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.*

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 1H-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.

Keywords: Microbiota; Biomarkers; Metabolomics; Dietary patterns; Mediterranean diet; High adherence

Abbreviations: AUC, area under the curve; CVD, cardiovascular disease; FFQ, food frequency questionnaire; 1H-NMR, proton nuclear magnetic resonance; HMDA, 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).

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 1H-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/m2) ≥40, and allergy or intolerance to olive oil or nuts. The trial was registered at http://www.controlledtrials. 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 scorewas calculated for all participants and usedfor their subsequent stratification, which was done by using the proposed cutoff values previously reported in

123 the PREDIMEDstudy. [15]:MedDiet score ≤7 indicated low MedDiet adherence (L-MDA) (n=63),
124 and MedDiet score ≥10 (n=56) indicated high MDA (H-MDA). All participants were also asked to
125 complete a validated semiquantitative 137-item food frequency questionnaire (FFQ) [22] and the
126 Spanish version of the Minnesota Leisure-Time Physical Activity Questionnaire [23]. Trained
127 dieticians in the PREDIMED study administered all questionnaires, including the PREDIMED
128 MedDiet score questionnaire. The nutrient composition and energy intakes were calculated from the
129 FFQ data by using Spanish food composition tables [24]. Also, anthropometrical measurements were
130 taken directly by qualified nurses.

131 *2.3. Metabolite profiling*

132 1H-NMR analysis and spectra processing were performed by following previous methodology [8].
133 Briefly, the urine samples were thawed, vortexed and centrifuged at 13,200 rpm for 5 min. From each
134 supernatant, a volume of 300 µl was taken and diluted with 200 µl of H2O/D2O and mixed with a
135 buffer solution [8]. The optimized pH of the buffer was set at 7.0, with a potassium deuteroxide
136 solution, to minimize variations in the chemical shifts of the 1H-NMR resonances. This mixture was
137 transferred to a 5-mm NMR tube. 1H-NMR spectra acquisition was performed using a Varian-Inova
138 500-MHz NMR spectrometer with presaturation of the water resonance using a NOESYPRESAT
139 pulse sequence. The spectra data processed were binned to 1165 variables with bin widths of 0.005
140 ppm and integrated with ACD/NMR Processor 12.0 software (Advanced Chemistry Development,
141 Toronto, Canada). The spectral region containing water (δ 4.75–5.00 ppm) was excluded before
142 normalization to avoid spectral interference. Integrated spectra were row-wise normalized by sum
143 using MetaboAnalyst 3.0 (www.metaboanalyst.ca), a web server designed to permit comprehensive
144 metabolomics data analysis. Metabolites were identified following a multistep procedure as
145 previously reported [8]. This multistep includes (a) comparison of the experimentalNMRspectra with
146 those in the library of ChenomxSuite 8.1 profiler software (Chenomx, Inc., Edmonton, Canada),
147 which includes access to the Human Metabolome Database (HMDB) library [25]. Other databases
148 such as the Biological Magnetic Resonance Data Bank and the MadisonMetabolomics

149 ConsortiumDatabase were also consulted along with existing NMR-basedmetabolomics literature.
150 (b) At the same time, the NMR peak assignments were correlated using Pearson's correlation
151 coefficient ($r \geq 0.7$, $P \leq .05$) to confirmthemultiplicity and identify clusters of metabolites which were
152 then compared to databases with NMR data [8,25]. Biological interpretation was made by consulting
153 the HMDB and Kyoto Encyclopedia of Genes and Genomes databases. *2.4. Statistical analyses*
154 Demographic characteristics, medication usage and dietary intake were compared between groups by
155 conducting Student's t tests and $\chi 2$ analyses for continuous and categorical variables, respectively.
156 Variables with a non-normal distribution were log transformed before analyses. Multivariate data
157 analysis was performed using SIMCAP+ 13.0 (Umetrics, Umeå, Sweden) software. Data sets
158 containing the integrated NMR spectral bins were log transformed and Pareto scaled before
159 performing a principal component analysis to explore the quality of data acquisition. An orthogonal
160 signal correction (OSC) filter was applied to the data sets in order to reduce the variability not
161 associated with the dietary classification [8], in this case, the category of adherence to a MedDiet.
162 Afterward, a partial least squares discriminant analysis (PLS-DA) was performed to examine the
163 difference in metabolite profile between subjects at H-MDA and L-MDA. The quality and validation
164 of resultant model were also appraised, the first through the R2Y (cum) and Q2 (cum) parameters,
165 and the second with a permutation test (n=200). Discriminant variables between groups were selected
166 based on their variable importance in projection (VIP) value N1.0,which is a generally accepted
167 threshold [26]. The normality of discriminant variables was assessed using a Kolmogorov–Smirnov
168 testwith Lilliefors significance correction, and additionally Mann–Whitney or independent Student's
169 t tests were performed according to the normality of the data. To account for multiple comparisons
170 in the metabolomic analysis,we used a corrected P value with the Benjamini–Hochberg procedure.
171 Aiming to identify the metabolites with the best discriminant capability between H-MDA and L-
172 MDA, we designed and assessed the performance of a prediction model with the H-MDA as
173 dependent variable and the discriminant metabolites identified from the OSCPLS- DA model as
174 independent variables [6]. To this end, first the data set of individuals (n=119) was randomly split
175 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/m2, 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 (Pb.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 R2Y (cum) and Q2Y (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 intercept R2 and Q2 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 N1.0 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 fromthe OSC-PLS-DAmodel 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 fromthe 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 combinedmodel 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

229 and L-MDA) in comparisonwith the use of each metabolite individually. In the analysis of
230 correlation, the combined metabolite model showed a strong correlation with the MedDiet score
231 (r=0.7; Pb.001), aswell aswith the intake of vegetables, fruits, legumes, fish (r=0.2–0.3; Pb.01)
232 (Supplementary Table 3) and dietary fiber (r=0.3; Pb.01).Otherwise, the three individualmetabolites
233 had good correlation with the MedDiet score (r=0.3–0.6; Pb.001) but weaker or not significant
234 correlations with the intake of individual foods.

235 **4. Discussion**

236 In the present metabolomic study, we identified the urinary metabolite profile consisting of 34
237 metabolites that enable discrimination between 2 groups of individuals with high or low adherence
238 to the score of Mediterranean diet adherence validated in the PREDIMED study [13]. Participants in
239 the L-MDA group had a cutoff of ≤7 points and those in the H-MDAgroup ≥10 points, as previously
240 proposed [15]. The set of metabolites that discriminated between H-MDA and LMDA suggested the
241 metabolic modulation of the MedDiet. These metabolites are involved in multiple molecular
242 mechanisms and metabolic pathways, which together provide a holistic view of variations in the urine
243 metabolome due to the effect of follow this dietary pattern. Some of these metabolites have previously
244 been proposed as putative biomarkers of single food intake and also related to foods included in
245 MedDiet, as well as up- and down-regulated endogenous metabolites. We found that the higher
246 excretion of proline betaine and scyllo-inositol in H-MDA correlated significantly with the intake of
247 citrus fruits (r=0.36 and r=0.35; Pb.001, respectively), which is consistent with previous studies that
248 have proposed them as markers of citrus fruit intake [27]. Recently, in our previous work, we also
249 found positive correlations of proline betaine signals with orange consumption in long-term MedDiet
250 intervention and low-fat diet groups in a subsample of nondiabetic participants of the PREDIMED
251 study [8]. In this regard, in the H-MDA group, we also identified other metabolites derived from the
252 microbial metabolism: 4-hydroxyhippurate, 3-(3-hydroxyphenyl)-3-hydroxypropanoate and 4-
253 hydroxyphenylacetate. These metabolites have been described after interventions with mixed red
254 wine/grape juice extracts [28], as well as linked to the intake of other polyphenol-rich foods such as
255 cocoa and almond, among others (http://phenol-explorer.eu/). It is interesting to note that these

256 microbial metabolites were not found in our previous work evaluating the intervention with MedDiet
257 [8]. Other metabolites have been related to the intake of meat or fish. Both dimethylamine and
258 trimethylamine-N-oxide (TMAO), which were higher in the H-MDA group than in the L-MDA, have
259 been related to the intake of fish and seafood [29,30]. Nevertheless, only TMAO had a significant
260 correlation with the intake of total fish (r=0.24, Pb.01) in our study sample, and this correlation was
261 not previously observed in the subset of samples of nondiabetics in the PREDIMED study [8]. In
262 addition, we identified some compounds related to the histidine pathway: carnosine, anserine and 3-
263 methylhistidine. While 3- methylhistidine and carnosine are proposed as good biomarkers of red meat
264 intake [31–33], anserine has alternatively been proposed as a marker of white meat (i.e., poultry) [34]
265 or salmon intake [35]. In our study, subjects in the H-MDA group excreted lower amounts of 3-
266 methylhistidine and higher amounts of carnosine and anserine than those in the L-MDA group. We
267 found significant correlations between anserine, with antioxidant properties, and the intake of white
268 meats (r=0.231, Pb.01) but not with fish (PN.05). In addition, we did not find significant correlations
269 between methylhistidine or carnosine and foods. Nevertheless, it should be noted that carnosine is a
270 normal constituent in human urine, which occurs naturally in the skeletal muscle of mammals and
271 has antioxidant properties and therapeutic potential against numerous diseases [36]. To the best of
272 our knowledge, there is limited information about which mechanisms could increase these
273 antioxidants in the organism, but a recent study has shown that dietary vitamin B6 could determine
274 the carnosine concentration in the skeletal muscle of rats [37]. In this regard, we observed
275 significantly higher ingested concentrations of vitamin B6 in subjects in the H-MDA group (2.9±0.7
276 mg/d) than in those in the LMDA group (2.4±0.5 mg/d, Pb.001). Moreover, 3-methylhistidine is a
277 metabolite of anserine and carnosine in the histidine pathway. 3- Methylhistidine is also found in
278 urine as anserine and carnosine, and its concentration was inversely correlated with them in our study
279 (r=−0.32, Pb.05; r=−0.27; Pb.05, respectively). Thus, the contribution of a MedDiet in the modulation
280 of the histidine pathway could be an interesting field to further explore. Besides the metabolites
281 related to the intake of foods, we also identified other endogenousmetabolites involved in
282 energymetabolism. In particular, individuals in the H-MDA group showed lower excretion of

283 glucose, lactate and succinate than those in the L-MDA group, whose metabolites are related to
284 pathways affected in diabetic patients and other diseases [38]. This finding is interesting because
285 although our population included diabetic individuals, their distribution and the use ofmedication
286 between the H-MDA and L-MDA groups were well balanced (Table 1). This fact suggested therefore
287 that, in comparison to the L-MDA group, individuals in the H-MDA group could show relatively
288 better glycemic control. Supporting this notion, several studies have previously found an inverse
289 association between adherence to a MedDiet and indices of glucose homeostasis in the general
290 population, including elderly people, and high-risk patients [39]. Previous works comparing
291 intervention with MedDiet and low-fat diet in nondiabetic PREDIMED participants did not find
292 changes in these metabolites [8]. After identifying 34 metabolites in the H-MDA pattern, we studied
293 the prediction of high adherence to MedDiet pattern. To improve the prediction of H-MDA, a model
294 with a combination of more than one discriminatory metabolite was developed. For this purpose,
295 population was split into training and validation sets. The model included PAGN, 4-
296 hydroxyphenylacetate and p-cresol. Interestingly and in accordance with this finding, in our previous
297 study, we also reported associations of PAGN and p-cresol with long-term interventions with
298 MedDiet [8]. Currently, the use of multimetabolite biomarkers is still limited. In a recent study,
299 Marklund et al. (2014) combined six serum metabolites to create a dietary biomarker score [19]. This
300 score was aimed at assessing the compliance with a healthy Nordic diet in a population with metabolic
301 syndrome (MetS) and subsequently used to analyze the effects of diet on cardiometabolic risk factors
302 among those individuals with highest compliance to this diet [19]. In our study, the inclusion of
303 PAGN, 4-hydroxyphenylacetate and p-cresol in the model, as well as the identification of other
304 metabolites derived from gut microbiota in the previous set of discriminant metabolites, highlights
305 the role of a MedDiet in the modulation of gut microbiota. Supporting this notion, De Filippis et al.
306 (2015) demonstrated the relationship between the level of MedDiet adherence, gut microbiota and
307 microbial metabolites [18]. Later, Haro et al. (2016) demonstrated that a long-term intervention with
308 a MedDiet partially restores the alteration in the gut microbiota composition in individuals with MetS
309 [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 postintervention 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 foodderived 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 PREDIMEDMedDiet 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 themain tool to assess dietary patterns adherence and, second, because this robust analytical platformallows 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 MINECOand 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 dietdisease 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.

386 [8] Vazquez-Fresno R, Llorach R, Urpi-Sarda M, Lupianez-Barbero A, Estruch R, Corella D, et al.
387 Metabolomic pattern analysis after Mediterranean diet intervention in a nondiabetic population: a 1-
388 and 3-year follow-up in the PREDIMED study. J Proteome Res 2015;14:531–40.
389 http://dx.doi.org/10.1021/pr5007894.

390 [9] Bondia-Pons I, Martinez JA, de la Iglesia R, Lopez-Legarrea P, Poutanen K, Hanhineva K, et al.
391 Effects of short- and long-term Mediterranean-based dietary treatment on plasma LC-QTOF/MS
392 metabolic profiling of subjects with metabolic syndrome features: the Metabolic Syndrome
393 Reduction in Navarra (RESMENA) randomized controlled trial. Mol Nutr Food Res 2015;59:711–
394 28. http://dx.doi. org/10.1002/mnfr.201400309.

395 [10] Altmaier E, Kastenmüller G, Römisch-Margl W, Thorand B, Weinberger KM, Illig T, et al.
396 Questionnaire-based self-reported nutrition habits associate with serum metabolism as revealed by
397 quantitative targeted metabolomics. Eur J Epidemiol 2011;26:145–56.
398 http://dx.doi.org/10.1007/s10654-010-9524-7.

399 [11] O'Sullivan A, Gibney MJ, Brennan L. Dietary intake patterns are reflected in metabolomic
400 profiles: potential role in dietary assessment studies. Am J Clin Nutr 2011;93:314–21.
401 http://dx.doi.org/10.3945/ajcn.110.000950.

402 [12] Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and
403 survival in a Greek population. N Engl J Med 2003;348: 2599–608.
404 http://dx.doi.org/10.1056/NEJMoa025039.

405 [13] Schroder H, Fito M, Estruch R, Martinez-Gonzalez MA, Corella D, Salas-Salvado J, et al. A
406 short screener is valid for assessing Mediterranean diet adherence among older Spanish men and
407 women. J Nutr 2011;141:1140–5. http://dx.doi.org/10. 3945/jn.110.135566.

408 [14] Gutierrez-Diaz I, Fernandez-Navarro T, Salazar N, Bartolome B, Moreno-Arribas MV, de
409 Andres-Galiana EJ, et al. Adherence to a Mediterranean diet influences the fecal metabolic profile of

410 microbial-derived phenolics in a Spanish cohort of middle-age and older people. J Agric Food Chem
411 2017;65:586–95. http://dx.doi. org/10.1021/acs.jafc.6b04408.

412 [15] Martinez-González MA, Garcia-Arellano A, Toledo E, Salas-Salvado J, Buil-Cosiales P, Corella
413 D, et al. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects:
414 the PREDIMED trial. PLoS One 2012;7:e43134. http://dx.doi.org/10.1371/journal.pone.0043134.

415 [16] Talegawkar SA, Bandinelli S, Bandeen-Roche K, Chen P, Milaneschi Y, Tanaka T, et al. A
416 higher adherence to a Mediterranean-style diet is inversely associated with the development of frailty
417 in community-dwelling elderly men and women. J Nutr 2012;142:2161–6.
418 http://dx.doi.org/10.3945/jn.112.165498.

419 [17] Guasch-Ferre M, Zheng Y, Ruiz-Canela M, Hruby A, Martinez-Gonzalez MA, Clish CB, et al.
420 Plasma acylcarnitines and risk of cardiovascular disease: effect of Mediterranean diet interventions.
421 Am J Clin Nutr 2016. http://dx.doi.org/10.3945/ ajcn.116.130492.

422 [18] De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, et al. High-level
423 adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated
424 metabolome. Gut 2015. http://dx.doi.org/10.1136/gutjnl-2015- 309957.

425 [19] Marklund M, Magnusdottir OK, Rosqvist F, Cloetens L, Landberg R, Kolehmainen M, et al. A
426 dietary biomarker approach captures compliance and cardiometabolic effects of a healthy Nordic diet
427 in individuals with metabolic syndrome. J Nutr 2014;144:1642–9.
428 http://dx.doi.org/10.3945/jn.114.193771.

429 [20] Martínez-González MÁ, Corella D, Salas-Salvadó J, Ros E, Covas MI, Fiol M, et al. Cohort
430 profile: design and methods of the PREDIMED study. Int J Epidemiol 2012; 41:377–85.
431 http://dx.doi.org/10.1093/ije/dyq250.

432 [21] Estruch R, Ros E, Salas-Salvado J, Covas M-I, Corella D, Aros F, et al. Primary prevention of
433 cardiovascular disease with a Mediterranean diet. N Engl J Med 2013;368:1279–90.
434 http://dx.doi.org/10.1056/NEJMoa1200303.

435 [22] Fernandez-Ballart JD, Pinol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity
436 of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain.
437 Br J Nutr 2010;103:1808–16. http://dx.doi. org/10.1017/S0007114509993837.

438 [23] Elosua R, Marrugat J, Molina L, Pons S, Pujol E. Validation of the Minnesota Leisure Time
439 Physical Activity Questionnaire in Spanish men. The MARATHOM investigators. Am J Epidemiol
440 1994;139:1197–209.

441 [24] Mataix Verdú J. Tabla de Composición de Alimentos. Universidad de Granada; 2003.

442 [25] Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0 — the Human
443 Metabolome Database in 2013. Nucleic Acids Res 2013;41:D801–7. http://
444 dx.doi.org/10.1093/nar/gks1065.

445 [26] FengQ, Liu Z, Zhong S, Li R,XiaH, Jie Z, et al. Integratedmetabolomics andmetagenomics
446 analysis of plasma and urine identified microbial metabolites associated with coronary heart disease.
447 Sci Rep 2016;6:22525. http://dx.doi.org/10.1038/srep22525.

448 [27] Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, et al.
449 Metabolomics in nutritional epidemiology: identifying metabolites associated with diet and
450 quantifying their potential to uncover diet-disease relations in populations. Am J Clin Nutr
451 2014;100:208–17. http://dx.doi.org/10. 3945/ajcn.113.078758.

452 [28] van Dorsten FA, Grun CH, van Velzen EJJ, Jacobs DM, Draijer R, van Duynhoven JPM. The
453 metabolic fate of red wine and grape juice polyphenols in humans assessed by metabolomics. Mol
454 Nutr Food Res 2010;54:897–908. http://dx.doi. org/10.1002/mnfr.200900212.

455 [29] Mitchell SC, Zhang AQ, Smith RL. Dimethylamine and diet. Food Chem Toxicol
456 2008;46:1734–8. http://dx.doi.org/10.1016/j.fct.2008.01.010.

457 [30] Schmedes M, Aadland EK, Sundekilde UK, Jacques H, Lavigne C, Graff IE, et al. Leanseafood
458 intake decreases urinarymarkers ofmitochondrial lipid and energymetabolism in healthy subjects:
459 metabolomics results from a randomized crossover intervention study. Mol Nutr Food Res 2016.

460 http://dx.doi.org/10.1002/mnfr.201500785. [31] Stella C, Beckwith-hall B, Cloarec O, Holmes E,
461 Lindon JC, Powell J, et al. Susceptibility of human metabolic phenotypes to dietary modulation. J
462 Proteome Res 2006;5:2780–8. http://dx.doi.org/10.1021/pr060265y.

463 [32] Altorf-van der Kuil W, Brink EJ, Boetje M, Siebelink E, Bijlsma S, Engberink MF, et al.
464 Identification of biomarkers for intake of protein from meat, dairy products and grains: a controlled
465 dietary intervention study. Br J Nutr 2013;110:810–22.
466 http://dx.doi.org/10.1017/S0007114512005788.

467 [33] Cross AJ, Major JM, Sinha R. Urinary biomarkers of meat consumption. Cancer Epidemiol
468 Biomarkers Prev 2011;20:1107–11. http://dx.doi.org/10.1158/1055- 9965.EPI-11-0048.

469 [34] Yeum K-J, Orioli M, Regazzoni L, Carini M, Rasmussen H, Russell RM, et al. Profiling histidine
470 dipeptides in plasma and urine after ingesting beef, chicken or chicken broth in humans. Amino Acids
471 2010;38:847–58. http://dx.doi.org/10.1007/ s00726-009-0291-2.

472 [35] Lloyd AJ, Beckmann M, Haldar S, Seal C, Brandt K, Draper J. Data-driven strategy for the
473 discovery of potential urinary biomarkers of habitual dietary exposure. Am J Clin Nutr 2013;97:377–
474 89. http://dx.doi.org/10.3945/ajcn.112.048033.

475 [36] Boldyrev AA, Aldini G, Derave W. Physiology and pathophysiology of carnosine. Physiol Rev
476 2013;93:1803–45. http://dx.doi.org/10.1152/physrev.00039.2012. [37] Suidasari S, Stautemas J,
477 Uragami S, Yanaka N, Derave W, Kato N. Carnosine content in skeletal muscle is dependent on
478 vitamin B6 status in rats. Front Nutr 2015;2:39. http://dx.doi.org/10.3389/fnut.2015.00039.

479 [38] Urpi-Sarda M, Almanza-Aguilera E, Tulipani S, Tinahones FJ, Salas-Salvadó J, Andres-Lacueva
480 C. Metabolomics for biomarkers of type 2 diabetes mellitus: advances and nutritional intervention
481 trends. Curr Cardiovasc Risk Rep 2015;9. http://dx.doi.org/10.1007/s12170-015-0440-y.

482 [39] Salas-Salvadó J, BullóM, Estruch R, Ros E, CovasM-I, Ibarrola-Jurado N, et al. Prevention of
483 diabetes with Mediterranean diets: a subgroup analysis of a randomized trial. Ann Intern Med
484 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 stressrelated metabolism in free-living subjects. J Proteome Res 2009;8:5568–79. http://dx.doi.org/10.1021/pr900607v.

509 [48] Liu H, Tayyari F, Edison AS, Su Z, Gu L. NMR-based metabolomics reveals urinary
510 metabolomemodifications in female Sprague–Dawley rats by cranberry procyanidins. J Nutr
511 Biochem 2016;34:136–45. http://dx.doi.org/10.1016/j.jnutbio.2016.05.007

512

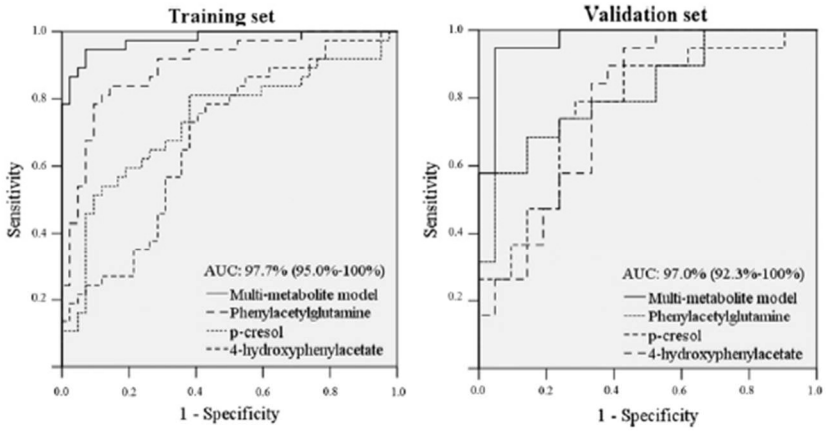

Fig. 1. ROC curves of the combined multimetabolite model and individual metabolites included for discrimination between high and low Mediterranean diet adherence in both training and validation sets.

TABLES

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 ($P<.05$).

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