F., 2016. The gut microbiota and metabolic disease: current understanding and future perspectives. Journal of Internal Medicine, 280(4), pp.339–349.

Asghari, G. et al., 2013. Dietary quality among Tehranian adults in relation to lipid profile: findings from the Tehran Lipid and Glucose Study. Journal of health, population, and nutrition, 31(1), pp.37–48.

Babio, N. et al., 2014. Mediterranean diets and metabolic syndrome status in the PREDIMED randomized trial. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne, 186(17), pp.E649-57.

Bach-Faig, A. et al., 2011. Mediterranean diet pyramid today. Science and cultural updates. Public Health Nutr, 14.

Badoud, F. et al., 2015. Molecular insights into the role of white adipose tissue in metabolically unhealthy normal weight and metabolically healthy obese individuals. The FASEB Journal, 29(3), pp.748–758.

Bain, J.R. et al., 2009. Metabolomics applied to diabetes research: moving from information to knowledge. Diabetes, 58(11), pp.2429–2443.

Barrea, L. et al., 2015. Nutrition and psoriasis: is there any association between the severity of the disease and adherence to the Mediterranean diet? Journal of translational medicine, 13, p.18.

Barton, S. et al., 2015. Targeted plasma metabolome response to variations in dietary glycemic load in a randomized, controlled, crossover feeding trial in healthy adults. Food and Function, 6(9), pp.2949–2956.

Beaulac, J. & Sandre, D., 2017. Critical review of bariatric surgery, medically supervised diets, and behavioural interventions for weight management in adults. Perspectives in public health, 137(3), pp.162–172.

Beckman, J. a et al., 2013. Diabetes and vascular disease: pathophysiology, clinical consequences, and medical therapy: part II. European heart journal, 34(31), pp.2444–52.

Beckonert, O. et al., 2007. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nature protocols, 2(11), pp.2692–2703.

Beisken, S., Eiden, M. & Salek, R.M., 2015. Getting the right answers: understanding metabolomics challenges. Expert review of molecular diagnostics, 15(1), pp.97–109.

Bell, J.A. & Hamer, M., 2016. Healthy obesity as an intermediate state of risk: a critical review. Expert Review of Endocrinology & Metabolism, 11(5), pp.403–413.

Bertoli, S. et al., 2015. Adherence to the Mediterranean diet is inversely associated with visceral abdominal tissue in Caucasian subjects. Clinical Nutrition, 34(6), pp.1266–1272.

Blekherman, G. et al., 2011. Bioinformatics tools for cancer metabolomics. Metabolomics : Official journal of the Metabolomic Society, 7(3), pp.329–343.

Blüher, M., 2014. Mechanisms in endocrinology: Are metabolically healthy obese individuals really healthy? European Journal of Endocrinology, 171(6), pp.R209–R219.

Bondia-Pons, I. et al., 2015. Effects of short- and long-term Mediterranean-based dietary treatment on plasma LC-QTOF/MS metabolic profiling of subjects with metabolic syndrome features: The Metabolic Syndrome Reduction in Navarra (RESMENA) randomized controlled trial. Molecular nutrition & food research, 59(4), pp.711–28.

Boto-Ordonez, M. et al., 2013. Microbial metabolomic fingerprinting in urine after regular dealcoholized red wine consumption in humans. Journal of agricultural and food chemistry, 61(38), pp.9166–9175.

Bouchard-Mercier, A. et al., 2013. The metabolic signature associated with the Western dietary pattern: a cross-sectional study. Nutrition journal, 12, p.158.

Brehm, B.J. et al., 2005. The role of energy expenditure in the differential weight loss in obese women on low-fat and low-carbohydrate diets. The Journal of clinical endocrinology and metabolism, 90(3), pp.1475–1482.

Breiman, L., 2001. Random forests. Machine Learning, 45(1), pp.5–32.

Brennan, L., 2014. NMR-based metabolomics: from sample preparation to applications in nutrition research. Progress in nuclear magnetic resonance spectroscopy, 83, pp.42–49.

Brennan, L., Gibbons, H. & O'Gorman, A., 2015. An Overview of the Role of Metabolomics in the Identification of Dietary Biomarkers. Current Nutrition Reports, 4(4), pp.304–312.

Browning, J.D. et al., 2008. Alterations in hepatic glucose and energy metabolism as a result of calorie and carbohydrate restriction. Hepatology (Baltimore, Md.), 48(5), pp.1487–1496.

Buckland, G. et al., 2009. Adherence to the Mediterranean diet and risk of coronary heart disease in the Spanish EPIC Cohort Study. American journal of epidemiology, 170(12), pp.1518–1529.

Buckland, G. et al., 2010. Adherence to a Mediterranean diet and risk of gastric adenocarcinoma within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study. The American journal of clinical nutrition, 91(2), pp.381–390.

Buckland, G. et al., 2011. Adherence to the Mediterranean diet reduces mortality in the Spanish cohort of the European Prospective Investigation into Cancer and Nutrition (EPIC-Spain). The British journal of nutrition, 106(10), pp.1581–1591.

Bujak, R. et al., 2015. Metabolomics for laboratory diagnostics. Journal of pharmaceutical and biomedical analysis, 113, pp.108–120.

Carter, T.C. et al., 2016. Validation of a metabolite panel for early diagnosis of type 2 diabetes. METABOLISM-CLINICAL AND EXPERIMENTAL, 65(9), pp.1399–1408.

Cheng, S. et al., 2012. Metabolite profiling identifies pathways associated with metabolic risk in humans. Circulation, 125(18), pp.2222–2231.

Ciccarone, E. et al., 2003. A high-score Mediterranean dietary pattern is associated with a reduced risk of peripheral arterial disease in Italian patients with Type 2 diabetes. Journal of thrombosis and haemostasis : JTH, 1(8), pp.1744–1752.

Claus, S.P. & Swann, J.R., 2013. Nutrimetabonomics:applications for nutritional sciences, with specific reference to gut microbial interactions. Annual review of food science and technology, 4, pp.381–399.

Cobb, J. et al., 2015. A novel test for IGT utilizing metabolite markers of glucose tolerance. Journal of diabetes science and technology, 9(1), pp.69–76.

Corella, D. & Ordovas, J.M., 2015. Biomarkers: background, classification and guidelines for applications in nutritional epidemiology. Nutricion hospitalaria, 31 Suppl 3, pp.177–188.

Cotter, D.G., Schugar, R.C. & Crawford, P.A., 2013. Ketone body metabolism and cardiovascular disease. American journal of physiology. Heart and circulatory physiology, 304(8), pp.H1060-76.

Couto, E. et al., 2011. Mediterranean dietary pattern and cancer risk in the EPIC cohort. British journal of cancer, 104(9), pp.1493–1499.

Craig, A. et al., 2006. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. Analytical chemistry, 78(7), pp.2262–2267.

Cross, A.J., Major, J.M. & Sinha, R., 2011. Urinary biomarkers of meat consumption. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 20(6), pp.1107–1111.

Cuenca-Garcia, M. et al., 2014. Dietary indices, cardiovascular risk factors and mortality in middle-aged adults: findings from the Aerobics Center Longitudinal Study. Annals of epidemiology, 24(4), p.297–303.e2.

Dattilo, A.M. & Kris-Etherton, P.M., 1992. Effects of weight reduction on blood lipids and lipoproteins: a meta-analysis. The American journal of clinical nutrition, 56(2), pp.320–328.

Davis, C. et al., 2015. Definition of the Mediterranean Diet; a Literature Review. Nutrients, 7(11), pp.9139–9153.

De Filippis, F. et al., 2015. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut.

De Lorgeril, M. et al., 1994. Mediterranean alpha-linolenic acid-rich diet in secondary prevention of coronary heart disease. Lancet (London, England), 343(8911), pp.1454–1459.

De Lorgeril, M. et al., 1999. Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction: final report of the Lyon Diet Heart Study. Circulation, 99(6), pp.779–785.

Del Chierico, F. et al., 2014. Mediterranean diet and health: food effects on gut microbiota and disease control. International journal of molecular sciences, 15(7), pp.11678–11699.

Di Giuseppe, R. et al., 2008. Adherence to Mediterranean diet and anthropometric and metabolic parameters in an observational study in the “Alto Molise” region: the MOLI-SAL project. Nutrition, metabolism, and cardiovascular diseases : NMCD, 18(6), pp.415–421.

Dinu, M. et al., 2017. Mediterranean diet and multiple health outcomes: An umbrella review of meta-analyses of observational studies and randomized trials. Nutrition, Metabolism and Cardiovascular Diseases, 27(1), p.e21.

Dona, A.C. et al., 2016. A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. Computational and Structural Biotechnology Journal, 14, pp.135–153.

Donovan, C.M. & Sumida, K.D., 1997. Training enhanced hepatic gluconeogenesis: the importance for glucose homeostasis during exercise. Medicine and science in sports and exercise, 29(5), pp.628–634.

Duranti, S. et al., 2017. Obesity and microbiota: an example of an intricate relationship. Genes & nutrition, 12, p.18.

Edmands, W.M.B. et al., 2011. Identification of human urinary biomarkers of cruciferous vegetable consumption by metabonomic profiling. Journal of proteome research, 10(10), pp.4513–4521.

Eguaras, S. et al., 2015. Does the Mediterranean diet counteract the adverse effects of abdominal adiposity? Nutrition, Metabolism and Cardiovascular Diseases, 25(6), pp.569–574.

Eid, H.M. et al., 2017. Significance of Microbiota in Obesity and Metabolic Diseases and the Modulatory Potential by Medicinal Plant and Food Ingredients. , 8(June), pp.1–29.

Elliott, P. et al., 2015. Urinary metabolic signatures of human adiposity. Science translational medicine, 7(285), p.285ra62.

Elosua, R. et al., 1994. Validation of the Minnesota Leisure Time Physical Activity Questionnaire in Spanish men. The MARATHOM Investigators. American journal of epidemiology, 139(12), pp.1197–1209.

Esposito, K. et al., 2011. Mediterranean diet and weight loss: meta-analysis of randomized controlled trials. Metabolic syndrome and related disorders, 9(1), pp.1–12.

Estruch, R. et al., 2013. Primary prevention of cardiovascular disease with a Mediterranean diet. The New England journal of medicine, 368(14), pp.1279–1290.

Estruch, R. et al., 2016. Effect of a high-fat Mediterranean diet on bodyweight and waist circumference: a prespecified secondary outcomes analysis of the PREDIMED randomised controlled trial. The Lancet Diabetes & Endocrinology, 4(8), pp.666–676.

Feng, Q. et al., 2016. Integrated metabolomics and metagenomics analysis of plasma and urine identified microbial metabolites associated with coronary heart disease. Scientific reports, 6, p.22525.

Fernandez-Ballart, J.D. et al., 2010. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. The British journal of nutrition, 103(12), pp.1808–1816.

Filzmoser, P., Liebmann, B. & Varmuza, K., 2009. Repeated double cross validation. Journal of Chemometrics, 23(4), pp.160–171.

Floegel, A. et al., 2013. Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. European journal of clinical nutrition, 67(10), pp.1100–1108.

Forshed, J., Schuppe-Koistinen, I. & Jacobsson, S.P., 2003. Peak alignment of NMR signals by means of a genetic algorithm. Analytica Chimica Acta, 487(2), pp.189–199.

Fung, T.T. et al., 2005. Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction. The American journal of clinical nutrition, 82(1), pp.163–173.

Funtikova, A.N. et al., 2014. Mediterranean diet impact on changes in abdominal fat and 10-year incidence of abdominal obesity in a Spanish population. The British journal of nutrition, (2014), pp.1–7.

Gall, W.E. et al., 2010. alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. PloS one, 5(5), p.e10883.

Garcia-Aloy, M. et al., 2014. Novel Multimetabolite Prediction of Walnut Consumption by a Urinary Biomarker Model in a Free-Living Population: the PREDIMED Study. Journal of Proteome Research, 13(7), pp.3476–3483.

Garcia-Aloy, M. et al., 2015. A metabolomics-driven approach to predict cocoa product consumption by designing a multimetabolite biomarker model in free-living subjects from the PREDIMED study. Molecular nutrition & food research, 59(2), pp.212–220.

Garcia-Aloy, M., 2014. Biomarcadores de exposición dietética en estudios nutricionales de intervención y observacionales . Aplicación de una aproximación metabolómica HPLC- q-ToF-MS para la mejora de la capacidad predictiva a través de modelos combinados multi-metabolito OBSERV. University of Barcelona.

Garcia-Perez, I. et al., 2017. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. The lancet. Diabetes & endocrinology, 5(3), pp.184–195.

George, M.D., McGill, N.-K. & Baker, J.F., 2016. Creatine kinase in the U.S. population: Impact of demographics, comorbidities, and body composition on the normal range. Medicine, 95(33), p.e4344.

Gibbons, H. et al., 2017. Metabolomic based identification of clusters that reflect dietary patterns. Molecular nutrition & food research.

Gibbons, H., O'Gorman, A. & Brennan, L., 2015. Metabolomics as a tool in nutritional research. Current opinion in lipidology, 26(1), pp.30–34.

Giesbertz, P. & Daniel, H., 2016. Branched-chain amino acids as biomarkers in diabetes. Current Opinion in Clinical Nutrition and Metabolic Care, 19(1), pp.48–54.

Gonzalez-Franquesa, A. et al., 2016. What Have Metabolomics Approaches Taught Us About Type 2 Diabetes? Current Diabetes Reports, 16(8).

Goodacre, R. et al., 2007. Proposed minimum reporting standards for data analysis in metabolomics. Metabolomics, 3(3), pp.231–241.

Greenberg, I. et al., 2009. Adherence and success in long-term weight loss diets: the dietary intervention randomized controlled trial (DIRECT). Journal of the American College of Nutrition, 28(2), pp.159–168.

Gromski, P.S. et al., 2015. A tutorial review: Metabolomics and partial least squares-discriminant analysis--a marriage of convenience or a shotgun wedding. Analytica chimica acta, 879, pp.10–23.

Grosso, G. et al., 2017. A comprehensive meta-analysis on evidence of Mediterranean diet and cardiovascular disease: Are individual components equal? Critical reviews in food science and nutrition, 57(15), pp.3218–3232.

Gu, Y. et al., 2013. Very low carbohydrate diet significantly alters the serum metabolic profiles in obese subjects. Journal of proteome research, 12(12), pp.5801–5811.

Guan, M. et al., 2013. Systemic perturbations of key metabolites in diabetic rats during the evolution of diabetes studied by urine metabonomics. PloS one, 8(4), p.e60409.

Guasch-Ferre, M. et al., 2013. Mediterranean diet and risk of hyperuricemia in elderly participants at high cardiovascular risk. The journals of gerontology. Series A, Biological sciences and medical sciences, 68(10), pp.1263–1270.

Guasch-Ferre, M. et al., 2016. Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. Diabetes care, 39(5), pp.833–846.

Guertin, K.A. et al., 2014. Metabolomics in nutritional epidemiology: Identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. American Journal of Clinical Nutrition, 100(1), pp.208–217.

Gutierrez-Diaz, I. et al., 2017. Adherence to a Mediterranean Diet Influences the Fecal Metabolic Profile of Microbial-Derived Phenolics in a Spanish Cohort of Middle-Age and Older People. Journal of agricultural and food chemistry, 65(3), pp.586–595.

Haan, Y.C. et al., 2017. Creatine kinase as a marker of obesity in a multi-ethnic population. Molecular and cellular endocrinology, 442, pp.24–31.

Harmon, B.E. et al., 2015. Associations of key diet-quality indexes with mortality in the Multiethnic Cohort: the Dietary Patterns Methods Project. The American journal of clinical nutrition, 101(3), pp.587–597.

Haro, C. et al., 2016. The gut microbial community in metabolic syndrome patients is modified by diet. The Journal of nutritional biochemistry, 27, pp.27–31.

Haro, C., Garcia-Carpintero, S., et al., 2016. The gut microbial community in metabolic syndrome patients is modified by diet. The Journal of nutritional biochemistry, 27, pp.27–31.

Haro, C., Montes-Borrego, M., et al., 2016. Two Healthy Diets Modulate Gut Microbial Community Improving Insulin Sensitivity in a Human Obese Population. The Journal of clinical endocrinology and metabolism, 101(1), pp.233–242.

Haslam, D.W. & James, W.P.T., 2005. Obesity. Lancet (London, England), 366(9492), pp.1197–1209.

Hassan, Y. et al., 2016. Lifestyle interventions for weight loss in adults with severe obesity: a systematic review. Clinical obesity, 6(6), pp.395–403.

Heinzmann, S.S. et al., 2010. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. The American journal of clinical nutrition, 92(2), pp.436–443.

Heinzmann, S.S. et al., 2012. Stability and robustness of human metabolic phenotypes in response to sequential food challenges. Journal of proteome research, 11(2), pp.643–55.

Herder, C. et al., 2014. The potential of novel biomarkers to improve risk prediction of type 2 diabetes. Diabetologia, 57(1), pp.16–29.

Hernandez-Ruiz, A. et al., 2015. DESCRIPTION OF INDEXES BASED ON THE ADHERENCE TO THE MEDITERRANEAN DIETARY PATTERN: A REVIEW. Nutricion hospitalaria, 32(5), pp.1872–1884.

Hoevenaar-Blom, M.P. et al., 2012. Mediterranean style diet and 12-year incidence of cardiovascular diseases: the EPIC-NL cohort study. PloS one, 7(9), p.e45458.

Hu, F.B., 2002. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol, 13.

International Diabetes Dederation, 2015. Atlas de la Diabetes de la FID. Séptima Edición.,

Jacobs, S. et al., 2015. A priori-defined diet quality indexes and risk of type 2 diabetes: the Multiethnic Cohort. Diabetologia, 58(1), pp.98–112.

Kanikarla-Marie, P. & Jain, S.K., 2016. Hyperketonemia and ketosis increase the risk of complications in type 1 diabetes. Free radical biology & medicine, 95, pp.268–277.

Kastorini, C.-M. et al., 2011. The effect of Mediterranean diet on metabolic syndrome and its components: a meta-analysis of 50 studies and 534,906 individuals. Journal of the American College of Cardiology, 57(11), pp.1299–313.

Kell, D.B. et al., 2005. Metabolic footprinting and systems biology: the medium is the message. Nature reviews. Microbiology, 3(7), pp.557–565.

Keys A, Aravanis C, Blackburn H, Buzina R, Djordjevic BS, Dontas AS, Fidanza F, Karvonen MJ, Kimura N, Menotti A, Mohacek I, Nedeljkovic S, Puddu V, Punsar S, Taylor HL, V.B.F., 1980. A multivariate analysis of death and coronary heart disease, Cambridge, MA: Harvard University Press.

Khakimov, B. et al., 2016. New Nordic Diet versus Average Danish Diet: A Randomized Controlled Trial Revealed Healthy Long-Term Effects of the New Nordic Diet by GC-MS Blood Plasma Metabolomics. Journal of proteome research, 15(6), pp.1939–1954.

Klein, M.S. & Shearer, J., 2016. Metabolomics and Type 2 Diabetes: Translating Basic Research into Clinical Application. Journal of diabetes research, 2016, p.3898502.

Koeth, R.A. et al., 2013. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. Nature medicine, 19(5), pp.576–585.

Koloverou, E. et al., 2014. The effect of Mediterranean diet on the development of type 2 diabetes mellitus: A meta-analysis of 10 prospective studies and 136,846 participants. Metabolism, 63(7), pp.903–911.

Lancha, A.H.J., Poortmans, J.R. & Pereira, L.O., 2009. The effect of 5 days of aspartate and asparagine supplementation on glucose transport activity in rat muscle. Cell biochemistry and function, 27(8), pp.552–557.

Lankinen, M. et al., 2016. A Healthy Nordic Diet Alters the Plasma Lipidomic Profile in Adults with Features of Metabolic Syndrome in a Multicenter Randomized Dietary Intervention. The Journal of nutrition.

Lavie, C.J. et al., 2016. Obesity and Prevalence of Cardiovascular Diseases and Prognosis-The Obesity Paradox Updated. Progress in cardiovascular diseases, 58(5), pp.537–547.

Lean, M.E., 2000. Is long-term weight loss possible? The British journal of nutrition, 83 Suppl 1, pp.S103-11.

Lees, H.J. et al., 2013. Hippurate: the natural history of a mammalian-microbial cometabolite. Journal of proteome research, 12(4), pp.1527–46.

Ley, R.E. et al., 2005. Obesity alters gut microbial ecology. Proceedings of the National Academy of Sciences of the United States of America, 102(31), pp.11070–11075.

Liu, R. et al., 2017. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. Nature medicine.

Lloyd, A.J. et al., 2013. Data-driven strategy for the discovery of potential urinary biomarkers of habitual dietary exposure. The American journal of clinical nutrition, 97(2), pp.377–89.

Madrid-Gambin, F. et al., 2017. Urinary 1H Nuclear Magnetic Resonance Metabolomic Fingerprinting Reveals Biomarkers of Pulse Consumption Related to Energy-Metabolism Modulation in a Subcohort from the PREDIMED study. Journal of proteome research, 16(4), pp.1483–1491.

Magnusson, M. et al., 2013. A diabetes-predictive amino acid score and future cardiovascular disease. European heart journal, 34(26), pp.1982–1989.

Mahendran, Y. et al., 2013. Association of ketone body levels with hyperglycemia and type 2 diabetes in 9,398 Finnish men. Diabetes, 62(10), pp.3618–3626.

Manach, C. et al., 2004. Polyphenols: food sources and bioavailability. The American journal of clinical nutrition, 79(5), pp.727–747.

Mandviwala, T., Khalid, U. & Deswal, A., 2016. Obesity and Cardiovascular Disease: a Risk Factor or a Risk Marker? Current atherosclerosis reports, 18(5), p.21.

Marklund, M. et al., 2014. A dietary biomarker approach captures compliance and cardiometabolic effects of a healthy Nordic diet in individuals with metabolic syndrome. The Journal of nutrition, 144(10), pp.1642–9.

Martin, F.-P.J. et al., 2009. Metabolic effects of dark chocolate consumption on energy, gut microbiota, and stress-related metabolism in free-living subjects. Journal of proteome research, 8(12), pp.5568–5579.

Martinez-Gonzalez, M.A. & Martin-Calvo, N., 2016. Mediterranean diet and life expectancy; beyond olive oil, fruits, and vegetables. Current opinion in clinical nutrition and metabolic care, 19(6), pp.401–407.

Martinez-Gonzalez, M.A. et al., 2002. Mediterranean diet and reduction in the risk of a first acute myocardial infarction: an operational healthy dietary score. European journal of nutrition, 41(4), pp.153–160.

Martinez-Gonzalez, M.A. et al., 2009. Mediterranean food pattern and the primary prevention of chronic disease: recent developments. Nutrition reviews, 67 Suppl 1, pp.S111-6.

Martinez-González, M.A. et al., 2012. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. PloS one, 7(8), p.e43134.

Martínez-González, M.Á. et al., 2012. Cohort profile: design and methods of the PREDIMED study. International journal of epidemiology, 41(2), pp.377–85.

Martinez-Gonzalez, M.A. et al., 2016. Intervention Trials with the Mediterranean Diet in Cardiovascular Prevention: Understanding Potential Mechanisms through Metabolomic Profiling. The Journal of nutrition.

Matoulek, M. et al., 2014. Post-exercise changes of beta hydroxybutyrate as a predictor of weight changes. Physiological research, 63 Suppl 2, pp.S321-5.

Meijers, B.K.I. et al., 2010. p-Cresol and cardiovascular risk in mild-to-moderate kidney disease. Clinical journal of the American Society of Nephrology : CJASN, 5(7), pp.1182–1189.

Menni, C. et al., 2013. Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. Metabolomics : Official journal of the Metabolomic Society, 9(2), pp.506–514.

Mitchell, S.C., Zhang, A.Q. & Smith, R.L., 2008. Dimethylamine and diet. Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association, 46(5), pp.1734–1738.

Moreno-Indias, I. et al., 2014. Impact of the gut microbiota on the development of obesity and type 2 diabetes mellitus. Frontiers in microbiology, 5, p.190.

Morris, C. et al., 2013. The relationship between aerobic fitness level and metabolic profiles in healthy adults. Molecular nutrition & food research, 57(7), pp.1246–1254.

Mosharraf, S. et al., 2013. Impact of the components of Mediterranean nutrition regimen on long-term prognosis of diabetic patients with coronary artery disease. ARYA atherosclerosis, 9(6), pp.337–342.

Muñoz-Garach, A., Cornejo-Pareja, I. & Tinahones, F.J., 2016. Does metabolically healthy obesity exist? Nutrients, 8(6).

Najafi, M. & Sheikhvatan, M., 2013. Gender differences in coronary artery disease: correlational study on dietary pattern and known cardiovascular risk factors. International cardiovascular research journal, 7(4), pp.124–129.

Newgard, C.B. et al., 2009. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell metabolism, 9(4), pp.311–26.

Newgard, C.B., 2012. Interplay between lipids and branched-chain amino acids in development of insulin resistance. Cell metabolism, 15(5), pp.606–14.

Nigg, C.R. et al., 1999. Stages of change across ten health risk behaviors for older adults. The Gerontologist, 39(4), pp.473–482.

Norhammar, A. et al., 2016. Incidence, prevalence and mortality of type 2 diabetes requiring glucose-lowering treatment, and associated risks of cardiovascular complications: a nationwide study in Sweden, 2006-2013. Diabetologia, 59(8), pp.1692–1701.

O'Gorman, A. & Brennan, L., 2015. Metabolomic applications in nutritional research: A perspective. Journal of the Science of Food and Agriculture, 95(13), pp.2567–2570.

O'Gorman, A. et al., 2014. Habitual dietary intake impacts on the lipidomic profile. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences, 966, pp.140–146.

O'Sullivan, A., Gibney, M.J. & Brennan, L., 2011. Dietary intake patterns are reflected in metabolomic profiles: Potential role in dietary assessment studies. American Journal of Clinical Nutrition, 93(2), pp.314–321.

Odriozola, L. & Corrales, F.J., 2015. Discovery of nutritional biomarkers: future directions based on omics technologies. International journal of food sciences and nutrition, 66 Suppl 1, pp.S31-40.

Organización Mundial de la Salud (OMS), 2016. Obesidad y Sobrepeso. Available at: http://www.who.int/mediacentre/factsheets/fs311/es/ [Accessed May 16, 2017].

Organización Mundial de la Salud (OMS), 2017. Enfermedades cardiovasculares. ¿Qué son las enfermedades cardiovasculares? Available at: http://www.who.int/cardiovascular_diseases/about_cvd/es/ [Accessed May 16, 2017].

Organización Mundial de la Salud, 2016. 10 datos sobre la diabetes. Available at: http://www.who.int/features/factfiles/diabetes/es/ [Accessed June 28, 2017].

Organization, W.W.H.F.W.S. ed., 2011. Global atlas on cardiovascular disease prevention and control. Policies, strategies and interventions,

Ostojic, S.M. et al., 2013. Creatine Metabolism and Safety Profiles after Six-Week Oral Guanidinoacetic Acid Administration in Healthy Humans. International Journal of Medical Sciences, 10(2), pp.141–147.

Palau-Rodriguez, M. et al., 2015. Metabolomic insights into the intricate gut microbial-host interaction in the development of obesity and type 2 diabetes. Frontiers in microbiology, 6, p.1151.

Panagiotakos, D.B. et al., 2007. Adherence to the Mediterranean food pattern predicts the prevalence of hypertension, hypercholesterolemia, diabetes and obesity, among healthy adults; the accuracy of the MedDietScore. Preventive medicine, 44(4), pp.335–340.

Paneni, F. et al., 2013. Diabetes and vascular disease: pathophysiology, clinical consequences, and medical therapy: part I. European heart journal, 34(31), pp.2436–2443.

Paneni, F., Costantino, S. & Cosentino, F., 2014. Insulin resistance, diabetes, and cardiovascular risk. Current atherosclerosis reports, 16(7), p.419.

Park, S., Sadanala, K.C. & Kim, E.-K., 2015. A Metabolomic Approach to Understanding the Metabolic Link between Obesity and Diabetes. Molecules and cells, 38(7), pp.587–596.

Park, Y.-M. et al., 2017. Obesity Mediates the Association between Mediterranean Diet Consumption and Insulin Resistance and Inflammation in US Adults. The Journal of Nutrition, 147(4), pp.563–571.

Patalay, M. et al., 2005. The lowering of plasma lipids following a weight reduction program is related to increased expression of the LDL receptor and lipoprotein lipase. The Journal of nutrition, 135(4), pp.735–739.

Patti, G.J., Yanes, O. & Siuzdak, G., 2012. Innovation: Metabolomics: the apogee of the omics trilogy. Nature reviews. Molecular cell biology, 13(4), pp.263–269.

Penn, L. et al., 2013. Importance of weight loss maintenance and risk prediction in the prevention of type 2 diabetes: analysis of European Diabetes Prevention Study RCT. PloS one, 8(2), p.e57143.

Perez-Martinez, P. et al., 2017. Lifestyle recommendations for the prevention and management of metabolic syndrome: an international panel recommendation. Nutrition reviews, 75(5), pp.307–326.

Petelin, A. et al., 2014. Low-grade inflammation in overweight and obese adults is affected by weight loss program. Journal of endocrinological investigation, 37(8), pp.745–755.

Phillips, C.M., 2017. Metabolically healthy obesity across the life course: epidemiology, determinants, and implications. Annals of the New York Academy of Sciences, 1391(1), pp.85–100.

Piccolo, B.D. et al., 2015. Habitual physical activity and plasma metabolomic patterns distinguish individuals with low vs. high weight loss during controlled energy restriction. The Journal of nutrition, 145(4), pp.681–690.

Pitsavos, C. et al., 2005. Adherence to the Mediterranean diet is associated with total antioxidant capacity in healthy adults: the ATTICA study. The American journal of clinical nutrition, 82(3), pp.694–699.

Playdon, M.C. et al., 2017. Identifying biomarkers of dietary patterns by using metabolomics. American Journal of Clinical Nutrition, 105(2), pp.450–465.

Poirier, P. et al., 2006. Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical. Circulation, 113(6), pp.898–918.

Praud, D. et al., 2014. Adherence to the Mediterranean diet and gastric cancer risk in Italy. International journal of cancer, 134(12), pp.2935–2941.

Psaltopoulou, T. et al., 2004. Olive oil, the Mediterranean diet, and arterial blood pressure: the Greek European Prospective Investigation into Cancer and Nutrition (EPIC) study. The American journal of clinical nutrition, 80(4), pp.1012–1018.

Pujos-Guillot, E. et al., 2017. Systems Metabolomics for Prediction of Metabolic Syndrome. Journal of proteome research.

Quideau, S. et al., 2011. Plant polyphenols: chemical properties, biological activities, and synthesis. Angewandte Chemie (International ed. in English), 50(3), pp.586–621.

Radjursoga, M. et al., 2017. Metabolic profiles from two different breakfast meals characterized by 1H NMR-based metabolomics. Food chemistry, 231, pp.267–274.

Ramakrishna, B.S. et al., 1991. Impaired sulphation of phenol by the colonic mucosa in quiescent and active ulcerative colitis. Gut, 32(1), pp.46–49.

Rasmussen, L.G. et al., 2012. Assessment of dietary exposure related to dietary GI and fibre intake in a nutritional metabolomic study of human urine. Genes & nutrition, 7(2), pp.281–293.

Reedy, J. et al., 2014. Higher diet quality is associated with decreased risk of all-cause, cardiovascular disease, and cancer mortality among older adults. The Journal of nutrition, 144(6), pp.881–889.

Remely, M. et al., 2015. Gut microbiota composition correlates with changes in body fat content due to weight loss. Beneficial microbes, 6(4), pp.431–439.

Romaguera, D. et al., 2010. Mediterranean dietary patterns and prospective weight change in participants of the EPIC-PANACEA project. The American journal of clinical nutrition, 92(4), pp.912–921.

Romaguera, D. et al., 2011. Mediterranean diet and type 2 diabetes risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) study: the InterAct project. Diabetes care, 34(9), pp.1913–1918.

Romieu, I. et al., 2017. Energy balance and obesity: what are the main drivers? Cancer Causes and Control, 28(3), pp.247–258.

Saccenti, E. et al., 2014. Reflections on univariate and multivariate analysis of metabolomics data. Metabolomics, 10(3), pp.361–374.

Salas-Salvadó, J. et al., 2014. Prevention of diabetes with Mediterranean diets: a subgroup analysis of a randomized trial. Annals of internal medicine, 160(1), pp.1–10.

Salas-Salvadó, J. et al., 2016. Protective effects of the mediterranean diet on type 2 diabetes and metabolic syndrome. J Nutr.

Salek, R.M. et al., 2007. A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. , pp.99–108.

Sanchez-Villegas, A. et al., 2006. Adherence to a Mediterranean dietary pattern and weight gain in a follow-up study: the SUN cohort. International journal of obesity (2005), 30(2), pp.350–358.

Scalbert, A. et al., 2002. Absorption and metabolism of polyphenols in the gut and impact on health. Biomedicine & pharmacotherapy = Biomédecine & pharmacothérapie, 56(6), pp.276–82.

Scalbert, A. et al., 2014. The food metabolome: a window over dietary exposure. The American journal of clinical nutrition, 99(6), pp.1286–1308.

Schmedes, M. et al., 2016. Lean-seafood intake decreases urinary markers of mitochondrial lipid and energy metabolism in healthy subjects: metabolomics results from a randomized crossover intervention study. Molecular nutrition & food research.

Schmidt, J.A. et al., 2015. Metabolic profiles of male meat eaters, fish eaters, vegetarians, and vegans from the EPIC-Oxford cohort. The American journal of clinical nutrition, 102(6), pp.1518–26.

Schroder, H. et al., 2004. Adherence to the traditional mediterranean diet is inversely associated with body mass index and obesity in a spanish population. The Journal of nutrition, 134(12), pp.3355–3361.

Schroder, H. et al., 2011. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. J Nutr, 141.

Schroder, H., Fito, M. & Covas, M.I., 2007. Association of fast food consumption with energy intake, diet quality, body mass index and the risk of obesity in a representative Mediterranean population. The British journal of nutrition, 98(6), pp.1274–1280.

Schroder, H., Marrugat, J. & Covas, M.I., 2006. High monetary costs of dietary patterns associated with lower body mass index: a population-based study. International journal of obesity (2005), 30(10), pp.1574–1579.

Schulte, P.A., 2005. The use of biomarkers in surveillance, medical screening, and intervention. Mutation research, 592(1–2), pp.155–163.

Schwingshackl, L. et al., 2015. Adherence to a Mediterranean diet and risk of diabetes: a systematic review and meta-analysis. Public health nutrition, 18(7), pp.1292–1299.

Scientific Report of the 2015 Dietary Guidelines Advisory Committee [DGAC], 2015. Part D. Chapter 2: Dietary Patterns, Foods and Nutrients, and Health Outcomes. Available at: http://health.gov/dietaryguidelines/2015-scientific-report/ [Accessed April 1, 2016].

Shen, J. et al., 2015. Mediterranean Dietary Patterns and Cardiovascular Health. Annual review of nutrition, 35, pp.425–449.

Sinkeler, S.J. et al., 2013. Creatinine excretion rate and mortality in type 2 diabetes and nephropathy. Diabetes care, 36(6), pp.1489–1494.

Sofi, F. et al., 2013. Mediterranean diet and health. BioFactors (Oxford, England), 39(4), pp.335–342.

Sofi, F. et al., 2014. Mediterranean diet and health status: an updated meta-analysis and a proposal for a literature-based adherence score. Public health nutrition, 17(12), pp.2769–82.

Sonnenburg, J.L. & Bäckhed, F., 2016. Diet–microbiota interactions as moderators of human metabolism. Nature, 535(7610), pp.56–64.

Sotos-Prieto, M. et al., 2015. Design and development of an instrument to measure overall lifestyle habits for epidemiological research: the Mediterranean Lifestyle (MEDLIFE) index. Public health nutrition, 18(6), pp.959–967.

Sousa, S.A.A., Magalhães, A. & Ferreira, M.M.C., 2013. Optimized bucketing for NMR spectra: Three case studies. Chemometrics and Intelligent Laboratory Systems, 122, pp.93–102.

Stella, C. et al., 2006. Susceptibility of Human Metabolic Phenotypes to Dietary Modulation. J Proteome Res, 5(10), pp.2780–2788.

Strimbu, K. & Tavel, J.A., 2010. What are biomarkers? Current opinion in HIV and AIDS, 5(6), pp.463–466.

Takashina, C. et al., 2016. Associations among the plasma amino acid profile, obesity, and glucose metabolism in Japanese adults with normal glucose tolerance. Nutrition & metabolism, 13, p.5.

Tang, W.H.W. et al., 2013. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. The New England journal of medicine, 368(17), pp.1575–1584.

Teixeira, T.F.S. et al., 2015. Main characteristics of metabolically obese normal weight and metabolically healthy obese phenotypes. Nutrition Reviews, 73(3), pp.175–190.

Thomson, C.A. et al., 2003. Measuring dietary change in a diet intervention trial: comparing food frequency questionnaire and dietary recalls. American journal of epidemiology, 157(8), pp.754–762.

Tresserra-Rimbau, A. et al., 2016. Intake of Total Polyphenols and Some Classes of Polyphenols Is Inversely Associated with Diabetes in Elderly People at High Cardiovascular Disease Risk. The Journal of nutrition.

Trichopoulou, A. et al., 1995. Diet and overall survival in elderly people. BMJ (Clinical research ed.), 311(7018), pp.1457–1460.

Trichopoulou, A. et al., 2003. Adherence to a Mediterranean diet and survival in a Greek population. The New England journal of medicine, 348(26), pp.2599–2608.

Trichopoulou, A. et al., 2005. Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. BMJ (Clinical research ed.), 330(7498), p.991.

Trichopoulou, A. et al., 2014. Definitions and potential health benefits of the Mediterranean diet: views from experts around the world. BMC medicine, 12, p.112.

Trygg, J., Holmes, E. & Lundstedt, T., 2007. Chemometrics in metabonomics. Journal of proteome research, 6(2), pp.469–479.

Urpi-Sarda, M. et al., 2015. Metabolomics for Biomarkers of Type 2 Diabetes Mellitus: Advances and Nutritional Intervention Trends. Current Cardiovascular Risk Reports, 9(3).

van den Berg, R.A. et al., 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC genomics, 7, p.142.

van der Voet, H., 1994. Comparing the predictive accuracy of models using a simple randomization test. Chemometrics and Intelligent Laboratory Systems, 25(2), pp.313–323.

van Doorn, M. et al., 2007. Evaluation of metabolite profiles as biomarkers for the pharmacological effects of thiazolidinediones in Type 2 diabetes mellitus patients and healthy volunteers. British journal of clinical pharmacology, 63(5), pp.562–574.

Van Weyenberg, S. et al., 2008. The effect of weight loss by energy restriction on metabolic profile and glucose tolerance in ponies. Journal of animal physiology and animal nutrition, 92(5), pp.538–545.

Vazquez-Fresno, R. et al., 2012. (1)H-NMR-based metabolomic analysis of the effect of moderate wine consumption on subjects with cardiovascular risk factors. Electrophoresis, 33(15), pp.2345–2354.

Vazquez-Fresno, R. et al., 2015. Metabolomic pattern analysis after mediterranean diet intervention in a nondiabetic population: a 1- and 3-year follow-up in the PREDIMED study. Journal of proteome research, 14(1), pp.531–540.

Vazquez-Fresno, R. et al., 2016. Clinical phenotype clustering in cardiovascular risk patients for the identification of responsive metabotypes after red wine polyphenol intake. The Journal of nutritional biochemistry, 28, pp.114–120.

Vu, T.N. & Laukens, K., 2013. Getting your peaks in line: a review of alignment methods for NMR spectral data. Metabolites, 3(2), pp.259–276.

Wang, Z. et al., 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature, 472(7341), pp.57–63.

Westerhuis, J.A. et al., 2008. Assessment of PLSDA cross validation. Metabolomics, 4(1), pp.81–89.

Woo, J. et al., 2001. The Mediterranean score of dietary habits in Chinese populations in four different geographical areas. European journal of clinical nutrition, 55(3), pp.215–220.

Woting, A. & Blaut, M., 2016. The intestinal microbiota in metabolic disease. Nutrients, 8(4).

Wurtz, P. et al., 2015. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. Circulation, 131(9), pp.774–785.

Wyss, M. & Kaddurah-Daouk, R., 2000. Creatine and creatinine metabolism. Physiological reviews, 80(3), pp.1107–1213.

Xia, J. & Wishart, D.S., 2016. Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. Current protocols in bioinformatics, 55, p.14.10.1-14.10.91.

Xia, J. et al., 2013. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics : Official journal of the Metabolomic Society, 9(2), pp.280–299.

Xia, J. et al., 2015. MetaboAnalyst 3.0--making metabolomics more meaningful. Nucleic acids research, 43(W1), pp.W251-7.

Yamamoto, H. et al., 2014. Statistical hypothesis testing of factor loading in principal component analysis and its application to metabolite set enrichment analysis. BMC Bioinformatics, 15(1), p.51.

Yang, J. et al., 2014. Modified Mediterranean diet score and cardiovascular risk in a North American working population. PloS one, 9(2), p.e87539.

Ye, X. et al., 2013. Mediterranean diet, healthy eating index 2005, and cognitive function in middle-aged and older Puerto Rican adults. Journal of the Academy of Nutrition and Dietetics, 113(2), pp.273–276.

Zeng, F.F. et al., 2014. Diet-quality scores and risk of hip fractures in elderly urban Chinese in Guangdong, China: a case-control study. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA, 25(8), pp.2131–2141.

Zheng, Y. et al., 2016. Weight-loss diets and 2-y changes in circulating amino acids in 2 randomized intervention trials. The American journal of clinical nutrition, 103(2), pp.505–511.

IX. ANEXOS

1 PUBLICACIONES DERIVADAS DE LA TESIS DOCTORAL

La presente Tesis Doctoral ha representado un proceso de investigación predoctoral a lo largo de 4 años y del cual se han concretado en la producción de 4 artículos científicos: 2 de ellos publicados y el resto en vías de publicación en revistas del área. De estos artículos, 3 se derivan directamente de las actividades enmarcadas en la Tesis Doctoral y uno más de una colaboración en el mismo grupo de investigación siguiendo la misma línea de la nutrimetabolómica de los alimentos. En las siguientes páginas se describen las contribuciones específicas del doctorando en cada uno de estos artículos, los cuales a su vez son anexados a continuación.

Publicación original 1:

Mireia Urpi-Sarda[#], Enrique Almanza Aguilera[#], Rafael Llorach, Rosa Vázquez-Fresno, Ramon Estruch, Dolores Corella, Jose Vicente Sorli, Francesc Carmona, Alex Sanchez-Pla, Jordi Salas-Salvadó, and Cristina Andres-Lacueva. Nontargeted metabolomics biomarkers and metabotypes of type 2 diabetes prevalence: a cross-sectional study in the Predimed trial. *Diabetologia* [*In process of publication*]. [#]*Equal contribution*.

Participación del doctorando:

• Procesado y tratamiento de los datos y aplicación de herramientas bioestadísticas para corroborar las hipótesis de partida

• Identificación y confirmación de los metabolitos discriminantes

• Interpretación biológica

• Elaboración del primer borrador de la publicación

• Seguimiento y trabajo del manuscrito hasta su envío a la revista.

Non-targeted metabolomics biomarkers and metabotypes of type 2 diabetes: a cross-sectional study in the Predimed trial

Mireia Urpi-Sarda,[1,2*,#] Enrique Almanza,[1,2#] Rafael Llorach,[1,2] Rosa Vázquez-Fresno,[1,3] Ramon Estruch,[4,5] Dolores Corella,[5,6] Jose Vicente Sorli,[5,6] Francesc Carmona,[7] Alex Sanchez-Pla,[7] Jordi Salas-Salvadó,[5,8] Cristina Andres-Lacueva,[1,2,*]

[1]Biomarkers and Nutrimetabolomic Laboratory, Department of Nutrition, Food Sciences and Gastronomy, XaRTA, INSA, Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain. [2]CIBER de Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Barcelona, Spain. [3]Department of Computing Science and Biological Sciences, University of Alberta, Edmonton, Canada. [4]Department of Internal Medicine, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi Sunyer (IDIBAPS), Barcelona, Spain. [5]CIBER Pathophysiology of Obesity and Nutrition (CIBERobn), Instituto de Salud Carlos III, Madrid, Spain. [6]Department of Preventive Medicine and Public Health, University of Valencia, Valencia, Spain. [7]Statistics Department, Biology Faculty, University of Barcelona, 08028, Spain. [8]Human Nutrition Unit, Biochemistry and Biotechnology Department. Hospital Universitari de Sant Joan de Reus, Institut d'Investigació Sanitària Pere Virgili, Universitat Rovira i Virgili, Reus, Spain.

* Corresponding authors: Mireia Urpi-Sarda, Av. Joan XXIII s/n, 08028 Barcelona, Spain; Phone: +34934034844, Fax: +34934035931, murpi@ub.edu; Cristina Andrés Lacueva, Av. Joan XXIII s/n, 08028 Barcelona, Spain; Phone: +34934034840, Fax: +34934035931, candres@ub.edu.

# Equal contribution

Short running title: Metabolomic biomarkers and metabotypes of T2D

Word count abstract: 250

Word count main text: 3421

Number of references: 43

Tables: 7 (3 in main manuscript, and 4 as supplemental information)

Figures: 5 (2 in main manuscript, and 3 as supplemental information)

ABSTRACT

Aims: To characterize the urinary metabolomic fingerprint and the biomarker associated with type 2 diabetes (T2D) and classify the population into metabotypes related to diabetes.

Methods: We conducted a metabolomics analysis using a ^{1}H-NMR-based non-targeted metabolomic approach to determine the urinary metabolomic fingerprint of T2D compared to non-T2D participants in the Predimed trial. The discriminant metabolite fingerprint was subjected to logistic regression analysis and receiver operating curve analyses to establish and assess a biomarker of T2D prevalence, respectively. Metabotypes associated with diabetes were identified using the *k*-means algorithm.

Results: Thirty-three metabolites were found to be significantly different ($p<0.05$) between T2D and non-T2D participants. The biomarker of T2D was composed of high levels of methylsuccinate, alanine, dimethylglycine and guanidoacetate, and decreased levels of glutamine, methylguanidine, 3-hydroxymandelate and hippurate. This biomarker had a 96.4% of area under the curve which was higher than individual metabolites and glucose. Amino acid and carbohydrate metabolism are the main metabolic alterations in T2D. We identified two metabotypes in the T2D participants. The metabotype that had higher levels of phenylalanine, phenylacetylglutamine, *p*-cresol and acetoacetate also had significantly higher levels of plasma glucosa.

Conclusions: The biomarker of T2D highlights the altered metabolic fingerprint associated mainly with amino acid, carbohydrate and microbiota metabolism. The metabotype with higher plasma glucose levels, had a metabolic profile that could be related to higher risk of long-term cardiovascular events. Metabolomics is a useful tool in the understanding of the metabolic complexity and interindividual variation in T2D for stratified and precision nutrition and medicine.

Trial registration www.controlled-trials.com ISRCTN35739639

Keywords diabetes type 2, metabolomics, NMR, Predimed, metabotypes, biomarkers

Abbreviations: type 2 diabetes (T2D); cardiovascular disease (CVD); the PREvention with MEDiterranean Diet (PREDIMED) study; free induction decay (FID); orthogonal signal correction (OSC); partial least squares discriminant analysis (PLS-DA); variable importance in projection (VIP); receiver operating characteristic (ROC); area under the curve from ROC curves (AROC); dimethylglycine (DMG); phenylacetylglutamine (PAG)

INTRODUCTION

Type 2 diabetes (T2D) encompasses individuals who have hyperglycaemia resulting from defects in insulin secretion, insulin action or both [1]. Moreover, hyperglycaemia and insulin resistance are risk factors for cardiovascular disease (CVD) [2]. Besides an understanding of the pathophysiology of T2D, identification of individuals at high risk, as well as knowledge of metabolic alterations produced in patients with T2D, are crucial for prevention and management strategies. In recent years, improvements in the development of biomarkers of T2D have been reached due to advances in the emergent "-omics" technologies, including metabolomics [3]. Successful applications of metabolomics in T2D research include the discovery of biomarkers for diagnoses, prognoses, altered metabolic pathways and drug mechanisms of action [4]. Currently, among the high-throughput analytical techniques, HPLC-MS and ^{1}H-NMR are the most widely employed in metabolomics for diabetes, mainly due to their advantages in the analysis and identification of a broad range of metabolites in biofluids [3, 4]. Furthermore, ^{1}H-NMR is frequently used in non-targeted metabolomic approaches in order to determine the metabolite profiling in studies comparing T2D and non-diabetic populations, as well as for the elucidation and confirmation of metabolic pathways altered as a consequence of T2D [3, 5, 6].

Several accurate prediction models have been used in recent years, including variables such as age, gender and lifestyle factors [7]. However, recent "-omics" technologies have the potential to be used as accurate analytical techniques aimed at discovering novel biomarkers that could be involved in prediction models for diabetes [4]. It is worth pointing out that these models have mainly been tested with plasma samples, and few of them with urine samples and in addition, comparing healthy versus T2D subjects. Thus, the use of these models in metabolomics is essential for identifying molecular signatures and phenotypic variation in order to improve the prediction of disease risk and better manage patient care and outcome [8]. Recently, the term "stratified medicine" has emerged as the concept that some groups of individuals should be treated differently from others due to intervariability [8]. This intervariability can be characterized by metabolomics through the study of metabolic phenotypes or metabotypes, which is a starting point for future stratified medicine programmes and lifestyle interventions [9]. The patterns of variation or metabotypes have previously been used in diabetes data sets to separate controls from diabetics [10], or healthy from diseased groups [11], in studying metabolomic differences among clinical phenotypes [12] and in developing strategies for the delivery of dietary advice [11, 13].

In the current study, we aimed to characterize the urinary metabolomic fingerprint and the biomarker associated with T2D prevalence in the PREDIMED population by using a ^{1}H-NMR-based non-

targeted metabolomic approach and classifying the population into metabolic phenotypes (metabotypes) related to diabetes.

METHODS

Study population The PREvention with MEDiterranean Diet (PREDIMED) study parallel-group, single-blind, multicenter, randomized, controlled 5-year clinical trial aimed to assess the effects of the Mediterranean diet on the primary prevention of CVD [14]. Full details of the design study and protocol have been published elsewhere [15] (www.predimed.es). Briefly, the participants were men (55–80 years of age) and women (60–80 years of age) with T2D and/or at least three of the following cardiovascular risk factors: hypertension, overweight (BMI≥25 kg/m^2), current smoker, LDL-cholesterol level ≥4.14 mmol/L, HDL-cholesterol level ≤1.03 mmol/L or a family history of premature CVD. The trial was registered at www.controlled-trials.com as ISRCTN35739639.

This substudy used data collected from 154 consecutive participants from two centres (Hospital Clinic of Barcelona and University of Valencia) in whom we determined the urinary metabolome using the H-NMR approach at baseline. Out of these 154 participants, 85 were T2D subjects and 69 were non-diabetic subjects. T2D subjects were diagnosed as previously reported [14, 16]. The institutional review boards of the centres agreed to the study protocols, and written informed consent was provided by the participants.

Urine collection and measurements Spot urine samples were collected at baseline and immediately stored at -80 °C until analysis. Trained personnel conducted anthropometric and blood pressure measurements. Validated questionnaires were employed to record physical activity [15], lifestyle, disease history and medication use.

Metabolomic analysis: ^{1}H-NMR sample preparation, data acquisition and processing Urine samples were thawed at 4 °C and gently vortexed before analysis. A procedure based on previously published methodology [17] was applied for the metabolomic analysis. Briefly, 300 microlitres of urine were diluted in 200 µl H_2O/D_2O (8:2) and mixed with an internal standard solution [(0.1%, chemical shift reference 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt, TSP), 2 mM of sodium azide (NaN_3) and 1.5 M KH_2PO_4, in 99% deuterated water (D_2O)]. The pH was set at 7.0 with a KOD solution. ^{1}H-NMR experiments were conducted using a 500 MHz Varian INOVA spectrometer, with presaturation of the water resonance using a NOESYPRESAT pulse sequence. The internal temperature was kept constant at 298 K during the acquisition. Spectra were acquired by collecting 128 scans into 32 K data points with a spectral width of 14 ppm, an acquisition time of 2 s, a relaxation delay of 5 s and a mixing time of 100 ms. For spectral processing, the free induction decay (FID) was multiplied by an exponential function corresponding to a line broadening of 0.3 Hz before

Fourier transformation. All spectra were phased, baseline corrected and referenced to TSP (δ 0.0) using TopSpin software (version 3.2, Bruker, BioSpin, Germany). The spectral data processed were intelligent, bucketed in domains of 0.005 ppm [17] and integrated using ACD/NMR Processor 12.0 software (Advanced Chemistry Development, Inc.). The spectral region between 4.75 and 5.00 ppm was excluded from the data set to avoid spectral interference from residual water.

Statistical analysis A data set containing integrals of NMR spectra was imported into MetaboAnalyst 3.0, a web-based platform for extensive analysis of metabolomic data [18]. Data were filtered using interquantile range and row-wise normalized by the sum of the intensities of spectra. The normalized data set was then imported into SIMCA-P+13.0 (Umetrics, Umeå, Sweden) before being log-transformed and range-scaled prior to performing a principal component analysis to explore the data distribution [17].

In order to reduce the variability not associated with the T2D classification we applied an orthogonal signal correction (OSC) to the data set and subsequently conducted a partial least squares discriminant analysis (OSC-PLS-DA) to determine the differences in metabolite profiling between the T2D and non-T2D groups. The quality and validation of the resultant OSC-PLS-DA model were assessed through the R^2Y (cum) and Q^2 (cum) parameters, as well as by performing a permutation test (n=200). Discriminant features between T2D and non-T2D individuals were taken from their variable importance in projection (VIP) values >1.0 [17].

Characteristics between participants were compared by applying a Student's t-test and X^2 test for continuous and qualitative variables, respectively. Differences in metabolites between the T2D and non-T2D groups were tested using a Student's t-test and a Benjamini–Hochberg procedure for adjusting *p* values. The significance level was set at $P<0.05$. All univariate analyses were performed using IBM SPSS 21® statistics software (IBM Corp., Armonk, NY, USA).

Metabolite identification The identification of metabolites was achieved using Chenomx NMR Suite 7.6 Profiler (Chenomx Inc, Edmonton, Canada). In addition, NMR spectra libraries were consulted in databases such as the Human Metabolome Database [19] and the Biological Magnetic Resonance Data Bank [20], together with existing literature using NMR-based metabolomics [4, 17].

Metabolic pathways analysis Identified metabolites were submitted to the Pathway Analysis and Network Analysis modules in Metaboanalyst 3.0 [18] and in MetaCore™ (GeneGo, division of Thomson Reuters, St. Joseph, MI, USA), respectively, to conduct analyses of metabolic pathways and biological interpretation of metabolites related to T2D.

Study of combined urinary biomarker model for T2D prevalence The results obtained from the OSC-PLS-DA analysis were subjected to a forward-conditional stepwise logistic regression analysis to design a biomarker model of diabetes prevalence. The prediction model was used on the training set

(2/3 of participants) and subsequently validated on the validation set (1/3 of participants). The quality of the models was evaluated by calculating the sensitivity, specificity and area under the curve from receiver operating characteristic (ROC) curves (AROC). Urinary glucose was not included in this analysis due to high levels of AROC. The optimal cut-off for calculating the sensitivity and specificity was determined as the minimum distance to the top-left corner [21]. The significance level was set at $P<0.05$. IBM SPSS 21® statistics software was used to perform logistic regression and ROC analyses.

Metabolic phenotypes by k-means algorithm Cluster analysis to identify metabolic phenotypes or metabotypes was performed by using the *k*-means cluster algorithm in MetaboAnalyst 3.0 [12, 22]. MetaboAnalyst generated two unique clusters in the diabetic and non-diabetic participants by taking as an input the identified metabolites from the OSC-PLS-DA analysis and by applying the *k*-means clustering algorithm, where k=2 had been chosen by visual inspection of the heatmap [12].

RESULTS

Participant characteristics The participants were 67 ± 6 years old and nearly one-third were men (Table 1). Fifty-five per cent of participants had T2D and 47% were obese. The participants were distributed according to the diagnosis of T2D as previously reported [14, 16]. Both groups (T2D and non-T2D) were well balanced in terms of demographic characteristics and other cardiovascular risk factors such as blood pressure, plasma lipids, and antihypertensive and hypolipidaemic medication (p>0.05). Otherwise, waist circumference, plasma glucose and antidiabetic agents were significantly higher in diabetic participants, as expected.

Profile of discriminant metabolites of T2D biomarkers using ^{1}H-NMR metabolomic approach An OSC-PLS-DA model was conducted in order to determine the profile of discriminant metabolites between T2D and non-T2D individiuals. The OSC-PLS-DA model resulted in one latent component with R^2Y (cum) and Q^2Y (cum) values of 0.815 and 0.737, respectively, indicating a good ability to classify individuals according to their T2D prevalence status. Additionally, a permutation test (n=200), with intercept R^2Y and Q^2Y values of 0.226 and -0.139, respectively, showed the validity of the model (Figure S1). The VIP$\geq$1.0 values and t test analyses identified 33 metabolites that were significantly different between T2D and non-T2D individuals (Table 2). Among these metabolites, 17 were identified as being significantly increased in T2D subjects compared to non-diabetic participants, while the other 16 metabolites decreased in the group of T2D subjects.

A comprehensive analysis of the metabolic pathway (P and impact values) showed that the carbohydrate and aminoacid metabolism pathways were the most altered among T2D participants (Figure S2, ESM Table 1). The metabolites involved in these pathways are shown in Figure S3 as up- and downregulated, and the pathways related to each metabolite are shown in Table S2.

Combined urinary biomarker model for T2D prevalence The multi-metabolite biomarker for better discrimination of T2D prevalence included higher levels of methylsuccinic acid, alanine, dimethylglycine and guanidoacetic acid, as well as lower levels of glutamine, methylguanidine, 3-hydroxymandelate and hippuric acid (Table S3). The specificity and sensitivity of the multi-metabolite biomarker were 87% and 96.4%, respectively, in the validation set, whereas the AROC was 96.4% (CI 92.0–100 %, p<0.001) in the same set. The specificity, sensitivity and AROC values of individual metabolites, as well as of glucose, were lower than those of the multi-metabolite biomarker (Figure 1 and Table 3).

Characterization of metabotypes *K*-means cluster analysis identified two metabotypes of diabetic participants and two metabotypes of non-diabetic participants (Table S4 and Figure 2) from data of the 33 identified metabolites. The four clusters were represented in the heatmap analysis (Figure 2) where samples were shown on the x-axis and metabolites were displayed on the y-axis. Most of the up- and

downregulated metabolites observed in these clusters were similar to those reported in Table 2 between T2D and non-T2D participants except for four metabolites: acetoacetate, p-cresol, phenylalanine and phenylacetylglutamine. The levels of these four metabolites were significantly higher in clusters 2 and 3 than in clusters 1 and 4 (Figure 2) ($p<0.05$). Thus, the two metabotypes of T2D (clusters 1 and 2) differed in these four metabolites. Cluster differences for the characteristics of the participants, the concentrations of biochemical parameters and the use of medication are presented in Table S4. The main difference was that cluster 2 had the highest plasma glucose levels followed by cluster 1, and these were significantly different from clusters 3 and 4 ($p<0.001$). As expected, the use of insulin and oral diabetic agents was significantly different between the T2D and non-diabetic participants ($p<0.001$), but there were no differences between T2D metabotypes ($p=0.2$).

DISCUSSION

In this study, we found significant differences in the profile of 33 urinary metabolites between T2D subjects and non-diabetic participants by using a ^{1}H-NMR-based non-targeted metabolomic approach. Specifically, a model of eight metabolites was the multi-metabolite biomarker that discriminated between T2D and non-T2D participants after a stepwise logistic regression analysis and ROC evaluation. To the best of our knowledge, this is the first study to use spot urine to determine altered pathways in T2D in a free-living population jointly with the determination of a biomarker of T2D prevalence that highlights the key implied metabolites. This metabolomic study in humans confirms disturbances in the amino acid metabolism, with some amino acids being used as substrates for gluconeogenesis. Also, the increased excretion of amino acids could indicate an increase in protein degradation [23]. This is observed in our study and corroborated in other metabolomic studies by the enhanced excretion of the glucogenic amino acids alanine [5, 6] and phenylalanine [6], or derived metabolites such as guanidinoacetate, and by the decreased excretion of glutamine [6] and histidine [5]. Previously, it has been found that levels of phenylalanine and glutamine were positively and inversely, respectively, associated with the risk of prediabetes and T2D [7, 24].

In addition, deregulations in BCAA metabolites (valine, leucine and isoleucine) were also associated with the risk of diabetes and insulin resistance [7, 24]. We observed these changes in urine through the increased excretion of metabolites from their degradation pathways such as 3-hydroxyisovalerate [25] and methylsuccinate, whose increase may reflect higher isoleucine catabolism [26].

This metabolomic study in humans shows an increase in the glycolysis and gluconeogenesis pathways in the liver associated with increased excretion of metabolites including lactate, glucose and pyruvate, as has also been observed in previous studies [4, 27]. Increased amounts of some carboxylic acids such as cis-aconitate, which is an intermediate in the tricarboxylic acid (TCA) cycle, dicarboxylic suberic acid and DTEA were observed in the urine of T2D participants. The increased excretion of *cis*-aconitate reflects the systemic stress caused by hyperglycaemia or by local effects on the tubular transport in the kidneys [5]. Metabolites related to methylamine metabolism such as DMG and trimethylamine N-oxide are systemic breakdown products of choline [5], and due to their osmoregulator properties, when found in high concentrations, they may be linked to a hyperosmotic effect of glucose or indicate renal papillary dysfunction [28]. In addition, they were related to risk of CVD [29].

Alterations of the glomerular filtration rate (GFR) in the diabetic, along with a possible decrease of muscle mass, could be related to a decreased creatinine excretion rate [30] and also to a decreased excretion of its metabolite methylguanidine [31]. In addition, in a recent work, a lower creatinine excretion rate was associated with all-cause mortality in diabetic subjects and nephropathy [30].

Diabetes and obesity are lifestyle-related diseases that could cause increased incidence of gut dysbiosis [32], which is directly related to the alteration of gut microbial-related metabolites [33]. Therefore, the results of our study have shown a reduction in the excretion of well-known microbial metabolites, such as hippurate, phenylacetylglutamine (PAG) and *p*-cresol [32], as well as trigonelline and 3-hydroxymandelate. While PAG and *p*-cresol are related to protein putrefaction, hippurate is a breakdown product of polyphenol and fibre metabolism [9], and 3-hydroxymandelate is a metabolite of tyrosine [19]. Previous studies observed that individuals with impaired glucose tolerance and T2D subjects had lower levels of hippuric acid and PAG [5, 34]. Therefore, this study supports previous findings that an imbalance of microbiota could be key in the pathogenesis of a diabetic state and that diets or healthy lifestyle patterns directed towards a change in order to improve the microbiota quality are essential for preventing advanced states of pathology.

In this study, we identified two distinct metabotypes in the T2D participants (clusters 1 and 2) and two in the non-diabetic participants (clusters 3 and 4) by using *k*-means cluster analysis on the basis of the identified metabolic profile. With regard to clustering of T2D participants, the two clusters were different in regard to four metabolite levels: phenylalanine, phenylacetylglutamine, *p*-cresol and acetoacetate. It is worth noting that those participants with higher levels of these four metabolites also had higher levels of plasma glucose but without differences in the use of diabetic medication or other characteristics. Thus, our hypothesis is that the diabetics in cluster 2 could have a greater lack of control over their diabetes, which in the long term could cause a higher number of complications including myocardial infarction, stroke, heart failure and kidney disease [35]. Certainly, phenylalanine has been described as a marker of higher risk of diabetes [7, 24], and furthermore, it has also been used jointly with tyrosine and isoleucine to predict long-term future cardiovascular events, increased disposition to atherosclerosis and could predict inducible myocardial ischaemia [36]. In addition, phenylalanine was identified as a biomarker associated with future cardiovascular events in meta-analyses [37]. PAG and *p*-cresol are metabolites of microbial origin [38]. PAG comes from the conversion of phenylalanine to phenylacetate by microbiota and subsequent conjugation with glutamine [39]. *p*-cresol is the most widely studied uremic retention solute and is formed from microbial metabolism of tyrosine [39]. PAG has been described as a strong and independent risk factor for mortality and cardiovascular diseases in patients with chronic kidney disease [38] and *p*-cresol has been described as a predictor of cardiovascular events independent of GFR in mild-to-moderate kidney disease patients [40]. The fourth metabolite that was different between T2D clusters was the ketone body acetoacetate (AA). It is generated from the ketogenic amino acid lysine or could derive from β-oxidation of fatty acids.

AA and 3-hydroxybutyrate (3HB) are in a proportion of 1:1 in the physiological state but 3HB increase its excretion in ketoacidosis [41]. Recent evidence has highlighted the association of elevated

levels of ketone bodies with hyperglycaemia and T2D [42]. It is interesting to note that the participants in clusters 1 and 2 had similar mean ratios of AA:3HB (1:2), while clusters 3 and 4 had mean ratios of 1:1 but without statistical significance. Meanwhile, there were statistically significant differences between ratios of T2D (1:2) and non-diabetic participants (1:1) (p=0.007). Hiperketonaemia and ketosis have been related to complications of the liver, brain and microvasculature, which increase the risk of morbidity and mortality [43]. Thus, those participants in cluster 2 that had increased levels of these four metabolites and also higher levels of plasma glucose could have a higher risk of CVD. It should be noted that the participants in cluster 3 also had similar levels of these four metabolites to those in cluster 2, but although they were non-diabetic participants, they had cardiovascular risk factors. Thus, based on our metabolomic results, this metabotype (cluster 3) could have a higher risk of CVD than the participants in cluster 4.

One of the limitations of this study is that the panel of biomarkers and the model with the biomarker imprinting of T2D was obtained from a high cardiovascular risk population and it should be validated and replicated in other populations. On the other hand, this panel should be tested in patients with different grades of T2D as prediabetes states in order to assay the limit value for prediction. Another limitation of this study is that there is no microbial composition for these participants to correlate with identified metabolites. One of the strengths of our study is the reproduction of real-life conditions of participants.

In conclusion, the results of this cross-sectional study using a non-targeted ^{1}H-NMR metabolomics approach showed a biomarker of T2D prevalence including eight metabolites. These metabolites belong to pathways related mainly to glucogenic and ketogenic amino acids, glycolysis and gluconeogenesis, carboxylic acid metabolism and alteration of gut microbiota metabolism. This is the first study to identify two metabotypes in T2D where those participants with higher levels of phenylalanine, PAG, p-cresol and AA, metabolites related to a higher risk of long-term cardiovascular events, also had higher plasma glucose levels.

This study reinforces the use of metabolomics to discover and evaluate the main metabolic pathways altered in T2D and the metabotypes in individuals. Hereafter, it would be useful to carry out further investigation of its diagnosis and treatment, which would support the development of stratified and precise medicine.

ACKNOWLEDGMENTS

M.U.-S. would like to thank the "Ramón y Cajal" program from MINECO and the Fondo Social Europeo. EAA would like to thank CONACYT (Mexico) for the Ph.D. fellowship.

Coflict of interest: No potential conflicts of interest relevant to this article were reported.

Funding This research was supported by Spanish National Grants from the Ministry of Economy and Competitiveness (MINECO), and co-funded by FEDER (Fondo Europeo de Desarrollo Regional): AGL2009-13906-C02-01, RYC-2011-09677, Associació Catalana de Diabetis (ACD2015), CIBERFES, CIBER 06/03 and CNIC-06 as well as PI13/01172 Project (Plan Nacional de I+D+i 2013–2016), and PI11/02505 by ISCII-Subdirección General de Evaluación y Fomento de la Investigación. We also thank the EU Joint Programming Initiative A Healthy Diet for a Healthy Life on Biomarkers BioNHFOODBALL (PCIN-2014-133-MINECO-Spain) and the award of 2014SGR1566 from the Generalitat de Catalunya's Agency AGAUR.

Author contribution M.U.-S., R.L., R.E., D.C., J.V.S., J.S.-S and C.A.-L. conceived and designed the study; M.U.-S., E.A., R.L. and R.V.-F. performed the analyses; M.U.-S., E.A., R.L., R.V.-F., F.C., C.A.-L. and A.S.-P. analysed the results; M.U.-S. and E.A. drafted the article. All the authors conducted a critical revision of the article for important intellectual content. M.U.-S. and C.A.-L. are the guarantors of this work and, as such, had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

[1] American Diabetes Association (2017) Classification and Diagnosis of Diabetes. Diabetes Care 40: S11-S24

[2] Paneni F, Beckman JA, Creager MA, Cosentino F (2013) Diabetes and vascular disease: pathophysiology, clinical consequences, and medical therapy: part I. Eur Heart J 34: 2436-2443

[3] Adamski J (2016) Key elements of metabolomics in the study of biomarkers of diabetes. Diabetologia 59: 2497-2502

[4] Urpi-Sarda M, Almanza-Aguilera E, Tulipani S, Tinahones F, Salas-Salvado J, Andres-Lacueva C (2015) Metabolomics for biomarkers of Type 2 Diabetes Mellitus: advances and nutritional intervention trends. Curr Cardiovasc Risk Rep 9:12

[5] Salek RM, Maguire ML, Bentley E, et al. (2007) A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. Physiological genomics 29: 99-108

[6] van Doorn M, Vogels J, Tas A, et al. (2007) Evaluation of metabolite profiles as biomarkers for the pharmacological effects of thiazolidinediones in Type 2 diabetes mellitus patients and healthy volunteers. British journal of clinical pharmacology 63: 562-574

[7] Herder C, Kowall B, Tabak AG, Rathmann W (2014) The potential of novel biomarkers to improve risk prediction of type 2 diabetes. Diabetologia 57: 16-29

[8] Pearson ER (2016) Personalized medicine in diabetes: the role of 'omics' and biomarkers. Diabet Med 33: 712-717

[9] Heinzmann SS, Merrifield CA, Rezzi S, et al. (2012) Stability and robustness of human metabolic phenotypes in response to sequential food challenges. J Proteome Res 11: 643-655

[10] Cuperlovic-Culf M, Belacel N, Culf A, et al. (2009) NMR metabolic analysis of samples using fuzzy k-means clustering. Magn Reson Chem 47: S96-104

[11] Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC (2012) Metabolic phenotyping in clinical and surgical environments. Nature 491: 384-392

[12] Vazquez-Fresno R, Llorach R, Perera A, et al. (2016) Clinical phenotype clustering in cardiovascular risk patients for the identification of responsive metabotypes after red wine polyphenol intake. J Nutr Biochem 28: 114-120

[13] O'Donovan C, Walsh M, Nugent A, et al. (2015) Use of metabotyping for the delivery of personalised nutrition. Mol Nutr Food Res 59: 377-385

[14] Estruch R, Ros E, Salas-Salvado J, et al. (2013) Primary prevention of cardiovascular disease with a Mediterranean diet. The New England journal of medicine 368: 1279-1290

[15] Estruch R, Martinez-Gonzalez MA, Corella D, et al. (2006) Effects of a Mediterranean-style diet on cardiovascular risk factors: a randomized trial. Ann Intern Med 145: 1-11

[16] Salas-Salvado J, Bullo M, Estruch R, et al. (2014) Prevention of diabetes with Mediterranean diets: a subgroup analysis of a randomized trial. Ann Intern Med 160: 1-10

[17] Vazquez-Fresno R, Llorach R, Urpi-Sarda M, et al. (2014) Metabolomic Pattern Analysis after Mediterranean Diet Intervention in a Nondiabetic Population: A 1- and 3-Year Follow-up in the PREDIMED Study. J Proteome Res

[18] Xia J, Mandal R, Sinelnikov IV, Broadhurst D, Wishart DS (2012) MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. Nucleic acids research 40: W127-133

[19] Wishart DS, Jewison T, Guo AC, et al. (2013) HMDB 3.0--The Human Metabolome Database in 2013. Nucleic acids research 41: D801-807

[20] Ulrich EL, Akutsu H, Doreleijers JF, et al. (2008) BioMagResBank. Nucleic acids research 36: D402-408

[21] Garcia-Aloy M, Llorach R, Urpi-Sarda M, et al. (2014) Novel multimetabolite prediction of walnut consumption by a urinary biomarker model in a free-living population: the PREDIMED study. J Proteome Res 13: 3476-3483

[22] Gawlik A, Shmoish M, Hartmann MF, Malecka-Tendera E, Wudy SA, Hochberg Z (2016) Steroid Metabolomic Disease Signature of Nonsyndromic Childhood Obesity. J Clin Endocrinol Metab 101: 4329-4337

[23] Connor SC, Hansen MK, Corner A, Smith RF, Ryan TE (2010) Integration of metabolomics and transcriptomics data to aid biomarker discovery in type 2 diabetes. Molecular bioSystems 6: 909-921

[24] Guasch-Ferre M, Hruby A, Toledo E, et al. (2016) Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. Diabetes Care 39: 833-846

[25] Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014) Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic acids research 42: D199-205

[26] Nowaczyk MJ, Lehotay DC, Platt BA, et al. (1998) Ethylmalonic and methylsuccinic aciduria in ethylmalonic encephalopathy arise from abnormal isoleucine metabolism. Metabolism 47: 836-839

[27] Zhang S, Nagana Gowda GA, Asiago V, Shanaiah N, Barbas C, Raftery D (2008) Correlative and quantitative 1H NMR-based metabolomics reveals specific metabolic pathway disturbances in diabetic rats. Analytical biochemistry 383: 76-84

[28] Guan M, Xie L, Diao C, et al. (2013) Systemic perturbations of key metabolites in diabetic rats during the evolution of diabetes studied by urine metabonomics. PLoS One 8: e60409

[29] Wang Z, Klipfell E, Bennett BJ, et al. (2011) Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 472: 57-63

[30] Sinkeler SJ, Kwakernaak AJ, Bakker SJ, et al. (2013) Creatinine excretion rate and mortality in type 2 diabetes and nephropathy. Diabetes care 36: 1489-1494

[31] Ienaga K, Hum Park C, Yokozawa T (2014) Daily hydroxyl radical scavenging capacity of mammals. Drug discoveries & therapeutics 8: 71-75

[32] Nicholson JK, Holmes E, Kinross J, et al. (2012) Host-gut microbiota metabolic interactions. Science 336: 1262-1267

[33] Herrema H, RG IJ, Nieuwdorp M (2017) Emerging role of intestinal microbiota and microbial metabolites in metabolic control. Diabetologia 60: 613-617

[34] Zhao X, Fritsche J, Wang J, et al. (2010) Metabonomic fingerprints of fasting plasma and spot urine reveal human pre-diabetic metabolic traits. Metabolomics : Official journal of the Metabolomic Society 6: 362-374

[35] Norhammar A, Bodegard J, Nystrom T, Thuresson M, Eriksson JW, Nathanson D (2016) Incidence, prevalence and mortality of type 2 diabetes requiring glucose-lowering treatment, and associated risks of cardiovascular complications: a nationwide study in Sweden, 2006-2013. Diabetologia 59: 1692-1701

[36] Magnusson M, Lewis GD, Ericson U, et al. (2013) A diabetes-predictive amino acid score and future cardiovascular disease. Eur Heart J 34: 1982-1989

[37] Wurtz P, Havulinna AS, Soininen P, et al. (2015) Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. Circulation 131: 774-785

[38] Poesen R, Claes K, Evenepoel P, et al. (2016) Microbiota-Derived Phenylacetylglutamine Associates with Overall Mortality and Cardiovascular Disease in Patients with CKD. J Am Soc Nephrol 27: 3479-3487

[39] Swann JR, Spagou K, Lewis M, et al. (2013) Microbial-mammalian cometabolites dominate the age-associated urinary metabolic phenotype in Taiwanese and American populations. J Proteome Res 12: 3166-3180

[40] Meijers BK, Claes K, Bammens B, et al. (2010) p-Cresol and cardiovascular risk in mild-to-moderate kidney disease. Clin J Am Soc Nephrol 5: 1182-1189

[41] Sheikh-Ali M, Karon BS, Basu A, et al. (2008) Can serum beta-hydroxybutyrate be used to diagnose diabetic ketoacidosis? Diabetes care 31: 643-647

[42] Mahendran Y, Vangipurapu J, Cederberg H, et al. (2013) Association of ketone body levels with hyperglycemia and type 2 diabetes in 9,398 Finnish men. Diabetes 62: 3618-3626

[43] Kanikarla-Marie P, Jain SK (2016) Hyperketonemia and ketosis increase the risk of complications in type 1 diabetes. Free Radic Biol Med 95: 268-277

Table 1 Characteristics of the 154 participants and according to diabetic prevalence status.

	Total (n=154)	T2D (n=85)	non-T2D (n=69)	*P*-value
Age (years)	67 ± 5.7	67 ± 6.0	67 ± 5.5	0.67
Men, n (%)	48 (31)	30 (35)	18 (22)	0.15
Current smokers, n (%)	20 (13)	11 (13)	9 (13)	0.58
BMI, kg/m^2	30.18 ± 4.29	30.40 ± 4.60	30.0 ± 3.92	0.50
Obesity (BMI≥25 kg/m^2), n (%)	71 (47)	37 (44)	34 (49)	0.29
Weight, kg	76.0 ± 12.67	77.71 ± 13.27	73.90 ± 11.65	0.063
Waist circumference, cm	101.31 ± 11.38	103.71 ± 11.34	98.35 ± 10.81	0.003
Physical activity, MET·min/d	261.27 ± 268.32	241.22 ± 285.51	285.96 ± 245.31	0.30
SBP, mmHg	144.01 ± 19.72	140.70 ± 17.97	148.0 ± 21.20	0.08
DBP, mmHg	81.94 ± 10.22	81.06 ± 10.46	83.0 ± 9.95	0.38
Mediterranean dietary score	8.5 ± 1.8	8.383 ± 1.8	8.7 ± 1.8	0.19
Plasma biomarkers				
Glucose, mg/dl	128.2 ± 48.5	159.3 ± 45.3	89.9 ± 7.2	<0.001
HDL cholesterol, mg/dl	55.1 ± 14.4	51.5 ± 14.4	58.4 ± 13.9	0.09
LDL cholesterol, mg/dl	131.7 ± 28.8	123.3 ± 29.7	139.0 ± 26.5	0.06
Cholesterol total, mg/dl	213. 6 ± 14.4	206.9 ± 35.0	218.8 ± 32.9	0.21
Triglycerides, mg/dl	150.0 ± 113.6	176.6 ± 153.5	125.8 ± 49.3	0.13
Medication, n ± (%)				
Antihypertensive agents	65 (42)	32 (38)	33 (48)	0.13
Hypolipidaemic agents	85 (55)	42 (49)	43 (62)	0.07
Oral diabetic agents	65 (42)	65 (76)	0	<0.01
Insulin	18 (12)	18 (21)	0	<0.01

Data are shown as mean ± SD for continuous variables or % for categorical variables, as appropriate.

Abbreviations: BMI, body mass index; DBP, diastolic blood pressure; HDL, high-density lipoprotein; LDL, low-density lipoprotein; MEDAS, Mediterranean diet adherence screener; MET, metabolic equivalent; SBP, systolic blood pressure.

Table 2 Urinary metabolites identified by ^{1}H-NMR among T2D and non-T2D participants of the Predimed study

Metabolite	ppm (multiplicity† and J coupling in Hz) †	Excretion T2D vs non-T2D	P-value¶	VIP
Higher in T2D participants				
3-Hydroxyhippuric acid	7.12 (d, J=8.92), 7.29 (m), 7.37 (m), 7.41 (m)	↑	0.011	1.11
3-Hydroxybutyrate	1.19 (d, J= 6.24)	↑	0.047	0.94
3-Hydroxyisovaleric acid	1.27 (s)	↑	0.004	1.26
4-Deoxythreonic acid	1.22 (d, J=6.47)	↑	0.039	1.00
Alanine	1.48 (d, J=7.27)	↑	0.013	1.45
cis-Aconitate	3.13 (s), 5.75 (s)	↑	0.004	1.44
Citrate	2.54 (d, J=15.16), 2.68 (d, J=14.67)	↑	0.047	1.28
Dimethylglicine	2.93 (s)	↑	0.014	1.46
Glucose	3.25 (m), 3.45 (m), 3.50 (m), 3.72 (m), 4.64 (d, J=7.93), 5.25 (d, J=3.71)	↑	<0.001	3.58
Guanidoacetic acid	3.79 (s)	↑	<0.001	2.82
Isopropyl alcohol	1.16 (d, J=6.24)	↑	0.002	1.29
Lactic acid	1.33 (d, J=6.90), 4.11 (q)	↑	<0.001	2.03
Methylsuccinic acid	1.07 (d, J=7.01)	↑	<0.001	1.77
Phenylalanine	7.32 (d, J=7.15), 7.36 (m), 7.43 (t)	↑	0.009	1.20
Pyruvic acid	2.38 (s)	↑	<0.001	1.61
Suberic acid	1.29 (m), 1.59 (t)	↑	<0.001	1.77
Trimethylamine N-oxide /betaine‡	3.27 (s), 3.90 (s)	↑	<0.001	2.66

Table 2. Continued

Lower in T2D participants				
3-Hydroxymandelate	6.87 (d, J=8.27), 6.93 (t), 6.99 (d, J=7.64)	↓	0.003	1.68
Acetoacetate	2.28 (s)	↓	0.007	1.12
Acetylcartinine	2.15 (s), 2.19 (s)	↓	0.015	1.26
Creatinine	4.06 (s)	↓	<0.001	2.36
Dimethylsulphone	3.16 (s)	↓	<0.001	1.87
Glutamine	2.14 (m), 2.47 (m)	↓	0.001	1.28
Hippurate	3.98 (d, J=5.82) 7.55 (t) 7.63 (tt) 7.83 (dd) 8.54 (bb)	↓	<0.001	2.80
Histidine	7.08 (s)	↓	0.011	1.35
Methylguanidine	2.82 (s)	↓	0.001	1.68
N-acetylglutamine	2.09 (m), 4.16 (m), 2.25 (t), 7.36 (m), 7.98 (m)	↓	0.004	1.09
N-acetylglycoproteins	2.04 (s)	↓	0.017	1.09
p-Cresol	2.34 (s), 7.22 (d, J=8.37), 7.27 (d. J=8.90)	↓	0.009	1.70
Phenylacetylglutamine	2.27 (t), 4.17 (m), 7.42 (m), 7.26 (m), 7.98 (bb)	↓	0.026	1.14
Scyllo-inositol	3.36 (s)	↓	<0.001	2.13
Trigonelline	4.43 (s) 8.08 (m) 8.84 (m) 9.12 (s)	↓	0.002	1.44
Trimethylamine	2.86 (s)	↓	0.005	1.30

*Metabolites were identified by VIP≥1.0 in OSC-PLS-DA model. ¶adjusted value using Benjamini–Hochberg procedure. † Multiplicities are: s=singlet, d=doublet, t=triplet, q=quadruplet, m=multiplet, bb=broad band. ‡Overlapping peaks.

Table 3 Receiver operating characteristic (ROC) curve parameters for prediction model and for individual metabolites

Training set	AUC (95% CI)	*P*-value	Specificity (%)	Sensitivity (%)
Combined model	96.1 (92.3-100)	1.11×10^{-15}	96.0	93.0
Glucose	83.5 (76.0-91.1)	5.50×10^{-9}	91.3	66.7
Validation set				
Combined model	96.4 (92.0-100)	1.11×10^{-8}	87.0	96.4
Methylsuccinic acid	69.4 (54.8-84.0)	0.018	74.0	57.1
Alanine	66.3 (51.5-81.1)	0.047	78.3	53.6
Glutamine	37.6 (21.8-53.4)	0.130	34.8	67.8
Methylguanidine	31.8 (17.1-46.5)	0.027	34.8	50.0
Dimethylglycine	64.0 (48.7-79.3)	0.088	65.2	64.3
Guanidoacetic acid	81.0 (68.8-93.4)	<0.001	74.0	82.1
3-hydroxymandelate	40.2 (24.3-56.2)	0.233	30.0	64.3
Hippuric acid	15.0 (3.4-36.4)	<0.001	8.69	75.0
Glucose	89.8 (80.2-99.3)	1.26×10^{-6}	82.6	93.0

FIGURE LEGENDS

Fig. 1 ROC curves comparing the biomarker and individual metabolites of prevalence of T2D.

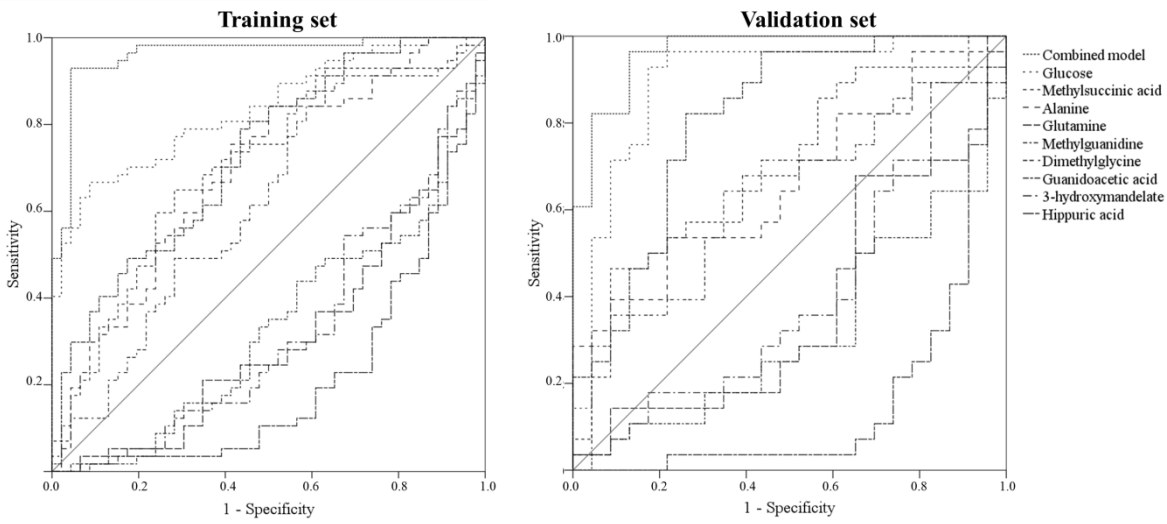
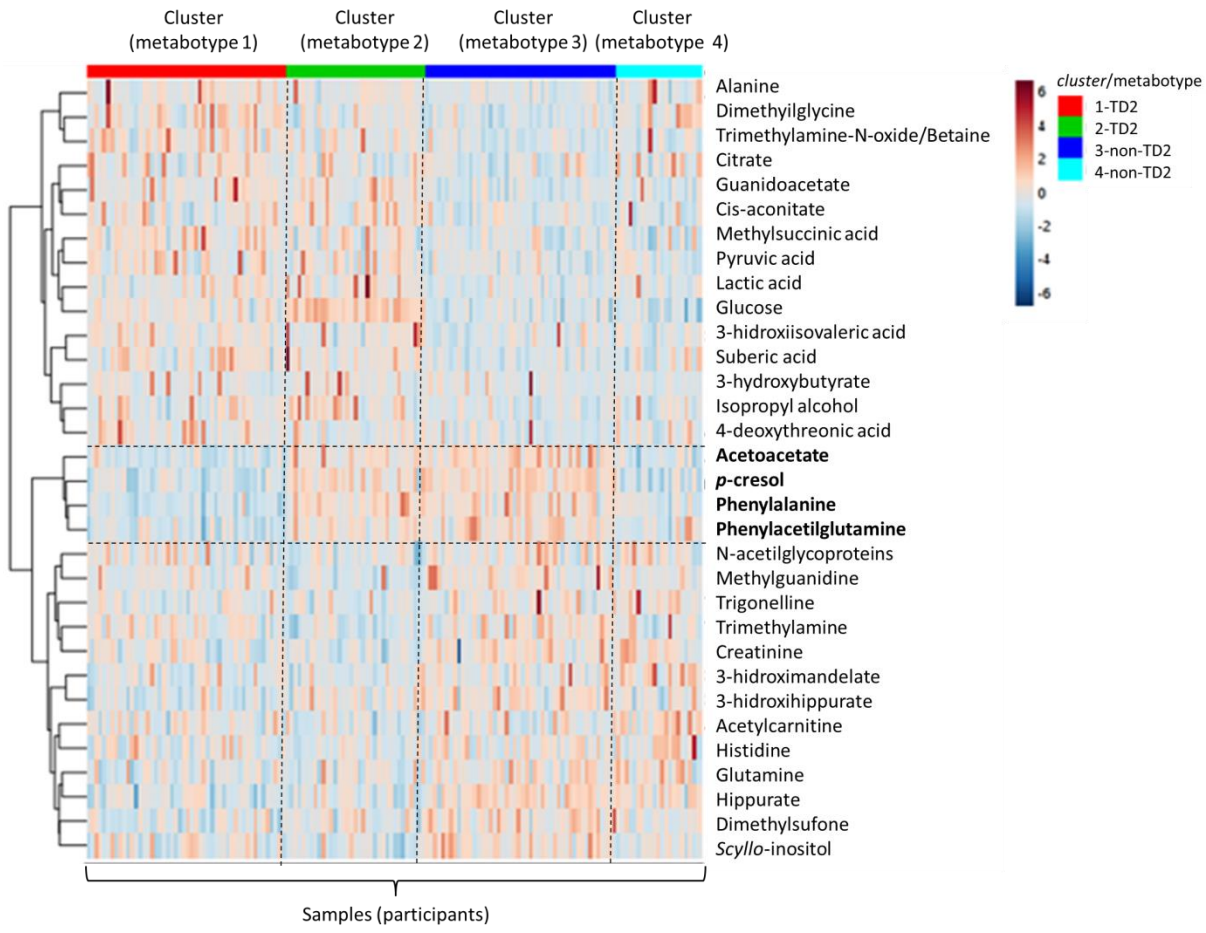


Fig. 2 Heatmap of metabotypes of participants in the Predimed study.



Note: T2D, diabetic participant; non-T2D, non-diabetic participant

Table S1 List of metabolic pathways potentially altered in individuals with T2D¶.

Pathway Name	Total	Expected	Hits	Raw p	-log(p)	Holm adjust	FDR	Impact
Citrate cycle (TCA cycle)	20	0.207	3	0.001	6.902	0.080	0.034	0.208
Phenylalanine metabolism	45	0.467	4	0.001	6.894	0.080	0.034	0.150
Glycine, serine and threonine metabolism	48	0.498	4	0.001	6.649	0.101	0.034	0.062
Alanine, aspartate and glutamate metabolism	24	0.249	3	0.001	6.356	0.133	0.034	0.264
Glycolysis or Gluconeogenesis	31	0.321	3	0.003	5.606	0.279	0.058	0.095
Aminoacyl-tRNA biosynthesis	75	0.778	4	0.006	5.006	0.501	0.073	0.0
Nitrogen metabolism	39	0.405	3	0.007	4.951	0.523	0.073	0.0
Arginine and proline metabolism	77	0.799	4	0.007	4.913	0.536	0.073	0.030
Glyoxylate and dicarboxylate metabolism	50	0.519	3	0.014	4.264	1.0	0.125	0.006
Taurine and hypotaurine metabolism	20	0.207	2	0.017	4.042	1.0	0.140	0.053
Pentose phosphate pathway	32	0.332	2	0.042	3.158	1.0	0.283	0.0
Pyruvate metabolism	32	0.332	2	0.042	3.158	1.0	0.283	0.320
Propanoate metabolism	35	0.363	2	0.050	2.995	1.0	0.307	0.0
Nicotinate and nicotinamide metabolism	44	0.457	2	0.075	2.588	1.0	0.429	0.0
D-glutamine and D-glutamate metabolism	11	0.114	1	0.108	2.219	1.0	0.567	0.026
Cysteine and methionine metabolism	56	0.581	2	0.113	2.175	1.0	0.567	0.016
Selenoamino acid metabolism	22	0.228	1	0.206	1.579	1.0	0.969	0.0
Pantothenate and CoA biosynthesis	27	0.280	1	0.246	1.399	1.0	0.970	0.0
Phenylalanine, tyrosine and tryptophan biosynthesis	27	0.280	1	0.246	1.399	1.0	0.970	0.000

Table S1. Continued

Valine, leucine and isoleucine biosynthesis	27	0.280	1	0.246	1.399	1.0	0.970	0.021
beta-Alanine metabolism	28	0.290	1	0.254	1.367	1.0	0.970	0.0
Vitamin B6 metabolism	32	0.332	1	0.285	1.253	1.0	1.0	0.019
Terpenoid backbone biosynthesis	33	0.342	1	0.293	1.227	1.0	1.0	0.0
Methane metabolism	34	0.353	1	0.300	1.202	1.0	1.0	0.0
Butanoate metabolism	40	0.415	1	0.343	1.068	1.0	1.0	0.085
Galactose metabolism	41	0.425	1	0.350	1.048	1.0	1.0	0.002
Histidine metabolism	44	0.457	1	0.370	0.991	1.0	1.0	0.139
Ascorbate and aldarate metabolism	45	0.467	1	0.377	0.973	1.0	1.0	0.016
Starch and sucrose metabolism	50	0.519	1	0.409	0.891	1.0	1.0	0.017
Pentose and glucuronate interconversions	53	0.550	1	0.428	0.847	1.0	1.0	0.0
Pyrimidine metabolism	60	0.623	1	0.469	0.755	1.0	1.0	0.0
Tyrosine metabolism	76	0.789	1	0.553	0.591	1.0	1.0	0.0
Amino sugar and nucleotide sugar metabolism	88	0.914	1	0.607	0.497	1.0	1.0	0.0
Purine metabolism	92	0.955	1	0.624	0.470	1.0	1.0	0.0

¶From pathway enrichment analysis in MetaboAnalyst. Significance (p-values) of pathways according to matches of the number of metabolites found in data set and those which compose the metabolic pathway.

Table S2. Significant ^{1}H-NMR-derived urinary metabolites among diabetic and non-diabetic participants and the metabolic pathways implicated.

Metabolite	Pathway *
Higher in diabetic participants	
3-Hydroxyisovaleric acid	Leucine, isoleucine and valine metabolism
4-Deoxythreonic acid	Carboxylic acid
3-Hydroxybutyrate	Synthesis and degradation of ketone bodies; butanoate metabolism
Alanine	Alanine, cysteine and methionine metabolism; Glycine, serine, cysteine and threonine metabolism; Histidine, glutamate, glutamine and proline metabolism
cis-Aconitate	Tricarbonic acid cycle
Citrate	Metabolite in the tricarboxylic acid cycle
Dimethylglycine	Glycine, serine, cysteine and threonine metabolism
Glucose	Glycolysis and gluconeogenesis; Fructose and galactose metabolism; Regulation of lipid metabolism and insulin regulation of fatty acid metabolism
Guanidoacetic acid	Urea cycle; Amino acid metabolism
Isopropyl alcohol	Pyruvate metabolism; Glycolysis and gluconeogenesis
Lactate	Glycolysis and gluconeogenesis
Methylsuccinic acid	Isoleucine metabolism
Phenylalanine	Phenylalanine metabolism; Tyrosine metabolism
Pyruvate	Alanine, cysteine and methionine metabolism; Glycine, serine, cysteine and threonine metabolism; Glycolysis and gluconeogenesis; Histidine, glutamate, glutamine and proline metabolism; Pyruvate metabolism; Tricarbonic acid cycle; Regulation of lipid metabolism and insulin regulation of fatty acids metabolism
Suberic acid	Dicarboxylic acid
Trimethylamine N-oxide /betaine‡	Choline metabolism

Table S2. Continued

Lower in diabetic participants	
3-hydroxyhippuric acid	Microbial aromatic acid metabolite; acyl glycine metabolite
3-Hydroxymandelate	Catecholamine metabolism; Breakdown product of m-octopamine, m-synephrine (phenylephrine) and m-tyrosine
Acetoacetate	Synthesis and degradation of ketone bodies; Leucine, isoleucine and valine metabolism; Butanoate and propanoate metabolism; Tyrosine metabolism
Acetylcarnitine	Oxidation of fatty acids; insulin resistance pathway
Creatinine	Arginine and proline metabolism
Dimethylsulfone	Intestinal bacterial metabolism; human endogenous methanethiol metabolism; sulfur metabolism
Glutamine	Histidine, glutamate, glutamine and proline metabolism; Nitrogen metabolism
Hippurate	Microbial metabolism; Phenylalanine metabolism
Histidine	Beta-alanine metabolism; Histamine metabolism; Histidine, glutamate, glutamine and proline metabolism
Methylguanidine	Protein catabolism
N-acetylglutamine	Amino sugar and nucleotide sugar metabolism
N-acetylglycoproteins	Protein catabolic products
p-Cresol ¶	Microbial metabolism; Protein breakdown product
Phenylacetylglutamine	Phenylalanine metabolism
Scyllo-inositol	Inositol phosphate metabolism
Trigonelline ¶	Nicotinate and nicotinamide metabolism
Trimethylamine	Methane metabolism, Microbial metabolism, Carbon metabolism

*Pathway data were obtained from Pathway Analysis in MetaCore^{TM}, and from KEGG and/or HMDB databases.

Table S3 Metabolites selected by stepwise logistic regression model for T2D discrimination.

Metabolite	Coefficient	SE	P-value
Methylsuccinic acid	9.253	2.852	0.001
Alanine	2.868	.998	0.004
Glutamine	-5.287	2.783	0.057
Methylguanidine	-5.610	2.159	0.009
Dimethylglycine	2.469	1.247	0.048
Guanidoacetic acid	3.317	1.299	0.011
3-hydroxymandelate	-5.092	1.888	0.007
Hippuric acid	-10.424	2.815	0.000

Table S4 Characteristics of the participants according to the identified metabotype

	Cluster 1 – T2D (n=50)	Cluster 2 – T2D (n=35)	Cluster 3 – non-T2D (n=48)	Cluster 4 – non-T2D (n=21)	P
Age (years)	67 ± 5	68 ± 6	67 ± 5	66 ± 6	0.42
Men, n (%)	20 (40)	10 (29)	12 (25)	6 (29)	0.42
Current smokers, n (%)	6 (12)	5 (14)	6 (13)	3 (14)	0.87
BMI, kg/m^2	30.8 ± 3.6	29.9 ± 5.8	30.0 ± 3.9	29.7 ± 4.1	0.70
Obesity (BMI≥25 kg/m^2), n (%)	25 (50)	12 (34)	23 (48)	11 (52)	0.45
Weight, kg	79.4 ± 11.7	75.3 ± 15.1	74.2 ± 11.7	73.3 ± 11.8	0.13
Waist circumference, cm	104.9 ± 9.8 [a]	102.0 ± 13.2 [ab]	98.0 ± 10.4 [b]	99.2 ± 11.9 [ab]	0.018
Physical activity, MET·min/d	258.3 ± 323.4	216.8 ± 222.8	305.0 ± 257.5	242.5 ± 214.3	0.51
SBP, mmHg	138.6 ± 16.5	143.4 ± 19.8	143.9 ± 20.7	156.5 ± 20.3	0.06
DBP, mmHg	80.9 ± 10.0	81.3 ± 11.3	82.4 ± 10.3	84.2 ± 9.5	0.79
Mediterranean diet score	8.2 ± 1.7	8.5 ± 1.9	9.0 ± 1.6	8.1 ± 1.9	0.08
Plasma biomarkers					
Glucose, mg/dl	147.9 ± 34.8 [b]	175.5 ± 53.5 [a]	89.9 ± 7.6 [c]	90.0 ± 6.4 [c]	<0.001
Cholesterol total, mg/dl	197.1 ± 29.7	219.7±38.7	219.7±28.5	216.6 ± 44.7	0.25
Triglycerides, mg/dl	150.2 ± 53.3	138.6± 56.6	127.0 ± 53.3	121.2 ± 34.2	0.59
Medication, n ± (%)					
Antihypertensive agents	22 (44)	10 (29)	24(50)	9 (43)	0.27
Hypolipidaemic agents	25 (50)	17 (49)	32 (67)	11 (52)	0.29
Oral diabetic agents	40 (80) [a]	25 (71) [a]	0 (0) [b]	0 (0) [b]	<0.001
Insulin	8 (16) [a]	10 (29) [a]	0 (0) [b]	0 (0) [b]	<0.001

Data are shown as mean ± SD for continuous variables or % for categorical variables, as appropriate.

Abbreviations: BMI, body mass index; DBP, diastolic blood pressure; HDL, high-density lipoprotein; LDL, low-density lipoprotein; MEDAS, Mediterranean diet adherence screener; MET, metabolic equivalent; SBP, systolic blood pressure

Figure S1 Permutation test (n=200) plot of OSC-PLS-DA model for discriminant analysis between T2D and non-T2D individuals.

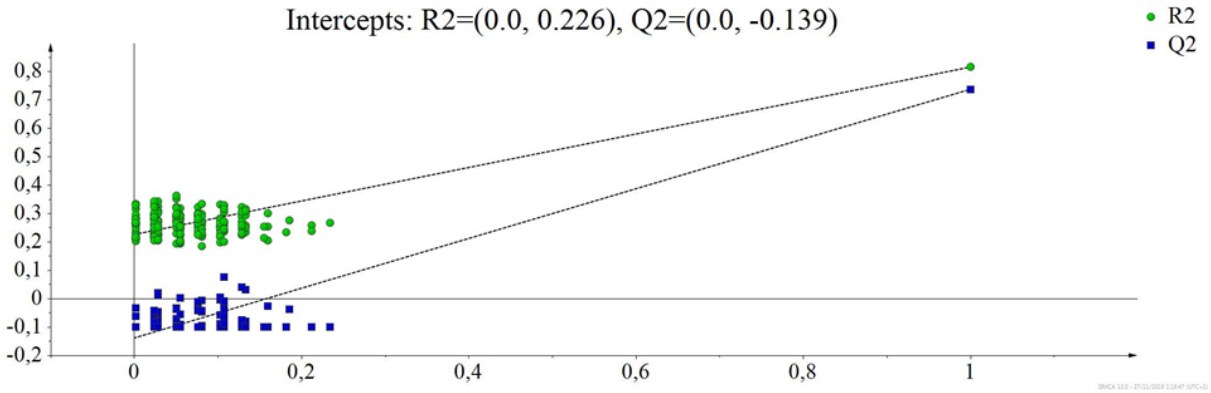


Figure S2 Metabolome view map of significant metabolic pathways of T2D prevalence in Predimed participants. The x-axis represents pathway enrichment analysis, and the y-axis represents pathway impact values (values shown in Table S1). Circled colours (varying from yellow to red) indicate the different levels of significance.

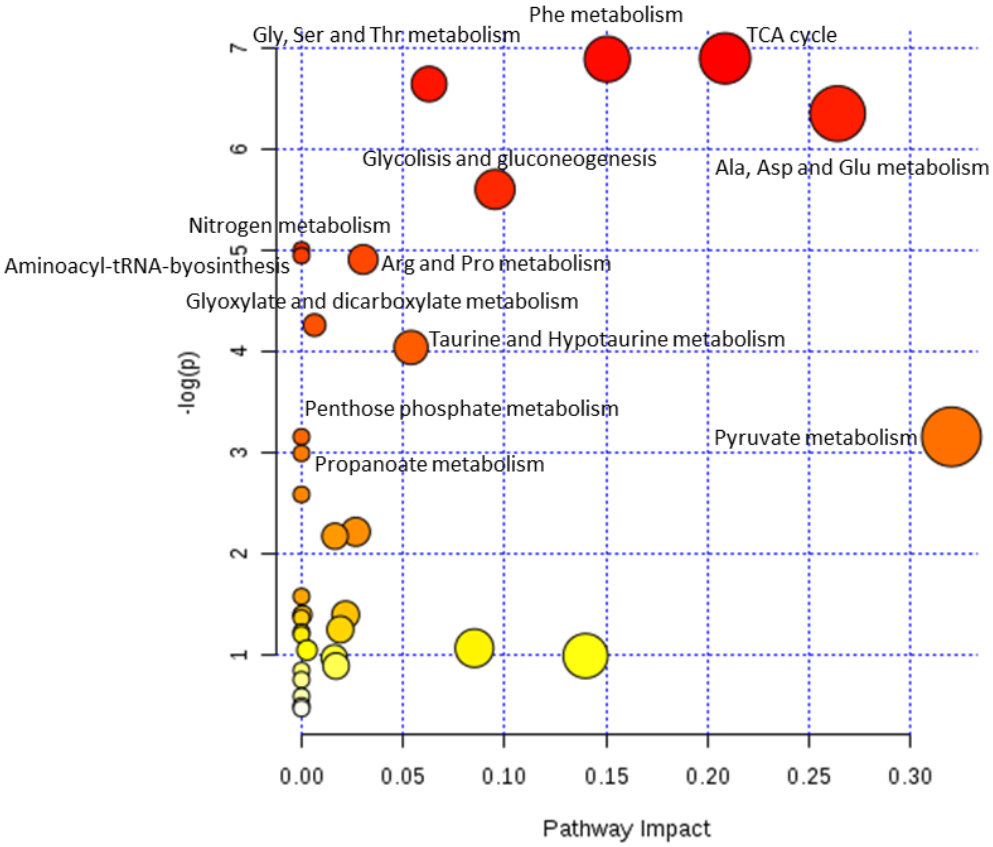
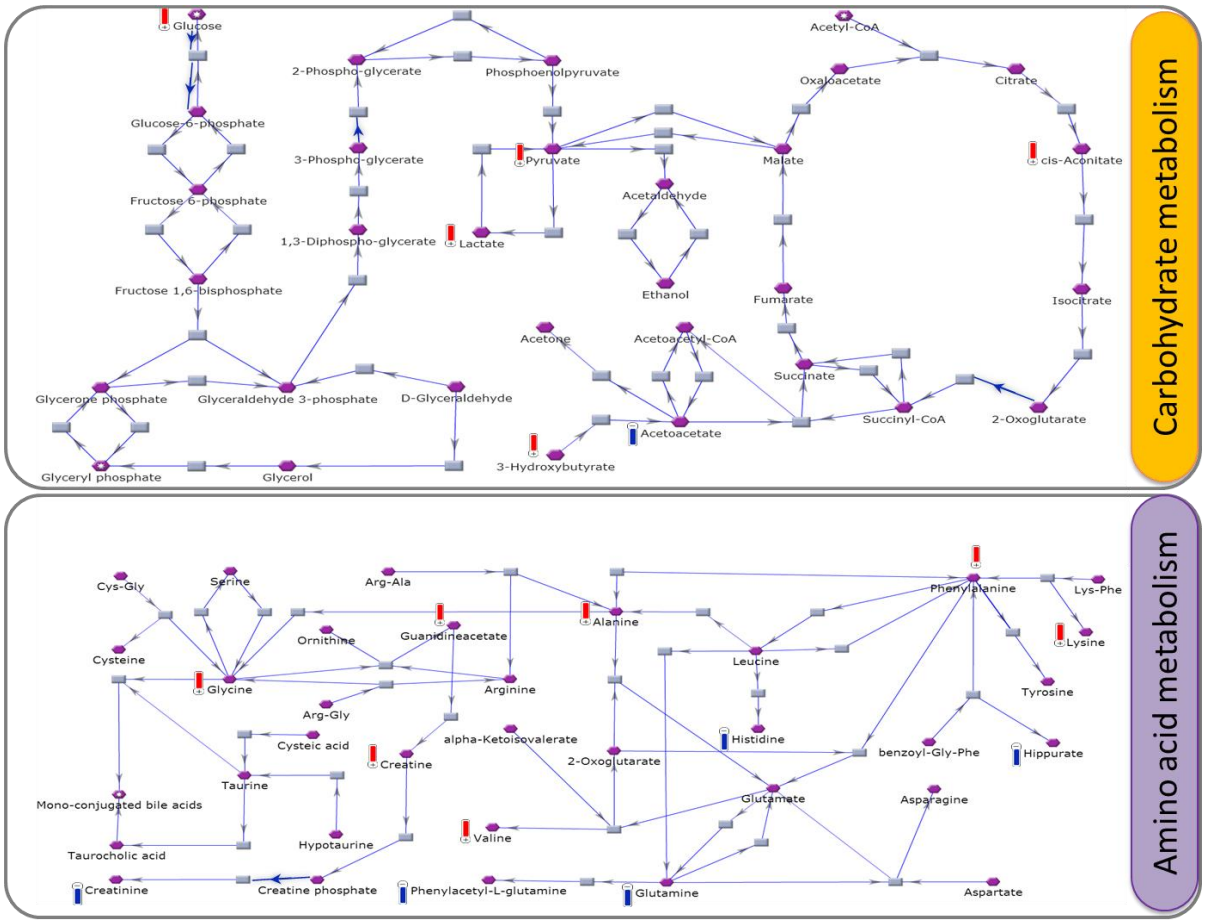


Figure S3 Up- and downregulated urinary metabolites (red and blue bars, respectively) in T2D participants in carbohydrate metabolism and amino acid metabolism pathways. Networks were built in MetaCore™.



Publicación original 2:

Enrique Almanza-Aguilera, Mireia Urpi-Sarda, Rafael Llorach, Rosa Vázquez-Fresno, Mar Garcia-Aloya, Francesc Carmona, Alex Sanchez, Francisco Madrid-Gambin, Ramon Estruch, Dolores Corella, Cristina Andres-Lacueva. Microbial metabolites are associated with a high adherence to a Mediterranean dietary pattern using a 1H-NMR-based untargeted metabolomics approach. *Journal of Nutricional Biochemistry*, 48 (2017) 36-43.

Participación del doctorando:

• Procesado y tratamiento de los datos y aplicación de herramientas bioestadísticas para corroborar las hipótesis de partida

• Identificación y confirmación de los metabolitos discriminantes

• Interpretación biológica

• Elaboración del primer borrador de la publicación

• Seguimiento y trabajo del manuscrito hasta su publicación

Available online at www.sciencedirect.com

ScienceDirect

Journal of Nutritional Biochemistry 48 (2017) 36–43

Journal of Nutritional Biochemistry

Microbial metabolites are associated with a high adherence to a Mediterranean dietary pattern using a ^{1}H-NMR-based untargeted metabolomics approach☆

Enrique Almanza-Aguilera[a], Mireia Urpi-Sarda[a,b,*], Rafael Llorach[a,b], Rosa Vázquez-Fresno[a], Mar Garcia-Aloy[a,b], Francesc Carmona[c], Alex Sanchez[c,d], Francisco Madrid-Gambin[a,b], Ramon Estruch[e,f], Dolores Corella[e,g], Cristina Andres-Lacueva[a,b,*]

[a]Biomarkers and Nutrimetabolomics Laboratory, Nutrition, Food Science and Gastronomy Department, XaRTA, INSA, Campus Torribera, Faculty of Pharmacy and Food Science, University of Barcelona, Barcelona 08028, Spain
[b]CIBER Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Madrid 28028, Spain
[c]Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain
[d]Statistics and Bioinformatics Unit. Vall d'Hebron Institut de Recerca, Barcelona, Spain
[e]CIBER Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Instituto de Salud Carlos III, Madrid 28028, Spain
[f]Department of Internal Medicine, Institut d'Investigacions Biomediques August Pi Sunyer (IDIBAPS), Hospital Clinic, University of Barcelona, Barcelona, 08036, Spain
[g]Department of Preventive Medicine, University of Valencia, Valencia 46010, Spain

Received 30 December 2016; received in revised form 9 May 2017; accepted 1 June 2017

Abstract

The study of biomarkers of dietary patterns including the Mediterranean diet (MedDiet) is scarce and could improve the assessment of these patterns. Moreover, it could provide a better understanding of health benefits of dietary patterns in nutritional epidemiology. We aimed to determine a robust and accurate biomarker associated with a high adherence to a MedDiet pattern that included dietary assessment and its biological effect. In this cross-sectional study, we included 56 and 63 individuals with high (H-MDA) and low (L-MDA) MedDiet adherence categories, respectively, all from the *Prevención con Dieta Mediterránea* trial. A ^{1}H-NMR-based untargeted metabolomics approach was applied to urine samples. Multivariate statistical analyses were conducted to determine the metabolite differences between groups. A stepwise logistic regression and receiver operating characteristic curves were used to build and evaluate the prediction model for H-MDA. Thirty-four metabolites were identified as discriminant between H-MDA and L-MDA. The fingerprint associated with H-MDA included higher excretion of proline betaine and phenylacetylglutamine, among others, and decreased amounts of metabolites related to glucose metabolism. Three microbial metabolites — phenylacetylglutamine, *p*-cresol and 4-hydroxyphenylacetate — were included in the prediction model of H-MDA (95% specificity, 95% sensitivity and 97% area under the curve). The model composed of microbial metabolites was the biomarker that defined high adherence to a Mediterranean dietary pattern. The overall metabolite profiling identified reflects the metabolic modulation produced by H-MDA. The proposed biomarker may be a better tool for assessing and aiding nutritional epidemiology in future associations between H-MDA and the prevention or amelioration of chronic diseases.
© 2017 Elsevier Inc. All rights reserved.

Keywords: Microbiota; Biomarkers; Metabolomics; Dietary patterns; Mediterranean diet; High adherence

Abbreviations: AUC, area under the curve; CVD, cardiovascular disease; FFQ, food frequency questionnaire; ^{1}H-NMR, proton nuclear magnetic resonance; H-MDA, high Mediterranean diet adherence; HPHPA, 3-(3-hydroxyphenyl)-3-hydroxypropanoate; L-MDA, low Mediterranean diet adherence; MedDiet, Mediterranean diet; MetS, metabolic syndrome; PREDIMED, Prevención con Dieta Mediterránea; OSC-PLS-DA, orthogonal signal correction–partial least squares discriminant analysis; PAGN, phenylacetylglutamine; ROC, receiver operating characteristic; VIP, variable importance in projection.

☆ This work was supported by Spanish national grants from the Ministry of Economy and Competitiveness (MINECO) and cofunded by Fondo Europeo de Desarrollo Regional: the EU Joint Programming Initiative "A Healthy Diet for a Healthy Life" on Biomarkers BioNHFOODBALL (PCIN-2014-133-MINECO-Spain), AGL2009-13906-C02-01, (co-funded by the FEDER Program from EU)CIBER de Fragilidad y Envejecimiento Saludable (CIBERFES) (co-funded by the FEDER Program from EU) and MTM2015-64465-C2-1-R. This work was also supported by Associació Catalana de Diabetis (ACD2015) and awards of 2014SGR1566 and 2014SGR464 (GRBIO) from the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) Generalitat de Catalunya.

* Corresponding authors at: Biomarkers and Nutrimetabolomics Laboratory, Nutrition, Food Science and Gastronomy Department, Food Campus Torribera, Faculty of Pharmacy and Food Science, University of Barcelona, Av. Joan XXIII s/n, 08028 Barcelona, Spain. Tel.: +34 934034844, +34 934034840.
E-mail addresses: murpi@ub.edu (M. Urpi-Sarda), candres@ub.edu (C. Andres-Lacueva).

http://dx.doi.org/10.1016/j.jnutbio.2017.06.001
0955-2863/© 2017 Elsevier Inc. All rights reserved.

1. Introduction

High adherence to healthy dietary patterns is associated with lower risk of chronic diseases [1]. The measurement of dietary intake is an essential component in studies attempting to establish links between dietary exposure and health outcomes [1]. Currently, the quality and adherence of dietary patterns in nutritional epidemiology are measured by the use of self-reported questionnaires such as the dietary indexes/scores [2]. A healthy dietary score is based on dietary recommendations as a result of scientific consensus or as proposed by investigators using an evidence-based approach [3]. Several dietary scores have been developed and applied to populations to evaluate the role of diet in a more holistic perspective in the risk of mortality, cardiovascular disease (CVD) and cancer [4]. However, because dietary patterns are complex not only in terms of composition but also in terms of amounts and frequency of food intake, it is well recognized that, in addition to the conventional methods, the emergence of novel biomarkers of food exposure may help to improve the accuracy of the assessment of compliance and adherence [5]. In this regard, metabolomics has emerged as a valuable tool in nutrition research for the discovery of novel dietary biomarkers for both single foods [6,7] and food patterns [5,8] and, in addition, is able to evaluate their effect in the organism [9]. The use of these biomarkers is a more specific tool and complementary to traditional indexes/scores. Moreover, although progress has been made in the metabolite characterization of dietary patterns, most of the metabolomic studies have applied approaches based on multivariate analyses of food and nutritional data intake [10,11], leaving the complementary use of dietary scores and metabolomics as a new field to explore in dietary patterns characterization.

Up to now, several Mediterranean diet (MedDiet) adherence scores have been described in the literature, such as the 9-item MedDiet score [12] or the *Prevención con Dieta Mediterránea* (PREDIMED) 14-item MedDiet score [13]. These scores were developed with the aim of appraising the adherence to a traditional MedDiet of several populations as well as to evaluate the effect of adherence to a MedDiet on microbiota composition [14], CVD risk factors [15], aging diseases [16] or total mortality [12]. The application of targeted and untargeted metabolomics approaches in the study of the effects of the protective mechanisms of a MedDiet on CVD has been poorly studied but is now beginning to attract more interest [9,17]. Furthermore, currently, there are some reports about metabolic profiling in biological samples (feces or urine) that enable the characterization of high adherence to a MedDiet pattern [14,18], but to our knowledge, there are no reports calculating a prediction model of MedDiet adherence. In addition, the study of biomarkers to explain the assessment of the pattern and gain a better understanding of its health benefits in nutritional epidemiology is limited. The characterization of dietary patterns by using metabolomic approaches is important because it would allow insights into the relationship between diet, taking into account the bioavailability of MedDiet bioactives, and the risk of chronic diseases [14,19].

In the current study, we aimed to determine a robust and accurate urinary biomarker associated with a high adherence to MedDiet pattern that included dietary assessment and its biological effect on the organism by using a ^{1}H-NMR-based untargeted metabolomics approach which can be usefully applied in nutritional epidemiology.

2. Subjects and methods

2.1. Study population and study design

The PREDIMED study is a multicenter, randomized, parallel and controlled clinical trial conducted in Spain and aimed at assessing the effects of a MedDiet on primary prevention of CVD. Full details of the design and methods have been published elsewhere [20,21]. Briefly, the study population included men (55–80 years) and women (60–80 years) without a previous history of CVD at enrolment but with either type 2 diabetes mellitus or at least three or more of the following CVD risk factors: current smoking, hypertension, high low-density lipoprotein cholesterol, low high-density lipoprotein cholesterol, overweight/obesity or family history of premature CVD. Exclusion criteria were the presence of any severe chronic illness, alcohol or drug abuse, body mass index (BMI; in kg/m^2) ≥40, and allergy or intolerance to olive oil or nuts. The trial was registered at http://www.controlled-trials.com (ISRCTN35739639).

For the current work, we conducted a cross-sectional study with baseline dietary data and urine samples of 119 individuals recruited in 2 PREDIMED trial centers (Hospital Clinic of Barcelona and University of Valencia). At baseline, one morning urine sample was collected from all participants and immediately aliquoted and stored at −80°C until the day of analysis.

2.2. Assessment of Mediterranean diet adherence and other parameters

In order to appraise the adherence to a MedDiet among participants, the validated 14-item PREDIMED MedDiet score questionnaire was administered. In detail, the MedDiet score questionnaire consists of 12 questions on food consumption frequency and 2 questions on food intake habits considered characteristic of the Spanish MedDiet [13]. Each item/question is scored as 1 or 0 according to whether it is met or not, respectively (Supplementary Table 1). Thus, the total MedDiet score ranges from 0 to 14 points, meaning that the higher the score, the higher the adherence to a MedDiet. The MedDiet adherence score was calculated for all participants and used for their subsequent stratification, which was done by using the proposed cutoff values previously reported in the PREDIMED study. [15]: MedDiet score ≤7 indicated low MedDiet adherence (L-MDA) (n=63), and MedDiet score ≥10 (n=56) indicated high MDA (H-MDA).

All participants were also asked to complete a validated semi-quantitative 137-item food frequency questionnaire (FFQ) [22] and the Spanish version of the Minnesota Leisure-Time Physical Activity Questionnaire [23]. Trained dieticians in the PREDIMED study administered all questionnaires, including the PREDIMED MedDiet score questionnaire. The nutrient composition and energy intakes were calculated from the FFQ data by using Spanish food composition tables [24]. Also, anthropometrical measurements were taken directly by qualified nurses.

2.3. Metabolite profiling

^{1}H-NMR analysis and spectra processing were performed by following previous methodology [8]. Briefly, the urine samples were thawed, vortexed and centrifuged at 13,200 rpm for 5 min. From each supernatant, a volume of 300 μl was taken and diluted with 200 μl of H_2O/D_2O and mixed with a buffer solution [8]. The optimized pH of the buffer was set at 7.0, with a potassium deuteroxide solution, to minimize variations in the chemical shifts of the ^{1}H-NMR resonances. This mixture was transferred to a 5-mm NMR tube. ^{1}H-NMR spectra acquisition was performed using a Varian-Inova 500-MHz NMR spectrometer with presaturation of the water resonance using a NOESYPRESAT pulse sequence. The spectra data processed were binned to 1165 variables with bin widths of 0.005 ppm and integrated with ACD/NMR Processor 12.0 software (Advanced Chemistry Development, Toronto, Canada). The spectral region containing water (δ 4.75–5.00 ppm) was excluded before normalization to avoid spectral interference. Integrated spectra were row-wise normalized by sum using MetaboAnalyst 3.0 (www.metaboanalyst.ca), a web server designed to permit comprehensive metabolomics data analysis.

Metabolites were identified following a multistep procedure as previously reported [8]. This multistep includes (a) comparison of the experimental NMR spectra with those in the library of Chenomx Suite 8.1 profiler software (Chenomx, Inc., Edmonton, Canada), which includes access to the Human Metabolome Database (HMDB) library [25]. Other databases such as the Biological Magnetic Resonance Data Bank and the Madison Metabolomics Consortium Database were also consulted along with existing NMR-based metabolomics literature. (b) At the same time, the NMR peak assignments were correlated using Pearson's correlation coefficient ($r \geq 0.7$, $P \leq .05$) to confirm the multiplicity and identify clusters of metabolites which were then compared to databases with NMR data [8,25]. Biological interpretation was made by consulting the HMDB and Kyoto Encyclopedia of Genes and Genomes databases.

2.4. Statistical analyses

Demographic characteristics, medication usage and dietary intake were compared between groups by conducting Student's *t* tests and χ^2 analyses for continuous and categorical variables, respectively. Variables with a non-normal distribution were log transformed before analyses. Multivariate data analysis was performed using SIMCA-P+13.0 (Umetrics, Umeå, Sweden) software. Data sets containing the integrated NMR spectral bins were log transformed and Pareto scaled before performing a principal component analysis to explore the quality of data acquisition. An orthogonal signal correction (OSC) filter was applied to the data sets in order to reduce the variability not associated with the dietary classification [8], in this case, the category of adherence to a MedDiet. Afterward, a partial least squares discriminant analysis (PLS-DA) was performed to examine the difference in metabolite profile between subjects at H-MDA and L-MDA. The quality and validation of resultant model were also appraised, the first through the R^2Y (*cum*) and Q^2 (*cum*) parameters, and the second with a permutation test ($n=200$). Discriminant variables between groups were selected based on their variable importance in projection (VIP) value >1.0, which is a generally accepted threshold [26]. The normality of discriminant variables was assessed using a Kolmogorov–Smirnov test with Lilliefors significance correction, and additionally Mann–Whitney or independent Student's *t* tests were performed according to the normality of the data. To account for multiple comparisons in the metabolomic analysis, we used a corrected *P* value with the Benjamini–Hochberg procedure.

Aiming to identify the metabolites with the best discriminant capability between H-MDA and L-MDA, we designed and assessed the performance of a prediction model with the H-MDA as dependent variable and the discriminant metabolites identified from the OSC-PLS-DA model as independent variables [6]. To this end, first the data set of individuals ($n=119$) was randomly split into two thirds to build one training set ($n=79$) and one third for a validation set ($n=40$). A stepwise binary logistic regression analysis was performed in the training set in order to identify the metabolites with the most significant predictive capacity, and from these, a combined model was built. Subsequently, a receiver operating characteristic (ROC) curve analysis was performed, first in the training set and then in the validation set, to evaluate the obtained model, as well as the individual metabolites included. Then the performance of both the combined model and the individual metabolites was determined by the area under the curve (AUC) of the ROC curves, as well as by the sensitivity and specificity at the optimal cutoff point defined as the minimum distance to the top-left corner in the ROC curve. Finally, the association between the combined multimetabolite model and individual metabolites with food groups was tested by a Spearman's rank correlation analysis with correction of *P* value using Benjamini–Hochberg procedure.

All univariate analyses, including normality, Student's *t* tests, Spearman's rank correlation, logistic regression and ROC curve analyses, were performed on IBM SPSS 21 statistics software (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Demographic and dietary intake measurements

For the current study, we included individuals who, according to their individual MedDiet score, were assigned to L-MDA (≤7 points, $n=63$) or H-MDA (≥10 points, $n=56$). Our population had a mean (±S.D.) age of 67±6 years and a mean (±S.D.) BMI of 30.3±4.5 kg/m^2, and 68.9% of the participants were women (Table 1). Cardiovascular risk factors as well as medication use were similar between both groups (Table 1). With regard to food and nutrient intake (Table 2), individuals in H-MDA consumed higher amounts of olive oil, nuts, vegetables, fruits, legumes and fish, and total dietary fiber than L-MDA participants ($P<.05$). However, no statistically significant differences were found for total energy, total fat, carbohydrates and protein intakes.

3.2. Discriminant metabolite profile of high adherence to a Mediterranean diet pattern

The OSC-PLS-DA analysis resulted in one latent component model with R^2Y (*cum*) and Q^2Y (*cum*) values of 0.913 and 0.764, respectively, indicating a good ability to classify individuals according to their MedDiet adherence. In addition, a permutation test ($n=200$), with

Table 1
Demographic characteristics and medication usage of 119 participants classified in low and high Mediterranean diet adherence at baseline

	Total	L-MDA (≤7 MedDiet score)	H-MDA (≥10 MedDiet score)	*P* value
Participants, *n*	119	63	56	
Age, y	66.7±5.9	66.8±6.1	66.6±5.7	.83
Women, *n* (%)	82 (68.9)	42 (66.7)	40 (71.4)	.58
BMI, kg/m^2	30.3±4.5	31.0±4.1	29.5±4.7	.08
Current smoker, *n* (%)	12 (10.1)	7 (11.1)	5 (8.9)	.69
Family history of CHD, *n* (%)	39 (32.8)	23 (36.5)	16 (28.6)	.34
Diabetes, *n* (%)	63 (52.9)	36 (57.1)	26 (46.4)	.24
Hypertension, *n* (%)	102 (85.7)	51 (81.0)	51 (91.1)	.12
Hypercholesterolemia, *n* (%)	85 (71.4)	44 (69.8)	41 (73.2)	.68
Medication use				
Oral antidiabetic agents, *n* (%)	46 (38.7)	26 (41.3)	20 (35.7)	.53
Insulin, *n* (%)	18 (15.1)	10 (16.0)	8 (14.3)	.81
Antihypertensive agents, *n* (%)	56 (47.1)	27 (42.9)	29 (51.8)	.33
Lipid-lowering medication, *n* (%)	58 (48.7)	29 (46.0)	29 (51.8)	.53

Values are mean ± S.D. or *n* (%) as appropriate. Differences between low and high Mediterranean diet adherence groups were tested by Student's *t* test and χ^2 test for continuous and categorical variables, respectively ($P<.05$). Abbreviation: CHD, coronary heart disease.

Table 2
Dietary intake of 119 participants with low and high adherence to Mediterranean diet

Food and nutrient intake	Total	L-MDA (≤7 MedDiet score)	H-MDA (≥10 MedDiet score)	*P* value
Participants, *n*	119	63	56	
Foods				
Olive oil, g/d	38±15	35±15	41±13	.024
Nuts, g/d	12±12	9±12	15±12	.018
Vegetables, g/d	402±192	377±215	430±160	.032
Fruits, g/d	480±242	419±235	550±233	.003
Legumes, g/d	20±10	17±9	22±10	.004
Cereals, g/d	269±111	275±112	263±111	.56
Fish, g/d	103±50	84±47	124±45	<.001
Meat, g/d	147±69	152±81	142±52	.41
Dairy, g/d	358±213	366±207	350±220	.70
Pastries, g/d	18±21	19±22	16±20	.44
Wine, ml/d	56±116	54±4	47±3	.71
Nutrients				
Total energy intake, kcal/d	2397±569	2399±635	2394±490	.96
Carbohydrates, % of total energy	43.3±6.8	43.4±6.6	43.2±7.0	.87
Proteins, % of total energy	16.7±2.6	16.4±2.8	17.1±2.4	.13
Total fat, % of total energy	38.0±6.2	37.8±6.5	38.2±5.8	.71
Dietary fiber, g/d	28.4±8.0	26.2±7.0	31.0±8.3	.001
MedDiet score	8.3±2.2	6.4±0.9	10.5±0.6	<.001

Values are mean ± S.D. Differences between low and high Mediterranean diet adherence groups were tested by Student's *t* test ($P<.05$).

intercept R^2 and Q^2 values of 0.346 and −0.154, respectively, showed the validity of the model. Table 3 shows a list of 34 metabolites that were identified after the selection of discriminant variables from the OSC-PLS-DA (based on VIP >1.0 values). Individuals in H-MDA group had a marked excretion of metabolites involved in protein/amino acid metabolism characterized by a higher excretion of anserine, carnosine, creatine, creatinine, guanidoacetate, histidine and *N*-acetylglutamine, as well as a lower excretion of 3-methylhistidine, alanine, glycine and lysine, than those in L-MDA. Similarly, another major group of metabolites derived from gut microbiota was identified. Concretely, H-MDA was characterized by a higher urinary amount of 4-hydroxyhippurate, 4-hydroxyphenylacetate, dimethylsulfone, 3-(3-hydroxyphenyl)-3-hydroxypropanoate (HPHPA), *p*-cresol and phenylacetylglutamine (PAGN) and lower urinary amounts of 3-indoxyl sulfate, hippurate and isobutyrate compared to the L-MDA. Furthermore, participants in the H-MDA group excreted lower levels of metabolites involved in the energy pathway, and in the propanoate and purine and caffeine pathways, than those in L-MDA. In addition, participants in the H-MDA group had higher levels of metabolites involved in the choline pathway (except betaine), as well as the inositol, niconitate, nicotidamide and pyrimidine pathways (Table 3).

3.3. Prediction model for high adherence to Mediterranean dietary pattern

We performed a model based on a stepwise binary logistic regression analysis including the previous 34 metabolites identified from the OSC-PLS-DA model and further ROC curve analyses to evaluate the resulting model and the individual metabolites included in this. To this end, the data set of individuals from the H-MDA and L-MDA groups was divided into training and validation sets, as indicated above. The resulting model included three metabolites derived from gut microbiota, namely, PAGN, *p*-cresol and 4-hydroxyphenylacetate (Supplementary Table 2). The specificity and sensitivity of the model were higher than 90% in both the training and validation sets. The PAGN had 85.7% and 68.4% specificity and sensitivity, respectively, in the validation set, while *p*-cresol and 4-hydroxyphenylacetate showed values between 66% and 84% in these parameters (Table 4) for the validation set. Furthermore, the global performance of the combined model in terms of AUC was 97.7% for the training and 97.0% for the validation set, while individually, each metabolite had values of AUC between 59% and 86% (Table 4). Fig. 1 illustrates that the model improves the classification of MedDiet adherence (H-MDA and L-MDA) in comparison with the use of

Table 3
List of discriminant urinary metabolites between individuals with low and high Mediterranean diet adherence[a]

Metabolite	^{1}H δ, ppm (multiplicity)	Excretion in H-MDA	*P* value[b]
Protein/amino acid metabolism			
3-Methylhistidine	7.95 (s)	↓	2.9×10^{-2}
Alanine	1.49 (d)	↓	9.4×10^{-3}
Anserine	3.79 (s), 8.29 (s)	↑	7.8×10^{-4}
Carnosine	7.18 (s), 8.15 (s)	↑	3.8×10^{-2}
Creatine	3.94 (s)	↑	7.6×10^{-3}
Creatinine	3.05 (s), 4.07 (s)	↑	2.3×10^{-3}
Glycine	3.57 (s)	↓	9.1×10^{-3}
Guanidoacetate	3.78 (s)	↑	6.7×10^{-3}
Histidine	7.09 (s), 7.93 (s)	↑	4.8×10^{-4}
Lysine	1.72 (m)	↓	2.1×10^{-2}
N-Acetylglutamine	2.08 (m), 2.30 (m)	↑	6.0×10^{-3}
Proline betaine	3.11 (s), 3.30 (s), 4.08 (m)	↑	6.2×10^{-7}
Gut microbiota metabolites			
3-Indoxyl sulfate	7.27 (t), 7.51 (d), 7.71 (d)	↓	1.6×10^{-4}
4-Hydroxyhippurate	6.98 (d), 7.75 (d)	↑	5.3×10^{-4}
4-Hydroxyphenylacetate	6.88 (d), 7.16 (s)	↑	5.0×10^{-4}
Dimethylsulfone	3.16 (s)	↑	6.7×10^{-3}
Hippurate	7.55 (tt), 7.82 (dd)	↓	2.0×10^{-3}
HPHPA	5.02 (t), 6.86 (d), 6.93 (br s), 7.30 (t)	↑	3.4×10^{-2}
Isobutyrate	1.07 (d)	↓	2.2×10^{-4}
p-Cresol	2.34 (s), 7.22 (d), 7.28 (d)	↑	9.6×10^{-5}
Phenylacetylglutamine	1.93 (m), 2.27 (m), 3.66 (m), 4.18 (m), 7.36 (t), 7.43 (t), 8.0 (d)	↑	1.5×10^{-11}
Energy metabolism (glycolysis/gluconeogenesis, TCA and ketone bodies)			
β-Glucose	3.39 (m), 3.44 (m), 3.49 (m), 3.90 (m), 4.66 (d)	↓	1.9×10^{-4}
Lactate	1.35 (d)	↓	1.1×10^{-2}
Succinate	2.42 (s)	↓	4.7×10^{-3}
Choline metabolism			
Dimethylamine	2.72 (s)	↑	5.2×10^{-4}
Betaine	3.27 (s), 3.91 (s)	↓	1.2×10^{-2}
TMAO	3.28 (s)	↑	2.6×10^{-3}
Inositol phosphate metabolism			
Scyllo-inositol	3.36 (s)	↑	7.8×10^{-4}
Nicotinate and nicotinamide metabolism			
N-Methylnicotinamide	4.49 (s), 8.85 (m), 9.13 (s)	↑	$9.3\text{x}10^{-3}$
Propanoate metabolism			
Isopropanol	1.15 (d)	↓	1.9×10^{-2}
Purine and caffeine metabolism			
Xanthosine	7.86 (s)	↓	1.1×10^{-2}
Pyrimidine metabolism			
Urea	5.82 (br.s)	↑	3.8×10^{-3}
Methylguanidine	2.82 (s)	↑	2.0×10^{-2}
Malonate	3.12 (s)	↑	2.7×10^{-2}

[a] Obtained from the OSC-PLS-DA model (VIP >1.0) and using the high Mediterranean diet adherence group as reference. (↑) and (↓) indicate a relatively higher and lower excretion, respectively, in H-MDA with respect to L-MDA.

[b] Differences between NMR-signals were tested by Student's *t* test or Mann–Whitney *U* tests according to their normal distribution. Significance was corrected by the Benjamini–Hochberg procedure. Abbreviations: multiplicity (s, singlet; br. s, broad singlet; d, doublet; dd, doublet of doublets; t, triplet; tt, triplet of triplets; m, multiplet).

Table 4
ROC curve parameters of combined multimetabolite model and individual metabolites for prediction of high Mediterranean diet adherence

	ROC curve parameters							
	Training set				Validation set			
	Specificity (%)	Sensitivity (%)	AUC (95% CI)	*P* value	Specificity (%)	Sensitivity (%)	AUC (95% CI)	*P* value
Multimetabolite model	93.0	94.6	97.7 (95.0–100)	<.001	95.2	94.7	97.0 (92.3–100)	<.001
Phenylacetylglutamine	85.7	83.8	89.7 (82.7–96.7)	<.001	85.7	68.4	81.7 (68.5–94.9)	.001
p-Cresol	59.5	81.1	73.2 (61.7–84.8)	<.001	76.2	73.7	77.2 (62.4–92.0)	.003
4-Hydroxyphenylacetate	62.0	73.0	68.1 (56.4–79.9)	.006	66.7	84.2	77.7 (63.1–92.3)	.003

each metabolite individually. In the analysis of correlation, the combined metabolite model showed a strong correlation with the MedDiet score (r=0.7; P<.001), as well as with the intake of vegetables, fruits, legumes, fish (r=0.2–0.3; P<.01) (Supplementary Table 3) and dietary fiber (r=0.3; P<.01). Otherwise, the three individual metabolites had good correlation with the MedDiet score (r=0.3–0.6; P<.001) but weaker or not significant correlations with the intake of individual foods.

4. Discussion

In the present metabolomic study, we identified the urinary metabolite profile consisting of 34 metabolites that enable discrimination between 2 groups of individuals with high or low adherence to the score of Mediterranean diet adherence validated in the PREDIMED study [13]. Participants in the L-MDA group had a cutoff of ≤7 points and those in the H-MDA group ≥10 points, as previously proposed [15].

The set of metabolites that discriminated between H-MDA and L-MDA suggested the metabolic modulation of the MedDiet. These metabolites are involved in multiple molecular mechanisms and metabolic pathways, which together provide a holistic view of variations in the urine metabolome due to the effect of follow this dietary pattern. Some of these metabolites have previously been proposed as putative biomarkers of single food intake and also related to foods included in MedDiet, as well as up- and down-regulated endogenous metabolites. We found that the higher excretion of proline betaine and scyllo-inositol in H-MDA correlated significantly with the intake of citrus fruits (r=0.36 and r=0.35; P<.001, respectively), which is consistent with previous studies that have proposed them as markers of citrus fruit intake [27]. Recently, in our previous work, we also found positive correlations of proline betaine signals with orange consumption in long-term MedDiet intervention and low-fat diet groups in a subsample of nondiabetic participants of the PREDIMED study [8]. In this regard, in the H-MDA group, we also identified other metabolites derived from the microbial metabolism: 4-hydroxyhippurate, 3-(3-hydroxyphenyl)-3-hydroxypropanoate and 4-hydroxyphenylacetate. These metabolites have been described after interventions with mixed red wine/grape juice extracts [28], as well as linked to the intake of other polyphenol-rich foods such as cocoa and almond, among others (http://phenol-explorer.eu/). It is interesting to note that these microbial metabolites were not found in our previous work evaluating the intervention with MedDiet [8].

Other metabolites have been related to the intake of meat or fish. Both dimethylamine and trimethylamine-*N*-oxide (TMAO), which were higher in the H-MDA group than in the L-MDA, have been related to the intake of fish and seafood [29,30]. Nevertheless, only TMAO had a significant correlation with the intake of total fish (r=0.24, P<.01) in our study sample, and this correlation was not previously observed in the subset of samples of nondiabetics in the PREDIMED study [8]. In addition, we identified some compounds related to the histidine pathway: carnosine, anserine and 3-methylhistidine. While 3-methylhistidine and carnosine are proposed as good biomarkers of red meat intake [31–33], anserine has alternatively been proposed as a marker of white meat (*i.e.*, poultry) [34] or salmon intake [35]. In our study, subjects in the H-MDA group excreted lower amounts of 3-methylhistidine and higher amounts of carnosine and anserine than those in the L-MDA group. We found significant correlations between anserine, with antioxidant properties, and the intake of white meats (r=0.231, P<.01) but not with fish (P>.05). In addition, we did not find significant correlations between methylhistidine or carnosine and foods. Nevertheless, it should be noted that carnosine is a normal constituent in human urine, which occurs naturally in the skeletal muscle of mammals and has antioxidant properties and therapeutic potential against numerous diseases [36]. To the best of our knowledge, there is limited information about which mechanisms could increase these antioxidants in the organism, but a recent study

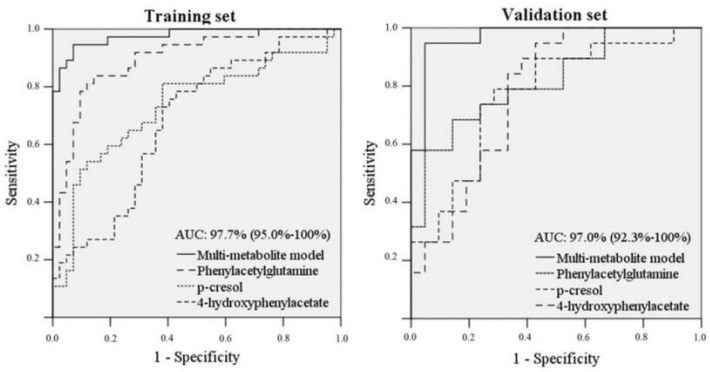


Fig. 1. ROC curves of the combined multimetabolite model and individual metabolites included for discrimination between high and low Mediterranean diet adherence in both training and validation sets.

has shown that dietary vitamin B6 could determine the carnosine concentration in the skeletal muscle of rats [37]. In this regard, we observed significantly higher ingested concentrations of vitamin B6 in subjects in the H-MDA group (2.9±0.7 mg/d) than in those in the L-MDA group (2.4±0.5 mg/d, $P<.001$). Moreover, 3-methylhistidine is a metabolite of anserine and carnosine in the histidine pathway. 3-Methylhistidine is also found in urine as anserine and carnosine, and its concentration was inversely correlated with them in our study ($r=-0.32$, $P<.05$; $r=-0.27$; $P<.05$, respectively). Thus, the contribution of a MedDiet in the modulation of the histidine pathway could be an interesting field to further explore.

Besides the metabolites related to the intake of foods, we also identified other endogenous metabolites involved in energy metabolism. In particular, individuals in the H-MDA group showed lower excretion of glucose, lactate and succinate than those in the L-MDA group, whose metabolites are related to pathways affected in diabetic patients and other diseases [38]. This finding is interesting because although our population included diabetic individuals, their distribution and the use of medication between the H-MDA and L-MDA groups were well balanced (Table 1). This fact suggested therefore that, in comparison to the L-MDA group, individuals in the H-MDA group could show relatively better glycemic control. Supporting this notion, several studies have previously found an inverse association between adherence to a MedDiet and indices of glucose homeostasis in the general population, including elderly people, and high-risk patients [39]. Previous works comparing intervention with MedDiet and low-fat diet in nondiabetic PREDIMED participants did not find changes in these metabolites [8].

After identifying 34 metabolites in the H-MDA pattern, we studied the prediction of high adherence to MedDiet pattern. To improve the prediction of H-MDA, a model with a combination of more than one discriminatory metabolite was developed. For this purpose, population was split into training and validation sets. The model included PAGN, 4-hydroxyphenylacetate and *p*-cresol. Interestingly and in accordance with this finding, in our previous study, we also reported associations of PAGN and *p*-cresol with long-term interventions with MedDiet [8]. Currently, the use of multimetabolite biomarkers is still limited. In a recent study, Marklund et al. (2014) combined six serum metabolites to create a dietary biomarker score [19]. This score was aimed at assessing the compliance with a healthy Nordic diet in a population with metabolic syndrome (MetS) and subsequently used to analyze the effects of diet on cardiometabolic risk factors among those individuals with highest compliance to this diet [19]. In our study, the inclusion of PAGN, 4-hydroxyphenylacetate and *p*-cresol in the model, as well as the identification of other metabolites derived from gut microbiota in the previous set of discriminant metabolites, highlights the role of a MedDiet in the modulation of gut microbiota. Supporting this notion, De Filippis et al. (2015) demonstrated the relationship between the level of MedDiet adherence, gut microbiota and microbial metabolites [18]. Later, Haro et al. (2016) demonstrated that a long-term intervention with a MedDiet partially restores the alteration in the gut microbiota composition in individuals with MetS [40]. Interestingly, by using the same PREDIMED MedDiet score, authors found a weak but significant correlation between this score and the abundance of *Faecalibacterium prausnitzii* and *Bifidobacterium adolescentis*, which were observed to be decreased in MetS patients compared with non-MetS patients [40]. More recently, Gutierrez-Díaz et al. (2017) found increased fecal concentrations of benzoic and 3-hydroxyphenylacetic acids in individuals with higher adherence to MedDiet [14] and Vazquez-Fresno et al. (2016) found that 4-hydroxyphenylacetate significantly increased in urine in a "healthier" cluster of participants compared to "obese and diabetic" cluster after intervention with wine polyphenol intake, exhibiting a distinct post-intervention metabolic response between groups possibly associated with differences/alteration in gut microbiota metabolism [41]. Interestingly and in accordance with our results, a hydroxyphenylacetic acid metabolite entered in our multimetabolite model, thus supporting the notion that H-MDA modulates gut microbiota metabolism. Nowadays, there exists an increasing interest in the study of the relationship between MedDiet adherence and food-derived alterations of the gut microbiota in order to use these data in the prevention of food-related diseases [42].

The three metabolites included in our multimetabolite model may arise from several food sources. PAGN is formed by the conjugation of glutamine with phenylacetate, which can arise from endogenous β-oxidation of phenyl-containing fatty acids or phenylalanine metabolism [43], or be obtained through the exogenous intake contained in plant-food sources [44]. In line with our data, O'Sullivan et al. (2011) found a positive association between the urinary concentration of PAGN and vegetable intake, suggesting that PAGN in urine may be a useful biomarker of vegetable intake [11]. *p*-Cresol is a product of microbial tyrosine breakdown *via* 4-hydroxyphenylacetate [45]. Moreover, 4-hydroxyphenylacetate has been found to be related to the intake of vegetarian diets [31] and polyphenol-rich foods, including red wine [46] and dark chocolate [47]. A further correlation analysis showed that although we found weak or nonsignificant correlations between individual metabolites and some food groups, the combined multimetabolite model was significantly associated with the intake of vegetables, fruits, legumes, fish and total fiber (Supplementary Table 3).

The main limitation of this cross-sectional study is that our results may not be generalized or extrapolated to other populations, mainly because of the age of our population as well as their high risk of CVD. However, the validity of the PREDIMED MedDiet score in distinguishing between individuals at high or low MedDiet adherence has been ascertained within the prominent PREDIMED study [15]. Moreover, the current complementary use of this score with an NMR-based untargeted metabolomics approach deserves mention, first, because dietary scores are still the main tool to assess dietary patterns adherence and, second, because this robust analytical platform allows us to identify differences in urinary metabolome at micro- to millimolar levels [48].

In conclusion, the model composed of microbial metabolites was the biomarker that defined high adherence to a Mediterranean dietary pattern. This fact highlights the role of microbiota in the study of the biomarkers associated to the MedDiet pattern. The effect of the MedDiet involves several interconnected molecular mechanisms through complex regulatory networks, which are reflected in the microbiota metabolism as the metabolic modulation of H-MDA. Future studies in nutritional research should have to include the measurement of these dietary biomarkers in order not only to improve the assessment of dietary intake but also to understand in-depth the molecular mechanisms involved in the effects associated with food intake. The proposed biomarker may assess and aid nutritional epidemiology in future associations between adherence to the MedDiet and the prevention or amelioration of chronic diseases.

Conflict of interest

The authors declare no financial or personal conflicts of interest.

Role of the funding sources

The funding sources had no role in the study design, data collection and analysis, decision to publish or the preparation of the manuscript.

Authorship

The authors' contributions to the manuscript were as follows: M.U.S., R.L. and C.A.L. designed the research; R.E. and D.C. provided the clinical samples from the PREDIMED study; R.V.F., E.A.A. and F.M.G. conducted the NMR experimental data and the NMR data analysis;

E.A.A., M.G.A., M.U.S., R.L., F.C. and A.S. conducted the statistical analyses; E.A.A. and M.U.S. wrote the paper; All authors provided critical revision. M.U.S. and C.A.L. have the primary responsibility for the final content. All authors read and approved the final manuscript.

Acknowledgments

M.U.S. would like to thank the "Ramón y Cajal" program from MINECO and the Fondo Social Europeo (RYC-2011-09677). E.A.A. would like to thank CONACYT (México) for the Ph.D. fellowship. F.M.-G. acknowledges the APIF Ph.D. fellowship (University of Barcelona).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jnutbio.2017.06.001.

References

[1] Boeing H. Nutritional epidemiology: new perspectives for understanding the diet-disease relationship? Eur J Clin Nutr 2013;67:424–9. http://dx.doi.org/10.1038/ejcn.2013.47.
[2] Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002;13. http://dx.doi.org/10.1097/00041433-200202000-00002.
[3] Scientific Report of the 2015 Dietary Guidelines Advisory Committee (DGAC). Part D. Chapter 2: dietary patterns, foods and nutrients, and health outcomes; 2015. http://dx.doi.org/10.1017/CBO9781107415324.004.
[4] Harmon BE, Boushey CJ, Shvetsov YB, Ettienne R, Reedy J, Wilkens LR, et al. Associations of key diet-quality indexes with mortality in the multiethnic cohort: the dietary patterns methods project. Am J Clin Nutr 2015;101:587–97. http://dx.doi.org/10.3945/ajcn.114.090688.
[5] Andersen M-BS, Rinnan A, Manach C, Poulsen SK, Pujos-Guillot E, Larsen TM, et al. Untargeted metabolomics as a screening tool for estimating compliance to a dietary pattern. J Proteome Res 2014;13:1405–18. http://dx.doi.org/10.1021/pr400964s.
[6] Garcia-Aloy M, Llorach R, Urpi-Sarda M, Tulipani S, Estruch R, Martínez-González MA, et al. Novel multimetabolite prediction of walnut consumption by a urinary biomarker model in a free-living population: the PREDIMED study. J Proteome Res 2014;13:3476–83. http://dx.doi.org/10.1021/pr500425r.
[7] Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas M-E, Kochhar S, et al. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. Am J Clin Nutr 2010;92:436–43. http://dx.doi.org/10.3945/ajcn.2010.29672.
[8] Vazquez-Fresno R, Llorach R, Urpi-Sarda M, Lupianez-Barbero A, Estruch R, Corella D, et al. Metabolomic pattern analysis after Mediterranean diet intervention in a nondiabetic population: a 1- and 3-year follow-up in the PREDIMED study. J Proteome Res 2015;14:531–40. http://dx.doi.org/10.1021/pr5007894.
[9] Bondia-Pons I, Martinez JA, de la Iglesia R, Lopez-Legarrea P, Poutanen K, Hanhineva K, et al. Effects of short- and long-term Mediterranean-based dietary treatment on plasma LC-QTOF/MS metabolic profiling of subjects with metabolic syndrome features: the Metabolic Syndrome Reduction in Navarra (RESMENA) randomized controlled trial. Mol Nutr Food Res 2015;59:711–28. http://dx.doi.org/10.1002/mnfr.201400309.
[10] Altmaier E, Kastenmüller G, Römisch-Margl W, Thorand B, Weinberger KM, Illig T, et al. Questionnaire-based self-reported nutrition habits associate with serum metabolism as revealed by quantitative targeted metabolomics. Eur J Epidemiol 2011;26:145–56. http://dx.doi.org/10.1007/s10654-010-9524-7.
[11] O'Sullivan A, Gibney MJ, Brennan L. Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies. Am J Clin Nutr 2011;93:314–21. http://dx.doi.org/10.3945/ajcn.110.000950.
[12] Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. N Engl J Med 2003;348:2599–608. http://dx.doi.org/10.1056/NEJMoa025039.
[13] Schroder H, Fito M, Estruch R, Martinez-Gonzalez MA, Corella D, Salas-Salvado J, et al. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. J Nutr 2011;141:1140–5. http://dx.doi.org/10.3945/jn.110.135566.
[14] Gutierrez-Diaz I, Fernandez-Navarro T, Salazar N, Bartolome B, Moreno-Arribas MV, de Andres-Galiana EJ, et al. Adherence to a Mediterranean diet influences the fecal metabolic profile of microbial-derived phenolics in a Spanish cohort of middle-age and older people. J Agric Food Chem 2017;65:586–95. http://dx.doi.org/10.1021/acs.jafc.6b04408.
[15] Martinez-González MA, Garcia-Arellano A, Toledo E, Salas-Salvado J, Buil-Cosiales P, Corella D, et al. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. PLoS One 2012;7:e43134. http://dx.doi.org/10.1371/journal.pone.0043134.
[16] Talegawkar SA, Bandinelli S, Bandeen-Roche K, Chen P, Milaneschi Y, Tanaka T, et al. A higher adherence to a Mediterranean-style diet is inversely associated with the development of frailty in community-dwelling elderly men and women. J Nutr 2012;142:2161–6. http://dx.doi.org/10.3945/jn.112.165498.
[17] Guasch-Ferre M, Zheng Y, Ruiz-Canela M, Hruby A, Martinez-Gonzalez MA, Clish CB, et al. Plasma acylcarnitines and risk of cardiovascular disease: effect of Mediterranean diet interventions. Am J Clin Nutr 2016. http://dx.doi.org/10.3945/ajcn.116.130492.
[18] De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut 2015. http://dx.doi.org/10.1136/gutjnl-2015-309957.
[19] Marklund M, Magnusdottir OK, Rosqvist F, Cloetens L, Landberg R, Kolehmainen M, et al. A dietary biomarker approach captures compliance and cardiometabolic effects of a healthy Nordic diet in individuals with metabolic syndrome. J Nutr 2014;144:1642–9. http://dx.doi.org/10.3945/jn.114.193771.
[20] Martínez-González MÁ, Corella D, Salas-Salvadó J, Ros E, Covas MI, Fiol M, et al. Cohort profile: design and methods of the PREDIMED study. Int J Epidemiol 2012;41:377–85. http://dx.doi.org/10.1093/ije/dyq250.
[21] Estruch R, Ros E, Salas-Salvado J, Covas M-I, Corella D, Aros F, et al. Primary prevention of cardiovascular disease with a Mediterranean diet. N Engl J Med 2013;368:1279–90. http://dx.doi.org/10.1056/NEJMoa1200303.
[22] Fernandez-Ballart JD, Pinol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. Br J Nutr 2010;103:1808–16. http://dx.doi.org/10.1017/S0007114509993837.
[23] Elosua R, Marrugat J, Molina L, Pons S, Pujol E. Validation of the Minnesota Leisure Time Physical Activity Questionnaire in Spanish men. The MARATHOM investigators. Am J Epidemiol 1994;139:1197–209.
[24] Mataix Verdú J. Tabla de Composición de Alimentos. Universidad de Granada; 2003.
[25] Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0 — the Human Metabolome Database in 2013. Nucleic Acids Res 2013;41:D801–7. http://dx.doi.org/10.1093/nar/gks1065.
[26] Feng Q, Liu Z, Zhong S, Li R, Xia H, Jie Z, et al. Integrated metabolomics and metagenomics analysis of plasma and urine identified microbial metabolites associated with coronary heart disease. Sci Rep 2016;6:22525. http://dx.doi.org/10.1038/srep22525.
[27] Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, et al. Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr 2014;100:208–17. http://dx.doi.org/10.3945/ajcn.113.078758.
[28] van Dorsten FA, Grun CH, van Velzen EJJ, Jacobs DM, Draijer R, van Duynhoven JPM. The metabolic fate of red wine and grape juice polyphenols in humans assessed by metabolomics. Mol Nutr Food Res 2010;54:897–908. http://dx.doi.org/10.1002/mnfr.200900212.
[29] Mitchell SC, Zhang AQ, Smith RL. Dimethylamine and diet. Food Chem Toxicol 2008;46:1734–8. http://dx.doi.org/10.1016/j.fct.2008.01.010.
[30] Schmedes M, Aadland EK, Sundekilde UK, Jacques H, Lavigne C, Graff IE, et al. Lean-seafood intake decreases urinary markers of mitochondrial lipid and energy metabolism in healthy subjects: metabolomics results from a randomized crossover intervention study. Mol Nutr Food Res 2016. http://dx.doi.org/10.1002/mnfr.201500785.
[31] Stella C, Beckwith-hall B, Cloarec O, Holmes E, Lindon JC, Powell J, et al. Susceptibility of human metabolic phenotypes to dietary modulation. J Proteome Res 2006;5:2780–8. http://dx.doi.org/10.1021/pr060265y.
[32] Altorf-van der Kuil W, Brink EJ, Boetje M, Siebelink E, Bijlsma S, Engberink MF, et al. Identification of biomarkers for intake of protein from meat, dairy products and grains: a controlled dietary intervention study. Br J Nutr 2013;110:810–22. http://dx.doi.org/10.1017/S0007114512005788.
[33] Cross AJ, Major JM, Sinha R. Urinary biomarkers of meat consumption. Cancer Epidemiol Biomarkers Prev 2011;20:1107–11. http://dx.doi.org/10.1158/1055-9965.EPI-11-0048.
[34] Yeum K-J, Orioli M, Regazzoni L, Carini M, Rasmussen H, Russell RM, et al. Profiling histidine dipeptides in plasma and urine after ingesting beef, chicken or chicken broth in humans. Amino Acids 2010;38:847–58. http://dx.doi.org/10.1007/s00726-009-0291-2.
[35] Lloyd AJ, Beckmann M, Haldar S, Seal C, Brandt K, Draper J. Data-driven strategy for the discovery of potential urinary biomarkers of habitual dietary exposure. Am J Clin Nutr 2013;97:377–89. http://dx.doi.org/10.3945/ajcn.112.048033.
[36] Boldyrev AA, Aldini G, Derave W. Physiology and pathophysiology of carnosine. Physiol Rev 2013;93:1803–45. http://dx.doi.org/10.1152/physrev.00039.2012.
[37] Suidasari S, Stautemas J, Uragami S, Yanaka N, Derave W, Kato N. Carnosine content in skeletal muscle is dependent on vitamin B6 status in rats. Front Nutr 2015;2:39. http://dx.doi.org/10.3389/fnut.2015.00039.
[38] Urpi-Sarda M, Almanza-Aguilera E, Tulipani S, Tinahones FJ, Salas-Salvadó J, Andres-Lacueva C. Metabolomics for biomarkers of type 2 diabetes mellitus: advances and nutritional intervention trends. Curr Cardiovasc Risk Rep 2015;9. http://dx.doi.org/10.1007/s12170-015-0440-y.
[39] Salas-Salvadó J, Bulló M, Estruch R, Ros E, Covas M-I, Ibarrola-Jurado N, et al. Prevention of diabetes with Mediterranean diets: a subgroup analysis of a randomized trial. Ann Intern Med 2014;160:1–10. http://dx.doi.org/10.7326/M13-1725.
[40] Haro C, Garcia-Carpintero S, Alcala-Diaz JF, Gomez-Delgado F, Delgado-Lista J, Perez-Martinez P, et al. The gut microbial community in metabolic syndrome patients is modified by diet. J Nutr Biochem 2016;27:27–31. http://dx.doi.org/10.1016/j.jnutbio.2015.08.011.

[41] Vazquez-Fresno R, Llorach R, Perera A, Mandal R, Feliz M, Tinahones FJ, et al. Clinical phenotype clustering in cardiovascular risk patients for the identification of responsive metabotypes after red wine polyphenol intake. J Nutr Biochem 2016;28:114–20. http://dx.doi.org/10.1016/j.jnutbio.2015.10.002.
[42] Del Chierico F, Vernocchi P, Dallapiccola B, Putignani L. Mediterranean diet and health: food effects on gut microbiota and disease control. Int J Mol Sci 2014;15: 11678–99. http://dx.doi.org/10.3390/ijms150711678.
[43] Nicholls AW, Nicholson JK, Haselden JN, Waterfield CJ. A metabonomic approach to the investigation of drug-induced phospholipidosis: an NMR spectroscopy and pattern recognition study. Biomarkers 2000;5:410–23. http://dx.doi.org/10.1080/135475000750052420.
[44] Wightman F, Lighty DL. Identification of phenylacetic acid as a natural auxin in the shoots of higher plants. Physiol Plant 1982;55:17–24. http://dx.doi.org/10.1111/j.1399-3054.1982.tb00278.x.
[45] Ramakrishna BS, Roberts-Thomson IC, Pannall PR, Roediger WE. Impaired sulphation of phenol by the colonic mucosa in quiescent and active ulcerative colitis. Gut 1991;32:46–9.
[46] Vazquez-Fresno R, Llorach R, Alcaro F, Rodriguez MA, Vinaixa M, Chiva-Blanch G, et al. (1)H-NMR-based metabolomic analysis of the effect of moderate wine consumption on subjects with cardiovascular risk factors. Electrophoresis 2012; 33:2345–54. http://dx.doi.org/10.1002/elps.201100646.
[47] Martin F-PJ, Rezzi S, Pere-Trepat E, Kamlage B, Collino S, Leibold E, et al. Metabolic effects of dark chocolate consumption on energy, gut microbiota, and stress-related metabolism in free-living subjects. J Proteome Res 2009;8:5568–79. http://dx.doi.org/10.1021/pr900607v.
[48] Liu H, Tayyari F, Edison AS, Su Z, Gu L. NMR-based metabolomics reveals urinary metabolome modifications in female Sprague–Dawley rats by cranberry procyanidins. J Nutr Biochem 2016;34:136–45. http://dx.doi.org/10.1016/j.jnutbio.2016.05.007.

Publicación original 3:

Enrique Almanza-Aguilera, Carl Brunius, M Rosa Bernal-Lopez, Mar Garcia-Aloy, Francisco J Madrid-Gambin, Francisco J Tinahones, Ricardo Gomez-Huelgas, Rikard Landberg, Cristina Andres-Lacueva. A lifestyle intervention for weight loss impacts on plasma metabolome of metabolically healthy obese women towards a healthier phenotype. [*Under revision*].

Participación del doctorando:

• Preparación y análisis de las muestras biológicas por RMN

• Procesado y tratamiento de los datos y aplicación de herramientas bioestadísticas para corroborar las hipótesis de partida

• Identificación y confirmación de los metabolitos discriminantes

• Interpretación biológica

• Elaboración del primer borrador de la publicación

• Seguimiento y trabajo del manuscrito hasta su envío a la revista

Plasma metabolomic profiling reveals metabolic benefits in healthy obese women after lifestyle intervention with Mediterranean diet and physical exercise

Enrique Almanza-Aguilera[1,2#], Carl Brunius[3,4#], M Rosa Bernal-Lopez[5,6]*, Mar Garcia-Aloy[1,2], Francisco J Madrid-Gambin[1,2], Francisco J Tinahones[6,7], Ricardo Gomez-Huelgas[5,6], Rikard Landberg[2,3], Cristina Andres-Lacueva[1,6]*

[1]Biomarkers and Nutrimetabolomics Laboratory, Department of Nutrition, Food Sciences and Gastronomy, Food Technology Reference Net (XaRTA), Nutrition and Food Safety Research Institute (INSA), Campus Torribera, Faculty of Pharmacy and Food Sciences, University of Barcelona, Barcelona, Spain. [2]CIBER Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Barcelona 08028, Spain. [3]Department of Food Science, Swedish University of Agricultural Sciences, Uppsala, Sweden. [4]Food and Nutrition Division, Department of Biology and Biological Engineering, Chalmers University of Technology, Göteborg, Sweden. [5]Internal Medicine Department, Biomedical Institute of Malaga (IBIMA), Regional University Hospital of Malaga (Carlos Haya Hospital), Spain. [6]Ciber Fisiopatología de la Obesidad y Nutrición, Instituto de Salud Carlos III, Madrid, Spain. [7]Endocrinology and Nutrition Department, Biomedical Institute of Malaga (IBIMA), Regional University Hospital of Malaga (Virgen de la Victoria Hospital), Spain.

#Equal contribution

*Correspondence to: **Dr. Cristina Andres-Lacueva**. Biomarkers and Nutrimetabolomics Laboratory, Department of Nutrition, Food Sciences and Gastronomy, Faculty of Pharmacy and Food Sciences, University of Barcelona, Av. Joan XXIII, 27-31, 08028, Barcelona, Spain. Phone +34.934034840; Fax: +34.93403593.

E-mail address: candres@ub.edu

*Dr. M Rosa Bernal Lopez. Internal Medicine Department, Institute of Biomedical Research in Málaga (IBIMA), Regional Hospital of Málaga, Avda. Hospital Civil s/n, 29009 Malaga, Spain.

Phone.: +34 951 29 03 46; Fax: +34 951 29 03 02. E-mail address: robelopajiju@yahoo.es

Running head: Weight loss intervention impacts on plasma metabolome

Word count abstract: 448

Word count in text: 4010

Figures: 5 (3 in main manuscript and 2 in supplemental material)

Tables: 3 (2 in main manuscript and 1 in supplemental material)

Keywords: metabolomics, NMR, hypocaloric diet, physical activity, metabolically healthy obesity

Abbreviations: 2-HB, 2-hydroxybutyrate; 3-HB, 3-hydroxybutyrate; BMI, body mass index; CK, creatine kinase; CVD, cardiovascular disease; DBP, diastolic blood pressure; FDR, false discovery rate; HbA1c, haemoglobin A1c; HDL, high-density lipoprotein cholesterol; ^{1}H-NMR, proton nuclear magnetic resonance; HOMA-IR, homeostasis model assessment of insulin resistance; LDL, low-density lipoprotein cholesterol; LWL, lifestyle weight loss; MHO, metabolically healthy obese; MedDiet, Mediterranean diet; MUO, metabolically unhealthy obese; PLS, partial least squares regression; rdCV, repeated double cross-validation; RF, random forest; SBP, systolic blood pressure; SFA, saturated fatty acids; T2D, type 2 diabetes; TMA, trimethylamine; VLDL, very low-density lipoprotein cholesterol; WC, waist circumference; WL, weight loss.

ABSTRACT

BACKGROUND: Metabolically healthy obese (MHO) individuals have lower cardiovascular risk than their unhealthy counterpart. However, little is known regarding metabolic benefits of a lifestyle weight-loss (LWL) intervention on the MHO phenotype.

OBJECTIVE: We aimed to examine impact of an LWL treatment on the plasma metabolomic profiling in MHO individuals, and further to explore its association with changes in cardiometabolic risk markers.

METHODS: Plasma samples from 57 MHO women (35–55 y, BMI $\geq$ 30) allocated either to an intensive LWL treatment group (hypocaloric Mediterranean diet and regular physical activity, n = 30) or to a control group (general recommendations of a healthy diet and physical activity, n = 27) were collected and analysed using untargeted ^{1}H-NMR metabolomics at baseline, and after 3 and 12 months of the intervention. In addition, changes in anthropometric measurements, and glycaemic and lipid profiles, were analysed at the same time points. The effects of the intervention on plasma metabolome were explored using multivariate modelling and bivariate correlations.

RESULTS: Compared to the control group, the treatment group experienced greater weight loss (WL, kg) at both 3 (9.10 ± 3.50 *vs* 3.04 ± 3.62, $p < 0.001$) and 12 months (9.98 ± 6.79 *vs* 4.62 ± 7.22, $p < 0.01$). Multivariate modelling of NMR data for the intervention (control or treatment) showed a high rate of correct classification at 3 months (86%, permutation test $p < 0.001$) but not at 12 months (65%). WL, but none of the clinical parameters, was moderately associated with the changes in metabolome at 3 months ($Q^2 = 0.257$, $p < 0.001$). Higher levels of formate and phosphocreatine, and lower levels of LDL/VLDL (signals) and trimethylamine, were found in the treatment group than in the control group after the intervention period and were associated with WL. In addition, higher myo-inositol, methylguanidine and 3-hydroxybutyrate, and lower proline, were found in the treatment group after 3 months of the intervention. Moreover, higher levels of hippurate and asparagine, and lower levels of 2-hydroxybutyrate (2-HB) and creatine, were found to be associated with WL.

Conclusions: A 3-month intensive LWL treatment based on a hypocaloric Mediterranean diet and regular physical activity resulted in greater WL than an LWL treatment based on general recommendations of a healthy lifestyle and concurrently modulated energy, amino acid, lipoprotein and microbial metabolism in MHO women. An inverse association between WL and an early biomarker of impaired glucose regulation, 2-HB, is reported, and supports the modulatory effect of WL on the diabetes risk of the MHO phenotype towards healthier phenotype. Although WL was still affected after 12 months of the intervention, the observed

metabolic regulation was attenuated. The study of metabolome changes and adaptation to long-term studies on WL is warranted.

INTRODUCTION

Obesity is a major worldwide health-care concern because it is one of the main determinants of non-communicable diseases such as type 2 diabetes (T2D) and cardiovascular disease (CVD) [1,2]. However, recent studies have shown that even within the obese phenotype, cardiometabolic risk may not necessarily vary primarily in relation to weight or body mass index (BMI), but to other subclinical alterations [3]. In this sense, there is a subset of obese individuals with lower risk of CVD or all-cause mortality [4,5], which has been referred to as the "obese healthy paradox" [6]. Although there is currently a lack of consensus on its definition, it has been suggested that metabolically healthy obese (MHO) individuals have ≤ 1 of the metabolic syndrome (MetS) criteria, in addition to waist circumference (WC) $\geq$ 102 cm, i.e. hypertension, hypertriglyceridaemia, hyperglycaemia or diagnosis of diabetes, and dyslipidaemia in comparison to their metabolically unhealthy obese (MUO) counterparts [7]. In addition, other parameters such as insulin sensitivity, and inflammatory markers, such as tumour necrosis factor alpha (TNF-α), have been suggested for inclusion in the definition of MHO [6,8]. The prevalence of MHO varies widely, partly depending on definition criteria, but seems to be higher in women and in people of younger ages and also dependent on region and lifestyle [9,10]. On the other hand, long-term studies have suggested that MHO is a transient state towards MUO [11,12], indicating that without proper care, MHO individuals may increase their risk of developing T2D and CVD. Furthermore, since a decisive feature of MHO is the absence of visceral fat accumulation [13], the promotion of lifestyle interventions aimed at minimizing visceral fat accumulation is of fundamental importance from a public health perspective. Although diet and physical activity are well-known and modifiable CVD risk factors, their potential beneficial impact on MHO under controlled conditions has only recently received attention [14,15].

Recent studies have demonstrated that intensive lifestyle weight loss (LWL) interventions based on calorie restriction and physical activity were effective as a means of improving body composition and several cardiometabolic risk markers in obese individuals [16,17]. In this sense, for instance, the adherence to healthy dietary patterns such as the Mediterranean diet (MedDiet), with or without energy restriction, has been widely recommended for reducing the incidence and lowering the prevalence of MetS and its components [18]. Currently, a better understanding of how an LWL intervention impacts on the metabolism of obese individuals has barely been addressed by the use of metabolomics. For example, in a recent study, Khakimov and colleagues (2016) reported that as result of an LWL intervention study with a healthy New Nordic Diet, participants with greater weight loss differed in their plasma metabolite composition, including metabolites related to energy metabolism and food intake [19].

In the current study, we therefore aimed to: i) determine the impact of an intensive lifestyle treatment for weight loss (based on a hypocaloric Mediterranean diet and regular physical activity) in comparison to a control group (general recommendations for a cardiometabolic healthy diet and physical activity) on the plasma metabolome, measured by untargeted ^{1}H-NMR metabolomics, in women defined as MHO; and to ii) investigate the associations of WL and changes in other cardiometabolic risk markers with changes in metabolome after following the intervention.

SUBJECTS AND METHODS

Participants and study design

A total of 115 women aged 35–55 defined as MHO were recruited from four health centres in the Malaga District of the Andalusian Health Service (Spain) [20]. Diagnosis of metabolic health status was based on the general criteria proposed by the International Diabetes Federation (IDF) [21]. Besides obesity (BMI ≥ 30 kg/m^2), participants were included if they had ≤ 1 of the following cardiovascular risk factors: elevated fasting glucose levels (plasma glucose ≥ 100 mg/dL), elevated blood pressure (systolic ≥ 135 and/or diastolic ≥ 85 mmHg or use of antihypertensive drugs), elevated triglycerides (≥ 150 mg/dL or treatment with lipid-lowering medication), or decreased high-density-lipoprotein cholesterol (HDL) (< 50 mg/dL). The exclusion criteria were previous diagnosis of diabetes, pregnancy or planning to become pregnant during the study, cardiovascular disease, presence of any severe chronic illness, alcohol or drug abuse, or undertaking a WL program that included physical activity and/or diet in the past three months.

Participants were randomly allocated to either the control (n = 48) or the treatment group (n = 67). Participants assigned to the control group received general recommendations on a cardiometabolic healthy diet and physical activity based on national dietary guidelines. Participants in the treatment group received an intensive intervention program for losing weight, consisting of a hypocaloric MedDiet and regular physical activity. The MedDiet included the intake of extra virgin olive oil and nuts but with an overall energy restriction of about 600 kcal (approximately 30% of estimated energy requirements). The distribution of the target daily total caloric intake for the intervention group was: 35–40 % fats (8–10 % saturated fatty acids; SFA), 40–45 % carbohydrates with low glycaemic index and 20% protein. In the physical activity program, participants were encouraged to practise a minimum of 150 min/wk of walking. Participants allocated to the treatment group attended visits with a certified nutritionist every week during the first 3 months, and then once at 12 months, whereas individuals in the control group attended these visits only after 3 and 12 months of the study. All participants provided

written informed consent. The study was conducted in accordance with the guidelines set out in the Declaration of Helsinki and all protocols were approved by an Ethics and Research Committee (Comité Coordinador de Ética de la Investigación Biomédica de Andalucía). This study was registered at https://www.isrctn.com/ as ID ISRCTN88315555.

Clinical measurements and sampling

Anthropometric measurements, including weight, height, waist circumference (WC) and BMI, were taken by trained nurses at baseline, and after 3 and 12 months, and blood pressure measurements were taken at baseline and at 12 months. Fasting blood samples were collected in tubes containing EDTA on the day of enrolment and after 3 and 12 months. Analyses of fasting glucose and lipid profile were conducted according to routine methods and within 12 hours of sample collection. For metabolomics analysis, plasma samples were collected at the same time points, aliquoted and immediately stored at -80 °C until analysis.

NMR metabolomics

Sample preparation

Plasma samples were thawed at 4 °C, briefly spun down and 150 µl were mixed with 150 µl ultrapure water and 600 µl pure cold (-20 °C) methanol in 96-deep-well plates. The mixtures were vortexed, incubated (first at 12 °C, 800 rpm for 10 min and then at -20 °C for 30 min) and centrifuged (2250 *g* at 4 °C for 60 min) to precipitate proteins. Supernatants (600 µl) were transferred into clean deep-well plates and lyophilized at -4 °C for 16 h. Dried samples were washed in 50 µl of deuterated methanol (MeOD) and again lyophilized to remove the excess of non-MeOD. The new pellets were reconstituted in 200 µl buffer (37.5 mM sodium phosphate, pH 6.95, 100% D_20, 0.02% NaN_3, 0.25 mM DSS-d_6 and 1 mM imidazole) and shaken in an Eppendorf ThermoMixer at 800 rpm, 22 °C, for 30 min. Samples in buffer (180 µl) were transferred into 3 mm SampleJet NMR tubes using a SamplePro L liquid handling robot (Bruker BioSpin, Rheinstetten, Germany).

^{1}H-NMR spectroscopy

All ^{1}H-NMR experiments were performed on an Oxford 800 MHz magnet equipped with a Bruker Avance III HD console and a 3 mm TCI cryoprobe, using a water suppression pulse program. Each spectrum was acquired at 298 K applying 128 scans, a spectral width of 20 ppm, a data size of 65 K points, an acquisition time of 2.05 s and a relaxation delay of 3 s. Spectra were processed using TopSpin 3.5pI6 (Bruker GmbH, Rheinstetten, Germany). Processed

spectral data were imported into MatLab (MathWorks Inc., Natick, MA) using in-house written scripts. Alignment was achieved using a combination of an in-house peak reference picking function and the "speaq" R-package (version 1.2.1) [22].

Statistical analyses

All statistical data analyses were performed within the R environment (version 3.3.1). Differences in anthropometric and clinical variables at baseline and after 3 and 12 months were assessed by independent or paired Student's t-tests, according to comparisons between or within groups, respectively. A false discovery rate (FDR) procedure was applied to adjust the p-value for multiple comparisons in all univariate analyses. Data are expressed as mean ± SD, unless otherwise stated.

In order to determine discriminant metabolites between control and treatment groups at 3 and 12 months of study we used NMR data of differences in metabolome between baseline and each time point (3 or 12 months) and conducted a supervised analysis based on Random Forest (RF) modelling within an in-house-developed repeated double cross-validation framework (rdCV) [23,24]. In brief, the in-house double CV procedure, which has been successfully used in untargeted metabolomics [25] and microbiota analysis [26], consists of nested loops (outer "testing" and inner "calibration" loops) to reduce bias from overfitting models to experimental data [23]. Feature ranking and selection is performed within the inner loop, to minimize statistical overfitting, by iteratively turning over successively fewer features, removing from each step in the loop the 10% least informative features [25]. The rdCV procedure was subjected to 30 repetitions to improve modelling accuracy and with misclassification as the fitness function. The overall validity and degree of overfitting of models were assessed by permutation analysis, reporting the cumulative probability of actual model fitness within a population of fitness measures of randomly permuted classifications (n = 200) based on the assumption of Student's t-distribution. The assumption was confirmed by visual inspection of the histograms of permuted distributions.

Secondary analyses of associations of changes in metabolome with changes in weight or other clinical parameters were performed in both the control and treatment groups, as well as in treatment group alone, by partial least squares (PLS) regression within a similar rdCV framework. Permutation tests (n = 200) were performed similarly to the analysis above, but with Q^2 as the fitness measure.

Differences in changes of metabolites between groups after the intervention were calculated by fold change, taking the control group as reference, and assessed by independent Student's t-

tests. Correlations between significant metabolites selected from multivariate modelling of weight change after the intervention were calculated by Spearman's rank correlation ("Hmisc" R-package version 4.0-2).

Metabolite identification

Identification of metabolites was achieved by matching experimental NMR spectra with those stored in Chenomx NMR Suite 8.2 software (Chenomx Inc., Canada) in combination with an in-house R script for statistical correlation spectroscopy [27] and through searching in the Human Metabolome Database (HMDB) compound reference library [28].

RESULTS

Baseline characteristics of the participants

Of the 115 participants recruited, 58 were excluded due to either loss of compliance, the occurrence of illness or an unavailable sample at some time point (at baseline, 3 or 12 months). Therefore, 57 participants were included in the present data analyses. Anthropometric measures and clinical parameters of participants at baseline and after 3 and 12 months, subdivided into control and treatment groups, are presented in Supplementary Table 1. At baseline, MHO participants had a mean (± SD) age of 45.1 ± 3.45 y, and a BMI of 35.8 ± 4.93 kg/m^2. No differences between the control and treatment groups were observed at baseline regarding weight, waist circumference, blood pressure, glycaemia or lipid profile (Supplementary Table 1).

Changes in anthropometric and clinical measurements

At both 3 and 12 months, the treatment group showed greater WL and more pronounced reductions in BMI and WC than in the control group (Supplementary Table 1). Compared to baseline, both groups showed a decrease in total cholesterol and changes in HDL at 3 and 12 months. In particular, at 3 months the levels of HDL were found to be decreased in the treatment group and increased in the control group. Moreover, at 3 months only participants in the treatment group showed decreases in LDL cholesterol and at 12 months decreases in SBP, glucose and triglycerides, whereas at 12 months only the control group showed decreases in LDL (Supplementary Table 1).

Multivariate modelling of intervention and weight change

The classification of participants as treatment or control based on the changes in metabolome is shown in Figure 1. The rdCV-RF models resulted in a high classification rate (86%, permutation test $p < 0.001$) of the individuals at 3 months and low classification rate (65%) at 12 months (Figure 1; Supplemental Figure 1).

With the exception of weight change, changes in other clinical parameters were not significantly associated with changes in the metabolome (data not shown). The rdCV-PLS regression of weight change based on changes in metabolome resulted in moderate associations when all participants ($Q^2 = 0.257$, permutation test $p < 0.001$), or only participants in the treatment group ($Q^2 = 0.298$, permutation test $p < 0.001$), were included in the analysis (Figure 2; Supplemental Figure 2).

Modulatory effect of intervention on plasmatic metabolites

Changes in metabolome after 3 months of intervention included higher levels in the treatment group of 3-hydroxybutyrate (3-HB), formate, methylguanidine, myo-inositol and phosphocreatine, as well as lower levels of LDL/VLDL signals, proline, trimethylamine (TMA) and three unassigned compounds (U3.32, U4.35 and U6.40) (Table 1). In accordance with these results, absolute fold changes in 3-HB, methylguanidine, phosphocreatine, myo-inositol, proline, U4.35 and U6.40 were ≥ 2 (two times or more) higher in the treatment group than in the control group. Due to the poor multivariate classification between groups at 12 months, discriminant metabolites at 3 months were further investigated by univariate analysis t-test after 12 months of intervention (Table 1). From this analysis, differences between the treatment and control groups at 12 months were only observed in U3.32 ($p < 0.05$) and phosphocreatine ($p < 0.05$). However, compared to the situation at 3 months, fold changes in these metabolites at 12 months indicated a more accentuated change in U3.32 and a change from upregulation to downregulation in phosphocreatine.

Changes in metabolome associated with weight loss

A total of 11 metabolites were moderately associated with a change in weight from baseline in both groups at 3 months (Table 2). Consequently, an association of metabolites with WL was established as an inverse association with weight change (i.e. a positive association with weight change means an inverse association with WL). Therefore, WL was inversely associated with 2-hydroxybutyrate (2-HB), creatine, LDL/VLDL signals, TMA and three unknown compounds (U.sugar, U2.96 and U3.32) and directly associated with asparagine, formate, hippurate and phosphocreatine. Interestingly, from this model, the changes in formate, phosphocreatine, LDL/VLDL signals, TMA and U3.32 were found to be in the same direction as those observed in the previous model with treatment (Figure 3).

DISCUSSION

Using untargeted ^{1}H-NMR-based metabolomics and multivariate modelling we were able to determine changes in the plasma metabolome associated with an LWL treatment, based on a hypocaloric diet and physical activity in MHO women. Furthermore, within this context, we further investigated the association of WL with changes in the metabolome. As expected, compared to the control group, individuals in the treatment group underwent greater WL.

Differences in the plasma metabolome between the treatment and control groups were more pronounced at 3 than at 12 months (Figure 1; Supplemental Figure 1). One reason could be the similar weight loss achieved during the period between the 3rd and 12th months after beginning the intervention (Supplementary Table 1) and/or compensatory mechanisms that attenuated the effects at 12 months. Among cardiometabolic risk markers, only weight change was associated with the changes in metabolome after 3 months of the study (Figure 2; Supplemental Figure 2).

We found distinct and common metabolites associated at the same with the intervention and WL (Figure 3), which together reflect a positive impact of an intensive LWL intervention as well as of WL on the metabolism of energy, amino acids, lipoproteins and microbiota. For instance, the higher 3-HB observed in the treatment group than in the control group is consistent with previous studies on weight loss [29,30]. High circulating levels of ketone bodies are observed under energy-restricted metabolic states caused by fasting and caloric restriction, through increased lipolysis of fatty acids in liver mitochondria [31]. Therefore, the increase of 3-HB in treatment may reflect energy homeostasis through increased lipid oxidation [32].

Interestingly, 2-HB, a well-known early biomarker of impaired glucose regulation in non-T2D individuals [33], was found to decrease with WL. We therefore speculate that as a result of WL, the MHO individuals decreased their risk of developing T2D [34].

The association of proline and asparagine with treatment and WL, respectively, reflects impacts on amino acid metabolism. The lower levels of proline in the treatment group than in the control group are in line with previously reported data, indicating that both caloric restriction [30] and increased physical activity [35] result in lower circulating proline in obese individuals. Being a glucogenic amino acid, proline may be decreased in plasma due to an increase of hepatic gluconeogenesis, which is characteristic of caloric restriction [36] and physical exercise [37]. Moreover, previous studies have also shown an association between long-term successful WL and lower plasma proline levels [38]. However, this was not supported in the present study, since proline was not directly associated with WL. The positive association between asparagine and WL found in our study is consistent with previous reports, which have shown an inverse association between this amino acid and obesity [39,40]. Circulating levels of asparagine can be obtained from dietary sources or synthesized from endogenous oxaloacetate via aspartate.

Studies conducted in animal models have demonstrated that supplementation with aspartate and asparagine increased the glucose uptake and glycogen content in skeletal muscle, possibly through the incorporation of glucose transporters type 4 or vesicles into the glycogen complex [41]. We therefore speculate that along with weight loss, an increase of asparagine may be associated, in part, with an improved glucose homeostasis. However, future studies are warranted to better determine the functional role of asparagine in WL. Taken together, the observed associations of 3-HB, 2-HB, proline and asparagine with the current LWL intervention and WL strongly suggest a positive impact on glucose homeostasis in the MHO phenotype, which could also be traduced into a decrease of their cardiovascular risk.

Also related to amino acid metabolism, we found that phosphocreatine increased with both treatment and WL, whereas creatine decreased with WL. Creatine is mainly produced in the liver and skeletal muscle from glycine and arginine, and can further be phosphorylated to form phosphocreatine by the enzymatic action of creatine kinase (CK) [42]. We therefore speculate that the contrasting association of creatine and phosphocreatine with WL may be related to a modulatory effect of WL on CK activity. Recent studies showing significant associations of CK with obesity [43,44] and weight loss [45] may support this hypothesis.

With regard to lipoprotein metabolism, we found lower intensity of signals corresponding to LDL/VLDL in the treatment than in the control group and this was also inversely associated with WL, which is consistent with previous studies [46] and probably related to increased expression of LDL receptor and lipoprotein lipase [47].

The contribution of four microbial metabolites, i.e. formate, hippurate, methylguanidine and TMA, to either treatment or weight change models strongly supports previous reports that highlight the role of host-microbiota interactions in body weight composition and WL; this latter promoted by either a LWL intervention based on diet [48] or bariatric surgery [49]. For example, TMA, an intermediate metabolite from the microbial metabolism of dietary carnitine and choline, was found to be lower after treatment and also with WL. TMA is oxidized by hepatic flavin-containing monooxygenases to form trimethylamine-N-oxide (TMAO), which has been shown to be both proatherogenic and associated with cardiovascular disease risk [50,51]. We hypothesize that the lower levels of TMA after treatment and inversely associated to WL are related to either a lower intake of its dietary precursors (i.e. eggs and meat) [51,52] or modulation of choline and carnitine metabolism, and consequently point to a lower risk of CVD due to a likely reduced synthesis of TMAO.

Some other unidentified metabolites, including unassigned signals corresponding to sugars, were found to be related to treatment and WL (Tables 2–3). Of particular interest is the unknown U3.32, which was not only found to be related to treatment and WL at 3 months, but

also remained significant in the treatment group at 12-month follow-up. Further research to identify this compound to understand its role in WL in the short and long term is needed.

The high number of subjects misclassified at 12 months suggests a larger similarity in the changes in metabolome between groups, presumably due to either loss of compliance or adaptation to changes after the first 3 months in the treatment group. Several factors, including physiological, behavioural and environmental ones, are key to both compliance and dropout in long-term programmes for WL [53]. It is well documented that although lifestyle interventions can be effective for long-term WL and improving cardiometabolic markers, maximum WL is normally achieved in between 1 and 6 months, followed by variable weight maintenance or weight regain [54]. However, in the present study, no differences in WL could be observed within the groups from 3 to 12 months, thereby suggesting a maintenance phase. Therefore, based on the poor multivariate predictions at 12 months combined with the maintained WL between 3 and 12 months, we hypothesize that a metabolic adaptation occurs during this maintenance stage. To the best of our knowledge, however, there are no reports on this type of metabolic/metabolome adaptation as a result of longer-term WL interventions. Furthermore, although participants were defined a priori as belonging to MHO, the combined results at 3 months, in terms of changes in clinical parameters and metabolome, indicate that the current weight loss intervention caused shifts towards a healthier phenotype or reduced the risk of CVD. However, the positive metabolic regulations appeared to be attenuated in the longer term, even though weight loss was maintained. The reasons for this attenuation remain unclear.

We recognize that our study has a number of limitations and strengths. For instance, the sample size is relatively small and the study participants were exclusively Caucasian, women and middle-aged. Thus, we cannot extrapolate our conclusions to the general population. Moreover, the large and maintained WL in the treatment group implies a generally high compliance with the intervention. Because we were not able to assign the identity of unknown compounds, potentially important information about metabolic perturbations in relation to intervention and WL was unavailable. Future research aimed at identifying these unknown compounds is warranted. As was also pointed out above, our study also had several strengths. The untargeted workflow employed peak picking instead of binning, thereby expanding and improving the available information content in the original data. The multivariate modelling procedure and validation framework employed a data-driven, robust approach to maximize information density while minimizing the likelihood of false-positive findings, thereby focusing automatically on the most relevant metabolic perturbations in relation to the WL intervention [25]. Finally, our findings reinforce the utility of metabolomics in the identification of biomarkers (beyond clinical parameters) of LWL interventions in individuals with moderate risk of CVD. These biomarkers could be used in future research as additional targets of LWL interventions.

In conclusion, using untargeted ^{1}H-NMR metabolomics and multivariate modelling, we determined the impact on plasma metabolome caused by lifestyle intervention for weight loss, based on diet and physical activity, in MHO women. Furthermore, we found that changes in the metabolome were associated with weight loss within the frame of the same intervention. Taken together, a hypocaloric Mediterranean diet and regular physical activity intervention modulated energy, amino acid, lipoprotein and microbial metabolism in MHO women towards a healthier phenotype. Such regulations were, however, only observed at 3 months. Although weight loss was maintained at 12 months, the metabolic changes driven by intervention were substantially attenuated at 12 months, suggesting metabolic adaptation. The inverse association between WL and 2-HB, in conjunction with the changes in other energy-related metabolites (3-HB, proline and asparagine), could be interpreted as a decreased risk of T2D as an effect of the current LWL treatment. Future research on metabolomic changes and adaptation in long-term studies is warranted.

CONFLICT OF INTEREST: All authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by grants from the Instituto de Salud Carlos III, cofinancial by the Fondo Europeo de Desarrollo Regional-FEDER (FIS PI14/00696, PI12/01373 and “Centros de Investigación En Red” (CIBERFES and CIBEROBN). The Spanish Ministry of the Economy and Competitiveness (MINECO) together with the Joint Programming Initiative “A Healthy Diet for a Healthy Life” (JPI HDHL, website: http://www.healthydietforhealthylife.eu) [grant number FOODBALL-PCIN-2014-133]; CIBERFES (co-funded by the FEDER Program from the EU); and the award of the Generalitat de Catalunya’s Agency AGAUR [grant number 2014SGR1566]. EAA would like to thank CONACYT (Mexico) for the PhD fellowship. M Rosa Bernal-Lopez was supported by “Miguel Servet Type I” program (CP15/00028) from the ISCIII-Madrid (Spain), cofinancial by the Fondo Europeo de Desarrollo Regional-FEDER.

Authors’ contribution

MRBL, RGH, FJT, and CAL planned and designed the study. MRBL, FJT and RGH provided the clinical samples of study. EAA and FJM performed metabolomics analyses. EAA performed statistical analyses and drafted the initial manuscript. MGA, CB supervised the statistical analyses and drafting of the manuscript. MGA, CB, RL and CAL interpreted data and revised manuscript for important intellectual content. CAL and MRBL have the primary responsibility for the final content. All authors contributed to the discussion of the results and to the critical revision of the manuscript.

REFERENCES

[1] Lee CMY, Huxley RR, Wildman RP, Woodward M. Indices of abdominal obesity are better discriminators of cardiovascular risk factors than BMI: a meta-analysis. J Clin Epidemiol 2008;61:646–53. doi:10.1016/j.jclinepi.2007.08.012.

[2] Lavie CJ, Milani R V, Ventura HO. Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. J Am Coll Cardiol 2009;53:1925–32. doi:10.1016/j.jacc.2008.12.068.

[3] Sims EA. Are there persons who are obese, but metabolically healthy? Metabolism 2001;50:1499–504. doi:10.1053/meta.2001.27213.

[4] Calori G, Lattuada G, Piemonti L, Garancini MP, Ragogna F, Villa M, et al. Prevalence, metabolic features, and prognosis of metabolically healthy obese Italian individuals: the Cremona Study. Diabetes Care 2011;34:210–5. doi:10.2337/dc10-0665.

[5] Hamer M, Stamatakis E. Metabolically healthy obesity and risk of all-cause and cardiovascular disease mortality. J Clin Endocrinol Metab 2012;97:2482–8. doi:10.1210/jc.2011-3475.

[6] Barbarroja N, Lopez-Pedrera R, Mayas MD, Garcia-Fuentes E, Garrido-Sanchez L, Macias-Gonzalez M, et al. The obese healthy paradox: is inflammation the answer? Biochem J 2010;430:141–9. doi:10.1042/BJ20100285.

[7] Alberti KGMM, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International . Circulation 2009;120:1640–5. doi:10.1161/CIRCULATIONAHA.109.192644.

[8] Phillips CM, Perry IJ. Does inflammation determine metabolic health status in obese and nonobese adults? J Clin Endocrinol Metab 2013;98:E1610-9. doi:10.1210/jc.2013-2038.

[9] Rey-Lopez JP, de Rezende LF, Pastor-Valero M, Tess BH. The prevalence of metabolically healthy obesity: a systematic review and critical evaluation of the definitions used. Obes Rev 2014;15:781–90. doi:10.1111/obr.12198.

[10] Wang B, Zhuang R, Luo X, Yin L, Pang C, Feng T, et al. Prevalence of Metabolically Healthy Obese and Metabolically Obese but Normal Weight in Adults Worldwide: A Meta-Analysis. Horm Metab Res 2015;47:839–45. doi:10.1055/s-0035-1559767.

[11] Hwang Y-C, Hayashi T, Fujimoto WY, Kahn SE, Leonetti DL, McNeely MJ, et al.

Visceral abdominal fat accumulation predicts the conversion of metabolically healthy obese subjects to an unhealthy phenotype. Int J Obes (Lond) 2015;39:1365–70. doi:10.1038/ijo.2015.75.

[12] Kramer CK, Zinman B, Retnakaran R. Are metabolically healthy overweight and obesity benign conditions?: A systematic review and meta-analysis. Ann Intern Med 2013;159:758–69. doi:10.7326/0003-4819-159-11-201312030-00008.

[13] Bluher M. The distinction of metabolically "healthy" from "unhealthy" obese individuals. Curr Opin Lipidol 2010;21:38–43. doi:10.1097/MOL.0b013e3283346ccc.

[14] Dalzill C, Nigam A, Juneau M, Guilbeault V, Latour E, Mauriege P, et al. Intensive lifestyle intervention improves cardiometabolic and exercise parameters in metabolically healthy obese and metabolically unhealthy obese individuals. Can J Cardiol 2014;30:434–40. doi:10.1016/j.cjca.2013.11.033.

[15] Arsenault BJ, Cote M, Cartier A, Lemieux I, Despres J-P, Ross R, et al. Effect of exercise training on cardiometabolic risk markers among sedentary, but metabolically healthy overweight or obese post-menopausal women with elevated blood pressure. Atherosclerosis 2009;207:530–3. doi:10.1016/j.atherosclerosis.2009.05.009.

[16] Petelin A, Bizjak M, Cernelic-Bizjak M, Jurdana M, Jakus T, Jenko-Praznikar Z. Low-grade inflammation in overweight and obese adults is affected by weight loss program. J Endocrinol Invest 2014;37:745–55. doi:10.1007/s40618-014-0102-9.

[17] Abbenhardt C, McTiernan A, Alfano CM, Wener MH, Campbell KL, Duggan C, et al. Effects of individual and combined dietary weight loss and exercise interventions in postmenopausal women on adiponectin and leptin levels. J Intern Med 2013;274:163–75. doi:10.1111/joim.12062.

[18] Perez-Martinez P, Mikhailidis DP, Athyros VG, Bullo M, Couture P, Covas MI, et al. Lifestyle recommendations for the prevention and management of metabolic syndrome: an international panel recommendation. Nutr Rev 2017;75:307–26. doi:10.1093/nutrit/nux014.

[19] Khakimov B, Poulsen SK, Savorani F, Acar E, Gurdeniz G, Larsen TM, et al. New Nordic Diet versus Average Danish Diet: A Randomized Controlled Trial Revealed Healthy Long-Term Effects of the New Nordic Diet by GC-MS Blood Plasma Metabolomics. J Proteome Res 2016;15:1939–54. doi:10.1021/acs.jproteome.6b00109.

[20] Rodriguez-Garcia E, Ruiz-Nava J, Santamaria-Fernandez S, Fernandez-Garcia JC, Vargas-Candela A, Yahyaoui R, et al. Characterization of lipid profile by nuclear

magnetic resonance spectroscopy (1H NMR) of metabolically healthy obese women after weight loss with Mediterranean diet and physical exercise. Medicine (Baltimore) 2017;96:e7040. doi:10.1097/MD.0000000000007040.

[21] Alberti KGMM, Zimmet P, Shaw J. The metabolic syndrome—a new worldwide definition. Lancet 2005;366:1059–62. doi:http://dx.doi.org/10.1016/S0140-6736(05)67402-8.

[22] Vu TN, Valkenborg D, Smets K, Verwaest KA, Dommisse R, Lemiere F, et al. An integrated workflow for robust alignment and simplified quantitative analysis of NMR spectrometry data. BMC Bioinformatics 2011;12:405. doi:10.1186/1471-2105-12-405.

[23] Filzmoser P, Liebmann B, Varmuza K. Repeated double cross validation. J Chemom 2009;23:160–71. doi:10.1002/cem.1225.

[24] Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, van Velzen EJJ, et al. Assessment of PLSDA cross validation. Metabolomics 2008;4:81–9. doi:10.1007/s11306-007-0099-6.

[25] Hanhineva K, Brunius C, Andersson A, Marklund M, Juvonen R, Keski-Rahkonen P, et al. Discovery of urinary biomarkers of whole grain rye intake in free-living subjects using nontargeted LC-MS metabolite profiling. Mol Nutr Food Res 2015;59:2315–25. doi:10.1002/mnfr.201500423.

[26] Buck M, Nilsson LKJ, Brunius C, Dabiré RK, Hopkins R, Terenius O. Bacterial associations reveal spatial population dynamics in Anopheles gambiae mosquitoes. Sci Rep 2016;6:22806.

[27] Cloarec O, Dumas M-E, Craig A, Barton RH, Trygg J, Hudson J, et al. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. Anal Chem 2005;77:1282–9. doi:10.1021/ac048630x.

[28] Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0--The Human Metabolome Database in 2013. Nucleic Acids Res 2013;41:D801-7. doi:10.1093/nar/gks1065.

[29] Brehm BJ, Spang SE, Lattin BL, Seeley RJ, Daniels SR, D'Alessio DA. The role of energy expenditure in the differential weight loss in obese women on low-fat and low-carbohydrate diets. J Clin Endocrinol Metab 2005;90:1475–82. doi:10.1210/jc.2004-1540.

[30] Gu Y, Zhao A, Huang F, Zhang Y, Liu J, Wang C, et al. Very low carbohydrate diet significantly alters the serum metabolic profiles in obese subjects. J Proteome Res 2013;12:5801–11. doi:10.1021/pr4008199.

[31] Cotter DG, Schugar RC, Crawford PA. Ketone body metabolism and cardiovascular disease. Am J Physiol Heart Circ Physiol 2013;304:H1060-76. doi:10.1152/ajpheart.00646.2012.

[32] Matoulek M, Svobodova S, Vetrovska R, Stranska Z, Svacina S. Post-exercise changes of beta hydroxybutyrate as a predictor of weight changes. Physiol Res 2014;63 Suppl 2:S321-5.

[33] Gall WE, Beebe K, Lawton KA, Adam K-P, Mitchell MW, Nakhle PJ, et al. alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. PLoS One 2010;5:e10883. doi:10.1371/journal.pone.0010883.

[34] Penn L, White M, Lindstrom J, den Boer AT, Blaak E, Eriksson JG, et al. Importance of weight loss maintenance and risk prediction in the prevention of type 2 diabetes: analysis of European Diabetes Prevention Study RCT. PLoS One 2013;8:e57143. doi:10.1371/journal.pone.0057143.

[35] Morris C, Grada CO, Ryan M, Roche HM, De Vito G, Gibney MJ, et al. The relationship between aerobic fitness level and metabolic profiles in healthy adults. Mol Nutr Food Res 2013;57:1246–54. doi:10.1002/mnfr.201200629.

[36] Browning JD, Weis B, Davis J, Satapati S, Merritt M, Malloy CR, et al. Alterations in hepatic glucose and energy metabolism as a result of calorie and carbohydrate restriction. Hepatology 2008;48:1487–96. doi:10.1002/hep.22504.

[37] Donovan CM, Sumida KD. Training enhanced hepatic gluconeogenesis: the importance for glucose homeostasis during exercise. Med Sci Sports Exerc 1997;29:628–34.

[38] Zheng Y, Ceglarek U, Huang T, Li L, Rood J, Ryan DH, et al. Weight-loss diets and 2-y changes in circulating amino acids in 2 randomized intervention trials. Am J Clin Nutr 2016;103:505–11. doi:10.3945/ajcn.115.117689.

[39] Cheng S, Rhee EP, Larson MG, Lewis GD, McCabe EL, Shen D, et al. Metabolite profiling identifies pathways associated with metabolic risk in humans. Circulation 2012;125:2222–31. doi:10.1161/CIRCULATIONAHA.111.067827.

[40] Takashina C, Tsujino I, Watanabe T, Sakaue S, Ikeda D, Yamada A, et al. Associations among the plasma amino acid profile, obesity, and glucose metabolism in Japanese

adults with normal glucose tolerance. Nutr Metab (Lond) 2016;13:5. doi:10.1186/s12986-015-0059-5.

[41] Lancha AHJ, Poortmans JR, Pereira LO. The effect of 5 days of aspartate and asparagine supplementation on glucose transport activity in rat muscle. Cell Biochem Funct 2009;27:552–7. doi:10.1002/cbf.1606.

[42] Wyss M, Kaddurah-Daouk R. Creatine and creatinine metabolism. Physiol Rev 2000;80:1107–213.

[43] Haan YC, Oudman I, Diemer FS, Karamat FA, van Valkengoed IG, van Montfrans GA, et al. Creatine kinase as a marker of obesity in a multi-ethnic population. Mol Cell Endocrinol 2017;442:24–31. doi:10.1016/j.mce.2016.11.022.

[44] George MD, McGill N-K, Baker JF. Creatine kinase in the U.S. population: Impact of demographics, comorbidities, and body composition on the normal range. Medicine (Baltimore) 2016;95:e4344. doi:10.1097/MD.0000000000004344.

[45] Van Weyenberg S, Hesta M, Buyse J, Janssens GPJ. The effect of weight loss by energy restriction on metabolic profile and glucose tolerance in ponies. J Anim Physiol Anim Nutr (Berl) 2008;92:538–45. doi:10.1111/j.1439-0396.2007.00744.x.

[46] Dattilo AM, Kris-Etherton PM. Effects of weight reduction on blood lipids and lipoproteins: a meta-analysis. Am J Clin Nutr 1992;56:320–8.

[47] Patalay M, Lofgren IE, Freake HC, Koo SI, Fernandez ML. The lowering of plasma lipids following a weight reduction program is related to increased expression of the LDL receptor and lipoprotein lipase. J Nutr 2005;135:735–9.

[48] Remely M, Tesar I, Hippe B, Gnauer S, Rust P, Haslberger AG. Gut microbiota composition correlates with changes in body fat content due to weight loss. Benef Microbes 2015;6:431–9. doi:10.3920/BM2014.0104.

[49] Liu R, Hong J, Xu X, Feng Q, Zhang D, Gu Y, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. Nat Med 2017. doi:10.1038/nm.4358.

[50] Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 2011;472:57–63. doi:10.1038/nature09922.

[51] Tang WHW, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med

2013;368:1575–84. doi:10.1056/NEJMoa1109400.

[52] Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. Nat Med 2013;19:576–85. doi:10.1038/nm.3145.

[53] Greenberg I, Stampfer MJ, Schwarzfuchs D, Shai I. Adherence and success in long-term weight loss diets: the dietary intervention randomized controlled trial (DIRECT). J Am Coll Nutr 2009;28:159–68.

[54] Lean ME. Is long-term weight loss possible? Br J Nutr 2000;83 Suppl 1:S103-11.

Table 1. Metabolites selected from multivariate modelling for optimum discrimination between treatment and control groups at 3 months of intervention, with follow-up at 12 months, sorted in order of decreasing variable importance

		3 months		12 months	
Chemical shift, ppm (multiplicity)	Compound	Fold change[1]	*P*[2]	Fold change[2]	*P*[2]
3.32 (s), 3.61 (s)	Unknown (U3.32)	0.62	< 0.001	0.20	0.01
3.03 (s), 3.94 (s)	Phosphocreatine	2.07	< 0.001	0.12	0.01
1.27 (m), 5.32 (m)	LDL/VLDL[3]	0.55	0.01	0.26	0.32
6.40 (m)	Unknown (U6.40)	0.14	< 0.01	0.10	0.11
2.90 (s)	Trimethylamine	0.53	< 0.01	0.38	0.28
1.20 (d), 2.30 (m), 2.40 (m), 4.16 (m)	3-hydroxybutyrate	2.21	0.04	0.66	0.69
2.82 (s)	Methylguanidine	3.17	0.10	0.98	0.97
4.35 (s)	Unknown (U4.35)	0.21	0.10	0.91	0.93
2.0 (m), 2.06 (m), 2.34 (m), 3.33 (m), 3.41 (m), 4.12 (dd)	Proline	0.38	0.01	0.41	0.14
3.28 (t), 3.52 (dd), 3.61 (t), 4.05 (t)	Myo-inositol	12.5	0.58	0.75	0.78
8.45 (s)	Formate	1.60	0.72	0.43	0.16

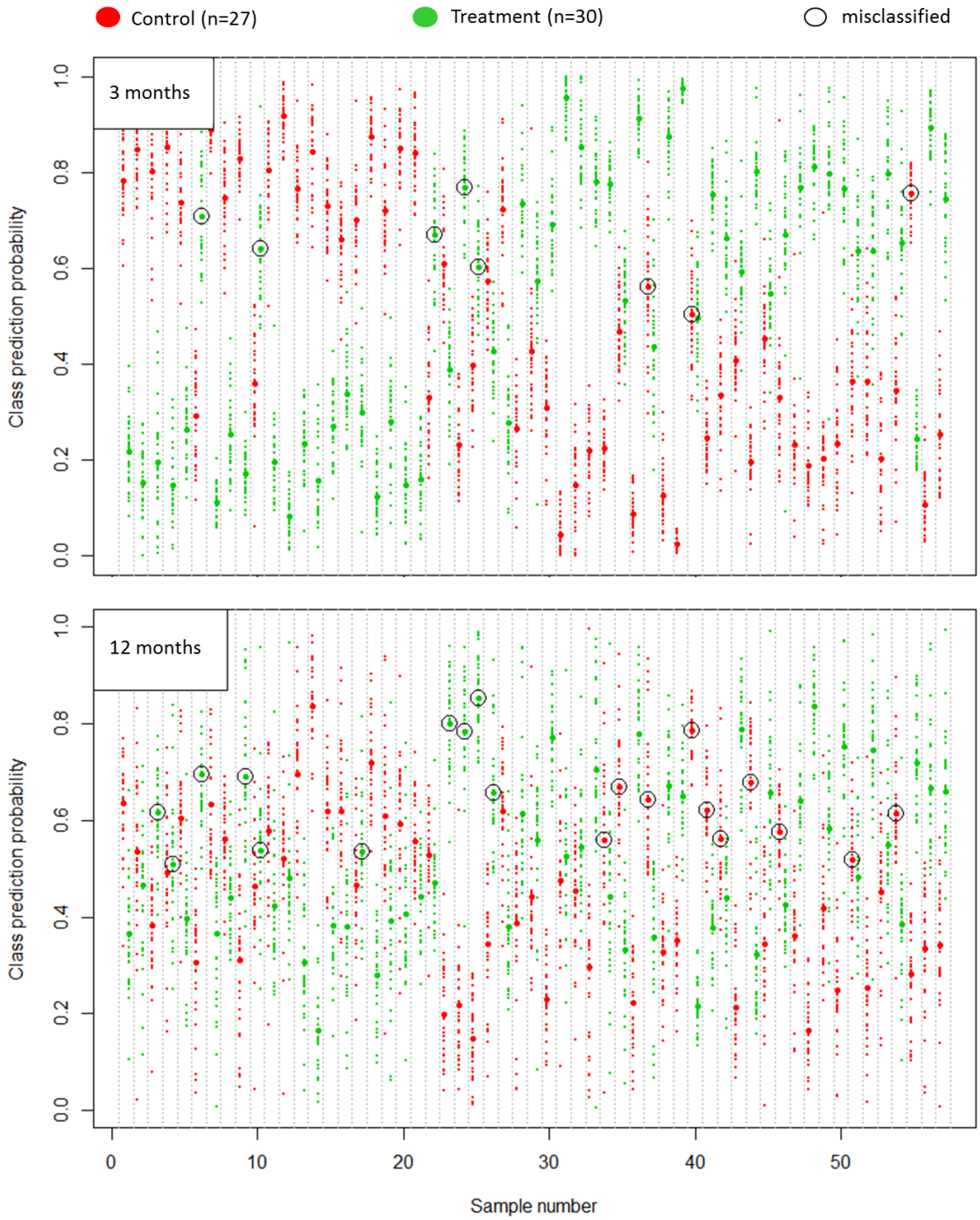
[1]Fold change of treatment group/control group mean values. [2]Student's t-test (2-tailed) between treatment and control groups (FDR-adjusted). [3]Methylene ($(CH_2)_n$) and olefinic (-CH=CH-) resonances at 1.27 ppm and 5.32 ppm, respectively. Abbreviations: s, singlet; d, doublet; m, multiplet; dd, double of doublets; t, triplet.

Table 2. Metabolites selected from multivariate modelling of weight change from baseline to 3 months in both control and treatment groups, sorted in order of decreasing variable importance

		Correlation	
Chemical shift, ppm (multiplicity)	Compound	Spearman	P^1
3.03 (s), 3.93 (s)	Phosphocreatine	-0.39	< 0.01
3.32 (s), 3.61 (t)	Unknown (U3.32)	0.42	< 0.001
2.85 (m), 2.94 (m), 4.0 (dd)	Asparagine	-0.27	0.04
8.44 (s)	Formate	-0.31	0.02
0.89 (m), 1.64 (m), 1.74 (m)	2-hydroxybutyrate	0.32	0.01
3.39 (m), 3.42 (m), 3.68 (m)	Unknown (U.sugar)	0.42	< 0.001
0.90 (m), 1.25 (m)	LDL/VLDL[2]	0.4	< 0.01
3.03 (s), 3.92 (s)	Creatine	0.32	0.01
2.90 (s)	Trimethylamine	0.28	0.04
2.96 (d)	Unknown (U2.96)	0.27	0.04
7.62 (t), 7.84 (d)	Hippurate	-0.29	0.03

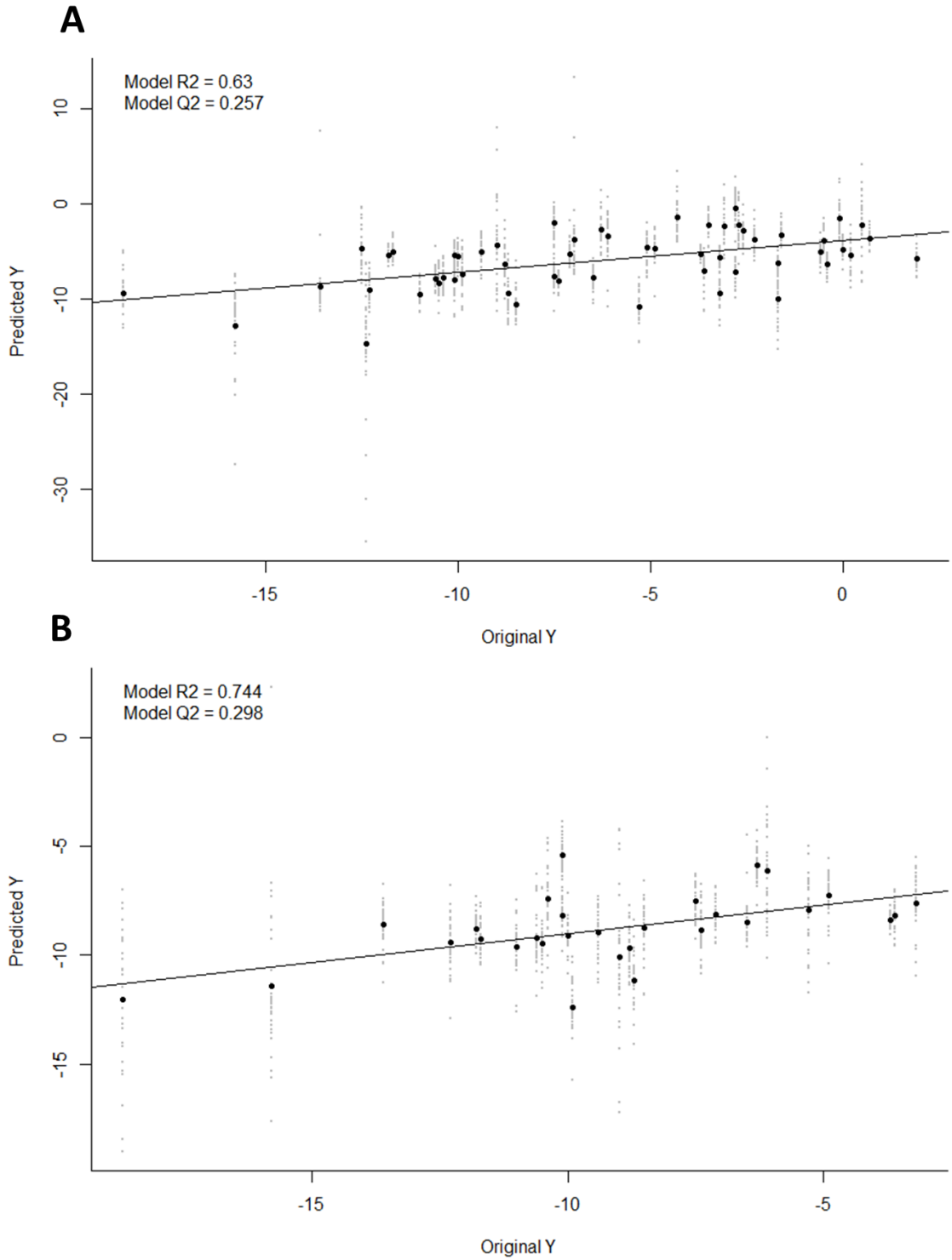
P^1, significance (FDR-adjusted) for the corresponding correlations. [2]Methyl (CH_3) and methylene ($(CH_2)_n$) resonances at 0.90 ppm and 1.25 ppm, respectively.

Figure 1. Predictive classification of individuals according to intervention group at 3 and 12 months[1].



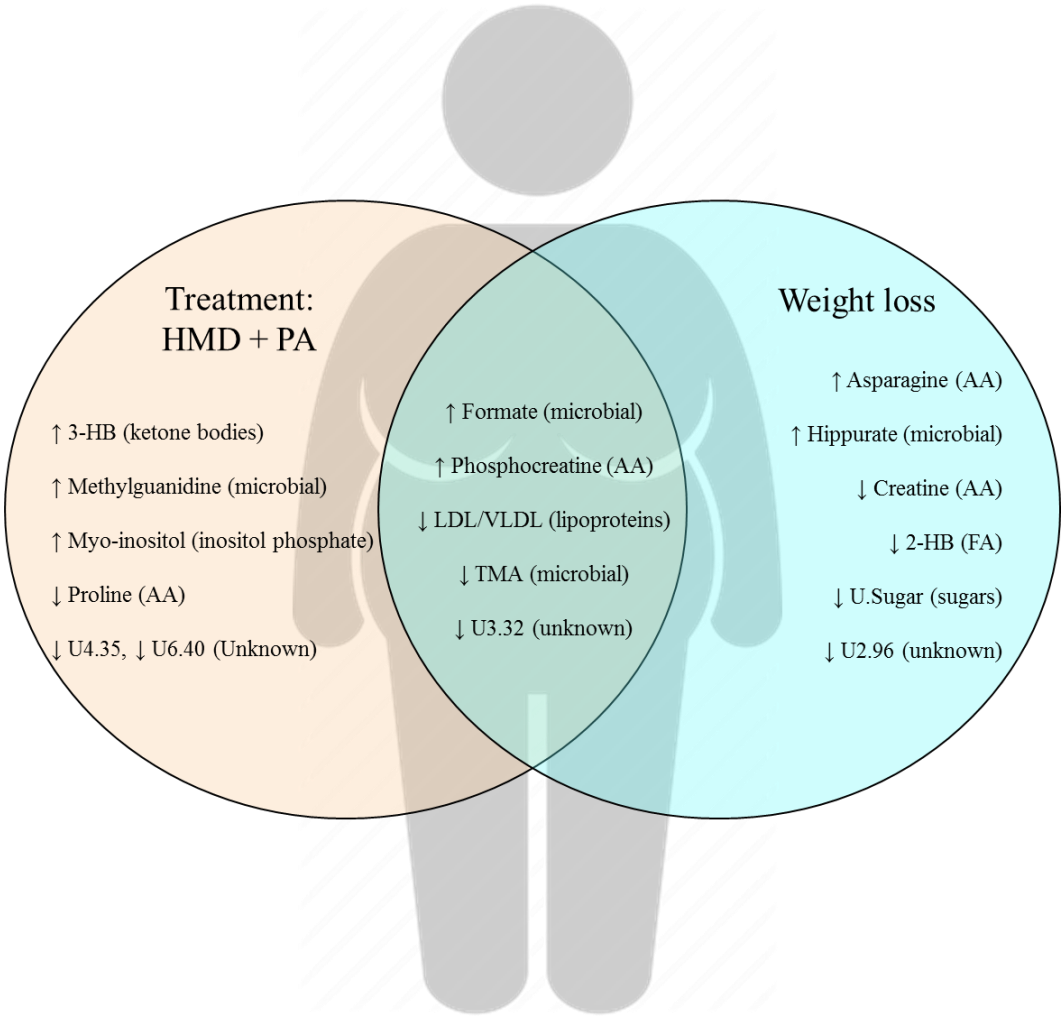
[1]Multiple random forest modelling was conducted applying a repeated double cross-validation algorithm as described in the methods section. Individual classification probability from each submodel (n = 30) is coloured by treatment group (see legend). Averaged classification probability per individual according to group is shown in larger size and similar colour. Misclassified individuals are marked by a black circle.

Figure 2. Regression analyses between actual and predicted weight change from baseline to 3 months according to rdCV-PLS modelling in control and treatment groups combined (A), and in treatment group alone (B)[1].



[1]Modelling differences in weight vs differences in untargeted ^{1}H-NMR metabolomics features after 3 months.

Figure 3. Distinct and common metabolic regulations caused by treatment and/or associated with weight loss in MHO participants[1].



[1]Metabolite's class is enclosed in parentheses; ↑ and ↓ denote up- and downregulation, respectively.

Abbreviations: 3-HB, 3-hydroxybutyrate; 2-HB, 2-hydroxybutyrate; AA, amino acids; FA, fatty acids; HD, hypocaloric diet; LDL, low-density lipoprotein cholesterol; PA, physical activity; VLDL, very low-density lipoprotein cholesterol; TMA, trimethylamine.

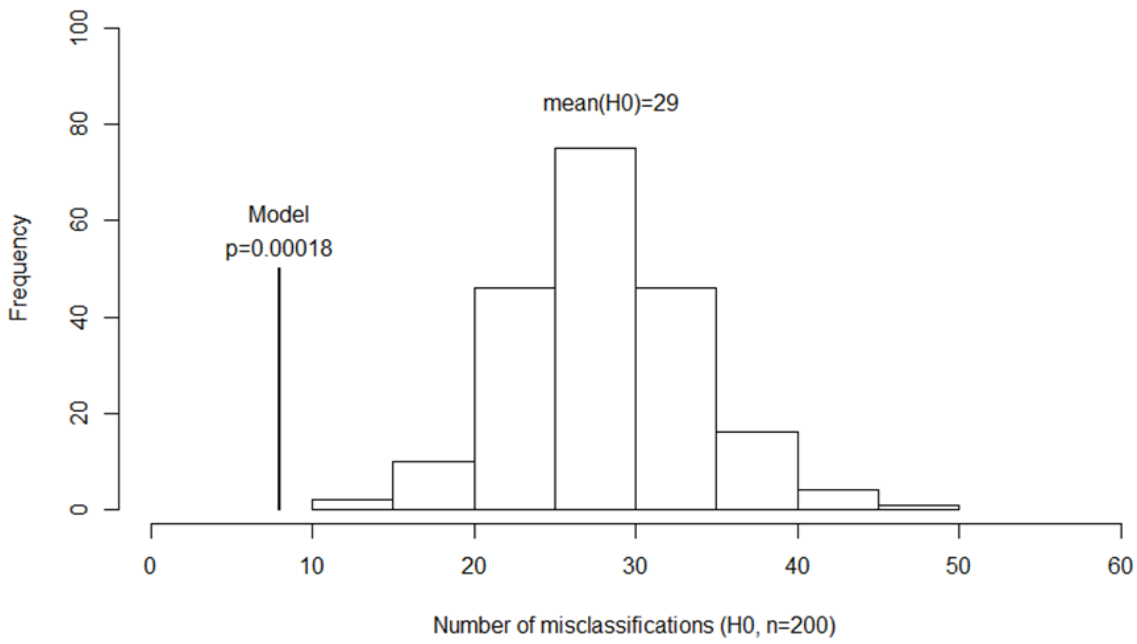
SUPPLEMENTAL MATERIAL

Table 1. Anthropometric and clinical characteristics of study participants at baseline, and after 3 and 12 months of intervention

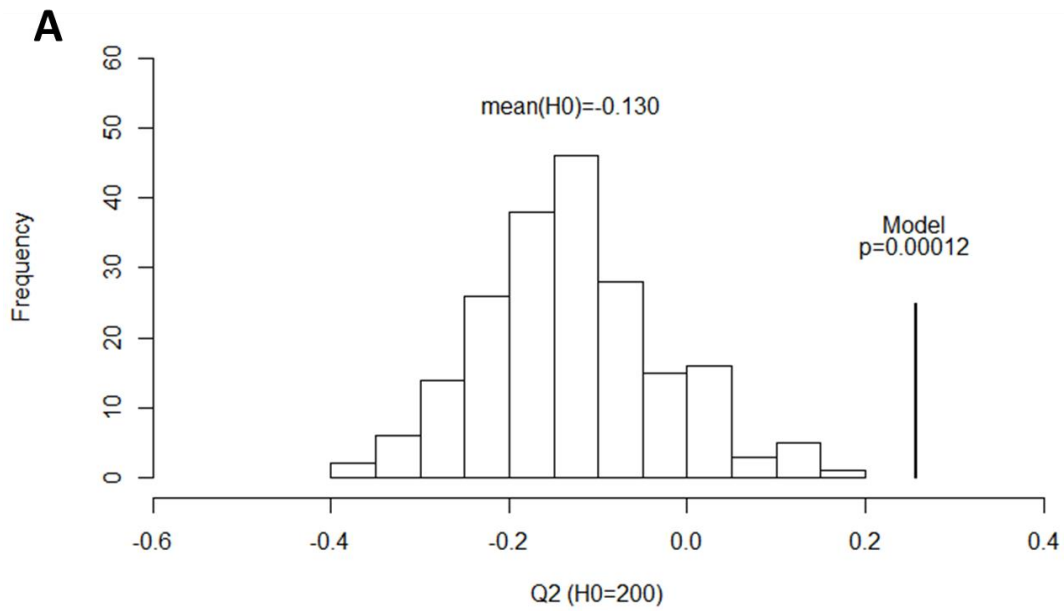
	Baseline				3 months			12 months		
	All	Control	Treatment	P[1]	Control	Treatment	P[2]	Control	Treatment	P[2]
Age, y	45.1 ± 3.45	44.4 ± 3.31	45.7 ± 3.51	0.16						
Weight (kg)	90.3 ± 13.5	91.4± 15.6	89.3 ± 11.5	0.69	88.3 ± 13.8***	80.2 ± 10.5***	<0.001	86.7 ± 13.5**	79.3 ± 12.3***	0.03
BMI, kg/m^2	35.8 ± 4.93	36.3 ± 5.74	35.4 ± 4.12	0.49	35.1 ± 4.89***	31.7 ± 3.67***	<0.001	34.5 ± 4.84**	31.3 ± 4.19***	<0.01
Waist circumference, cm	112 ± 11.2	110 ± 12.2	114 ± 10.1	0.19	105 ± 10.1***	104 ± 9.48***	<0.001	109 ± 12.5	103 ± 10.7***	<0.001
SBP, mmHg	114 ± 14.1	113 ± 15.0	115 ± 13.4	0.53	N/A	N/A		111 ± 12.1	110 ± 13.7*	0.35
DBP, mmHg	75.6 ± 9.29	74.7 ± 8.45	76.4 ± 10.1	0.50	N/A	N/A		74.9 ± 9.10	73.4 ± 10.4	0.25
Glycaemia, mg/dL	88.0 ± 8.38	89.5 ± 9.30	86.4 ± 7.29	0.15	87.7 ± 6.55	80.9 ± 5.78	0.24	86.1 ± 7.12	80.4 ± 9.72**	0.28
Total cholesterol, mg/mL	196 ± 28.6	195 ± 31.2	197 ± 26.5	0.72	184 ± 40.0*	175 ± 26.2***	0.16	186 ± 32.6*	189 ± 26.0*	0.93
LDL cholesterol, mg/mL	119 ± 27.5	123 ± 27.4	116 ± 27.7	0.43	135 ± 143	105.6 ± 21.8*	0.32	114 ± 27.1**	113 ± 21.6	0.98
HDL cholesterol, mg/mL	56.1 ± 12.1	52.9 ± 11.4	59.1 ± 12.2	0.05	57.0 ± 11.7**	53.2 ± 9.52***	<0.001	52.6 ± 11.0	57.8 ± 14.8	0.93
Triglycerides, mg/mL	93.2 ± 38.7	95.7 ± 39.4	91.0 ± 38.6	0.67	95.3 ± 48.2	81.4 ± 28.1	0.29	98.9 ± 60.6	82.3 ± 29.5*	0.84

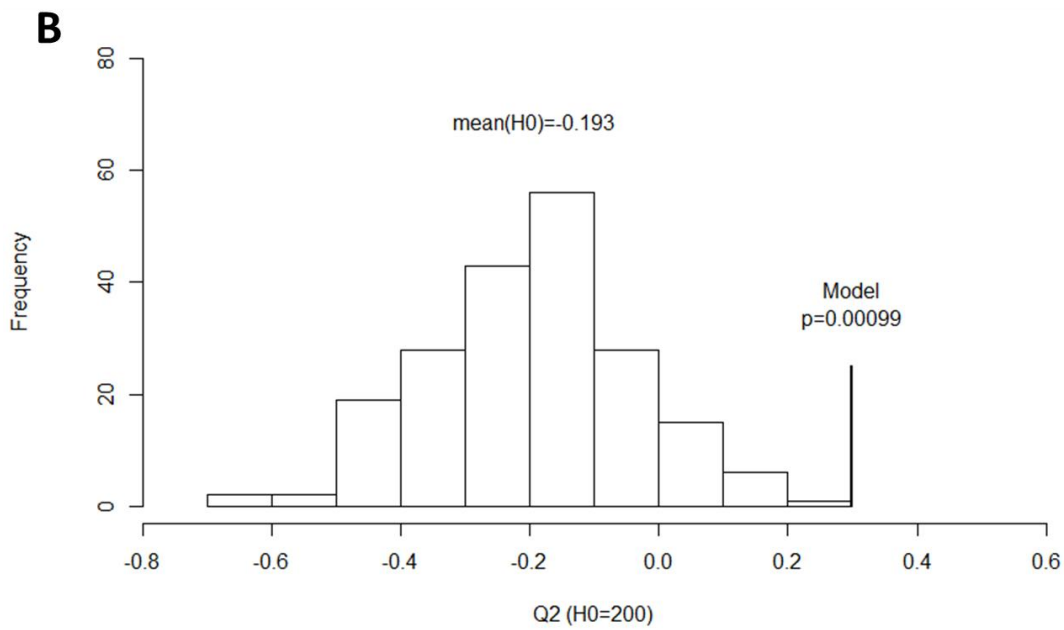
Data are presented as mean ± standard deviation. Differences from baseline were assessed by paired t-test (*p ≤ 0.05; **p ≤ 0.01;***p ≤ 0.001). *P*[1], significance for comparisons of mean values between groups. *P*[2], significance for comparisons of mean change from baseline between groups. Abbreviations: BMI, body mass index; DBP, diastolic blood pressure; HbA1c, haemoglobin A1c; HDL, high-density lipoprotein; HOMA-IR, homeostasis model assessment of insulin resistance; LDL, low-density lipoprotein; N/A, data not available; SBP, systolic blood pressure.

Supplemental Figure 1. Permutation test of the repeated double cross-validated random forest model for classification of individuals according to intervention group (control vs treatment) at 3 months.



Supplemental Figure 2. Permutation tests of the rdCV-PLS models for weight change from baseline to 3 months in control and treatment groups, together (A) and in the treatment group alone (B).





Publicación original 4:

Francisco Madrid-Gambin, Rafael Llorach, Rosa Vázquez-Fresno, Mireia Urpi-Sarda, Enrique Almanza-Aguilera, Mar Garcia-Aloy, Ramon Estruch, Dolores Corella, and Cristina Andres-Lacueva. Urinary ^{1}H Nuclear Magnetic Resonance Metabolomic Fingerprinting Reveals Biomarkers of Pulse Consumption Related to Energy-Metabolism Modulation in a Subcohort from the PREDIMED study. *Journal of Research Proteome*. 2017. 16(4):1483-1491.

Participación del doctorando:

• Aplicación de herramientas bioestadísticas para corroborar las hipótesis de partida
• Identificación y confirmación de los metabolitos discriminantes
• Seguimiento y trabajo del manuscrito hasta su publicación

Journal of proteome research

Article

pubs.acs.org/jpr

Urinary ^{1}H Nuclear Magnetic Resonance Metabolomic Fingerprinting Reveals Biomarkers of Pulse Consumption Related to Energy-Metabolism Modulation in a Subcohort from the PREDIMED study

Francisco Madrid-Gambin,[†] Rafael Llorach,[*,†,‡] Rosa Vázquez-Fresno,[†,§] Mireia Urpi-Sarda,[†,‡] Enrique Almanza-Aguilera,[†] Mar Garcia-Aloy,[†,‡] Ramon Estruch,[∥,⊥] Dolores Corella,[⊥,#] and Cristina Andres-Lacueva[*,†,‡]

[†]Biomarkers & Nutrimetabolomics Laboratory, Nutrition, Food Science and Gastronomy Department, XaRTA, INSA, Campus Torribera, Faculty of Pharmacy and Food Science, University of Barcelona, Barcelona 08028, Spain

[‡]CIBER Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Barcelona 08028, Spain

[§]Department of Biological Sciences and Department of Computing Science, University of Alberta, Edmonton AB T6G 2E9, Canada

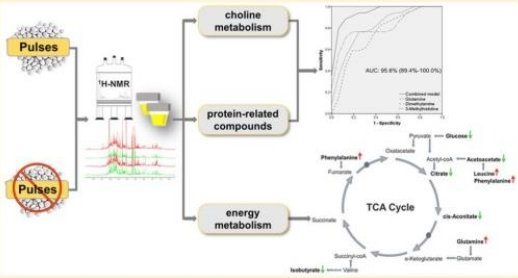
[∥]Department of Internal Medicine, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi Sunyer (IDIBAPS), Barcelona 08036, Spain

[⊥]CIBER OBN, The Spanish Biomedical Research Centre in Physiopathology of Obesity and Nutrition, Instituto de Salud Carlos III, Madrid 28029, Spain

[#]Department of Preventive Medicine and Public Health, University of Valencia, Valencia 46010, Spain

Supporting Information

ABSTRACT: Little is known about the metabolome fingerprint of pulse consumption. The study of robust and accurate biomarkers for pulse dietary assessment has great value for nutritional epidemiology regarding health benefits and their mechanisms. To characterize the fingerprinting of dietary pulses (chickpeas, lentils, and beans), spot urine samples from a subcohort from the PREDIMED study were stratified using a validated food frequency questionnaire. Urine samples of nonpulse consumers (≤4 g/day of pulse intake) and habitual pulse consumers (≥25 g/day of pulse intake) were analyzed using a ^{1}H nuclear magnetic resonance (NMR) metabolomics approach combined with multi- and univariate data analysis. Pulse consumption showed differences through 16 metabolites coming from (i) choline metabolism, (ii) protein-related compounds, and (iii) energy metabolism (including lower urinary glucose). Stepwise logistic regression analysis was applied to design a combined model of pulse exposure, which resulted in glutamine, dimethylamine, and 3-methylhistidine. This model was evaluated by a receiver operating characteristic curve (AUC > 90% in both training and validation sets). The application of NMR-based metabolomics to reported pulse exposure highlighted new candidates for biomarkers of pulse consumption and the impact on energy metabolism, generating new hypotheses on energy modulation. Further intervention studies will confirm these findings.



KEYWORDS: *pulses, legumes, metabolomics, NMR, choline metabolism, energy, biomarkers, ROC curve*

1. INTRODUCTION

The Mediterranean diet (MD) is a dietary pattern characterized by a high intake of vegetables, cereals, pulses, nuts, fish and olive oil; a low intake of red meat and processed meat products; and low to moderate consumption of poultry, wine, and dairy products.[1] Moreover, the MD has been demonstrated to be useful in the prevention of type 2 diabetes, obesity, inflammatory diseases, cardiovascular diseases (CVD), and even cancer.[2−5]

One of the components of the MD is pulses, which constitute an excellent food, providing protein, dietary fiber, and many vitamins and minerals as well as a great variety of phytochemicals.[6−8] Thus, they could contribute to the beneficial effects reported for this dietary pattern.[9] In addition, pulses are increasingly being recognized for their role in promoting good health.[6,10−12] Indeed, habitual pulse consumption is included in the main dietary guidelines worldwide, including the MD,[13] the Dietary Guidelines for Americans[14,15] and the Nordic Diet,[16] among others, and they are also

Received: September 28, 2016
Published: January 9, 2017

ACS Publications © 2017 American Chemical Society

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

advocated in view of their low environmental impact compared with other protein sources.[17]

Metabolomics is a powerful tool for identifying food-exposure biomarkers in humans[18] and provides new information on dietary components and dietary patterns.[19] In this regard, the evaluation of dietary exposure through a combination of biomarkers enables a better understanding of compliance to a dietary exposure.[20] Moreover, little is known about the metabolome fingerprint from pulse consumption either individually or as a complex food group, with only a few tentative biomarkers being described.[21,22]

Determining the changes in the urinary metabolome, new biomarkers of intake, and the effect may reveal potential modifications in diet-related physiology both in healthy and diseased individuals.[23] Furthermore, metabolomic approaches have been proposed for evaluating the relationship between nutrition and health status.[24] In light of this connection, recent scientific publications have pointed out the potential health benefits of legumes in chronic diet-related diseases, such as CVD and type 2 diabetes mellitus.[6,8,25,26] Thus, the application of metabolomics to a high-cardiovascular-risk population could provide new insights into this potential relationship.

In the present work, we compared the metabolome profiles among reported high versus low legume consumers in a free-living population to find putative biomarkers reflecting intake, the effect of intake on health status, or both. Analysis of individuals under free-living conditions enables more-representative data to be obtained on the metabolome fingerprints of pulse consumers. In light of this, a better understanding of the specific role of pulse consumption in terms of health benefits, beyond their excellent nutritional profile, is expected. Therefore, the aim of the present study was to investigate dietary pulse fingerprinting in spot urine using an untargeted ^{1}H nuclear magnetic resonance (NMR) metabolomic approach on a free-living subcohort from the PREDIMED study. For this purpose, we mainly focused on urinary biomarkers of a complex pulse exposure composed of chickpeas, lentils, and beans in a combined urinary biomarker model.

2. MATERIALS AND METHODS

2.1. PREDIMED Subcohort Study

For the present study, a subsample of 50 participants from the PREDIMED study (ISRCTN 35739639; http://www.predimed.org) was selected. The PREDIMED study is a large, parallel-group, multicenter, randomized and controlled clinical trial assessing the effects of a MD on the primary prevention of CVD. The trial protocol was conducted according to the Declaration of Helsinki and was approved by the Institutional Review Boards of all the centers involved. Briefly, free-living participants (55−80 years old) without CVD that fulfilled at least one of the two criteria (type 2 diabetes mellitus or three or more major cardiovascular risk factors) were included for a MD supplemented either with extra virgin olive oil or mixed nuts.[27] The exclusion criteria were CVD, any severe chronic illness, drug or alcohol addiction, a history of allergies, or intolerance to olive oil or nuts. The subcohort consisted of a sample of participants at high cardiovascular risk, recruited from the Barcelona and Valencia PREDIMED centers. The PREDIMED study design and 137 item validated food-frequency questionnaires (FFQs) used have been reported elsewhere.[28,29] Data reported from the FFQs included information on total legume consumption and disaggregated type of legume consumed.

2.2. Stratification of the Study Population

2.2.1. Defining Potential Consumers. Both the use of FFQs and the population stratification of a cohort of individuals by consumption have demonstrated an effective approach for the study of biomarkers of food consumption.[30−32] Participants were classified into two levels (consumers and nonconsumers) of habitual intake of dietary pulses (chickpeas, lentils, or beans) on the basis of the analysis of the validated FFQs (Table S1). Intake of pulses was calculated as the sum of consumed chickpeas, lentils, and beans. Nonpulse (NP) consumers were defined as subjects with nonconsumption or sporadic consumption (≤4.00 g/day) of pulses. Habitual pulse (HP) consumers were set at a consumption of ≥25.71 g/day, regularly. To explore global pulse consumption, individuals that did not consume the three kinds of pulses simultaneously were excluded. Additionally, the condition of sporadic or nonintake of peas (≤4 g/day) was taken into consideration because the features of this type of legume are not similar to the others.[33] No other legume types were considered.

2.2.2. Selecting Individuals by Consumption. Spot urine samples were matched to corresponding individual FFQ data. From a cohort of 828 individuals, 25 subjects were defined as NP consumers and 37 as HP consumers (none of the other participants from both pulse consumer groups fulfilled any criteria). To reduce the potential sources of variability not related to pulse exposure, the number of HP consumers was balanced against NP consumers (HP = 25, NP = 25). Finally, dietary data, anthropometry, biochemical parameters, health status, and medication were explored with a view to discarding any variability unrelated to pulse consumption.

2.3. Metabolomics Analysis

2.3.1. Urine Sample Analysis and Data Processing. Morning fasting spot urine samples were collected, aliquoted, encoded, and frozen at −80 °C until they were used. Sample preparation was based on the methodology previously published.[19] The ^{1}H NMR urinary spectra were acquired using a Varian Inova 500 MHz NMR spectrometer with presaturation of the water resonance using a NOESYPRESAT pulse sequence. During the acquisition, the internal temperature was kept constant at 298 K. An exponential window function was applied to the free induction decay (FID) with a line-broadening factor of 0.3 Hz prior to Fourier transformation. For each sample, a total of 128 scans were collected into 32 K data points with a spectral width of 14 ppm at 300 K, an acquisition time of 3.2 s, and a relaxation delay of 3 s.

^{1}H NMR spectra were phased, baseline-corrected, and calibrated (TSP, 0.0 ppm) using TopSpin software (version 3.0, Bruker, BioSpin, Germany). After baseline correction, original spectral data were bucketed in intelligent bucketing domains of 0.005 ppm with ACD NMR Processor 12.0 software (Advanced Chemistry Development, Toronto, Canada). The water signal and noise regions above 9.5 ppm and below 0.5 ppm were excluded from the analysis. Data were submitted to MetaboAnalyst 3.0 for interquartile range filtering and normalization by the sum of the intensities of the spectra.[34]

2.3.2. Statistical Analysis. The NMR data set was log-transformed, Pareto scaled, and posteriorly analyzed in a multivariate approach using SIMCA-P+13.0 software (Umetrics, Umeå, Sweden). Interindividual variation may confuse the effects of intervention, particularly in multivariate data of high

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

Journal of Proteome Research Article

Table 1. Tentative Discriminant Metabolites Derived from the Multi- and Univariate Analysis of ^{1}H NMR Signal Intensities in Urine From HP Consumers[a]

source	metabolite	HP vs NP	δ (multiplicity)	FDR p-value[b]	Cliff's delta[c]
choline metabolism	choline	↑	3.19 (s)	3.27×10^{-2}	0.475
	dimethylglycine	↑	2.93 (s)	3.81×10^{-2}	0.386
	TMAO	↑	3.27 (s)	7.29×10^{-3}	0.485
	dimethylamine	↑	2.72 (s)	1.05×10^{-2}	0.488
protein-related compounds	*N*-acetylglutamine	↑	2.04 (s)	2.55×10^{-2}	0.706
			2.08 (m)		
			2.26 (m)		
			4.18 (m)		
	glutamine	↑	2.12 (m)	1.17×10^{-6}	0.814
			2.46 (m)		
			3.76 (t)		
	phenylalanine	↑	3.19 (m)	3.21×10^{-2}	0.354
			3.98 (dd)		
			7.32 (d)		
			7.36 (m)		
			7.42 (m)		
	methylguanidine	↑	2.83 (s)	3.72×10^{-4}	0.635
	3-methylhistidine	↑	7.18 (s)	1.73×10^{-4}	0.658
			7.92 (s)		
energy metabolism	citric acid	↓	2.55 (dd)	8.43×10^{-5}	−0.690
			2.69 (dd)		
	cis-aconitic acid	↓	5.74 (s)	1.11×10^{-3}	−0.629
			3.12 (s)		
	glucose	↓	3.50 (m)	7.89×10^{-5}	−0.718
			4.66 (d)		
			5.25 (d)		
	acetoacetic acid	↓	2.27 (s)	1.95×10^{-2}	−0.408
	isovalerylglycine	↑	0.92 (d)	2.83×10^{-4}	0.635
			2.16 (d)		
			3.74 (d)		
	leucine	↑	0.94 (t)	1.18×10^{-3}	0.626
			1.70 (m)		
			3.72 (m)		
	isobutyric acid	↓	1.06 (d)	1.29×10^{-2}	−0.446

[a]All features have VIP values of ≥1.0 in the corresponding OSC-PLS-DA model. [b]The p-value of Student's t test with false discovery rate correction. [c]Estimation of the effect size by Cliff's delta with thresholds: $|n| < 0.330$ "small", $0.330 > |n| < 0.474$ "medium" and $|n| > 0.474$ "large". TMAO, trimethylamine-*N*-oxide. s: singlet, d: doublet, t: triplet, dd: double doublet, m: multiplet.

dimensionality. Therefore, partial least-squares discriminant analysis with orthogonal signal correction (OSC-PLS-DA) was used to explore the differences in metabolomes among the pulse consumption.[35] OSC filtration was used to reduce the variability not associated with dietary classification, as has been done in other published nutrimetabolomic studies.[19,31,36] The quality of the models was evaluated by the proportion of the variance of the response variable that is explained by the model (R^2Y) and the predictive ability (Q^2) parameters.[35] Validation of the models and the evaluation of the degree of overfitting were carried out using a permutation test ($n = 200$), and the correlation coefficient between the original Y and the permuted Y plotted against the cumulative R^2 and Q^2 was calculated. Those NMR signals with variable importance for projection (VIP) values ≥1 in the component of the OSC-PLS-DA model were selected as being relevant for explaining the differences in metabolic profiles. These variables were further studied through the univariate Student's t test among HP and NP consumers to assess the statistical significances. Multiple tests were controlled by the false discovery rate (FDR). Statistical significance was considered at a FDR-adjusted p-value of <0.05. Next, Cliff's delta was chosen for estimation of the effect size[37] and calculated for each feature.

2.3.3. Metabolite Identification. Metabolite identification was performed using the Chenomx NMR Suite Professional Software package (version 8.1; Chenomx Inc., Edmonton, Canada) and by comparing NMR spectral data to those available in databases such as the Human Metabolome Database (http://www.hmdb.ca), the Biological Magnetic Resonance Data Bank (http://www.bmrb.wisc.edu), and the Madison Metabolomics Consortium Database (www.mmcd.nmrfam.wisc.edu), along with the existing NMR-based metabolomics literature. Further, a Pearson's correlation test and clustering analysis with Pearson distance and Ward's minimum variance using PermutMatrix 1.9.3.0 software[38] were applied to identify the signals corresponding to the same metabolite.

2.4. Study of Combined Urinary Biomarker Model

The interaction between gender and the resulting metabolites was evaluated by a logistic regression for discarding any effect on the biomarkers. Then, the discriminant metabolites were submitted to a stepwise logistic regression analysis (IBM SPSS

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

Journal of Proteome Research Article

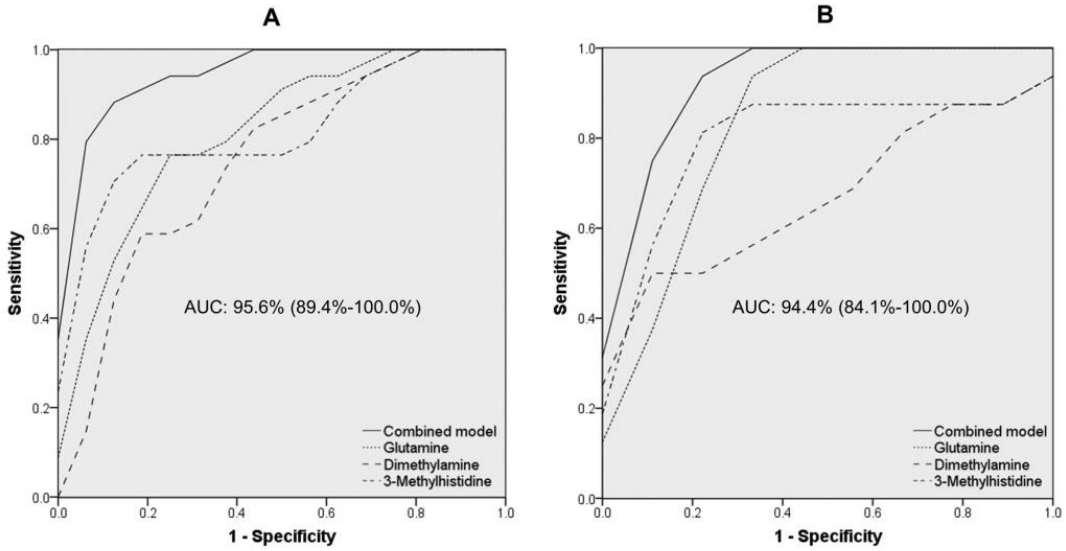


Figure 1. Receiver operating characteristic (ROC) curves of combined model (continuous line) with the area under the ROC curve and of included individual metabolites (discontinuous lines) in the training (A) and validation (B) sets.

Statistics 20 software, SPSS, Inc., Chicago, IL) to evaluate whether the combination of more than one biomarker improves the discrimination[20] of pulse consumption. The models were constructed through a dichotomous variable of pulse consumption as the dependent variable and identified metabolites as independent variables, with a p-value of <0.05 as a condition required for entering and remaining in the model. For validation of models, the analysis with a training set of 2/3 of the samples (removing 1/3 of the individuals as the validation set) was permuted 20 times. Spearman's rank correlation coefficient was used to assess correlations between the combined models and pulse consumption.

The global performance of the models was evaluated by receiver operating characteristic (ROC) curve and estimation of the area under the curve (AUC) values. The optimum cutoff for sensitivity and specificity of the biomarkers was determined as the minimum distance to the top-left corner.[39]

3. RESULTS

A flowchart of the participants allocated in the present study is presented in the Figure S1. Anthropometric measurements and biochemical analyses were performed using standardized methods.[28] HP consumers showed a pulse consumption of 38.45 ± 14.68 g/day, while NP consumers reported a consumption of 3.75 ± 3.95 g/day (mean ± SD). The characteristics of participants classified by pulse consumption are presented in Table S2. The stratified populations were not different in terms of disease (type 2 diabetes mellitus or cardiovascular risk factors), medications or biochemical parameters, among other data. Subjects who were HP consumers showed higher amounts of both dietary fiber (p < 0.01) and polyunsaturated fatty acid (p < 0.05) intakes, probably as a consequence of legume macronutrient composition.[6,40] No significances other than pulses were found with regard to food intake.

3.1. Selection of Significant Biomarkers Related to Pulse Consumption

For the analysis of the features belonging to pulse consumption in the urinary metabolome of the HP and NP consumers, an orthogonal signal correction was applied before PLS-DA analysis. The OSC-PLS-DA analysis of the two groups resulted in a latent variable model with R^2Y and Q^2 values of 0.954 and 0.809, respectively, indicating that the model was able to classify most of the individuals in the correct consumption group. The corresponding permutation tests showed negative Q^2 intercepts with a value of −0.164, implicating validation of the model.[35] With the purpose of selecting the most discriminative urinary markers of consumption, only the statistically significant variables coming from both multi- and univariate analyses simultaneously were considered.

3.2. Identified Biomarkers of Habitual Pulse Consumption

A total of 16 compounds were identified as discriminant metabolites of pulse consumption. Metabolites and chemical shifts identified corresponding to statistical analyses are presented in Table 1. The total number of metabolites related to pulse consumption was divided into categories as follows: (i) choline metabolism: choline, dimethylglycine, trimethylamine-*N*-oxide (TMAO) and dimethylamine; (ii) protein-related compounds: 3-methylhistidine, methylguanidine, phenylalanine, glutamine and *n*-acetylglutamine; and (iii) energy metabolism: glucose, leucine, isovalerylglycine, and isobutyric, acetoacetic, citric, and *cis*-aconitic acids.

3.3. Combined Urinary Biomarker Approach

Logistic regression analysis revealed that there was no significant interaction between gender and the metabolites (p > 0.05; all) as shown in Table S3. For the improvement of the discrimination between groups (HP and NP consumers), a conditional stepwise variable selection method, through a binary logistic regression analysis, was used on a combination of more than one discriminant metabolite. Table S4 shows the

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

Journal of Proteome Research Article

Table 2. Receiver Operating Characteristic Curve Parameters of Combined Models and of Individual Metabolites in Both Training and Validation Sets

	set[b]	sensitivity (%)	specificity (%)	AUC (95% CI)
combined model	training	88.2	93.7	95.6 (89.4–100.0)
	validation	87.5	88.9	94.4 (84.1–100.0)
3-methylhistidine	training	76.5	87.5	82.4 (67.7–97.0)
	validation	87.5	77.8	80.6 (56.2–100.0)
glutamine	training	76.5	81.2	81.6 (67.0–96.3)
	validation	87.5	77.8	84.7 (65.3–100.0)
dimethylamine	training	82.4	62.5	75.0 (57.8–92.2)
	validation	50.0	66.7	68.1 (40.4–100.0)

[a]AUC: area under the ROC curve. CI: confidence interval. [b]Corresponding to 2/3 of the population for the training and 1/3 for the validation set.

resulting metabolites included in all 20 permuted models and the contribution to the model. A total of three metabolites were included in the fitted model according to the maximum AUC, which contained two protein-related metabolites (glutamine and 3-methylhistidine) and one choline-related metabolite (dimethylamine). These three metabolites correlated individually with the pulse consumption. However, the combined model exhibited the strongest correlation ($r = 0.73$, $p < 0.01$) with the pulse exposure, as shown in Table S5.

The ROC curve analysis was used to evaluate the combined metabolite model and their metabolites using both training and validation sets separately. The highest AUC was for the combined metabolite model for both training (AUC = 95.6%) and validation (AUC = 94.4%) sets, including glutamine, 3-methylhistidine, and dimethylamine followed by the individual metabolites 3-methylhistidine (AUC = 82.4%), glutamine (AUC = 81.6%), and dimethylamine (AUC = 75.0%), as shown in Figure 1. The equations generated from the logistic regression and the AUCs from the models with their sensitivity and specificity are shown in Table 2.

4. DISCUSSION

In this study, we present a panel of different urinary metabolites related to habitual pulse exposure using a ^{1}H NMR based untargeted nutrimetabolomic approach in a free-living population. In addition, high correlations were found when the exposure was assessed as a continuous variable (defined by the combined biomarker panel).

4.1. Characterization of Pulse Fingerprinting in Urine

4.1.1. Pulse Metabolomic Fingerprinting and Choline Metabolism. Several compounds found in the spot urine of pulse consumers are related to choline. Thus, pulses, as a rich source of choline,[41] may be the precursor of additional metabolites that are susceptible to microbial degradation generating new compounds.[42] Therefore, the increase of several intermediates of choline metabolism, such as TMAO and dimethylamine, appears to be a consequence of the microbial activity in HP consumers. The increase of dimethylamine, which is also a downstream product of choline, supports the microbial degradation of TMAO from choline. Furthermore, TMAO was identified as a major source of urinary dimethylamine in humans,[43] directly related to gut microbiota metabolism.[44] However, the increase of urinary dimethylglycine may also come from the choline contained in pulses. The enzymes choline dehydrogenase, betaine aldehyde dehydrogenase, and betaine homocysteine methyltransferase lead to dimethylglycine from choline.[45] Therefore, the results of the present study suggest the impact on urinary metabolome by choline from pulses that is degraded via both (i) mammalian pathways in which choline is converted to dimethylglycine through betaine, and (ii) microbial metabolism in which choline is degraded to trimethylamine, TMAO, and dimethylamine. For this reason, we propose dimethylamine and dimethylglycine in spot urine as potential candidates of biomarkers of pulse consumption. Nevertheless, these choline-related metabolites need to be further explored in controlled studies confirming that they are food intake biomarkers instead of reflecting metabolic differences due to the pulse consumption. Figure 2 shows both proposed pathways for downstream products of choline.

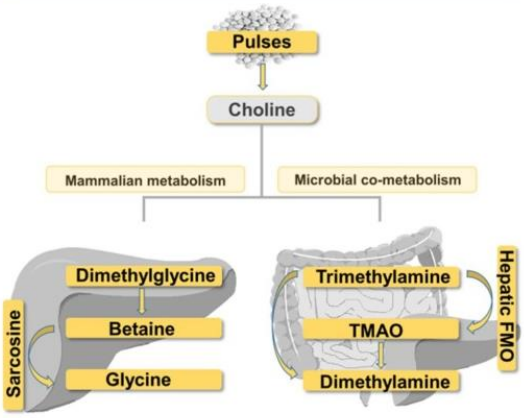


Figure 2. Proposed pathways for choline degradation from pulses including significant metabolites in HP consumers in the present study.

4.1.2. Pulse Metabolomic Fingerprinting and Protein-Related Compounds. With regard to the increases in glutamine and the acetylated form *n*-acetylglutamine, several explanations may be proposed. Glutamine and *n*-acetylglutamine could come from dietary sources because glutamine is found in high-protein foods, such as pulses.[46] Another explanation could be the alteration of urinary levels previously shown in this type of population,[47] affected by pulse consumption. There was a higher excretion of 3-methylhistidine in HP consumers. This metabolite is a biomarker of meat and fish consumption,[48] denoting a potential role as a biomarker of consumption of high-protein foods. Interestingly, all food sources of this metabolite are also protein sources, including pulses as a vegetable source, as highlighted in the

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483–1491

Journal of Proteome Research Article

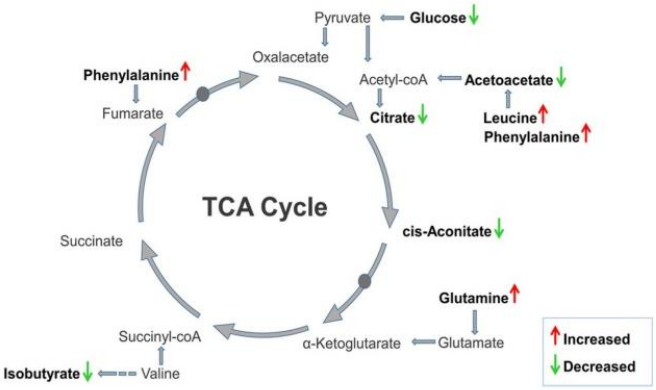


Figure 3. Modified metabolites found in HP consumers connected to energy metabolism.

present study. However, 3-methylhistidine is also a muscle protein breakdown that is sensitive to gender and age.[49] Therefore, its use as biomarker should be considered carefully. Methylguanidine is derived from protein catabolism and from the breakdown of creatinine;[50,51] therefore, it may be related to protein from pulses.

4.1.3. Pulse Metabolomic Fingerprinting and Energy Metabolism. The signals of several usual metabolites were altered between the two groups. However, their definition as food intake biomarkers is controversial. Instead, they probably reflect metabolic differences associated with being a low and high consumer, determined on the basis of the study design. Most of the biomarkers found in the present study are metabolites related to energy metabolism. The lower excretion of acetoacetic acid, glucose, and tricarboxylic acid (TCA) cycle intermediates (citric and *cis*-aconitic acids) appears to involve a different energy modulation according to the pulse consumption. This fact is in part reinforced by changes in BCAAs and subproducts, which are implicated in energy metabolism. For example, isobutyric acid is a short-chain fatty acid that is a product of BCAA catabolism of valine, which is a glucogenic BCAA metabolized via the methylmalonyl−CoA in the TCA cycle.[52] However, acetoacetic acid is a ketone body produced in the human liver for fatty acid breakdown,[53] which serves as a source of energy when normal glycolysis is altered. Interestingly, acetoacetic acid was shown to be increased in diabetes mellitus.[54] Hence, we hypothesize that gluconeogenesis may be diminished in pulse consumers, supported by the urinary reduction of acetoacetic and isobutyric acids, and the reduction of TCA cycle intermediates and urinary glucose. Furthermore, it was observed that pulse consumption has a glucose-lowering role in diabetes mellitus,[55,56] thereby explaining the lower plasma glucose concentration and lower urinary excretion. Figure 3 shows the resulting endogenous metabolites connected to the TCA cycle. Nevertheless, the small sample size that resulted after the stratification of the population leads to only exploratory results that should be confirmed.

The role of other findings such as increases of leucine and phenylalanine in pulse consumers is unclear. On the one hand, these habitual urinary compounds could be increased as a consequence of pulses being the source. On the other hand, another explanation of these findings could support the hypothesis above. Leucine, which is an acetoacetic acid precursor, may modulate glucose metabolism through oxidation, as well as insulin signaling and release. In addition, stimulation of glucose recycling via the glucose−alanine cycle by leucine may inhibit protein breakdown.[57,58] Alterations in urinary leucine have also been proposed for the prediction of diabetes mellitus, probably related to the perturbed energy metabolism.[54] The origin of increased phenylalanine is also uncertain. This ketogenic amino acid is also altered in an insulin-resistant state and obesity.[59] Overall, the consumption of pulses seems to affect the energy metabolism in the studied population.

4.2. New Biomarker Panel to Characterize Habitual Pulse Consumption

To delimit the prediction of habitual pulse intake, composed of lentils, chickpeas, and beans, a combination of more than one discriminatory metabolite had to be studied. The combination of three metabolites enhanced considerably the AUC and the confidence interval of the model in comparison with individual metabolites, as shown in Table 2. The developed model indicated that glutamine, 3-methylhistidine, and dimethylamine were strong candidates for exposure biomarkers. It is important to note that the role of the component coming from choline metabolism suggests the importance of this metabolite as a biomarker of intake. Interestingly, metabolites displaying changes in energy metabolism were scarcely considered by the stepwise logistic regression. None of the other metabolites entered the model, probably as a result of collinearity in the evidence provided by these compounds, which may originate from the same metabolic pathways, giving similar biological or dietary information.[36] Instead, two metabolites related to protein coming from pulses and one connected to microbiota choline degradation were established in the combined metabolite model, giving complementary information, showing a better discrimination (AUC > 90% in both training and validation sets) than each metabolite individually (AUC < 90% in all cases) and reinforcing the improved capacity of biomarker patterns to distinguish between different dietary exposures.

5. CONCLUSIONS

We applied an untargeted ^{1}H NMR-based metabolomic strategy to distinguish the urinary metabolome of habitual pulse consumption in a free-living population. Stepwise logistic regression analysis exhibited a useful approach to designing a combined urinary biomarker model, taking into consideration

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

the different characteristics of pulses. With regard to food metabolome, this study points to a central role of choline contained in pulses and breakdown products such as dimethylglycine, TMAO, and dimethylamine. Protein-related compounds such as glutamine, 3-methylhistidine, and methylguanidine were also increased in the urine of HP consumers. The combined metabolite model indicated that dimethylamine, 3-methylhistidine, and glutamine were the strongest candidates for exposure prediction. In relation to energy metabolism, numerous compounds connected to the TCA cycle, including BCAAs and acetoacetic acid, were modified, denoting a substantial impact on energy metabolism modulation and on urinary glucose in this population. However, because the status of type 2 diabetes mellitus or three or more major cardiovascular risk factors in the studied population could have a distinctive energy modulation, properly controlled interventions could confirm the findings observed in this cross-sectional study.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.6b00860.

Tables showing criteria for stratifying participants by frequency of consumption, characteristics of the study population according to pulse consumption, interaction between gender and the metabolites found in the present study, permuted models used in training and validation sets with the resulting metabolites, and correlations between legume consumption and the combined model for prediction of legume exposure and considered individual metabolites; and a flowchart of subjects from the PREDIMED subcohort included in the study. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: rafallorach@ub.edu. Tel: +34.934033798.
*E-mail: candres@ub.edu. Tel: +34.934034840. Fax: +34.93403593.

ORCID

Cristina Andres-Lacueva: 0000-0002-8494-4978

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study is supported by PCIN-2014-133 (FOODBALL project) from the Spanish Ministry of Economy and Competitiveness (MINECO) together with the Joint Programming Initiative "Healthy Diet for Healthy Life" (JPI-HDHL, http://www.healthydietforhealthylife.eu) and the award of 2014SGR1566 from the Generalitat de Catalunya's Agency AGAUR. F.M-G. acknowledges the APIF Ph.D. fellowship (University of Barcelona). M.U-S. thank thes "Ramón y Cajal" programme (RYC-2011-09677) from MINECO and the Fondo Social Europeo. E.A.-A. thanks CONACYT (Mexico) for the Ph.D. fellowship.

■ ABBREVIATIONS

AUC, area under the curve; FFQ, food frequency questionnaire; FID, free induction decay; HP, habitual pulses; ISRCTN, International Standard Randomized Controlled Trial Number; KOD, potassium deuteroxide; MD, Mediterranean diet; NMR, nuclear magnetic resonance; NP, nonpulses; OSC-PLS-DA, partial least-squares discriminant analysis with orthogonal signal correction; ROC, receiver operating characteristic; TCA, tricarboxylic acid; TMAO, trimethylamine-*N*-oxide; TSP, 3-(trimethylsilyl)-proprionate-2,2,3,3-d_4; VIP, variable importance projection

■ REFERENCES

(1) Bergouignan, A.; Momken, I.; Schoeller, D. A.; Simon, C.; Blanc, S. Metabolic fate of saturated and monounsaturated dietary fats: the Mediterranean diet revisited from epidemiological evidence to cellular mechanisms. *Prog. Lipid Res.* **2009**, *48* (3−4), 128−147.

(2) Salas-Salvadó, J.; Bulló, M.; Babio, N.; Martínez-González, M. Á.; Ibarrola-Jurado, N.; Basora, J.; Estruch, R.; Covas, M. I.; Corella, D.; Arós, F.; et al. Reduction in the incidence of type 2 diabetes with the Mediterranean diet: results of the PREDIMED-Reus nutrition intervention randomized trial. *Diabetes Care* **2011**, *34* (1), 14−19.

(3) Marlow, G.; Ellett, S.; Ferguson, I. R.; Zhu, S.; Karunasinghe, N.; Jesuthasan, A. C.; Han, D. Y.; Fraser, A. G.; Ferguson, L. R. Transcriptomics to study the effect of a Mediterranean-inspired diet on inflammation in Crohn's disease patients. *Hum. Genomics* **2013**, *7* (1), 24.

(4) Lopez-Legarrea, P.; Fuller, N. R.; Zulet, M. A.; Martinez, J. A.; Caterson, I. D. The influence of Mediterranean, carbohydrate and high protein diets on gut microbiota composition in the treatment of obesity and associated inflammatory state. *Asia Pac. J. Clin. Nutr.* **2014**, *23* (3), 360−368.

(5) Estruch, R.; Ros, E.; Salas-Salvadó, J.; Covas, M.-I.; Corella, D.; Arós, F.; Gómez-Gracia, E.; Ruiz-Gutiérrez, V.; Fiol, M.; Lapetra, J.; et al. Primary prevention of cardiovascular disease with a Mediterranean diet. *N. Engl. J. Med.* **2013**, *368* (14), 1279−1290.

(6) Bouchenak, M.; Lamri-Senhadji, M. Nutritional quality of legumes, and their role in cardiometabolic risk prevention: a review. *J. Med. Food* **2013**, *16* (3), 185−198.

(7) Roy, F.; Boye, J. I.; Simpson, B. K. Bioactive proteins and peptides in pulse crops: Pea, chickpea and lentil. *Food Res. Int.* **2010**, *43* (2), 432−442.

(8) Campos-Vega, R.; Loarca-Piña, G.; Oomah, B. D. Minor components of pulses and their potential impact on human health. *Food Res. Int.* **2010**, *43* (2), 461−482.

(9) Sofi, F.; Abbate, R.; Gensini, G. F.; Casini, A. Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis. *Am. J. Clin. Nutr.* **2010**, *92* (5), 1189−1196.

(10) Dilis, V.; Trichopoulou, A. Nutritional and health properties of pulses. *Mediterranean Journal of Nutrition and Metabolism*; IOS Press January 1, 2009; pp 149−157.

(11) Ramalingam, A.; Kudapa, H.; Pazhamala, L. T.; Weckwerth, W.; Varshney, R. K. Proteomics and Metabolomics: Two Emerging Areas for Legume Improvement. *Front. Plant Sci.* **2015**, *6*, 1116.

(12) Faris, M. A. I. E.; Takruri, H. R.; Issa, A. Y. Role of lentils (Lens culinaris L.) in human health and nutrition: A review. *Mediterranean Journal of Nutrition and Metabolism.* IOS Press: Amsterdam, The Netherlands, 2013; pp 3−16.

(13) Bach-Faig, A.; Berry, E. M.; Lairon, D.; Reguant, J.; Trichopoulou, A.; Dernini, S.; Medina, F. X.; Battino, M.; Belahsen, R.; Miranda, G.; et al. Mediterranean diet pyramid today. Science and cultural updates. *Public Health Nutr.* **2011**, *14* (12A), 2274−2284.

(14) U.S. Department of Health and Human Services and U.S. Department of Agriculture. 2015−2020 Dietary Guidelines for Americans. http://health.gov/dietaryguidelines/2015/ (accessed May 9, 2016).

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

(15) Britten, P.; Marcoe, K.; Yamini, S.; Davis, C. Development of food intake patterns for the MyPyramid Food Guidance System. *J. Nutr. Educ. Behav.* **2006**, *38* (6), S78−92.
(16) Mithril, C.; Dragsted, L. O.; Meyer, C.; Tetens, I.; Biltoft-Jensen, A.; Astrup, A. Dietary composition and nutrient content of the New Nordic Diet. *Public Health Nutr.* **2013**, *16* (5), 777−785.
(17) MacWilliam, S.; Wismer, M.; Kulshreshtha, S. Life cycle and economic assessment of Western Canadian pulse systems: The inclusion of pulses in crop rotations. *Agric. Syst.* **2014**, *123*, 43−53.
(18) O'Gorman, A.; Gibbons, H.; Brennan, L. Metabolomics in the Identification of Biomarkers of Dietary Intake. *Comput. Struct. Biotechnol. J.* **2013**, *4* (5), 1−7.
(19) Vázquez-Fresno, R.; Llorach, R.; Urpi-Sarda, M.; Lupianez-Barbero, A.; Estruch, R.; Corella, D.; Fitó, M.; Arós, F.; Ruiz-Canela, M.; Salas-Salvadó, J.; et al. Metabolomic pattern analysis after mediterranean diet intervention in a nondiabetic population: A 1- and 3-year follow-up in the PREDIMED study. *J. Proteome Res.* **2015**, *14* (1), 531−540.
(20) Garcia-Aloy, M.; Llorach, R.; Urpi-Sarda, M.; Tulipani, S.; Estruch, R.; Martínez-González, M. a.; Corella, D.; Fitó, M.; Ros, E.; Salas-Salvadó, J.; et al. Novel multimetabolite prediction of walnut consumption by a urinary biomarker model in a free-living population: The predimed study. *J. Proteome Res.* **2014**, *13* (7), 3476−3483.
(21) Perera, T.; Young, M. R.; Zhang, Z.; Murphy, G.; Colburn, N. H.; Lanza, E.; Hartman, T. J.; Cross, A. J.; Bobe, G. Identification and monitoring of metabolite markers of dry bean consumption in parallel human and mouse studies. *Mol. Nutr. Food Res.* **2015**, *59* (4), 795−806.
(22) Bonetti, A.; Marotti, I.; Dinelli, G. Urinary excretion of kaempferol from common beans (Phaseolus vulgaris L.) in humans. *Int. J. Food Sci. Nutr.* **2007**, *58* (4), 261−269.
(23) Scalbert, A.; Brennan, L.; Manach, C.; Andres-Lacueva, C.; Dragsted, L. O.; Draper, J.; Rappaport, S. M.; van der Hooft, J. J. J.; Wishart, D. S. The food metabolome: a window over dietary exposure. *Am. J. Clin. Nutr.* **2014**, *99* (6), 1286−1308.
(24) McNiven, E. M. S.; German, J. B.; Slupsky, C. M. Analytical metabolomics: nutritional opportunities for personalized health. *J. Nutr. Biochem.* **2011**, *22* (11), 995−1002.
(25) Souza, R. G. M.; Gomes, A. C.; Naves, M. M. V; Mota, J. F. Nuts and legume seeds for cardiovascular risk reduction: scientific evidence and mechanisms of action. *Nutr. Rev.* **2015**, *73* (6), 335−347.
(26) Wilson, C. Nutrition: Consumption of legumes might be beneficial in type 2 diabetes mellitus. *Nat. Rev. Endocrinol.* **2013**, *9* (1), 3.
(27) Martinez-Gonzalez, M. A.; Corella, D.; Salas-Salvado, J.; Ros, E.; Covas, M. I.; Fiol, M.; Warnberg, J.; Aros, F.; Ruiz-Gutierrez, V.; Lamuela-Raventos, R. M.; et al. Cohort Profile: Design and methods of the PREDIMED study. *Int. J. Epidemiol.* **2012**, *41* (2), 377−385.
(28) Estruch, R.; Martinez-Gonzalez, M. A.; Corella, D.; Salas-Salvado, J.; Ruiz-Gutierrez, V.; Covas, M. I. Effects of a Mediterranean-Style Diet on Cardiovascular Risk Factors. *Ann. Int. Med.* **2006**, *145* (1), 1−11.
(29) Fernández-Ballart, J. D.; Piñol, J. L.; Zazpe, I.; Corella, D.; Carrasco, P.; Toledo, E.; Perez-Bauer, M.; Martínez-González, M. A.; Salas-Salvadó, J.; Martín-Moreno, J. M. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. *Br. J. Nutr.* **2010**, *103*, 1808−1816.
(30) Khan, N.; Khymenets, O.; Urpí-Sardà, M.; Tulipani, S.; Garcia-Aloy, M.; Monagas, M.; Mora-Cubillos, X.; Llorach, R.; Andres-Lacueva, C. Cocoa polyphenols and inflammatory markers of cardiovascular disease. *Nutrients* **2014**, *6* (2), 844−880.
(31) Pujos-Guillot, E.; Hubert, J.; Martin, J. F.; Lyan, B.; Quintana, M.; Claude, S.; Chabanas, B.; Rothwell, J. a.; Bennetau-Pelissero, C.; Scalbert, A.; et al. Mass spectrometry-based metabolomics for the discovery of biomarkers of fruit and vegetable intake: Citrus fruit as a case study. *J. Proteome Res.* **2013**, *12* (4), 1645−1659.
(32) Vázquez-Fresno, R.; Llorach, R.; Urpi-Sarda, M.; Khymenets, O.; Bullo, M.; Corella, D.; Fito, M.; Martinez-Gonzalez, M. A.; Estruch, R.; Andres-Lacueva, C. An NMR metabolomics approach reveals a combined-biomarkers model in a wine interventional trial with validation in free-living individuals of the PREDIMED study. *Metabolomics* **2015**, *11* (4), 797−806.
(33) Food and Agriculture Organization. Definition and classification of commodities, 1994. http://www.fao.org/ES/faodef/fdef04e.htm (accessed May 1, 2016).
(34) Xia, J.; Sinelnikov, I. V.; Han, B.; Wishart, D. S. MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Res.* **2015**, *43* (W1), W251−W257.
(35) Llorach-Asunción, R.; Jauregui, O.; Urpi-Sarda, M.; Andres-Lacueva, C. Methodological aspects for metabolome visualization and characterization: a metabolomic evaluation of the 24 h evolution of human urine after cocoa powder consumption. *J. Pharm. Biomed. Anal.* **2010**, *51* (2), 373−381.
(36) Garcia-Aloy, M.; Llorach, R.; Urpi-Sarda, M.; Jáuregui, O.; Corella, D.; Ruiz-Canela, M.; Salas-Salvadó, J.; Fitó, M.; Ros, E.; Estruch, R.; et al. A metabolomics-driven approach to predict cocoa product consumption by designing a multimetabolite biomarker model in free-living subjects from the PREDIMED study. *Mol. Nutr. Food Res.* **2015**, *59* (2), 212−220.
(37) Cliff, N. Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin* **1993**, *114* (3), 494−509.
(38) Caraux, G.; Pinloche, S. PermutMatrix: A graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics* **2005**, *21* (7), 1280−1281.
(39) Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S. Translational biomarker discovery in clinical metabolomics: An introductory tutorial. *Metabolomics* **2013**, *9* (2), 280−299.
(40) De Almeida Costa, G. E.; Da Silva Queiroz-Monici, K.; Pissini Machado Reis, S. M.; De Oliveira, A. C. Chemical composition, dietary fibre and resistant starch contents of raw and cooked pea, common bean, chickpea and lentil legumes. *Food Chem.* **2006**, *94* (3), 327−330.
(41) Lewis, E. D.; Kosik, S. J.; Zhao, Y.-Y.; Jacobs, R. L.; Curtis, J. M.; Field, C. J. Total choline and choline-containing moieties of commercially available pulses. *Plant Foods Hum. Nutr.* **2014**, *69* (2), 115−121.
(42) Tremaroli, V.; Bäckhed, F. Functional interactions between the gut microbiota and host metabolism. *Nature* **2012**, *489* (7415), 242−249.
(43) Zhang, A. Q.; Mitchell, S. C.; Ayesh, R.; Smith, R. L. Dimethylamine formation in man. *Biochem. Pharmacol.* **1993**, *45* (11), 2185−2188.
(44) Dumas, M.-E.; Barton, R. H.; Toye, A.; Cloarec, O.; Blancher, C.; Rothwell, A.; Fearnside, J.; Tatoud, R.; Blanc, V.; Lindon, J. C.; et al. Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (33), 12511−12516.
(45) Friesen, R. W.; Novak, E. M.; Hasman, D.; Innis, S. M. Relationship of Dimethylglycine, Choline, and Betaine with Oxoproline in Plasma of Pregnant Women and Their Newborn Infants. *J. Nutr.* **2007**, *137* (12), 2641−2646.
(46) Marinangeli, C. P. F.; Jones, P. J. H. Pulse grain consumption and obesity: effects on energy expenditure, substrate oxidation, body composition, fat deposition and satiety. *Br. J. Nutr.* **2012**, *108* (S1), S46−S51.
(47) Roberts, L. D.; Koulman, A.; Griffin, J. L. Towards metabolic biomarkers of insulin resistance and type 2 diabetes: progress from the metabolome. *Lancet Diabetes Endocrinol.* **2014**, *2* (1), 65−75.
(48) Brennan, L.; Gibbons, H.; O'Gorman, A. An Overview of the Role of Metabolomics in the Identification of Dietary Biomarkers. *Curr. Nutr. Rep.* **2015**, *4* (4), 304−312.
(49) Aranibar, N.; Vassallo, J. D.; Rathmacher, J.; Stryker, S.; Zhang, Y.; Dai, J.; Janovitz, E. B.; Robertson, D.; Reily, M.; Lowe-Krentz, L.; et al. Identification of 1- and 3-methylhistidine as biomarkers of skeletal muscle toxicity by nuclear magnetic resonance-based metabolic profiling. *Anal. Biochem.* **2011**, *410* (1), 84−91.
(50) Yokozawa, T.; Fujitsuka, N.; Oura, H. Studies on the precursor of methylguanidine in rats with renal failure. *Nephron* **1991**, *58* (1), 90−94.

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

(51) Ando, A.; Orita, Y.; Nakata, K.; Tsubakihara, Y.; Takamitsu, Y.; Ueda, N.; Yanase, M.; Abe, H. Effect of low protein diet and surplus of essential amino acids on the serum concentration and the urinary excretion of methylguanidine and guanidinosuccinic acid in chronic renal failure. *Nephron* **1979**, *24* (4), 161−169.

(52) Hutson, S. M.; Sweatt, A. J.; Lanoue, K. F. Branched-chain amino acid metabolism: implications for establishing safe intakes. *J. Nutr.* **2005**, *135* (6), 1557S−64S.

(53) Stern, J. R. Enzymes of acetoacetate formation and breakdown. *Methods Enzymol.* **1955**, *1*, 573−585.

(54) Urpi-Sarda, M.; Almanza-Aguilera, E.; Tulipani, S.; Tinahones, F. J.; Salas-Salvadó, J.; Andres-Lacueva, C. Metabolomics for Biomarkers of Type 2 Diabetes Mellitus: Advances and Nutritional Intervention Trends. *Curr. Cardiovasc. Risk Rep.* **2015**, *9* (3), 12.

(55) Mudryj, A. N.; Yu, N.; Aukema, H. M. Nutritional and health benefits of pulses. *Appl. Physiol., Nutr., Metab.* **2014**, *39* (11), 1197−1204.

(56) Singhal, P.; Kaushik, G.; Mathur, P. Antidiabetic potential of commonly consumed legumes: a review. *Crit. Rev. Food Sci. Nutr.* **2014**, *54* (5), 655−672.

(57) Layman, D. K.; Walker, D. A. Potential importance of leucine in treatment of obesity and the metabolic syndrome. *J. Nutr.* **2006**, *136* (1), 319S−23S.

(58) Shearer, J.; Duggan, G.; Weljie, A.; Hittel, D. S.; Wasserman, D. H.; Vogel, H. J. Metabolomic profiling of dietary-induced insulin resistance in the high fat-fed C57BL/6J mouse. *Diabetes, Obes. Metab.* **2008**, *10* (10), 950−958.

(59) Adams, S. H. Emerging Perspectives on Essential Amino Acid Metabolism in Obesity and the Insulin-Resistant State. *Adv. Nutr.* **2011**, *2* (6), 445−456.

DOI: 10.1021/acs.jproteome.6b00860
J. Proteome Res. 2017, 16, 1483−1491

9.2 COMUNICACIONES EN CONGRESOS

Durante el periodo predoctoral se han presentado 8 comunicaciones en congresos, de los cuales 3 corresponden a comunicaciones en forma oral y 5 en forma de póster.

9.2.1 Comunicaciones Orales

Comunicación oral 1: Identificación de biomarcadores relacionados con factores de riesgo de enfermedad cardiovascular y el efecto de una intervención dietética mediante una aproximación metabolómica. Comunicación presentada en la *VII Jornada de Recerca*. Universidad de Barcelona, Facultad de Farmacia. 26 de mayo de 2014. Barcelona, España.

Comunicación oral 2: Nueva ciencia, nuevos métodos para el estudio de la interacción dieta-diabetes. Comunicación presentada en la *ExpAliments 2014*. Universidad de Barcelona, Campus de Alimentación Torribera. 18 de octubre de 2014. Santa Coloma de Gramanet, España.

Comunicación oral 3: Enrique Almanza-Aguilera, Mireia Urpi-Sarda y Cristina Andrés Lacueva. Perfil metabolómico urinario asociado a una alta adherencia a la Dieta Mediterránea en individuos con alto riesgo cardiovascular. Comunicación presentada en ocasión del *V Simposio Becarios CONACYT en Europa*. 2016. Estrasburgo, Francia.

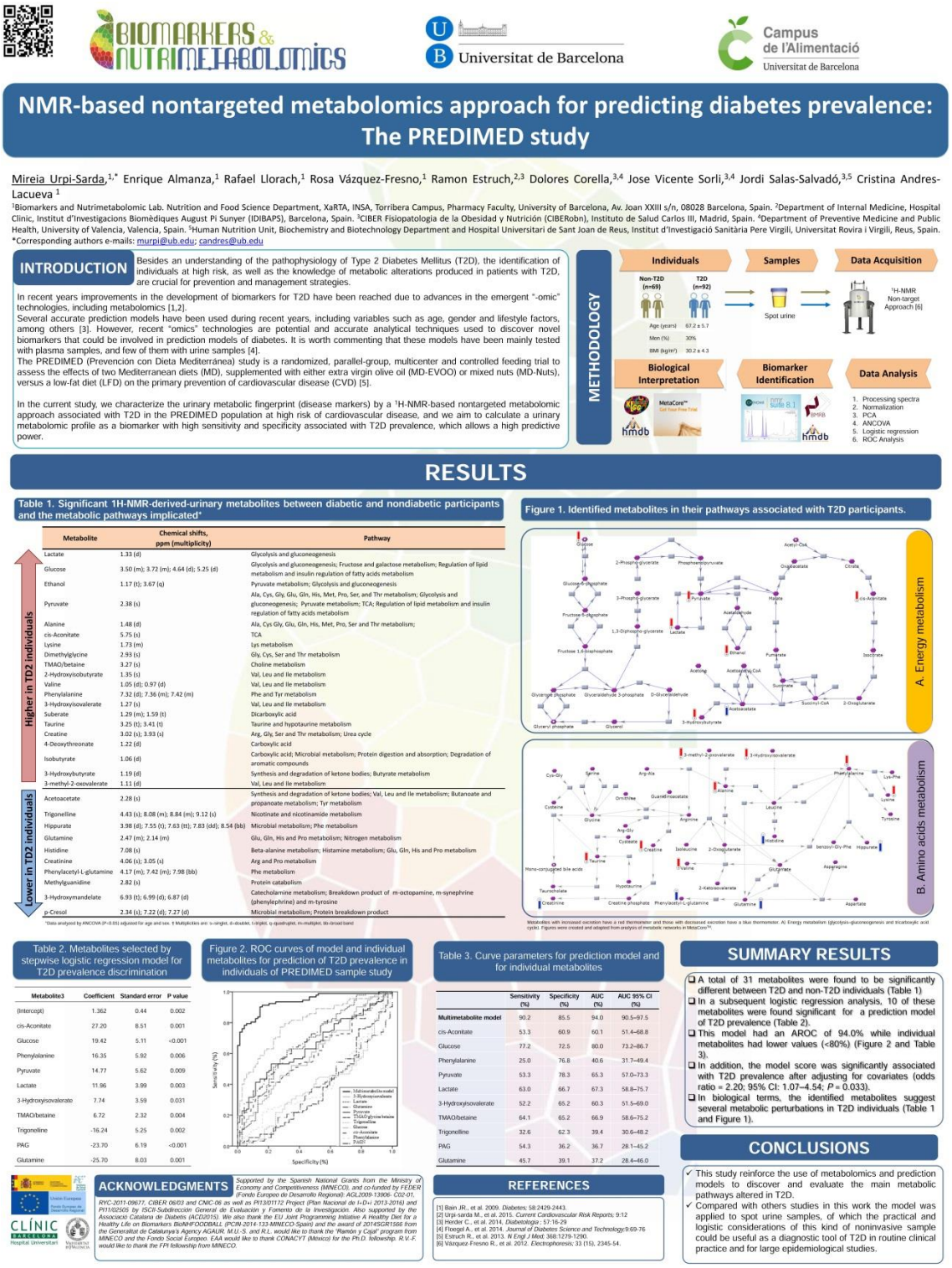
Comunicación oral 4: Enrique Almanza-Aguilera, Mireia Urpi-Sarda, Rafael Llorach, Rosa Vazquez-Fresno, Mar Garcia-Aloy, Francisco Madrid-Gambin, Ramón Estruch, Dolores Corella, Cristina Andres-Lacueva. Microbial metabolites and the exploration of the adherence to a Mediterranean dietary pattern by ^{1}H-NMR-based untargeted metabolomics approach. Comunicación presentada en ocasión del *XIV Congreso de la Asociación Catalana de Diabetes*. . 16-17 de marzo de 2017. Badalona, España.

Comunicación oral 5: Mireia Urpi-Sarda, Enrique Almanza-Aguilera, Rafael Llorach, Rosa Vázquez-Fresno, Ramón Estruch, Dolores Corella, Marisa Guillen, Jordi Salas-Salvadó, Cristina Andres-Lacueva. Urinary metabolomic biomarkers of prevalence of diabetes in predimed study. Comunicación presentada en ocasión del *XIV Congreso de la Asociación Catalana de Diabetes*. 16-17 de marzo de 2017. Badalona, España.

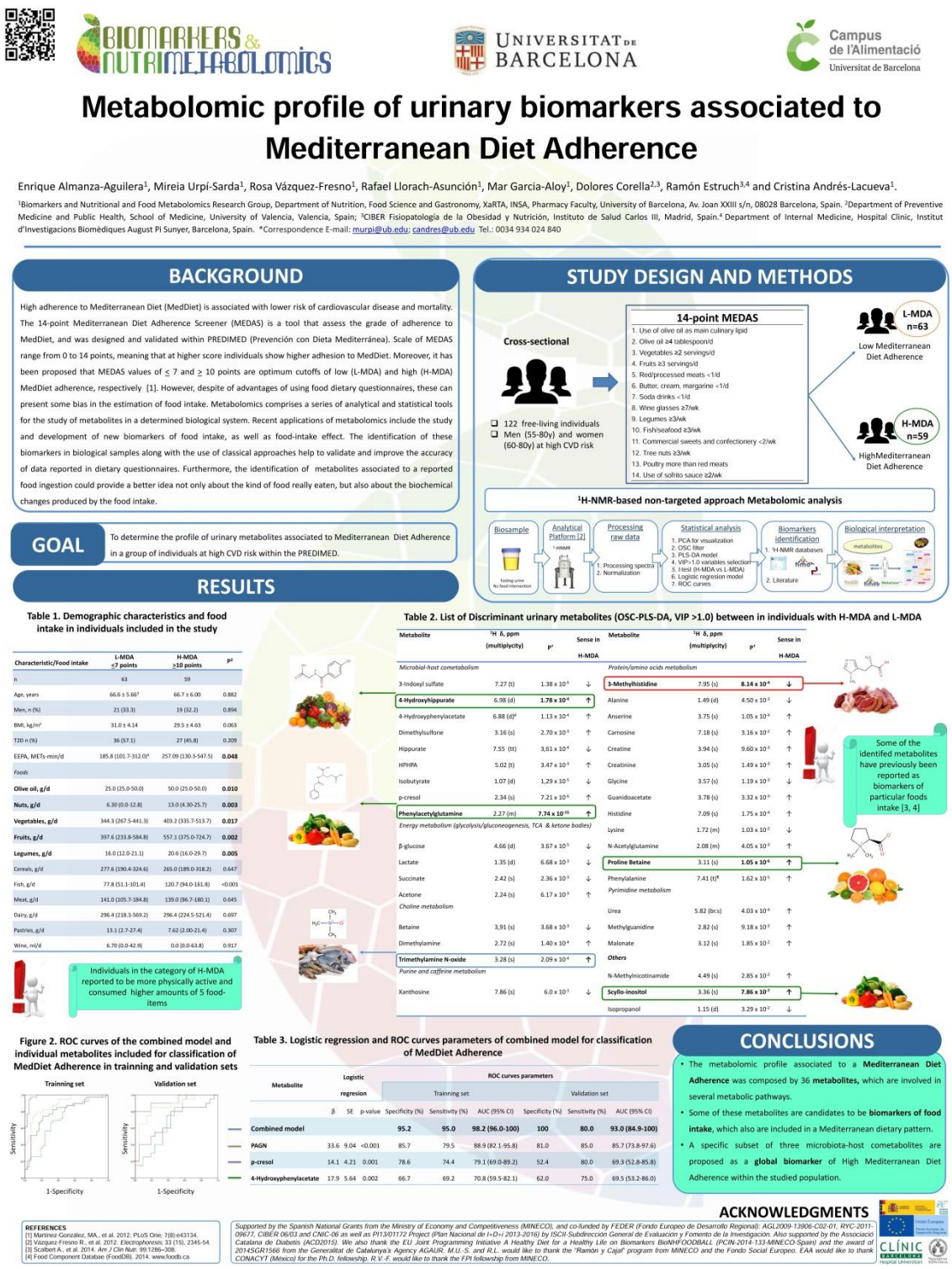
9.2.1 Comunicaciones en forma de póster

Comunicación póster 1: Enrique Almanza-Aguilera, Mireia Urpi-Sarda, Rosa Vázquez-Fresno, Rafael Llorach-Asunción, Mar García-Aloy, Dolores Corella, Ramón Estruch and Cristina Andres-Lacueva. Metabolomic biomarkers of adherence to Mediterranean Diet in subjects at high cardiovascular risk. Póster presentado en ocasión de la *NuGoweek 2015*. Barcelona, España.

Comunicación póster 2: Mireia Urpi-Sarda, Enrique Almanza, Rafael Llorach, Rosa Vázquez-Fresno, Ramon Estruch, Dolores Corella, Jose Vicente Sorli, Jordi Salas-Salvadó, Cristina Andres-Lacueva. NMR-based nontargeted metabolomics approach for predicting diabetes prevalence: The PREDIMED study. Póster presentado en ocasión de la *NuGoweek 2015*. Barcelona, España.



Comunicación póster 3: Enrique Almanza-Aguilera, Mireia Urpí-Sarda, Rosa Vázquez-Fresno, Rafael Llorach-Asunción, Mar Garcia-Aloy, Dolores Corella, Ramón Estruch and Cristina Andrés-Lacueva. Metabolomic profile of urinary biomarkers associated to Mediterranean Diet Adherence. Póster presentado en ocasión del *XI international conference on the Mediterranean Diet*. 2016. Barcelona, España.



Metabolomic profile of urinary biomarkers associated to Mediterranean Diet Adherence

Enrique Almanza-Aguilera[1], Mireia Urpí-Sarda[1], Rosa Vázquez-Fresno[1], Rafael Llorach-Asunción[1], Mar Garcia-Aloy[1], Dolores Corella[2,3], Ramón Estruch[3,4] and Cristina Andrés-Lacueva[1].

[1]Biomarkers and Nutritional and Food Metabolomics Research Group, Department of Nutrition, Food Science and Gastronomy, XaRTA, INSA, Pharmacy Faculty, University of Barcelona, Av. Joan XXIII s/n, 08028 Barcelona, Spain. [2]Department of Preventive Medicine and Public Health, School of Medicine, University of Valencia, Valencia, Spain; [3]CIBER Fisiopatología de la Obesidad y Nutrición, Instituto de Salud Carlos III, Madrid, Spain.[4] Department of Internal Medicine, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi Sunyer, Barcelona, Spain. *Correspondence E-mail: murpi@ub.edu; candres@ub.edu Tel.: 0034 934 024 840

BACKGROUND

High adherence to Mediterranean Diet (MedDiet) is associated with lower risk of cardiovascular disease and mortality. The 14-point Mediterranean Diet Adherence Screener (MEDAS) is a tool that assess the grade of adherence to MedDiet, and was designed and validated within PREDIMED (Prevención con Dieta Mediterránea). Scale of MEDAS range from 0 to 14 points, meaning that at higher score individuals show higher adhesion to MedDiet. Moreover, it has been proposed that MEDAS values of ≤ 7 and ≥ 10 points are optimum cutoffs of low (L-MDA) and high (H-MDA) MedDiet adherence, respectively [1]. However, despite of advantages of using food dietary questionnaires, these can present some bias in the estimation of food intake. Metabolomics comprises a series of analytical and statistical tools for the study of metabolites in a determined biological system. Recent applications of metabolomics include the study and development of new biomarkers of food intake, as well as food-intake effect. The identification of these biomarkers in biological samples along with the use of classical approaches help to validate and improve the accuracy of data reported in dietary questionnaires. Furthermore, the identification of metabolites associated to a reported food ingestion could provide a better idea not only about the kind of food really eaten, but also about the biochemical changes produced by the food intake.

GOAL

To determine the profile of urinary metabolites associated to Mediterranean Diet Adherence in a group of individuals at high CVD risk within the PREDIMED.

STUDY DESIGN AND METHODS

Cross-sectional

- 122 free-living individuals
- Men (55-80y) and women (60-80y) at high CVD risk

14-point MEDAS

1. Use of olive oil as main culinary lipid
2. Olive oil ≥4 tablespoon/d
3. Vegetables ≥2 servings/d
4. Fruits ≥3 servings/d
5. Red/processed meats <1/d
6. Butter, cream, margarine <1/d
7. Soda drinks <1/d
8. Wine glasses ≥7/wk
9. Legumes ≥3/wk
10. Fish/seafood ≥3/wk
11. Commercial sweets and confectionery <2/wk
12. Tree nuts ≥3/wk
13. Poultry more than red meats
14. Use of sofrito sauce ≥2/wk

L-MDA n=63 — Low Mediterranean Diet Adherence

H-MDA n=59 — HighMediterranean Diet Adherence

^{1}H-NMR-based non-targeted approach Metabolomic analysis

Biosample (Fasting urine, No food intervention) → Analytical Platform [2] (^{1}H-NMR) → Processing raw data (1. Processing spectra, 2. Normalization) → Statistical analysis (1. PCA for visualization, 2. OSC filter, 3. PLS-DA model, 4. VIP>1.0 variables selection, 5. t-test (H-MDA vs L-MDA), 6. Logistic regression model, 7. ROC curves) → Biomarkers identification (1. H-NMR databases, 2. Literature) → Biological interpretation (metabolites, Pathways, Metabolism)

RESULTS

Table 1. Demographic characteristics and food intake in individuals included in the study

Characteristic/Food intake	L-MDA ≤7 points	H-MDA ≥10 points	p[2]
n	63	59	
Age, years	66.6 ± 5.66[3]	66.7 ± 6.00	0.882
Men, n (%)	21 (33.3)	19 (32.2)	0.894
BMI, kg/m^2	31.0 ± 4.14	29.5 ± 4.63	0.063
T2D n (%)	36 (57.1)	27 (45.8)	0.209
EEPA, METs-min/d	185.8 (101.7-312.0)[4]	257.09 (130.3-547.5)	**0.048**
Foods			
Olive oil, g/d	25.0 (25.0-50.0)	50.0 (25.0-50.0)	**0.010**
Nuts, g/d	6.30 (0.0-12.8)	13.0 (4.30-25.7)	**0.003**
Vegetables, g/d	344.3 (267.5-441.3)	403.2 (335.7-513.7)	**0.017**
Fruits, g/d	397.6 (233.8-584.8)	557.1 (375.0-724.7)	**0.002**
Legumes, g/d	16.0 (12.0-21.1)	20.6 (16.0-29.7)	**0.005**
Cereals, g/d	277.6 (190.4-324.6)	265.0 (189.0-318.2)	0.647
Fish, g/d	77.8 (51.1-101.4)	120.7 (94.0-161.8)	<0.001
Meat, g/d	141.0 (105.7-184.8)	139.0 (96.7-180.1)	0.645
Dairy, g/d	296.4 (218.3-569.2)	296.4 (224.5-521.4)	0.697
Pastries, g/d	13.1 (2.7-27.4)	7.62 (2.00-21.4)	0.307
Wine, ml/d	6.70 (0.0-42.9)	0.0 (0.0-63.8)	0.917

Individuals in the category of H-MDA reported to be more physically active and consumed higher amounts of 5 food-items

Table 2. List of Discriminant urinary metabolites (OSC-PLS-DA, VIP >1.0) between in individuals with H-MDA and L-MDA

Metabolite	^{1}H δ, ppm (multiplycity)	p[1]	Sense in H-MDA
Microbial-host cometabolism			
3-Indoxyl sulfate	7.27 (t)	1.38×10^{-5}	↓
4-Hydroxyhippurate	6.98 (d)	$\mathbf{1.78 \times 10^{-4}}$	↑
4-Hydroxyphenylacetate	6.88 (d)[a]	1.13×10^{-4}	↑
Dimethylsulfone	3.16 (s)	2.70×10^{-5}	↑
Hippurate	7.55 (tt)	3.61×10^{-4}	↓
HPHPA	5.02 (t)	3.47×10^{-3}	↑
Isobutyrate	1.07 (d)	1.29×10^{-5}	↓
p-cresol	2.34 (s)	7.21×10^{-6}	↑
Phenylacetylglutamine	2.27 (m)	$\mathbf{7.74 \times 10^{-16}}$	↑
Energy metabolism (glycolysis/gluconeogenesis, TCA & ketone bodies)			
β-glucose	4.66 (d)	3.67×10^{-5}	↓
Lactate	1.35 (d)	6.68×10^{-3}	↓
Succinate	2.42 (s)	2.36×10^{-3}	↓
Acetone	2.24 (s)	6.17×10^{-3}	↑
Choline metabolism			
Betaine	3,91 (s)	3.68×10^{-3}	↓
Dimethylamine	2.72 (s)	1.40×10^{-4}	↑
Trimethylamine N-oxide	3.28 (s)	2.09×10^{-4}	↑
Purine and caffeine metabolism			
Xanthosine	7.86 (s)	6.0×10^{-3}	↓

Metabolite	^{1}H δ, ppm (multiplycity)	p[1]	Sense in H-MDA
Protein/amino acids metabolism			
3-Methylhistidine	7.95 (s)	$\mathbf{8.14 \times 10^{-4}}$	↓
Alanine	1.49 (d)	4.50×10^{-2}	↓
Anserine	3.75 (s)	1.05×10^{-4}	↑
Carnosine	7.18 (s)	3.16×10^{-2}	↑
Creatine	3.94 (s)	9.60×10^{-3}	↑
Creatinine	3.05 (s)	1.49×10^{-2}	↑
Glycine	3.57 (s)	1.19×10^{-3}	↓
Guanidoacetate	3.78 (s)	3.32×10^{-3}	↑
Histidine	7.09 (s)	1.75×10^{-4}	↑
Lysine	1.72 (m)	1.03×10^{-2}	↓
N-Acetylglutamine	2.08 (m)	4.05×10^{-3}	↑
Proline Betaine	3.11 (s)	$\mathbf{1.05 \times 10^{-6}}$	↑
Phenylalanine	7.41 (t)[a]	1.62×10^{-5}	↑
Pyrimidine metabolism			
Urea	5.82 (brs)	4.03×10^{-6}	↑
Methylguanidine	2.82 (s)	9.18×10^{-3}	↑
Malonate	3.12 (s)	1.85×10^{-2}	↑
Others			
N-Methylnicotinamide	4.49 (s)	2.85×10^{-2}	↑
Scyllo-inositol	3.36 (s)	$\mathbf{7.86 \times 10^{-7}}$	↑
Isopropanol	1.15 (d)	3.29×10^{-2}	↓

Some of the identifed metabolites have previously been reported as biomarkers of particular foods intake [3, 4]

Figure 2. ROC curves of the combined model and individual metabolites included for classification of MedDiet Adherence in trainning and validation sets

Trainning set — Validation set

Sensitivity — 1-Specificity

Table 3. Logistic regression and ROC curves parameters of combined model for classification of MedDiet Adherence

Metabolite	Logistic regresion			ROC curves parameters					
				Trainning set			Validation set		
	β	SE	p-value	Specificity (%)	Sensitivity (%)	AUC (95% CI)	Specificity (%)	Sensitivity (%)	AUC (95% CI)
Combined model				**95.2**	**95.0**	**98.2 (96.0-100)**	**100**	**80.0**	**93.0 (84.9-100)**
PAGN	33.6	9.04	<0.001	85.7	79.5	88.9 (82.1-95.8)	81.0	85.0	85.7 (73.8-97.6)
p-cresol	14.1	4.21	0.001	78.6	74.4	79.1 (69.0-89.2)	52.4	80.0	69.3 (52.8-85.8)
4-Hydroxyphenylacetate	17.9	5.64	0.002	66.7	69.2	70.8 (59.5-82.1)	62.0	75.0	69.5 (53.2-86.0)

CONCLUSIONS

- The metabolomic profile associated to a **Mediterranean Diet Adherence** was composed by 36 **metabolites,** which are involved in several metabolic pathways.
- Some of these metabolites are candidates to be **biomarkers of food intake**, which also are included in a Mediterranean dietary pattern.
- A specific subset of three microbiota-host cometabolites are proposed as a **global biomarker** of High Mediterranean Diet Adherence within the studied population.

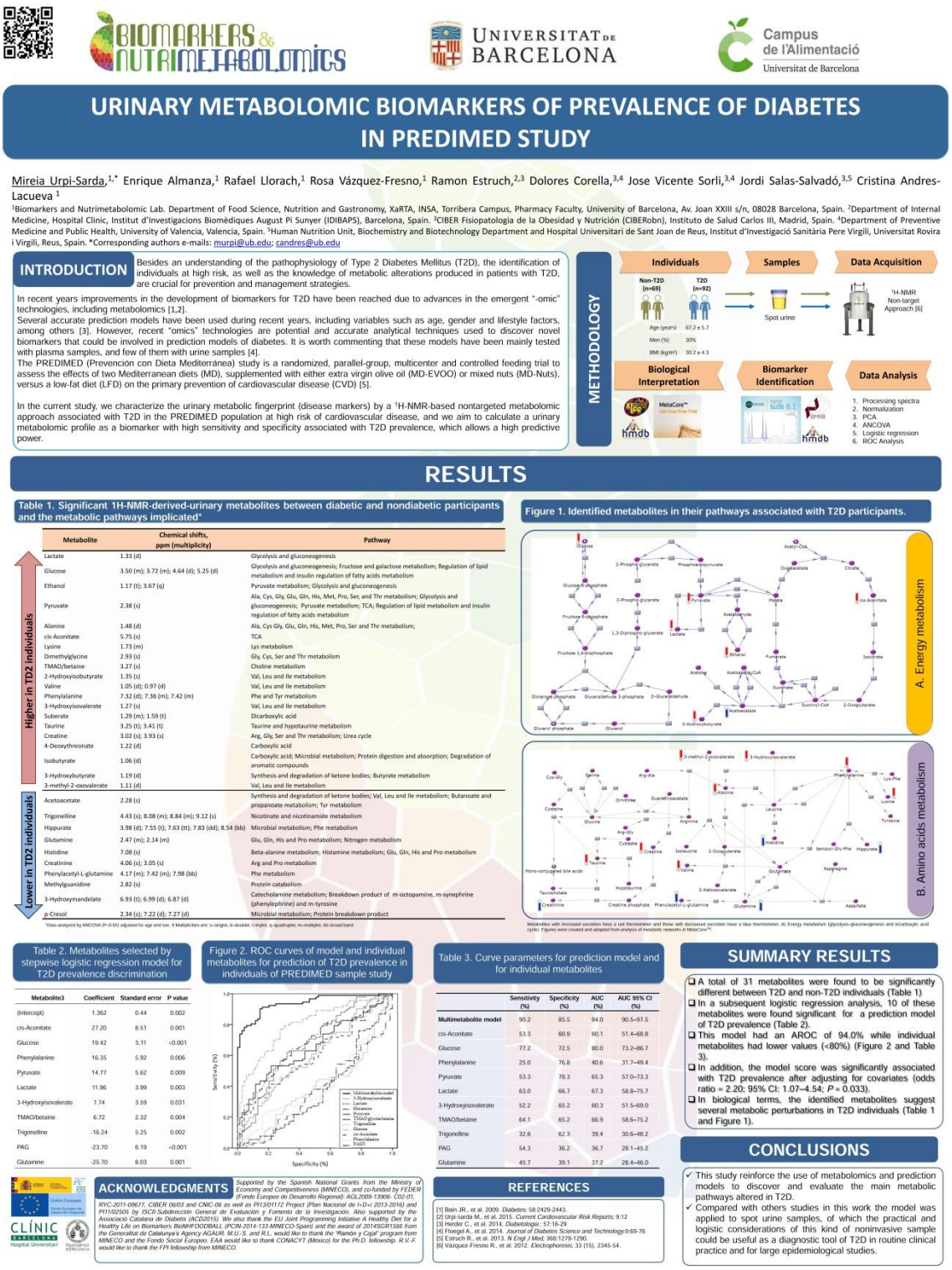
REFERENCES

[1] Martinez-González, MA., et al. 2012. PLoS One; 7(8):e43134.
[2] Vázquez-Fresno R., et al. 2012. *Electrophoresis*; 33 (15), 2345-54.
[3] Scalbert A., et al. 2014. *Am J Clin Nutr*; 99:1286–308.
[4] Food Component Database (FoodDB). 2014. www.foodb.ca.

ACKNOWLEDGMENTS

Supported by the Spanish National Grants from the Ministry of Economy and Competitiveness (MINECO), and co-funded by FEDER (Fondo Europeo de Desarrollo Regional): AGL2009-13906-C02-01, RYC-2011-09677, CIBER 06/03 and CNIC-06 as well as PI13/01172 Project (Plan Nacional de I+D+i 2013-2016) by ISCII-Subdirección General de Evaluación y Fomento de la Investigación. Also supported by the Associació Catalana de Diabetis (ACD2015). We also thank the EU Joint Programming Initiative A Healthy Diet for a Healthy Life on Biomarkers BioNHFOODBALL (PCIN-2014-133-MINECO-Spain) and the award of 2014SGR1566 from the Generalitat de Catalunya's Agency AGAUR. M.U.-S. and R.L. would like to thank the "Ramón y Cajal" program from MINECO and the Fondo Social Europeo. EAA would like to thank CONACYT (Mexico) for the Ph.D. fellowship. R.V.-F. would like to thank the FPI fellowship from MINECO.

Comunicación póster 4: Mireia Urpi-Sarda, Enrique Almanza, Rafael Llorach, Rosa Vázquez-Fresno, Ramón Estruch, Dolores Corella, José Vicente Sorli, Jordi Salas-Salvadó, Cristina Andres-Lacueva. Urinary Metabolomics biomarkers of prevalence od Diabetes in PREDIMED study. Póster presentado en ocasión del *XI international conference on the Mediterranean Diet*. 2016. Barcelona, España.



Comunicación póster 5: Enrique Almanza-Aguilera, Mireia Urpi-Sarda, Rafael Llorach, Rosa Vazquez-Fresno, Mar Garcia-Aloy, Francisco Madrid-Gambin, Ramon Estruch, Dolores Corella, Cristina Andres-Lacueva. Microbial metabolites and the exploration of the adherence to a Mediterranean dietary pattern by ^{1}H-NMR-based untargeted metabolomics approach. Póster presentado en ocasion de la *Max Rubner Conference* 2016. Karlsruhe, Alemania.

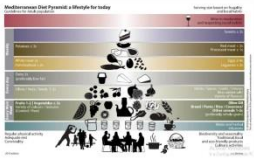
BIOMARKERS & NUTRIMETABOLOMICS — UNIVERSITAT DE BARCELONA — Campus de l'Alimentació, Universitat de Barcelona

Microbial metabolites and the exploration of the adherence to a Mediterranean dietary pattern by ^{1}H-NMR-based untargeted metabolomics approach

Enrique Almanza-Aguilera[1], Mireia Urpi-Sarda*[1], Rafael Llorach[1], Rosa Vazquez-Fresno[1], Mar Garcia-Aloy[1], Francisco Madrid-Gambin[1], Ramon Estruch[2], Dolores Corella[3], Cristina Andres-Lacueva*[1]

[1]Biomarkers and Nutrimetabolomics Laboratory, Nutrition, Food Science and Gastronomy Department, XaRTA, INSA, Campus Torribera, Faculty of Pharmacy and Food Science, University of Barcelona, Spain. [2]Department of Internal Medicine, IDIBAPS, University of Barcelona, Barcelona, Spain. [3]Department of Preventive Medicine, University of Valencia, Valencia, Spain

*Corresponding authors: murpi@ub.edu; candres@ub.edu



BACKGROUND

- The assessment of dietary pattern compliance is an essential component in nutritional research and epidemiology.
- To our knowledge, only dietary indexes or scores through dietary questionnaires have defined the adherence to the Mediterranean dietary (MedDiet) pattern.
- The 14-point Mediterranean Diet Adherence Screener (MEDAS) is a tool that assess the grade of adherence to MedDiet, and was designed and validated within PREDIMED (Prevención con Dieta Mediterránea) [1].
- Biological biomarkers are potentially powerful tools to assess MedDiet adherence and a better understanding of its health benefits.

GOAL

To identify a biomarker associated with high adherence to a MedDiet pattern through a ^{1}H-NMR-based untargeted metabolomics approach.

INDIVIDUALS

- 122 free-living individuals
- Men and women at high CVD risk

Table 1. The 14 point based Mediterranean Diet Assessment Screener score

14-point MEDAS	
1. Use of olive oil as main culinary lipid	8. Wine glasses ≥7/wk
2. Olive oil ≥4 tablespoon/d	9. Legumes ≥3/wk
3. Vegetables ≥2 servings/d	10. Fish/seafood ≥3/wk
4. Fruits ≥3 servings/d	11. Commercial sweets <2/wk
5. Red/processed meats <1/d	12. Tree nuts ≥3/wk
6. Butter, cream, margarine <1/d	13. Poultry more than red meats
7. Soda drinks <1/d	14. Use of sofrito sauce ≥2/wk

STUDY DESIGN

(n=63) **L-MDA** Low MediDiet Adherence MEDAS ≤ 7 points

(n=59) **H-MDA** High MediDiet Adherence MEDAS ≥ 10 points

METABOLOMICS WORKFLOW

1. Analytical platform [2]

2. Processing raw data
- Processing spectra
- Normalization data

3. Statistical analysis
- OSC-PLS-DA mode
- *t*-test (H-MDA vs L-MDA)
- Logistic regresion model
- ROC curves

4. Biomarkers identification
- NMR databases
- NMR literature

5. Biological interpretation
- HMDB
- FOODB
- KEGG

RESULTS

35 metabolites were found as discriminant between H-MDA and L-MDA (Table 2), including higher concentrations of metabolites related to the intake of fruits (¶) and vegetables (¥), fish (§) or white meat (*), antioxidant metabolites, microbial metabolites and lower concentration of metabolites related to energy metabolism pathway [3,4].

Table 2. List of Discriminant urinary metabolites (OSC-PLS-DA, VIP >1.0) between in individuals with H-MDA and L-MDA

Metabolite	^{1}H δ, ppm (multiplycity)	p-value	Sense in H-MDA	Metabolite	^{1}H δ, ppm (multiplycity)	p-value	Sense in H-MDA
Microbial-host cometabolism				*Protein/amino acids metabolism*			
3-Indoxyl sulfate	7.27 (t)	1.38×10^{-5}	↓	**3-Methylhistidine**	7.95 (s)	8.14×10^{-4}	↓
4-Hydroxyhippurate (¶)	6.98 (d)	1.78×10^{-4}	↑	Alanine	1.49 (d)	4.50×10^{-3}	↓
4-Hydroxyphenylacetate (4-HPA)	6.88 (d)	1.13×10^{-4}	↑	**Anserine (*)**	3.75 (s)	1.05×10^{-4}	↑
Dimethylsulfone	3.16 (s)	2.70×10^{-3}	↑	Carnosine	7.18 (s)	3.16×10^{-2}	↑
Hippurate	7.55 (tt)	3.61×10^{-4}	↓	Creatine	3.94 (s)	9.60×10^{-3}	↑
HPHPA	5.02 (t)	3.47×10^{-3}	↑	Creatinine	3.05 (s)	1.49×10^{-3}	↑
Isobutyrate	1.07 (d)	1.29×10^{-5}	↓	Glycine	3.57 (s)	1.19×10^{-3}	↓
p-cresol	2.34 (s)	7.21×10^{-6}	↑	Guanidoacetate	3.78 (s)	3.32×10^{-3}	↑
Phenylacetylglutamine (PAGN) (¥)	2.27 (m)	7.74×10^{-15}	↑	Histidine	7.09 (s)	1.75×10^{-4}	↑
Energy metabolism (glycolysis/gluconeogenesis, TCA & ketone bodies)				Lysine	1.72 (m)	1.03×10^{-2}	↓
β-glucose	4.66 (d)	3.67×10^{-5}	↓	N-Acetylglutamine	2.08 (m)	4.05×10^{-3}	↑
Lactate	1.35 (d)	6.68×10^{-3}	↓	**Proline Betaine (¶)**	3.11 (s)	1.05×10^{-6}	↑
Succinate	2.42 (s)	2.36×10^{-3}	↓	Phenylalanine	7.41 (t)[1]	1.62×10^{-5}	↑
Acetone	2.24 (s)	6.17×10^{-3}	↑	*Pyrimidine metabolism*			
Choline metabolism				Urea	5.82 (br.s)	4.03×10^{-4}	↑
Betaine	3,91 (s)	3.68×10^{-3}	↓	Methylguanidine	2.82 (s)	9.18×10^{-3}	↑
Dimethylamine	2.72 (s)	1.40×10^{-4}	↑	Malonate	3.12 (s)	1.85×10^{-2}	↑
Trimethylamine N-oxide (§)	3.28 (s)	2.09×10^{-4}	↑	***Others***			
Purine and caffeine metabolism				N-Methylnicotinamide	4.49 (s)	2.85×10^{-2}	↑
Xanthosine	7.86 (s)	6.0×10^{-3}	↓	**Scyllo-inositol (¥)**	3.36 (s)	7.86×10^{-7}	↑
				Isopropanol	1.15 (d)	3.29×10^{-2}	↓

Table 3. Demographics and food intake in individuals included in the study

	L-MDA ≤7 points	H-MDA ≥10 points	p[2]
Age, years	66.6 ± 5.66[3]	66.7 ± 6.00	0.882
BMI, kg/m^2	31.0 ± 4.14	29.5 ± 4.63	0.063
T2D n (%)	36 (57.1)	27 (45.8)	0.209
EEPA, METs-min/d	185.8 (101.7-312.0)[4]	257.09 (130.3-547.5)	**0.048**
Foods			
Olive oil, g/d	25.0 (25.0-50.0)	50.0 (25.0-50.0)	**0.010**
Nuts, g/d	6.30 (0.0-12.8)	13.0 (4.30-25.7)	**0.003**
Vegetables, g/d	344.3 (267.5-441.3)	403.2 (335.7-513.7)	**0.017**
Fruits, g/d	397.6 (233.8-584.8)	557.1 (375.0-724.7)	**0.002**
Legumes, g/d	16.0 (12.0-21.1)	20.6 (16.0-29.7)	**0.005**
Fish, g/d	77.8 (51.1-101.4)	120.7 (94.0-161.8)	**<0.001**

Individuals in the category of **H-MDA** reported to be more physically active and consumed **higher amounts of 6 typical food-items of MedDiet**

Table 4. Logistic regression and ROC curves parameters of combined model for classification of MedDiet Adherence

Metabolite	Logistic regresion			ROC curves parameters		
	β	SE	p-value	Specificity (%)	Sensitivity (%)	AUC (95% CI)
Multimetabolite model				90.5	95.0	96.9 (92.7–100)
PAGN	42.9	11.8	<0.001	81.0	80.0	89.0 (79.3–98.8)
p-cresol sulfate	10.5	2.99	<0.001	85.7	60.0	81.4 (68.0–94.8)
4-HPA	18.0	5.50	<0.001	66.7	80.0	79.0 (65.1–93.0)

The prediction model included **three microbial metabolites**, having together a **strong ability to discriminate between H-MDA and L-MDA** (Tablle 4). Moreover, the prediction model showed significant correlation with the intake of vegetables, fruits, legumes and fish (r=0.3, *p*<0.05).

CONCLUSIONS

- These results remark the **role of microbiota and microbial metabolites** in the study of the **biomarkers related to high adherence to a Mediterranean dietary pattern** which is rich in vegetable-derived foods.
- The study of these biological biomarkers is a more specific tool and complementary to traditional indexes/scores.
- This study may assess and aid to nutritional epidemiology in future associations between adherence to dietary patterns and prevention of diseases.

Supported by the Spanish National Grants from the Ministry of Economy and Competitiveness (MINECO), and co-funded by FEDER (Fondo Europeo de Desarrollo Regional): AGL2009-13906-C02-01, RYC-2011-09677, CIBER 06/03 and CNIC-08 as well as PI13/01172 Project (Plan Nacional de I+D+i 2013-2016) by ISCIII-Subdirección General de Evaluacion y Fomento de la Investigacion. Also supported by the Associació Catalana de Diabetis (ACD2015). We also thank the EU Joint Programming Initiative A Healthy Diet for a Healthy Life on Biomarkers BioNHFOODBALL (PCIN-2014-133-MINECO-Spain) and the award of 2014SGR1566 from the Generalitat de Catalunya's Agency AGAUR. M.U.-S. and R.L. would like to thank the "Ramón y Cajal" program from MINECO and the Fondo Social Europeo. EAA would like to thank CONACYT (México) for the Ph.D. fellowship. R.V.-F. would like to thank the FPI fellowship from MINECO.

REFERENCES
[1] Martínez-González, MA., et al. 2012. PLoS One; 7(8):e43134.
[2] Vázquez-Fresno R., et al. 2012. *Electrophoresis*; 33 (15), 2345-54.
[3] Scalbert A., et al. 2014. *Am J Clin Nutr.* 99:1286–308.
[4] Food Component Database (FoodDB). 2014. www.foodb.ca.