Intelligent data acquisition for untargeted metabolomics followed by high-throughput quantitative metabolomics utilizing high-resolution accurate mass measurements

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

Abstract

Purpose: Development of an intelligent data acquisition workflow for untargeted LC-MS metabolomics with deep metabolome coverage and confident compound annotation to identify components for a high-throughput and robust quality screening study in milk.

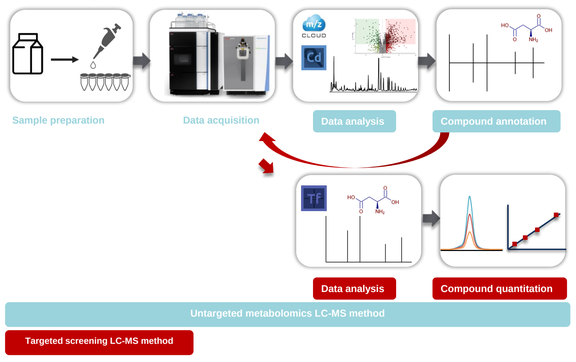
Methods: Reversed-phase LC-MS methods using Hypersil GOLD™ separation were developed utilizing a Thermo Scientific™ Vanquish™ Horizon UHPLC system coupled to a Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer to assess and quantitate metabolic variation among different milk samples; bovine milk with various fat content, almond, oat, coconut, and soy milk.

Results: Higher levels of amino acids were shown to classify plant-based milk from bovine milk. Hippuric acid and orotic acid were verified as markers for bovine milk compared to plant-based milk. Gluconic acid, however, was verified as a marker for soy milk.

Introduction

The goal of untargeted metabolomics is to comprehensively detect and annotate as many metabolites as possible in biological samples. Efforts are continuously made to improve analytical workflows in terms of sensitivity, mass accuracy, robustness, and metabolome coverage. Here we outline an untargeted metabolomics workflow using a Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer to assess metabolic variation among different milk samples (i.e., bovine and plant-based milk). This approach utilizes high-resolution accurate mass full scan data for robust and sensitive compound detection and an AcquireX™ intelligent data acquisition workflow to maximize the number of relevant compounds interrogated by MS/MS, resulting in higher confidence annotation. The untargeted workflow is used to identify components that are then targeted in a screening study (Figure 1).

Figure 1. An outline of an untargeted metabolomics workflow using a Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer to assess metabolic variation among different milk samples to identify components that could be then targeted in a high-throughput screening study using the same analytical platform.



Materials and methods

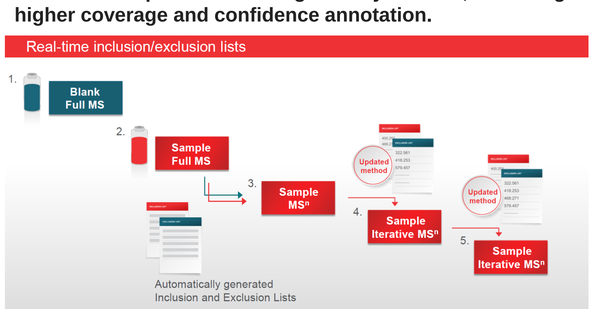
Sample Preparation – Untargeted and Targeted Workflows

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 – Untargeted Workflow

A full scan (70 – 800 m/z), polarity switching (ESI (+)/ESI (-)) MS-based method was developed for the untargeted analysis of extracted milk samples. Data were acquired on an Orbitrap Exploris 240 mass spectrometer using the Deep Scan AcquireX acquisition workflow (Figure 2). Two full-scan (70 – 800 m/z) MS-based methods were developed for the quantification of selected amino acids, via ESI (+ polarity ionization), and organic acids, via ESI (- polarity ionization).

Figure 2. Thermo Scientific™ 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: Thermo Scientific™ Vanquish™ Horizon UHPLC system. Autosampler temp.: 5 ˚ C.

HPLC Column: Thermo Scientific Hypersil GOLD™ C18 (2.1 x 150 mm, 1.9 µm) at 45 ˚ C (untargeted) and 55 ˚ C (targeted).

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

HPLC Gradient (targeted):

Time	A%	B%
0.00	100	0
3.50	80	20
3.70	2	98
4.00	2	98
4.60	100	0
5.00	100	0

Flow rate: 0.30 mL/min (untargeted) and 0.45 mL/min (targeted)
Divert valve: to waste = 0 – 0.2 min
to MS = 0.2 – 15.0 min (untargeted) and 0.2 – 5.0 min (targeted)

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. Thermo Scientific™ Compound Discoverer™ 3.3 software was used for data processing, unknown identification, and differential analysis for the untargeted metabolomics runs. Targeted compound standards and isotope-labeled internal standards were used to prepare calibration solutions. Quantitation data were processed in Thermo Scientific™ TraceFinder™ Software 5.1 using a 3-ppm mass tolerance filter.

Results

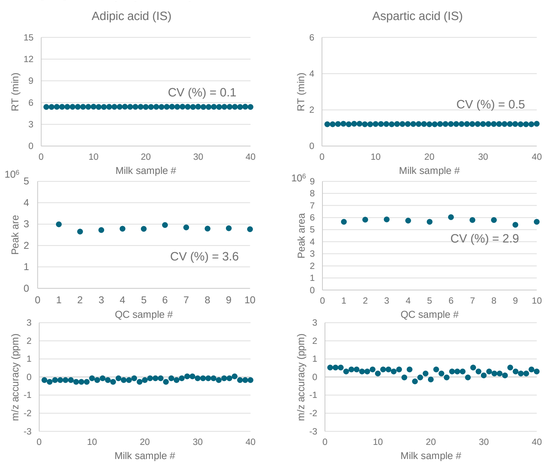
Data Acquisition

A 15-minute reversed-phase LC-MS method was developed to assess metabolic variation among different milk samples. In addition, two 5-minute screening methods were developed utilizing the same LC-MS system to quantify selected amino acids, via ESI(+), and organic acids, via ESI(-), in milk samples. 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 – Untargeted Workflow

Instrument data quality and robustness were assessed by evaluating the spiked IS 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.

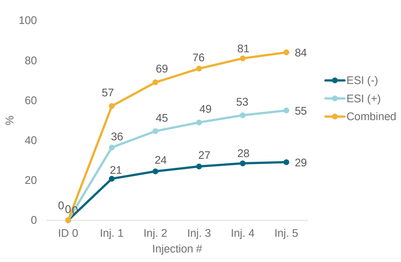
Figure 3. Reproducibility of retention time (RT), mass accuracy in ppm, and integrated peak areas of isotope-labeled internal standards (IS) spiked into milk and quality control (QC) samples. Adipic acid (IS): $^{13}C_6H_{10}O_4$ and aspartic acid (IS): $^{13}C_4H_7{}^{15}NO_4$.



AcquireX Deep Scan Intelligent Data Acquisition

The deep scan AcquireX workflow increased the percentage of fragmented compounds (Figure 4) 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 4. Percentage of compounds with MS/MS spectra utilizing AcquireX deep-scan intelligent data acquisition workflow.

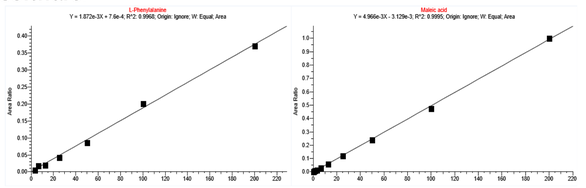


Method Validation – Targeted Workflow

High data quality, reliability, and robustness of measurement were observed by evaluating the isotopically labeled internal standards to assess instrument performance using metrics including retention time, peak area, and mass accuracy in milk and QC samples. Minimal chromatographic shift and consistent signal responses were observed.

Calibration curves were created for the quantified compounds using internal calibration. Linear fit (R2 > 0.99) was observed for all compounds. Figure 5 shows calibration curves for phenylalanine (3.13 – 200 µM) and maleic acid (0.39 – 200 µM) as two examples. All calibration levels showed a CV ≤ 10% and an average calculated difference CV ≤ 10%.

Figure 5. 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.

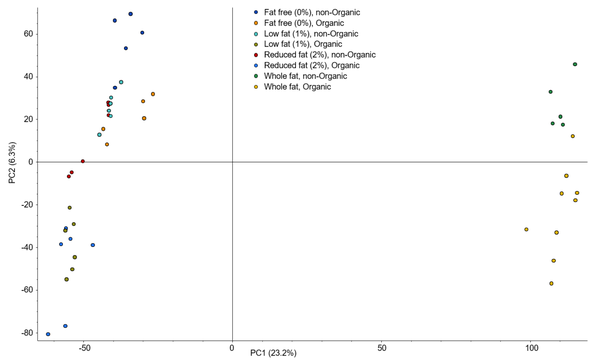
Table 1. Lower limits of quantitation (LOQ) and lower limits of detection (LOD) in µM of analyzed compounds using the developed targeted method.

Amino acids	LOQ (µM)	LOD (µM)	Organic acids	LOQ (µM)	LOD (µM)
Alanine	0.39	0.39	Maleic acid	0.78	0.39
Isoleucine	1.56	1.56	Succinic acid	0.39	0.39
Leucine	6.25	1.56	Gluconic acid	6.25	1.56
Phenylalanine	3.13	1.56	Malic acid	0.39	0.39
Proline	6.25	0.39	Hippuric acid	0.39	0.39
Valine	0.39	0.39	2-Hydroxyglutaric acid	0.78	0.39
			Orotic acid	0.39	0.39

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 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 6. Moreover, a clear separation was demonstrated between organic and non-organic milk in each milk type.

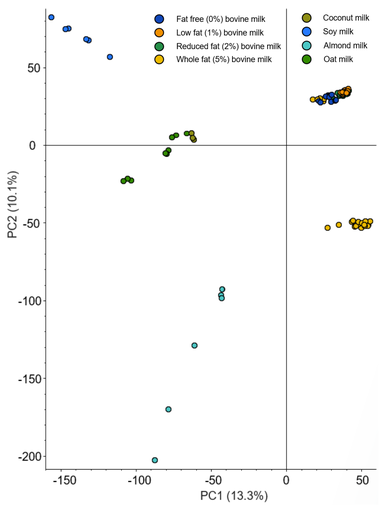
Figure 6. Scores plot of PCA analysis showing the distribution of analyzed bovine milk samples based on their polar metabolic profiles.



The performed PCA analysis facilitated selecting markers, which are responsible for the variation observed between the different bovine milk samples. Amino acids such as phenylalanine, isoleucine, leucine, valine, and proline, and organic acids such as maleic acid, succinic acid, and gluconic acid were among those milk components. These components are selected to be targeted in the high-throughput screening study to classify milk samples (Figure 1).

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 7. Plant-based milk samples were significantly discriminated against bovine milk. In addition, a clear separation was demonstrated among plant-based milk samples.

Figure 7. Scores plot of PCA analysis showing the distribution of analyzed bovine and plant-based milk samples based on their polar metabolic profiles.



Amino acids such as phenylalanine, isoleucine, leucine, valine, alanine, and proline, and organic acids such as 2-hydroxyglutaric acid, hippuric acid, maleic acid, succinic acid, gluconic acid, and orotic acid were among those milk components. These components are selected to be targeted in the high-throughput screening study to classify milk samples (Figure 1).

Amino Acids in Milk

Calculated concentrations of the investigated amino acids in analyzed bovine milk and plant-based milk samples are plotted in Figure 8 and Figure 9.

Figure 8. 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).

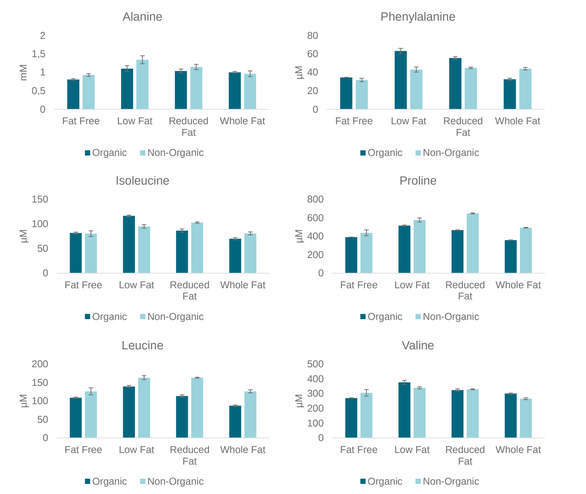
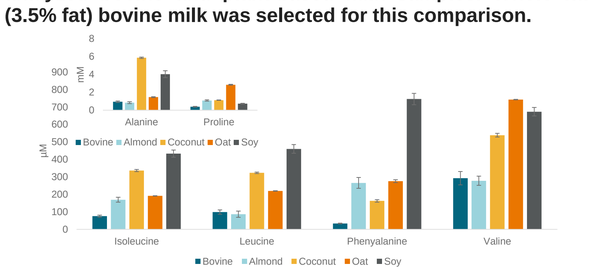


Figure 9. The concentration of targeted amino acids in analyzed bovine and plant-based milk samples. Whole fat (3.5% fat) bovine milk was selected for this comparison.



Organic Acids in Milk

Measured levels of the investigated organic acids in analyzed bovine milk and plant-based milk samples are plotted in Figure 10 and Figure 11.

Figure 10. The concentration of targeted organic acids level in analyzed bovine milk samples. Milk samples were grouped based on fat level and milk type (organic vs. non-organic; one milk brand and commercial source were selected for each type).

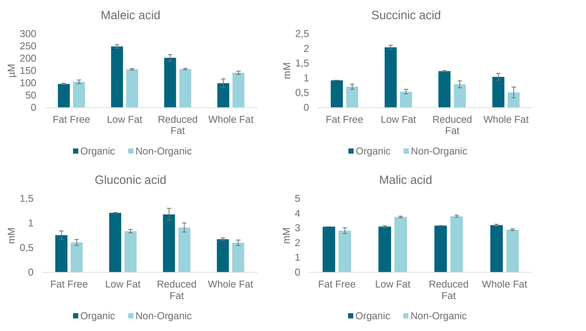
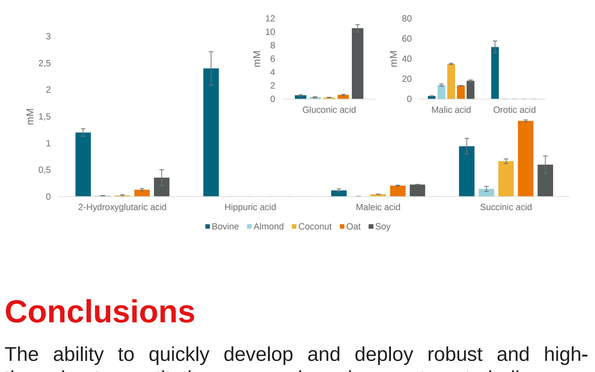


Figure 11. The concentration of targeted organic acids in analyzed bovine and plant-based milk samples. Whole fat (3.5% fat) bovine milk was selected for this comparison.



Conclusions

The ability to quickly develop and deploy robust and high-throughput quantitative assays based on untargeted discovery experiment results is essential for the validation and application of findings, which could be used to assess the quality and authenticate milk for increased food security and consumer protection.

Trademarks/licensing

Learn more at thermofisher.com/metabolomics