Journal of Breath Research

PURPOSE-LED PUBLISHING™

PAPER • OPEN ACCESS

Breath response following a nutritional challenge monitored by secondary electrospray ionization high-resolution mass spectrometry

To cite this article: Cedric Wüthrich *et al* 2022 *J. Breath Res.* **16** 046007

Journal of Breath Research

PAPER

Breath response following a nutritional challenge monitored by secondary electrospray ionization high-resolution mass spectrometry

Cedric Wüthrich[1], Miguel de Figueiredo[2], Kathryn Jane Burton-Pimentel[3], Guy Vergères[3], Fabian Wahl[3], Renato Zenobi[1,*] and Stamatios Giannoukos[1,*]

[1] Department of Chemistry and Applied Biosciences, ETHZ, Zurich, Switzerland
[2] School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland
[3] Food Microbial Systems Research Division, Agroscope, Bern, Switzerland
[*] Authors to whom any correspondence should be addressed.

E-mail: renato.zenobi@org.chem.ethz.ch **and** stamatios.giannoukos@org.chem.ethz.ch

Keywords: SESI, high-resolution MS, nutritional research, breath metabolomics, on-line sampling, CID

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
23 June 2022

REVISED
2 August 2022

ACCEPTED FOR PUBLICATION
12 August 2022

PUBLISHED
26 August 2022

Abstract

On-line breath analysis using secondary electrospray ionization coupled to high-resolution mass spectrometry (SESI-HRMS) is a sensitive method for biomarker discovery. The strengths of this technology have already been demonstrated in the clinical environment. For the first time, this study demonstrates the application of SESI-HRMS in the field of nutritional science using a standardized nutritional intervention, consisting of a high-energy shake (950 kcal, 8% protein, 35% sugar and 57% fat). Eleven subjects underwent the intervention on three separate days and their exhaled breath was monitored up to six hours postprandially. In addition, sampling was performed during equivalent fasting conditions for selected subjects. To estimate the impact of inter- and intra-individual variability, analysis of variance simultaneous component analysis was conducted, revealing that the inter-individual variability accounted for 30% of the data variation. To distinguish the effect of the intervention from fasting conditions, partial least squares discriminant analysis was performed. Candidate compound annotation was performed with pathway analysis and collision-induced dissociation (CID) experiments. Pathway analysis highlighted, among others, features associated with the metabolism of linoleate, butanoate and amino sugars. Tentative compounds annotated through CID measurements include fatty acids, amino acids, and amino acid derivatives, some of them likely derived from nutrients by the gut microbiome (e.g. propanoate, indoles), as well as organic acids from the Krebs cycle. Time-series clustering showed an overlap of observed kinetic trends with those reported previously in blood plasma.

1. Introduction

Nutrition is a major factor contributing to human health and consequently understanding the interaction of dietary intake with human metabolism is an important focus of life sciences research. Liquid chromatography–mass spectrometry (LC–MS) is one of the main analytical methods for identifying and monitoring biomarkers in body fluids, stemming from various types of foods [1–10]. While such biomarkers are mostly related to the intake of specific types of foods, there have been studies focusing on biomarkers that reflect the response of human metabolism to diet and the effect of diet on obesity, diabetes, nutritional disorders and metabolic flexibility/metabolic health [11–15]. These studies often use nutritional challenges in the form of high-calorie shakes to measure nutrition-specific metabolic molecules. While most of these studies are based on blood and urine sample collection and analysis, research utilizing off-line breath analysis (e.g. using sampling bags, adsorption tubes, or

exhaled breath condensate) has been published tracking intervention-related volatile organic compounds (VOCs) [11, 16].

Despite minimal use of breath sampling in nutritional sciences to-date, the non-invasive nature of the sampling and the associated possibility of high sampling frequency naturally lends itself to the field. As a direct on-line sampling technique, secondary electrospray ionization (SESI) coupled to high-resolution mass spectrometry (HRMS) could enable nutritional research to be conducted on exhaled breath samples with high sensitivity and frequency in real-time [17]. There is no precedent of utilizing SESI-MS for on-line breath analysis in nutritional studies, though proton transfer reaction-mass spectrometry (PTR-MS) has been used: Hageman and co-workers used PTR-MS for following the metabolic response on exhaled breath upon the consumption of a low and high-fat dairy drink [18]. Since SESI offers greater sensitivity and better coverage of the mass region above 100/mz (allowing the analysis of polar high-molecular-weight compounds) compared to PTR-MS [19], on-line breath analysis using SESI-HRMS could be a valuable complementary method for nutritional breath metabolomics investigations. For this purpose, questions about sources of variation, the typical postprandial breath metabolomics profile observed by SESI-HRMS and metabolites′ kinetics following dietary exposures should be considered.

This report uses a standardized meal in the form of a high-energy shake to address these questions. The shake used in our study was developed to assess the phenotypic resilience/flexibility of the postprandial response to a nutritional challenge and triggers 26 metabolic processes in more than seven organs, including among others the gut, adipose tissue, vasculature, muscle, liver, kidney, and pancreas [13]. The dynamic response to this intervention is a measure of metabolic health and systemic stress. Due to the wide impact on the human metabolome, the use of such an intervention in the context of nutritional breath research is highly relevant. This study therefore utilizes SESI-HRMS to characterize, for the first time, the chemical composition of the human postprandial breath metabolome in response to the ingestion of a standardized nutritional shake and compares it to the fasting state during the same time period. It investigates the intra- and inter-individual breath metabolome variability and monitors the kinetic response of features of nutritional interest over a postprandial period of six hours. It highlights metabolic pathways associated with the macronutrients involved and provides candidate compound annotation via tandem MS experiments.

2. Methodology

2.1. Study subjects and nutritional intervention

For this study, 11 subjects (three female, eight male) were recruited. Previous nutritional studies with similar group sizes have demonstrated the discovery of thousands of metabolites with statistical significance responding to different interventions [20, 21]. The average age of the subjects was 29 $\pm$ 3 years and the average body mass index was 24.2 $\pm$ 2.8 kg m^{-2}. All subjects were non-smokers and did not suffer from chronic diseases, food intolerances, or food allergies. Subjects underwent a nutritional challenge consisting of a shake adapted from Wopereis *et al* [13]. It consisted of water (tap water, 320 ml), dextrose (84 g, Lee Sports GmbH, CH), protein powder (20 g, Protifar® Nutricia, NL), sunflower oil (60 g, Migros Bio, CH) and vanilla aroma (20 droplets, Lee Sports GmbH, CH), providing a 950 kcal shake with the sunflower oil accounting for 57%, dextrose for 35% and the protein powder accounting for 8% of the total calorie count.

The nutritional challenge was tested with a day interval between each test for each subject to allow the investigation and evaluation of the intra- and inter-individual variability of their postprandial breath metabolome. This resulted in three measurement days in total per subject (figure 1(a)).

On the day preceding the intervention, subjects were asked to refrain from drinking alcoholic beverages, eating spices, allium vegetables, and to consume the last meal no later than 12 h before the start of the intervention. Before arriving at the test site, subjects followed their usual morning routine with the mentioned restrictions. Upon arrival in the test laboratory, the subjects delivered fasting breath samples for baseline measures, approximately 15 min before the intervention. In addition, the subjects filled out a questionnaire about their diet during the day before their participation. For each of the subjects, the standardized shake was consumed within 5 min. Following the ingestion of the shake, subjects rinsed their mouth with a standard quantity of water (1 l) to reduce the detection of shake-relevant residual molecules in the oral cavity. Postprandial breath samples were collected at predefined times after consumption: 15, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330, and 360 min, resulting in measurements over six hours after consumption. To normalize subjects' hydration, the subjects were allowed to drink water ab libitum during the post-ingestion period. No other foods or fluids were permitted for consumption during the laboratory testing. During the experiment period, subjects remained in close proximity to the instrumentation and were asked to minimize physical activity.

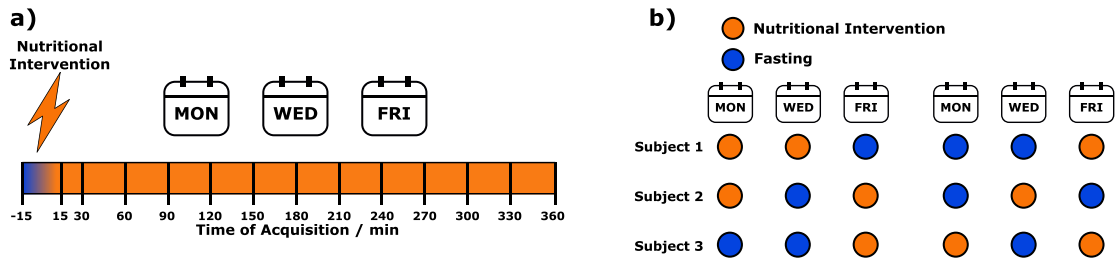

Figure 1. (a) Experimental design of the nutritional intervention and subsequent measurements. A baseline measurement designated at −15 min was recorded for each subject before the intervention. Each subject underwent this experiment on three days separated by a washout day (n = 11). (b) Design of a cross-over experiment to distinguish the response to the intervention from fasting. Each of the subjects who participated in this investigation were randomly assigned to complete the fasting and nutritional intervention on six different days.

The impact of fasting conditions on the observed metabolites was also explored to confirm that the observed metabolomic changes were due to the nutritional challenge. Thus, a further experiment was performed by a subgroup of three subjects to compare shake and no-shake consumption in a randomized crossover design, with each condition repeated three times (with a 1 d interval between each test). The same experimental conditions and sampling protocol were applied for the fasting test day with the sole difference being that the shake was not consumed. This resulted in six additional measurement days in total per participant (figure 1(b)).

The study was conducted in accordance with the Declaration of Helsinki and was approved by the ethics committee of the ETH Zurich (EK-2021 N-45). All subjects were provided with written information on the study design prior to inclusion in the investigation and provided (written) consent for participation.

2.2. SESI-HRMS experimental setup

On-line chemical analysis of exhaled breath was carried out using a commercial SESI source (Fossil Ion Tech, Spain) attached to a Q-Exactive Plus Orbitrap (Thermo Fischer, Germany). The MS vacuum was achieved and maintained in stable low-pressure values. As a forepump, a single-stage rotary pump (SOGEVAC Sv 40 BI, Oerlikon Leybold Vacuum) was used. A turbomolecular pump (SplitFlow 310, Pfeiffer Vacuum) created a stable vacuum in the C-trap, while the ultra-high vacuum in the analyzer chamber was obtained by a TURBOVAC TW 290/20/20-UHV (Oerlikon Leybold Vacuum). The operating pressure during experiments was maintained stable at 8×10^{-11} mbar. Pressure, breath flow rate and breath volume were simultaneously recorded using a flow meter (EXHALION, Fossil Ion Tech, Spain), connected at the front-end part of the sample transfer line. Both the positive and the negative ion production and detection modes were employed and for each ionization mode, three full exhalations (duration ∼20 s and volume ∼3 l per exhalation) were recorded per ionization mode, participant and time point; resulting in six exhalations per measurement time. During breath sampling, subjects exhaled through a spirometry filter (Vyaire Medical, Germany) and a heated sampling line at 130 °C into the ionization chamber kept at 90 °C constantly. The electrospray solution was a 0.1% aqueous formic acid solution passing through a nano-electrospray capillary (inner diameter = 20 μm, outer diameter = 365 μm, Fossil Ion Tech, Spain) with an overpressure of 0.8 bar. The water used for the preparation of the electrospray solution was LC-MS grade purchased from *Fisher Scientific* and the formic acid (99.5% purity) was obtained from *VWR Chemicals*. The electrospray solution was interacting horizontally with breath samples for charge transfer prior to introduction in the MS for analysis. Sheath and auxiliary gas values were adjusted to 15 psi and 2 a.u. respectively. The electrospray solution voltage was set to ±3.5 kV (depending on the ionization mode used) and the MS inlet capillary was heated to 250 °C. For the C-trap, the automatic gain control target was adjusted to 10^6 and the maximum injection time to 500 ms. The mass resolution of the Orbitrap was set to 140 000. The RF frequency and drive voltage for the C-trap and Obritrap used during our experiments were the standard settings provided by Thermo Fischer. Specifically, the voltage of the central electrode of the Q-Exactive Plus Orbitrap was 5 kV and the RF frequency was 1.5 Hz for a 140 000 resolution. For the C-trap, the drive voltage was 1 kV. To reduce ion competition within the C-trap, a spectral stitching technique was employed [22]. Specifically, during the first three exhalations, SESI-HRMS operated in the positive ion mode with the first scan covering the m/z range of 50–500 and the next ones the m/z ranges of 50–132, 132–195, 195–277 and 277–500. During the last three exhalations in the negative mode, it scanned the mass ranges m/z 50–500, and then a scan from 50 to 90, 90 to 161, 161 to 255 and 255 to 500. The mass windows were then stitched together using built-in-house pre-processing algorithms as described in section 2.3.

2.3. Data pre-processing

All mass spectra were processed with a custom-written Python (v3.7) script run on the Euler

cluster at the ETH Zurich. The individual raw files of the recorded measurements were converted to mzML-files with ProteoWizard [23]. The individual mass spectrometric scans were interpolated with a step of 10^{-5} from 50 to 500/mz and subsequently summed and averaged with all recorded scans over all conducted measurements to obtain a composite mass spectrum. Individual peaks were then detected in the composite mass spectrum and the peak width was determined at 90% of the peak's height. The intensity of the individual peaks was then determined in all scans through integration within the peak extends, yielding a time trace matrix for each measurement. The time traces of each measurement were subsequently aligned with the recorded exhalation patterns and if a feature had a significantly higher intensity during exhalation, the mean intensity was calculated and stored. The data pre-processing workflow is schematically described in figure S1.

2.4. Statistical analysis

Prior to statistical analyses, features were scaled to the total feature intensity within one measurement. For the assessment of various influences on the obtained data, analysis of variance simultaneous component analysis (ASCA) was conducted with the factors of the experimental design as potential sources. Briefly, ASCA separates data variance according to the experimental design factors of the measurement as well as the cross-terms and conducts principal component analysis on each variation source [24–26]. The experimental design was based on three factors: (a) day, (b) time and (c) subject. For the obtained data, the factors of each sample were the measurement day (1, 2, 3), subject (1, 2, up to 11) and the time point of measurement (−15 min (corresponds to the baseline measurement), and post-shake consumption at 15 min, 30 min, 60 min, and up to 360 min with 30 min intervals). Each feature was scaled according to its standard deviation before the application of ASCA. ASCA calculations were performed with a custom MATLAB (Version 2020b) script.

To distinguish a response to the nutritional intervention, partial least squares discriminant analysis (PLS-DA) was conducted with data obtained on the subgroup of three subjects who completed intervention and fasting test days utilizing the scikit-learn library [27]. By fitting a PLS-DA model to the data, measurements could be classified as being related to either fasting or intervention and the contributing features interpreted as being associated more with one of these states. For this purpose, features were scaled between different measurements to the 80% quantile and subsequently filtered with the condition that they contain less than 50% non-zero values in all measurements. For optimization of the latent variable number needed for the classification, the data were split into 100 iterations [training (60%) and test (40%) set]. PLS-DA was then performed with increasing latent variable numbers for each run applying *k*-folds for optimization. For performance evaluation, the mean accuracies over all runs were evaluated for each latent variable number. After choosing the number of latent variables, PLS-DA analysis was repeated with the set number. Features with a positive PLS-DA coefficient were deemed to be associated with the conducted intervention, whereas features with a negative coefficient were deemed not to be associated with the intervention. For each feature, a variable importance in the projection score (VIP-score) was calculated.

To gain an overview of average kinetic trends of the detected features, *k*-means cluster analysis was performed with the tslearn library [28] for Python, setting the number of clusters to five to be comparable to literature [13] and utilizing soft dynamic time warping as a metric (figure S2).

2.5. Annotation of metabolites

To obtain putative candidate structural information for selected features on level 3 annotation [17, 29], collision-induced dissociation (CID) experiments were conducted with the exhalations from two subjects (a subgroup of the whole study group) undergoing shake consumption on a separate experimentation day. These two subjects underwent an additional nutritional intervention and provided exhalations continuously for the duration of the respective CID-experiments. To prepare these CID-experiments, features were divided into three segments relating to the maximum intensity within the measurement time and grouped as features with early (0–2 h), middle (2–4 h), and late (4–6 h) responses, resulting in three-time intervals where subjects provided breath samples. Features whose maximum intensity was below 10^4 were excluded. The residual features were then divided into two groups depending on whether the intensity of the precursor made up at least 50% of total intensity within the 0.4/mz quadrupole isolation window or not. If a feature had a maximum intensity over the threshold (50% of the intensity within the window) direct fragmentation with the same settings as in section 2.3 and collision energies of 10, 35 and 50 was attempted. For features below the threshold, an incremental quadrupole acquisition for resolving spectra (IQAROS) sequence was set up [30]. Briefly, scans were set up with the same collision energies and a mass resolution of 17 500. For each precursor, the isolation window was moved over the precursor's *m/z* value in steps of 0.05/mz starting from −0.6/mz to +0.6/mz relative to the precursor mass. Through this modulation and linear regression, a reconstructed fragment spectrum was obtained for each precursor. These spectra were then processed with SIRIUS [31]. For molecular formula calculations, the elements C, H, O, N and S were considered to restrain the ionized species to $[M + H]^+$, $[M + H_2O + H]^+$, $[M–H_2O + H]^+$,

$[M–H]^-$ and $[M + H_2O–H]^-$. Database search for potential hits was restricted to the human metabolome and the KEGG databases [32, 33]. For pathway analysis based on MS1-data, the mummichog algorithm was used with a cut-off p-value of 10^{-3} [34]. Briefly, the mummichog algorithm is a pathway annotation tool with tolerance to technological noise and biological sampling, which predicts functional activities directly from spectral features bypassing metabolite identification. The input p and $\log_2$-fold change values were calculated through comparison with the baseline measurements.

3. Results and discussion

3.1. Impact of the experimental design factors

To assess the impact of the nutritional intervention on the postprandial human breath metabolome, the exhalations of 11 subjects were recorded over six hours with SESI-HRMS. Each subject underwent the intervention on three separate days. In total, 5083 features were detected in the positive and 4004 in the negative ion mode. Features were further filtered according to whether they were detected in at least 50% of all measurements, thus resulting in 3147 features in the positive and 1951 features in the negative ion mode.

Table S1 summarizes the variances of ASCA for each experimental factor and their interactions. Among the experimental parameters and their interactions with each other, the highest variance was explained by the individual subjects and the interaction of all three experimental parameters (both at 30%). (figure 2(a))

The first principal component of the subject contribution reveals the clustering of some subjects (figure 2(b)). Most of the clustered scores were from subjects measured in the same week (grouped within circles in figure 2(b)), therefore indicating a batch effect within the variation purely caused by subjects. The subject component could be interpreted as a measure of inter-individual variability and the day component as intra-individual variability. ASCA clearly revealed the strong difference between the pure effect of these factors, with variation between subjects being an order of magnitude larger than variation within subjects. Unfortunately, the separation of biological and instrumental variation was not possible with the experimental design of this study, as the number of measured subjects was not the same in different weeks. This means that the experimental factor of subject did not only contain information about the subject, but the week in which the subject underwent the intervention. Due to this imbalance, it was not possible to estimate with ASCA, how much of the 30% contribution to the total data variation related to a biological cause and how much to a change in instrumental conditions. Balancing the subject number within the different intervention weeks in future studies would at least enable ASCA to separate the corresponding effects. The reduction of variation contribution of week and subject factor would then be a question of standardization procedures.

The time effect itself was not responsible for most of the variation in the data, however its principal components reveal underlying longitudinal trends within all measurements (figure 2(c)). The initial three time-related principal components describe a change of measurements with the score being a measure of similarity between measurements. The first component describes a continuous change from baseline measurements towards measurements recorded at the end. The second component exhibits a significant change from baseline measurements to the ones directly after it. This sharp increase coincides with the nutritional intervention and is an indication of the intervention having an impact on the measured compounds. Component three shows the same trend as observed for component two in the early postprandial phase (0–90 min), but with a faster return towards the baseline, followed by a subsequent increase in signal intensity This suggests the reappearance of certain metabolites associated with the intervention towards the end of the experiment day.

3.2. Characterization of breath response after shake consumption

To distinguish potential breath metabolites arising due to the nutritional intervention, three subjects underwent a further cross-over intervention. They were randomly assigned to either consumed the shake or continued the overnight fasting for all six measurement days to minimize technical variation. The obtained intensity matrices after spectra preprocessing were evaluated by PLS-DA. Five latent variables were chosen to fit the data after optimization through splitting the data 100 times into training and test data sets and repetitive *k*-folds validation (figure S2).

The fitted PLS-DA model achieved a 0.96 area under the curve for the receiver operating characteristic (figure S2). Features with a mean positive PLS-DA loading coefficient were deemed to be associated with the nutritional intervention, whereas ones with a negative coefficient were seen as fasting-related compounds. Furthermore, features with VIP score of more than 1.3 and a positive coefficient were selected for structural elucidation through CID experiments. Applying the fitted PLS-DA model on the data set of all 11 subjects, a prediction accuracy of 84% for all post-intervention samples was obtained.

For the intervention-related compounds, the p-values and $\log_2$-fold changes were calculated. For the p-value calculation, a *Kruskal–Wallis* test was performed comparing the baseline measurements with the subsequent ones, log-fold changes were calculated

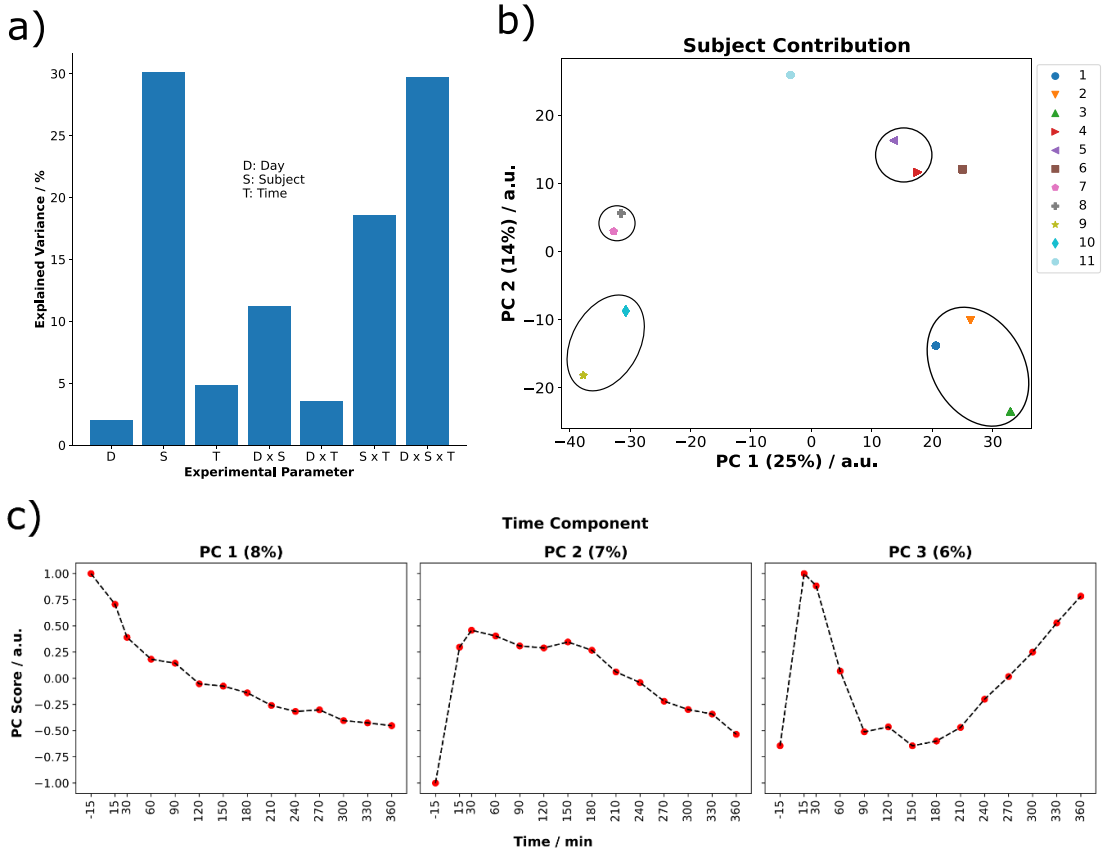

Figure 2. (a) Explained variance of the total data related to the factors of the experimental design as determined with ASCA. The highest variability was explained by the individual subjects. Other major contributions to data variation were the interaction terms of the experimental parameters with the three-term interaction being of the same magnitude as the participant factor. (b) First two principal components' scores of the subject effect. The circle contains subjects measured in the same week. (c) The first three principal components of the time effect. Each dot represents one time point with all the associated measurements. The first principal component describes a change over all measurements with the last measurements exhibiting the largest difference from the start. The second component differs strongly from the baseline measurement to the subsequent measurements with a slow return to the baseline measurement towards the end. A similar trend is observed for the third component, although measurements become more similar in value to the baseline after two hours and subsequently diverge again.

similarly, comparing the baseline with subsequent measurements [35]. The p-values were further utilized for pathway analysis. Figure 3 shows a discrimination of up- and down-regulated features following the nutritional intervention within all 11 subjects.

3.3. Pathway analysis and tentative compound annotation

To identify the metabolic pathways that are modulated by the nutritional intervention, pathway analysis was conducted applying the mummichog workflow [34] through MetaboAnalyst [36]. For this purpose, all intervention-related features (features with a positive PLS-DA coefficient in both positive and negative ionization mode) with their corresponding p-values were considered.

The mummichog algorithm annotated 74 metabolic pathways (figure 4 and table S2) with a total of 162 unique candidate metabolites linked with them. The resulting tentative pathways were expected to be intervention-related since the input for the mummichog analysis contained the features filtered with the PLS-DA loading coefficients. The comparison of the detected pathways with their classification in the KEGG pathway maps [33] supports this hypothesis as the *butanoate, amino sugar, propanoate, pyruvate, glyoxylate metabolism* and *citric acid cycle* are part of the carbohydrate, fat and protein metabolism, which was potentially modulated by the ingested shake. Pathways linked to the ingestion of fats were *linoleate* metabolism and *fatty acid oxidation*, both of which showed a large log fold change compared to the baseline measurement. The source of the fat (sunflower oil) used in the intervention can additionally be connected to the *limonene and pinene degradation* pathway since both these substances are found within the oil [37]. Pathways relating to the metabolism of most of the amino acids present the protein powder were enriched with *methionine, cysteine* and *glutamate* metabolic pathways showing the highest gamma-adjusted p values as determined by the algorithm.

An annotated compound given by the mummichog algorithm, alpha-ketoglutaric acid (AKG) was part of 18 out of the 74 annotated pathways, including *butanoate metabolism, methionine*

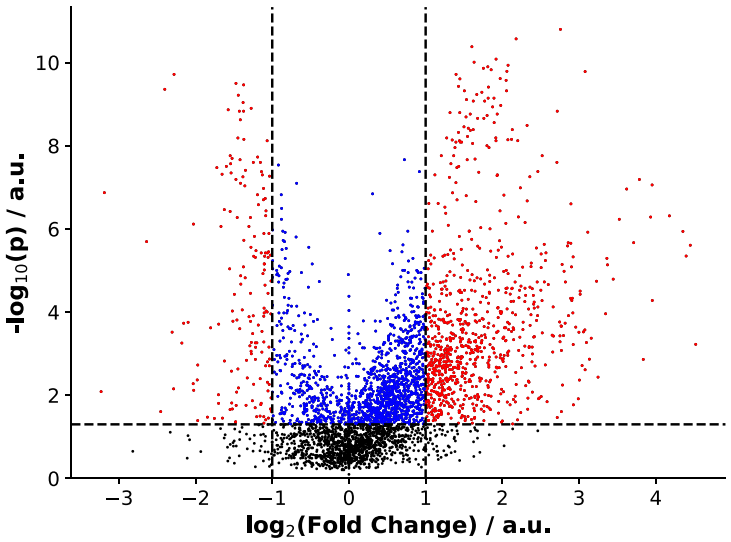

Figure 3. Volcano plot showing potential metabolites up- and down-regulated after the nutritional intervention. The $-\log_{10}(p)$-values were calculated through a Kruskal–Wallis test, $\log_2$-fold change through comparison with baseline measurements [35]. Features shown in red (n = 1011) exhibit both a significant group difference and an increase in detection compared to the baseline. Features depicted in blue show a significant difference, but not a sufficient increase, the ones in black exhibit neither.

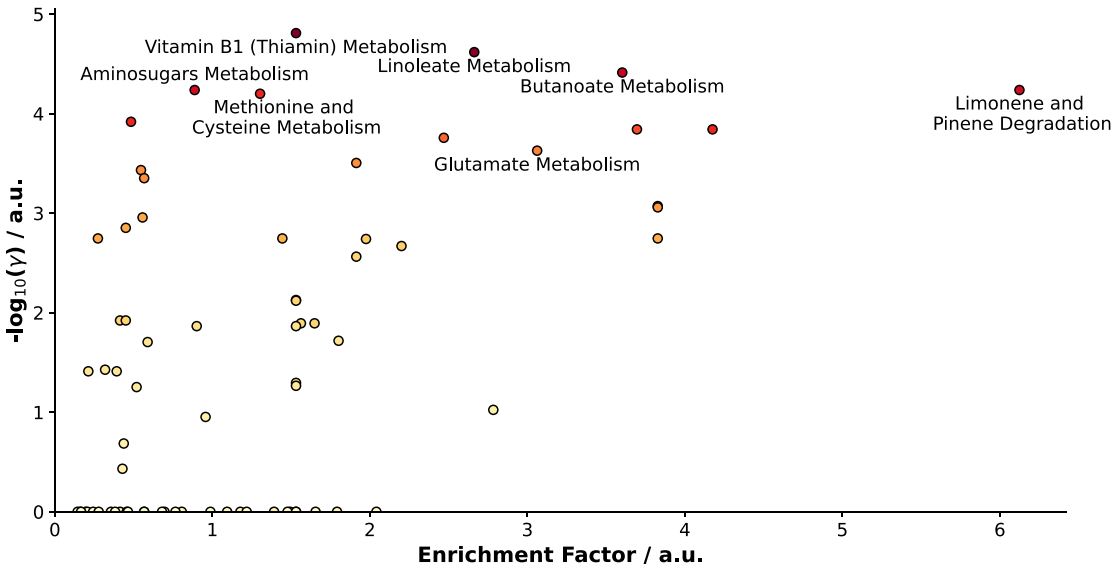

Figure 4. Mummichog enrichment analysis results as a function of the gamma-adjusted (γ) p values and enrichment factor (EF) with the MFN database serving as [36].

and *cysteine metabolism*, as well as *glutamate metabolism*. AKG has previously been detected in breath using SESI-HRMS [38] and is an interesting molecule not only due to being part of the *citric acid cycle* [39] but also serving multiple roles in the degradation of amino acids [40]. The second most frequently observed potential metabolite after AKG was acetate with 14 associated pathways, which among other endogenous sources, is a product of dietary amino acid fermentation by the gut microbiome to short-chain fatty acids (SCFAs) [41]. Among the pathways that produce acetate, the degradation of *methionine* and *cysteine* pathway obtained the highest gamma-adjusted p value in the mummichog analysis. Other pathways which included acetate were the *butanoate* and *pyruvate metabolism* which were both enriched in the analysis.

To gain additional insights into the identity of the detected features, CID experiments were conducted with two subjects, whose breath was analysed in real-time after undergoing the nutritional intervention, revealing 260 tentative annotated compounds. The precursors' ions were selected according to their VIP score (>1.3) within the PLS-DA model. As real-time analysis poses unique challenges connected to the number of different molecules

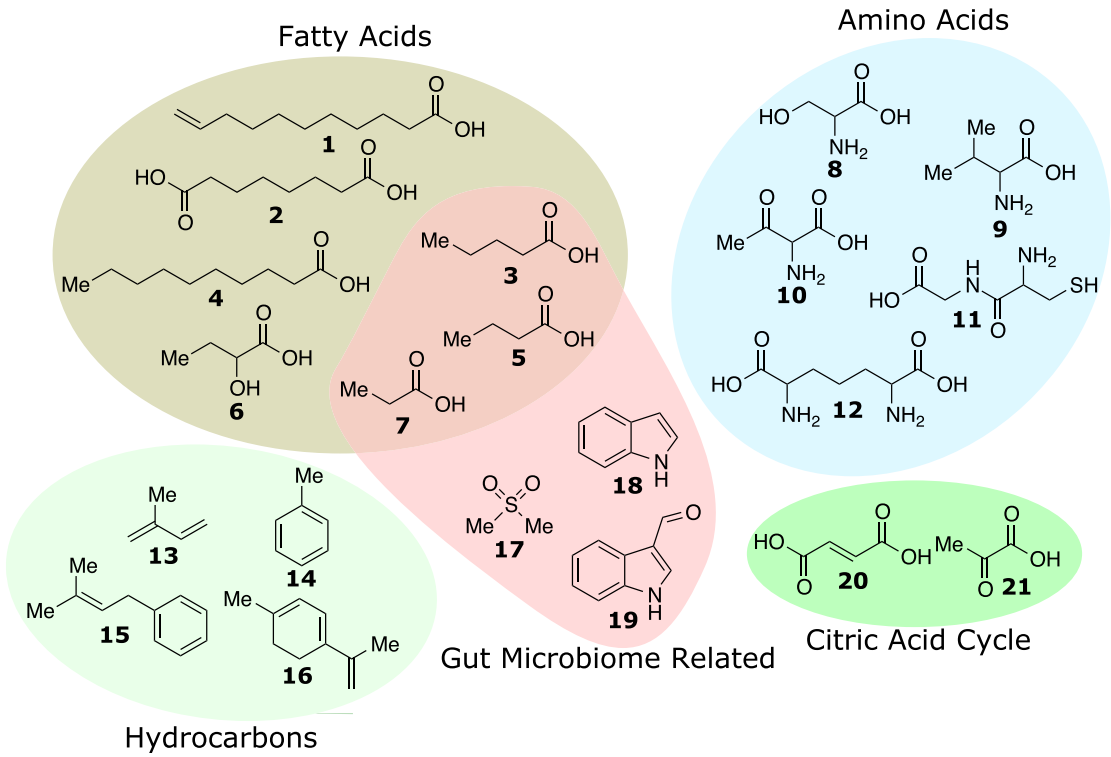

Figure 5. Selected structural candidates of intervention-related compounds as determined by PLS-DA after in-silico fragment analysis of MS^2-data ordered into different groups indicated by color. Compounds **1**–**7** are fatty acids of varying length and functionalization with propanoic acid **7** being the shortest that was annotated. The amino acids **8**–**12** represent the variety of these compounds, even two-amino acid compounds such as **11** were potentially observed. The category of hydrocarbons contains the substances **13**–**16**, representing the trend that the majority of these compounds contain six-membered rings or even aromatic moieties. The classes of gut microbiome-related (**5, 7, 17**–**19**), and citric acid cycle molecules (**20, 21**) show the effect of the intervention not only directly after it, but also late responses.

reaching the collision cell and consequently leading to chimeric fragment spectra, a recently developed methodology improving real-time MS^2-experiments called IQAROS was employed [30]. After processing the data, the reconstructed fragment spectra were processed within SIRIUS [31] to obtain structural candidates for each feature with the help of the HMDB [32] and KEGG [33] database. According to a proposed annotation confidence scheme [29], the workflow used in our experiments lead to level three structural candidates. This means that the structures obtained shown in figure 5 (CID-spectra reported in table S3) represent one possible isomer among several different ones derived from *in-silico* fragment matching.

Two hundred and sixty five structural candidates were found for the intervention-related metabolites which belong mostly to three chemical categories: fatty acids, amino acids, or hydrocarbons (figure 5). The suggested fatty acids cover the lengths from propanoic acid **7** to decanoic acid **4** and undecenoic acid **1**. Decanoic acid **4** is classified as a medium-chain fatty acid and the source was likely the sunflower oil with the acid being in the form of triacylglycerols, which are metabolised by enzymes in the saliva and stomach [42]. Both undecenoic acid **1** and octanedioic acid **2** are connected to the breakdown of longer fatty acids or have a more complicated origin. Among the candidate metabolites were also the SCFAs **3, 5** and **7**, that are likely products of the gut microbiome metabolism [43]. SCFAs are mainly produced from dietary fibers (though given the nutritional composition of the intervention this source is unlikely) or undigested sugars [44], as well as amino acids [45]. Within the structural candidates for intervention-related compounds, additional metabolites were found which relate to the gut microbiome. Dimethyl sulfone **17**, indole **18** and indole-3-carboxaldehyde **19** were among these metabolites. Dimethyl sulfone **17** has been proposed to be formed from methanethiol followed by subsequent oxidation and was confirmed to be a common substance to be found in the blood [46]. Both indole **18** and its aldehyde form **19** relate to the catabolism of tryptophan with varying bacterial strains playing a crucialpart [47–49].

Multiple amino acids structures and related metabolites were highlighted after the *in-silico* fragment comparison. These structures included serine **8** and valine **9**, as well as the 2-amino-3-oxabutanoic acid **10** and the dipeptide glycine-cysteine **11**. These compounds could be attributed to the exogenous source of amino acids from the protein powder of the nutritional intervention or could be endogenously

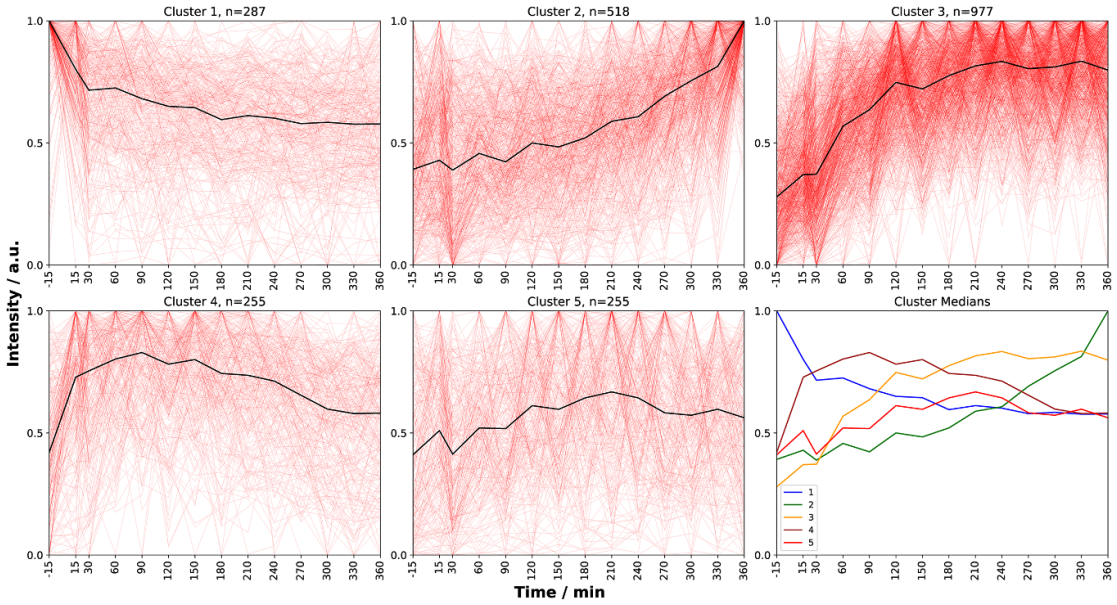

Figure 6. Five resulting clusters from K-means clustering with individual median feature traces, normalized to their maximum (in red) and the median of the cluster represented in black. Cluster 1 and 2 described a continuous decrease and increase, respectively, over the six-hour experimentation period. Most features fell into cluster 3, where an increase over two hours was observed, followed by a flattening. Cluster 4 exhibited a sharp rise after the intervention with a subsequent decrease after 90 min. In the final cluster 5, the trends showed a maximum in the period from two to four hours. The number (n) of the observed features per cluster is shown on the title of each subfigure.

produced compounds that were modulated by the intervention. On the other hand, the reason for the modulation of diaminopimelic acid **12,** which does not feature in postprandial amino acid metabolism, was unclear.

The proposed hydrocarbons (**13**–**16**) do not feature in any of the metabolic pathways typically associated with the consumed nutrients and likely relate to the components of the utilized oil [50] or may be contaminants derived from the sampling material. Isoprene **13** is a ubiquitous molecule in breath [51], similarly to toluene **14** [52] and are classified as belonging to the exposome.

In terms of glycolysis and the citric acid cycle (the major metabolic pathways of sugar), only fumaric acid **20** and pyruvic acid **21** were potential hits within the fragment data. Fumaric acid **20** was already shown to be detectable by SESI-HRMS [38]

3.4. Time trends related to the nutritional intervention

The experimental design included 14 measurements over the time period of six hours. It is, therefore, possible to follow multiple features over time and to compare their kinetic trends with previously reported ones [13–15]. For this purpose median kinetic trends of individual features were clustered with the help of the tslearn library [28]. To compare the clustering results with those reported [13] in blood plasma analysis, the number of clusters was set to five.

As shown in figure 6, cluster 1 contained 287 features that decreased after the intervention over the experimental period of over six hours as indicated by the mean of all time traces within this cluster. Within cluster 2, 518 features followed an opposite trend after the consumption of the shake, a steady increase over the experiment. In cluster 1, the features decreased immediately after the intervention. Some features within cluster 2 reacted intensely right after the intervention and a rise towards the end was observed. The 977 features of cluster 3 rise towards the two-hour mark and subsequently flatten, indicative of metabolic products of the nutrients present in the intervention. Cluster 4 exhibited wash-out kinetics of intervention-related components (255 features). The 255 features of cluster 5 show similar trends to those of cluster 3. The underlying features might have been secondary metabolites of the macronutrients and thus increase up to two hours, but they decreased again as they were further metabolized.

A comparison of the observed time clusters with the ones reported for blood plasma metabolites by Wopereis *et al* [13]. for the shake intervention revealed a large overlap of the observed clusters: except for cluster 5, all others had an analog within the time traces reported for blood plasma metabolites.

An important measure to follow the fatty acid metabolism is by monitoring carnitine and its acetylated derivatives [53, 54]. These molecules were previously detected in human breath using SESI-HRMS [55]. Acetylcarnitine reacts intensely to the intervention with its intensity dropping within the first 30 min, while carnitine increases (figure 7(a)). The trends of these molecules were opposite to each other as one would expect from the fact that acetylcarnitine is used by the body for energy storage and is formed

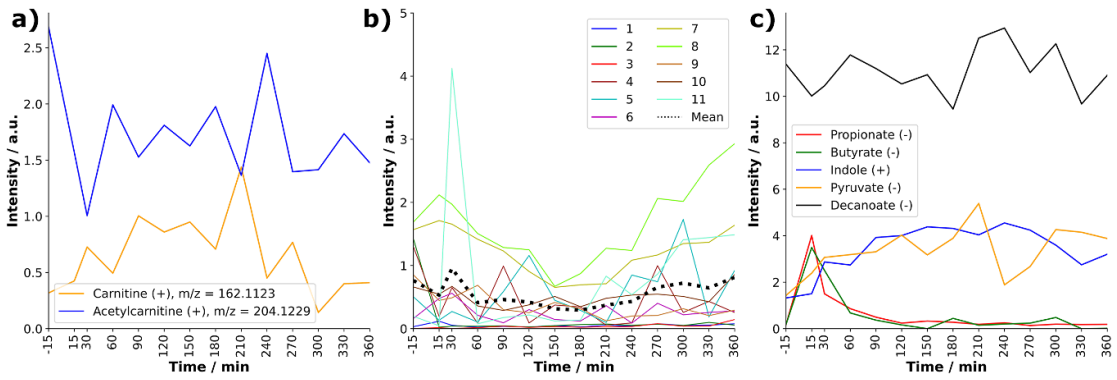

Figure 7. (a) Normalized mean time traces of carnitine and acetylcarnitine from all subjects over six hours. In terms of their behavior, these molecules were expected to show an inverse relationship [11]. This was true for most time intervals except the last hour. (b) Mean acetone time traces of each subject with the mean trend depicted as a dotted black line. (c) Mean time traces of selected compounds illustrating the general behavior of these metabolites after the nutritional intervention. Propionate and butyrate increased sharply after the intervention but fell quickly again. Indole and pyruvate rose slowly and their decrease was steadier compared to propionate and butyrate. The intensity for decanoate decreased within the first three hours, then rose and was followed by a follow-up decrease.

from carnitine [53, 54]. A significant inter-individual variation was observed for acetone (figure 7(b)). Between subjects, a large difference in signal intensity was distinguishable and sudden spikes of the acetone intensity were observed for multiple subjects. As the data were normalized to the total ion current of the measurement, a large acetone excess could suppress other ions, thus being overrepresented and leading to these observed spikes. As one of the major components of breath VOCs [51], interindividual differences would also lead to major differences between intensities and account for the observed differences.

Selected time traces of metabolites annotated through CID-experiments (figure 7(c)) illustrate the various trends captured through SESI-HRMS. The anions propionate and butyrate rose after the intervention within the first 15 min and decayed exponentially again. A study [56] on postprandial SCFA anion levels within blood after glucose consumption reported a rise in propionate levels after the four-hour mark [56]. Such a rise was not observed within the data reported here. Increasing the time of measurement after the intervention would be thus favourable for following gut microbiome activity with SESI-HRMS. Indoles, compounds that are associated with gut microbiome fermentation [49], did not react immediately to the intervention, in contrast to the SCFAs. Instead, the intensity increased until around three hours after which a slow decrease followed. Within the six-hour measurement period, the level did not return to the baseline measurement. Similar behavior was observed for pyruvate, although at the four-hour mark a decrease in intensity was detected, after which it increased again. A potential hypothesis of this drop could be the switch from catabolism to anabolism. Decanoate, as a representative of medium-chain fatty acids, followed a less clear trend, the intensity decreased after the intervention and subsequently underwent twice a cycle of rising and falling in levels.

The annotated breath metabolite trends registered with SESI-HRMS gave an accurate representation of metabolite kinetics reported for blood plasma after a nutritional intervention, although at the individual level the variation had a non-negligible influence. Targeted analysis of selected metabolites with comparison to blood plasma LC-MS analysis could deliver additional insight into the magnitude of variation between SESI-HRMS and standard LC-MS.

4. Conclusions

A nutritional intervention with subsequent on-line breath measurements to follow the impact over six hours was conducted showing the utility of on-line SESI-HRMS for nutritional metabolomics. To distinguish the effect of the intervention on metabolism, a cross-over experiment was conducted comparing the state of fasting with the intervention. Subsequent PLS-DA revealed features related to the intervention. These features were then mapped to metabolic pathways, indicating multiple pathways connected to chemical components of the intervention. Each macronutrient class represented in the nutritional challenge seemed to have caused a response within the human metabolome and thus confirms the relevance of SESI-HRMS breath analysis for following postprandial human behaviour. MS^2 experiments after the intervention indicated potential structural candidates for significant features. Among the identified candidates were the classes of fatty acids, amino acids, and unsaturated hydrocarbons, as well as molecules related to gut microbiome activity and the citric acid cycle. Further nutritional studies utilizing this technique could therefore rely on these classes of metabolites. The metabolites could be detected with targeted experiments as one way

to closely monitor selected representatives of these classes. Following the kinetics of selected metabolites upon the nutritional intervention revealed an overlap of the observed mean trends with the ones reported in the literature. Comparison of acetone kinetics between individuals revealed large fluctuations and differences, exemplifying the need for a higher degree of standardization and control of measurements to minimize potential instrumental variation to achieve better comparability of individual kinetics.

The results presented here serve as a basis for further endeavours at the important interface of nutritional science, metabolomics, and breath research. This work illustrates the potential of SESI-HRMS as a new important tool at the intersection of these fields as well as, more generally, for nutritional science.

Data availability statement

The original data used in this publication are made available in a curated data archive at ETH Zürich (www.research-collection. ethz.ch) under the DOI: https://doi.org/10.3929/ethz-b-000551788.

The data that support the findings of this study are openly available at the following URL/DOI: 10.3929/ethz-b-000551788.

Acknowledgments

The research leading to these results has received funding from the NutriExhalomics project (Grant Agreement No. 15668) and the Jubiläumsstiftung der Von Roll Holding AG. The authors would like to thank Pablo Sinues, Jiayi Lan, Bettina Streckenbach, and Jérôme Käslin for valuable discussion and ideas.

ORCID iDs

Renato Zenobi https://orcid.org/0000-0001-5211-4358
Stamatios Giannoukos https://orcid.org/0000-0002-5066-533X

References

[1] Gibbons H, O'Gorman A and Brennan L 2015 Metabolomics as a tool in nutritional research *Curr. Opin. Lipidol.* **26** 30–34
[2] Bas-Bellver C, Andrés C, Seguí L, Barrera C, Jiménez-Hernández N, Artacho A, Betoret N and Gosalbes M J 2020 Valorization of Persimmon and Blueberry byproducts to obtain functional powders: *in vitro* digestion and fermentation by gut microbiota *J. Agric. Food Chem.* **68** 8080–90
[3] Ulaszewska M M *et al* 2019 Nutrimetabolomics: an integrative action for metabolomic analyses in human nutritional studies *Mol. Nutr. Food Res.* **63** 1800384
[4] Saenger T, Hübner F and Humpf H-U 2017 Short-term biomarkers of apple consumption *Mol. Nutr. Food Res.* **61** 1600629
[5] Praticò G, Gao Q, Manach C and Dragsted L O 2018 Biomarkers of food intake for allium vegetables *Genes Nutr.* **13** 34
[6] Lawson L D and Gardner C D 2005 Composition, stability, and bioavailability of garlic products used in a clinical trial *J. Agric. Food Chem.* **53** 6254–61
[7] Rosen R T, Hiserodt R D, Fukuda E K, Ruiz R J, Zhou Z, Lech J, Rosen S L and Hartman T G 2001 Determination of allicin, S-allylcysteine and volatile metabolites of garlic in breath, plasma or simulated gastric fluids *J. Nutr.* **131** 968S–71S
[8] Cuparencu C, Rinnan Å and Dragsted L O 2019 Combined markers to assess meat intake-human metabolomic studies of discovery and validation *Mol. Nutr. Food Res.* **63** e1900106
[9] Khodorova N V, Rutledge D N, Oberli M, Mathiron D, Marcelo P, Benamouzig R, Tomé D, Gaudichon C and Pilard S 2019 Urinary metabolomics profiles associated to bovine meat ingestion in humans *Mol. Nutr. Food Res.* **63** e1700834
[10] Trimigno A, Münger L, Picone G, Freiburghaus C, Pimentel G, Vionnet N, Pralong F, Capozzi F, Badertscher R and Vergères G 2018 GC-MS based metabolomics and NMR spectroscopy investigation of food intake biomarkers for milk and cheese in serum of healthy humans *Metabolites* **8** 26
[11] Krug S *et al* 2012 The dynamic range of the human metabolome revealed by challenges *FASEB J.* **26** 2607–19
[12] van Ommen B, van der Greef J, Ordovas J M and Daniel H 2014 Phenotypic flexibility as key factor in the human nutrition and health relationship *Genes Nutr.* **9** 423
[13] Wopereis S, Stroeve J H M, Stafleu A, Bakker G C M, Burggraaf J, van Erk M J, Pellis L, Boessen R, Kardinaal A A F and van Ommen B 2017 Multi-parameter comparison of a standardized mixed meal tolerance test in healthy and type 2 diabetic subjects: the PhenFlex challenge *Genes Nutr.* **12** 21
[14] Pellis L *et al* 2012 Plasma metabolomics and proteomics profiling after a postprandial challenge reveal subtle diet effects on human metabolic status *Metabolomics* **8** 347–59
[15] Fiamoncini J *et al* 2018 Plasma metabolome analysis identifies distinct human metabotypes in the postprandial state with different susceptibility to weight loss-mediated metabolic improvements *FASEB J.* **32** 5447–58
[16] Raninen K J, Lappi J E, Mukkala M L, Tuomainen T-P, M M H, Poutanen K S and Raatikainen O J 2016 Fiber content of diet affects exhaled breath volatiles in fasting and postprandial state in a pilot crossover study *Nutr. Res.* **36** 612–9
[17] Bruderer T, Gaisl T, Gaugg M T, Nowak N, Streckenbach B, Müller S, Moeller A, Kohler M and Zenobi R 2019 On-line analysis of exhaled breath *Chem. Rev.* **119** 10803–28
[18] Hageman J H J, Nieuwenhuizen A G, van Ruth S M, Hageman J A and Keijer J 2019 Application of volatile organic compound analysis in a nutritional intervention study: differential responses during five hours following consumption of a high- and a low-fat dairy drink *Mol. Nutr. Food Res.* **63** 1900189
[19] Bruderer T, Gaugg M T, Cappellin L, Lopez-Hilfiker F, Hutterli M, Perkins N, Zenobi R and Moeller A 2020 Detection of volatile organic compounds with secondary electrospray ionization and proton transfer reaction high-resolution mass spectrometry: a feature comparison *J. Am. Soc. Mass Spectrom.* **31** 1632–40
[20] Pimentel G, Burnand D, H M L, Pralong F P, Vionnet N, Portmann R and Vergères G 2020 Identification of milk and cheese intake biomarkers in healthy adults reveals high interindividual variability of Lewis system–related oligosaccharides *J. Nutr.* **150** 1058–67
[21] Brouwer-Brolsma E M *et al* 2017 Combining traditional dietary assessment methods with novel metabolomics techniques: present efforts by the food biomarker alliance *Process Nutr. Soc.* **76** 619–27
[22] Lan J, Kaeslin J, Greter G and Zenobi R 2021 Minimizing ion competition boosts volatile metabolome coverage by secondary electrospray ionization orbitrap mass spectrometry *Anal. Chim. Acta* **1150** 338209
[23] Chambers M C *et al* 2012 A cross-platform toolkit for mass spectrometry and proteomics *Nat. Biotechnol.* **30** 918–20

[24] Smilde A K, Jansen J J, Hoefsloot H C J, Lamers R J A N, van der Greef J and Timmerman M E 2005 ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data *Bioinformatics* **21** 3043–8
[25] Jansen J J, Hoefsloot H C J, van der Greef J, Timmerman M E, Westerhuis J A and Smilde A K 2005 ASCA: analysis of multivariate data obtained from an experimental design *J. Chemom.* **19** 469–81
[26] Zwanenburg G, Hoefsloot H C J, Westerhuis J A, Jansen J J and Smilde A K 2011 ANOVA-principal component analysis and ANOVA-simultaneous component analysis: a comparison *J. Chemom.* **25** 561–7
[27] Pedregosa F *et al* 2011 Scikit-learn: machine learning in python *J. Mach. Learn. Res.* **12** 2825–30
[28] Tavenard R *et al* 2020 Tslearn, a machine learning toolkit for time series data *J Mach. Learn. Res.* **21** 1–6
[29] Schymanski E L, Jeon J, Gulde R, Fenner K, Ruff M, Singer H P and Hollender J 2014 Identifying small molecules via high resolution mass spectrometry: communicating confidence *Environ. Sci. Technol.* **48** 2097–8
[30] Kaeslin J and Zenobi R 2022 Resolving isobaric interferences in direct infusion tandem mass spectrometry *Rapid Commun. Mass Spectrom.* **36** e9266
[31] Dührkop K, Fleischauer M, Ludwig M, Aksenov A A, Melnik A V, Meusel M, Dorrestein P C, Rousu J and Böcker S 2019 SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information *Nat. Methods* **16** 299–302
[32] Wishart D S *et al* 2007 HMDB: the human metabolome database *Nucleic Acids Res.* **35** 521–6
[33] Kanehisa M, Goto S, Furumichi M, Tanabe M and Hirakawa M 2010 KEGG for representation and analysis of molecular networks involving diseases and drugs *Nucleic Acids Res.* **38** D355–60
[34] Li S, Park Y, Duraisingham S, Strobel F H, Khan N, Soltow Q A, Jones D P and Pulendran B 2013 Predicting network activity from high throughput metabolomics ed C A Ouzounis *PLoS Comput. Biol.* **9** e1003123
[35] Van H T 2012 Power study of anova versus Kruskal–Wallis test *J. Stat. Manage. Syst.* **15** 241–7
[36] Chong J, Wishart D S and Xia J 2019 Using metaboanalyst 4.0 for comprehensive and integrative metabolomics data analysis *Curr Protoc. Bioinforma* **68** 1–128
[37] Bocci F and Frega N 1996 Analysis of the volatile fraction from sunflower oil extracted under pressure *J. Am. Oil Chem. Soc.* **73** 713–6
[38] Tejero Rioseras A, Singh K D, Nowak N, Gaugg M T, Bruderer T, Zenobi R and Sinues P M-L 2018 Real-time monitoring of tricarboxylic acid metabolites in exhaled breath *Anal. Chem.* **90** 6453–60
[39] Wu N, Yang M, Gaur U, Xu H, Yao Y and Li D 2016 Alpha-ketoglutarate: physiological functions and applications *Biomol. Ther.* **24** 1–8
[40] Pons G, Raefsky-Estrin C, Carothers D J, Pepin R A, Javed A A, Jesse B W, Ganapathi M K, Samols D and Patel M S 1988 Cloning and cDNA sequence of the dihydrolipoamide dehydrogenase component human alpha-ketoacid dehydrogenase complexes *Process Natl Acad. Sci.* **85** 1422–6
[41] Blachier F, Mariotti F, Huneau J F and Tomé D 2007 Effects of amino acid-derived luminal metabolites on the colonic epithelium and physiopathological consequences *Amino Acids* **33** 547–62
[42] Romano A, Koczwara J B, Gallelli C A, Vergara D, Micioni Di Bonaventura M V, Gaetani S and Giudetti A M 2017 Fats for thoughts: an update on brain fatty acid metabolism *Int J. Biochem. Cell Biol.* **84** 40–45
[43] Morrison D J and Preston T 2016 Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism *Gut Microbes* **7** 189–200
[44] Schönfeld P and Wojtczak L 2016 Short- and medium-chain fatty acids in energy metabolism: the cellular perspective *J. Lipid Res.* **57** 943–54
[45] Lin R, Liu W, Piao M and Zhu H 2017 A review of the relationship between the gut microbiota and amino acid metabolism *Amino Acids* **49** 2083–90
[46] Engelke U F H, Tangerman A, Willemsen M A A P, Moskau D, Loss S, Mudd S H and Wevers R A 2005 Dimethyl sulfone in human cerebrospinal fluid and blood plasma confirmed by one-dimensional 1H and two-dimensional 1H-13C NMR *NMR Biomed.* **18** 331–6
[47] Zelante T *et al* 2013 Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22 *Immunity* **39** 372–85
[48] Wlodarska M *et al* 2017 Indoleacrylic acid produced by commensal peptostreptococcus species suppresses inflammation *Cell Host Microbe* **22** 25–37.e6
[49] Roager H M and Licht T R 2018 Microbial tryptophan catabolites in health and disease *Nat. Commun.* **9** 3294
[50] Guillen M D and Goicoechea E 2008 Formation of oxygenated α,β-unsaturated aldehydes and other toxic compounds in sunflower oil oxidation at room temperature in closed receptacles *Food Chem.* **111** 157–64
[51] Drabinska N, Flynn C, Ratcliffe N, Belluomo I, Myridakis A, Gould O, Fois M, Smart A, Devine T and de Lacy Costello B P J 2021 A literature survey of volatiles from the healthy human breath and bodily fluids: the human volatilome *J. Breath Res.* **15** 034001
[52] Mochalski P, King J, Klieber M, Unterkofler K, Hinterhuber H, Baumann M and Amann A 2013 Blood and breath levels of selected volatile organic compounds in healthy volunteers *Analyst* **138** 2134–45
[53] Hoppel C 2003 The role of carnitine in normal and altered fatty acid metabolism *Am. J. Kidney Dis.* **41** S4–12
[54] Bene J, Hadzsiev K and Melegh B 2018 Role of carnitine and its derivatives in the development and management of type 2 diabetes *Nutr. Diabetes* **8** 1–10
[55] Nowak N *et al* 2021 Rapid and reversible control of human metabolism by individual sleep states *Cell Rep.* **37** 109903
[56] Tarini J and Wolever T M S 2010 The fermentable fibre inulin increases postprandial serum short-chain fatty acids and reduces free-fatty acids and ghrelin in healthy subjects *Appl. Physiol. Nutr. Metab.* **35** 9–16