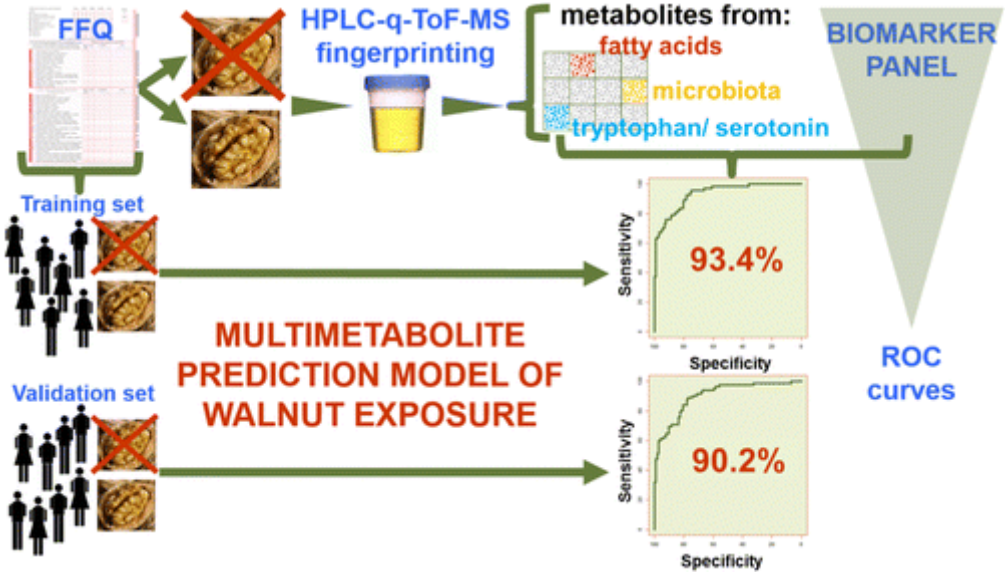
Novel Multimetabolite Prediction of Walnut Consumption by a Urinary Biomarker Model in a Free-Living Population: the PREDIMED Study

Mar Garcia-Aloy[†‡], Rafael Llorach[*†‡], Mireia Urpi-Sarda[†‡], Sara Tulipani[†‡§], Ramon Estruch[‖⊥], Miguel A. Martínez-González[#⊥], Dolores Corella[¶⊥], Montserrat Fitó[□⊥], Emilio Ros[Δ⊥], Jordi Salas-Salvadó[◇⊥], and Cristina Andres-Lacueva[*†‡]

[†]Biomarkers & Nutrimetabolomic Lab, Nutrition and Food Science Department, XaRTA, INSA, Campus Torribera, Pharmacy Faculty, University of Barcelona, Barcelona 08028, Spain
[‡]INGENIO-CONSOLIDER Program, Fun-C-Food CSD2007-063, Barcelona, Spain
[§]Biomedical Research Institute (IBIMA), Service of Endocrinology and Nutrition, Hospital Complex (Virgen de la Victoria), Campus de Teatinos s/n, University of Málaga, Malaga 29010, Spain
[‖]Department of Internal Medicine, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi Sunyer (IDIBAPS), Barcelona 08036, Spain
[⊥]CIBER Fisiopatología de la Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III (ISCIII), Madrid 28029, Spain
[#]Department of Preventive Medicine and Public Health, Medical School, University of Navarra, Pamplona 31080, Spain
[¶]Department of Preventive Medicine and Public Health, University of Valencia, Valencia 46010, Spain
[□]Cardiovascular Risk and Nutrition Research Group, IMIM-Institut de Recerca del Hospital del Mar, Barcelona 08003, Spain
[Δ]Lipid Clinic, Endocrinology and Nutrition Service, Biomedical Research Institute August Pi i Sunyer (IDIBAPS), Hospital Clínic, Barcelona 08036, Spain
[◇]Human Nutrition Unit, Hospital Universitari de Sant Joan de Reus, Institut d'Investigació Sanitària Pere Virgili (IISPV), Universitat Rovira i Virgili, Reus 43204, Spain

*Correspondence: Cristina Andres-Lacueva and Rafael Llorach. Biomarkers & Nutrimetabolomic Lab., Nutrition and Food Science Department, Campus Torribera, Pharmacy Faculty, University of Barcelona, Av. Joan XXIII s/n, 08028 Barcelona, Spain. E-mail: candres@ub.edu and rafallorach@ub.edu. Phone: +34.934034840. Fax: +34.934035931.

ABSTRACT



The beneficial impact of walnuts on human health has been attributed to their unique chemical composition. In order to characterize the dietary walnut fingerprinting, spot urine samples from

two sets of 195 (training) and 186 (validation) individuals were analyzed by an HPLC-q-ToF-MS untargeted metabolomics approach, selecting the most discriminating metabolites by multivariate data analysis (VIP ≥ 1.5). Stepwise logistic regression analysis was used to design a multimetabolite prediction biomarker model. The global performance of the model and each included metabolite in it was evaluated by receiver operating characteristic curves, using the area under the curve (AUC) values. Dietary exposure to walnuts was characterized by 18 metabolites, including markers of fatty acid metabolism, ellagitannin-derived microbial compounds, and intermediate metabolites of the tryptophan/serotonin pathway. The predictive model of walnut exposure included at least one compound of each class. The AUC (95% CI) for the combined biomarker model was 93.4% (90.1–96.8%) in the training set and 90.2% (85.9–94.6%) in the validation set. The AUCs for individual metabolites were ≤85%. As far as we know, this is the first study proposing a combination of biomarkers of walnut exposure in a population under free-living conditions, as considered in epidemiological studies examining associations between diet and health outcomes.

Keywords: biomarkers; HPLC-q-ToF-MS; metabolic fingerprinting; nutrimetabolomics; walnuts

INTRODUCTION

Nut consumption has been associated with a reduced risk of all-cause mortality in different populations.[1-3] Observational studies and clinical trials have reported beneficial effects of nut consumption on chronic diseases and their mediators.[4-6] These effects on human health have been attributed to their unique nutritional composition. They are considered nutrient-dense foods containing several bioactive compounds, such as unsaturated fatty acids (e.g., α-linolenic acid), different types of dietary fiber, plant protein (e.g., arginine), vitamins (e.g., folate, niacin, and vitamin E), minerals (e.g., calcium, magnesium and potassium), and numerous phytochemicals (e.g., phenolic compounds and phytosterols).[4,7,8] Indeed, regular nut consumption is included in the main dietary guidelines worldwide, including the Mediterranean Diet,[9] the Dietary Guidelines for Americans,[10] the Nordic Diet,[11] and the DASH eating plan,[12] among others.

In the study of the associations between dietary exposure and health outcomes, the use of robust biomarkers of food exposure has been proposed as an accurate measurement to estimate real intake.[13] Biomarkers cannot completely replace traditional methods of dietary assessment, but

there is a broad consensus that the application of metabolomics to identify new biomarkers of dietary intake can improve and complement traditional methods, while providing an insight into the underlying mechanisms of diet–health associations.[14,15] In a field of increasing interest, untargeted metabolomic studies have reported a wide variety of new biomarkers of dietary exposure, leading to the discovery of bioactive compounds with potential applications in the design of novel functional foods or dietary supplements.[15] As a result of these types of studies, the interest in using metabolomics for the discovery of new dietary biomarkers for application in nutritional epidemiology has risen remarkably. However, their replication in free-living populations has hardly been tested so far.[14] In this context, replication allows the level of evidence from observed associations to be increased, as has been suggested for genomic studies.[16] Recently a critical review highlighted the new prospects that untargeted metabolomics displays for providing added value in the current nutritional research field to enable us to go much further than we have been able to do using traditional approaches.[15]

Given that walnuts are one of the most widely consumed nut varieties,[17] the aim of the present study was to characterize dietary walnut fingerprinting in spot urine using an untargeted high-performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry (HPLC-q-ToF-MS) metabolomic approach, in two differentiated free-living populations with various levels of walnut intake from the PREDIMED study before changes induced by interventions occurred. For this purpose, we mainly focused on urinary biomarkers of nut exposure proposed in a previous clinical trial analyzing the urinary metabolomic changes that occurred after long-term consumption of mixed nuts by subjects with metabolic syndrome under specific controlled dietary conditions.[18] Therefore, the identified biomarkers need further research on their validity in an uncontrolled free-living population. This type of study could contribute to more realistic data on the differences in the urinary metabolome of habitual walnut consumers.

MATERIALS AND METHODS

The present study was designed and conducted on baseline dietary data and spot urine samples from two differentiated subsamples [subsample 1 (S1), n = 275 subjects; subsample 2 (S2), n = 327 subjects] randomly selected from the cohort of the PREDIMED study (http://www.predimed.es), a clinical trial registered at http://www.controlled-trials.com as ISRCTN35739639. The PREDIMED study is a large randomized clinical trial carried out in Spain, aiming to assess the effects of the traditional Mediterranean diet on the primary

prevention of cardiovascular disease in a high-risk population. Recruitment took place between October 2003 and January 2009. The trial protocol was in accordance with the Helsinki Declaration and was approved by the institutional review boards of all the centers involved. All participants provided written informed consent. Participants were men (55–80 years) and women (60–80 years) without cardiovascular disease at baseline and fulfilling at least one of the two following criteria: presence of type 2 diabetes mellitus or three or more major cardiovascular risk factors.[19]

Sample Selection According to Walnut Consumption

For the current study, baseline spot urine samples from three PREDIMED centers (Barcelona, Valencia, and Navarra) were used and matched to corresponding data obtained from validated semiquantitative food-frequency questionnaires (FFQ) with 9 categories for the frequency of consumption.[20] For statistical analyses purposes, which aim was to identify biomarkers of habitual walnut exposure, subjects were selected from the lowest and highest quintiles of walnut consumption. Thus, two walnut consumer categories were defined as follows: (i) non-walnut consumers (NC), subjects who never consumed walnuts (0 g/day) and (ii) habitual walnut consumers (WN_C), subjects with an intake of at least three servings/week of walnuts (serving size, 30 g) during the preceding year.

Urine Sample Analysis and Data Processing

Spot urine samples used in the present study were collected in the morning in fasting conditions. They were aliquoted, encoded, and kept frozen at −80 °C until used. Sample preparation was based on methodology previously published.[18,21]

HPLC-q-ToF-MS analysis was conducted using an Agilent 1200 Series Rapid Resolution HPLC system coupled to a hybrid quadrupole TOF QSTAR Elite mass spectrometer (Applied Biosystems/MDS Sciex, Framingham, MA, USA) following our protocols published elsewhere.[18] The quality and reproducibility of acquired data were evaluated according to previously reported procedure.[18]

All analyses were carried out in two separate and independent subsamples (S1 and S2). Considering the criteria for sample selection by walnut consumption category, a subgroup of 195 subjects (128 NC and 67 WN_C) was included in the current study and considered as S1.

From the second subsample (S2) of analyzed samples, 186 other subjects were selected (104 NC and 82 WN_C). S1 was used as the training set, while S2 was used as the validation set.

The HPLC-q-ToF-MS raw data were extracted and aligned using MarkerView TM 1.2.1. software (Applied Biosystems, MDS Sciex, Toronto, Ontario, Canada). The parameters used for the processing of raw data are listed in Table S1 (Supporting Information). Data from positive and negative ionization modes were included in two separate data sets for each subsample in order to analyze them separately. Thus, a total of four data sets (two from S1 and two from S2) were used in the current study.

Multivariate Data Analysis and Biomarker Selection

Mass feature data sets were log-transformed and Pareto-scaled before their multivariate statistical analysis using SIMCA-P+13.0 software (Umetrics, Umeå, Sweden). Partial least-squares discriminant analysis with orthogonal signal correction (OSC-PLS-DA) was used to explore differences between groups.[22] As this study involved samples from free-living individuals, OSC filtration was used to reduce the variability not associated with dietary classification.[23] The quality of the models was evaluated by $R^2Y(cum)$ and $Q^2(cum)$ parameters. Validation of the models was evaluated by a permutation test (n = 200).[22] Those mass features with the highest variable importance for projection (VIP) values (cutoff $\geq$1.5) were selected as the most determinants of the differences in urinary metabolomic profiles. VIP value is defined as the influence that each variable has in the PLS-DA model.[24] Thus, the higher the VIP value, the more relevant is the variable in the model. Bearing in mind that usually a threshold of VIP scores $\geq$1 is considered as appropriate for metabolomic studies,[24] the cutoff applied in this study is more restrictive, reducing the possibility of obtaining false positive results.

Identification of Metabolites

Selected mass features were identified by a multistep procedure.[18] First, clustering analysis with Pearson distance and Ward's method to aggregate the observations (PermutMatrix 1.9.3.0 software) was applied in order to identify the mass features corresponding to the same metabolite: (de)protonated molecules, ^{13}C isotopes, adducts, and in-source fragments mainly derived from the loss of the corresponding glucuronide moiety (−176 Da) or sulfate moiety (−80 Da). Then, metabolites were tentatively identified on the basis of their exact mass (±5 mDa of accepted mass difference) and fragmentation patterns using an in-house database

mainly focused on the metabolites expected from the intake of walnuts and the databases Human Metabolome Database (HMDB),[25] LIPID MAPS Structure Database (LMSD),[26] and MetFrag.[27] Finally, the biological interpretation was carried out using information from published research reports and from HMDB.[25]

Building Models of Combined Urinary Markers

To design a multimetabolite prediction biomarker model of walnut exposure, a forward-conditional stepwise logistic regression analysis was conducted with the results obtained from the OSC-PLS-DA analysis using S1 as the training set, while S2 was used as the validation set (IBM SPSS Statistics 20 software, SPSS Inc., Chicago, IL, USA). Metabolites identified as discriminant for walnut exposure in both subsamples were used as independent variables, with a p-value of <0.05 as a condition required for entering and remaining in the model. The dichotomous variable on walnut consumption was used as dependent variable. The correlation between walnut intake and the combined model was evaluated using Spearman's rank-correlation coefficient.

The global performance of the model, as well as of each metabolite included in it, was evaluated by constructing receiver operating characteristic (ROC) curves and estimating the area under the curve (AUC) values.[28] The optimal cutoff for calculating the sensitivity and specificity of biomarkers was determined as the minimum distance to the top-left corner.[29]

RESULTS

The mean (±SD) walnut consumption of the consumer group was 18.6 ± 7.5 and 21.2 ± 11.6 g/day for S1 and S2, respectively. In S1 there were 68 (35%) men and 127 (65%) women with 67.7 ± 6.2 years and BMI 29.6 ± 3.9 kg/m^2, whereas the S2 group comprised 60 (32%) men and 126 (68%) women with 67.2 ± 5.5 years and BMI 30.3 ± 4.1 kg/m2. The OSC-PLS-DA analyses resulted in four models with one component. All models held robust modeling and prediction results,[23] suggesting that they were able to classify each subject in the correct consumption group (NC or WN_C). The quality parameters of all OSC-PLS-DA models, as well as the corresponding permutation tests, are summarized in Supplementary Table S2.

Selection and Identification of Discriminatory Metabolites Related to Regular Walnut Exposure

A total of 18 metabolites were identified or tentatively identified. All of them were higher in the WN_C group compared to the NC group. Twelve of the identified metabolites were discriminative for both subsamples (Table 1), whereas the other six were only characteristic of walnut exposure in one of the two analyzed subsamples (Supplementary Table S3). Supplementary Table S4 includes information about the MS pattern of identified metabolites (isotopes, adducts, and source-generated ion fragments). The identified metabolites related to walnut exposure were grouped in three differentiated classes as follows: (i) fatty acid metabolites: 10-hydroxy-decene-4,6-diynoic acid sulfate, tridecadienoic/tridecynoic acid glucuronide, and dodecanediodic acid; (ii) polyphenol microbial metabolites: three conjugates of urolithin A, urolithin B and its glucuronide, two conjugates of urolithin C, pyrogallol sulfate, methylpyrogallol sulfate, and enterolactone glucuronide; and (iii) compounds related to the tryptophan/serotonin (methoxyindoles) pathway: 3-indolecarboxylic acid glucuronide, hydroxyindoleacetic acid sulfate, N-acetylserotonin sulfate, and hydroxyindoleacetic acid.

Table 1. Urinary Metabolites Tentatively Identified As Walnut Exposure Biomarkers[a]

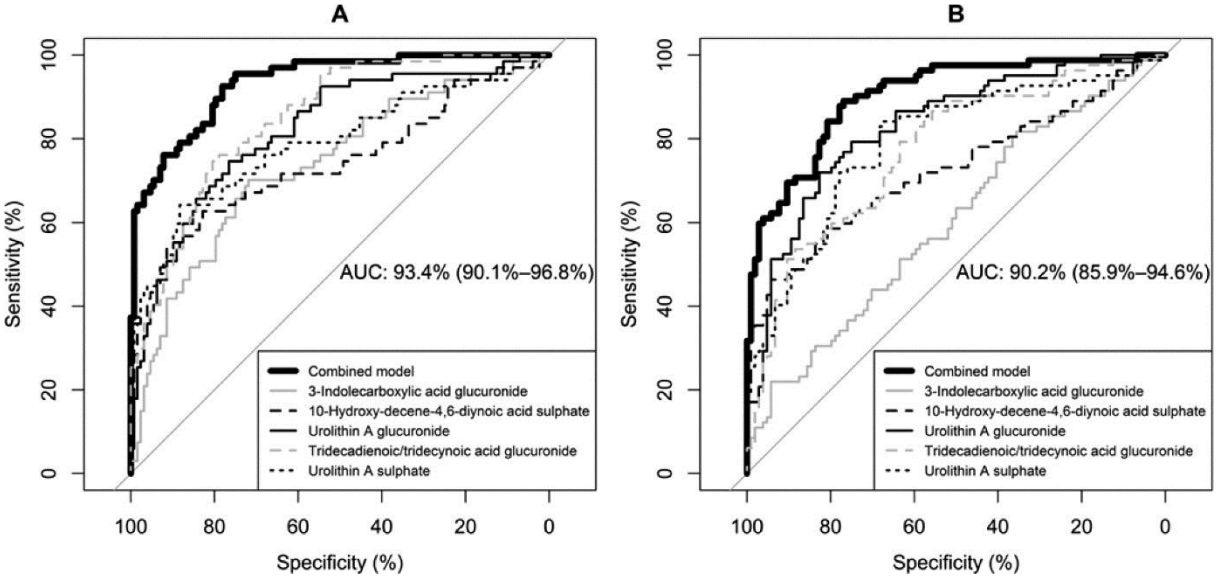
RT	detected mass (m/z)	assignation	tentative metabolite identification	biological source	level of identify-cation[b]
3.23	336.0751 338.0854	$[M–H]^–$ $[M+H]^+$	3-indolecarboxylic acid glucuronide[c]	tryptophan/serotonin metabolic pathway[f]	II
3.83	270.0081	$[M–H]^–$	hydroxyindoleacetic acid sulfate[d]	tryptophan/serotonin metabolic pathway[f]	II
4.20	297.0561	$[M–H]^–$	N-acetylserotonin sulfate[d]	tryptophan/serotonin metabolic pathway[f]	II
4.62	257.0149	$[M–H]^–$	10-hydroxy-decene-4,6-diynoic acid sulfate[d]	fatty acid metabolism	II
5.22	419.0618	$[M–H]^–$	urolithin C glucuronide	polyphenol microbial metabolism	III
5.25	403.0662 405.0830	$[M–H]^–$ $[M+H]^+$	urolithin A glucuronide[d]	polyphenol microbial metabolism	I
5.35	483.0227	$[M–H]^–$	urolithin A sulfoglucuronide[d]	polyphenol microbial metabolism	II
6.20	385.1838 387.1995	$[M–H]^–$ $[M+H]^+$	tridecadienoic/tridecynoic acid glucuronide[d]	fatty acid metabolism	II
6.25	387.0770 389.0864	$[M–H]^–$ $[M+H]^+$	urolithin B glucuronide[e]	polyphenol microbial metabolism	II
6.34	473.1491	$[M–H]^–$	enterolactone glucuronide[e]	polyphenol microbial metabolism	III
6.67	243.0295	$[M–H–sulfate]^–$	urolithin C sulfate[e]	polyphenol microbial metabolism	III
6.72	306.9915	$[M–H]^–$	urolithin A sulfate[d]	polyphenol microbial metabolism	II

[a]All features have VIP values ≥1.5 in the corresponding OSC-PLS-DA model in both subsamples. [b]Level of identification has been assigned according to Sumner et al.[30] [c]Structural identification agrees with Dong et al.[31] [d]Structural identification agrees with Tulipani et al.[18] [e]Structural identification agrees with MetFrag[27] (LC–MS pattern information in Supplementary Table S4). [f]Endogenous and/or exogenous origin.

Designed Model of Combined Urinary Metabolites As More Robust Biomarker of Walnut Exposure

In order to improve the discrimination between groups (NC and WN_C), a combination of more than one of the identified metabolites was developed using S1 as the training set. For this purpose, all metabolites identified as sufficiently discriminative in both subsamples were subjected to a conditional stepwise variable selection method through a binary logistic regression analysis. Data from the negative ionization mode were considered for this analysis, since all metabolites were detected in this mode. The results of the designed model are shown in Supplementary Table S5. Five metabolites were included in the fitted model, which contained at least one metabolite of each class, namely, two markers of fatty acid metabolism (10-hydroxy-decene-4,6-diynoic acid sulfate and tridecadienoic/tridecynoic acid glucuronide), two metabolites from microbial-derived polyphenol metabolism (glucuronide and sulfate conjugates of urolithin A), and one compound related to the tryptophan metabolic pathway (3-indolecarboxylic acid glucuronide). Reported daily walnut consumption correlated [r (95% CI)] significantly with values of the combined model [r = 0.71 (0.63–0.77) for training set and r = 0.67 (0.58–0.74) for validation set]. The multimetabolite prediction biomarker model and the metabolites included in it were evaluated individually and compared through ROC analyses using both training and validation sets separately. Supplementary Table S6 shows the results of ROC analyses for each combined biomarker model and of the corresponding included metabolites individually. All sensitivity and specificity values from the combined model for both subsets were nearly higher or higher than 80%, while none of the analyzed individual metabolites displayed at the same time a sensitivity and specificity both ≥80%. Accordingly, the AUC for the combined biomarker model was >90% for both S1 and S2 sets, whereas this value for individual metabolites was ≤85%. Figure 1 also shows that the combination of metabolites significantly improved the discrimination of walnut consumers compared to the use of each metabolite individually ($p < 0.05$, Supplementary Table S6). We also assessed the AUC values according to gender and age in order to evaluate if these variables could affect the predictability of the multimetabolite biomarker panel. However, there were no statistically significant differences in the AUC values in either of the two subsamples between men and women or according to median age (≤67 or >67 years).

Figure 1. Receiver operating characteristic (ROC) curves of multimetabolite prediction biomarker model (broad line) and of included individual metabolites (narrow lines) in the training (A) and validation (B) sets.



DISCUSSION

In this study, a panel of different urinary metabolites related to walnut exposure in the habitual diet was confirmed as discriminatory biomarkers using an untargeted nutrimetabolomic approach in a free-living population. In addition, sufficiently high correlations were found when the exposure was assessed as a continuous variable (defined by the combined biomarker panel). Markers of fatty acid metabolism were identified, together with several ellagitannin-microbial metabolites and other microbial-derived compounds, as well as intermediate metabolites of the tryptophan/serotonin pathway (Table 1 and Supplementary Table S2), reinforcing our previously published results obtained using 24-h urine samples.[18] While 24 h urine has been described as a more robust method to monitor daily intake than spot urine,[32,33] collecting 24-h urine samples is extremely difficult and impractical, especially in large-scale epidemiological studies.[33,34] Thus, the replication of this biomarker panel using spot urines from free-living subjects reinforces its discriminatory power of walnut exposure.

Characterization of Dietary Walnut Fingerprinting in Urine

The presence of different types of fatty acids as markers of walnut consumption has been explained by the substantial and specific content of polyunsaturated fatty acids in walnuts.[18]

Urolithins are gut microbiota products from ellagic acid and ellagitannins, which are characterized by low absorption rates. They usually reach the lower gastrointestinal tract, where they are converted to urolithins,[35] which are absorbed and metabolized to finally circulate in blood reaching different tissues prior to excretion.[36] Urolithins have been considered as bioactive compounds associated with relevant health effects.[36] Tryptophan/serotonin-derived compounds have been related to both exogenous and endogenous sources.[18] Hydroxyindoleacetic acid and N-acetylserotonin are involved in the tryptophan methoxyindole pathway. This pathway comprises serotonin as an intermediate metabolite, which has also been detected in significant amounts in walnuts.[37] 3-Indolecarboxylic acid has also been related to tryptophan metabolism.[38,39] In this sense, bacterial degradation of tryptophan generates different indole acid derivatives.[40] Walnuts are one of the food products with relatively high amounts of tryptophan,[40] and it has been demonstrated that indole-containing metabolite levels are highly impacted by gut microbiota.[41] The presence of 3-indolecarboxylic acid glucuronide has also recently been reported in human urine by an untargeted metabolomics study.[42]

Multimetabolite Prediction Biomarker Model of Walnut Exposure

Among all identified significant metabolites, half of them matched proposed biomarkers of nut consumption that were observed in a controlled dietary clinical trial,[18] allowing their replication in free-living subjects. One possible explanation may be that in the previous clinical trial mixed nuts were used, with half of the supplemented serving composed of walnuts. Thus, the matched compounds in both studies could be those characteristic of walnut exposure. These metabolites could be the strongest candidates as exposure biomarkers to be applied as routine measurements of walnut consumption,[23] since they are characteristic of subjects from both highly and less controlled studies, i.e., consuming both walnuts on their own and in a controlled clinical trial. The presence of the same metabolites in two independent studies with different designs and dietary conditions further extends the robustness of their validity as biomarkers of walnut exposure.

Our findings indicate that the use of the proposed biomarkers individually is not as specific for walnut exposure as the multimetabolite panel. One possible explanation may be that ellagic acid and ellagitannins (precursors of urolithins) are present in a variety of plant foods, such as walnuts (pedunculagins), berries (sanguiin H6, sanguiin H10, and lambertianin C), and pomegranates (punicalagins and punicalins), among others.[35,36] Still, even though various foods contain ellagitanins,[43] it must be underlined that walnuts are the most important source

throughout the year in our Mediterranean population. Accordingly, although the observed fatty acid compounds and metabolites of the tryptophan/serotonin pathway have great discriminatory power, their specificity as robust biomarkers of habitual walnut consumption seems to be suboptimal. For this reason, we proposed combining the most discriminating metabolites in a unique multimetabolite model using a stepwise logistic regression procedure to improve the ability of walnut exposure prediction using morning fasting spot urine samples. Additionally, this combined panel allows us to take into account the correlations between metabolites,[44] as well as covering a wide range of characteristic compounds of walnuts, thus providing a precise walnut-exposure fingerprint. The discriminatory capacity of the designed model was significantly better than models using each metabolite individually (Supplementary Table S6). These results reinforce the improved ability of multimetabolite biomarker models to selectively define dietary exposure. Interestingly, proposed biomarkers have been applied using habitual and recommended walnut portion size consumption and morning fasting spot urines, which reflect real conditions of nutritional epidemiologic studies.

Because it is a particular component of this food, plasma α-linolenic acid levels determined by gas chromatography have been previously proposed as markers of walnut intake.[45] In this context, as has been commented previously, walnuts are characterized by a unique chemical composition. Thus, considering their whole composition could provide more accurate information about their consumption. Indeed in this study untargeted metabolomics has been used to obtain a broad picture of several bioactive constituents of walnuts in human urine, specifically a combination of polyphenols, fatty acids, and serotonin/tryptophan metabolites. For this reason, the approach presented here could be considered as complementary information to that offered by the other markers.

Given the increasing interest of the food industry in developing new functional foods, there is a need for objective biomarkers of food exposure that enable accurate measurements of their bioavailability. One of the strategic plans of the international community, including both the U.S. Food and Drug Administration (FDA)[46] and the European Commission,[47] concerns the need for the development and validation of biomarkers of food intake using omics-based approaches such as food metabolomics. Therefore, the development of new biomarkers of food exposure and even the proposal of new strategies to obtain novel biomarker patterns such as that developed in the current study could contribute to the advancement in this important field in terms of both health and economics.

That data used to categorize walnut consumers was derived from a FFQ is a potential weakness of the study, as participants could have under- or over-reported their usually intakes. However, the results obtained from this FFQ can be considered reliable since it was previously validated[20] and was administrated by trained dieticians.[19] Additionally, the participants of the present study were at high cardiovascular risk, from a Mediterranean region, and not representative of the general population, so the results should be extrapolated with some caution to other populations. Finally, widespread application of biomarkers in dietary studies could be limited by the cost of analytical technique used for their determination.[48] However, everyday advances in technology should soon lead to improved efficiency and precision in biomarker determinations, making them more economical and, therefore, more accessible for application in dietary studies.[14] An important strength of the study is that it included free-living volunteers following their customary diet and with varying walnut consumption levels, providing more reliable data than highly controlled studies. It is also important to note the use of a combination of multivariate and univariate analyses as complementary statistical techniques to better characterize discriminant metabolites.

CONCLUSIONS

Dietary walnut fingerprinting in urine is characterized by a combination of metabolites of different classes, including markers of fatty acid metabolism, ellagitannin-derived microbial compounds, and intermediate metabolites of the tryptophan/serotonin pathway. When merged in a multimetabolite prediction biomarker, a better prediction performance was obtained than from any metabolite individually. As far as we know, this is the first study to propose a combination of more than one biomarker of walnut exposure in an uncontrolled population under free-living conditions, which is characteristic of epidemiological studies examining dietary exposures in relation to health outcomes. This approach might also be of interest for the further discovery of dietary exposure biomarkers as determinants of compliance in long-term intervention trials conducted in free-living individuals.

ASSOCIATED CONTENT

Supporting Information

Parameters used for processing of raw data in MarkerView TM 1.2.1. (Table S1). Summary of parameters for assessing the OSC-PLS-DA modeling quality (Table S2). Urinary metabolites

tentatively identified as walnut exposure biomarkers in one of the two analyzed subsamples (Table S3). HPLC–MS pattern of tentatively identified metabolites (Table S4). Metabolites selected by stepwise logistic regression model for the discrimination of walnut exposure using subsample 1 (Table S5). Receiver operating characteristic (ROC) curve parameters of combined models and of corresponding included metabolites in subsample 1 (training set) and subsample 2 (validation set) (Table S6). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: rafallorach@ub.edu

*E-mail: candres@ub.edu

Notes

E.R. and J.S.-S. are nonpaid members of the Scientific Advisory Board of the California Walnut Commission (California, USA) and of the International Nut and Dried Fruit Foundation (Reus, Spain), respectively. E.R.'s institution has received research funds from the California Walnut Commission. J.S.-S.'s institution has received research funds from the International Nut and Dried Fruit Foundation. None of the other authors had any conflict of interest to report.

ACKNOWLEDGMENTS

This work has been supported by Spanish public grants (AGL2009-13906-C02-01, AGL2010-10084-E and the CONSOLIDER-INGENIO 2010 Program, FUN-C-FOOD (CSD2007-063)) from the Spanish Ministry of Economy and Competitiveness (MINECO), as well as FEDER (Fondo Europeo de Desarrollo Regional); Merck Serono Research Grants 2010 (Fundación Salud 2000); CIBERobn and RD 06/0045 from the Spanish Ministry of Health (ISCIII). M.G.-A. thanks the Generalitat de Catalunya's Agency AGAUR for the predoctoral FI-DGR 2011 fellowship; R.Ll. and M.U.-.S thank the "Ramón y Cajal" program (RYC-2010–07334 and RYC-2011-09677, respectively) from the MINECO and Fondo Social Europeo; and S.T. thanks the "Juan de la Cierva" program from the MINECO. We would like to thank Dr. J. C. Espín (Murcia) for his generous gift providing urolithin metabolites.

ABBREVIATIONS

AUC, area under the curve; FFQ, food frequency questionnaire; HPLC-q-ToF-MS, high-performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry; NC, non-walnut consumers; OSC-PLS-DA, partial least-squares discriminant analysis with orthogonal signal correction; VIP, variable importance for projection; ROC, receiver operating characteristic; WN_C, walnut consumers

REFERENCES

(1) Bao, Y.; Han, J.; Hu, F. B.; Giovannucci, E. L.; Stampfer, M. J.; Willett, W. C.; Fuchs, C. S. Association of nut consumption with total and cause-specific mortality N. Engl. J. Med. 2013, 369 (21) 2001–11.

(2) Guasch-Ferre, M.; Bullo, M.; Martinez-Gonzalez, M. A.; Ros, E.; Corella, D.; Estruch, R.; Fito, M.; Aros, F.; Warnberg, J.; Fiol, M. Frequency of nut consumption and mortality risk in the PREDIMED nutrition intervention trial BMC Med. 2013, 11, 164.

(3) Fernández-Montero, A.; Bes-Rastrollo, M.; Barrio-López, M. T.; de la Fuente-Arrillaga, C.; Salas-Salvadó, J.; Moreno-Galarraga, L.; Martínez-González, M. A. Nut consumption and 5-year all-cause mortality in a Mediterranean cohort: The SUN Project. Nutrition 2014, 10.1016/j.nut.2014.02.018.

(4) Sabate, J.; Oda, K.; Ros, E. Nut consumption and blood lipid levels: a pooled analysis of 25 intervention trials Arch. Intern. Med. 2010, 170 (9) 821–7.

(5) Ros, E.; Hu, F. B. Consumption of plant seeds and cardiovascular health: epidemiological and clinical trial evidence Circulation 2013, 128 (5) 553–65.

(6) Kris-Etherton, P. M. Walnuts decrease risk of cardiovascular disease: a summary of efficacy and biologic mechanisms J. Nutr. 2014, 144 (4) 547S–54S.

(7) Ros, E.; Mataix, J. Fatty acid composition of nuts--implications for cardiovascular health Br. J. Nutr. 2006, 96 (Suppl 2) S29–35.

(8) Segura, R.; Javierre, C.; Lizarraga, M. A.; Ros, E. Other relevant components of nuts: phytosterols, folate and minerals Br. J. Nutr. 2006, 96 (Suppl 2) S36–44.

(9) Bach-Faig, A.; Berry, E. M.; Lairon, D.; Reguant, J.; Trichopoulou, A.; Dernini, S.; Medina, F. X.; Battino, M.; Belahsen, R.; Miranda, G. Mediterranean diet pyramid today. Science and cultural updates Public Health Nutr. 2011, 14 (12A) 2274–84.

(10) U.S. Department of Agriculture; U.S. Department of Health and Human Services; Dietary Guidelines for Americans. U.S. Government Printing Office: Washington, DC, 2010. http://www.health.gov/dietaryguidelines/dga2010/DietaryGuidelines2010.pdf (accessed Jan 3, 2014).

(11) 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) 777–85.

(12) U.S. Department of Health and Human Services. Your Guide to Lowering Your Blood Pressure with DASH, 2006. http://www.nhlbi.nih.gov/health/public/heart/hbp/dash/new_dash.pdf (accessed Jan 3, 2014).

(13) Jenab, M.; Slimani, N.; Bictash, M.; Ferrari, P.; Bingham, S. A. Biomarkers in nutritional epidemiology: applications, needs and new horizons Hum. Genet. 2009) 125 (5–6) 507–25.

(14) Llorach, R.; Garcia-Aloy, M.; Tulipani, S.; Vazquez-Fresno, R.; Andres-Lacueva, C. Nutrimetabolomic strategies to develop new biomarkers of intake and health effects J. Agric. Food Chem. 2012, 60 (36) 8797–808.

(15) Scalbert, A.; Brennan, L.; Manach, C.; Andres-Lacueva, C.; Dragsted, L. O.; Draper, J.; Rappaport, S. M.; Van der Hooft, J. J.; Wishart, D. The food metabolome: a window over dietary exposure Am. J. Clin. Nutr. 2014, 99 (6) 1286–308.

(16) Chanock, S. J.; Manolio, T.; Boehnke, M.; Boerwinkle, E.; Hunter, D. J.; Thomas, G.; Hirschhorn, J. N.; Abecasis, G.; Altshuler, D.; Bailey-Wilson, J. E. Replicating genotype-phenotype associations Nature 2007, 447 (7145) 655–60.

(17) Jenab, M.; Sabate, J.; Slimani, N.; Ferrari, P.; Mazuir, M.; Casagrande, C.; Deharveng, G.; Tjonneland, A.; Olsen, A.; Overvad, K. Consumption and portion sizes of tree nuts, peanuts and seeds in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts from 10 European countries Br. J. Nutr. 2006, 96 (Suppl 2) S12–23.

(18) Tulipani, S.; Llorach, R.; Jauregui, O.; Lopez-Uriarte, P.; Garcia-Aloy, M.; Bullo, M.; Salas-Salvado, J.; Andres-Lacueva, C. Metabolomics unveils urinary changes in subjects with

metabolic syndrome following 12-week nut consumption J. Proteome Res. 2011, 10 (11) 5047–58.

(19) 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. Cohort profile: design and methods of the PREDIMED study Int. J. Epidemiol. 2012, 41 (2) 377–85.

(20) Fernandez-Ballart, J. D.; Pinol, J. L.; Zazpe, I.; Corella, D.; Carrasco, P.; Toledo, E.; Perez-Bauer, M.; Martinez-Gonzalez, M. A.; Salas-Salvado, J.; Martin-Moreno, J. M. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain Br. J. Nutr. 2010, 103 (12) 1808–16.

(21) Llorach, R.; Urpi-Sarda, M.; Jauregui, O.; Monagas, M.; Andres-Lacueva, C. An LC-MS-based metabolomics approach for exploring urinary metabolome modifications after cocoa consumption J. Proteome Res. 2009, 8 (11) 5060–8.

(22) Llorach-Asuncion, 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) 373–81.

(23) Pujos-Guillot, E.; Hubert, J.; Martin, J. F.; Lyan, B.; Quintana, M.; Claude, S.; Chabanas, B.; Rothwell, J. A.; Bennetau-Pelissero, C.; Scalbert, A. 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) 1645–59.

(24) Bryan, K.; Brennan, L.; Cunningham, P. MetaFIND: A feature analysis tool for metabolomics data BMC Bioinformatics. 2008, 9 (1) 470.

(25) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E. HMDB 3.0—The Human Metabolome Database in 2013 Nucleic Acids Res. 2013, 41 (Database issue) D801–7.

(26) Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. LIPID MAPS online tools for lipid research Nucleic Acids Res. 2007, 35 (Web Server issue) W606–12.

(27) Wolf, S.; Schmidt, S.; Muller-Hannemann, M.; Neumann, S. In silico fragmentation for computer assisted identification of metabolite mass spectra BMC Bioinf. 2010, 11, 148.

(28) Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J. C.; Muller, M. pROC: an open-source package for R and S+ to analyze and compare ROC curves BMC Bioinf. 2011, 12, 77.

(29) Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S. Translational biomarker discovery in clinical metabolomics: an introductory tutorial Metabolomics 2013, 9 (2) 280–99.

(30) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.; Fiehn, O.; Goodacre, R.; Griffin, J. L. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI) Metabolomics 2007, 3 (3) 211–21.

(31) Dong, H.; Zhang, A.; Sun, H.; Wang, H.; Lu, X.; Wang, M.; Ni, B.; Wang, X. Ingenuity pathways analysis of urine metabolomics phenotypes toxicity of Chuanwu in Wistar rats by UPLC-Q-TOF-HDMS coupled with pattern recognition methods Mol. Biosyst. 2012, 8 (4) 1206–21.

(32) Spencer, J. P.; Abd El Mohsen, M. M.; Minihane, A. M.; Mathers, J. C. Biomarkers of the intake of dietary polyphenols: strengths, limitations and application in nutrition research Br. J. Nutr. 2008, 99 (1) 12–22.

(33) Zamora-Ros, R.; Rabassa, M.; Cherubini, A.; Urpi-Sarda, M.; Llorach, R.; Bandinelli, S.; Ferrucci, L.; Andres-Lacueva, C. Comparison of 24-h volume and creatinine-corrected total urinary polyphenol as a biomarker of total dietary polyphenols in the Invecchiare InCHIANTI study Anal. Chim. Acta 2011, 704 (1–2) 110–5.

(34) Potischman, N. Biologic and methodologic issues for nutritional biomarkers J. Nutr. 2003, 133 (Suppl 3) 875S–80S.

(35) Garcia-Munoz, C.; Vaillant, F. Metabolic fate of ellagitannins: implications for health, and research perspectives for innovative functional foods Crit. Rev. Food Sci. Nutr. 2014, 54 (12) 1584–98.

(36) Espin, J. C.; Larrosa, M.; Garcia-Conesa, M. T.; Tomas-Barberan, F. Biological significance of urolithins, the gut microbial ellagic acid-derived metabolites: The evidence so far J. Evidence-Based Complementary Altern. Med. 2013, DOI: 10.1155/2013/270418.

(37) Feldman, J. M.; Lee, E. M. Serotonin content of foods: effect on urinary excretion of 5-hydroxyindoleacetic acid Am. J. Clin. Nutr. 1985, 42 (4) 639–43.

(38) Byrd, D. J.; Kochen, W.; Idzko, D.; Knorr, E. The analysis of indolic tryptophan metabolites in human urine. Thin-layer chromatography and in situ quantitation J. Chromatogr. 1974, 94 (0) 85–106.

(39) Magnus, V.; Simaga, S.; Iskric, S.; Kveder, S. Metabolism of tryptophan, indole-3-acetic acid, and related compounds in parasitic plants from the genus Orobanche Plant Physiol. 1982, 69 (4) 853–8.

(40) Keszthelyi, D.; Troost, F. J.; Masclee, A. A. Understanding the role of tryptophan and serotonin metabolism in gastrointestinal function Neurogastroenterol. Motil. 2009, 21 (12) 1239–49.

(41) Wikoff, W. R.; Anfora, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites Proc. Natl. Acad. Sci. U.S.A. 2009, 106 (10) 3698–703.

(42) Roux, A.; Xu, Y.; Heilier, J. F.; Olivier, M. F.; Ezan, E.; Tabet, J. C.; Junot, C. Annotation of the human adult urinary metabolome and metabolite identification using ultra high performance liquid chromatography coupled to a linear quadrupole ion trap-Orbitrap mass spectrometer Anal. Chem. 2012, 84 (15) 6429–37.

(43) Perez-Jimenez, J.; Hubert, J.; Hooper, L.; Cassidy, A.; Manach, C.; Williamson, G.; Scalbert, A. Urinary metabolites as biomarkers of polyphenol intake in humans: a systematic review Am. J. Clin. Nutr. 2010, 92 (4) 801–9.

(44) Altorf-van der Kuil, W.; Brink, E. J.; Boetje, M.; Siebelink, E.; Bijlsma, S.; Engberink, M. F.; van ’t Veer, P.; Tome, D.; Bakker, S. J.; van Baak, M. A. Identification of biomarkers for intake of protein from meat, dairy products and grains: a controlled dietary intervention study Br. J. Nutr. 2013, 110 (5) 810–22.

(45) Segovia-Siapco, G.; Singh, P.; Jaceldo-Siegl, K.; Sabate, J. Validation of a food-frequency questionnaire for measurement of nutrient intake in a dietary intervention study Public Health Nutr. 2007, 10 (2) 177–84.

(46) U.S. Food and Drug Administration Guidance for industry: evidence-based review system for the scientific evaluation of health claims. http://www.fda.gov/food/guidanceregulation/guidancedocumentsregulatoryinformation/labelingnutrition/ucm073332.htm (accessed Mar 06, 2014)..

(47) Joint Programming Initiative Strategic Research Agenda 2012–2020 and beyond. http://www.healthydietforhealthylife.eu/images/documents/jpi_sra.pdf (accessed Mar 06, 2014).

(48) Zamora-Ros, R.; Rabassa, M.; Llorach, R.; Gonzalez, C. A.; Andres-Lacueva, C. Application of dietary phenolic biomarkers in epidemiology: past, present, and future J. Agric. Food Chem. 2012, 60 (27) 6648–57.