Semi-targeted metabolomics

An end-to-end robust semi-targeted metabolomics workflow to facilitate deeper coverage and confident annotation of metabolites in milk

Bashar Amer, Rahul Ravi Deshpande, Daniel Hermanson, Andreas Hühmer, and Susan Bird
Thermo Fisher Scientific, 355 River Oaks Parkway, San Jose, California, 95134, United States

Abstract

Purpose: Development of an intelligent data acquisition workflow for semi-targeted LC-MS metabolomics with deep metabolome coverage, accurate metabolite quantitation, and confident compound annotation in milk.

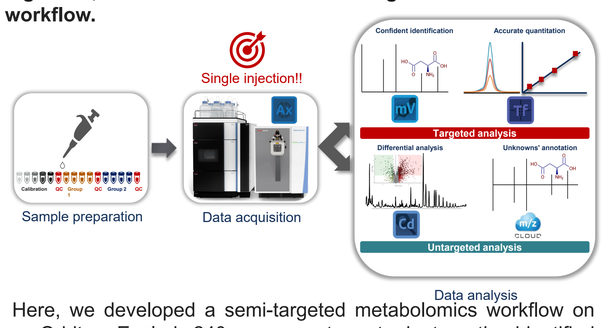
Methods: A semi-targeted metabolomics workflow was developed on a Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer coupled to a Thermo Scientific™ Vanquish™ Horizon UHPLC system by targeting identified metabolites (i.e., amino acids) in milk, processed with Thermo Scientific™ TraceFinder™ Software 5.1. The data was, then, reanalyzed using an untargeted approach to assess other metabolic variations among the different samples utilizing Thermo Scientific™ Compound Discoverer™ 3.3 software.

Results: Plant-based milk showed, in general, higher concentrations of amino acids compared to bovine milk. Additionally, higher levels of organic acids were reported in bovine milk.

Introduction

Targeted metabolomics is used in hypothesis-driven research to annotate and quantify a biologically relevant subset of known metabolites. However, targeted analysis has limited coverage of the metabolome. In comparison, untargeted metabolomics offers a wider overview of metabolites and their relative levels, providing the opportunity to find unexpected changes not part of the original goals. The varying physiochemical properties of the metabolome require specific detection criteria for different metabolite groups. Semi-targeted metabolomics recently emerged as a promising alternative, offering researchers a middle ground between targeted and untargeted approaches in one single experiment (Figure 1). Semi-targeted workflows begin with annotation and quantifying a pre-selected group of metabolites in a sample. In addition, the data can then be reanalyzed (retro-mined) to look for global metabolic changes that were not part of the original focus, therefore, identifying other biologically meaningful metabolite changes.

Figure 1, Illustration of the semi-targeted metabolomics workflow.



Here, we developed a semi-targeted metabolomics workflow on an Orbitrap Exploris 240 mass spectrometer by targeting identified metabolites (i.e., amino acids) in animals and plant-based milk. The data was, then, reanalyzed using an untargeted approach to assess other metabolic variations among the different samples utilizing Compound Discoverer 3.3 software for data processing, unknown identification, and differential analysis.

Materials and methods

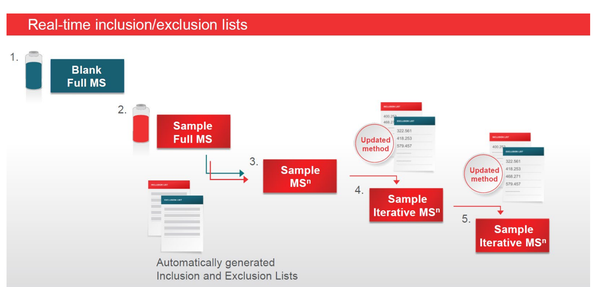
Sample preparation

Animal and plant-based milk samples were obtained from local markets (San Jose, California). Pooled samples were prepared, by mixing 100 µL of each sample, to be used for quality control (QC). Aliquots of milk and QC samples were collected in 3 mL Eppendorf tubes and kept at -80˚ C until the time of analysis. Metabolites were extracted after thawing samples in an ice bath using the modified Folch method by adding 1 mL of chloroform:methanol (2:1 v/v) solution and 300 µL of water to 200 µL of milk. The organic solvents mix contained isotope-labeled standards (IS) to evaluate LC-MS data acquisition quality (untargeted and targeted methods) and for calibration and concentration calculations (targeted method). The mixture was then vortexed for 3 minutes at room temperature and centrifuged for 15 minutes (21 k x g) at 4˚ C to separate the two extraction layers. An aliquot, 500 µL, of the methanol:water, the upper layer, was transferred to 3 mL Eppendorf tubes and evaporated under nitrogen flow at 37˚ C for 60 minutes using a TurboVap® LV nitrogen evaporator from Biotage. Finally, samples were resuspended in 500 µL of 5% methanol solution in LC-MS water, vortexed for 3 minutes at room temperature, and centrifuged for 10 minutes (21 k x g) at 4˚ C before submitting an aliquot of the supernatant to LC-MS analysis.

Data acquisition

A full scan (70 – 800 m/z), polarity switching (ESI (+)/ESI (-)) MS-based method was developed for the semi-targeted workflow in this study that is focusing on the quantitation of amino acids while performing an untargeted analysis of the extracted milk samples in one injection. Data were acquired on an Orbitrap Exploris 240 mass spectrometer using the Thermo Scientific™ Deep Scan AcquireX acquisition workflow (Figure 2).

Figure 2. AcquireX Deep Scan mode for intelligent data acquisition to maximize the number of relevant compounds interrogated by MS/MS, resulting in higher coverage and confidence annotation.



Liquid Chromatography

LC system: Vanquish Horizon UHPLC system. Autosampler temp.: 5 ˚ C.

HPLC Column: Thermo Scientific™ Hypersil GOLD™ C18 column (2.1 x 150 mm, 1.9 µm) at 45 ˚ C.

Injection Volume: 2 µL.

Mobile Phase: (A) 0.1% (v) formic acid (FA) in LC-MS grade water

(B) 0.1% (v) FA in LC-MS grade methanol

HPLC Gradient (untargeted):

Time	A%	B%
0.00	100	0
8.00	50	50
9.00	2	98
13.00	2	98
13.10	100	0
15.00	100	0

Flow rate: 0.30 mL/min.

Divert valve: to waste = 0 – 0.2 min, to MS = 0.2 – 15.0 min.

Mass Spectrometry

Mass spectrometer: Orbitrap Exploris 240 mass spectrometer equipped with heated ESI probe. Ion source settings: polarity switching mode with spray Voltage = 3.5 and 3.0 kV, positive and negative polarity, respectively. Vaporizer = 320 ˚ C, Transfer Tube = 275 ˚ C, RF Lens = 35 %, Sheath Gas = 40, Aux. gas = 8, Sweep Gas = 1. Scan range: 70 – 800 m/z, at 120 k orbitrap resolution. Scan-to-scan Easy-IC™ internal calibration.

Data Analysis

All data were acquired using Thermo Scientific™ Xcalibur™ Software. Targeted amino acids (i.e., alanine, isoleucine, leucine, phenylalanine, proline, and valine) standards and isotope-labeled internal standards were used to prepare calibration solutions. Quantitation data were processed in TraceFinder Software 5.1 using a 3-ppm mass tolerance filter.

Thermo Scientific™ Compound Discoverer™ 3.3 software was used for data processing, unknown annotation, and differential analysis of the untargeted portion of the workflow.

Results

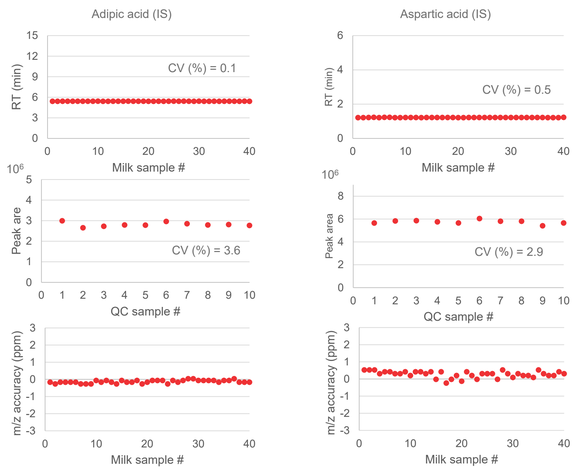
Data acquisition

A 15-minute reversed-phase LC-MS method was developed to quantify amino acids and assess other unknown variations among different milk samples; bovine milk with various fat content, almond, oat, coconut, and soy milk. The high resolution and high mass accuracy of the orbitrap lead to improved discrimination between signals derived from analytes and those resulting from co-eluting isobaric compounds or matrix interferences.

Method validation

Instrument data quality and robustness were assessed by evaluating the spiked adipic acid and aspartic acid isotopically labeled internal standards using metrics including retention time, mass accuracy, and signal response. Sub-ppm mass accuracy was detected for the two internal standards over the entire acquisition period. Minimal chromatographic shift and consistent signal responses were observed as evidenced by low %CV for quality control samples, which were run intermittently throughout the sequence, Figure 3.

Figure 3. Reproducibility of retention time (RT), mass accuracy in ppm, and integrated peak areas of isotope-labeled internal standards (IS) spiked into milk and quality control (QC) samples. Adipic acid (IS): $^{13}C_6H_{10}O_4$ and aspartic acid (IS): $^{13}C_4H_7{}^{15}NO_4$.



Calibration data

Calibration curves were created for the quantified compounds using internal calibration. Linear fit (R^2 > 0.99) was observed for all targeted amino acids. Figure 4 shows the calibration curve for phenylalanine (3.13 – 200 µM) as an example. All calibration levels showed a CV ≤ 10% and an average calculated difference CV ≤ 10%.

Figure 4. Calibration curves of phenylalanine and maleic acid were created and used for quantitation via TraceFinder 5.1 software.

Lower limits of quantification (LOQ) and lower limits of detection (LOD) are presented in Table 1 for targeted compounds.

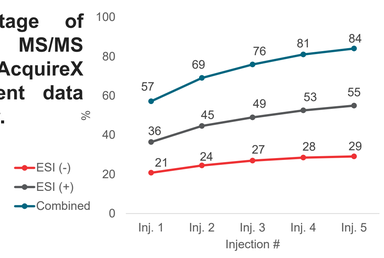
Table 1. Lower limits of quantitation (LOQ) and lower limits of detection (LOD) in µM of analyzed compounds using the developed targeted method.

	LOQ (µM)	LOD (µM)
Alanine	0.39	0.39
Isoleucine	1.56	1.56
Leucine	6.25	1.56
Phenylalanine	3.13	1.56
Proline	6.25	0.39
Valine	0.39	0.39

AcquireX Deep Scan intelligent data acquisition

The deep scan AcquireX workflow increased the percentage of fragmented compounds (Figure 5) while reducing the number of fragmented background compounds, increasing instrument utilization, and enabling the fragmentation of lower abundance compounds. This results in improved annotation capabilities on a wider dynamic range of compounds across the different varieties of milk.

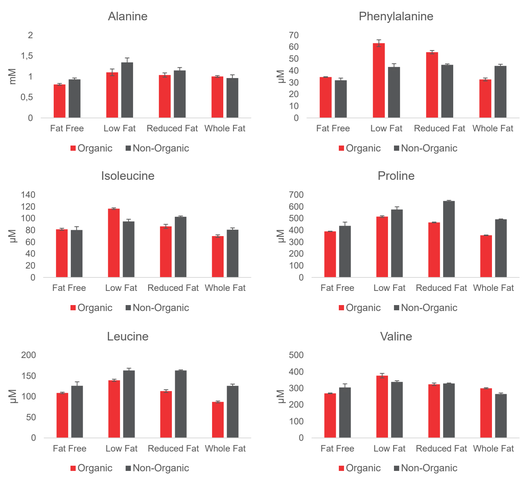
Figure 5. Percentage of compounds with MS/MS spectra utilizing the AcquireX deep scan intelligent data acquisition workflow.



Amino acids in milk

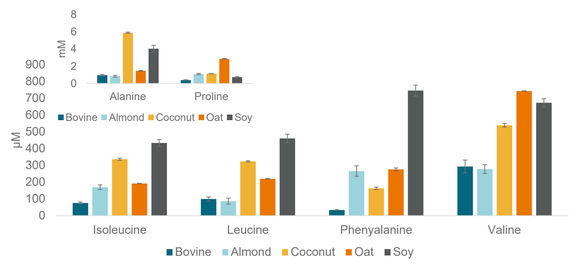
Calculated concentrations of the investigated amino acids in analyzed bovine milk and plant-based milk samples are plotted in Figure 6 and Figure 7.

Figure 6. The concentration of targeted amino acids in analyzed bovine milk samples. Milk samples were grouped based on fat level and milk type (organic vs. non-organic; one milk brand was selected for each type).



Interestingly, different fat-level milk samples showed variation in targeted amino acid levels; where low-fat (1% fat) and reduced-fat (2% fat) milk showed higher content of those amino acids compared to fat-free (0% fat) and whole-fat (3.5% fat) milk, which might be a result of the physical treatment of milk during production. In addition, variations in the concentration of phenylalanine and proline were observed comparing organic with non-organic bovine milk being higher in organic for the former and lower for the latter (Figure 6). This might be a result of variation in protein content of feeding material for each type. These variations could facilitate the differentiation between fat-free and low-fat milk for example, and between organic and non-organic milk.

Figure 7. The concentration of targeted amino acids in analyzed bovine and plant-based milk samples. Whole fat (3.5% fat) bovine milk was selected for this comparison.



Bovine milk (whole fat) showed lower levels of amino acid concentrations compared to all plant-based milk samples, which also demonstrated variation among each other (Figure 7). Soy milk, for example, contained ~20-fold higher levels of phenylalanine compared to bovine milk, but only 2-fold higher levels of proline compared to bovine milk. Noticeably, phenylalanine concentration in soy milk is significantly higher than in any other milk type. On another end, higher levels of alanine can be considered as a distinguishing marker for almond milk compared to all other milk types analyzed in this study. In general, the relative concentration of measured amino acids was lower in bovine milk compared to plant-based milk. Almond milk shared a similar pattern with bovine milk for alanine, leucine, and valine.

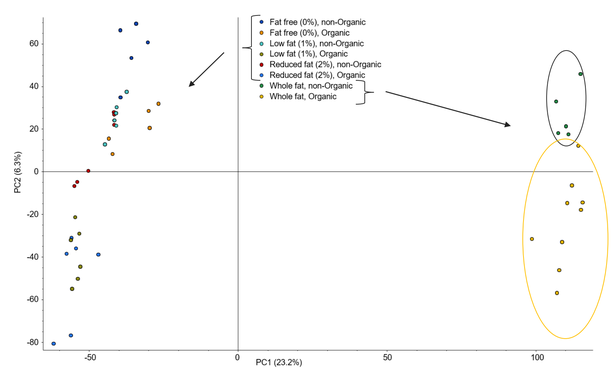
Differential analysis and compound annotation

Differential analysis and compound annotation using Compound Discoverer 3.3 software revealed relative differences among the milk samples and provided a wide array of annotation tools to leverage the acquired data.

Bovine milk

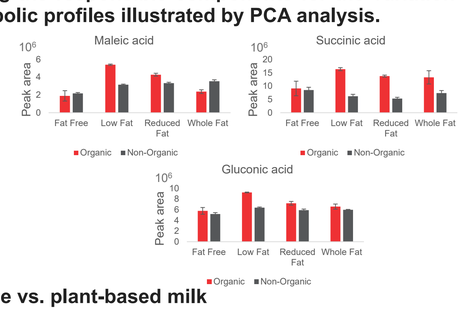
Bovine milk samples showed significant variation in their polar metabolic profiles based on their fat content as illustrated by the scores plot of PCA analysis in Figure 8. Moreover, a clear separation was demonstrated between organic and non-organic milk in each milk type.

Figure 8. Scores plot of PCA analysis showing the distribution of analyzed bovine milk samples based on their polar metabolic profiles.



The performed PCA analysis facilitated selecting markers, which are responsible for the variation observed between the different bovine milk samples. Organic acids such as maleic acid, succinic acid, and gluconic acid were among those milk components (Figure 9).

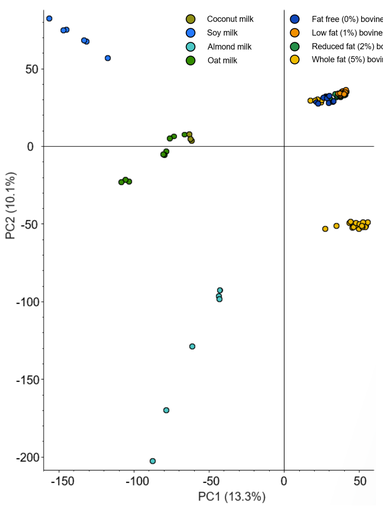
Figure 9. Variations in levels of organic acids, which are among the responsible components for the variation in milk metabolic profiles illustrated by PCA analysis.



Bovine vs. plant-based milk

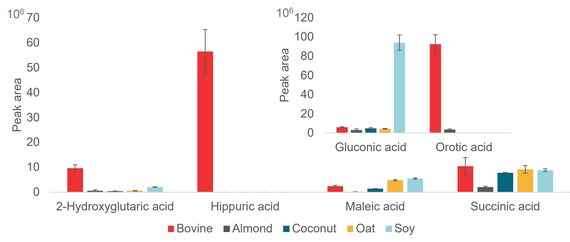
Further analysis revealed relative differences between bovine milk (whole milk was selected for this comparison) and plant-based milk samples (almond, oat, coconut, and soy) as shown in the scores plot of PCA analysis in Figure 10. Plant-based milk samples were significantly discriminated against bovine milk. In addition, a clear separation was demonstrated among plant-based milk samples.

Figure 10. Scores plot of PCA analysis showing the distribution of analyzed bovine and plant-based milk samples based on their polar metabolic profiles.



Organic acids such as 2-hydroxyglutaric acid, hippuric acid, maleic acid, succinic acid, gluconic acid, and orotic acid were among those milk components as shown in Figure 11. this observed data from the untargeted analysis of the semi-targeted workflow can facilitate the identification of markers, which can be then used to assess the quality and to authenticate milk for increased food security and consumer protection

Figure 11. Variations in levels of organic acids, which are among the responsible components for the variation between polar metabolic profiles of bovine and milk-based milk.



Conclusions

An end-to-end robust semi-targeted metabolomics workflow enables the ability to perform targeted and untargeted analysis in a single sample injection allowing scientists to gain more knowledge about biological samples.

Trademarks/licensing

Learn more at https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/metabolomics-lipidomics.html