Nutrimetabolomics fingerprinting to identify biomarkers of bread exposure in a free-living population

Mar Garcia-Aloy[1,2], Rafael Llorach[1,2,*], Mireia Urpi-Sarda[1,2], Sara Tulipani[2,3], Jordi Salas-Salvadó[4,5], Miguel Angel Martínez-González[5,6], Dolores Corella[5,7], Montserrat Fitó[5,8], Ramon Estruch[5,9], Lluis Serra-Majem[5,10], Cristina Andres-Lacueva[1,2,*]

[1]Biomarkers & Nutrimetabolomic Lab., Nutrition and Food Science Department, XaRTA, INSA, Campus Torribera, Pharmacy Faculty, University of Barcelona, Spain; [2]INGENIO−CONSOLIDER Program, Fun-C-Food CSD2007-063, Ministry of Science and Innovation, Spain. [3]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. [4]Human Nutrition Unit, Hospital Universitari de Sant Joan de Reus, Institut d'Investigació Sanitària Pere Virgili (IISPV), Universitat Rovira i Virgili, Reus, Spain. [5]CIBER Fisiopatología de la Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III (ISCIII), Madrid, Spain. [6]Department of Preventive Medicine and Public Health, Medical School-Clinica, University of Navarra, Pamplona, Spain. [7]Department of Preventive Medicine and Public Health, University of Valencia, Valencia, Spain. [8]Cardiovascular Risk and Nutrition Research Group, IMIM-Institut de Recerca del Hospital del Mar, Barcelona, Spain. [9]Department of Internal Medicine, Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi Sunyer (IDIBAPS), Barcelona, Spain. [10]Research Institute of Biomedical and Health Sciences, University of Las Palmas de Gran Canaria, Las Palmas, 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

Bread is one of the most widely consumed foods. Its impact on human health is currently of special interest for researchers. We aimed to identify biomarkers of bread consumption by applying a nutrimetabolomic approach to a free-living population. An untargeted HPLC-q-TOF-MS and multivariate analysis was applied to human urine from 155 subjects stratified by habitual bread consumption in three groups: non-consumers of bread (NCB, n=56), white-bread consumers (WHB, n=48) and whole-grain bread consumers (WGB, n=51). The most differential metabolites (VIP ≥1.5) included compounds originating from cereal plant phytochemicals such as benzoxazinoids and alkylresorcinol metabolites, and compounds produced by gut microbiota (such as enterolactones, hydroxybenzoic and dihydroferulic acid metabolites). Pyrraline, riboflavin, 3-indolecarboxylic acid glucuronide, 2,8-dihydroxyquinoline glucuronide and N-α-acetylcitrulline were also tentatively identified. In order to combine multiple metabolites

in a model to predict bread consumption, a stepwise logistic regression analysis was used. Receiver operating curves were constructed to evaluate the global performance of individual metabolites and their combination. The area under the curve values [AUC (95% CI)] of combined models ranged from 77.8% (69.1%-86.4%) to 93.7% (89.4%-98.1%), whereas the AUC for the metabolites included in the models had weak values when they were evaluated individually: from 58.1 % (46.6-69.7 %) to 78.4% (69.8%-87.1%). Our study showed that a daily bread intake significantly impacted on the urinary metabolome, despite being examined under uncontrolled free-living conditions. We further concluded that a combination of several biomarkers of exposure is better than a single biomarker for the predictive ability of discriminative analysis.

Keywords: Nutrimetabolomics, Food metabolome, Biomarkers, Bread, HPLC-q-TOF-MS, Metabolic Fingerprinting

1. INTRODUCTION

Recently, metabolomics started being applied in nutritional sciences (i.e., nutrimetabolomics) in order to discover new biomarkers of intake and to explore the complex relationships between the consumption of dietary compounds and health outcomes (Ismail et al. 2013). In this regard, the study of food metabolome, being a data-driven approach, is expected to be more objective in the measurement of dietary exposure (Llorach et al. 2012) than traditional methods, such as food frequency questionnaires (FFQs), 24-hour dietary recalls or food diaries.

Bread is a fundamental component of the Mediterranean Diet (MeDiet) belonging to the cereal group, which is one of the three basic elements of the MeDiet included in all daily meals (Bach-Faig et al. 2011). Thus, it could contribute to the beneficial effects reported for this dietary pattern (Sofi et al. 2010). In Spain, bread, mainly white bread, is the most widely consumed cereal (Spanish Agency for Food Safety and Nutrition 2011). Bread provides important amounts of complex carbohydrates, fibre and B vitamins, among other bioactive compounds (Quilez and Salas-Salvado 2012). Recent scientific publications suggest some potential health benefits of cereals and bread with respect of chronic diet-related diseases, such as cardiovascular disease (CVD) and type-2 diabetes mellitus. However, these relationships have been mainly observed for whole-grain bread or cereals

(Gil et al. 2011), whereas the associations for white bread are not clear, and have been associated with increased adiposity and weight gain (Bautista-Castano et al. 2012a; Bautista-Castano and Serra-Majem 2012b; Williams 2012).

Thus the application of nutrimetabolomics to a high-cardiovascular-risk population could provide new insights into this potential relationship. Recently, changes in metabolomic profiles due to dietary intervention with different types of bread were evaluated in some studies, providing a list of potential biomarkers of dietary intake (Beckmann et al. 2013; Bondia-Pons et al. 2013; Bondia-Pons et al. 2011; Lankinen et al. 2011; Moazzami et al. 2012; Nielsen et al. 2014). Due to the intervention character of these studies, they were performed on a relatively small number of subjects and under specifically controlled dietary conditions. Therefore, the identified biomarkers need further research on their validity in epidemiological studies, in which diet is evaluated in uncontrolled conditions.

In the present work we evaluated the differences in metabolomic profiles of habitual bread consumers from a general uncontrolled population of volunteers. Analysing a larger population and under free-living conditions, we expect to obtain more representative data on the metabolome fingerprints of bread consumers, with interest especially in the identification of biomarkers of habitual bread consumption. Since biomarkers may be affected by subject-specific characteristics, it should be necessary to consider the specifications of population used in the present study when proposed biomarkers will be applied elsewhere. Therefore, the principal objective of the study was to discover novel biomarkers of habitual bread exposure applying an untargeted HPLC-q-TOF-MS-based metabolomic approach. We anticipated that biomarkers identified in this study would help to provide a more accurate estimation of intake and to generate new insights into biological mechanisms underlying the health-related effects of bread consumption.

2. MATERIALS AND METHODS

This study is a cross-sectional assessment of baseline data obtained from a subset of 275 participants from the PREDIMED Study (ISRCTN 35739639; http://www.predimed.org).

2.1. PREDIMED study

The PREDIMED study is a large, parallel-group, multicentre, randomized, and controlled clinical trial aimed at assessing the effects of the Mediterranean diet on the primary prevention of CVD. Protocol details of the PREDIMED trial have been reported elsewhere (Estruch et al. 2013; Martinez-Gonzalez et al. 2012). The trial protocol was conducted according to the Declaration of Helsinki and was approved by the institutional review boards of all the centres involved. All participants had signed an informed consent. Briefly, participants were free-living men (55–80 years) and women (60–80 years), without CVD and fulfilling at least one of the two following criteria: type 2 diabetes mellitus, or three or more major cardiovascular risk factors (Martinez-Gonzalez et al. 2012). Urine samples were collected at baseline visit without any previous dietary restrictions in addition to demographic and social data of the volunteers according to the PREDIMED study protocol (Martinez-Gonzalez et al. 2012). All participants completed a validated semi-quantitative 137-item FFQ (Fernandez-Ballart et al. 2010), and nutrient and energy intakes were calculated using Spanish food composition tables (Mataix Verdú 2003).

2.2. Stratification of study population

The current study was focused on establishing markers of habitual bread exposure in a free-living population. For this purpose, participants from three PREDIMED trial centres (Barcelona, Valencia and Navarra) were stratified on the basis of their FFQ data in three bread consumer groups: (i) non-consumers of bread (NCB) (n=56), subjects who consume ≤3 portions of any type of bread monthly (1 portion=75gr); (ii) white-bread consumers (WHB) (n=48), subjects with at least 1 portion/day intake of white bread (≤3 portions of whole-grain bread monthly); and (iii) whole-grain bread consumers (WGB) (n=51), subjects with at least 1 portion/day intake of whole-grain bread (≤3 portions of white bread monthly). The selection of subjects took into account that all groups were homogeneous by gender, age (±5 years), smoking habit, and fruit and vegetable consumption (±200 gr/day). Considering the criteria for sample selection by bread consumption category, a subgroup of 155 subjects were evaluated in the study: 56 males

and 99 females with a mean age (±SD) of 68±6 years and a mean BMI (±SD) of 29±4 kg/m^2.

2.3. Urine samples collection and preparation

Morning fasting spot urine samples were collected, aliquoted, encoded and frozen at -80 ºC until used (Martinez-Gonzalez et al. 2012). On the day of analysis, urine samples were placed in the refrigerator for gradual thawing and then a 150 μL aliquot of each sample was centrifuged at 12,000 rpm for 5 min at 4 ºC. A 50 μL of the supernatant was diluted 1:1 with Milli-Q water, vortexed and placed into a 96-well plate for HPLC-q-TOF-MS analysis (Llorach et al. 2009). No preservatives were used at any stage of the collection and manipulation of the urine samples.

2.4. HPLC-q-TOF-MS analysis and data acquisition

HPLC-q-TOF-MS analysis was performed using an Agilent 1200 Series Rapid Resolution HPLC system coupled to a hybrid quadrupole TOF QSTAR Elite (AB Sciex). 15 μL of the prepared samples were injected using a thermostatted auto-injector at 4 ºC into an RP Luna® C18(2) column (5 μm, 50 x 2.0mm; Phenomenex, Torrance, CA, USA). The MS acquisition was performed in positive and negative modes using a TurboIonSpray source in full scan mode, within the m/z range 70-700. The parameters of chromatography and mass spectrometer used in the present study were developed previously by our group and reported elsewhere (Llorach et al. 2009; Tulipani et al. 2011). In order to avoid possible bias, batches were equilibrated to represent the variability of all the samples to be analysed (i.e. samples of each batch had similar characteristics, like the number of males or the number of each bread consumption group, among others), and sequences of injections were randomized. Throughout the whole data analysis process, four types of quality control (QC), described and applied in previous studies (Tulipani et al. 2011), were analysed in order to monitor the system stability and functionality, in addition to evaluation of the quality and the reproducibility of the acquired data (Tulipani et al. 2011).

The MS/MS confirmatory experiments were performed using the same mass and chromatographic conditions applying collision energy from +20 V to +38 V, depending on the analysed compound.

2.5. Data processing

The HPLC-q-TOF-MS raw data were extracted and aligned using MarkerView TM 1.2.1. software (AB Sciex, Toronto, Ontario, Canada). Peak detection was performed using a minimum spectral peak width of 1 ppm, a minimum RT peak width of 3 scans, a noise threshold of 5, and a subtraction multiple factor of 1.5. Alignment used 0.05 Da mass tolerance and 0.07 min RT tolerance windows. Consequently, a data set containing 5000 mass features present in at least 12 samples was obtained and then subjected to multivariate analysis. Data from positive and negative ionization modes were included in two separate data sets in order to analyse them individually.

2.6. Multivariate statistical analysis

The mass signals data sets were log-transformed and Pareto-scaled before their multivariate statistical analysis using SIMCA-P+13.0 software (Umetrics, Umeå, Sweden). Principal component analysis (PCA) was applied to evaluate the quality of the data acquisition. Partial least squares discriminant analysis with orthogonal signal correction (OSC-PLS-DA) was used to explore the differences in metabolomes among the three bread consumer groups (Llorach-Asuncion et al. 2010). OSC filtration was used to reduce the variability not associated with dietary classification, as has been done in other recently published nutrimetabolomic studies (Pujos-Guillot et al. 2013). The quality of the models was evaluated by $R^2Y(cum)$ and $Q^2(cum)$ parameters (Llorach-Asuncion et al. 2010). Validation of the models was evaluated by a permutation test (n=200). Those mass signals with variable importance for projection (VIP) values ≥ 1.5 in at least one component of the OSC-PLS-DA model were selected as the most relevant to explain the differences in metabolic profiles.

2.7. Identification of metabolites

The differential mass features selected after multivariate analysis were identified by a multistep procedure (Tulipani et al. 2011). First, clustering analysis with Pearson distance and Ward's minimum variance (PermutMatrix 1.9.3.0 software) was applied in order to identify the mass features corresponding to the same metabolite (source-generated ion fragments, isotopes and adducts). Then, metabolites were tentatively identified on the basis of their exact mass information (±5 mDa of accepted mass difference) for corresponding source-generated ion fragments. Following this, MS/MS confirmatory experiments were carried out. These data were compared to those from the scientific bibliography and the following databases: MetFrag (http://msbi.ipb-halle.de/MetFrag/), MassBank (http://www.massbank.jp/), Metlin (http://metlin.scripps.edu/index.php) and Human Metabolome Database (HMDB; http://www.hmdb.ca). Additionally, an in-house database mainly focused on the metabolites expected from the intake of dietary phytochemicals was also queried. Finally, the biological interpretation was performed using information from published research reports and from online databases such as HMDB and Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp). The level of identification was assigned according to the classification of the Metabolomics Standard Initiative (Sumner et al. 2007).

2.8. Univariate analysis of identified metabolites and development of combined urinary biomarker models

Data from OSC-PLS-DA analyses were further analysed by univariate methods using R 3.0.2 software (R Core Team 2014). The Kolmogorov-Smirnov test was used to check the normal data distribution. Features with skewed distributions were analysed by the Kruskal-Wallis test, whereas for features with normal distributions the ANOVA test was used. Differences among groups were determined using the Mann-Whitney or the t-Test comparisons for non-parametric or parametric tests, respectively.

Finally, to investigate whether the combination of more than one biomarker from food metabolome improves the discrimination of bread consumption groups, we conducted forward stepwise logistic regression analyses (IBM SPSS Statistics 20 software, SPSS Inc., Chicago, IL, USA). Three separate models were performed in order to compare NCB

versus WHB, NCB *versus* WGB, and WHB *versus* WGB. Based on Qi et al. (2012), each model was constructed using a dichotomous variable on bread consumption as dependent variable and identified metabolites from food metabolome as independent variable. Correlations between bread intake and the combined models were evaluated using Spearman's rank-correlation coefficient.

2.9. Evaluation of biomarkers

The global performance of each identified metabolite and biomarker model was evaluated by receiver operating (ROC) curves (Robin et al. 2011), using area under the curve (AUC) value. The optimum cut-off for sensitivity and specificity of the biomarkers was determined as the minimum distance to the top-left corner (Xia et al. 2013).

3. RESULTS AND DISCUSSION

3.1 Data processing and quality control

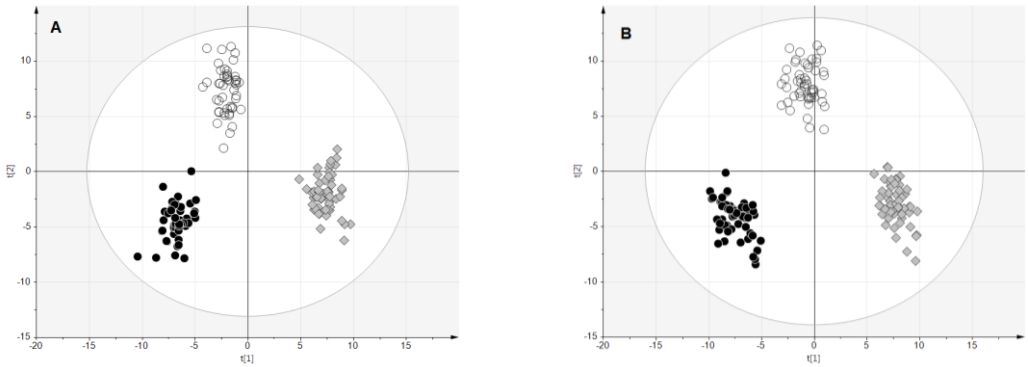
A PCA was used for the evaluation of the quality of the data acquisition analysing the variability of the urine and QC samples. As shown in Supporting Information Figure S1, replicates of each QC type were positioned clustered among themselves. The PCA on the urine samples showed 8 outlying samples in the positive ionization mode data set (data not shown). To avoid a possible effect of these samples during data processing, they were removed from the raw data (before peak picking process).

3.2. Urinary metabolome analysis

To identify differences associated with bread consumption in the urinary metabolome among the NCB, WHB and WGB groups an OSC filter was applied before PLS-DA analysis. The OSC-PLS-DA analysis (Figure 1) resulted in a 2-latent variable model with $R^2Y(cum)$ and $Q^2(cum)$ values of 0.93 and 0.73, respectively, for positive ionization mode, and 0.94 and 0.70 for negative ionization mode, indicating that both models were able to classify each subject in the correct bread consumption group. The permutation tests showed negative Q^2 intercepts (-0.23 and -0.22 for positive and negative ionization

modes, respectively), validating the model (Llorach-Asuncion et al. 2010; Pujos-Guillot et al. 2013). In both OSC-PLS-DA models, the NBC group was separated from the two bread consumption groups in the first latent variable, whereas the second one separated the WHB group from the other two.

Fig. 1 Scores plot of OSC-PLS-DA model obtained from urine samples for (A) positive and (B) negative ionization modes. The coloured dots correspond to each bread consumption group (NCB: grey diamonds; WHB: white circles; WGB: black circles).



3.3. Detection and identification of bread consumption biomarkers

In order to select the most discriminative urinary metabolome markers of each bread consumption group, a cut-off of ≥1.5 for VIP value was applied. On the basis of clustered mass features, exact masses for in-source fragments and experimental MS/MS fragmentations, a total of eighteen metabolites were tentatively identified. They are summarized in Table 1. Additionally, Table S1 (Supporting Information) summarizes clustered mass features of tentatively identified compounds and Figure S2 (Supporting Information) provides the mass spectra of the MS/MS experiments conducted with the urine samples. These discriminating metabolites were classified into 2 groups, namely food metabolome compounds and endogenous metabolites. All identified markers from food metabolome showed higher urinary levels in regular-bread consumers, whereas one metabolite from endogenous metabolome (2,8-dihydroxyquinoline glucuronide) was more abundant in WGB consumers and another one (N-α-acetylcitrulline) was more abundant in the NCB group. Table 1 also shows the significant identified metabolite differences among the three groups.

Table 1. Metabolites identified in human subjects with different bread consumption levels.

RT (min)	Detected mass (m/z)	Assignation	Metabolite tentatively identification	Level of identification[1]	NBC vs WHB	NBC vs WGB	WHB vs WGB
			Benzoxazinoid-related compounds				
0.88	188.0049	$[M - H]^-$	2-aminophenol sulphate	III	↑	↑	-
1.48	328.1036	$[M + H]^+$	HPAA glucuronide[2]	III	↑	↑	-
	326.0851	$[M - H]^-$			-	↑	-
2.07	168.0609	$[M + H]^+$	HHPAA	III	-	↑	↑
3.40	372.0925	$[M + H]^+$	HMBOA glucuronide[3]	III	↑	-	-
	370.0772	$[M - H]^-$			↑	↑	-
3.68	326.0922	$[M - H]^-$	HBOA glycoside	III	-	↑	↑
3.72	152.0671	$[M + H]^+$	HPPA	III	-	↑	-
4.78	196.0596	$[M + H]^+$	HMBOA[3]	III	↑	↑	-
	194.0410	$[M - H]^-$			↑	↑	-
			Alkylresorcinol metabolites				
2.85	357.0791	$[M - H]^-$	DHPPA glucuronide	III	↑	↑	↑
3.12	233.0118	$[M - H]^-$	3,5-Dihydroxyphenylethanol sulphate	III	-	↑	-
5.75	289.0412	$[M - H]^-$	DHPPTA sulphate	III	-	↑	↑
			Microbial-derived metabolites				
3.67	313.0558	$[M - H]^-$	Hydroxybenzoic acid glucuronide	III	↑	↑	-
4.72	275.0219	$[M - H]^-$	Dihydroferulic acid sulphate	III	-	↑	↑
6.32	299.1278	$[M + H - \text{glucuronide}]^+$	Enterolactone glucuronide[4]	III	-	↑	↑
	473.1447	$[M - H]^-$			-	↑	↑

RT (min)	Detected mass (m/z)	Assignation	Metabolite tentatively identification	Level of identification[1]	NBC vs WHB	NBC vs WGB	WHB vs WGB
			Markers of heat-treated food products				
2.73	255.1345	$[M + H]^+$	Pyrraline[5]	II	-	↑	-
	253.1172	$[M – H]^-$			-	↑	↑
			Other exogenous metabolites				
3.25	338.0871	$[M + H]^+$	3-Indolecarboxylic acid glucuronide[6]	II	-	↑	↑
	336.0697	$[M – H]^-$			-	↑	↑
4.65	377.1475	$[M + H]^+$	Riboflavin[7]	II	↑	↑	↑
			Endogenous metabolites		-		
0.63	218.1140	$[M + H]^+$	N-α-Acetylcitrulline[3,8]	III	-	↓	-
4.20	338.0882	$[M + H]^+$	2,8-Dihydroxyquinoline glucuronide[9]	II	-	↑	↑
	160.0382	$[M – H – \text{glucuronide}]^-$			-	↑	↑

All features in the table were statistically significant. The significance level was obtained by Kruskal-Wallis analysis or by analysis of variance (ANOVA), according to their distribution. Pairwise comparisons were analysed by the Mann-Whitney test or the t-Test for non-parametric and parametric tests, respectively. ↑, indicates significantly higher levels in the second group of the comparison; ↓, indicates significantly lower levels in the second group of the comparison.

[1]Level of identification has been assigned according to Sumner et al. 2007. Structural identification agrees with Beckmann et al. 2013[2], MetFrag[3], Metlin: MID64734[4], Hellwig and Henle 2012[5], Andersen et al. 2014[5], Dong et al. 2012[6], MassBank Record PR100399[7], MassBank Record BML01052[8], and Zhen et al. 2007[9] (MS/MS information in Supporting Information Figure S2; and LC-MS pattern information in Supporting Information Table S1).

Abbreviations: DHPPA, 3-(3,5-dihydroxyphenyl) propanoic acid; DHPPTA, 5-(3,5-dihydroxyphenyl) pentanoic acid; HBOA, 2-hydroxy-1,4-benzoxazin-3-one; HHPAA, 2-hydroxy-N-(2-hydroxyphenyl) acetamide ; HMBOA, 2-hydroxy-7-methoxy-2H-1,4-benzoxazin-3-one; HPAA, N-(2-hydroxyphenyl) acetamide; HPPA, 2-hydroxy-N-(2-hydroxyphenyl) acetamide; NCB, non-consumers of bread; RT, retention time; WGB, whole-grain bread consumers; WHB, white-bread consumers.

3.4. Biological interpretation

With regard to the markers belonging to food metabolome, it is important to highlight that most of them were related to compounds derived from cereal plant phytochemicals such as benzoxazinoids, alkylresorcinols and other compounds produced by gut microbiota. Specifically, urinary levels of 2-hydroxy-7-methoxy-2H-1,4-benzoxazin-3-one (HMBOA) glucuronide, 2-hydroxy-1,4-benzoxazin-3-one (HBOA) glycoside and HMBOA were higher in both types of regular bread consumers than in NCB. These compounds are secondary metabolites of wheat and belong to the benzoxazinoids (Moraes et al. 2008). Recent metabolomic studies have reported an increase in urinary excretions of some benzoxazinoid metabolites in human intervention studies with different types of whole-grain breads (Beckmann et al. 2013; Bondia-Pons et al. 2013) and after a 5-week dietary intervention with high fibre doses (Johansson-Persson et al. 2013). Furthermore, Adhikari et al. (2013) analysed for the first time the absorption, metabolism and excretion of dietary benzoxazinoids in humans and reported that these compounds are bioavailable in the human organism after their dietary intake, adding more scientific evidence of their role as potential bioactive compounds from bread (Adhikari et al. 2013). Additionally, some precursors of benzoxazinoids such as 2-aminophenol sulphate and hydroxylated phenylacetamides [N-(2-hydroxyphenyl) acetamide (HPAA) glucuronide, 2-hydroxy-N-(2-hydroxyphenyl) acetamide (HHPAA) and HPAA] were also increased in bread consumption groups.

Interestingly, two alkylresorcinol metabolites [3-(3,5-dihydroxyphenyl) propanoic acid (DHPPA) glucuronide and 5-(3,5-dihydroxyphenyl) pentanoic acid (DHPPTA) sulphate] were also tentatively identified. Their levels were higher among WGB consumers. Related to this type of compound, 3,5-dihydroxyphenylethanol sulphate and enterolactone glucuronide were also characteristic of the WGB group in concordance with recently published results by Bondia-Pons et al. (2013). The other microbial-derived compounds excreted in higher amounts by bread consumers were hydroxybenzoic acid glucuronide and dihydroferulic acid sulphate. Microbial metabolites are directly related to dietary consumption (Moco et al. 2012). In addition, they may also be influenced by individual gut microbiota composition. Due to interindividual variation in microbiota composition (Lozupone et al. 2012), the sensibility and specificity of the microbial metabolites used for bread exposure prediction can be reduced. However, their

combination in biomarker panels from food metabolome could provide more accurate measurements of dietary exposure, as it is displayed in the next section.

Pyrraline, a compound related to the Maillard reaction that can be absorbed after its intake through heat-treated foodstuffs, such as bakery products like bread (Hellwig et al. 2012), was also tentatively identified. It is a derivative compound from the amino acid lysine formed during the Maillard reaction which provides the aroma and palatable traits specific for heat-treated foods like bread (Hellwig and Henle 2012). Thus, since pyrraline can be characteristic for various types of heat-treated foods, it alone cannot be a specific biomarker of bread consumption (see Table S3 in the next section). In fact, bread from whole flour has a high protein content and low level of lysine blockage (Michalska et al. 2008). This could explain why pyrraline was more characteristic of WGB group than WHB. Its inclusion in the combined biomarker models (discussed in the next section) provides additional and complementary information reinforcing the capacity to distinguish between different habitual bread consumption patterns and improving the sensibility and specificity to predict habitual bread exposure.

Another detected marker that could be related to food metabolome was riboflavin. This compound is a B vitamin with an important role in the intermediate energy metabolism (Depeint et al. 2006) present in bread and wheat compositions (Spanish Food Composition Database (BEDCA) 2010). Therefore, the higher levels of riboflavin detected in regular-bread consumers could come directly from bread consumption. Finally, WGB group had greater levels of 3-indole carboxylic acid (3-ICA) glucuronide, a compound related to tryptophan metabolism (Roux et al. 2012). In this context, 3-ICA has been described as an end metabolite synthesized from conversions of tryptophan to indole-3-carboxaldehyde (Magnus et al. 1982). However, this amino acid has also been detected in whole-grain bread (Jiang and Peterson 2013). Thus, the 3-ICA glucuronide could be generated exogenously and/or endogenously.

Regarding urinary endogenous metabolites, this study showed that the compound 2,8-dihydroxyquinoline-ß-D-glucuronide (2,8-DHQ glucuronide) is present in a higher proportion in WGB consumers than in other two groups. Variations of this compound have been observed in previous metabolomics studies (Li et al. 2013; Zhen et al. 2007). This compound was significantly elevated in the urine of mice treated with a peroxisome proliferator-activated receptor (PPARα) agonist, suggesting that this metabolite could be

a potential urinary biomarker of PPARα agonists (Zhen et al. 2007). In this context, the activation of PPARα through diet or by synthetic agonists may cause an improvement in lipid profile and by extension impact on cardiovascular health. More recently, Li *et al.* also observed an increase in 2,8-DHQ-β-D-glucuronide urinary excretion after treatment with a drug that previously had shown effects on glucose and lipid metabolism (Li et al. 2013). The greater levels of 2,8-DHQ-β-D-glucuronide observed among WGB consumers suggests a potential mechanism for previously observed associations between whole-grain foods and lipid metabolism in epidemiological studies (McKeown et al. 2002) through PPARα. In contrast with the other tentatively identified metabolites, N-a-acetylcitrulline (NAC) urine excretion was higher among NCB than in WGB. In the urea cycle, NAC takes part in the arginine synthesis. Arginine is an amino acid that has been related to improvements in parameters related to cardiovascular disease (Tousoulis et al. 2002). The higher concentrations of NAC observed in NCB could be associated with a lower activity of the arginine synthesis pathway, which could have some mechanistic link with alterations in some atherosclerotic risk factors.

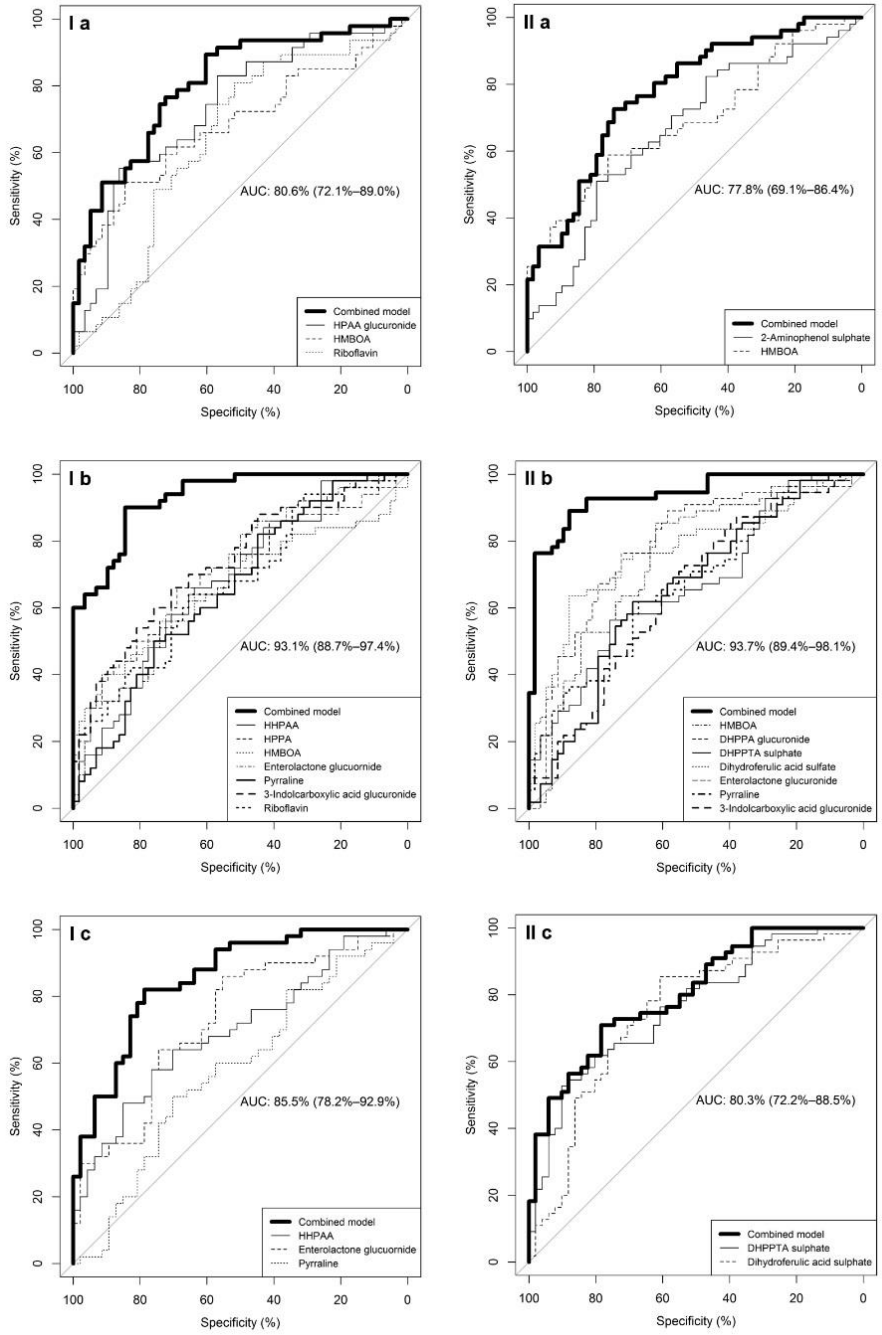
3.5. Combination of detected biomarkers to improve the prediction of bread intake

In order to assess whether one of the food metabolome compounds or a combination of them can effectively discriminate among groups of bread consumers, stepwise logistic regressions were performed. As a result, three different combination models for each ionization mode were developed. Table S2 (Supporting Information) shows the metabolites included in all combined biomarker models. As can be seen in this table, in the positive ionization mode HMBOA and riboflavin were selected for NCB *versus* both bread consumption groups (in negative ionization mode only HMBOA fulfils this characteristic), whereas HHPAA, enterolactone glucuronide and pyrraline contributed to the discrimination of WGB consumers from the other two groups (DHPPTA sulphate and dihydroferulic acid sulphate for negative ionization mode). The reported daily bread intake significantly correlated with values of combined models (in all cases, $p<0.001$). This association was strong for NCB *versus* WGB models (0.70 and 0.74 for positive and negative ionization modes, respectively), and moderate for NCB *versus* WHB (0.46 and 0.44 for positive and negative ionization modes, respectively) and WHB *versus* WGB (0.60 and 0.48 for positive and negative ionization models, respectively) models

(Martinez-Gonzalez et al. 2006). Figure 2 displays ROC curves for the designed models and for all individual metabolites that took part in each combined model. As can be seen, the best discrimination was observed between the NCB and WGB groups (AUC>90%), whereas the performance of the other combined models was lower than the former (AUC between 75% and 90%) (Xia et al. 2013). However, all combined biomarker models had higher AUC values than any individual metabolite, whose AUC were <80%. Table S3 summarizes the parameters of all ROC curves presented in Figure 2. It is important to point out that the combination of the detected food metabolome compounds in biomarker models by stepwise logistic regression improved the predictive power of discriminative analysis for bread consumer groups. The use of these combinations of metabolites could provide more accurate information about the habitual bread intake in free-living populations or the level of exposition to a dietary intervention with bread.

Finally, for the interpretation of these results it is important to take into account that no specification either on cereal origin (e.g., wheat, rye, barley…), or on technological treatments (e.g., bakery bread, baguette, sliced bread…) of consumed breads was discerned in the administered FFQ. Moreover, the participants of the present study were elderly at high cardiovascular risk and not representative of the general population, so results should be extrapolated with some caution to other populations. Thus, it could be interesting to quantify these metabolites in different populations in order to know if they could be markers of bread intake whatever individuals were studied. In spite of these limitations, it is important to underline that the included subjects were free-living volunteers, who provided highly reliable data. Furthermore, it is also important to highlight the relatively large number of subjects included in this non-targeted HPLC-q-TOF-MS nutrimetabolomics study.

Fig. 2 Receiver operating characteristic (ROC) curves of combined models and of identified metabolites from food metabolome in the positive (I) and negative (II) ionization modes for (a) NCB *versus* WHB, (b) NCB *versus* WGB, and (c) WHB *versus* WGB comparisons.



Abbreviations: DHPPA, 3-(3,5-dihydroxyphenyl) propanoic acid; DHPPTA, 5-(3,5-dihydroxyphenyl) pentanoic acid; HHPAA, 2-hydroxy-N-(2-hydroxyphenyl) acetamide; HMBOA, 2-hydroxy-7-methoxy-2H-1,4-benzoxazin-3-one; HPAA, N-(2-hydroxy-phenyl)acetamide; HPPA, 2-hydroxy-N-(2-hydroxyphenyl) acetamide.

4. CONCLUDING REMARKS

The results of the present study showed that a daily bread intake is reflected in urinary metabolome. The variation in metabolome of the regular consumers of bread was associated with eighteen metabolites, which were related mainly to food metabolome and to endogenous metabolomic differences.

Globally, the higher urinary excretion of food metabolome metabolites could discriminate regular bread consumers from non-consumers in a free-living uncontrolled population. The combination of more than one of these metabolites into a complex biomarker model, improved the predicting capacity of habitual bread consumption exposure. Therefore, we anticipate that the proposed food metabolome biomarkers, especially as combined models, could be of great support in accurate and precise evaluation and confirmation of bread intake in addition to traditional dietary assessment methods.

Differences in endogenous metabolites might reveal unexplored mechanisms responsible for the beneficial effects of bread consumption previously observed in epidemiological studies and, thus, would help to generate new hypotheses on the relationship between the intake of bread and health. In this case, the greater levels of 2,8-DHQ-β-D-glucuronide observed among WGB consumers could suggest novel hypotheses about the biological mechanisms involved in the beneficial effects observed for whole-grain cereals through PPARα receptors, which requires further investigation.

ACKNOWLEDGEMENTS

This research was supported by Spanish national grants from the Ministry of Economy and Competitiveness (MINECO), as well as FEDER (Fondo Europeo de Desarrollo Regional): AGL2009-13906-C02-01, AGL2010-10084-E, CONSOLIDER INGENIO 2010 Programme: FUN-C-FOOD-CSD2007-063, and ISCIII-CIBEROBN, Merck Serono Research Grants 2010 from Fundación Salud 2000, and by “Pan cada día” open call promoted by the Scientific Committee of Bread and by INCERHPAN. The "CIBER de Fisiopatología de la Obesidad y Nutrición" (CIBEROBN) is an initiative of the Instituto de Salud Carlos III, Madrid, Spain. 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” (RYC-2010–07334 and RYC-2011-09677, respectively) and ST the “Juan de la Cierva” program, both programmes from MINECO and Fondo Social Europeo (FSE). MF was funded by a contract from the Catalan Government and the Instituto de Salud Carlos III FEDER (FIS CP06/00100). None of the funding sources had any involvement in the study or data analysis.

The authors declare that they have no conflict of interest.

REFERENCES

Adhikari, K. B., B. B. Laursen, P. L. Gregersen, et al. (2013). Absorption and metabolic fate of bioactive dietary benzoxazinoids in humans. *Mol Nutr Food Res.* 57, 1847-1858.

Andersen, M. B., A. Rinnan, C. Manach, et al. (2014). Untargeted Metabolomics as a Screening Tool for Estimating Compliance to a Dietary Pattern. *J Proteome Res.* 13, 1405-1418.

Bach-Faig, A., E. M. Berry, D. Lairon, et al. (2011). Mediterranean diet pyramid today. Science and cultural updates. *Public Health Nutr.* 14, 2274-2284.

Bautista-Castano, I., A. Sanchez-Villegas, R. Estruch, et al. (2012a). Changes in bread consumption and 4-year changes in adiposity in Spanish subjects at high cardiovascular risk. *Br J Nutr.* 110, 337-346.

Bautista-Castano, I., L. Serra-Majem (2012b). Relationship between bread consumption, body weight, and abdominal fat distribution: evidence from epidemiological studies. *Nutr Rev.* 70, 218-233.

Beckmann, M., A. J. Lloyd, S. Haldar, et al. (2013). Hydroxylated phenylacetamides derived from bioactive benzoxazinoids are bioavailable in humans after habitual consumption of whole grain sourdough rye bread. *Mol Nutr Food Res.* 57, 1859-1873.

Bondia-Pons, I., T. Barri, K. Hanhineva, et al. (2013). UPLC-QTOF/MS metabolic profiling unveils urinary changes in humans after a whole grain rye versus refined wheat bread intervention. *Mol Nutr Food Res.* 57, 412-422.

Bondia-Pons, I., E. Nordlund, I. Mattila, et al. (2011). Postprandial differences in the plasma metabolome of healthy Finnish subjects after intake of a sourdough fermented endosperm rye bread versus white wheat bread. *Nutr J.* 10, 116.

Depeint, F., W. R. Bruce, N. Shangari, et al. (2006). Mitochondrial function and toxicity: role of the B vitamin family on mitochondrial energy metabolism. *Chem Biol Interact.* 163, 94-112.

Dong, H., A. Zhang, H. Sun, et al. (2012). 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.* 8, 1206-1221.

Estruch, R., E. Ros, J. Salas-Salvado, et al. (2013). Primary prevention of cardiovascular disease with a Mediterranean diet. *N Engl J Med.* 368, 1279-1290.

Fernandez-Ballart, J. D., J. L. Pinol, I. Zazpe, et al. (2010). Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. *Br J Nutr.* 103, 1808-1816.

Gil, A., R. M. Ortega, J. Maldonado (2011). Wholegrain cereals and bread: a duet of the Mediterranean diet for the prevention of chronic diseases. *Public Health Nutr.* 14, 2316-2322.

Hellwig, M., T. Henle (2012). Quantification of the Maillard reaction product 6-(2-formyl-1-pyrrolyl)-l-norleucine (formyline) in food. *Eur Food Res Technol.* 235, 99-106.

Ismail, N. A., J. M. Posma, G. Frost, et al. (2013). The role of metabonomics as a tool for augmenting nutritional information in epidemiological studies. *Electrophoresis.* 34, 2276-2286.

Jiang, D., D. G. Peterson (2013). Identification of bitter compounds in whole wheat bread. *Food Chem.* 141, 1345-1353.

Johansson-Persson, A., T. Barri, M. Ulmius, et al. (2013). LC-QTOF/MS metabolomic profiles in human plasma after a 5-week high dietary fiber intake. *Anal Bioanal Chem.* 405, 4799-4809.

Lankinen, M., U. Schwab, T. Seppanen-Laakso, et al. (2011). Metabolomic analysis of plasma metabolites that may mediate effects of rye bread on satiety and weight maintenance in postmenopausal women. *J Nutr.* 141, 31-36.

Li, F., X. Pang, K. W. Krausz, et al. (2013). Stable Isotope- and Mass Spectrometry-based Metabolomics as Tools in Drug Metabolism: A Study Expanding Tempol Pharmacology. *J Proteome Res.* 12, 1369-1376.

Llorach-Asuncion, R., O. Jauregui, M. Urpi-Sarda, C. Andres-Lacueva (2010). 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.* 51, 373-381.

Llorach, R., M. Garcia-Aloy, S. Tulipani, et al. (2012). Nutrimetabolomic strategies to develop new biomarkers of intake and health effects. *J Agric Food Chem.* 60, 8797-8808.

Llorach, R., M. Urpi-Sarda, O. Jauregui, et al. (2009). An LC-MS-based metabolomics approach for exploring urinary metabolome modifications after cocoa consumption. *J Proteome Res.* 8, 5060-5068.

Lozupone, C. A., J. I. Stombaugh, J. I. Gordon, et al. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature.* 489, 220-230.

Magnus, V., S. Simaga, S. Iskric, S. Kveder (1982). Metabolism of Tryptophan, Indole-3-acetic Acid, and Related Compounds in Parasitic Plants from the Genus Orobanche. *Plant Physiol.* 69, 853-858.

Martinez-Gonzalez, M. A., D. Corella, J. Salas-Salvado, et al. (2012). Cohort profile: design and methods of the PREDIMED study. *Int J Epidemiol.* 41, 377-385.

Martinez-Gonzalez, M. A., S. Palma, E. Toledo (2006). Correlación y regresión. In: Martinez-Gonzalez, M. A. (ed) *Bioestadística Amigable*. Spain: Díaz de Santos.

Mataix Verdú, J. (2003). Tabla de Composición de Alimentos. Granada, Universidad de Granada.

McKeown, N. M., J. B. Meigs, S. Liu, et al. (2002). Whole-grain intake is favorably associated with metabolic risk factors for type 2 diabetes and cardiovascular disease in the Framingham Offspring Study. *Am J Clin Nutr.* 76, 390-398.

Michalska, A., M. Amigo-Benavent, H. Zielinski, M. D. del Castillo (2008). Effect of bread making on formation of Maillard reaction products contributing to the overall antioxidant activity of rye bread. 48, 123-132.

Moazzami, A. A., I. Bondia-Pons, K. Hanhineva, et al. (2012). Metabolomics reveals the metabolic shifts following an intervention with rye bread in postmenopausal women--a randomized control trial. *Nutr J.* 11, 88.

Moco, S., F. P. Martin, S. Rezzi (2012). Metabolomics view on gut microbiome modulation by polyphenol-rich foods. *J Proteome Res.* 11, 4781-4790.

Moraes, M. C., M. A. Birkett, R. Gordon-Weeks, et al. (2008). cis-Jasmone induces accumulation of defence compounds in wheat, Triticum aestivum. *Phytochemistry.* 69, 9-17.

Nielsen, K. L., M. L. Hartvigsen, M. S. Hedemann, et al. (2014). Similar metabolic responses in pigs and humans to breads with different contents and compositions of dietary fibers: a metabolomics study. *Am J Clin Nutr.* 99, 941-949.

Pujos-Guillot, E., J. Hubert, J. F. Martin, et al. (2013). Mass spectrometry-based metabolomics for the discovery of biomarkers of fruit and vegetable intake: citrus fruit as a case study. *J Proteome Res.* 12, 1645-1659.

Quilez, J., J. Salas-Salvado (2012). Salt in bread in Europe: potential benefits of reduction. *Nutr Rev.* 70, 666-678.

R Core Team (2014). R: A language and environment for statistical computing. In: R Foundation for Statistical Computing. http://www.R-project.org/

Robin, X., N. Turck, A. Hainard, et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 12, 77.

Roux, A., Y. Xu, J. F. Heilier, et al. (2012). 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.* 84, 6429-6437.

Sofi, F., R. Abbate, G. F. Gensini, A. Casini (2010). Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis. *Am J Clin Nutr.* 92, 1189-1196.

Spanish Agency for Food Safety and Nutrition (2011). ENIDE: National Survey of Dietary Intake (2009-2010): Results of consumer data. In: AESAN.

http://www.aesan.msc.es/AESAN/docs/docs/evaluacion_riesgos/datos_consumo/ENIDE.pdf Accessed 23 January 2013

Spanish Food Composition Database (BEDCA) (2010). BEDCA Database v1.0. In: RedBEDCA, AESAN. http://www.bedca.net Accessed 23 January 2013

Sumner, L. W., A. Amberg, D. Barrett, et al. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics.* 3, 211-221.

Tousoulis, D., C. Antoniades, C. Tentolouris, et al. (2002). L-arginine in cardiovascular disease: dream or reality? *Vasc Med.* 7, 203-211.

Tulipani, S., R. Llorach, O. Jauregui, et al. (2011). Metabolomics unveils urinary changes in subjects with metabolic syndrome following 12-week nut consumption. *J Proteome Res.* 10, 5047-5058.

Williams, P. G. (2012). Evaluation of the evidence between consumption of refined grains and health outcomes. *Nutr Rev.* 70, 80-99.

Xia, J., D. I. Broadhurst, M. Wilson, D. S. Wishart (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics.* 9, 280-299.

Zhen, Y., K. W. Krausz, C. Chen, et al. (2007). Metabolomic and genetic analysis of biomarkers for peroxisome proliferator-activated receptor alpha expression and activation. *Mol Endocrinol.* 21, 2136-2151.