

Identification of Milk and Cheese Intake Biomarkers in Healthy Adults Reveals High Interindividual Variability of Lewis System–Related Oligosaccharides

Grégory Pimentel,¹ David Burnand,¹ Linda H Münger,¹ François P Pralong,² Nathalie Vionnet,² Reto Portmann,¹ and Guy Vergères¹

¹Federal Department of Economic Affairs, Education, and Research, Agroscope, Bern, Switzerland; and ²Service of Endocrinology, Diabetes, and Metabolism, Lausanne University Hospital, Lausanne, Switzerland

ABSTRACT

Background: The use of biomarkers of food intake (BFIs) in blood and urine has shown great promise for assessing dietary intake and complementing traditional dietary assessment tools whose use is prone to misreporting.

Objectives: Untargeted LC-MS metabolomics was applied to identify candidate BFIs for assessing the intake of milk and cheese and to explore the metabolic response to the ingestion of these foods.

Methods: A randomized controlled crossover study was conducted in healthy adults [5 women, 6 men; age: 23.6 ± 5.0 y; BMI (kg/m²): 22.1 ± 1.7]. After a single isocaloric intake of milk (600 mL), cheese (100 g), or soy-based drink (600 mL), serum and urine samples were collected postprandially up to 6 h and after fasting after 24 h. Untargeted metabolomics was conducted using LC-MS. Discriminant metabolites were selected in serum by multivariate statistical analysis, and their mass distribution and postprandial kinetics were compared.

Results: Serum metabolites discriminant for cheese intake had a significantly lower mass distribution than metabolites characterizing milk intake ($P = 4.1 \times 10^{-4}$). Candidate BFIs for milk or cheese included saccharides, a hydroxy acid, amino acids, amino acid derivatives, and dipeptides. Two serum oligosaccharides, blood group H disaccharide (BGH) and Lewis A trisaccharide (LeA), specifically reflected milk intake but with high interindividual variability. The 2 oligosaccharides showed related but opposing trends: subjects showing an increase in either oligosaccharide did not show any increase in the other oligosaccharide. This result was confirmed in urine.

Conclusions: New candidate BFIs for milk or cheese could be identified in healthy adults, most of which were related to protein metabolism. The increase in serum of LeA and BGH after cow-milk intake in adults calls for further investigations considering the beneficial health effects on newborns of such oligosaccharides in maternal milk. The trial is registered at clinicaltrials.gov as NCT02705560. *J Nutr* 2020;150:1058–1067.

Keywords: biomarker, dairy, cheese, milk, soy, postprandial, metabolomics, serum metabolome, Lewis antigen

Introduction

The validity of nutritional studies relies on an accurate assessment of participants' food intake and dietary exposure. However, traditional tools of dietary assessment such as 24-h recalls, food-frequency questionnaires or food records are prone to under-/overestimation bias as they rely on subjects to self-report (1–3). Biomarkers of food intake (BFIs) have shown great promise as a tool for addressing this issue. BFIs are compounds measured within the body that accurately reflect the intake of a food or food constituent (4, 5). BFIs can be considered as unbiased and objective surrogate estimators of food intake, but also as a tool to evaluate metabolic status and to highlight associations between diet and metabolic disorders (2, 6). For the discovery of BFIs, the Joint Programming

Initiative “A Healthy Diet for a Healthy Life” Food Biomarkers Alliance (FoodBALL) has conducted standardized intervention studies consisting of postprandial sampling after an acute intake of a specific test food, completed with 24-h fasting sampling (6). The consortium opted for an untargeted metabolomics analysis of urine and serum samples as this method allows the simultaneous measurement of a large number of low-molecular-weight metabolites. Using this approach, the consortium has successfully identified candidate BFIs for several foods, including cereals (7), fruits (8, 9), meat (10), peas (11), and milk and cheese (12, 13). These studies highlighted the importance of the postprandial phase for the detection of candidate BFIs, as well as the challenges particular to the validation of BFIs (i.e., the lack of specificity for the tested

food and the confirmation of the observed postprandial trends in the fasting state), and therefore the necessity to validate results in cohort studies under free-living conditions (6, 14). Few nutritional interventions have aimed to identify BFIs for dairy products despite their prominent place in Western diets and their widely discussed effects on health. A systematic review published by members of the FoodBALL consortium listed and evaluated putative BFIs for dairy, including metabolites related to lipid (C17:0, C:15:0, C17:1, myristoyl-sphingomyelin or methyl palmitate), galactose (galactonate), or amino acid metabolism for cheese intake (15). However, most of the reviewed studies were not specifically designed for BFI discovery and the authors recommended further investigation to validate the proposed biomarkers. Within the framework of the FoodBALL project, an acute intervention study focused on the discovery of dairy BFIs. GC-MS and NMR analytical platforms characterized milk intake by an increase in lactose-derived compounds (lactose, galactose, galactitol, galactonate, and galactono-1,5-lactone), while most BFIs for cheese were related to protein and amino acid metabolism (3-phenyllactic acid, methionine, proline, leucine, tyrosine, valine, and isoleucine) (12, 13).

In addition to BFI identification, postprandial untargeted metabolomics can comprehensively characterize the specific metabolic response of individuals to the ingestion of a test food. The interindividual variability observed in the postprandial phase can be used to 1) estimate the capacity of a subject to respond to a dietary challenge and thus evaluate overall metabolic health (16, 17), 2) identify dysregulation of metabolic pathways that is not necessarily visible in the fasting state (18, 19), and 3) identify metabolic subgroups of subjects according to their response to the ingestion of a test food and further investigate these groups in the context of other nutrigenomic tools (20, 21).

This report presents results obtained from the FoodBALL nutritional intervention for dairy products for serum and urine samples analyzed by LC-MS-based metabolomics. By covering a specific spectrum of the metabolome, LC-MS analysis aimed to identify new putative BFIs for milk or cheese (compared with a soy drink as a nondairy control), and therefore complement the previous results obtained by GC-MS and NMR (12, 13). By further characterizing the postprandial kinetics and interindividual variability of BFIs, we evaluated their potential to reflect not only food intake but also metabolic phenotypes.

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 European Healthy Diet for a Healthy Life Joint Programming Initiative "The Food Biomarkers Alliance" (FoodBALL).

Author disclosures: The authors report no conflicts of interest.

Supplemental Tables 1–4, Supplemental Figures 1–15, and Supplemental Methods are available from the "Supplementary data" link the online posting of the article and from the same link in the online table of contents at <https://academic.oup.com/jn/>.

Address correspondence to GV (e-mail: guy.vergeres@agroscope.admin.ch).

Abbreviations used: BFI, biomarker of food intake; BGH, blood group H disaccharide; CV-ANOVA, ANOVA of the cross-validated residuals; FoodBALL, Food Biomarkers Alliance; FP, phenylalanyl-proline; FUT, fucosyltransferase; iAUC, incremental AUC; ILA, indole-3-lactic acid; LeA, Lewis A trisaccharide; nparLD, nonparametric analysis of longitudinal data; OPLS-DA, orthogonal projections to latent structures discriminant analysis; PLS-DA, partial least-square discriminant analysis; VIP, variable importance in projection for the predictive component; VT, valyl-threonine.

Methods

Study design

The present project was conducted within the framework of the FoodBALL project. Given the exploratory nature of the untargeted metabolomics analysis, the desired sample size for the study could not be calculated a priori and is therefore based on similar published crossover nutritional studies that used untargeted metabolomics analyses on postprandial samples. A sample size in the range of 10–15 subjects has been shown to give sufficient statistical power to identify regulated metabolites and, in particular, potential BFIs (6, 10, 22, 23). The study was a controlled, crossover study and conducted in 11 participants (5 women and 6 men; Supplemental Figure 1). The subjects were regular consumers of dairy products, aged 19–31 y, and with a BMI (in kg/m²) ranging between 18 and 30. Details regarding exclusion and inclusion criteria and characteristics of study population are given by Münger et al. (12). The study was conducted in accordance with the Declaration of Helsinki, received ethical approval from the Commission Cantonale d'Ethique de la Recherche sur l'Être Humain (Vaud, Switzerland), and was registered at clinicaltrials.gov as NCT02705560. All subjects provided written informed consent beforehand.

The study followed the general design defined within the FoodBALL project (Figure 1). Briefly, the subjects came to the study center after 12 h of fasting for an intervention day. The 3 foods tested were 600 mL organic pasteurized full fat milk (3.9% fat, Coop Naturaplan, Switzerland), 100 g hard cheese (Le Gruyère AOP, Fromagerie Bullet, Bullet, Switzerland) plus 500 mL of water, and 600 mL soy drink, which was a mixture of soy milk and added vegetable fat (Soja Line, Migros, Switzerland). All 3 test products were isocaloric (~400 kcal); details regarding their composition are given by Münger et al. (12). Blood samples were taken before ($t = 0$ h) and postprandially after the ingestion of 1 of the 3 test products ($t = 1, 2, 4,$ and 6 h). Urine samples were taken postprandially during the time intervals 0–1, 1–2, 2–4, and 4–6 h. After the postprandial phase, urine was collected during the time intervals 6–12 and 12–24 h. The following morning, blood samples were collected under fasting conditions ($t = 24$ h). During the 2-d run-in phase, diet was restricted to non-bovine-, non-dairy-, and non-soy-containing foods. In addition, participants were given standardized meals the evening before and during the intervention day. Details regarding blood and urine sample preparation for LC-MS analysis are given in the Supplemental Methods.

Untargeted LC-MS-based metabolomics analysis

The UltiMate 3000 HPLC system (Thermo Fisher Scientific), coupled with the maXis 4G+ quadrupole time-of-flight mass spectrometer (Bruker Daltonik GmbH) were used for the untargeted metabolomics analysis. The mass spectrometer electrospray interface was operating in positive ion mode and spectra were recorded from m/z 75 to m/z 1500. Collision-induced dissociation was performed using energies from 20 to 70 eV. Detailed serum and urine sample preparation as well as chromatography and spectrometer settings are given in the Supplemental Methods.

Data processing and statistical analysis

Reduction and filtering of serum and urine datasets.

Progenesis QI (v.2.3.6198.24128; NonLinear Dynamics Ltd.) was used for retention time correction, peak-picking, deconvolution, adducts annotation, and normalization (default automatic sensitivity and without minimum peak width). The dataset was corrected to account for signal drift, and reduced via multiple filtering steps to remove metabolites with poor repeatability and potential contaminants (Supplemental Methods). The net incremental AUC (iAUC) was calculated for each metabolite in the postprandial phase using R (v.3.1.2, MESS R package; R Foundation for Statistical Computing). After principal components analysis, a group of samples obtained from one male subject were considered as technical outliers (outside the Hotelling T² ellipse with 95% CI) and removed from the dataset, resulting in ten subjects remaining out of eleven subjects. The dataset was further reduced by retaining only metabolites

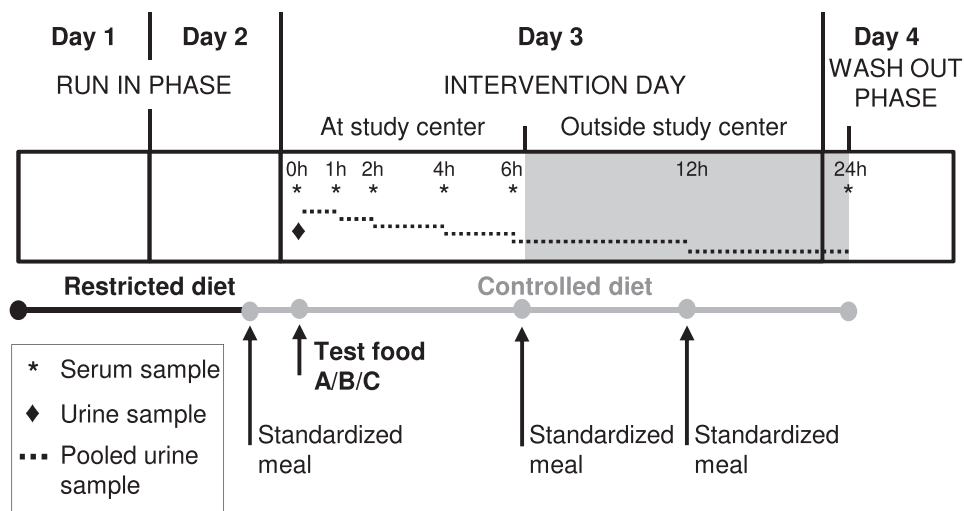


FIGURE 1 Overview of the controlled, crossover study design. During the intervention day, 11 healthy subjects ingested isocaloric doses of three test products in random order (A = milk, B = cheese, C = soy drink). Serum and pooled urine samples were collected after an overnight fast and postprandially, up to 6 h for serum and up to 12 h for urine. The following day after fasting, serum and pooled urine samples were collected. The participants followed a 2-d restricted diet before the intervention day and consumed a controlled diet composed of standardized meals during the study.

that presented a significant postprandial response after ≥ 1 of the three test foods [nonparametric analysis of longitudinal data (nparLD); R package; 0.05 as the P -value significance cutoff] (24). Similarly, a paired Wilcoxon signed-rank test was used to identify and filter for metabolites presenting a significant change in fasting serum concentrations after 24 h ($t = 0$ h vs. $t = 24$ h). The same data-reduction procedure was used on the dataset from urine samples.

Selection of discriminant metabolites by multivariate statistical analysis.

Partial least-square discriminant analysis (PLS-DA) was performed to differentiate serum samples collected after milk, cheese, and soy drink intake (SIMCA-P software v.14.0; Umetrics). Orthogonal PLS-DA (OPLS-DA) was carried out for 1 by 2 comparisons (milk vs. cheese/soy drink, cheese vs. milk/soy drink, and soy drink vs. cheese/milk). The dataset was scaled using the Pareto method. Validity of the models was evaluated by the goodness-of-fit parameter (R^2Y), the predictive ability parameter (Q^2 ; calculated by 10-fold cross-validation, $Q^2 > 0.50$ as a cutoff value), ANOVA of the cross-validated residuals to confirm the reliability of the model (CV-ANOVA; $P < 0.05$ as a significance threshold) (25), and permutation tests with 999 random permutations to exclude any random separation of the sample groups (26). Finally, discriminant metabolites were selected based on variable importance in projection (VIP) scores (VIP > 1.5 as a cutoff value).

Univariate statistical analysis, postprandial kinetics, and mass distribution of discriminant metabolites.

For the discriminant metabolites selected by multivariate analysis (VIP score > 1.5), the Kruskal-Wallis test was also applied to assess the effect of test foods, at each time point and on the 6-h postprandial iAUC ($P < 0.05$ as the significance threshold). When significant, the Kruskal-Wallis test was completed by a post-hoc Conover-Iman pairwise comparison test. The same approach was used to compare the levels of selected discriminant metabolites in different subgroups of subjects. Metabolites discriminant for milk, cheese, or soy intake were clustered according to their postprandial kinetics after the intake of each of the three foods. Hierarchical clustering analysis was performed using the R *heatmap* package (version 1.0.12) with maximum distance measure and Ward's linkage. The mass distributions of discriminant metabolites for milk and cheese intake (smoothed densities of m/z values using R) were compared using the Kolmogorov-Smirnov test ($P < 0.05$ as the significance threshold).

Identification of discriminant metabolites

The Human Metabolome Database (27), the MassBank of North America (28), and the National Institute of Standards and Technology database (NIST v14) were used with a 5-ppm mass accuracy threshold for the identification of discriminant metabolites. Identity suggestions from databases were then confirmed by MS fragmentation data (when available) and/or with the injection of pure standards solutions. The list of standards suppliers is given in the Supplemental Methods.

The level of identification of each discriminant metabolite is defined according to the Metabolomics Standards Initiative recommendations (29), as follows—level 1: compounds were identified by comparison to a pure reference based on spectral data (molecular weight with a 5-ppm accuracy threshold, fragmentation pattern when available, isotopic distribution, and retention time with 10% accuracy threshold); level 2: without chemical standards, based on spectral data; level 3: putatively characterized compound classes; and level 4: unknown compound. Discriminant metabolites were considered as putative BFIs if they showed an increase in postprandial or fasting serum collected after the intake of the food of interest. The presence of the putative BFIs selected in serum was investigated in the filtered urine dataset.

Results

Serum metabolome postprandial response

Following data filtering, a total of 2988 unique serum metabolites were detected. A total of 1639 metabolites (55%) showed a significant postprandial kinetic after the ingestion of ≥ 1 of the three test foods (nparLD test, nonadjusted $P < 0.05$). Interestingly, one-third of the postprandial metabolome (973 metabolites, 32.6%) showed a postprandial response specifically after the intake of one of the three foods (382 metabolites showed a dynamic response only after the intake of milk, 426 only after cheese intake, and 165 only after soy intake). A total of 430 metabolites (14%) showed a significant difference in fasting serum 24 h after the intake of ≥ 1 of the test foods (Wilcoxon test, nonadjusted $P < 0.05$). Detailed numbers of metabolites showing a significant postprandial response or significant change in fasting serum are given in Table 1.

TABLE 1 Serum metabolites presenting a significant postprandial response, or a significant change in 24 h fasting serum after the ingestion of milk, cheese, or soy in healthy adults¹

	Significant postprandial response ²		Significant change in 24-h fasting serum ³	
	Serum metabolites, <i>n</i>	% of total	Serum metabolites, <i>n</i>	% of total
After ≥1 food	1639	54.9	430	14.4
Including milk	901	30.2	184	6.2
Including cheese	1005	33.6	155	5.2
Including soy	736	24.6	220	7.4
After 1 food only	973	32.6	333	11.1
Milk only	382	12.8	102	3.4
Cheese only	426	14.3	96	3.2
Soy only	165	5.5	135	4.5
After ≥2 foods	666	22.3	98	3.3
Milk and cheese	432	14.5	43	1.4
Milk and soy	424	14.2	69	2.3
Cheese and soy	484	16.2	46	1.5
After 2 foods only	329	11.0	68	2.3
Milk and cheese only	95	3.2	13	0.4
Milk and soy only	87	2.9	39	1.3
Cheese and soy only	147	4.9	16	0.5
After all 3 foods	337	11.3	30	1.0

¹ *n* = 10 subjects. Serum metabolome measured by LC-MS untargeted metabolomics, 2988 unique metabolites detected.

² Postprandial time effect evaluated by nonparametric longitudinal data analysis (*P* < 0.05 as the significance threshold).

³ Change in fasting serum evaluated by paired Wilcoxon signed-rank test (*P* < 0.05 as the significance threshold).

Differentiation of serum samples by multivariate statistical analysis

PLS-DA analysis using the selection of 1639 postprandial serum metabolites showed a clear separation of serum samples taken postprandially after milk, cheese, or soy intake (Supplemental Figure 2). The model showed good fit and predictability ($R^2 = 0.88$, $Q^2 = 0.62$, CV-ANOVA $P = 1.3 \times 10^{-3}$).

Postprandial 1-by-2 OPLS-DA comparisons indicated that serum taken after a test food could be differentiated from serum taken after the 2 other foods (milk vs. soy/cheese, cheese vs. milk/soy, or soy vs. milk/cheese). The strongest separation was obtained when comparing serum samples taken after soy drink intake against nonsoy intake (soy vs. milk/cheese) ($R^2 = 0.95$, $Q^2 = 0.78$, CV-ANOVA $P = 5.3 \times 10^{-8}$). The lowest predictability was observed when comparing milk intake against nonmilk intake (milk vs. cheese/soy) ($R^2 = 0.88$, $Q^2 = 0.55$, CV-ANOVA $P = 3.4 \times 10^{-3}$) (Supplemental Figures 3–5). Conversely, PLS-DA analysis using the selection of 430 metabolites could not differentiate serum samples taken under fasting condition 24 h after the ingestion of milk, cheese, or soy ($R^2 = 0.36$, $Q^2 = 0.16$, CV-ANOVA $P = 3.8 \times 10^{-1}$; data not shown). OPLS-DA 1-by-2 comparisons showed that only soy intake compared with nonsoy intake led to a difference in 24-h fasting serum, the Q^2 value reaching almost the cut-off value of 0.50 ($R^2 = 0.91$, $Q^2 = 0.45$, CV-ANOVA $P = 8.8 \times 10^{-3}$) (Supplemental Figure 6). Serum samples taken 24 h after milk or cheese intake could not be differentiated from samples taken after nonmilk and noncheese intake, respectively ($R^2 = 0.97$, $Q^2 = 0.37$, CV-ANOVA $P = 2.1 \times 10^{-1}$, and $R^2 = 0.99$, $Q^2 = 0.31$, CV-ANOVA $P = 5.8 \times 10^{-1}$, respectively; data not shown).

A total of 297 metabolites were considered discriminant (VIP > 1.5): 261 in the postprandial phase, including 65 metabolites for milk intake (25%); 124 metabolites for cheese intake (47%); and 72 for soy intake (28%). Forty-five metabolites

were considered discriminant for soy intake in 24-h fasting serum.

Metabolites discriminant for milk, cheese, and soy drink intake showed different postprandial kinetics and mass distributions

The postprandial kinetics of metabolites discriminant for milk, cheese, or soy drink intake were classified by hierarchical clustering analyses (Figure 2). Interestingly, heatmaps showed that metabolites discriminant for milk, cheese, or soy drink intake present different postprandial kinetics. While metabolites discriminant for cheese intake (Figure 2B) are mostly metabolites that increased in serum during the early (cluster 1) or late (cluster 2) postprandial phase, metabolites discriminant for milk intake (Figure 2A) are essentially metabolites that, compared with cheese and soy drink intake, remain unchanged (cluster 2) or decreased (cluster 3) postprandially. Only one-third of the metabolites discriminant for milk intake showed increased iAUCs compared with cheese and soy intake (cluster 1). Most of metabolites discriminant for soy drink intake (Figure 2C) either decreased (cluster 1) or increased only in the late postprandial phase (cluster 2).

In addition, to further characterize the selected BFIs, and in particular the influence of the fermentation step, metabolites discriminant for milk and cheese intake were compared according to their mass distribution. This analysis revealed bimodal distributions of the masses centered at masses (m/z) of ~300 and 600 Da. However, compared with the response after milk intake, the relative distribution of these masses was shifted toward the lower masses after cheese intake. Overall, a statistical analysis of these distributions showed that metabolites discriminant for cheese intake had a significantly lower mass distribution than metabolites discriminant for milk intake ($P = 4.1 \times 10^{-4}$; Figure 2D).

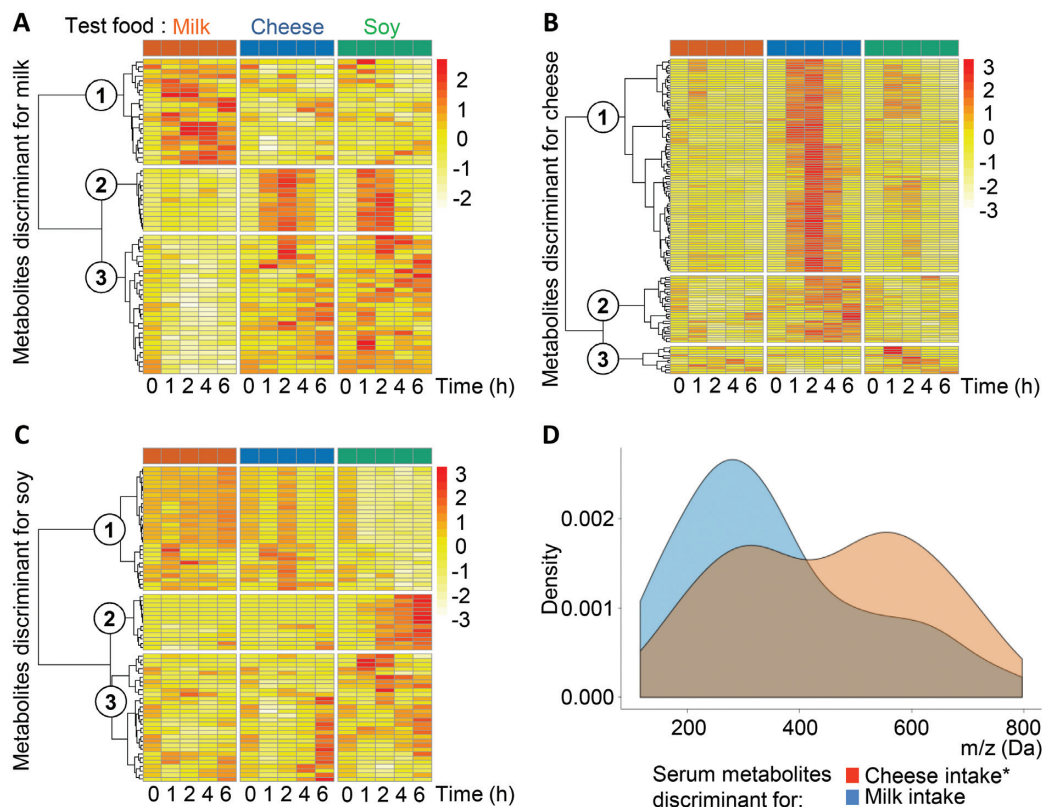


FIGURE 2 Postprandial kinetics and mass distributions of metabolites discriminant for milk, cheese, or soy drink intake in serum of healthy adults ($n = 10$). Metabolites discriminant for milk (A), cheese (B), and soy drink (C) intake were selected by 1-by-2 OPLS-DA analysis (VIP score > 1.5 as the significance threshold). Hierarchical clustering analysis according to postprandial kinetics after the intake of each of the 3 foods was performed using Maximum's distance measure and Ward linkage. The circled numbers 1 to 3 in the figures refer to the cluster numbering. Mass distribution density plots of discriminant metabolites (D) were compared using the Kolmogorov-Smirnov test. *Significant difference between mass distributions ($P < 0.05$). OPLS-DA, orthogonal projections to latent structures discriminant analysis; VIP score, variable importance in projection for the predictive component.

Identification of discriminant metabolites and selection of putative BFIs

Of the 297 discriminant metabolites that were selected in postprandial and/or fasting serum, 9 have been identified at level 1, 23 at level 2, and 30 at level 3, and 199 at level 4 remained unknown. Supplemental Table 1 details all of the discriminant metabolites.

Among the discriminant metabolites with level 1 identification, 8 were considered as putative BFIs (Table 2), with 5 also being detected in urine. Two BFIs were specific to milk intake: blood group H disaccharide (2-O-a-L-fucopyranosyl-galactose; BGH) and Lewis A trisaccharide (3-O-b-D-galactopyranosyl-4-O-a-L-fucopyranosyl-N-acetyl-D-glucosamine; LeA). The analytical method used could not differentiate galactonic acid from gluconic acid (stereoisomers) as the standards of both compounds had identical retention time, but this metabolite was specific to milk intake. Six discriminant metabolites were considered as putative BFIs for cheese: amino adipic acid, citrulline, valyl-threonine (VT), phenylalanyl-proline (FP), indole-3-lactic acid (ILA), and proline. The specificity of amino adipic acid is, however, limited as it showed a significant postprandial increase after the intake of each of the 3 foods (nparLD, $P < 0.05$). The most important change was observed after cheese intake, especially after 4 h, resulting in a higher iAUC (VIP cheese vs. non-cheese = 1.84).

Supplemental Figures 7–15 show the serum and urine postprandial kinetics of the 8 putative BFIs, and

Supplemental Tables 2 and 3 present the P values for the Kruskal-Wallis and the Conover-Iman tests at each time point.

LeA and BGH after milk intake

LeA and BGH were identified as putative BFIs for milk. Both showed significant postprandial responses (serum nparLD: $P = 1.6 \times 10^{-2}$ and $P = 1.9 \times 10^{-2}$, respectively), although their kinetics indicated high interindividual variability (Supplemental Figures 7 and 8). Interestingly, individual kinetics revealed that this variability was due to the presence of three distinct groups of subjects (Figure 3): subjects showing a postprandial increase for LeA but not for BGH (LeA increase group), subjects showing a postprandial increase for BGH but not for LeA (BGH increase group), and subjects who did not show any increase for either of the two metabolites (No increase group). The postprandial increase in LeA and BGH reached a peak at ~ 2 h, and returned to baseline after 6 h. In both cases, no difference was visible in 24-h fasting serum. Interestingly, the patterns observed in serum for BGH were confirmed in urine (Figure 3C, F) but not for LeA as it was not detected in urine. Significant differences were observed between the three groups of subjects when comparing serum or urine iAUCs using a Kruskal-Wallis test completed by a Conover-Iman pairwise comparison test (Figure 3D–F; Supplemental Table 4).

TABLE 2 Identified BFIs for milk or cheese selected by 1-by-2 OPLS-DA analysis in serum of healthy adults¹

Retention time, min	Adduct mass, Da/charge	Neutral mass, Da	Identification	HMDB ID	Adducts	Identification level ²	Food source of BFIs ³
1.07	349.110	326.121	Blood group H disaccharide*	HMDB06590	+Na	1	Milk
1.14	530.207	529.200	Lewis A trisaccharide	HMDB06582	+H, +Na	1	Milk
1.02	219.047	196.058	Galactonic acid/gluconic acid*	HMDB00565/HMDB00625	+Na	3 (hydroxy acid)	Milk
1.07	162.076	161.068	Aminoadipic acid*	HMDB00510	+H, +2Na-H, +Na, +H-H ₂ O	1	Cheese
0.99	176.103	175.095	Citrulline	HMDB00904	+H, +2Na-H, +Na	1	Cheese
1.18	219.134	218.127	Valyl-threonine	HMDB29137	+H, 2M+H	1	Cheese
5.08	263.139	262.131	Phenylalanyl-proline*	HMDB11177	+H	1	Cheese
6.55	206.081	205.074	Indolelactic acid*	HMDB00671	+H, +Na, +NH ₄ , +2Na-H	1	Cheese
1.07	116.071	115.063	Proline	HMDB00162	+H, +Na, +2Na-H	1	Cheese

¹ $n = 10$ subjects. The isomers galactonic acid and gluconic acid could not be differentiated. *BFIs also detected in urine. Postprandial kinetics in serum and urine are detailed in Supplemental Figures 7–15. BFI, biomarker of food intake; HMDB ID, Human Metabolome Database Identification number; OPLS-DA, orthogonal projections to latent structures discriminant analysis.

² Identification levels: 1, identified by comparison to a pure reference; 3, putatively characterized compound classes.

³ Selected as a BFI when VIP score > 1.5 in 1-by-2 OPLS-DA analyses.

Discussion

General trends

Using untargeted metabolomics, we observed that the intake of milk, cheese, or soy leads to distinct postprandial serum metabolomes. Likely due to the plant origin of the test food, soy intake results in the most specific response, in accordance with the observations of M \ddot{u} nger et al. (12) and Trimigno et al. (13). The results show that the majority of metabolites have a

postprandial response that is not specific to the intake of a single food. In fact, we expect that the fraction of the postprandial metabolome responding specifically to the intake of a single food would further decrease as the number of control foods is increased, rendering the identification of specific BFIs difficult.

The postprandial phase after acute intake showed its efficiency to identify putative BFIs; however, their validity was not confirmed in fasting serum 24 h after ingestion, as

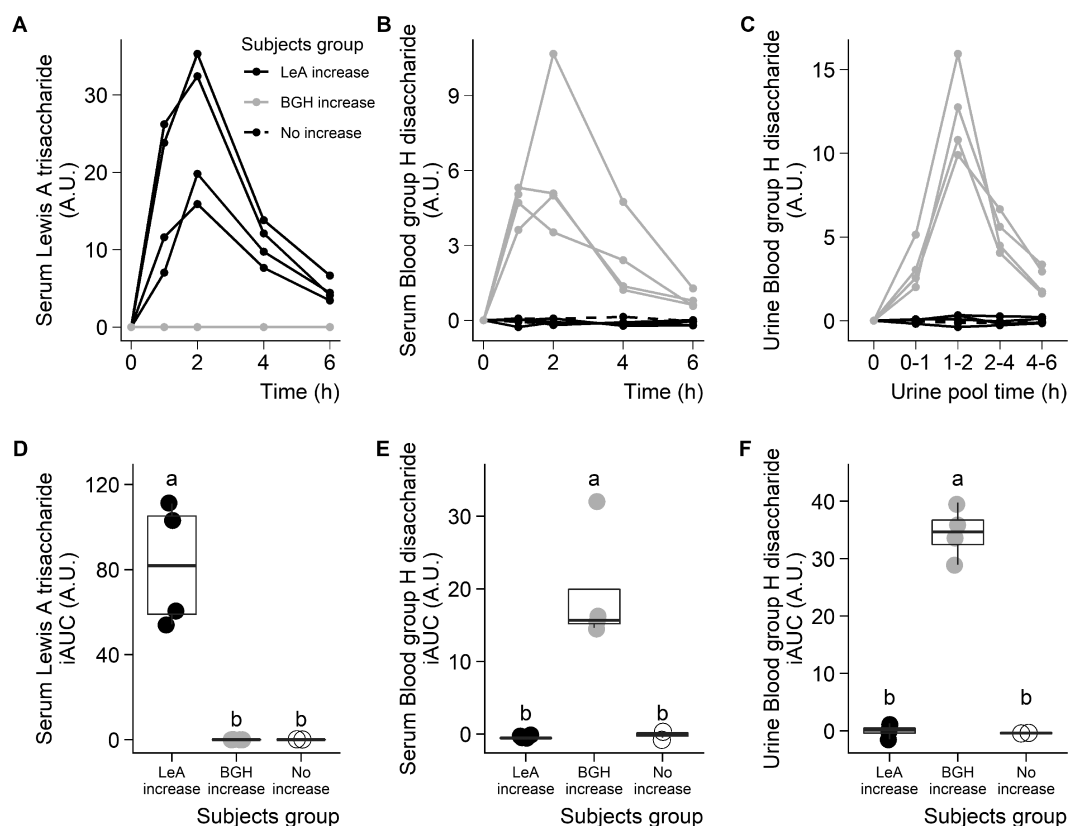


FIGURE 3 Interindividual variability in LeA and the BGH after milk intake in healthy adults. Serum LeA (A), serum BGH (B), and urine BGH (C) measured 6 h postprandially after milk intake in ten healthy subjects. Subjects were separated into three groups according to their postprandial response: LeA increase ($n = 4$), BGH increase ($n = 4$), and no increase ($n = 2$). Comparison of postprandial responses of the three subject groups used iAUCs (D–F), IQRs plotted with medians. Means with different superscript letters (a, b, c) denote significant differences using a Kruskal-Wallis test completed by a Conover-Iman pairwise comparison test ($P < 0.05$). A.U., arbitrary units; BGH, blood group H disaccharide; iAUC, incremental AUC; LeA, Lewis A trisaccharide.

previously observed (12). As the heatmap of kinetics suggests, the concentrations of many of the discriminant metabolites returned to baseline values 6 h after ingestion, in line with the very nature of nutrition, which is characterized by the repetition of multiple daily meals across the lifespan. The transient nature of the postprandial metabolome generally indicates that the identified metabolites can likely only be used as short-term BFIs or as markers of the quality of the metabolic response of the human organism to food intake. Nevertheless, the small sample size of the study and the single intake tested here may have prevented confirmation of significant differences in fasting serum (if any). Therefore, a validation of the postprandial BFIs in fasting samples from observational cohorts is essential and may be efficient as such studies provide larger sample sizes and samples taken after repeated intakes (14, 30, 31).

Heatmaps also reveal that metabolites discriminant for milk, cheese, or soy drink intake differ in their overall postprandial kinetics (early or late increase or decrease). Surprisingly, about one-third of the metabolites discriminant for milk intake decrease postprandially while increasing after cheese or soy intake. This observation suggests that milk intake downregulates endogenous metabolites and pathways. Conversely, most of metabolites discriminant for cheese intake increase postprandially and, therefore, are most likely absorbed exogenous compounds. Interestingly, the BFIs for cheese have a significant lower mass distribution than the BFIs for milk. A similar observation was made when comparing milk and yogurt intake (32). This shift in the postprandial mass distribution of discriminant metabolites could be indicative of a “predigestive” effect of fermentation on milk, resulting in the production and the absorption of specific low-molecular-weight compounds (e.g., SCFAs, oligonucleotides, or free amino acids).

Biomarkers of cheese intake

While amino adipic acid was discriminant for cheese intake, its use as a BFI is not recommended as increases were also observed after milk and soy intake. Amino adipic acid is an intermediate in lysine metabolism (33), and a significant postprandial increase in serum lysine was observed after the intake of all three foods (data not shown). The higher concentrations of amino adipic acid in serum after cheese intake could be attributed to higher amounts in cheese as it has been identified as a product of fermentation (34, 35). Nevertheless, higher concentrations of amino adipic acid have also been reported after meat and fish intake (36) and after purple grumixama intake (37).

In contrast to amino adipic acid, citrulline was a clear BFI for cheese, especially after 1 and 2 h. Citrulline can be produced from dietary arginine, proline, or glutamine by enterocytes (38), and consequently its concentration in blood is used by clinicians to evaluate gut integrity (39). Stanstrup et al. (22) reported higher serum citrulline concentrations at 2 h postprandially after whey protein intake compared with casein, gluten, or cod protein intake, suggesting greater arginine or proline uptake. Although we did not observe a significant difference in arginine concentrations (data not shown), we did see higher proline concentrations after the intake of cheese. The production of citrulline from arginine by lactic acid bacteria has been described during soy, cabbage, or milk fermentation (40–42). It is thus likely that cheese had higher citrulline contents than milk or soy. The use of this molecule as a BFI for cheese has not been reported and needs to be further explored.

Two dipeptides with level 1 identification showed a clear increase after cheese intake. “Predigestion” of milk proteins

by lactic acid bacteria during fermentation might explain the higher concentrations of peptides in serum after the ingestion of fermented dairy products compared with after milk intake (43). VT and FP are both present in milk protein sequences: β -lactoglobulin (VT), serum albumin (VT and FP), κ -casein (VT), α -S2-casein (FP), and β -casein (FP) (44). Given that any dipeptide sequence should appear randomly every 400 amino acids along proteomes, the specificity of the two dipeptides as BFIs for cheese is questionable. Their relevance as a BFI has therefore to be confirmed since they are likely to be found in other dietary proteins. However, the frequency of valine, threonine, phenylalanine, and proline in major dairy proteins is higher compared with that in nondairy proteins (45), suggesting that the presence in serum or urine of these dipeptides, possibly in combination with others, might be useful biomarkers to differentiate nonfermented from fermented dairy consumption.

The release of free amino acids during cheese ripening has been well described (46) and would directly explain the higher concentration in serum of free proline after cheese intake [also observed using NMR (13) and GC-MS (12)]. It could also, indirectly, explain the increase in serum ILA observed in our study as ILA is an end-product of tryptophan metabolism via the indole pathway. Free tryptophan might be metabolized into ILA either by lactic acid bacteria during cheese ripening (32, 47) or by the host microbiota as intestinal *Lactobacilli* use tryptophan as a source of energy in high tryptophan conditions, notably through the indole pathway (48, 49). A similar increase in postprandial serum ILA has been described after yogurt intake compared with non-fermented milk intake, with a persisting trend in fasting serum after 2 wk of daily consumption (32). These observations qualify ILA (and potentially other indole derivatives) as candidate BFIs for fermented dairy consumption.

Biomarkers of milk intake

The presence of galactonic acid in serum samples is likely due to the oxidation of serum galactose, which is known to increase postprandially following the ingestion of a high quantity of milk (20, 21). Galactonic acid was previously detected in postprandial serum and urine samples from the same study using GC-MS (12, 13), confirming its potential use as a biomarker of milk intake, as proposed by Vionnet et al. (21) in postprandial samples and Playdon et al. (50) in fasting serum samples. While measurement using LC-MS could not differentiate galactonic acid from its isomer gluconic acid, the above-mentioned literature suggests that the biomarker would indeed be galactonic acid.

LeA and BGH as biomarkers of milk intake subject to interindividual variability

The presence of LeA and BGH in serum is particularly interesting as it demonstrates, for the first time, their appearance in adults after the ingestion of bovine milk. LeA and BGH are two fucosylated oligosaccharides composed of galactose-fucose-N-acetylglucosamine and fucose-galactose, respectively (Figure 4). Contrary to human milk oligosaccharides, bovine milk oligosaccharides are not fucosylated (51, 52), or are in a very limited proportion (<1%) (53); the two compounds might therefore be of endogenous origin. In humans, fucosylation of carbohydrates occurs in the epithelial lining of the pulmonary, urinary, reproductive, and most importantly, gastrointestinal tract. The resulting fucosylated compounds are then found at the tissue surface, in the respective mucosal secretions, as well as in plasma (54, 55). Intestinal fucosylation takes place in

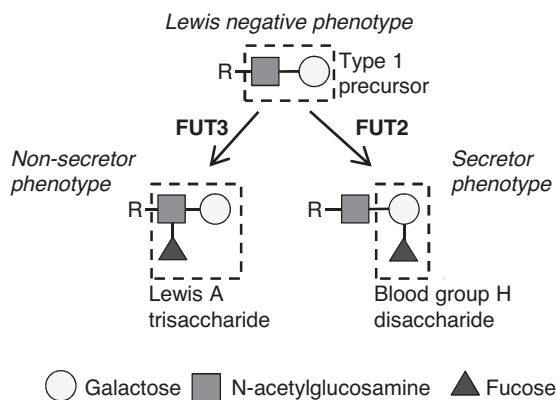


FIGURE 4 Schematic diagram for the formation of BGH or LeA motives on a type 1 precursor chain, in secretor and nonsecretor subjects. BGH, blood group H disaccharide; FUT2, fucosyl transferase 2; FUT3, fucosyl transferase 3; LeA, Lewis A trisaccharide.

the brush-border enterocytes and mainly relies on expression of the fucosyltransferase (*FUT*) 2 and 3 genes (*FUT2*, *FUT3*), which control the addition of a fucose molecule, respectively, on the galactose or the N-acetylglucosamine of a type 1 precursor chain (Figure 4) (55). *FUT2* and *FUT3* enzymes act in competition, with *FUT2* having greater activity. This antigenic pattern consisting of a fucosylated type 1 precursor is at the basis of the Lewis antigen system. In humans, three main Lewis phenotypes have been described due to polymorphisms on the *FUT2* and *FUT3* genes: secretor status (*FUT2* active, resulting in fucosylated galactose, ~72% in Caucasians), nonsecretor status (*FUT2* inactive, *FUT3* active, resulting in fucosylated N-acetylglucosamine, ~22% in Caucasians), and Lewis negative (*FUT2* and *FUT3* both inactive, resulting in no fucosylation of the precursor chain, ~6% in Caucasians) (55). Polymorphism in *FUT* genes could therefore explain the interindividual variability observed in our study for the BGH and LeA response after milk intake and the presence of three clear groups of subjects: postprandial increase in BGH (fucosylated galactose), postprandial increase in LeA (fucosylated N-acetylglucosamine), or no increase.

The biological benefits of milk oligosaccharides and fucosylation have been extensively investigated, in particular regarding the benefits of human breast milk for newborns (56–58). Fucosylated oligosaccharides at the intestinal surface are used as targets by various pathogenic bacteria and viruses to adhere to the epithelium (59–61). The high concentrations of free LeA in human milk might thus act as an analog to inhibit the binding of pathogens in newborns (62). The link between Lewis antigen phenotype and susceptibility to various gut infections has been demonstrated in children and adults. In particular, minor alleles of *FUT3* polymorphisms appear to increase susceptibility to ulcerative colitis (63), whereas minor alleles of *FUT2* polymorphism are associated with increased susceptibility to celiac disease (64), as well as increased risks of *Escherichia coli* (65) and *Staphylococcus aureus* (66) infections. On the other hand, the same *FUT2* polymorphisms are associated with resistance to norovirus (67–69) and rotavirus (70) infections and reduced risk of diarrhea (71). Furthermore, the secretor/nonsecretor status has been associated with susceptibility to various diseases such as type 1 diabetes (72), asthma (73, 74), and cardiovascular diseases (75). Finally, intestinal fucosylated glycan structures can be used as a substrate by specific colonic bacteria such as *Bifidobacterium*

species or *Ruminococcus*, *Clostridium*, and *Akkermansia* genera (76, 77). Consequently, the secretor/nonsecretor status has a key influence on the microbiota composition, in particular on the *Bifidobacteria* population (78, 79), an effect that could be modulated by diet (80).

The limited sample size of our study did not allow for powerful statistical comparison of the three subject groups observed here; however, the clear postprandial kinetics with five consecutive time points and the dichotomic behavior of the two metabolites, which is in line with current knowledge on the biochemical pathways and genetic polymorphisms leading to the production of Lewis A antigens, underline the plausibility of our findings and encourage further investigations in independent larger studies.

In conclusion, by extending the analysis of the serum and urine metabolomes using LC-MS, we could identify new candidate BFIs for milk, cheese, and soy. Among the identified BFIs, only two (proline and galactonic/gluconic acid) were previously identified using GC-MS and NMR, highlighting the relevance of multiplatform metabolomics. However, the validation of BFIs is a complex process and requires testing their plausibility, dose–response, time–response, robustness, reliability, stability, analytical performance, and interlaboratory reproducibility (14). The presence of LeA and BGH in serum of adults after milk intake is of interest as such metabolites have mainly been discussed in the context of maternal milk and newborns. However, the link between the presence of LeA and BGH and the subjects' secretor or nonsecretor status still has to be confirmed. Nevertheless, considering the extensive effects on health related to the presence or absence of such antigenic motifs, it would be of interest to investigate if secretors benefit from drinking milk differently from nonsecretor subjects.

Acknowledgments

We thank Doreen Gille, Alexandra Schmid, Kathryn Burton, and Barbara Walther for their assistance in designing and in conducting the clinical trial; Edith Beutler for the preparation of standardized meals and Pascal Fuchsmann for refractive index analyses (all Agroscope); as well as Ismael Turk, MD, for the organization and conduct of the clinical trial and Marie-Jeanne Voirol and Alexia Spoerl (all from Lausanne University Hospital) for contributions to laboratory work and sample collection. The authors' responsibilities were as follows—LHM, NV, FPP, and GV: designed and conducted the clinical study; DB, RP, and GP: performed the LC-MS analysis of urine, serum samples, and standards; GP: analyzed LC-MS data, performed the statistical analysis and wrote the article; GP and GV: had primary responsibility for final content; and all authors: read and approved the final manuscript.

References

- Bingham S, Luben R, Welch A, Tasevska N, Wareham N, Khaw KT. Epidemiologic assessment of sugars consumption using biomarkers: comparisons of obese and nonobese individuals in the European Prospective Investigation of Cancer Norfolk. *Cancer Epidemiol Biomarkers Prev* 2007;16:1651–4.
- Gormley IC, Bai Y, Brennan L. Combining biomarker and self-reported dietary intake data: a review of the state of the art and an exposition of concepts. *Stat Methods Med Res* 2019;29:617–35.
- Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Rappaport SM, van der Hooft JJJ, Wishart DS. The food metabolome: a window over dietary exposure. *Am J Clin Nutr* 2014;99:1286–308.

4. The Food Biomarker Alliance. The FOODBALL portal [Internet]. [Cited 2019 Jan]. Available from: <http://foodmetabolome.org>.
5. Dragsted LO, Gao Q, Pratico G, Manach C, Wishart DS, Scalbert A, Feskens EJM. Dietary and health biomarkers—time for an update. *Genes Nutr* 2017;12:24.
6. Brouwer-Brolsma EM, Brennan L, Drevon CA, van Kranen H, Manach C, Dragsted LO, Roche HM, Andres-Lacueva C, Bakker SJL, Bouwman J, et al. Combining traditional dietary assessment methods with novel metabolomics techniques: present efforts by the Food Biomarker Alliance. *Proc Nutr Soc* 2017;76:619–27.
7. Landberg R, Hanhineva K, Tuohy K, Garcia-Aloy M, Biskup I, Llorach R, Yin X, Brennan L, Kolehmainen M. Biomarkers of cereal food intake. *Genes Nutr* 2019;14:28.
8. Ulaszewska M, Vázquez-Manjarrez N, Garcia-Aloy M, Llorach R, Mattivi F, Dragsted LO, Pratico G, Manach C. Food intake biomarkers for apple, pear, and stone fruit. *Genes Nutr* 2018;13:29.
9. Vázquez-Manjarrez N, Weinert CH, Ulaszewska MM, Mack CI, Micheau P, Pétéra M, Durand S, Pujos-Guillor E, Egert B, Mattivi F, et al. Discovery and validation of banana intake biomarkers using untargeted metabolomics in human intervention and cross-sectional studies. *J Nutr* 2019;149(10):1701–13.
10. Cuparencu C, Rinnan Å, Dragsted LO. Combined markers to assess meat intake-human metabolomic studies of discovery and validation. *Mol Nutr Food Res* 2019;63:e1900106.
11. S C Sri Harsha P, Abdul Wahab R, Cuparencu C, Dragsted LO, Brennan L. A metabolomics approach to the identification of urinary biomarkers of pea intake. *Nutrients* 2018;10(12):pii: E1911.
12. Münger LH, Trimigno A, Picone G, Freiburghaus C, Pimentel G, Burton KJ, Pralong FP, Vionnet N, Capozzi F, Badertscher R, et al. 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;16:3321–35.
13. 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(2):pii: E26.
14. Dragsted LO, Gao Q, Scalbert A, Vergères G, Kolehmainen M, Manach C, Brennan L, Afman LA, Wishart DS, Andres Lacueva C, et al. Validation of biomarkers of food intake—critical assessment of candidate biomarkers. *Genes Nutr* 2018;13:14.
15. Münger LH, Garcia-Aloy M, Vázquez-Fresno R, Gille D, Rosana ARR, Passerini A, Soria-Florido M-T, Pimentel G, Sajed T, Wishart DS, et al. Biomarker of food intake for assessing the consumption of dairy and egg products. *Genes Nutr* 2018;13:26.
16. Fiamoncini J, Rundle M, Gibbons H, Thomas EL, Geillinger-Kästle K, Bunzel D, Trezzi J-P, Kiselova-Kaneva Y, Wopereis S, Wahrheit J, et al. Plasma metabolome analysis identifies distinct human metabolotypes in the postprandial state with different susceptibility to weight loss-mediated metabolic improvements. *FASEB J* 2018;32:5447–58.
17. Huber M, Knottnerus JA, Green L, Horst Hvd, Jadad AR, Kromhout D, Leonard B, Lorig K, Loureiro MI, Meer JWMvd, et al. How should we define health? *BMJ* 2011;343:d4163.
18. Pimentel G, Burton KJ, Pralong FP, Vionnet N, Portmann R, Vergères G. The postprandial metabolome—a source of Nutritional Biomarkers of Health. *Curr Opin Food Sci* 2017;16:67–73.
19. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, Ben-Yacov O, Lador D, Avnit-Sagi T, Lotan-Pompan M, et al. Personalized nutrition by prediction of glycemic responses. *Cell* 2015;163:1079–94.
20. Pimentel G, Burton KJ, Rosikiewicz M, Freiburghaus C, von Ah U, Münger LH, Pralong FP, Vionnet N, Greub G, Badertscher R, et al. Blood lactose after dairy product intake in healthy men. *Br J Nutr* 2017;118:1070–7.
21. Vionnet N, Münger LH, Freiburghaus C, Burton KJ, Pimentel G, Pralong FP, Badertscher R, Vergères G. Assessment of lactase activity in humans by measurement of galactitol and galactonate in serum and urine after milk intake. *Am J Clin Nutr* 2019;109:470–7.
22. Stanstrup J, Rasmussen JE, Ritz C, Holmer-Jensen J, Hermansen K, Dragsted LO. Intakes of whey protein hydrolysate and whole whey proteins are discriminated by LC-MS metabolomics. *Metabolomics* 2014;10:719–36.
23. Anesi A, Mena P, Bub A, Ulaszewska M, Del Rio D, Kulling SE, Mattivi F. Quantification of urinary phenyl-gamma-valerolactones and related valeric acids in human urine on consumption of apples. *Metabolites* 2019;9:254.
24. Noguchi K, Gel YR, Brunner E, Konietzschke F. nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments. *J Stat Softw* 2012;50:1–23.
25. Eriksson L, Trygg J, Wold S. CV-ANOVA for significance testing of PLS and OPLS[®] models. *J Chemometrics* 2008;22:594–600.
26. Triba MN, Le Moyec L, Amathieu R, Goossens C, Bouchemal N, Nahon P, Rutledge DN, Savarin P. PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Molecular bioSystems* 2015;11:13–9.
27. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res* 2018;46:D608–17.
28. MassBank of North America. [Internet]. [Cited 2019 Jan]. Available from: <http://mona.fiehnlab.ucdavis.edu>.
29. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW-M, Fiehn O, Goodacre R, Griffin JL, et al. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 2007;3:211–21.
30. Gibbons H, Michielsen CJR, Rundle M, Frost G, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. Demonstration of the utility of biomarkers for dietary intake assessment; proline betaine as an example. *Mol Nutr Food Res* 2017;61:1700037.
31. Lloyd AJ, Beckmann M, Fave G, Mathers JC, Draper J. Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. *Br J Nutr* 2011;106:812–24.
32. Pimentel G, Burton KJ, von Ah U, Bütikofer U, Pralong FP, Vionnet N, Portmann R, Vergères G. Metabolic footprinting of fermented milk consumption in serum of healthy men. *J Nutr* 2018;148:851–60.
33. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 2000;28:27–30.
34. Mozzi F, Ortiz ME, Bleckwedel J, De Vuyst L, Pescuma M. Metabolomics as a tool for the comprehensive understanding of fermented and functional foods with lactic acid bacteria. *Food Res Int* 2013;54:1152–61.
35. Nangung H-J, Park H-J, Cho IH, Choi H-K, Kwon D-Y, Shim S-M, Kim Y-S. Metabolite profiling of doenjang, fermented soybean paste, during fermentation. *J Sci Food Agric* 2010;90:1926–35.
36. Ross AB, Svelander C, Undeland I, Pinto R, Sandberg A-S. Herring and beef meals lead to differences in plasma 2-amino adipic acid, β -alanine, 4-hydroxyproline, cetoleic acid, and docosaheptaenoic acid concentrations in overweight men. *J Nutr* 2015;145:2456–63.
37. Teixeira LdL, Dörr F, Dias CTS, Pinto E, Lajolo FM, Villas-Bôas SG, Hassimotto NMA. Human urine metabolomic signature after ingestion of polyphenol-rich juice of purple grumixama (*Eugenia brasiliensis* Lam.). *Food Res Int* 2019;120:544–52.
38. Curis E, Crenn P, Cynober L. Citrulline and the gut. *Curr Opin Clin Nutr Metab Care* 2007;10:620–6.
39. Crenn P, Messing B, Cynober L. Citrulline as a biomarker of intestinal failure due to enterocyte mass reduction. *Clin Nutr* 2008;27:328–39.
40. D’Incecco P, Gatti M, Hogenboom JA, Bottari B, Rosi V, Neviani E, Pellegrino L. Lysozyme affects the microbial catabolism of free arginine in raw-milk hard cheeses. *Food Microbiol* 2016;57:16–22.
41. Fang F, Zhang J, Zhou J, Zhou Z, Li T, Lu L, Zeng W, Du G, Chen J. Accumulation of citrulline by microbial arginine metabolism during alcoholic fermentation of soy sauce. *J Agric Food Chem* 2018;66:2108–13.
42. Hwang H, Lee J-H. Characterization of arginine catabolism by lactic acid bacteria isolated from kimchi. *Molecules* 2018;23:pii: E3049.
43. Boutrou R, Gaudichon C, Dupont D, Jardin J, Airinei G, Marsset-Baglieri A, Benamouzig R, Tomé D, Leonil J. Sequential release of milk protein-derived bioactive peptides in the jejunum in healthy humans. *Am J Clin Nutr* 2013;97:1314–23.
44. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 2019;47:D506–15.

45. Khaldi N, Holton TA, Shields DC. Amino acid enrichment and compositional changes among mammalian milk proteins and the resulting nutritional consequences. *J Dairy Sci* 2014;97:1248–58.
46. Bütikofer U, Fuchs D. Development of free amino acids in Appenzeller, Emmentaler, Gruyère, Raclette, Sbrinz and Tilsiter cheese. *Le Lait* 1997;77:91–100.
47. Gummalla S, Broadbent JR. Tryptophan Catabolism by *Lactobacillus casei* and *Lactobacillus helveticus* Cheese Flavor Adjuncts. *J Dairy Sci* 1999;82:2070–7.
48. Aragozzini F, Ferrari A, Pacini N, Gualandris R. Indole-3-lactic acid as a tryptophan metabolite produced by *Bifidobacterium* spp. *Appl Environ Microbiol* 1979;38:544–6.
49. Zelante T, Iannitti RG, Cunha C, De Luca A, Giovannini G, Pieraccini G, Zecchi R, D'Angelo C, Massi-Benedetti C, Fallarino F, et al. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* 2013;39:372–85.
50. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, Albanes D, Gu F, Kontto J, Lassale C, et al. Identifying biomarkers of dietary patterns by using metabolomics. *Am J Clin Nutr* 2017;105:450–65.
51. Gopal PK, Gill HS. Oligosaccharides and glycoconjugates in bovine milk and colostrum. *Br J Nutr* 2000;84:69–74.
52. Tao N, DePeters EJ, Freeman S, German JB, Grimm R, Lebrilla CB. Bovine milk glycome. *J Dairy Sci* 2008;91:3768–78.
53. Aldredge DL, Geronimo MR, Hua S, Nwosu CC, Lebrilla CB, Barile D. Annotation and structural elucidation of bovine milk oligosaccharides and determination of novel fucosylated structures. *Glycobiology* 2013;23:664–76.
54. de Mattos LC. Structural diversity and biological importance of ABO, H, Lewis and secretor histo-blood group carbohydrates. *Rev Bras Hematol Hemoter* 2016;38:331–40.
55. Henry S, Oriol R, Samuelsson B. Lewis histo-blood group system and associated secretory phenotypes. *Vox Sang* 1995;69:166–82.
56. Kunz C, Rudloff S, Baier W, Klein N, Strobel S. Oligosaccharides in human milk: structural, functional, and metabolic aspects. *Annu Rev Nutr* 2000;20:699–722.
57. Korpela K, Salonen A, Hickman B, Kunz C, Sprenger N, Kukkonen K, Savilahti E, Kuitunen M, de Vos WM. Fucosylated oligosaccharides in mother's milk alleviate the effects of caesarean birth on infant gut microbiota. *Sci Rep* 2018;8:13757.
58. Ruiz-Palacios GM, Cervantes LE, Ramos P, Chavez-Munguia B, Newburg DS. *Campylobacter jejuni* binds intestinal H(O) antigen (Fuc alpha 1, 2Gal beta 1, 4GlcNAc), and fucosyloligosaccharides of human milk inhibit its binding and infection. *J Biol Chem* 2003;278:14112–20.
59. Borén T, Falk P, Roth KA, Larson G, Normark S. Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. *Science* 1993;262:1892–5.
60. Henry SM. Molecular diversity in the biosynthesis of GI tract glycoconjugates: a blood-group-related chart of microorganism receptors. *Transfus Clin Biol* 2001;8:226–30.
61. Le Pendu J. Histo-blood group antigen and human milk oligosaccharides: genetic polymorphism and risk of infectious diseases. *Adv Exp Med Biol* 2004;554:135–43.
62. Newburg DS. Neonatal protection by an innate immune system of human milk consisting of oligosaccharides and glycans. *J Anim Sci* 2009;87:26–34.
63. Hu D, Zhang D, Zheng S, Guo M, Lin X, Jiang Y. Association of ulcerative colitis with FUT2 and FUT3 polymorphisms in patients from Southeast China. *PLoS One* 2016;11:e0146557.
64. Parmar AS, Alakulppi N, Paavola-Sakki P, Kurppa K, Halme L, Färkkilä M, Turunen U, Lappalainen M, Kontula K, Kaukinen K, et al. Association study of FUT2 (rs601338) with celiac disease and inflammatory bowel disease in the Finnish population. *Tissue Antigens* 2012;80:488–93.
65. Stapleton A, Nudelman E, Clausen H, Hakomori S, Stamm WE. Binding of uropathogenic *Escherichia coli* R45 to glycolipids extracted from vaginal epithelial cells is dependent on histo-blood group secretor status. *J Clin Invest* 1992;90:965–72.
66. Nurjadi D, Lependu J, Kremsner PG, Zanger P. *Staphylococcus aureus* throat carriage is associated with ABO-/secretor status. *J Infect* 2012;65:310–7.
67. Carlsson B, Kindberg E, Buesa J, Rydell GE, Lidón MF, Montava R, Abu Mallouh R, Grahn A, Rodríguez-Díaz J, Bellido J, et al. The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 Norovirus infection. *PLoS One* 2009;4:e5593.
68. Kindberg E, Akerlind B, Johnsen C, Knudsen JD, Heltberg O, Larson G, Böttiger B, Svensson L. Host genetic resistance to symptomatic norovirus (GII.4) infections in Denmark. *J Clin Microbiol* 2007;45:2720–2.
69. Thorven M, Grahn A, Hedlund K-O, Johansson H, Wahlfrid C, Larson G, Svensson L. A homozygous nonsense mutation (428G→A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GII) infections. *J Virol* 2005;79:15351–5.
70. Imbert-Marcille B-M, Barbé L, Dupé M, Le Moullac-Vaidye B, Besse B, Peltier C, Ruvoën-Clouet N, Le Pendu J. A FUT2 gene common polymorphism determines resistance to rotavirus A of the P[8] genotype. *J Infect Dis* 2014;209:1227–30.
71. Bustamante M, Standl M, Bassat Q, Vilor-Tejedor N, Medina-Gomez C, Bonilla C, Ahluwalia TS, Bacelis J, Bradfield JP, Tiesler CMT, et al. A genome-wide association meta-analysis of diarrhoeal disease in young children identifies FUT2 locus and provides plausible biological pathways. *Hum Mol Genet* 2016;25:4127–42.
72. Ihara K, Fukano C, Ayabe T, Fukami M, Ogata T, Kawamura T, Urakami T, Kikuchi N, Yokota I, Takemoto K, et al. FUT2 non-secretor status is associated with type 1 diabetes susceptibility in Japanese children. *Diabet Med* 2017;34:586–9.
73. Innes AL, McGrath KW, Dougherty RH, McCulloch CE, Woodruff PG, Seibold MA, Okamoto KS, Ingmundson KJ, Solon MC, Carrington SD, et al. The H antigen at epithelial surfaces is associated with susceptibility to asthma exacerbation. *Am J Respir Crit Care Med* 2011;183:189–94.
74. Ronchetti F, Villa MP, Ronchetti R, Bonci E, Latini L, Pascone R, Bottini N, Gloria-Bottini F. ABO/Secretor genetic complex and susceptibility to asthma in childhood. *Eur Respir J* 2001;17:1236–8.
75. Hein HO, Sørensen H, Suadicani P, Gynzelberg F. The Lewis blood group—a new genetic marker of ischaemic heart disease. *J Intern Med* 1992;232:481–7.
76. Hoskins LC, Boulding ET. Degradation of blood group antigens in human colon ecosystems. II. A gene interaction in man that affects the fecal population density of certain enteric bacteria. *J Clin Invest* 1976;57:74–82.
77. van Passel MWJ, Kant R, Zoetendal EG, Plugge CM, Derrien M, Malfatti SA, Chain PSG, Woyke T, Palva A, de Vos WM, et al. The genome of *Akkermansia muciniphila*, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. *PLoS One* 2011;6:e16876.
78. Wacklin P, Mäkituokko H, Alakulppi N, Nikkilä J, Tenkanen H, Rabinä J, Partanen J, Aranko K, Mättö J. Secretor genotype (FUT2 gene) is strongly associated with the composition of *Bifidobacteria* in the human intestine. *PLoS One* 2011;6:e20113.
79. Wacklin P, Tuimala J, Nikkilä J, Sebastian Tims N, Mäkituokko H, Alakulppi N, Laine P, Rajlic-Stojanovic M, Paulin L, de Vos WM, et al. Faecal microbiota composition in adults is associated with the FUT2 gene determining the secretor status. *PLoS One* 2014;9:e94863.
80. Kashyap PC, Marcobal A, Ursell LK, Smits SA, Sonnenburg ED, Costello EK, Higinbottom SK, Domino SE, Holmes SP, Relman DA, et al. Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota. *Proc Natl Acad Sci* 2013;110:17059–64.