Journal of proteome research

pubs.acs.org/jpr

Article

Nutrivolatilomics of Urinary and Plasma Samples to Identify Candidate Biomarkers after Cheese, Milk, and Soy-Based Drink Intake in Healthy Humans

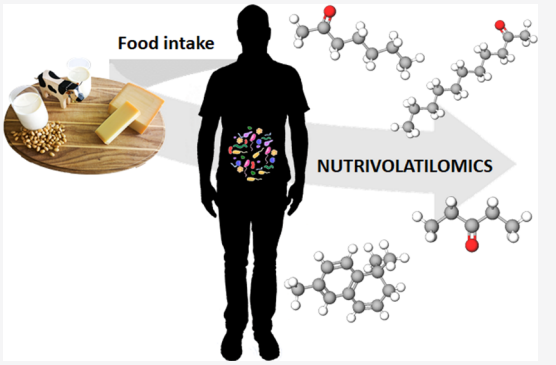
Pascal Fuchsmann,* Mireille Tena Stern, Linda H. Münger, Grégory Pimentel, Kathryn J. Burton, Nathalie Vionnet, and Guy Vergères

Cite This: https://dx.doi.org/10.1021/acs.jproteome.0c00324

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The characterization of volatile compounds in biological fluids offers a distinct approach to study the metabolic imprint of foods on the human metabolome, particularly to identify novel biomarkers of food intake (BFIs) that are not captured by classic metabolomics. Using a combination of dynamic headspace vacuum transfer In Trap extraction and gas chromatography coupled with mass spectrometry, we measured volatile compounds (the "volatilome") in plasma and urine samples from a randomized controlled crossover intervention study in which 11 healthy subjects ingested milk, cheese, or a soy-based drink. More than 2000 volatile compounds were detected in plasma, while 1260 compounds were detected in urine samples. A postprandial response in plasma was confirmed for 697 features. Univariate and multivariate analyses identified four molecules in plasma and 31 molecules in urine samples differentiating the ingestion of the foods, of which three metabolites in plasma and nine in urine were specific to the dairy products. Among these molecules, heptan-2-one, 3,5-dimethyloctan-2-one, and undecan-2-one in plasma and 3-ethylphenol, heptan-2-one, 1-methoxy-2-propyl acetate, and 9-decenoic acid were highly discriminative for dairy or cheese intake. In urine, 22 volatile compounds were highly discriminative for soy-based drink intake. The majority of these molecules have not been reported in humans. Our findings highlight the potential of plasma and urinary volatilomics for detection of novel dietary biomarkers.



KEYWORDS: *biomarker, volatilomics, VTT extraction, nutrition, plasma metabolome, urinary metabolome, cheese, milk, soy-based drink*

INTRODUCTION

Linking dietary patterns to the health status of human populations or individuals requires that food intake be accurately determined both qualitatively and quantitatively. The methods classically used to monitor food intake, including food frequency questionnaires, food record (weighted/unweighted), and 24 h diet recall, have widely contributed to supporting dieticians in their practice and researchers in nutritional science.[1,2] These methods, however, suffer from limitations primarily derived from the fact that the analytical instrument used to produce the data is a human being delivering a subjective self-report of food intake. Modern tools that combine specific aspects of the eating process (e.g., visualizing food, moving food from plate to mouth, chewing) and sensory technologies (e.g., cameras, movement sensors, microphones) have now entered this field and promise significant breakthroughs in nutrition sciences for the next decade.[3,4] However, the molecular information of the ingested foods is not directly delivered by these methods, which rely on food composition databases for the interpretation of the collected intake data.

A complementary strategy, which can be applied to obtain information on food intake, is the assessment of the food metabolome. Scalbert et al. define the food metabolome as "the part of the human metabolome directly derived from the digestion and biotransformation of foods and their constituents".[5] Indeed, a large fraction of the molecules present in human tissues, in particular body fluids such as urine and blood, is derived from nutrients present in food and processed by the gastrointestinal tract (GIT) system through chewing, gastric and intestinal digestion, gut microbiota metabolism, intestinal transport, and liver metabolism. The sequential nature and complexity of these biological processes introduce variability in the ability of the organism to process food between individuals,[6] during aging,[7] and disease. The food

Received: May 14, 2020
Published: September 3, 2020

metabolome therefore provides a key source of dietary biomarkers, which integrate the processing of food by the human organism.

The last decade has seen dramatic improvement in metabolomic technologies[8,9] and their application to characterize the human metabolome.[10,11] Given the close relationship between the human and food metabolomes, these technologies are well-placed to identify, characterize, and validate specific biomarkers of food intake (BFIs). The Food Biomarker Alliance (FoodBAll) project funded by the Joint Programming Initiative "A Healthy Diet for a Healthy Life" (JPI-HDHL) has consequently taken on the task of developing,[12] validating,[13] and harmonizing[14] the field of nutrimetabolomics as a tool to establish BFIs. Three technologies cover most of the field of nutrimetabolomics, namely, liquid chromatography−mass spectrometry (LC−MS) for molecules of moderate sizes, which can be separated and ionized in their native form;[14,15] gas chromatography−mass spectrometry (GC−MS) for semi-volatile molecules, which can be transferred to the gas phase upon chemical modification;[14,16] and nuclear magnetic resonance (NMR) for molecules whose electromagnetic properties can be specifically filtered from the molecular background of the matrix investigated.[17,18]

The fourth version of the human metabolome database contains 114,100 entries,[10] and each of the above technologies, combined with specific methods of molecular separation, ionization, and mass analysis, can only cover parts of this metabolome. The use of a range of complementary analytical strategies is therefore key to identify a spectrum of biomarkers that reflect the molecular complexity of foods. The characterization of small volatile molecules in human body fluids, i.e., the human volatilome,[19,20] has recently attracted the interest of clinical researchers to identify new markers associated with a range of pathologies, including cancer[21] and asthma.[22]

Given that volatile molecules are often odor or aroma compounds, food scientists have frequently made use of the volatilome to develop foods with attractive sensory properties.[23−25] However, this class of molecules has, so far, surprisingly escaped the scrutiny of nutrition researchers with few human intervention studies reporting on the metabolic fate of odor and aroma compounds after dietary exposures such as asparagus intake (*S*-methyl thioesters),[26] strawberry (2,5-dimethyl-4-hydroxy-3[2*H*]-furanone beta-D-glucuronide),[27] or fennel tea (estragole).[28]

These studies have investigated dietary volatile compounds in a targeted manner. In 2019, Mack et al.[29] reported for the first time the successful application of untargeted volatilomics to the identification in urine of candidate BFIs, in particular for coffee. This approach was limited by the efficiency of the solid-phase microextraction (SPME) method, and compared to the classical LC−MS and GC−MS technologies, a relatively modest number of molecules (n = 138) could be reliably measured, thus limiting the potential ability of this analytical strategy to identify novel discriminatory molecules. In this context, we have recently developed a novel method using a dynamic headspace vacuum transfer In Trap extraction method (DHS-VTT) to more efficiently extract volatile compounds from food matrices or biological fluids.[30] This extraction method could, therefore, foster the application of volatilomics for the identification of BFIs.

During the FoodBAll project, we applied LC−MS, GC−MS, and NMR technologies to characterize the postprandial metabolome of blood and urine samples of healthy human subjects in response to the intake of milk, cheese, and a soy-based drink.[31−33] We now apply the DHS-VTT extraction method to investigate new candidate BFIs for these three foods in urine and blood.

EXPERIMENTAL SECTION

Study Design and Test Products

The human intervention study was conducted according to the protocols set up for the international FoodBAll project and in accordance with the Declaration of Helsinki. It was approved by the Ethical Committee of the Canton of Vaud (Switzerland) and was registered at clinicaltrials.gov (NCT02705560). Eleven volunteers (five women and six men) agreed to participate in the study and complied with the protocol (Figure 1). The recruitment procedure and selection criteria were

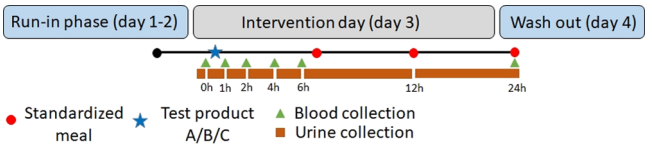


Figure 1. Design of the randomized, controlled, crossover study. The test products (milk A, cheese B, and soy-based drink C) were administered sequentially in random order during the intervention days and separated by one-day wash-out and two-day run-in phases with a restricted diet devoid of dairy products. Blood and urine samples were collected before and after food intake for up to 24 h.

described by Münger et al.[32] and follow the protocol of the FoodBAll consortium.[12] The study consisted of three two-day run-in phases, one for each test food in random order, characterized by a restricted diet (no dairy products, fermented products, foods of bovine origin, and soy products) and a standardized meal (rice and chicken), devoid of the foods tested in the study, for the diner preceding the intervention. Each of the three test phases (A, B, and C) consisted of the ingestion of the test food and the sampling of fasting and postprandial blood and urine for up to 24 h. The test products were as follows: A: 600 mL of organic pasteurized full fat milk (3.9% fat), B: 100 g of hard cheese (Le Gruyère AOP, Bulle, Switzerland), consumed with 500 mL of water, and C: 600 mL of a mixture of soy drink and vegetable fat (hydrogenated palm oil). All products consumed were isocaloric (400 kcal) and slightly differed in amounts of fat and protein.[32] Consumption of the test products had to be completed within 15 min from the start of the test. Blood samples were collected at fixed times at the study center. The first blood sampling (t = 0 h) was conducted shortly before ingestion of the test products began. Postprandial blood sampling was then conducted 1, 2, 4, and 6 h, as well as 24 h, after ingestion of the test foods. All participants had a standardized meal for lunch (after the 6 h sampling) and for dinner during the test days. Urine samples were collected at the study center during the postprandial time intervals 0−1, 1−2, 2−4, and 4−6 h and outside the study center at the time intervals 6−12 and 12−24 h. To obtain a control urine after fasting time, the participants were first asked to drink two glasses of tap water at home, after having discarded their first morning urine, and to collect their second urine at the study center prior to ingestion of the test foods. The processing and the storage of blood[31] and urine[32] samples were conducted as published previously.

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

Materials and Methods

Chemicals. All required chemical compounds were purchased from Sigma-Aldrich Chemie GmbH (Buchs, Switzerland).

Sample Preparation. An overview of the workflow for the preparation and analysis of the plasma and urine sample is shown in Figure 2. The plasma samples did not require

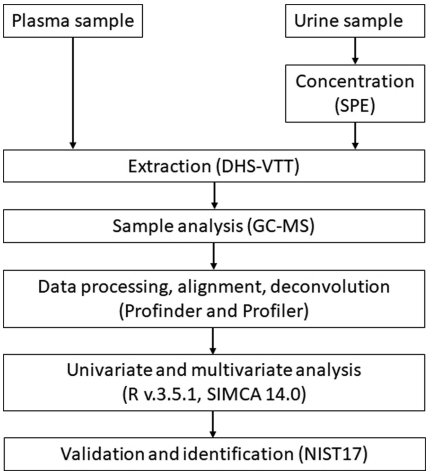


Figure 2. Workflow for the processing and data analysis of plasma and urine samples. SPE: solid-phase extraction, DHS-VTT: dynamic headspace vacuum transfer In Trap extraction, GC−MS: gas chromatography−mass spectrometry, NIST: National Institute of Standards and Technology.

normalization or preconcentration prior to analysis so that all samples could be directly analyzed in headspace. The plasma samples were prepared for analysis by defrosting at 4 °C overnight. After vortex agitation at room temperature, 500 μL of plasma was collected, added to a 20 mL headspace vial that was hermetically closed with a septum silicone Teflon (Interchim, France), and stored at 4 °C until analysis.

The urine samples prepared for the GC-volatile analyses were normalized based on the lowest specific gravity of all samples (SG = 1.0008) as described previously.[32] Specific gravity as a preacquisition normalization method was motivated by instrumental availability in our laboratories as well as the study conducted by Edmands et al.,[34] which demonstrates the efficiency of this method for the analysis of BFIs in urine samples. A 6 h pool was prepared for each subject by mixing equivalent volumes of normalized urine samples collected during the time intervals 0−1, 1−2, 2−4, and 4−6 h; a 24 h pool was prepared by adding equivalent volumes of the normalized samples collected during the time intervals 6−12 and 12−24 h to the 6 h pool. The separation of these two pools at 6 h was motivated by the fact that, in contrast to the 6 h pool collected under highly controlled conditions, the 24 h pool contained urine collected after ingestion of the standardized meals that was consumed ad libitum. Only the 6 and 24 h urine pools were prepared in order to maximize the possibility of finding candidate BFIs in sufficient quantities. These urine samples were defrosted at 4 °C overnight.

Pretests in our laboratory showed that the urine samples were too diluted after the normalization for headspace analysis without additional sample preparation (i.e., preconcentration). The samples thus required preconcentration on a solid-phase extraction (SPE) cartridge in order to be measured under optimal conditions. The preconcentration procedure, however, requested volumes of urine that were not available with all samples. The urine samples were therefore combined in two pools (6 and 24 h pools). The protocol for sample preparation using the same SPE cartridges was discussed elsewhere[35] and was applied with minor modifications to our samples.

For each pool, 700 μL of urine was acidified to pH 2.0 by addition of 700 μL of phosphoric acid (H_3PO_4 0.1 M). The samples were then concentrated through an SPE cartridge containing 30 mg of divinylbenzene-polymer Atoll Xtrem Capacity cartridge (Interchim, France). The cartridges were dried according to the manufacturer's instructions, and the compounds were eluted with 150 μL of methanol and before storage at 4 °C in 20 mL headspace vials hermetically sealed with a septum silicone Teflon (Interchim, France) until analysis.

Food samples were processed by two methods. To allow for comparison with plasma samples, 2 g of cheese, milk, and soy-based drink were directly extracted using DHS-VTT without pretreatment. To allow for comparison with urine samples, 2 g of each test food was acidified with 2 mL of H_3PO_4 (0.1 M). The samples were then centrifuged, and 1 mL of supernatant was concentrated through the SPE cartridge containing 30 mg of divinylbenzene polymer. The cartridges were dried, and the compounds were eluted with 150 μL of methanol and stored at 4 °C in 20 mL headspace vials hermetically sealed with a septum silicone Teflon until analysis.

Sample Analysis. Untargeted volatile analyses were conducted using an MPS2 autosampler (Gerstel, Sursee, Switzerland) and an Agilent 7890B gas chromatography (GC) system coupled with an Agilent 5977A mass selective detector (MSD) (Agilent Technology, Santa Clara, CA, USA). Samples were fully randomized for measurement using the Excel function RAND.

For plasma samples, containing water, a piece of cleaned swab Topper 8 10 × 10 cm (Systagenix, North Yorkshire, United Kingdom) was added to the vial to avoid boiling and foaming. The headspace was extracted for 60 min at 45 °C under reduced pressure (10 mbar) using a vacuum pump Buchi V-300 (Büchi, Flawil, Switzerland) and in-tube extraction materials equipped with a trap filled with Cabosieve S III/Tenax TA (ITEX2, Brechbühler, Switzerland) as described by Fuchsmann et al.,[30] without agitation. The sorbent and syringe were dried under a nitrogen stream for 17 min at 220 mL min^{-1}. For urine samples after SPE extraction, the headspace was extracted for 5 min at 40 °C under the same pressure conditions as for plasma because the solvent in this case is methanol.

The bound volatiles were desorbed from the sorbent for 2 min under a nitrogen flow of 220 mL min^{-1} at the recommended temperature for the employed polymer in a PTV injector with the vent mode at 50 mL min^{-1} and 0 kPa for 120 s. The injector was equipped with a glass liner filled with Tenax TA and cooled with liquid nitrogen at 10 °C. The injector was then heated, at a rate of 12 °C s^{-1}, to 240 °C. The purge flow to split vent was set at 100 mL min^{-1} after 2 min. The reconditioning of the trap was achieved at 300 °C under a nitrogen flow of 100 mL min^{-1} for 15 min.

Volatile compounds were separated on a TRB-FFAP fused silica capillary column (100% polyethylene glycol PEG with nitroterephthalic acid, bonded and cross-linked, 60 m × 0.32 mm × 1.0 μm film; Teknokroma, Barcelona, Spain) with

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

helium as the carrier gas at a constant flow of 2.1 mL min^{-1} (37 cm s^{-1}).

The oven temperature was programmed as follows: 4 min at 40 °C, then heated to 210 °C at a rate of 3 °C min^{-1} with a final hold time of 45 min. The MS settings were as follows: transfer line at 230 °C, source temperature at 230 °C. The analytes were monitored in SCAN mode between 29 and 250 amu with a gain at 15 without solvent delay.

The autosampler was controlled with the Cycle Composer V. 1.5.4 (CTC Analytics, Zwingen, Switzerland) and the CIS 4 injector with Maestro1 software V.1.4.8.14/3.5 (Gerstel).

Data Preprocessing and Statistical Analysis

Reduction and Filtering of Plasma and Urine Data Sets. Deconvolution and grouping of MS signals were achieved using Masshunter Profinder software version 10.0 in recursive mode (Agilent Technologies, Santa Clara, CA, USA). The deconvolution was based on six main parameters (peaks height: >300 counts, RT tolerance: 0.3 min, Min dot product value: 0.4, retention times: ±0.3 min, integration mode: Agile 2, Smoothing: gaussian). Metabolites with a signal less than 3 times the median height of the background noise were excluded during the deconvolution.

The incremental area under the curve (iAUC) was calculated for each metabolite in the postprandial phase with R (v.3.5.1; R Foundation for Statistical Computing, Vienna, Austria) and the MESS package (v.0.5.6). All samples were successfully collected according to the study protocol. The presence of missing values after automatic deconvolution was due to an analyte signal below the detection limit of the instrument in the analysis in question. As a result, missing values were replaced by zero in the data set as described by Xia et al.[36] and used for the calculation of iAUC. A principal component analysis (PCA) combining iAUC for all samples from each of the three groups was performed. Samples outside of the Hotelling T2 ellipse with a 95% confidence interval were considered outliers. A PCA of the iAUC was favored over a PCA of the individual time points to identify potential outliers as the candidate BFIs are statistically identified based on their iAUCs. The PCA analysis resulted in the elimination of two postprandial responses from further analysis, namely, the response of subject ID: 15 to the soy-based drink and the response of subject ID: 21 to milk.

The choice to remove two complete subsets of postprandial responses, instead of individual samples, was motivated by the fact that the two outlying iAUCs were due to abnormally elevated intensities of metabolites in the fasting samples ($t = 0$) used to calculate the corresponding iAUCs. Noteworthy, aside from the elevated fasting values, the postprandial kinetics of the candidate BFIs followed the regular shape expected from postprandial metabolites. Also, in contrast to the two outlying fasting values, the corresponding postprandial samples were not identified as outliers when analyzed by PCA. Taken together, these results indicate that the two fasting samples were technical, rather than biological, outliers. In our study, the first postprandial samples were collected after 1 h, i.e., well into the postprandial phase when a change in concentration is already expected. Given the importance of the fasting time point for the calculation of the iAUC and the methodological limitations in estimating missing values, we chose to exclude the two corresponding postprandial responses from the statistical analysis. To identify potential outliers in the urine samples, the relative signal of each metabolite in the 6 and 24 h pools was directly assessed by PCA analysis. Based on the PCA results that were within the 95% confidence interval, no subjects and samples needed to be eliminated.

Statistical Analysis. Using the plasma features that were retained after data preprocessing, the data sets were further filtered to retain only features demonstrating a postprandial response after the consumption of at least one of the test foods. Features were selected using the nonparametric analysis of longitudinal data package nparLD (v. 2.1) with a cutoff of $p < 0.05$. The features obtained in urine after data preprocessing were not filtered further. We chose to conduct a robust statistical analysis of the data. Metabolites in plasma or urine were therefore only considered to discriminate for the intake of the test foods if they were significantly differentiated by both univariate and multivariate testing.

Univariate analysis first applied was a Kruskal−Wallis test to compare the effect of the test meals on iAUCs (plasma samples) as well as 6 and 24 h urine pools using the signal area. The iAUCs and signal area of the features demonstrating significant differences in this test after false discovery rate (FDR) correction (FDR < 0.05) were then assessed by a Conover−Iman test to compare the effect of the test meals in a pairwise approach. Significant features ($p < 0.05$) were manually integrated using Masshunter quantitative analysis, and the refined iAUCs and signal areas were used to assess significant differences using the Conover−Iman test ($p < 0.05$).

A multivariate analysis using SIMCA-P software (V 14.0, Umetrics, Umea, Sweden) was also applied using the iAUC for plasma samples and the signal areas for the urine samples (6 and 24 h pools) to identify the metabolites that discriminated the responses to cheese, milk, or soy-based drink intake. The data were centered with unit variance (UV), and orthogonal PLS discriminant analysis (OPLS-DA) was conducted to model one by two comparisons on the test products (milk vs cheese/soy-based drink, cheese vs milk/soy-based drink, and soy-based drink vs cheese/milk). The quality of the models was evaluated on four parameters: (i) goodness-of-fit (R2Y); (ii) the predictive ability parameter Q2, calculated by 15-fold cross-validation and using Q2 > 0.3 as cutoff criteria; (iii) ANOVA of the cross-validated residuals (CV-ANOVA) ($p < 0.05$); and (iv) permutation tests with 999 random permutations. Discriminating metabolites were selected using the Variable Importance in Projection for the predictive component (VIP_{pred}) with a cutoff set to 2.0.

Identification of Biomarkers of Food Intake (BFIs)

The National Institute of Standards and Technology NIST/EPA/NIH mass spectral library (NIST17) version 2.3 (NIST, Gaithersburg, MD, USA) was used for peak identification.

The identification of features applied the standard criteria for identification levels (1−4) as recommended by the metabolomics standards initiative (MSI):[37,38] Level 1 consists of an identification using the injection of the pure standard and the comparison of the spectrum obtained with a database (minimal match factor of 90%) and the retention index (RI) (maximal relative difference of about ±10−15). Level 2 corresponds to spectra with a match factor >80% and a maximal relative difference in an RI of ±15 RI. Level 3 implies that the compound has a putative attribute of a compound class and has physicochemical properties and spectral similitude consistent with the compound from a reference library. Level 4 corresponds to unknown compounds.

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

The extraction of urinary molecules in the presence of methanol under acidic conditions is known to result in the transformation of carboxylic acids into methyl esters. The original carboxylic acids are therefore presented in the Results and Discussion section.

RESULTS AND DISCUSSION

Univariate and Multivariate Statistical Results

A total of 2074 features were detected in plasma samples after automatic deconvolution and grouping of the signals. The nparLD test filtered the data set to 697 postprandial features. Four potential BFIs corresponding to 0.19% of features detected in plasma samples were confirmed by the univariate (Kruskal–Wallis test; $p < 0.05$) and multivariate analyses (OPLS-DA, $VIP_{pred} > 2.0$) (Table 1).

A total of 1194 and 1265 features were detected in the urine samples from the 6 and 24 h pools, respectively, after automatic deconvolution and grouping of the signals. The univariate (Kruskal–Wallis test; $p < 0.05$) and multivariate analyses (OPLS-DA, $VIP_{pred} > 2.0$) identified 31 potential BFIs for the 6 h urine pool including 15 that were also identified for the 24 h pool (Table 2). This corresponds to 2.8% of the features detected in urine pools.

Candidate BFIs for Milk, Cheese, and Soy-Based Drink Intake. The following sections present the 35 potential BFIs detected in plasma and urine samples. The sampling of the plasma samples allows kinetics to be observed over six time points (0, 1, 2, 4, 6, and 24 h after food intake). Two pools (6 and 24 h) were prepared and measured for urine samples, and no kinetics could therefore be followed for urine samples. The 31 potential BFIs detected in the 6 h urine pool after milk, cheese, and soy-based drink intake are represented by a heatmap (R package pheatmap v. 1.0.12) separating the soy-based drink from dairy products (Figure 3).

Candidate BFIs for Milk. No molecule discriminant for milk intake was found in the plasma samples. However, one metabolite, 3-ethylphenol (UM33), was discriminant for milk intake in 6 h urine pools (Figure 3; Figure S4). This metabolite is generally present in cow's milk and cheese made from cow's milk[39] and can therefore be considered an exogenous marker appearing in urine after food intake. Although 3-ethylphenol was identified in the feces of adults of both sexes,[40] its presence in human urine has, until now, not been reported.

Candidate BFIs for Cheese. Heptan-2-one (PM1) and undecan-2-one (PM2) were discriminant for cheese intake in plasma (Figure 4), attaining a maximum level after approximately 6 h. The identification of the metabolites was confirmed at levels 1 and 2, respectively. Heptan-2-one was also present after milk intake with measurable kinetics between 1 and 4 h after food intake. The signal in the plasma of this compound returns to baseline values 24 h after cheese intake and 6 h after milk intake. No significant iAUC was observed for this compound after soy-based drink intake compared to baseline. Heptan-2-one is a well-known volatile compound produced during the fermentation of milk,[41] and it was detected in the Gruyère cheese. This metabolite has been detected in human blood[42] and urine[43] as an oxidation product of 2-ethylhexanoic acid, a molecule derived from the metabolism of the plasticizers used in medical devices or of plastic films in contact with food.[44] However, 2-ethylhexanoic acid is also produced from lipolysis and from breakdown of amino acids in cheese during the ripening process,[45,46] and this

Table 1. Identification Parameters of Potential Biomarkers of Food Intake Identified in Plasma Samples

metabolite no	compound identification	HMDB ID	CAS no	discriminant BFIs	presence in test food	retention time (min)	RI (sample)	RI (reference)	fragment used[a] (m/z)	spectral similarity (%)	VIP predictive	p value (6 h iAUC) Kruskal–Wallis test	LI
PM1	heptan-2-one	HMDB0003671	110-43-0	cheese	cheese	29.39	1211	1215[b]	114	91	3.64	8.8×10^{-5}	1
PM2	undecan-2-one	HMDB0033713	112-12-9	cheese	dairy	47.51	1632	1631[b]	58	84	2.23	2.4×10^{-3}	2
PM3	3,5-dimethyloctan-2-one	NA	19781-14-7	dairy	ND	36.93	1381	1375[b]	72	87	4.29	8.8×10^{-5}	2
PM4	NI	NA	NA	soy-based drink	ND	44.45	1560	UND	81		3.38	2.6×10^{-3}	4

[a]Specific fragment used for calculation of the relative peak area of a specific compound. [b]Standard injection. Abbreviations: PM, plasma marker; NI, not identified; ND, not detected; UND, undetermined; NA, not available; iAUC, incremental area under the curve; BFI, biomarker of food intake; LI, level of identification.

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 2. Identification Parameters of Candidate BFIs Identified in Urine Samples[a]

metabolite no	compound identification	HMDB ID	CAS no	discriminant BFIs	presence in test food	retention time (min)	RI (sample)	RI (reference)	fragment used[a] (m/z)	spectral similarity (%)	VIP predictive	p value (6 h pool)	p value (24 h pool)	LI
UM1	1,3-octadiene	HMDB0061900	13643-08-8	soy-based drink	ND	5.26	976	956[c]	81	85	3.80	8.3×10^{-5}	1.2×10^{-3}	2
UM3	2,4-octadiene	HMDB0040966	1002-33-4	soy-based drink	soy-based drink	6.36	1004	1014[c]	81	82	2.99	2.7×10^{-4}	1.7×10^{-1}	2
UM4	1-octen-3-ol, methyl ether[tent]	NA	NA	soy-based drink	soy-based drink	9.97	1094	UND	71	89	3.38	2.4×10^{-5}	1.7×10^{-1}	3
UM5	NI	NA	NA	soy-based drink	ND	11.14	1120	UND	85		3.36	6.0×10^{-5}	3.8×10^{-1}	4
UM6	monoterpene	NA	NA	soy-based drink	soy-based drink	11.40	1126	UND	136		4.24	8.3×10^{-5}	1.2×10^{-3}	3
UM7	heptan-4-one	HMDB0004814	123-19-3	cheese	ND	12.31	1145	1152[b]	114	94	2.88	1.7×10^{-2}	2.7×10^{-1}	1
UM8	methoxycyclooctane[tent]	NA	13212-32-6	soy-based drink	soy-based drink	17.15	1248	UND	71	91	3.42	1.1×10^{-4}	7.4×10^{-2}	3
UM9	1-methoxy-2-propyl acetate	NA	108-65-6	dairy	ND	17.31	1251	1254[b]	72	88	2.20	1.2×10^{-3}	2.7×10^{-1}	2
UM10	naphthalene derivative	NA	NA	soy-based drink	soy-based drink	21.86	1349	UND	159		3.24	1.3×10^{-3}	7.8×10^{-3}	4
UM11	medium-chain fatty acid ester	NA	NA	dairy	ND	22.22	1357	UND	172		2.63	2.2×10^{-3}	2.8×10^{-1}	3
UM12	medium-chain fatty acid ester	NA	NA	cheese	ND	23.72	1390	UND	110		3.06	1.1×10^{-3}	4.4×10^{-1}	3
UM13	naphthalene derivative	NA	NA	soy-based drink	soy-based drink	25.98	1440	UND	117		3.04	3.1×10^{-3}	2.8×10^{-1}	3
UM14	medium-chain fatty acid ester	NA	NA	cheese	ND	26.63	1455	UND	87		3.50	1.7×10^{-3}	2.8×10^{-1}	3
UM15	NI	NA	NA	soy-based drink	soy-based drink	27.04	1465	UND	143		3.59	1.1×10^{-4}	7.4×10^{-2}	4
UM17	1-octen-3-ol	HMDB0031299	3391-86-4	soy-based drink	soy-based drink & dairy	27.50	1475	1474[b]	57	90	3.19	2.2×10^{-3}	1.7×10^{-1}	1
UM18	naphthalene derivative	NA	NA	soy-based drink	soy-based drink	28.49	1498	UND	159		3.81	8.3×10^{-5}	3.7×10^{-4}	3
UM19	9-decenoic acid, methyl ester	HMDB0031003 (9-decenoic acid)	14436-32-9 (9-decenoic acid)	dairy	dairy	35.73	1673	1685[b]	74	90	2.97	8.3×10^{-5}	2.4×10^{-1}	1
UM20	acetophenone	HMDB33910	98-86-2	soy-based drink	soy-based drink & dairy	35.96	1678	1680[b]	77	91	2.64	2.0×10^{-3}	1.7×10^{-1}	1
UM21	naphthalene derivative	NA	NA	soy-based drink	soy-based drink	36.56	1694	UND	159		3.65	8.6×10^{-6}	3.7×10^{-4}	3
UM22	NI	NA	NA	dairy	ND	36.83	1701	UND	108		2.72	5.3×10^{-4}	2.3×10^{-1}	4
UM23	NI	NA	NA	soy-based drink	soy-based drink	38.99	1759	UND	112		4.10	6.0×10^{-5}	3.7×10^{-4}	4
UM24	naphthalene derivative	NA	NA	soy-based drink	soy-based drink	39.15	1764	UND	173		4.01	8.8×10^{-5}	1.3×10^{-1}	3
UM25	1,1,6-trimethyl-1,2-dihydronaphthalene[tent]	HMDB0040284	30364-38-6	soy-based drink	soy-based drink	39.55	1774	1751[c]	157	92	3.20	1.3×10^{-3}	4.0×10^{-2}	3
UM26	medium-chain fatty acid ester	NA	NA	dairy	ND	39.86	1783	UND	69		2.59	7.1×10^{-3}	1.7×10^{-1}	3

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

Table 2. continued

metabolite no	compound identification	HMDB ID	CAS no	discriminant BFIs	presence in test food	retention time (min)	RI (sample)	RI (reference)	fragment used[a] (*m*/*z*)	spectral similarity (%)	VIP predictive	*p* value (6 h pool)	*p* value (24 h pool)	LI
UM27	coumarin derivative	NA	NA	soy-based drink	soy-based drink	43.12	1875	UND	175		4.04	8.3×10^{-5}	3.0×10^{-2}	3
UM28	naphthalene derivative	NA	NA	soy-based drink	ND	44.43	1914	UND	157		3.79	8.3×10^{-5}	1.6×10^{-2}	3
UM29	methyl tetradecanoate	HMDB00806 (tetradecanoic acid)	544-63-8 (tetradecanoic acid)	soy-based drink	ND	48.77	2036	2035[b]	157	94	3.11	6.1×10^{-4}	2.1×10^{-3}	1
UM30	NI	NA	NA	soy-based drink	soy-based drink	49.07	2044	UND	175		3.38	7.5×10^{-3}	1.2×10^{-3}	4
UM31	naphthalene derivative	NA	NA	soy-based drink	ND	50.41	2070	UND	157		2.52	1.7×10^{-2}	5.0×10^{-2}	3
UM32	naphthalene derivative	NA	NA	soy-based drink	soy-based drink	52.87	2150	UND	173		3.01	2.2×10^{-3}	5.2×10^{-3}	3
UM33	3-ethylphenol	HMDB0059873	620-17-7	milk	dairy	55.07	2226	2220[b]	107	90	2.44	5.0×10^{-5}	1.7×10^{-1}	1

[a]Specific fragment used for calculation of the relative peak area of a specific compound. [b]Standard injection. [c]Retention index from the literature on FFAP column (NIST17); [tent]: tentatively identified. Abbreviations: UM, urine marker; UND, undetermined; NI, not identified; ND, not detected; NA, not available; iAUC, incremental area under the curve; BFI, biomarker of food intake; LI, level of identification. Esterification of some candidates' BFIs results from chemical modification of their carboxylic acid during the extraction procedure.

molecule can be transformed into heptan-2-one and heptan-4-one by beta-oxidation in humans. Heptan-2-one can thus be considered both an exogenous and endogenous metabolite.

The origin of undecan-2-one can be considered exogenous in plasma samples. The molecule is already present in the cheese matrix but was found in trace amounts in milk and soy-based drink. The formation of undecan-2-one in cheese is due to the fermentation of milk during the ripening process due to the degradation of esterified β-ketoalkanoic acids of milk fat.[47,48] Undecan-2-one can be present naturally in the milk but is mainly due to the heat treatment it has undergone as much higher concentrations have been measured after ultra-high-temperature processing of milk whereas it is only present in trace amounts in pasteurized milk.[49] This volatile compound has anti-inflammatory and antibacterial properties.[50] Another study also showed its benefit as a repellent for insects, including mosquitoes.[51]

Heptan-4-one was identified at level 1 and discriminated cheese intake from milk intake, but not from soy-based drink intake, in the 6 h urine pool (Figure S5). The presence of this metabolite in the test product B (Gruyère cheese) was not confirmed. Notably, the presence of this compound in urine samples may have the same metabolic origin as heptan-2-one.

Two additional features, UM12 and UM14, were discriminant for cheese intake in the 6 h, but not the 24 h, urine pools (Figure S5). These metabolites belong to the family of medium-chain fatty acids, as indicated by their ion specificity and spectral similarity. Medium-chain fatty acids originate mainly from the metabolism of triglycerides by intestinal bacteria and the liver. These molecules are essential for cellular energy metabolism and can control cell death and survival.[47] The medium-chain fatty acids present in the urine of the participants after consumption of cheese are not present in any of the three test products. We therefore hypothesize here that they are formed after consumption of the test cheese.

Candidate BFIs for Dairy. Several molecules discriminating dairy products from soy-based drink intake were detected: one in plasma (3,5-dimethyloctan-2-one PM3) (Figure 5) and five in the urine (1-methoxy-2-propyl acetate UM9, 9-decenoic acid UM19, two medium-chain fatty acids UM11 and UM26, and an unknown compound UM22) (Figure 3; Figure S6).

3,5-Dimethyloctan-2-one was detected with measurable kinetics after milk or cheese intake. The identification was determined at level 2. Kinetics return to approximately baseline values after 24 h, reaching a maximum level between 4 and 6 h after dairy intake. No significant iAUC was measured after soy-based drink intake. This compound was not detected in any food samples measured. 3,5-Dimethyloctan-2-one has been reported in the urine of wolves.[52] Apart from this report, the presence of 3,5-dimethyloctan-2-one in samples isolated from living organisms has not been reported. The mechanism leading to the excretion of 3,5-dimethyloctan-2-one in urine after dairy intake remains unknown, although one could speculate that this molecule is derived from the metabolism of dairy fats by the human organism.

9-Decenoic acid is a medium-chain fatty acid present in urine after the ingestion of both dairy products. This compound was discriminant in the 6 h urine pool, and its identity was confirmed at level 1. It is a constituent of milk fat and has also been measured in parmesan cheese, butter, wine, and beer.[53,54] As for the medium-chain fatty acids identified after cheese intake, 9-decenoic acid was found in the test

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

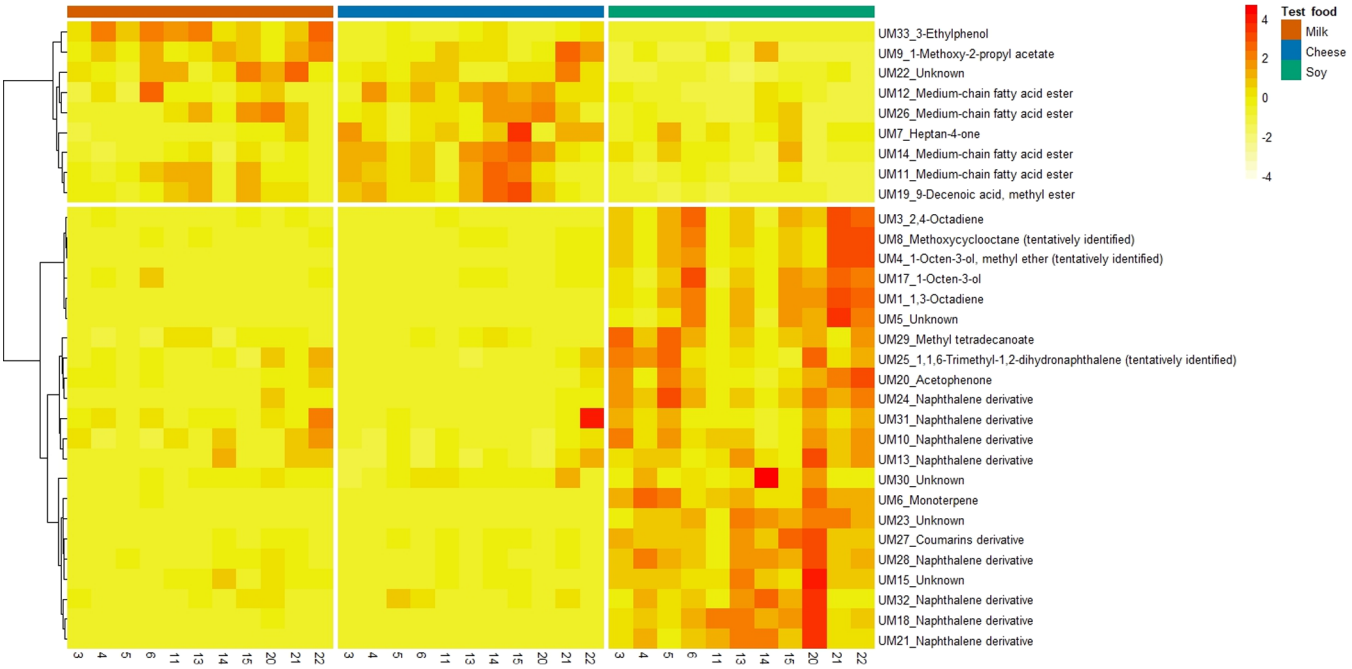


Figure 3. Postprandial distributions of 31 potential BFIs for milk (orange), cheese (blue), or soy-based drink (green) intake in 6 h urine pool. The 11 subjects and the test food consumed are in columns. Each row represents one of the 31 potential BFIs. BFIs were grouped by hierarchical clustering analysis using Euclidean distance measure and Ward's linkage. Color scale corresponds to intensities that have been centered-scaled by row.

products and is likely to be derived from the liver metabolism of triglycerides present in the dairy products.

1-Methoxy-2-propyl acetate was determined at level 2. This compound was not present in the tested foods but was discriminant for dairy intake in the 6 h urine pool. 1-Methoxy-2-propyl acetate has, until now, not been detected in biological fluids. Interestingly, this metabolite could potentially be formed from an esterification reaction between 1-methoxy-2-propanol and acetic acid, two molecules naturally present in cheeses.

Candidate BFIs for Soy-Based Drink. Only one unknown (PM4) metabolite was found in plasma samples after soy-based drink intake (Figure 6). This compound was not present in plasma after milk or cheese intake but showed a postprandial response to soy-based drink intake, attaining maximum levels between 4 and 6 h after soy-based drink intake.

Conversely, 22 BFIs that were discriminant for soy-based drink intake were detected in urine samples (Figure 3; Figure S-7). Two alkenes (UM1, UM3) were determined as 1,3-octadiene and 2,4-octadiene at level 2. Both metabolites showed a clear postprandial response in urine after soy-based drink intake but not after dairy products. 1,3-Octadiene was detected only in the 6 h urine pool and was not detected in the test foods. 1,3-Octadiene, a metabolite produced by the oxidation of the fatty acids present in plants, is described as a natural constituent of soybean[55] although it has also been identified in oysters.[56] In humans, 1,3-octadiene was previously identified in saliva[10] but, to our knowledge, not in urine. 2,4-Octadiene was only detected in 6 h urine pools as well as in the test product soy-based drink. 2,4-Octadiene was identified in human breath and was used to diagnose wheezy preschool children[57] but has not been previously detected in urine.

One monoterpene (UM6) was detected in 6 and 24 h urine pools. This metabolite was also detected in the test product soy-based drink. As several terpenes, in particular monoterpenes, have previously been identified in plants, including soybean,[55,58,59] terpenes might be interesting markers of the intake of plant foods.

UM8 was tentatively identified as methoxycyclooctane with a spectral similarity of 91% but classified at level 3. It could not be confirmed by the injection of the pure standard, as it was unavailable, or by comparison with the literature. UM8 was present in urine after soy-based drink intake but not after dairy intake. This compound has also been identified previously in onion seed oil.[60] Its presence in human samples has, until now, not been reported.

1-Octen-3-ol (UM17) was identified (level 1) in the 6 h urine pool after soy-based drink intake. This metabolite was present in all three products tested, but the concentration was much higher in soy-based drink, which explains the increased signal after consumption of the soy-based drink. 1-Octen-3-ol has already been identified in soybeans and is produced from the enzymatic degradation of fatty acids in plants.[61] 1-Octen-3-ol is naturally found at high concentrations in soymilk and is responsible in part for a strong flavor that is not generally appreciated in these beverages.[62] This compound has also been detected in human urine and feces after consumption of various animal and plant foods.[10] Although an exogenous nature of this metabolite is likely in our study, the presence of 1-octen-3-ol in human samples after ingestion of food of animal origin does not rule out that this compound can be produced from the metabolism of a precursor compound.

Acetophenone (UM20) was present at significantly higher levels after soy-based drink intake in the 6 h urine pool. Acetophenone was also present in much smaller amounts after consumption of the two other test foods. The molecule was detected in all three test products, therefore suggesting an exogenous origin of its presence in urine. Acetophenone has been previously identified in soybean[63] and heat-treated cow's

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

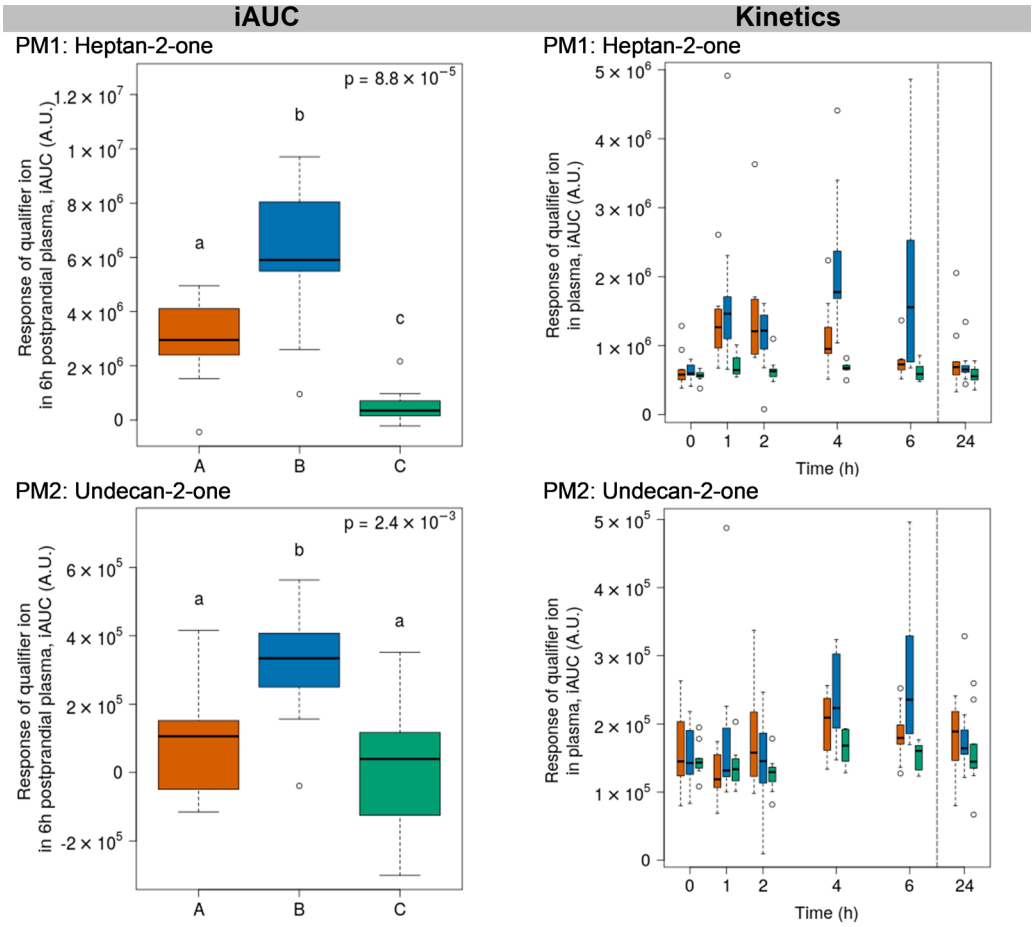


Figure 4. Candidate BFIs for cheese intake in plasma samples. Boxplots of iAUC (left panels) and postprandial kinetics (right panels) of candidate BFIs in plasma samples. Different letters (a, b, c) between the treatment groups in the left graphs denote statistically different iAUCs based on the Conover−Iman test. The *p* value denotes the significance of the Kruskal−Wallis test comparing the postprandial effects of the three test products, namely, milk (A), cheese (B), and soy-based drink (C). PM: plasma marker; iAUC: incremental area under the curve; A.U.: arbitrary unit.

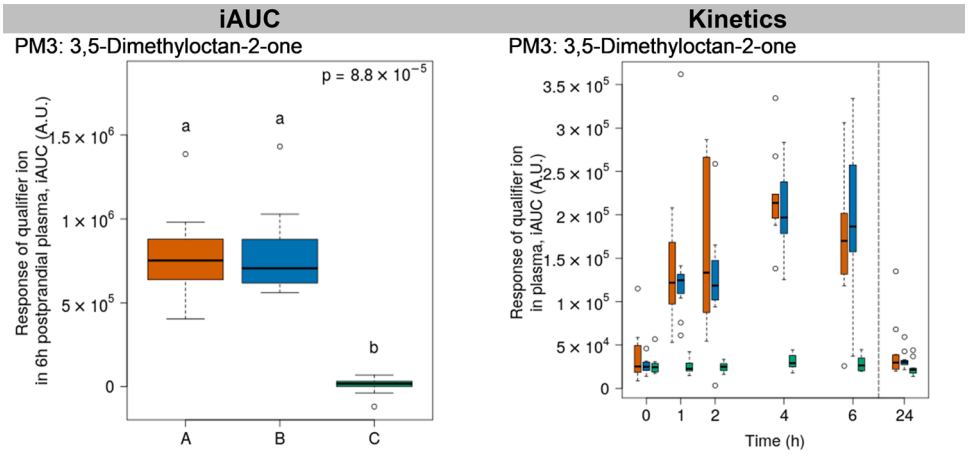


Figure 5. Candidate BFIs for dairy intake in plasma samples. Boxplots of iAUC (left panels) and postprandial kinetics (right panels) of candidate BFIs in plasma samples. Different letters (a, b, c) between the treatment groups in the left graphs denote statistically different iAUCs based on the Conover−Iman test. The *p* value denotes the significance of the Kruskal−Wallis test comparing the postprandial effects of the three test products, namely, milk (A), cheese (B), and soy-based drink (C). PM: plasma marker; iAUC: incremental area under the curve; A.U.: arbitrary unit.

milk and Gruyère cheese.[64,65] However, its presence in human urine has, until now, not been reported.

Tetradecanoic acid (UM29) was detected in urine after consumption of both dairy foods, but it was significantly higher after the consumption of the soy-based drink. The difference was most marked in the 6 h urine pool, although it was still significant in the 24 h urine pool. Tetradecanoic acid is present in both plants and animals and has been identified in various

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

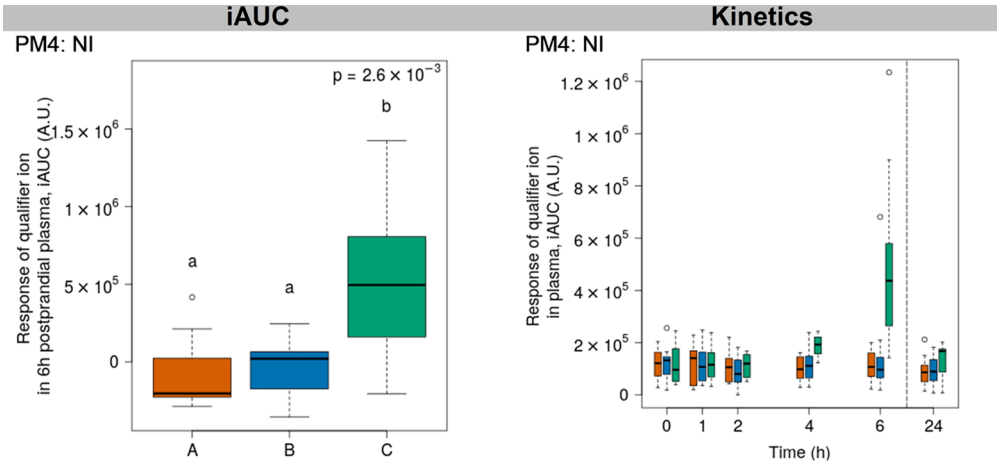


Figure 6. Candidate BFIs for soy-based drink intake in plasma samples. Boxplots of iAUC (left panels) and postprandial kinetics (right panels) of candidate BFIs in plasma samples. Different letters (a, b, c) between the treatment groups in the left graphs denote statistically different iAUCs based on the Conover−Iman test. The *p* value denotes the significance of the Kruskal−Wallis test comparing the postprandial effects of the three test products, namely, milk (A), cheese (B), and soy-based drink (C). PM: plasma marker; iAUC: incremental area under the curve; A.U.: arbitrary unit; NI: not identified.

biological fluids such as blood, urine, feces, and saliva. The source of the carboxylic acid may therefore be endogenous as well as absorbed directly from food.[10]

Nine naphthalene derivatives (UM10, UM13, UM18, UM21, UM24, UM25, UM28, UM31, and UM32) were detected in urine samples, each significantly discriminating for the consumption of the soy-based drink. One of these metabolites (UM25) was tentatively identified with 92% spectral similarity as 1,1,6-trimethyl-1,2-dihydronaphthalene. Seven of these metabolites were detected in the soy-based drink, which makes them interesting candidate BFIs. Naphthalene derivatives, such as 1,1,6-trimethyl-1,2-dihydronaphthalene, can be derived from the degradation of carotenoids (*β*-carotene and lutein) in the presence of oxygen, light, and elevated temperatures.[66,67] Lutein is a xanthophyll lipophilic antioxidant and is found in high concentration in soybeans.[68] We have also identified another polycyclic aromatic compound derived from coumarin (UM27), which significantly discriminates for the consumption of the soy-based drink. In this context, it is noteworthy that coumestrol is a phytoestrogen present in soybean and demonstrates a large spectrum of biological activities, including a cytotoxic activity against cancer cells.[69]

Assessment of the Postprandial Volatilome as a Source of BFIs

A relatively large number of molecules present in plasma and urine samples could discriminate the consumption of the three test foods when compared among each other. The majority of these metabolites were already present in the test foods, absorbed by the digestive tract and/or excreted in the urine without being metabolized. As the spectrum of molecular masses detected by volatilomics is shifted toward lower values compared to LC−MS, or even the semivolatile compounds analyzed by GC−MS after derivatization, the relatively high percentage of unmetabolized molecules detected by volatilomics might reflect the fact that small molecules tend to be at the end of metabolic processes. A smaller fraction of metabolites showed evidence of an endogenous character, being formed during the digestive process or degraded in the liver.

Out of the 35 discriminating metabolites characterized in this study, the vast majority of them (23 metabolites) were found after ingestion of the soy-based drink. This observation might reflect the fact that foods of plant origin possess a catalogue of molecules that significantly differs from animal tissues, in particular human tissues. On the other hand, only one molecule (3-ethylphenol) discriminated for milk, albeit not specifically, this molecule also being present in urine after cheese intake. This observation is in line with the knowledge that cow's milk, being a mammary physiological fluid, has an elevated compositional homology to human fluids, which makes the detection of specific BFIs difficult. However, transforming milk to cheese through bacterial activity increased the number of discriminating metabolites, four metabolites being potential BFIs for cheese. Interestingly, six metabolites were characterized as discriminative for dairy products, e.g., milk and cheese, compared to the soy-based drink. As the candidates' BFIs measured for both products in plasma or urine were subject to microbiological transformation (e.g., by the technological flora for cheese and by the intestinal microbiota for milk and cheese), the metabolic origin of these markers should be carefully evaluated.

Comparing the candidate BFIs identified for milk and cheese intake in the products as well as in the biofluids may actually help deciphering the relative contribution of both microbial processes: molecules derived from processing of the milk matrix by the human microbiota should not be detected in the dairy products but should be detected in humans after the ingestion of both milk and cheese. On the other hand, molecules derived from the fermentation of milk to cheese should only be found in cheese as well as after cheese intake, provided that the intestinal microbiota do not also contribute to the production of these compounds. This scenario is, however, likely to take place for some of the metabolites. However, clarifying the contribution of each of the two microbial transformations to the formation of such metabolites would request the use of a fermented food product having not undergone additional modifications, which is not the case with cheese making as this process removes whey proteins. Of note, we have already provided direct evidence for such dual microbial processing activity by following a range of indoles in

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

milk and yoghurt as well as in blood after the ingestion of these dairy products.[70]

The kinetics analysis of the three potential BFIs for the intake of dairy products showed a rapid increase in plasma concentration levels within the first hour after the test meal with a maximum after 4 to 6 h. In contrast, the single BFI identified in plasma for soy-based drink increased only after 4 h to reach its maximum after 6 h. Given that these molecules are exclusively present in either the animal or plant matrix, it is not clear whether the different postprandial kinetics are due to differences in the release of the precursor molecules from the food matrix or during subsequent steps (i.e., digestion, modification by the microbiota, intestinal absorption, liver metabolism), indicating a slower assimilation of molecules derived from the plant matrix.[71,72] The analysis of a larger number of molecules would be necessary to provide convincing evidence for different postprandial kinetics in blood between the dairy and soy matrices.

Out of the 31 candidate BFIs characterized in the 6 h urine pool (nine for dairy products, 22 for the soy-based drink), statistical significance for 17 of them was lost in the 24 h pool (nine for dairy products, eight for the soy-based drink). This effect is likely due to the rapid excretion of these BFIs or the pooling of urine samples over a longer period (6 to 24 h).

The interindividual variability was evaluated by visual inspection of each of the identified metabolites in the plasma. For all candidate BFIs, an increase compared to 0 h in plasma after ingestion of the relevant test products was confirmed in each of the subjects enrolled in the study. A representative example is 3,5-dimethyloctan-2-one, which showed, for each of the 11 subjects, a consistent increase in the signal after ingestion of milk and cheese between 1 and 6 h. These observations suggest that the postprandial behavior of the compounds characterized in this study might be discriminative in a broader population, which justifies their characterization in further studies.

Dragsted et al.[13] proposed a validation scheme for candidate BFIs, which includes biological as well as analytical criteria (plausibility, dose−response, time-response, robustness, reliability, stability, analytical performance, reproducibility). Given the exploratory nature of our study, being one of the first reporting the use of the volatilome as a source of BFIs, many of these criteria are yet to be fulfilled in order for the identified molecules to be validated as true markers of intake. Among these criteria, the data presented in this report provides convincing evidence that the origin of a majority of the identified postprandial metabolites is associated with their presence, or the presence of precursors thereof, in the respective food matrices. A key issue, however, relates to the specificity of these candidate markers, i.e., their ability to reflect on the intake of the specific foods investigated in this study, in a more complex dietary environment. Insights into this issue can be gained by an inspection of the relative behavior of the candidate BFIs under different conditions. For illustration, the strong postprandial response of 3,5-dimethyloctan-2-one after milk and cheese intake, but not after soy drink intake, does not prove that this molecule is a marker of dairy intake but only that it can differentiate dairy from soy drink intake under the study conditions. Also, as the levels of this molecule are different from zero, although low, under fasting conditions (0, 24 h), it is not clear whether its presence is reflecting an endogenous origin, the regular consumption of dairy products, or the consumption of other foods. The same caveat holds true for the candidate BFIs for the soy drink. For illustration, molecules such as 1,3-octadiene, are elevated in urine after soy drink intake but very low after dairy products intake, even in the 24 h pools, which integrate the postprandial phase after intake of the standardized meals at 6 and 12 h. The standardized meal containing rice and chicken and the absence of 1,3-octadiene in the 24 h pool after the soy drink intake therefore provide additional information on the specificity of this molecule. However, information on the specificity remains limited by the frame of the study design (e.g., the three test foods and the standardized meals) so that relevant information on the specificity of this marker can only be efficiently gained from observational cohorts.

Strengths and Limitations of the Study

Limited information is currently available on potential volatile biomarkers that are present in biological samples after food intake. The analysis of volatile compounds using the DHS-VTT method makes it possible to preserve the integrity of the samples during measurement and thus to provide an overview of these compounds in the biological matrices studied. SPE extraction was conducted on urine samples in order to concentrate them prior to DHS-VTT extraction. The molecular composition of the samples retained by the SPE cartridge depends on the pH of the loaded samples. We have selected an acidic pH in order to maximize the analysis of carboxylic acids in light of the relevance of fatty acids in nutritional studies. Neutral compounds such as alcohols, ketones, aldehydes, lactones, and even esters can be extracted in acidic conditions because they are also in an uncharged form and are extracted in the same way as carboxylic acids at the working pH. We cannot exclude, however, that a subset of molecules, evidently basic compounds such as amines in a charged form, are not adsorbed on the SPE cartridges due to the specific extraction conditions.

Nine of the candidate BFIs measured in urine belong to the naphthalene family. These plant-derived compounds could undergo phase I oxidation by liver cytochromes P450 and conjugation by phase II reactions with glucuronosyltransferase or sulfotransferase.[73,74] Glucuronide and sulfate conjugates are not volatile and, therefore, should not be measurable with our workflow. However, the use of an acidic pH during SPE extraction could promote hydrolysis of these molecules to their nonconjugated forms what would allow their measurement. The limitation, however, is that this method does not allow determination of whether these molecules are secreted in their conjugated or unconjugated form in urine. Noteworthy, the presence of conjugated naphthalene derivatives was not confirmed in LC−MS analyses[33] suggesting that the soy-based drink may contain some nonconjugated naphthalene derivatives, which are sufficiently water-soluble to be eliminated directly in the urine. Of note, LC−MS would be an appropriate complementary technology to measure conjugated forms of volatile compounds. For illustration, 3-ethylphenol, which is present in dairy foods as well as in urine after their ingestion, could be expected to be conjugated with glucuronide or sulfate by the human organism. These conjugated molecules were, however, not found in the urine of the subjects.[33]

In addition, compared to conventional headspace extraction techniques, such as SPME,[29] the DHS-VTT technique significantly increases the number of volatile compounds that can be extracted from the matrix, thus facilitating the discovery

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

of new potential biomarkers. Taken together, the extraction strategy was successful as indicated by the large number of volatile compounds measured in plasma (>2000) and urine (>1000).

Candidate blood and urinary BFIs for milk, cheese, or soy-based drink intake have already been identified using samples from the study reported here by GC−MS (after derivatization) and nuclear magnetic resonance (NMR),[31,32] as well as LC−MS.[33] However, none of the discriminating molecules presented in the current report was previously detected by the conventional metabolomics platforms. Our study therefore shows that the application of GC−MS volatilomics to nutritional research, in this particular case biomarker discovery, is an innovative method that complements classical approaches by characterizing changes in new classes of biomarkers. Of note, our volatilomics method should be sufficiently sensitive to allow for the future analysis of smaller volumes of urine and, consequently, more detailed analyses of the postprandial kinetics of the candidate BFIs. The volatile nature of the molecules detected by this technology offers an interesting link to existing food science data as many of these molecules have been already characterized by food scientists for their organoleptic properties, as shown, for example, for 3-ethylphenol in milk,[39] heptan-2-one in fermented milk,[41] and 1-octen-3-ol in soy products.[62]

This somewhat addresses a challenge faced in other MS-based metabolomics technologies where the identification of the molecules at level 1 remains a bottleneck: Indeed, even in the current work, out of the 35 discriminating molecules measured in this study, we could only identify seven potential BFIs, namely, one for milk (3-ethylphenol in urine), two for cheese (heptan-2-one in plasma and heptan-4-one in urine), one for dairy (9-decenoic acid in urine), and three for soy-based drink (1-octen-3-ol, acetophenone, and tetradecanoic acid in urine). Another important limitation of the study is related to the automatic deconvolution and alignment of the mass spectrometer signals with the commercially available software. Many parameters need to be taken into account to optimize peak recognition and integration, which are sensitive to variations in the performance of the instrument.

Intriguingly, most of the candidate BFIs were detected in urine and in the test foods, suggesting that, related to their small size, these molecules pass through the gastrointestinal and metabolic organs without being metabolized and are finally excreted in the urine. Furthermore, we could not detect these molecules in the postprandial plasma samples. Although a biological explanation for this observation should not be excluded, differences in the extraction efficiency of these molecules in the two fluids with different physicochemical properties could also explain this observation.

The contribution of nutrimetabolomics to a more objective quantification of food intake is steadily increasing as barriers representing the technical limitations associated with the detection, integration, and identification of metabolites are removed by improvement of methodologies. This report shows that volatilomics should complement the current battery of metabolomics platforms used to reach this goal.

CONCLUSIONS

Metabolomics studies in nutritional research almost exclusively focus on the analysis of the nonvolatile compounds present in biological tissues and fluids. This study has measured the volatilome of plasma and urine samples to identify new potential biomarkers of the consumption of milk, cheese, and soy-based drink. To this end, we used a method that preserved the samples as well as maximized the extraction of the volatile analytes from the measured matrices.

Our report demonstrates the benefits of analyzing the postprandial volatilome of human blood and urine. Of the 35 molecules reported here, several of them have, to our knowledge, not been reported in human tissues and fluids, and all of them are new candidate BFIs. Most of the candidate volatile BFIs identified in this study originate from consumption of the soy-based drink. Evidently, the precursors of the candidate markers of soy intake are specific molecules derived from plant secondary metabolism, which are not produced endogenously by the human organism. In contrast, the composition of cow milk is globally much closer to human biofluids, such as blood, which complicates the identification of candidate BFIs for dairy products. Nonetheless, despite this compositional similarity, candidate volatile markers of dairy intake, mostly derived from lipids, were identified. The candidate BFIs identified here should be validated in a cohort study to confirm their specificity and relevance as BFIs.

Nutrivolatilomics is an essential addition to nutrimetabolomics analysis, providing a broader picture of the food metabolome.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00324.

Figure S1. Multivariate analysis of 6 h urine pools; score plots, loading plots with most discriminating features in red, permutation tests (n = 999), and model parameters of the OPLS-DA model; Figure S2. Multivariate analysis of 24 h urine pools; score plots, loading plots with most discriminating features in red, permutation tests (n = 999), and model parameters of the OPLS-DA model; Figure S3. Multivariate analysis of plasma samples; score plots, loading plots with most discriminating features in red, permutation tests (n = 999), and model parameters of the OPLS-DA model. Figure S4. Boxplots of postprandial signal areas for candidate BFIs for milk in 6 (left panels) and 24 h urine pools. Figure S5. Boxplots of postprandial signal areas for cheese candidate BFIs in 6 (left panels) and 24 h urine pools. Figure S6. Boxplots of postprandial signal areas for dairy candidate BFIs in 6 (left panels) and 24 h urine pools. Figure S7. Boxplots of postprandial signal areas for soy-based drink candidate BFIs in 6 (left panels) and 24 h urine pools. Figure S8. Mass spectrum of selected compounds detected in plasma and urine samples at levels 1 to 3 in comparison with the injected reference molecules or the mass spectrum of the NIST17 library if possible (PDF)

AUTHOR INFORMATION

Corresponding Author

Pascal Fuchsmann − *Agroscope, 3003 Bern, Switzerland;* orcid.org/0000-0003-1013-5657; Phone: +41 58 463 82 60; Email: pascal.fuchsmann@agroscope.admin.ch

Authors

Mireille Tena Stern − *Agroscope, 3003 Bern, Switzerland*

Linda H. Münger − *Agroscope, 3003 Bern, Switzerland*

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

Grégory Pimentel − *Agroscope, 3003 Bern, Switzerland*
Kathryn J. Burton − *Agroscope, 3003 Bern, Switzerland*
Nathalie Vionnet − *Service of Endocrinology, Diabetes and Metabolism, Lausanne University Hospital, 1011 Lausanne, Switzerland*
Guy Vergères − *Agroscope, 3003 Bern, Switzerland;* orcid.org/0000-0003-4574-0590

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00324

Funding

The study was funded by the Swiss National Science Foundation (40HD40_160618) in the frame of the national research program "Healthy nutrition and sustainable food protection (NRP69)" and the HDHL Joint Programming Initiative "The Food Biomarker Alliance (FoodBAll)".

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Barbara Walther for helpful comments on the manuscript.

■ ABBREVIATIONS

BFI, biomarker of food intake; FoodBAll, Food Biomarker Alliance; GC, gas chromatography; GMD, Golm Metabolome Database; HMDB, Human Metabolome Database; IStd, internal standard; LAB, lactic acid bacteria; MSD, mass spectrometer detector; OPLS-DA, Orthogonal Projection to Latent Structures Discriminant Analysis; PCA, principal component analysis; QC, quality control; RI, retention index; RT, retention time; SPE, solid-phase extraction; VIP, variable importance of projection; VTT, Vacuum Transfer In Trap

■ REFERENCES

(1) Johnson, R. K. Dietary intake–how do we measure what people are really eating? *Obes. Res.* **2002**, *10*, 63s−68s.
(2) Subar, A. F.; Freedman, L. S.; Tooze, J. A.; Kirkpatrick, S. I.; Boushey, C.; Neuhouser, M. L.; Thompson, F. E.; Potischman, N.; Guenther, P. M.; Tarasuk, V.; Reedy, J.; Krebs-Smith, S. M. Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J. Nutr.* **2015**, *145*, 2639−2645.
(3) Archundia Herrera, M.; Chan, C. Narrative Review of New Methods for Assessing Food and Energy Intake. *Nutrients* **2018**, *10*, 1−19.
(4) Rusin, M.; Årsand, E.; Hartvigsen, G. Functionalities and input methods for recording food intake: a systematic review. *Int. J. Med. Inform.* **2013**, *82*, 653−664.
(5) Scalbert, A.; Brennan, L.; Manach, C.; Andres-Lacueva, C.; Dragsted, L. O.; Draper, J.; Rappaport, S. M.; van der Hooft, J. J.; Wishart, D. S. The food metabolome: a window over dietary exposure. *Am. J. Clin. Nutr.* **2014**, *99*, 1286−1308.
(6) Walther, B.; Lett, A. M.; Bordoni, A.; Tomás-Cobos, L.; Nieto, J. A.; Dupont, D.; Danesi, F.; Shahar, D. R.; Echaniz, A.; Re, R.; Fernandez, A. S.; Deglaire, A.; Gille, D.; Schmid, A.; Vergères, G. GutSelf: Interindividual Variability in the Processing of Dietary Compounds by the Human Gastrointestinal Tract. *Mol. Nutr. Food Res.* **2019**, *63*, 1−28.
(7) Rémond, D.; Shahar, D. R.; Gille, D.; Pinto, P.; Kachal, J.; Peyron, M.-A.; Dos Santos, C. N.; Walther, B.; Bordoni, A.; Dupont, D.; Tomás-Cobos, L.; Vergères, G. Understanding the gastrointestinal tract of the elderly to develop dietary solutions that prevent malnutrition. *Oncotarget* **2015**, *6*, 13858−13898.
(8) Cho, K.; Mahieu, N. G.; Johnson, S. L.; Patti, G. J. After the feature presentation: technologies bridging untargeted metabolomics and biology. *Curr. Opin. Biotechnol.* **2014**, *28*, 143−148.
(9) Putri, S. P.; Yamamoto, S.; Tsugawa, H.; Fukusaki, E. Current metabolomics: technological advances. *J. Biosci. Bioeng.* **2013**, *116*, 9−16.
(10) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **2018**, *46*, D608−d617.
(11) Crunkhorn, S. Applying the human metabolome. *Nat. Rev. Drug Discovery* **2019**, *18*, 668.
(12) Brouwer-Brolsma, E. M.; Brennan, L.; Drevon, C. A.; van Kranen, H.; Manach, C.; Dragsted, L. O.; Roche, H. M.; Andres-Lacueva, C.; Bakker, S. J. L.; Bouwman, J.; Capozzi, F.; De Saeger, S.; Gundersen, T. E.; Kolehmainen, M.; Kulling, S. E.; Landberg, R.; Linseisen, J.; Mattivi, F.; Mensink, R. P.; Scaccini, C.; Skurk, T.; Tetens, I.; Vergeres, G.; Wishart, D. S.; Scalbert, A.; Feskens, E. J. M. Combining traditional dietary assessment methods with novel metabolomics techniques: present efforts by the Food Biomarker Alliance. *Proc Nutr Soc* **2017**, *76*, 619−627.
(13) Dragsted, L. O.; Gao, Q.; Scalbert, A.; Vergères, G.; Kolehmainen, M.; Manach, C.; Brennan, L.; Afman, L. A.; Wishart, D. S.; Andres Lacueva, C.; Garcia-Aloy, M.; Verhagen, H.; Feskens, E. J. M.; Praticò, G. Validation of biomarkers of food intake-critical assessment of candidate biomarkers. *Genes Nutr.* **2018**, *13*, 1−14.
(14) Ulaszewska, M. M.; Weinert, C. H.; Trimigno, A.; Portmann, R.; Andres Lacueva, C.; Badertscher, R.; Brennan, L.; Brunius, C.; Bub, A.; Capozzi, F.; Cialie Rosso, M.; Cordero, C. E.; Daniel, H.; Durand, S.; Egert, B.; Ferrario, P. G.; Feskens, E. J. M.; Franceschi, P.; Garcia-Aloy, M.; Giacomoni, F.; Giesbertz, P.; Gonzalez-Dominguez, R.; Hanhineva, K.; Hemeryck, L. Y.; Kopka, J.; Kulling, S. E.; Llorach, R.; Manach, C.; Mattivi, F.; Migne, C.; Munger, L. H.; Ott, B.; Picone, G.; Pimentel, G.; Pujos-Guillot, E.; Riccadonna, S.; Rist, M. J.; Rombouts, C.; Rubert, J.; Skurk, T.; Sri Harsha, P. S. C.; Van Meulebroek, L.; Vanhaecke, L.; Vazquez-Fresno, R.; Wishart, D.; Vergeres, G. Nutrimetabolomics: An Integrative Action for Metabolomic Analyses in Human Nutritional Studies. *Mol. Nutr. Food Res.* **2019**, *63*, 1−38.
(15) Gika, H.; Virgiliou, C.; Theodoridis, G.; Plumb, R. S.; Wilson, I. D. Untargeted LC/MS-based metabolic phenotyping (metabonomics/metabolomics): The state of the art. *J. Chromatogr., B* **2019**, *1117*, 136−147.
(16) Beale, D. J.; Pinu, F. R.; Kouremenos, K. A.; Poojary, M. M.; Narayana, V. K.; Boughton, B. A.; Kanojia, K.; Dayalan, S.; Jones, O. A. H.; Dias, D. A. Review of recent developments in GC-MS approaches to metabolomics-based research. *Metabolomics* **2018**, *14*, 1−31.
(17) Emwas, A.-H.; Roy, R.; McKay, R. T.; Tenori, L.; Saccenti, E.; Gowda, G. A. N.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M.; Wishart, D. S. NMR Spectroscopy for Metabolomics Research. *Metabolites* **2019**, *9*, 1−39.
(18) Brennan, L. NMR-based metabolomics: from sample preparation to applications in nutrition research. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 42−49.
(19) Amann, A.; de Lacy Costello, B.; Miekisch, W.; Schubert, J.; Buszewski, B.; Pleil, J.; Ratcliffe, N.; Risby, T. The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J. Breath Res.* **2014**, *8*, 1−17.
(20) Pleil, J. D.; Williams, A. Centralized resource for chemicals from the human volatilome in an interactive open-sourced database. *J. Breath Res.* **2019**, *13*, 1−3.
(21) Taware, R.; Taunk, K.; Pereira, J. A. M.; Shirolkar, A.; Soneji, D.; Câmara, J. S.; Nagarajaram, H. A.; Rapole, S. Volatilomic insight of head and neck cancer via the effects observed on saliva metabolites. *Sci. Rep.* **2018**, *8*, 17725.

(22) Cavaleiro Rufo, J.; Paciência, I.; Mendes, F. C.; Farraia, M.; Rodolfo, A.; Silva, D.; de Oliveira Fernandes, E.; Delgado, L.; Moreira, A. Exhaled breath condensate volatilome allows sensitive diagnosis of persistent asthma. *Allergy* **2019**, *74*, 527−534.
(23) Chen, L.; Capone, D. L.; Jeffery, D. W. Analysis of Potent Odour-Active Volatile Thiols in Foods and Beverages with a Focus on Wine. *Molecules* **2019**, *24*, 1−24.
(24) Diez-Simon, C.; Mumm, R.; Hall, R. D. Mass spectrometry-based metabolomics of volatiles as a new tool for understanding aroma and flavour chemistry in processed food products. *Metabolomics* **2019**, *15*, 1−20.
(25) Sgorbini, B.; Cagliero, C.; Liberto, E.; Rubiolo, P.; Bicchi, C.; Cordero, C. Strategies for Accurate Quantitation of Volatiles from Foods and Plant-Origin Materials: A Challenging Task. *J. Agric. Food Chem.* **2019**, *67*, 1619−1630.
(26) White, R. H. Occurrence of S-methyl thioesters in urines of humans after they have eaten asparagus. *Science* **1975**, *189*, 810−811.
(27) Roscher, R.; Koch, H.; Herderich, M.; Schreier, P.; Schwab, W. Identification of 2,5-dimethyl-4-hydroxy-3[2H]-furanone beta-D-glucuronide as the major metabolite of a strawberry flavour constituent in humans. *Food Chem. Toxicol.* **1997**, *35*, 777−782.
(28) Zeller, A.; Horst, K.; Rychlik, M. Study of the metabolism of estragole in humans consuming fennel tea. *Chem. Res. Toxicol.* **2009**, *22*, 1929−1937.
(29) Mack, C. I.; Egert, B.; Liberto, E.; Weinert, C. H.; Bub, A.; Hoffmann, I.; Bicchi, C.; Kulling, S. E.; Cordero, C. Robust markers of coffee consumption identified among the volatile organic compounds in human urine. *Mol. Nutr. Food Res.* **2019**, *63*, 1−12.
(30) Fuchsmann, P.; Tena Stern, M.; Bischoff, P.; Badertscher, R.; Breme, K.; Walther, B. Development and performance evaluation of a novel dynamic headspace vacuum transfer "In Trap" extraction method for volatile compounds and comparison with headspace solid-phase microextraction and headspace in-tube extraction. *J. Chromatogr. A* **1601**, *1601*, 60−70.
(31) Trimigno, A.; Münger, L.; Picone, G.; Freiburghaus, C.; Pimentel, G.; Vionnet, N.; Pralong, F.; Capozzi, F.; Badertscher, R.; Vergères, G. GC-MS Based Metabolomics and NMR Spectroscopy Investigation of Food Intake Biomarkers for Milk and Cheese in Serum of Healthy Humans. *Metabolites* **2018**, *8*, 1−27.
(32) Münger, L. H.; Trimigno, A.; Picone, G.; Freiburghaus, C.; Pimentel, G.; Burton, K. J.; Pralong, F. P.; Vionnet, N.; Capozzi, F.; Badertscher, R.; Vergères, G. Identification of Urinary Food Intake Biomarkers for Milk, Cheese, and Soy-Based Drink by Untargeted GC-MS and NMR in Healthy Humans. *J. Proteome Res.* **2017**, *16*, 3321−3335.
(33) Pimentel, G.; Burnand, D.; Münger, L. H.; Pralong, F. P.; Vionnet, N.; Portmann, R.; Vergères, G. Identification of Milk and Cheese Intake Biomarkers in Healthy Adults Reveals High Interindividual Variability of Lewis System-Related Oligosaccharides. *J. Nutr.* **2020**, *150*, 1058−1067.
(34) Edmands, W. M. B.; Ferrari, P.; Scalbert, A. Normalization to specific gravity prior to analysis improves information recovery from high resolution mass spectrometry metabolomic profiles of human urine. *Anal. Chem.* **2014**, *86*, 10925−10931.
(35) Déporte, R.; Amiand, M.; Moreau, A.; Charbonnel, C.; Campion, L. High-performance liquid chromatographic assay with UV detection for measurement of dihydrouracil/uracil ratio in plasma. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2006**, *834*, 170−177.
(36) Xia, J.; Psychogios, N.; Young, N.; Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **2009**, *37*, W652−W660.
(37) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3*, 211−221.
(38) Viant, M. R.; Kurland, I. J.; Jones, M. R.; Dunn, W. B. How close are we to complete annotation of metabolomes? *Curr. Opin. Chem. Biol.* **2017**, *36*, 64−69.
(39) Kilic, M.; Lindsay, R. C. Distribution of Conjugates of Alkylphenols in Milk from Different Ruminant Species. *J. Dairy Sci.* **2005**, *88*, 7−12.
(40) Phua, L. C.; Koh, P. K.; Cheah, P. Y.; Ho, H. K.; Chan, E. C. Global gas chromatography/time-of-flight mass spectrometry (GC/TOFMS)-based metabonomic profiling of lyophilized human feces. *J. Chromatogr., B* **2013**, *937*, 103−113.
(41) Dan, T.; Wang, D.; Wu, S.; Jin, R.; Ren, W.; Sun, T. Profiles of Volatile Flavor Compounds in Milk Fermented with Different Proportional Combinations of Lactobacillus delbrueckii subsp. bulgaricus and Streptococcus thermophilus. *Molecules* **2017**, *22*, 1−14.
(42) Rossbach, B.; Kegel, P.; Letzel, S. Application of headspace solid phase dynamic extraction gas chromatography/mass spectrometry (HS-SPDE-GC/MS) for biomonitoring of n-heptane and its metabolites in blood. *Toxicol. Lett.* **2012**, *210*, 232−239.
(43) Mochalski, P.; Unterkofler, K. Quantification of selected volatile organic compounds in human urine by gas chromatography selective reagent ionization time of flight mass spectrometry (GC-SRI-TOF-MS) coupled with head-space solid-phase microextraction (HS-SPME). *Analyst* **2016**, *141*, 4796−4803.
(44) Loftus, N. J.; Woollen, B. H.; Steel, G. T.; Wilks, M. F.; Castle, L. An assessment of the dietary uptake of di-2-(ethylhexyl) adipate (DEHA) in limited population study. *Food Chem. Toxicol.* **1994**, *32*, 1−5.
(45) Ha, J. K.; Lindsay, R. C. Volatile Branched-chain Fatty Acids and Phenolic Compounds in Aged Italian Cheese Flavors. *J. Food Sci.* **1991**, *56*, 1241−1247.
(46) Urbach, G. Relations Between Cheese Flavour and Chemical Composition. *Int. Dairy J.* **1993**, *3*, 389−422.
(47) Forney, F. W.; Markovetz, A. J. The biology of methyl ketones. *J. Lipid Res.* **1971**, *12*, 383−395.
(48) Walker, N. J.; Keen, A. R. Formation of methyl ketones during ripening of Cheddar cheese. *J. Dairy Res.* **2009**, *41*, 73−80.
(49) Vazquez-Landaverde, P. A.; Velazquez, G.; Torres, J. A.; Qian, M. C. Quantitative Determination of Thermally Derived Off-Flavor Compounds in Milk Using Solid-Phase Microextraction and Gas Chromatography. *J. Dairy Sci.* **2005**, *88*, 3764−3772.
(50) Chen, J.; Wang, W.; Shi, C.; Wan, J.; Deng, L.; Fang, J. Determination of Four Volatile Compounds with Anti-Inflammatory Biological Activity inHouttuynia chordataThunb. by Gas Chromatography and Gas Chromatography−Mass Spectrometry. *Anal. Lett.* **2014**, *47*, 730−741.
(51) Bohbot, J. D.; Dickens, J. C. Insect repellents: modulators of mosquito odorant receptor activity. *PLoS One* **2010**, *5*, 1−11.
(52) Raymer, J.; Wiesler, D.; Novotny, M.; Asa, C.; Seal, U. S.; Mech, L. D. Volatile constituents of wolf (CaMs lupus) urine as related to gender and season. *Experientia* **1984**, *40*, 707−709.
(53) Chen, E. C. H.; Ho, C.-T. Identification of 9-Decenoic Acid in Beer and Yeast. *J. Am. Soc. Brew. Chem.* **2018**, *39*, 70−71.
(54) Kim Ha, J.; Lindsay, R. C. Method for the Quantitative Analysis of Volatile Free and Total Branched-Chain Fatty Acids in Cheese and Milk Fat. *J. Dairy Sci.* **1990**, *73*, 1988−1999.
(55) Del Rosario, R.; De Lumen, B. O.; Habu, T.; Flath, R. A.; Richard Mon, T.; Teranishi, R. Comparison of Headspace Volatiles from Winged Beans and Soybeans. *J. Agric. Food Chem.* **1984**, *32*, 1011−1015.
(56) Piveteau, F.; le Guen, S.; Gandemer, G.; Baud, J.-P.; Prost, C.; Demaimay, M. Aroma of Fresh Oysters Crassostrea gigas: Composition and Aroma Notes. *J. Agric. Food Chem.* **2000**, *48*, 4851−4857.
(57) van de Kant, K. D. G.; van Berkel, J. J. B. N.; Jobsis, Q.; Lima Passos, V.; Klaassen, E. M. M.; van der Sande, L.; van Schayck, O. C. P.; de Jongste, J. C.; van Schooten, F. J.; Derks, E.; Dompeling, E.;

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX

Dallinga, J. W. Exhaled breath profiling in diagnosing wheezy preschool children. *Eur Respir J* **2013**, *41*, 183−188.
(58) Ravi, R.; Taheri, A.; Khandekar, D.; Millas, R. Rapid Profiling of Soybean Aromatic Compounds Using Electronic Nose. *Biosensors* **2019**, *9*, 1−13.
(59) Vezzaro, A.; Krause, S. T.; Nonis, A.; Ramina, A.; Degenhardt, J.; Ruperti, B. Isolation and characterization of terpene synthases potentially involved in flavor development of ripening olive (Olea europaea) fruits. *J. Plant Physiol.* **2012**, *169*, 908−914.
(60) Yalcin, H. K. H. Physical, chemical and bioactive properties of onion (Allium cepa L.) seed and seed oil. *J. Appl. Bot. Food Qual.* **2014**, *87*, 87−92.
(61) Khrisanapant, P.; Kebede, B.; Leong, S. Y.; Oey, I. A Comprehensive Characterisation of Volatile and Fatty Acid Profiles of Legume Seeds. *Foods* **2019**, *8*, 1−19.
(62) Xia, N.; Yang, M.; Zhao, J.; Shao, Y.; Shi, Y.; Yan, W.; Wang, X.; Han, Y.; Wang, Z. Genome-wide association analysis of 1-octen-3-ol content related to soymilk off-flavor in soybean seed. *Crop Pasture Sci* **2019**, *70*, 133−139.
(63) Sethi, S.; Nanda, R.; Chakraborty, T. Clinical application of volatile organic compound analysis for detecting infectious diseases. *Clin Microbiol Rev* **2013**, *26*, 462−475.
(64) Scanlan, R. A.; Lindsay, R. C.; Libbey, L. M.; Day, E. A. Heat-Induced Volatile Compounds in Milk. *J. Dairy Sci.* **1968**, *51*, 1001−1007.
(65) Bosset, J. O.; Liardon, R. Thr aroma composition of Swiss Gruyere cheese. II. The neutral volatile components. *LWT* **1984**, *17*, 359−362.
(66) Marais, J. 1,1,6-Trimethyl-l ,2-dihydronaphthalene (TDN): A Possible Degradation Product of Lutein and beta-Carotene. *S. Afr. J. Enol. Vitic.* **1992**, *13*, 52−55.
(67) Kanamaru, K.; Wang, S.; Abe, J.; Yamada, T.; Kitamura, K. Identification and Characterization of Wild Soybean (Glycine soja Sieb. et Zecc.) Strains with High Lutein Content. *Breed Sci* **2006**, *56*, 231−234.
(68) Seguin, P.; Tremblay, G.; Pageau, D.; Liu, W.; Turcotte, P. Soybean Lutein Concentration: Impact of Crop Management and Genotypes. *Crop Science* **2011**, *51*, 1151−1160.
(69) Zafar, A.; Singh, S.; Naseem, I. Cytotoxic activity of soy phytoestrogen coumestrol against human breast cancer MCF-7 cells: Insights into the molecular mechanism. *Food Chem. Toxicol.* **2017**, *99*, 149−161.
(70) Pimentel, G.; Burton, K. J.; von Ah, U.; Bütikofer, U.; Pralong, F. P.; Vionnet, N.; Portmann, R.; Vergères, G. Metabolic Footprinting of Fermented Milk Consumption in Serum of Healthy Men. *J. Nutr.* **2018**, *148*, 851−860.
(71) Trošt, K.; Ulaszewska, M. M.; Stanstrup, J.; Albanese, D.; De Filippo, C.; Tuohy, K. M.; Natella, F.; Scaccini, C.; Mattivi, F. Host: Microbiome co-metabolic processing of dietary polyphenols - An acute, single blinded, cross-over study with different doses of apple polyphenols in healthy subjects. *Food Res. Int.* **2018**, *112*, 108−128.
(72) Rotches-Ribalta, M.; Andres-Lacueva, C.; Estruch, R.; Escribano, E.; Urpi-Sarda, M. Pharmacokinetics of resveratrol metabolic profile in healthy humans after moderate consumption of red wine and grape extract tablets. *Pharmacol. Res.* **2012**, *66*, 375−382.
(73) Tang, C.; Tan, J.; Fan, R.; Zhao, B.; Tang, C.; Ou, W.; Jin, J.; Peng, X. Quasi-targeted analysis of hydroxylation-related metabolites of polycyclic aromatic hydrocarbons in human urine by liquid chromatography-mass spectrometry. *J. Chromatogr. A* **2016**, *1461*, 59−69.
(74) Zamek-Gliszczynski, M. J.; Hoffmaster, K. A.; Nezasa, K.-I.; Tallman, M. N.; Brouwer, K. L. R. Integration of hepatic drug transporters and phase II metabolizing enzymes: Mechanisms of hepatic excretion of sulfate, glucuronide, and glutathione metabolites. *Eur. J. Pharm. Sci.* **2006**, *27*, 447−486.

https://dx.doi.org/10.1021/acs.jproteome.0c00324
J. Proteome Res. XXXX, XXX, XXX−XXX