# Development of Chemical Isotope Labeling (CIL) LC-MS Techniques for Untargeted Analysis Milk Metabolomics and Targeted Analysis of Propionic Acid

By

Dorothea Mung

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry

University of Alberta

# Abstract

Metabolomics is the study of all the small molecules in a given sample or organism. It is used to understand individual variations caused by the relationship between the individual's genetics, exposome and metabolic activity. There is interest in using metabolomics for personalized care leading to the merging of metabolomics with other fields such as nutritional sciences. These nutrimetabonomic studies focus on the individual's metabolic response to a food or diet. As a note, metabolomics aims to study the effects of a stimuli on a metabolome while metabolomics involves metabolome fingerprinting of various samples. They are considered interchangeable with the more common use of the word metabolomics to describe most studies.

Metabolomics can be used in a targeted or non-targeted manner. For non-targeted metabolomics, the entire set of metabolites is of interest. However, the metabolome is complex containing metabolites with various chemical properties. Profiling the metabolome with one analytical platform is impossible. However, the metabolome can be grouped by similar chemical properties (submetabolomes). Chemical isotope labeling (CIL) can be used to target submetabolomes by addition of an isotope tag to improve separation, ionization efficiency and provide the ability for relative or absolute quantification. In this way, the entire metabolome can be comprehensively studied using one or few analytical platforms such as positive mode reverse phase liquid chromatography mass spectrometry (RPLC-MS).

In this work, I used CIL LC-MS to profile the amine and phenol submetabolome of cow milk. Milk is one of the main food groups in North America and is consumed multiple times a day. I first determined a sample cleanup protocol involving protein precipitation with methanol and lipid removal with dichloromethane. Dansylation labeling was utilized to target the amine and phenol submetabolome of cow milk. The differences between brands, batches and fat percentages

of cow milk was investigated. Once the metabolome for cow milk was determined, the short-term effects of cow milk consumption on the human urine metabolome was investigated. Each individual was analyzed independently due to large inter-individual differences that masks metabolome changes when the volunteers were analyzed together. By comparing the urine dataset with the previously determined cow metabolome, subtle changes to the urine metabolome after milk consumption were able to be detected with confidence.

The ability of metabolomics to generate metabolome fingerprints can also be useful for quality control or product authentication purposes. Human milk is known as the best source of nutrition for an infant. However, not all mothers are able to produce milk. Human milk sharing has become popular due to easy access and low cost of online purchasing. With the high demand, there is potential for human milk adulteration via dilution with foreign milk in order to increase profits. I used CIL RPLC-MS to profile the metabolome of human milk along with 5 potential human milk adulterants (soy, goat, cow, almond, infant formula). Once the metabolome fingerprints were obtained, the human milk was investigated with varying levels of adulteration (5%, 10%, 25%, 50% and 75%). Overall, I was able to detect as low as 5% adulteration of human milk with foreign milk

When utilizing metabolomics in a targeted manner, a particular metabolite (i.e. propionic acid) or class of metabolites (i.e. short chain fatty acids) are of interest. Absolute quantification can be done for targeted metabolomics of one or a few metabolites if their respective standards are available. In this work, I used p-dimethyaminoacyl (DmPA) bromide to target and absolutely quantify propionic acid. The labeling method was optimized and a RPLC-MS method was tailored for the analysis of propionic acid.

# Preface

Chapter 2 has been published as "Development of Chemical Isotope Labeling LC-MS for Milk Metabolomics: Comprehensive and Quantitative Profiling of the Amine/Phenol Submetabolome" Mung, D.; Li, L. Anal. Chem. 2017, 89 (8), pp 4435-4443. I prepared all the samples, ran the samples, performed data processing and statistical analysis, generated the figures and wrote and edited the manuscript. Dr Liang Li was involved in the generation of the study along with the experimental design and the revision of the manuscript.

Chapter 3 and 5 were all reviewed by the University of Alberta health ethics review board and the study was approved with all participants providing informed consent. The samples used in chapter 5 were obtained from Dr. Juan Gonzalez Abraldes.

Chapter 4 has been submitted for publishing as "Development of a Quantitative Metabolomics Method for Detecting Potential Milk Adulterant in Human Milk" Mung, D.; Li, L. on August 26, 2017. I prepared all the samples, ran the samples, performed data processing and statistical analysis, generated the figures and wrote and edited the manuscript. Dr Liang Li was involved in the generation of the study along with the experimental design and the revision of the manuscript. The University of Alberta health ethics review board has approved this study and informed consent was provided by all participants.

All Supplemental Tables can be obtained from Professor Li's Lab. Contact Dr. Liang Li (liang.li@ualberta.ca) for information.

## Acknowledgements

I would like to give my utmost thanks to my supervisor, Dr. Liang Li, for his immense patience, support, professional and academic advice and guidance in revealing to me the world of metabolomics and LC-MS analysis. I am very grateful for the experience and knowledge gained from working in Dr. Li's lab these past 3 years. Thank you for the opportunities you've provided, the yearly Mass Spectrometry conferences and the encouragement for the acceptance of my oral presentation.

I would also like to thank my supervisory committee, Dr. Robert Campbell and Dr. Glen Loppnow for their academic advice and guidance. I would especially like to thank Dr. Campbell for taking the time to discuss academic and professional opportunities and the advice and reassurances received from him over these past years. I would also like to thank Dr. Jens Walter for participating in my oral defense and reviewing my thesis.

I would like to thank my group members for their unwavering encouragement, friendship and assistance over the years. I would like to thank the all the current members with special thanks to Shuang Zhao, Wei Han, Kevin Hooton, Yunong Li and Xian Luo for always extending a hand and open ear to help overcome the countless obstacles I've faced over the course of this degree. To all the past group members, Dr. Nan Wang, Jaspaul Tatlay, Dr. Zhendong Li, Dr. Tao Huan, Dr. Chiao-Li Tseng, thank you for your extreme patience and mentorship in my early years.

Finally, I would like to thank my family and friends for their support, reassurance and understanding through the course of my degree. The utmost thanks to the some of the most important people in my life, my parents Winnie and Malcolm Mung, along with my siblings, Vivien and Cory Ng, Manfred and Emilie Mung. Without all of you, I would not have been able

to get this far in my academic career. I owe it all to you and am forever grateful. My aunts, uncles and grandparents that have supported me near and overseas, I thank you. Lastly, thank you to my friends, Leonie Tiam, Amanda Tam, Fiona Nkala and Eye Vorapattanapong, for your unwavering support and keeping me sane throughout the stressful times. You have kept me grounded and stood by me from beginning to end. Finally, thank you to all those who have believed in me!

# Table of Contents

# List of Figures

# List of Abbreviations

| | |
|---|---|
| %RSD | Percent relative standard deviation |
| ACN | Acetonitrile |
| BPI | Base peak ion |
| CE | Capillary electrophoresis |
| CIL | Chemical Isotope Labeling |
| DCM | Dichloromethane |
| DHM | Donor human milk |
| DmPA | P-dimethylaminophenacyl (DmPA) |
| EML | Evidence-based Metabolome Library |
| ESI | Electrospray ionization |
| FTICR | Fourier Transform Ion Cyclotron Resonance |
| GC | Gas chromatography |
| HPLC | High performance liquid chromatography |
| HMDB | Human metabolome database |
| HMBANA | Human Milk Banking Association of North America |
| HILIC | Hydrophilic interactions liquid chromatography (HILIC) |
| LC | Liquid chromatography |
| MeOH | Methanol |
| Min | Minutes |
| MTBE | Methyl tert-butyl ether |
| MPA | Mobile phase A |
| MPB | Mobile phase B |

| | |
|---|---|
| MS | Mass spectrometry |
| NaOH | Sodium hydroxide |
| NICU | Neonatal Intensive Care Unit |
| NMR | Nuclear magnetic resonance spectroscopy |
| PA | Propionic Acid |
| PC1 | First principal component |
| PC2 | Second principal component |
| PCA | Principal component analysis |
| PDA | Photodiode array |
| PLS-DA | Partial least squares discriminant analysis |
| PRSL | Potential renal solute load |
| QTOF | Quadrupole time of flight |
| RP | Reversed phase |
| RT | Retention time |
| SCFA | Short chain fatty acids |
| TIC | Total ion count |
| TEOA | Triethanolamine |
| UHPLC | Ultra high-performance liquid chromatography |
| UPLC | Ultra performance liquid chromatography |
| USD | United States Dollar |
| UV | Ultraviolet |
| WHO | World Health Organization |

# List of Symbols

| | |
|---|---|
| °C | Degrees Celsius |
| µL | Microliter |
| µm | Micron |
| cm | Centimeter |
| Da | Dalton |
| M | Molar |
| mg/mL | Milligrams per milliliter |
| mm | Millimetre |
| mM | Millimolar |
| ppm | Parts per million |
| RPM | Revolutions per minute |
| S/N | Signal to noise ratio |
| v/v | Volume to volute ratio |

# Chapter 1 Introduction

## 1.1 Multi "omics"

Multi-cellular organisms are quite complex. Intricate interplay between various molecular networks and the adjustment of the system to perturbations to ensure physiological homeostasis allow for the proper functioning of the organism. Systems biology involves the study of the functional properties of various components of an organism at the molecular level.[1] This is made possible through the generation of "omics" technologies which aims to holistically study the different levels of "omes" which include the genome, transcriptome, proteome and metabolome.[2,3] At the top of the "omes" cascade is the genome which encompasses all the DNA within a cell or organism, essentially the genetic makeup or genotype.[4] However, the genotype does not always match/predict the phenotype due to other influencing factors such as the environment.[5] This lead to the study of the transcriptome. The transcriptome is the result of gene expression and refers to all the mRNA in a cell or organism.[4] Transcriptomics allows for the investigation of gene activity.[4] It links the genome to the proteome as the mRNAs are templates for protein synthesis through the process of translation.[4] Proteomics is the collective study of all the proteins in a biological sample. These proteins can be post-translationally modified and are involved in many protein-protein interactions, various pathways, signaling systems and molecular networks.[6] Due to these factors the transcriptome is not well correlated with the proteome.[7] The last of this cascade is metabolome which includes all the small molecules, <1500 Da, in a biological sample. These small molecules are the intermediates or products of metabolic pathways, the end products of genomics, transcriptomics, and proteomics. Each tier in the "omics" cascade adds unique information that is unable to be obtained from the upstream "omics".

## 1.2 Metabolomics

### 1.2.1 Advantages of Metabolomics

The most complex and dynamic of the "omes" is the metabolome. It is the final product of the "omes" cascade and is influenced by intrinsic genomic factors as well as external factors such as the exposome (the collective exposure an individual is subjected to throughout their lifetime) which includes the xenometabolome (exogenous compounds like drugs, chemical toxins or pollutants), diet, environmental conditions, gut microbiota, exercise and stress etc.[4] In this way, it is more reflective of the phenotype.[4] Metabolomics aims to quantitively profile the entire set of metabolites in a given sample at a given time. It also aims to investigate metabolome perturbations caused by external factors such as drugs, diet, environmental conditions and diseases. Depending on the biofluids studied, different information may be obtained. Blood collection is more invasive but can give a snapshot of the metabolome and physiological state of the organism at the time of collection whereas the less invasive urine gives insights of waste-end products from metabolic pathways.[8]

### 1.2.2 Platforms of Analysis for Metabolomics

For metabolomics, there are two main platforms used for analysis: Nuclear Magnetic Resonance (NMR) spectroscopy and Mass Spectrometry (MS). $^{1}$H NMR has been one of the very first platforms for metabolomics studies. This mode of analysis is able to analyze unprocessed samples and does not need chromatographic separation techniques for complex mixtures.[8,9] NMR is also gentle on the sample, the sample is not consumed and may be used for further analysis.[8,9] It is also quick, robust, reproducible and quantitative. As well, it allows for the detection and identification of compounds due to the structural information obtained from analysis.[8,9] The

drawbacks of $^{1}$H NMR analysis is the low sensitivity, typically only detecting metabolites as low as the micromolar range. [8,9] It also requires a larger sample volume and the spectral information may be hard to interpret for complex mixtures.[8,9]

Mass spectrometry (MS) is an analytical platform that provides information about the exact mass as well as fragmentation patterns for metabolites.[8] MS, unlike NMR, does require sample preparation and chromatographic separation (liquid or gas chromatography) for analysis of complex mixtures. With the need for derivatization and generally lower-throughput associated with gas chromatography (GC), liquid chromatography (LC-MS) has been the more popular combination of analytical platforms due to its high sensitivity, easier sample preparation, quick analysis and robustness.[10] With LC-MS, different column chemistries can be used for separation such as reversed phase liquid chromatography (RPLC) for separation of non-polar compounds and hydrophilic interactions liquid chromatography (HILIC) for analysis of polar compounds. As well, different ionization techniques can be utilized, such as electrospray ionization (ESI) which is more commonly used due to its softer ionization that does not result in fragmented metabolites.[11] Various mass analyzers can also be used for the detection of metabolites. Mass analyzers differ by sensitivity and resolution and range from simple MS to tandem MS/MS (for capturing structural information) capabilities. Some high-resolution mass analyzers include the Quadrupole Time of Flight (QTOF) and the Fourier Transform Ion Cyclotron Resonance (FTICR) mass spectrometer. In terms of mass accuracy and resolution, the FTICR is still the best, however the instrument has a slow scan rate and relatively low sensitivity.[12] QTOFs are hybrid mass spectrometers beneficial for both targeted and non-targeted analysis.[12] They have a fast scan rate, relatively high resolution, mass accuracy, sensitivity and a good dynamic range.[13]

### 1.2.3 Approaches for Metabolomics

Metabolomics can have differing approaches. It can be targeted; whereby the detection and quantitative analysis of a specific metabolite or class of metabolites is of interest. For targeted analysis, absolute quantification can be made possible for one or few metabolites if their standards are available. An clinical example of this is the quantitative analysis of vitamin D in blood samples.[14] Metabolomics can also be untargeted where as many metabolites are detected as possible for that sample (biofluid, tissue etc.). For untargeted work, metabolomics can be utilized to profile different types of samples to obtain metabolome fingerprints. The metabolome fingerprints can be used for comparisons such as diseased versus healthy samples. It can also be used to track changes to the metabolome over the progression of a disease/condition and be used to find biomarkers (for early and accurate diagnosis, for drug targets or drug therapies). As well, metabolomics can also be used clinically or toxicologically to study the effects of a toxin, vitamin, natural product, prescription or illicit drug on the individual's health or disease state.[11]

More recently, metabolomics for personalized care has become of interest as each individual is different (genetically and in terms of their exposome).[15] Population studies have shown large inter-individual metabolic differences and shown the need for personalized metabolomics as each individual may respond to the same factors differently.[8,16] Metabolomics and the study of the exposome/xenometabolome investigates the effects of various environmental exposures (including diet) to the individual's metabolome. Metabolomics can also be combined with nutritional sciences to generate what is known is nutrimetabonomics which studies the metabolic responses to a particular food or diet on an individual and how that may affect their health/disease status.[8]

In regards to food sciences, metabolomics is being used to characterize food metabolomes. The respective metabolome fingerprints can be used for quality control, food safety and product authentication purposes.[17,18] Milk metabolomics is a relatively new direction of study, with increasing numbers of publications every year. Milk metabolomics can range from using metabolomics in conjunction with agricultural and food sciences for quality assurance purposes to nutrimetabonomics whereby the effects of milk on the human body is studied. Many longitudinal population studies have been performed to assess the effects of milk consumption on the human body.[19,20] However there are limited studies involving the investigation of the effects of milk intake on the human body at the molecular level. The use of metabolomics in studying human breast milk is also of importance. There is great inter-individual variability in the milk produced by different individuals due to differences in genetics, diet, health state, social and economic status and overall lifestyle.[21] Interestingly enough, there is also large intra-individual variability in the milk produced by the same individual with differences being observed within the same feeding session.[21] Understanding the changes to human milk may give insight on the proper feeding managements for infants or the generation of similar infant formula products.[21]

In general metabolomics can be used in a variety of ways, linking various fields of studies including some of the following: food, agricultural, pharmacological, nutrition and life sciences.[11] Overall, many of these studies are performed in hopes of better understanding of the human metabolome and the biological processes associated with a healthy versus diseased state. This will hopefully allow for biomarkers to be discovered to assist with the early and accurate diagnosis of diseases and the development of accurate risk assessments associated with exposomic factors (food, drugs, chemicals) all working towards the final goal of utilizing metabolomics for personalized care in order to achieve or maintain a healthy state.

### 1.2.4 Workflow for Metabolomics

For an RPLC-MS metabolomics study, the sample of interest first needs to undergo sample preparation. Sample clean up may be required prior to analysis to avoid clogging the separation column and dirtying the mass spectrometer. To remove large particles such as cells or debris, centrifugation is performed to pellet out the large compounds and the supernatant is passed through a molecular weight cut off filter. To get rid of proteins, solvents such as acetonitrile, methanol or acetone is used to pellet out the proteins from the samples. If the sample also contains lipids, those are removed with dichloromethane or chloroform (or retained if lipids are of interest). Once sample preparation is complete the sample is normalized prior to LC-MS analysis. Some samples, such as urine, feces, and sweat have large concentration ranges. Urine can have large concentration differences because of factors such a water intake, dehydration due to fever/diarrhea, diet, exercise/activity and sweating.[22] It would be inaccurate to quantitively assess the sample based on similar injection volume. For example, one way to normalize urine is to normalize based on the amount of creatinine in the sample as there is a correlation between creatinine and the urine concentration.[22] Other methods can also be used such as measuring the ultraviolet (UV) absorbance, osmolality, specific gravity or using the total ion count (TIC).[22,23,24] After sample normalization, the sample is analyzed by LC-MS.

### 1.2.5 Chemical Isotope Labeling Metabolomics

One of the difficulties involved with studying the metabolome is the complexity due to the large diversity in physical and chemical properties of the metabolites.[4] There is no single technique or analytical platform that can account for all these metabolites. Rather, it would require the complementary use of many platforms to analyze the metabolome. In our lab, we utilize a “divide

and conquer" strategy whereby the metabolome is grouped into submetabolomes based on a shared chemical property. In this way, better coverage of the overall metabolome can be achieved. Chemical Isotope Labeling (CIL) is used to target a particular functional group (submetabolome) of interest. With CIL, a chemical reagent that has been differentially labeled (i.e. with $^{12}C$ or $^{13}C$), is used to react with a particular group of metabolites. By labeling an experimental sample with one isotope form and a reference sample with the other, after mixing and LC-MS analysis, peak pairs can be identified. A peak pair are two peaks (the lighter peak and the heavier peak) differing by the mass between the two isotope tags. The ratio between the two peaks can be used for relative or absolute quantification of metabolites. Two examples of labeling chemistries include dansylation used to target the amine and phenol submetabolome and the use of p-dimethylaminophenacyl (DmPA) bromide to react with organic acids.[25,26] There are many benefits to CIL including better retention and therefore separation of polar metabolites on a RP column, better ionization efficiency, the ability to detect "real" metabolites by means of peak pair picking and the ability for relative or absolute quantification from the peak ratios.[25]

### 1.3 Scope of the Thesis

The objective of this research is to develop CIL LC-MS techniques for untargeted and targeted metabolomics.

In chapter 2, CIL LC-MS is used to profile the cow milk amine and phenol submetabolome. A sample cleanup/metabolite extraction method is determined for fluid milk.

In chapter 3, CIL LC-MS is used to determine any short-term effects of cow milk consumption on the human urine metabolome. The data obtained from the urine analysis work is compared to the milk profiling data shown in chapter 2.

In chapter 4, CIL LC-MS is used to profile the milk metabolomes of human milk along with five potential human milk adulterants. It is then used to detect varying levels of human milk adulteration (5%, 10%, 25%, 50% and 75%).

Lastly in chapter 5, CIL LC-MS is used to determine a method for the absolute quantification of propionic acid in plasma samples.

## 1.4 Literature Cited


(1) Rotroff, D. M.; Motsinger-Reif, A. A. *Int. J. Genomics* **2016**, *2016*, 1-5.

(2) Franklin, S.; Vondriska, T. M. *Circ. Cardiovasc. Genet.* **2011**, *4*, 576.

(3) Rochfort, S. *J. Nat. Prod.* **2005**, *68*, 1813–1820.

(4) Horgan, R. P.; Kenny, L. C. *Obstet. Gynaecol.* **2011**, *13*, 189–195.

(5) Gagneur, J.; Stegle, O.; Zhu, C.; Jakob, P.; Tekkedil, M. M.; Aiyar, R. S.; Schuon, A.-K.; Pe'er, D.; Steinmetz, L. M. *PLoS Genet.* **2013**, *9*, e1003803.

(6) Wang, K.; Huang, C.; Nice, E. *Biomed. Chromatogr.* **2014**, *28*, 848–857.

(7) Zhang, H. M.; Nan, Z. R.; Hui, G. Q.; Liu, X. H.; Sun, Y. *Genet. Mol. Res.* **2014**, *13*, 198–204.

(8) Claus, S. P.; Swann, J. R. *Annu. Rev. Food Sci. Technol.* **2013**, *4*, 381–399.

(9) Lenz, E. M.; Wilson, I. D. *Journal of Proteome Research* **2007,** *6*, 443–458.

(10) Fang, Z.-Z.; Gonzalez, F. J. *Arch. Toxicol.* **2014**, *88*, 1491–1502.

(11) Gowda, G. A. N.; Djukovic, D. *Methods Mol. Biol.* **2014**, *1198*, 3–12.

(12) Forcisi, S.; Moritz, F.; Kanawati, B.; Tziotis, D.; Lehmann, R.; Schmitt-Kopplin, P. *J. Chromatogr. A* **2013**, *1292*, 51–65.

(13) Binkley, J. *Chromatogr. Online* **2010**, 1–7.

(14) Crutchfield, C. A.; Thomas, S. N.; Sokoll, L. J.; Chan, D. W. *Clin. Proteomics* **2016**, *13*, 1.

(15) Ahmed, M. U.; Saaem, I.; Wu, P. C.; Brown, A. S. *Crit. Rev. Biotechnol.* **2014**, *34*, 180–196.

(16) Dinis-Oliveira, R. J. *Drug Metab. Rev.* **2016**, *48*, 568–576.

(17) Cevallos-Cevallos, J. M.; Reyes-De-Corcuera, J. I. In *Advances in food and nutrition research* **2012**, *67*, 1–24.

(18) Putri, S. P.; Nakayama, Y.; Matsuda, F.; Uchikata, T.; Kobayashi, S.; Matsubara, A.; Fukusaki, E. *J. Biosci. Bioeng.* **2013**, *115*, 579–589.

(19) Rozenberg, S.; Body, J.-J.; Bruyère, O.; Bergmann, P.; Brandi, M. L.; Cooper, C.; Devogelaer, J.-P.; Gielen, E.; Goemaere, S.; Kaufman, J.-M.; Rizzoli, R.; Reginster, J.-Y. *Calcif. Tissue Int.* **2016**, *98*, 1–17.

(20) Okada, T. *Am. J. Clin. Nutr.* **2004**, *80*, 1088-9-90.

(21) Cesare Marincola, F.; Dessì, A.; Corbu, S.; Reali, A.; Fanos, V. *Clin. Chim. Acta* **2015**, *451*, 103–106.

(22) Wu, Y.; Li, L. *J. Chromatogr. A* **2016**, *1430*, 80–95.

(23) Zhang, T.; Watson, D. G. *Analyst* **2015**, *140*, 2907–2915.

(24) Wu, Y.; Li, L. *Anal. Chem.* **2014**, *86*, 9428–9433.

(25) Guo, K.; Li, L. *Anal. Chem.* **2009**, *81*, 3919–3932.

(26) Guo, K.; Li, L. *Anal. Chem.* **2010**, *82*, 8789–8793.

# Chapter 2 Chemical Isotope Labeling (CIL) Liquid Chromatography Mass Spectrometry (LC-MS) for Milk Metabolomics: Comprehensive and Quantitative Profiling of the Amine/Phenol Submetabolome

## 2.1 Introduction

In recent years, metabolomics has been increasingly used for nutritional studies ranging from studies on the relations of diet, health and diseases to the development of improved nutritional products.[1-4] Nutritional metabolomics involves characterizing the metabolomes of foods and diets and investigating their effects on the metabolomic profiles of a particular organism. Milk is an important component of many diets. In North America, dairy, which incorporates milk and all milk products, is one of the four main food groups. According to Statistics Canada (www.dairyinfo.gc.ca), individuals in North America consumed an average of 70 L of fluid milk per person in 2014. European countries (Finland, Ireland, UK) and Oceanian countries (Australia, New Zealand) consumed upwards of 100 L per person in the same year. In all species of mammals, milk has great importance as it is the first and only source of nutrition immediately after birth.[5] Milk is comprised of a variety of proteins, peptides, essential/free amino acids, fats, vitamins and minerals, making it a balanced source of nutrition for proper growth and development through all stages of life.[5]

Compositional analysis of milk and milk products is critical in many areas of applications including characterization of nutrients,[6] determination of health benefits of dairy consumption,[7] developing new processing technologies[8] and safety and quality control of dairy products.[9, 10]

*A version of this has been published as "Development of Chemical Isotope Labeling LC-MS for Milk Metabolomics: Comprehensive and Quantitative Profiling of the Amine/Phenol Submetabolome" Mung, D.; Li, L. Anal. Chem. 2017, 89 (8), pp 4435-4443. I prepared all the samples, ran the samples, performed data processing and statistical analysis, generated the figures and wrote and edited the manuscript.

Current analytical methods are mainly focused on the measurement of the high abundant components including proteins, lipids and oligosaccharides, often expressed as the total concentrations of individual groups.[11] Targeted analysis of a small number of compounds such as vitamins is also carried out using a variety of analytical tools.[12, 13] Untargeted analysis of milk and milk products at the individual molecular level using a metabolomics approach represents a relatively new development in compositional analysis.[1, 2] Several studies on milk metabolomics have been reported using NMR,[14-21] GC-MS[14, 15, 22, 23] and LC-MS[15, 16, 20, 23] with varying degrees of metabolomic coverage. Recent reports from Villasenor et al[23] and Andreas et al[15] demonstrated the significant advances in analytical methods for milk metabolome profiling, which also highlighted the current analytical challenges that include the analysis of polar metabolites in milk. NMR and capillary electrophoresis (CE) MS were found to be more suitable for polar metabolite analysis, while GC-MS and LC-MS were particularly useful for analyzing non-polar metabolites such as lipids.[15]

In this work, we report a high-performance chemical isotope labeling (CIL) LC-MS method for profiling the relatively polar metabolites in milk with very high coverage. CIL LC-MS is a general strategy of using chemical labeling to improve separation, detection and quantification of metabolites.[24] CIL targets a particular submetabolome based on a shared chemical group. For example, dansylation labeling has been shown to be effective in analysing the amine/phenol submetabolome.[24] Different labeling chemistries can be used to profile different submetabolomes in order to provide better coverage of the entire metabolome of interest.[25-31] Because of the complexity of the milk matrix, we have examined a number of sample preparation conditions to develop an optimal workflow. In this report, we first describe the workflow for analyzing polar metabolites in milk using dansylation LC-MS. We then apply this method for profiling different

groups (brands, batches and fat percentages) of milk to demonstrate the overall analytical performance and showcase the potential utilities of the method in milk metabolomics.

## 2.2 Experimental Section

### 2.2.1 Milk Sample Collection and Preparation

All cow and goat milk were purchased from a local grocery store and the human breast milk was donated from a healthy volunteer. Ethics approval for working with human samples was given by the University of Alberta Ethics Committee. Three brands of milk (Beatrice, Dairyland and Lucerne) along with each of the different fat percentages of the milk (skim, 1%, 2% and 3.25%) were purchased at three time points, i.e., in January, March and April of 2015. All milk samples were aliquoted into 1.5 mL vials and stored in a -80 °C freezer until use. 100 µL of milk was mixed with 300 µL of methanol, vortexed and incubated at -20 °C for 15 min. The sample was centrifuged at 14 000 rpm for 15 min and 350 µL of the supernatant was taken into a new vial. 500 µL of dichloromethane and 50 µL of water were added to the supernatant, vortexed and left at room temperature for 10 min. 350 µL of the aqueous layer was then taken into a new vial and dried down using a SpeedVac. The sample was re-dissolved in 50 µL of water for dansylation labeling.

### 2.2.2 Dansylation Protocol

Dansyl chloride was used as the labeling reagent to react mainly amine- and phenol-containing metabolites to form dansyl-amine or dansyl-phenol derivative.[24] The reactivity, specificity and type of reaction products generated have been described previously.[24, 32] 50 µL of a processed milk sample was mixed with 25 µL of 250 mM sodium bicarbonate buffer and 25 µL of acetonitrile and the solution was vortexed. 50 µL of 18 mg/mL $^{12}$C- or $^{13}$C-dansyl chloride in

acetonitrile was added, vortexed and incubated at 40 °C for 45 min. To quench the excess dansyl chloride, 10 µL of 250 mM sodium hydroxide was added and the solution was incubated at 40 °C for 10 min. Finally, 50 µL of 425 mM formic acid was added to acidify the sample. Individual samples were labeled with $^{12}$C-dansyl chloride and a pooled sample was labeled with $^{13}$C-danysl chloride. We note that for labeling amines and phenols, the reaction was carried out in an aqueous solution. Under anhydrous system, many metabolites such as some amino acids cannot be dissolved and thus not labeled by dansyl chloride. Reagent hydrolysis took place during the labeling, but slowly under the slightly basic buffer condition. It was only when we added NaOH that the excess reagent at the end of labeling reaction was consumed quickly.

### 2.2.3 LC-UV Quantification

For LC-UV, a Waters ACQUITY UPLC system with a photodiode array (PDA) detector was used for the quantification of dansyl labeled metabolites for sample amount normalization as described earlier.[33] Briefly, 4 µL of each labeled sample was injected onto a Phenomenex Kinetex C18 column (2.1 mm × 5 cm, 1.7 µm particle size) for a fast step-gradient run. Solvent A was 0.1% (v/v) formic acid in 5% (v/v) ACN, and solvent B was 0.1% (v/v) formic acid in ACN. The gradient started with 0% B for 1 min and was increased to 95% within 0.01 min and held at 95% B for 1 min to ensure complete elution of all labeled metabolites. The flow rate used was 0.45 mL/min. The peak area related to the total labeled metabolite concentration in the sample was integrated using the Empower software (6.00.2154.003). Based on the quantification results, the $^{12}$C-labeled sample and the $^{13}$C-labeled pool were mixed in equal amounts.

### 2.2.4 LC-MS

All LC-MS experiments were performed on an Agilent 1100 HPLC system (Palo Alto, CA) connected to a Bruker Impact HD quadrupole time-of-flight (QTOF) mass spectrometer (Billercia, MA) with an ESI source.

### 2.2.5 Data Analysis

Bruker DataAnalysis software 4.2 was used to extract MS spectral peaks. An in-house software tool, IsoMS, was used to process the raw data generated from multiple LC-MS runs by peak picking, peak pairing, peak-pair filtering to remove redundant peaks of the same metabolite, such as adduct ions, dimer, multimers, to retain only $[M+H]^+$ for a labeled metabolite (i.e., one peak pair corresponds to one metabolite), and peak-pair intensity ratio calculation.[34] IsoMS does not pick up isotopmers of a molecular ion as a peak pair as their intensity pattern is different from that of the labeled peak pair.[34] The same peak pairs detected from multiple samples were then aligned to produce a CSV file containing the metabolite information and peak ratios relative to a control (i.e., a pooled sample). A zero-fill program was used to find missing peak pairs from the raw mass spectral data, filling in the missing values.[35] Finally, peak ratios were re-calculated by using chromatographic peaks, instead of mass spectral peaks, with IsoMS-Quant.[36] The final metabolite-intensity data file was then exported to SIMCA-P+ 12.0 software (Umetrics, Umea, Sweden) for multivariate statistical analysis. Principal component analysis (PCA) and Partial Least Squares discriminant analysis (PLS-DA) were used to analyze the data after unit variance scaling.

Positive metabolite identification was performed based on mass and retention time matching to the dansyl standards library consisting of 273 unique standards with 315 entries.[32] This library with information on MS, MS/MS and ion chromatogram for each dansyl labeled

standard is freely accessible at www.MyCompoundID.org. Putative identification was done based on accurate mass match to the metabolites in the human metabolome database (HMDB) (8,021 known human endogenous metabolites) and the Evidence-based Metabolome Library (EML) (375,809 predicted human metabolites with one reaction) using MyCompoundID.[37] The mass accuracy tolerance window was set at 5 ppm and the retention time tolerance window set to 30 s for the searches. Since we were interested in the endogenous metabolites in milk, we did not use other databases such as Metlin and KEGG containing many exogenous compounds for our search. We also note that, for multi-function compounds (e.g., containing two amines), labeling by one or more reagent molecules to generate multiple products may happen to some metabolites, although in most cases complete labeling was found. If multiple products were found from one metabolite, they could be readily spotted in the final list of significant metabolites for differentiating different groups: they would be matched to the same metabolite.

### 2.3 Results and Discussion

#### 2.3.1 Removal of Proteins and Lipids

Milk is composed of a variety of proteins and lipids that may interfere with the chemical labeling procedure and LC-MS tailored to the analysis of polar metabolites. Based on our experience of working with serum and plasma, as well as other reported protocols for protein and lipid removal,[38] we have examined several experimental conditions to extract metabolites from milk. All method development experiments were chosen with due consideration of speed and convenience in sample handling. For protein precipitation, three solvents were compared which included acetone, acetonitrile and methanol at varying solvent-to-milk (v/v) ratios (3:1 and 5:1). The different solutions underwent the same sample work-up for incubation, centrifugation and the removal of the aqueous layer in quadruplicates. Both the aqueous layer and the pellet were dried

down using a SpeedVac, after which only the aqueous layer was re-dissolved in 50 µL of water. Three of the four replicates for each sample were labeled with $^{12}$C-dansyl chloride and the remaining sample with $^{13}$C-dansyl chloride. All samples were normalized using the concentration values determined by LC-UV, mixed at a 1:1 molar ratio of $^{12}$C- and its corresponding $^{13}$C-dansylated sample and injected into the LC-MS. Two measured properties were used to determine solvent extraction efficiency: the weight of the protein pellet and the corresponding number of peak pairs or metabolites detected from the aqueous layer.

Figure 2.1A and 2.1B show the plots of protein weight and peak pair number, respectively, for different protein precipitation conditions. The heavier the protein pellet (corresponding to better protein precipitation efficiency) the lower the number of peak pairs detected. Methanol at a 3:1 (v/v) ratio corresponded to the lightest protein pellet observed but also the highest number of metabolites detected. In addition, we have compared the absolute intensities of individual metabolites using a heat map (see Appendix Figure A2.1), which also indicated that 3:1 methanol gave a visually higher overall intensity map. Thus, methanol at a 3:1 solvent-to-sample ratio was chosen for protein precipitation as it resulted in the highest number of metabolites detected and was sufficient in protein removal. Note that, after methanol precipitation, some proteins may remain in the sample. However, during dansylation labeling, we observed some precipitates at the start of the reaction. This was from the dansyl labeling of multiple amine/phenol groups in the remaining proteins, causing precipitation of the labeled proteins out of solution.[39] In LC-MS analysis of labeled metabolites, we did not detect any proteins or large peptides. In addition, proteins would retain strongly on a C18 column, causing graduate increase in LC column pressure after running samples containing an appreciable amount of proteins; however, we did not observe any column pressure change after running hundreds of samples on a C18 column.

Due to the presence of a relatively large amount of lipids in milk, lipid removal was found to be essential to reduce column-washing time significantly and prolong the reversed phase (RP) column lifetime. Two extraction solvents, dichloromethane (DCM) and methyl tert-butyl ether (MTBE), were compared and the results are shown in Figure 2.1C. Overall, the addition of the lipid removal step resulted in a decrease in the number of peak pairs detected (e.g., from about 1100 peak pairs in methanol without lipid extraction to 870 peak pairs with lipid extraction), likely due to the loss of some hydrophobic metabolites during the lipid extraction process. However, the need for lipid removal is essential especially for the analysis of milk containing higher fat content. Of the two lipid-extraction solvents tested, dichloromethane resulted in the detection of higher numbers of metabolites when performed in addition to protein precipitation using the three protein precipitation solvents. DCM was also found to be more convenient during the removal of the aqueous layer for analysis due to the large density difference between this solvent and water/methanol. In addition, from the heat map of the absolute intensities of individual metabolites (see Appendix Figure A2.2), DCM combined with methanol gave a visually higher overall intensity map.

Based on the above results, we concluded that the optimal sample preparation for the dansylation LC-MS workflow was to use methanol at a 3:1 ratio of solvent to milk for protein precipitation and dichloromethane for lipid removal.

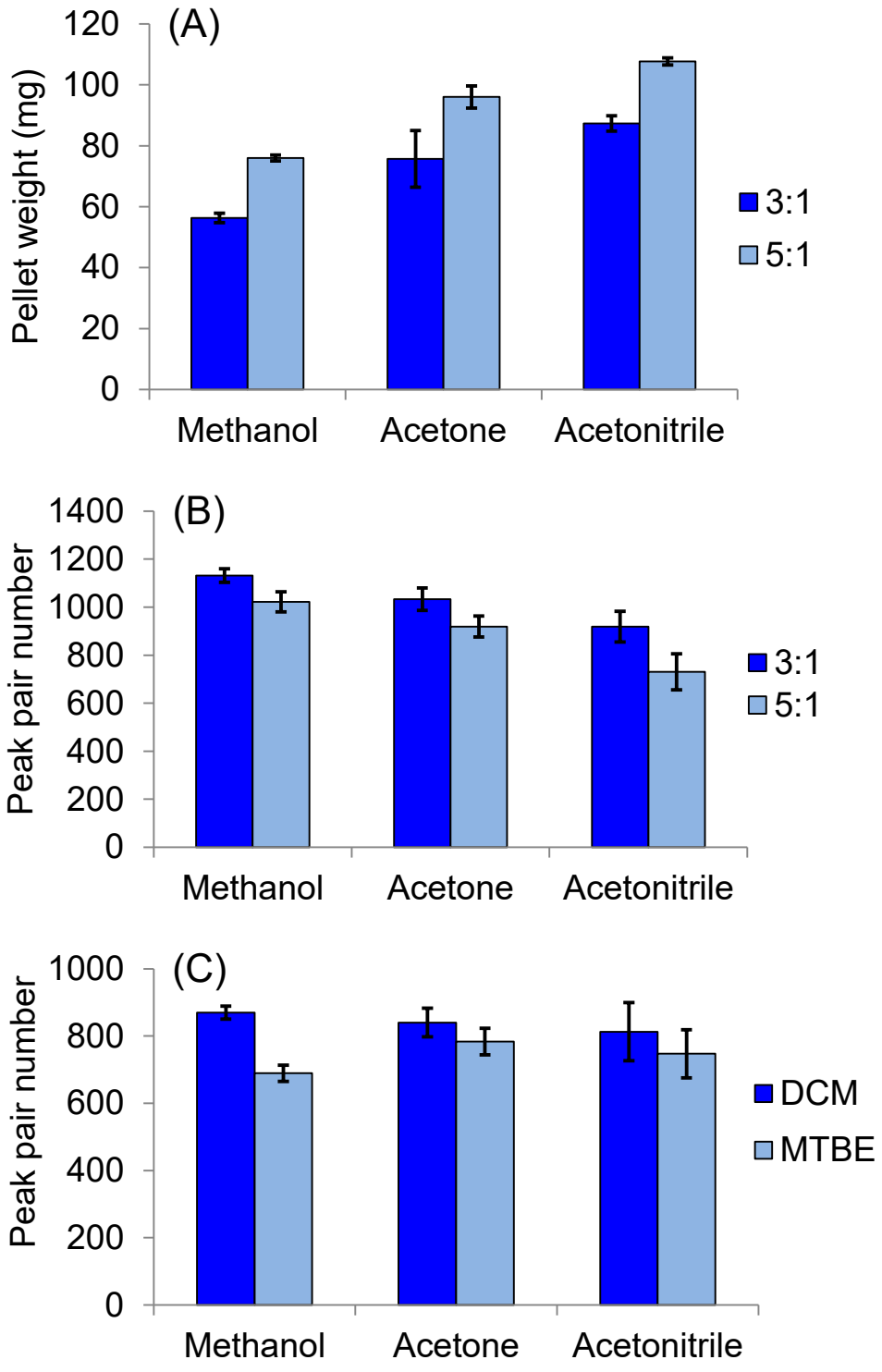



Figure 2.1 Method comparisons for protein removal: (A) protein pellet weight and (B) peak pair number detected as a function of solvent type and solvent-to-sample ratio. Method comparisons for lipid removal: (C) peak pair number detected as a function of protein-removal solvent and lipid-removal solvent types. Error bars correspond to standard deviation (n=3).

### 2.3.2 Milk Metabolite Detection

After removing the proteins and lipids, the aqueous solution was taken for dansyl labeling. Dansyl labeling adds a hydrophobic tag to a metabolite and, as a result, the labeled product becomes relatively hydrophobic even for very polar or ionic metabolites such as amino acids. Thus, the labeled milk metabolites can be readily retained and efficiently separated on a reversed phase column. Figure 2.2A shows a representative base-peak-ion (BPI) chromatogram of 1:1 $^{12}$C-/$^{13}$C-dansyl labeled milk. Many chromatographic peaks distributed along the gradient elution window were observed, indicating the complexity of the milk metabolome with diverse chemical and physical properties. The analytical separation takes 32 min, followed by 15-min column washing and re-equilibration. This gives an overall sample throughput of 47 min per sample. The washing/equilibration step was necessary to ensure that small molecule residues in the column were washed out and the column was properly equilibrated before the next run.

In LC-MS operation, using an optimal sample injection amount is critical in order to detect the maximum number of labeled metabolites. With dansyl labeling, the total concentration of labeled metabolites in each sample can be measured by LC-UV as described in the Experimental Section. Knowing the concentration of each sample, the exact amount of sample injected into LC-MS can be controlled. To determine the optimal injection amount, a 1:1 $^{12}$C-/$^{13}$C-labeled milk sample with a known concentration measured by LC-UV was injected in 10 µL increments from 10 to 60 µL. Experimental triplicate runs were performed for gauging the technical reproducibility. Figure 2.2B shows the plot of the peak pair number detected as a function of sample amount injected. Peak pair number saturation occurred when 20.9 nmol of sample in 30 µL was injected. Thus, in subsequent experiments, we injected 20 nmol of labeled sample into the LC-MS for analysis.

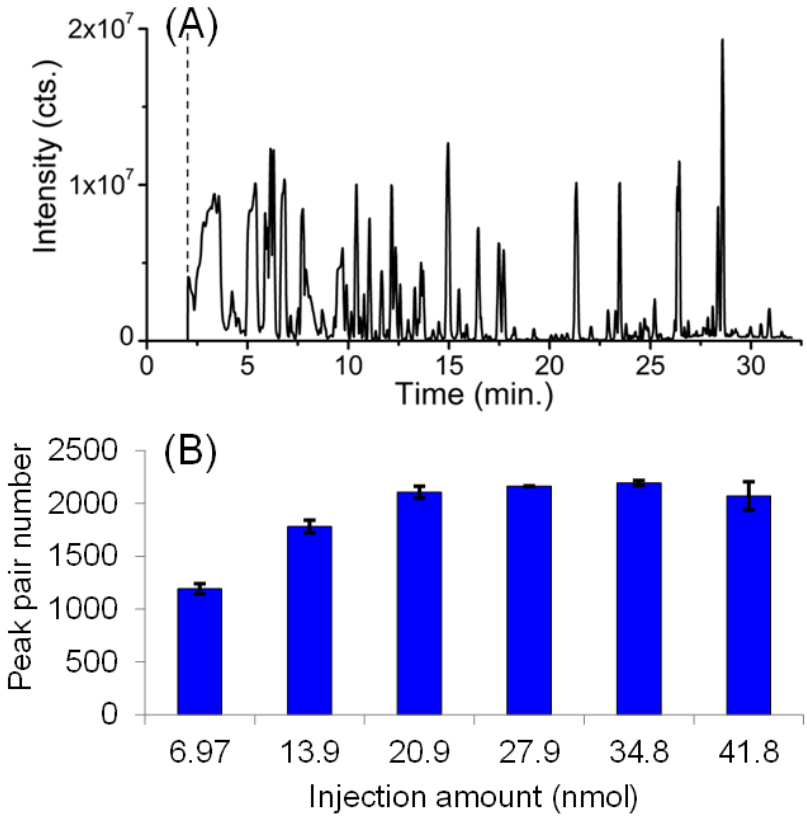



Figure 2.2 (A) Base peak ion chromatogram of a 1:1 12C-/13C-dansyl labeled milk sample and (B) peak pair number detected as a function of the injection amount (6 x n=3). The number of peak pairs was determined within the data set for injection-amount optimization data set alone.

Note that the number of peak pairs shown in Figure 2.2B is different from those shown in Figure 2.1 where less than the optimal amount of sample was injected. In addition, the number of peak pairs detected is dependent on the number of samples analyzed. Each sample adds some unique metabolites to the total number of peak pairs within a data set. An increase in the total number of metabolites will also translate to an increase in the number of metabolites detected in each individual sample run as the entire data set is used for peak picking and filling of missing ratio values during Zero-fill.[34, 35] For the initial peak-pair picking, we used a signal-to-noise (S/N)

threshold of 10 in IsoMS to minimize false positives. For Zero-fill, we considered all the peak pairs with S/N of greater than 3, as filling the peak pair with lower S/N in a sample initially with a missing ratio was judged based on the presence of a true peak pair in other sample(s). In the following studies investigating the different aspects of the analytical work and application, the number of samples or runs used for each data set was different and thus the numbers of peak pairs detected varied.

### 2.3.3 Accuracy and Precision

In the dansylation isotope labeling LC-MS workflow, relative quantification of a metabolite is based on the measured peak ratio between the $^{12}$C-labeled metabolite in an individual sample and its $^{13}$C-labeled counterpart in the pooled sample. To gauge the overall accuracy and precision of the workflow for profiling the amine/phenol submetabolome, experimental five-replicates (n=5) analysis of 1:1, 2:1 and 1:2 $^{12}$C-/$^{13}$C-labeled cow milk was used as an example. Six extractions were performed in parallel, five of which were dansylated with $^{12}$C-dansyl chloride and the sixth labeled with $^{13}$C-dansyl chloride.

Figure 2.3A-C shows the plots of the peak pair number as a function of the measured peak ratio. A total of 2245 common peak pairs were detected in the data set of these three mixtures. For the 1:1 mixture shown in Figure 2.3A, 1998 (88.9%) pairs gave peak ratios in the range of 0.8 and 1.25 or within ±25% of the expected ratio of 1.0. Only 98 pairs (4.3%) had peak ratios that exceeded the ±50% accuracy range (0.67>peak ratio>1.5). For the 2:1 mixture, Figure 3B shows 1906 (84.8%) peak pairs in the range of 1.6 and 2.5 (i.e., within ±25% of the expected ratio of 2.0). 130 pairs (5.7%) were found outside the 50% range of the expected peak ratio of 2.0. Lastly, the peak pair distribution shown in Figure 2.3C for the 1:2 mixture indicates 1932 pairs (86.0%)

found within ±25% of the expected ratio of 0.5. 153 pairs (6.8%) were outside the 50% accuracy range. These examples illustrate that accurate measurement of the peak ratio can be achieved using the workflow for majority of the peak pairs. This is not surprising considering that the individual samples and the pooled sample had a similar matrix and metabolite composition. The samples were also processed in parallel using the same sample work-up protocol, and analyzed as a $^{12}$C-/$^{13}$C-mixture in LC-MS where the matrix, ion suppression and instrumental effects were taken into account by measuring the peak ratio relative to the $^{13}$C-pool.

Figure 2.3D-F shows the plots of the peak pair number as a function of the relative standard deviation (RSD) of the measured peak ratio values from the $^{12}$C-/$^{13}$C-mixture samples (n=5). Out of the 2245 common peak pairs, 2037 (90.7%), 2081 (92.6%) and 2040 (90.8%) peak pairs found in 1:1, 2:1 and 1:2 mixtures, respectively, had RSD values of less than 20%. These results indicate that reproducible peak ratio measurements can be achieved.

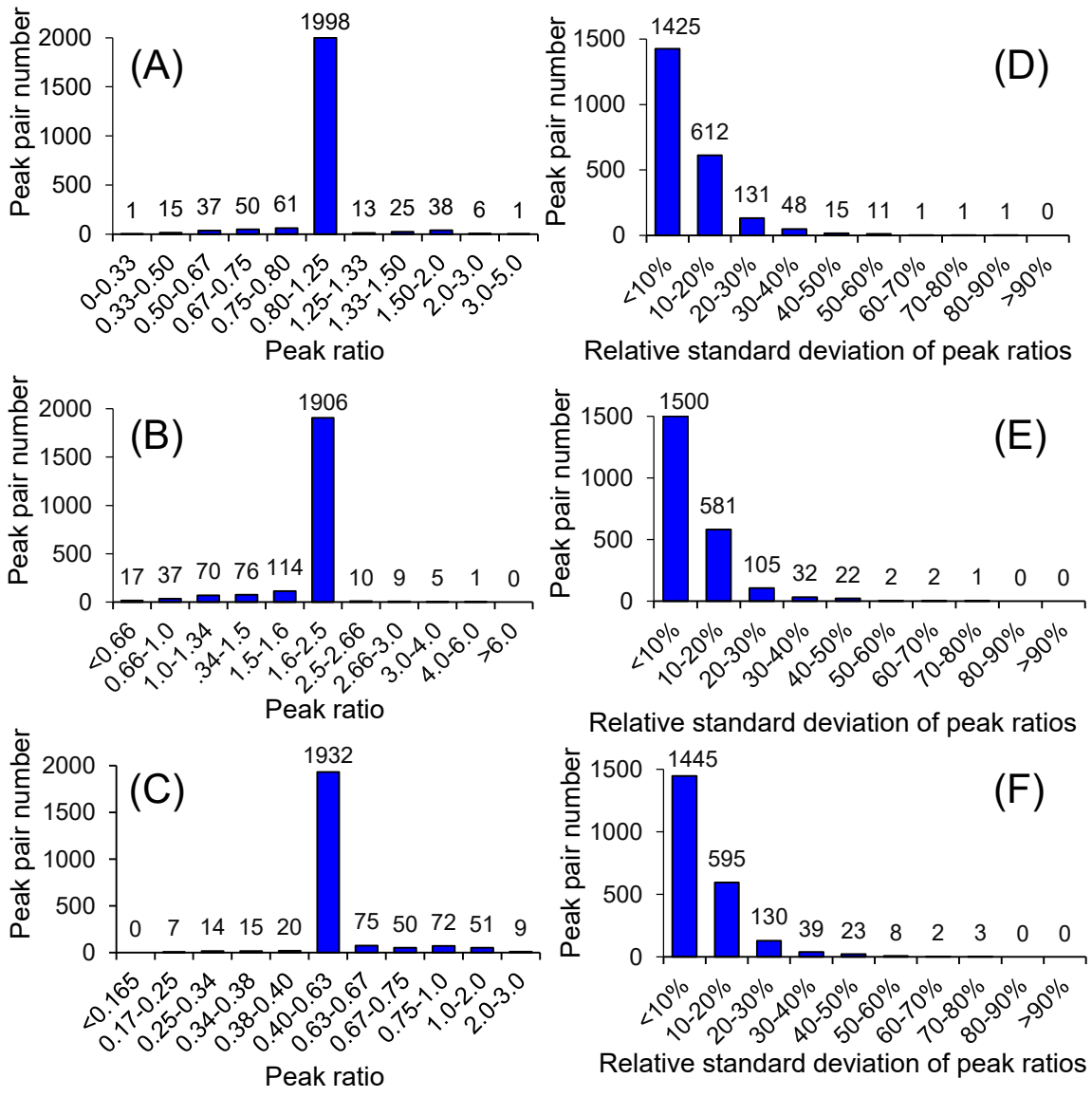



Figure 2.3 Accuracy comparison using the distribution of peak pairs as a function of the peak ratios for (A) 1:1, (B) 2:1 and (C) 1:2 mixtures of five 12C-/13C-dansyl labeled milk samples (n=5). Precision comparison using the distribution of peak pairs as a function of the relative standard deviation of peak ratios for (D) 1:1, (E) 2:1, and (F) 1:2 mixtures (n=5). The number of peak pairs was determined within the dataset of studying accuracy and precision alone**.**

### 2.3.4 General Applicability

To gauge the general applicability of the sample preparation and LC-MS method for milk metabolomics, different milk samples including goat milk, human milk, and cow milk with and without heat treatment were analyzed. For the heat treatment, two types of cow milk, commonly called homo-milk (homogenized milk containing 3.25% fat) and skim milk (0% fat) in the US and Canada, were brought to a boil, heated for 5 minutes, and then cooled down. A pooled sample was prepared by taking an aliquot from each sample and mixing them. After $^{12}$C-dansyl labeling of the individual samples and $^{13}$C-labeling of the pooled sample, LC-UV was used to measure the total concentration of labeled metabolites in all labeled samples. Figure 2.4A shows the average concentration of labeled metabolites in different samples. Goat milk had significantly higher concentrations than human and cow milk. Thus, for metabolome comparison of different species of milk, sample normalization is necessary. In our work, we took an aliquot of appropriate volume based on the measured concentration to ensure the same molar amount of all the individual samples was mixed with an equal mole of the $^{13}$C-labeled pool. The resultant 1:1 mixture was analyzed by LC-MS. Peak pair ratio values of individual metabolites determined from all the samples reflect the relative concentration differences of these metabolites in different samples. For each sample, experimental triplicate analyses were performed.

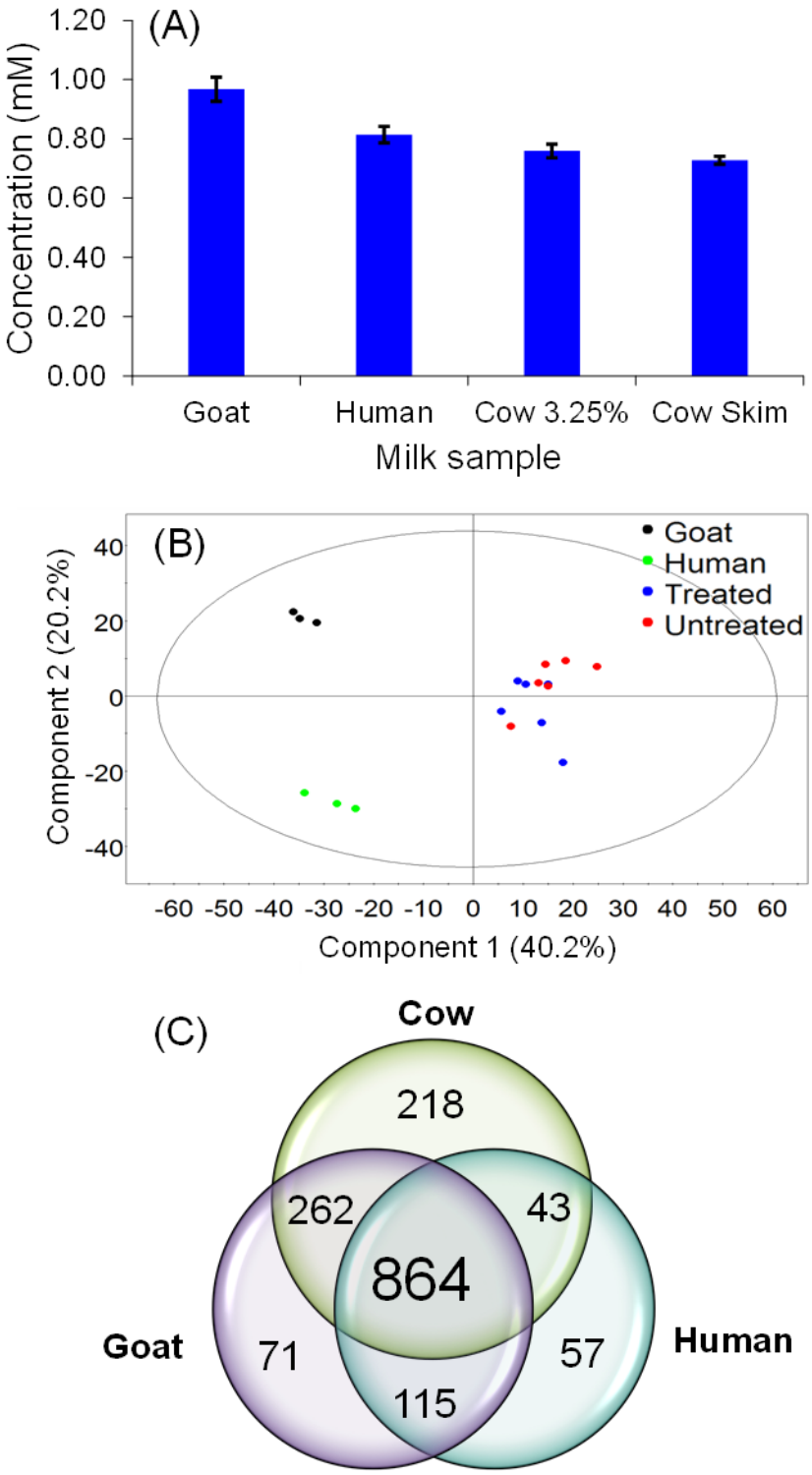



Figure 2.4 (A) Average concentration of labeled metabolites from various types of milk (n=3), (B) PCA plot of inter-species and intra-species mammalian milk including goat milk, cow milk (heat treated vs non-heat treated) and human milk, and (C) metabolite distribution of the different milk.

Figure 2.4B shows the PCA plot of the metabolomic data generated from these samples analyzed in one batch. The triplicate data points within a sample cluster together tightly, indicating that good reproducibility was achieved. Separation of the milk samples from three different species is much greater than the separation between the heat-treated and untreated cow milk. The inter-species difference is also reflected in Figure 2.4C where the numbers of peak pairs detected from goat, human and untreated cow milk are compared. The average number of peak pairs and the number of common peak pairs found in >50% of the experimental triplicates were 1290±14 (n=3) and 1312 pairs for goat milk, 1074±34 (n=3) and 1079 pairs for human milk, and 1398±16 (n=6) and 1387 pairs for the untreated cow milk. As Figure 2.4C shows, 864 peak pairs were commonly detected in the milk samples of the three species. We note that in working with the milk samples of different species, we did not observe any abnormity in preparing these samples including the cow milk with different fat contents (see below).

Out of 1312 peak pairs detected in goat milk, 69 pairs could be positively identified based on accurate mass and retention time matches to the dansyl standards library (Supplemental Table S2.1A) and 1130 pairs could be matched based on accurate mass searching against the HMDB and EML libraries (Supplemental Table S2.1B). For the 1079 human breast milk metabolites detected, 61 were positively identified by the dansyl-library (Supplemental Table S2.2A) and 963 were putatively identified by HMDB and EML (Supplemental Table S2.2B). For the 1387 cow milk metabolites detected, 68 were positively identified (Supplemental Table S2.3A) and 1230 were putatively identified (Supplemental Table S2.3B). It is interesting to note that many common peak pairs or metabolites were detected in goat, cow and human milk. Among the 864 common pairs detected in the three species, 60 metabolites including many amino acids were positively identified (Supplemental Table S2.4). We did not carry out MS/MS analysis of the labeled metabolites as

the MS/MS spectra of dansyl labeled metabolites usually do not provide adequate fragment ion information from the metabolite moiety for unknown metabolite identification.

The above results indicate that the workflow described herein should be generally applicable for handling different types of milk samples in milk metabolomics. It should be useful for future studies of comparing milk metabolomes of different species using a large cohort of samples in both hypothesis-generating and hypothesis-driving studies (e.g., how nutritional values are related to the metabolome composition). In this regard, applying other labeling chemistries targeting different chemical-group-submetabolomes (e.g., carboxyl submetabolome[25] and hydroxyl/thiol submetabolome[40]) will also increase the overall metabolome coverage for these studies.

### 2.3.5 Metabolomic Profiling of Different Cow Milk

To demonstrate the applicability of the workflow for milk metabolomics, we examined the metabolome profile differences among different groups of cow milk samples. This profiling work involved the analysis of three batches of milk with each batch consisting of 12 different milk samples from three brands (Dairyland, Lucerne and Beatrice) and four fat percentages (skim, 1%, 2% and 3.25%). Experimental triplicates were performed for each sample. In total, 108 labeled samples were analyzed by LC-MS.

From the 108 sample runs, a total of 7104 peak pairs were detected with an average of 4573±505 peak pairs (n=108) per sample. To determine how consistent the same peak pairs could be detected across different samples, Figure 2.5 shows the plot of the number of peak pairs and the percentage of common peak pairs as a function of the number of samples. A total of 2189 pairs were commonly detected in all the samples, while 3820 pairs were detected in more than 80% of

the samples. It is clear that a large number of metabolites could be consistently detected and quantified in the milk amine/phenol submetabolome profiling by the dansylation LC-MS method.

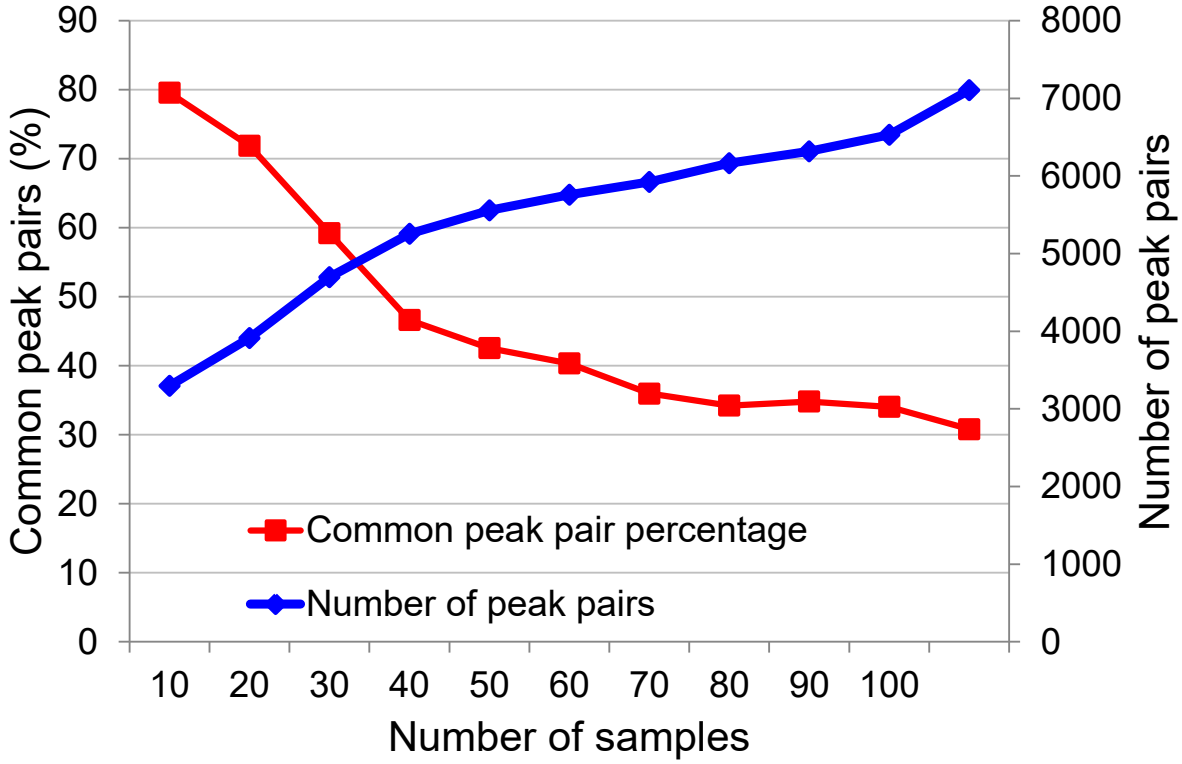



Figure 2.5 The inverse relationship between the percentages of common peak pairs detected and the total peak pair numbers generated by processing various numbers of samples (5, 10, 15, etc.) from the 108-sample milk profiling data sets.

The 3820 peak pairs commonly detectable in more than 80% of the samples were searched against the dansyl library, by which 70 metabolites were positively identified (Supplemental Table S2.5A). By using accurate mass search against HMDB and EML libraries, 954 and 2987 peak pairs were matched to one or a few chemical structures, respectively (Supplemental Tables S2.5B and S2.5C). It should be noted that the cow milk metabolites were searched against the human metabolome database, as an equivalent resource for the cow metabolome was not available.

Figure 2.6 shows the PLS-DA multivariate analysis plots generated from the milk metabolome data set containing the metabolites detectable in more than 80% of the samples. Appendix Figures A2.3 shows their corresponding PCA plots. Appendix Figure A2.4 shows the PCA plot including the QC samples; very close clustering of the QC data indicates CIL LC-MS did not introduce any instrument bias in sample analysis. In Figure 2.6A and 2.6B, there is no apparent visual separation between the different fat percentages of milk with the exception of the 3.25% milk. However, the plot had a poor $R^2$ value (representative of goodness of fit) and $Q^2$ value (predictability value) of 0.241 and 0.114, respectively, which did not meet the acceptable separation-threshold of 0.5. Therefore, the visual separation observed in the model was rejected. Even the comparison between skim milk and 3.25% milk showed no significant separation between the two after statistical analysis. The model generated in that comparison did not pass the validation test (see Appendix Figure A2.5 B where the $Q^2$ y-intercept value is greater than 0); thus, the visual separation observed (Appendix Figure A 2.5 A) was also rejected. These results indicate that the amine/phenol submetabolomes of different fat percentages of milk within a batch were similar, which is not surprising considering that all samples had undergone lipid removal prior to chemical labeling and analysis.

Differences between milk brands were also investigated. Majority of all milk purchased within Alberta comes from Alberta dairy farms due to the Canadian Dairy Commission's supply management on domestic production and consumption. A dairy farm may house more than one breed of cattle (Jersey, Holstein, etc.). They may use different combinations of cattle feed and sell their raw milk to more than one processor. Essentially, each brand of milk may be produced from a combination of different local dairy farms. Lastly, milk companies have variations in their production methods and processing techniques of saleable milk. From the PLS-DA plots shown

in Figure 2.6C and 2.6D, there are some separations between the three brands of cow milk. The 3-dimentional (3D) plot in Figure 2.6D shows a clear separation with $R^2$ and $Q^2$ values of 0.971 and 0.927. Using a 50-permutation test, the model passed validation with a $Q^2$ y-intercept value of -0.0891, which indicates no issue of over-fitting (Appendix Figure A2.6 A). These findings are consistent with previous studies that showed relationships between external/internal factors and differences in individual cow milk composition. External factors can affect milk protein and lipid composition.[41, 42] Factors such as nitrogen fertilization of the feed or the forage particle size can affect the composition of the cow milk.[42, 43] Internal factors such as the breed of the cow also affects milk composition. Jersey cows were found to have higher percentages of protein and fat in their milk as compared to Holstein cows.[41, 44] Overall, variances between the amine/phenol submetabolomes of different brands of grocery store cow milk were observed.

Lastly, the milk samples were grouped by batch number (or collection date). In Figure 2.6E and 6F, there are significant separations between the different batches of milk with $R^2$=0.978 and $Q^2$=0.947. The model was validated using a 50-permutation test that provided a passing $Q^2$ y-intercept of -0.257 (Appendix Figure A2.6 B). This separation is more clearly shown in the 3D PLS-DA plot in Figure 2.6F. The three batches of milk samples were obtained at different time points during the year. Any changes in the milk metabolome could be attributed to changes in factors such as season, batches of feeds and water or stage in the lactation cycle. Previous studies observed changes in milk composition correlated to changes in season.[45, 46] Authors observed changes in factors such as temperature and sunlight hours having a negative and positive correlation, respectively, to protein and fat concentration.[45, 46] Milk protein and fat concentration also showed fluctuations during lactation with the highest concentrations of both near the end of the lactation period.[46]

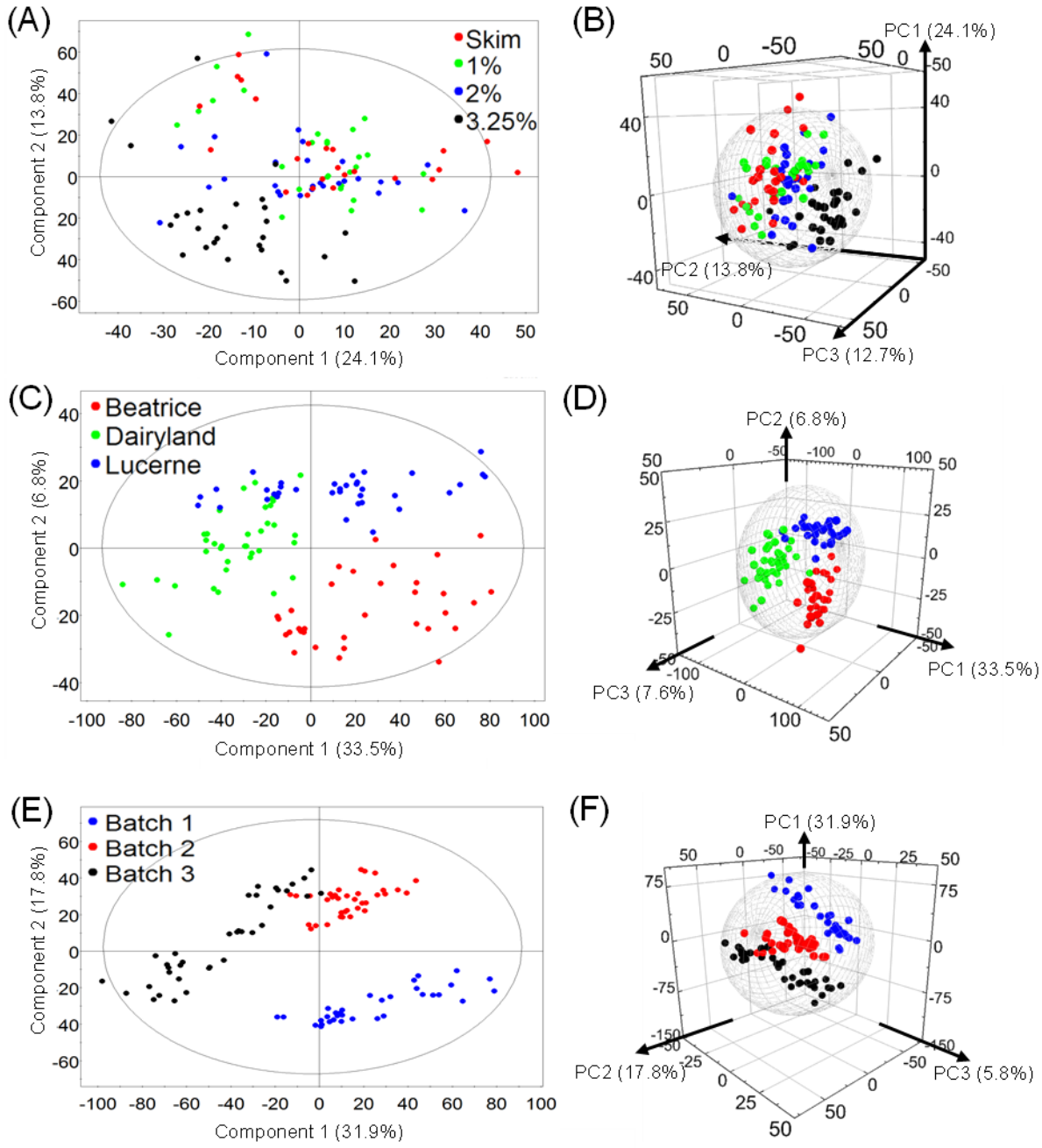



Figure 2.6 Two-dimensional and three-dimensional PLS-DA plots of cow milk samples grouped based on (A, B) fat percentage, (C, D) brand, and (E, F) and batch number.

It should be noted that none of the significant metabolites found to separate the different groups could be positively identified using the dansyl standard library, although many of them could match to some metabolite structures in the HMDB and EML database (i.e., they are likely real metabolites). Without positive IDs, we are unable to discuss their biological significance.

Overall, the above results of analyzing different samples show the ability of the dansylation LC-MS workflow to generate high-coverage and quantitative metabolome data that can reveal small differences in milk metabolome profiles. We note that direct comparison of metabolite detectability between our method and other reported work cannot be done as our focus is on the analysis of the amine/phenol submetabolome while other work used one or several combined analytical methods to detect as many different metabolites and lipids as possible. In a comprehensive profiling work reported recently,[15] a total of 710 metabolites were detected using multiple-platforms including LC-MS, UPLC-MS, GC-MS, CE-MS and NMR. Majority of them were lipids; however, 26 amine/phenol-containing metabolites were listed. In our work of analyzing 108 milk samples using the dansylation LC-MS platform alone, as discussed earlier, a total of 3820 peak pairs or metabolites were commonly detectable in more than 80% of the samples with 70 amine/phenol-containing metabolites positively identified. Among the positively identified metabolites from our study and the reported work, there were 21 metabolites in common (see Supplemental Table S2.6A) with 5 unique to the reported study (i.e., Gly-Gly, cytosine, creatine phosphate and creatine, cytidine; see Supplemental Table S2.6B). Among the 5 unique metabolites, cytosine, creatine and cytidine are in our dansyl standard library; but we did not detect them in our samples. The other two, Gly-Gly and creatine phosphate, are not in our library so we could not positively identify them in our samples. Thus, most of the amine/phenol-containing metabolites reported were identified in our samples. More significantly, we identified 49

additional amine/phenol-containing metabolites (see Supplemental Table S2.6C), including some nucleosides (e.g., guanosine and uridine) that are bioactive compounds found in milk of several species.[47, 48] In addition, there were 2988 putatively matched metabolites in our data set. These results clearly show the advantage of dansylation LC-MS for in-depth profiling of the amine/phenol submetabolome. We expect that, using other labeling chemistries, many more metabolites will be detected in milk. In addition, CIL LC-MS is a quantitative method for milk metabolome profiling as demonstrated in this work. We are not aware of any other studies of quantification accuracy and precision of profiling a large number of milk metabolites using other MS techniques.

## 2.4 Conclusions

In summary, we have developed an analytical workflow to address the current challenge of performing comprehensive and quantitative profiling of the milk metabolome. It involves the use of an optimized sample preparation method for protein and lipid removal from milk, followed by chemical isotope labeling of the milk metabolites and subsequent LC-MS analysis. We envisage the application of this workflow for comprehensive and quantitative milk metabolomics in future studies involving the dairy industry (e.g., new processing technologies, quality control and product authentication) and nutritional sciences with implications of dairy consumption on human health.

## 2.5 Literature Cited


(1) Zheng, H.; Clausen, M. R.; Dalsgaard, T. K.; Bertram, H. C. *Nutrients* **2015**, *7*, 4875-4896.

(2) Kim, S.; Kim, J.; Yun, E. J.; Kim, K. H. *Curr Opin Biotechnol* **2016**, *37*, 16-23.

(3) Gibbons, H.; O'Gorman, A.; Brennan, L. *Curr Opin Lipidol* **2015**, *26*, 30-34.

(4) Mayne, S. T.; Playdon, M. C.; Rock, C. L. *Nat Rev Clin Oncol* **2016**, *13*, 504-515.

(5) Pereira, P. C. *Nutrition* **2014**, *30*, 619-627.

(6) Cilla, A.; Quintaes, K. D.; Barbera, R.; Alegria, A. *Crit Rev Food Sci Nutr* **2016**, *56*, 1880-1892.

(7) Hess, J. M.; Jonnalagadda, S. S.; Slavin, J. L. *Compr. Rev. Food. Sci. Food Saf.* **2016**, *15*, 251-268.

(8) Sousa, S. G.; Delgadillo, I.; Saraiva, J. A. *Crit Rev Food Sci Nutr* **2016**, *56*, 1043-1060.

(9) Oliver, S. P.; Boor, K. J.; Murphy, S. C.; Murinda, S. E. *Foodborne Pathog Dis* **2009**, *6*, 793-806.

(10) Munir, M. T.; Yu, W.; Young, B. R.; Wilson, D. I.; I2C *Trends Food Sci Technol* **2015**, *43*, 205-218.

(11) Cappozzo, J. C.; Koutchma, T.; Barnes, G. *J Dairy Sci* **2015**, *98*, 5068-5079.

(12) Rocchi, S.; Caretti, F.; Gentili, A.; Curini, R.; Perret, D.; Perez-Fernandez, V. *Food Chem* **2016**, *211*, 455-464.

(13) Hampel, D.; Allen, L. H. *Crit Rev Food Sci Nutr* **2016**, *56*, 494-511.

(14) Klein, M. S.; Almstetter, M. F.; Schlamberger, G.; Nurnberger, N.; Dettmer, K.; Oefner, P. J.; Meyer, H. H. D.; Wiedemann, S.; Gronwald, W. *J Dairy Sci* **2010**, *93*, 1539-1550.

(15) Andreas, N. J.; Hyde, M. J.; Gomez-Romero, M.; Lopez-Gonzalvez, M. A.; Villasenor, A.; Wijeyesekera, A.; Barbas, C.; Modi, N.; Holmes, E.; Garcia-Perez, I. *Electrophoresis* **2015**, *36*, 2269-2285.

(16) Yang, Y. X.; Zheng, N.; Zhao, X. W.; Zhang, Y. D.; Han, R. W.; Yang, J. H.; Zhao, S. G.; Li, S. L.; Guo, T. J.; Zang, C. J., et al. *J. Proteomics* **2016**, *136*, 174-182.

(17) Sundekilde, U. K.; Downey, E.; O'Mahony, J. A.; O'Shea, C. A.; Ryan, C. A.; Kelly, A. L.; Bertram, H. C. *Nutrients* **2016**, *8*, 15.

(18) Wu, J. F.; Domellof, M.; Zivkovic, A. M.; Larsson, G.; Ohman, A.; Nording, M. L. *Biochem Biophys Res Commun* **2016**, *469*, 626-632.

(19) Spevacek, A. R.; Smilowitz, J. T.; Chin, E. L.; Underwood, M. A.; German, J. B.; Slupsky, C. M. *J Nutr* **2015**, *145*, 1698-1708.

(20) Tian, H.; Zheng, N.; Wang, W. Y.; Cheng, J. B.; Li, S. L.; Zhang, Y. D.; Wang, J. Q. *Sci Rep* **2016**, *6*, 10, DOI: 10.1038/srep24208.

(21) Smilowitz, J. T.; O'Sullivan, A.; Barile, D.; German, J. B.; Lonnerdal, B.; Supsky, C. M. *J Nutr* **2013**, *143*, 1709-1718.

(22) Caboni, P.; Murgia, A.; Porcu, A.; Demuru, M.; Pulina, G.; Nudda, A. *J Dairy Sci* **2016**, *99*, 6046-6051.

(23) Villasenor, A.; Garcia-Perez, I.; Garcia, A.; Posma, J. M.; Fernandez-Lopez, M.; Nicholas, A. J.; Modi, N.; Holmes, E.; Barbas, C. *Anal Chem* **2014**, *86*, 8245-8252.

(24) Guo, K.; Li, L. *Anal Chem* **2009**, *81*, 3919-3932.

(25) Guo, K.; Li, L. *Anal Chem* **2010**, *82*, 8789-8793.

(26) Liu, P.; Huang, Y. Q.; Cai, W. J.; Yuan, B. F.; Feng, Y. Q. *Anal Chem* **2014**, *86*, 9765-9773.

(27) Dai, W. D.; Huang, Q.; Yin, P. Y.; Li, J.; Zhou, J.; Kong, H. W.; Zhao, C. X.; Lu, X.; Xu, G. W. *Anal Chem* **2012**, *84*, 10245-10251.

(28) Tayyari, F.; Gowda, G. A. N.; Gu, H. W.; Raftery, D. *Anal Chem* **2013**, *85*, 8715-8721.

(29) Fan, R. J.; Guan, Q.; Zhang, F.; Leng, J. P.; Sun, T. Q.; Guo, Y. L. *Anal Chim Acta* **2016**, *908*, 132-140.

(30) Hao, L.; Zhong, X. F.; Greer, T.; Ye, H.; Li, L. J. *Analyst* **2015**, *140*, 467-475.

(31) Bruheim, P.; Kvitvang, H. F. N.; Villas-Boas, S. G. *J Chromatogr A* **2013**, *1296*, 196-203.

(32) Huan, T.; Wu, Y.; Tang, C.; Lin, G.; Li, L. *Anal Chem* **2015**, *87*, 9838-9845.

(33) Wu, Y.; Li, L. *Anal Chem* **2012**, *84*, 10723-10731.

(34) Zhou, R.; Tseng, C. L.; Huan, T.; Li, L. *Anal Chem* **2014**, *86*, 4675-4679.

(35) Huan, T.; Li, L. *Anal Chem* **2015**, *87*, 1306-1313.

(36) Huan, T.; Li, L. *Anal Chem* **2015**, *87*, 7011-7016.

(37) Li, L.; Li, R. H.; Zhou, J. J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y. M.; Huan, T.; Zheng, J. M.; Shi, Y.; Wishart, D. S., et al. *Anal Chem* **2013**, *85*, 3401-3408.

(38) Vuckovic, D. *Anal Bioanal Chem* **2012**, *403*, 1523-1548.

(39) Tang, Y. A.; Li, L. *Anal Chim Acta* **2013**, *792*, 79-85.

(40) Zhao, S.; Luo, X.; Li, L. *Anal Chem* **2016**, *88*, 10617-10623.

(41) Croissant, A. E.; Washburn, S. P.; Dean, L. L.; Drake, M. A. *J Dairy Sci* **2007**, *90*, 4942-4953.

(42) Nasrollahi, S. M.; Imani, M.; Zebeli, Q. *J Dairy Sci* **2015**, *98*, 8926-8939.

(43) Warner, D.; Podesta, S. C.; Hatew, B.; Klop, G.; van Laar, H.; Bannink, A.; Dijkstra, J. *J Dairy Sci* **2015**, *98*, 3383-3393.

(44) Palladino, R. A.; Buckley, F.; Prendiville, R.; Murphy, J. J.; Callan, J.; Kenny, D. A. *J Dairy Sci* **2010**, *93*, 2176-2184.

(45) Bernabucci, U.; Basirico, L.; Morera, P.; Dipasquale, D.; Vitali, A.; Piccioli Cappelli, F.; Calamari, L. *J Dairy Sci* **2014**, *98*, 1815-1827.

(46) Roche, J. R.; Turner, L. R.; Lee, J. M.; Edmeades, D. C.; Donaghy, D. J.; Macdonald, K. A.; Penno, J. W.; Berry, D. P. *Anim. Prod. Sci* **2009**, *49*, 222-232.

(47) Gill, B. D.; Indyk, H. E.; Manley-Harris, M. *International Dairy Journal* **2012**, *24*, 40-43.

(48) Gill, B. D.; Indyk, H. E. *International Dairy Journal* **2007**, *17*, 596-605.

# Chapter 3 CIL LC-MS for Profiling the Dietary Effects of Cow Milk Consumption on the Human Urine Metabolome

## 3.1 Introduction

Metabolomics is the large-scale analysis of all the small molecules (<1500 Da) in a given biological sample.[1] Metabolomics is a powerful analytical technique that is used to distinguish one set of samples from another set (i.e. diseased vs healthy) through comparison. It can also be used to determine the effects of a substance on a particular metabolome (i.e. drug treatment vs no drug treatment) within the same individuals. As we know, the metabolome is complex and dynamic. Along with genomic variability, the exposome of a particular individual will also influence their metabolome. What we put into our bodies will affect us at a molecular level. The exposome refers to things like nutrition, microbiota, xenobiotics, pollution and environmental exposures.[2] In recent years, due to the interest in metabolomics for personalised care, metabolomics is being readily combined with nutritional sciences generating what's more commonly known as nutrimetabonomics.[3]

It is well known that diet heavily influences health and can play a role in maintaining a healthy state or developing a diseased state. Food is complex.[4] In the exposome, food is one of the most influencing factors of human health due to the large number of nutritional and non nutritional components along with a variety of food and non-food related compounds individuals are exposed to through eating.[4,5] Non food/nutritional compounds can include pesticides and preservatives that provide no nutritional value to the individual but can, however, still affect their health state.[4] Each individual will respond differently to the same foods.

*I prepared the samples, analyzed the samples, data processed and performed statistical analysis on the dataset, wrote and edited this chapter.

Two individuals eating high fat containing diets may have vastly different outcomes. One may develop obesity, hypertension and heart disease while the other may remain fairly healthy. The reason behind this inter-individual variation is still unclear.

Nutrimetabonomics incorporates four main goals. Firstly, it aims to understand the interactions of foods or diet on the metabolic system of the individual at the molecular level.[1] Secondly it looks at the effects of a nutritional regime or diet on human health and disease.[1] Thirdly, it can be used to identify food biomarkers for diet tracking and quantitively assessing the amounts of substances consumed.[1] Lastly, it can be used to investigate the interactions between diet, the gut microbial composition and its overall effect on the individual's metabolic status (health and disease states).[1] By obtaining the information illustrated above, metabolomics can help play a part in developing a personalized nutrition regime for each individual that will help maintain or obtain a healthy state.[1] It can also help predict the effects of various foods and food compounds to allow for changes to nutritional requirements and quality control of commercially available foods.[1]

According to the World Health Organization (WHO), diet is one of the main factors that contribute to the outcome and risk of chronic diseases.[6,7] In Canada, healthy nutrition (according to the Canadian Food Guide) encompasses eating a balanced combination of foods from the four main food groups which include: fruits and vegetables, grain products, milk and alternatives and meat and alternatives. Milk is one of the main food groups and is consumed multiple times through different milk and milk-containing products daily from toddler to adulthood. Many population studies have been performed to investigate the effects of milk consumption on human health such as height, weight and bone strength.[8–10] However, there are limited studies involving the study of the dietary effects of milk at the molecular level and how it can modulate metabolic responses.

In this chapter, CIL LC-MS was used to investigate the dietary perturbations of cow milk consumption on the human urine metabolome one and two hours after intake.

## 3.2 Experimental

### 3.2.1 Chemical and Reagents

All the chemicals and reagents, unless otherwise stated, were purchased from Sigma-Aldrich Canada. For dansylation labeling, the $^{12}$C-labeling reagent (dansyl chloride) was purchased from Sigma-Aldrich and the $^{13}$C-labeling reagent was synthesized according to the method published previously.[11]

### 3.2.2 Sample Collection and Processing

Informed consent was obtained from individual volunteers, and ethics approval was obtained from the University of Alberta in compliance with the University of Alberta Heath Information policy. Urine samples were collected from 6 healthy individuals, 2 females and 4 males. The short-term urine study involved 4 urine samples per collection day. The first sample was collected after 12 hours of overnight fasting and was deemed the void sample. The next sample was collected 1 hour after the first and was categorized as the "before" milk intake sample. After collection of the second sample the individual consumed 250 mL of Dairyland 1% milk. The third sample was collected 1 hour after which the milk was consumed. This sample was labelled as "1 hour" after milk intake. Lastly the fourth urine sample was collected 2 hours after the initial milk consumption and labelled as such.

The study was repeated on 3 separate days for each individual. All urine samples were stored in the 4 °C fridge immediately after collection. Within the same day, the urine samples were vortexed at 4000 rpm for 10 minutes. The supernatant was filtered by 0.22 μm-pore-sized

Millipore filter (Millipore Corp., MA) and aliquoted into 0.6 mL vials. 12.5 µL of urine sample was aliquoted out and diluted 4-fold by adding 37.5 µL of water. The 50 µL diluted urine solutions were then ready for dansylation and stored in the -80 °C freezer until further use.

### 3.2.3 Dansyl Labeling

Dansyl chloride was used as the labeling reagent to react mainly amine- and phenol-containing metabolites to form dansyl-amine or dansyl-phenol derivative.[11] 50 µL of a processed urine sample was mixed with 25 µL of 250 mM sodium bicarbonate buffer and 25 µL of acetonitrile and the solution was vortexed. 50 µL of 18 mg/mL $^{12}$C- or $^{13}$C-dansyl chloride in acetonitrile was added, vortexed and incubated at 40 °C for 45 min. To quench the excess dansyl chloride, 10 µL of 250 mM sodium hydroxide was added and the solution was incubated at 40 °C for 10 min. Finally, 50 µL of 425 mM formic acid was added to acidify the sample. Individual samples were labeled with $^{12}$C-dansyl chloride and a pooled sample was labeled with $^{13}$C-danysl chloride.

### 3.2.4 LC-UV Quantification

For LC-UV, a Waters ACQUITY UPLC system with a photodiode array (PDA) detector was used for the quantification of dansyl labeled metabolites for sample amount normalization as described earlier.[12] Briefly, 4 µL of each labeled sample was injected onto a Phenomenex Kinetex C18 column (2.1 mm × 5 cm, 1.7 µm particle size) for a fast step-gradient run. Solvent A was 0.1% (v/v) formic acid in 5% (v/v) ACN, and solvent B was 0.1% (v/v) formic acid in ACN. The gradient started with 0% B for 1 min and was increased to 95% within 0.01 min and held at 95% B for 1 min to ensure complete elution of all labeled metabolites. The flow rate used was 0.45 mL/min. The peak area related to the total labeled metabolite concentration in the sample was

integrated using the Empower software (6.00.2154.003). Based on the quantification results, the $^{12}C$-labeled sample and the $^{13}C$-labeled pool were mixed in equal amounts.

### 3.2.5 LC-MS

All LC-MS experiments were performed on an Agilent 1100 HPLC system (Palo Alto, CA) connected to a Bruker Impact HD quadrupole time-of-flight (QTOF) mass spectrometer (Billercia, MA) with an ESI source. A reverse phase column (Agilent Eclipse Plus C18 column, 2.1 mm x 10 cm, 1.8 μm particle size, 95 Å pore size) was used for liquid chromatography separation of labeled metabolites. Mobile phase A was made up of 5% (v/v) acetonitrile and 0.1% (v/v) formic acid in water. Mobile phase B consists of 0.1% (v/v) formic acid in acetonitrile. The 32-min gradient conditions were: 0 min (20% B), 0-3.5 min (20-35% B), 3.5-18 min (35-65% B), 18-24 min (65-99% B) and 24-32 min (99% B). The column was then re-equilibrated at 20% B for 15 min. The flow rate was 180 μL/min.

### 3.2.6 Data Processing, Statistical Analysis and Metabolite Identification

Bruker DataAnalysis software 4.2 was used to extract MS spectral peaks. An in-house software tool, IsoMS was used to process the raw data generated from multiple LC− MS runs by peak picking, peak pairing, peak-pair filtering, and peak-pair intensity ratio calculation.[13] The same peak pairs detected from multiple samples were then aligned to produce a CSV file that contains the metabolite information and peak ratios relative to a control (i.e., a pooled sample). A zero-fill program was then used to find missing peak pairs from the raw mass spectral data, filling in the missing values.[14] Peak ratios were further optimized through the use of IsoMS-Quant.[15] Volcano plots were generated by Origin 2016 Graphing and Analysis (OriginLab Corporation, Northampton, USA).

## 3.3 Results and Discussion

### 3.3.1 Urine Collection

Six healthy individuals, 2 females and 4 males, volunteered for this study. This study investigates the short-term effects of milk consumption on the urine metabolome. As mentioned before, your exposome directly affects your metabolome. The exposome changes daily due to things like consumption of different foods, different environmental conditions, chemicals, pollutants etc. Therefore, a short-term nutrimetabonomics study was designed. With a short-term study, a control sample can be obtained from within the same day and within a few hours from the experimental samples. This limits the chances of other factors influencing the urine metabolome of interest and avoids the masking of the diet effects by the day-to-day variation commonly observed in urine metabolomic studies. Figure 3.1 illustrates the urine collection process. Briefly, all individuals fasted for 12 hours overnight. The first pass of urine was collected after 12 hours and labeled “void”. One hour after the void sample, the next urine sample was collected and labeled “before”. Immediately after the second urine collection, the individual then consumed 250 mL of 1% Dairyland cow milk. One hour after drinking milk, the individuals collected their third urine sample labeled “1 hour after milk”. Lastly, 2 hours after milk intake the individuals collected their final urine sample labeled “2 hours after milk”.

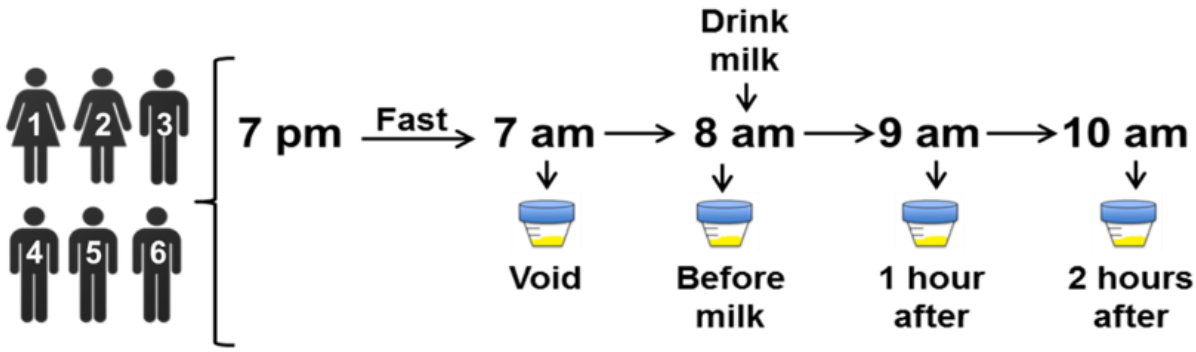



Figure 3.1 The urine collection workflow for the investigation of the dietary effects of cow milk consumption on the urine metabolome of 6 individuals, 1 and 2 hours after cow milk intake.

All individuals were able to drink water throughout the duration of the collection period. The urine samples were collected in 50 mL sterile Eppendorf tubes. After collection, all samples were stored upright in the 4°C fridge. The urine was processed the same day it was collected (detailed process in the experimental section) and was stored in the -80°C freezer until usage.

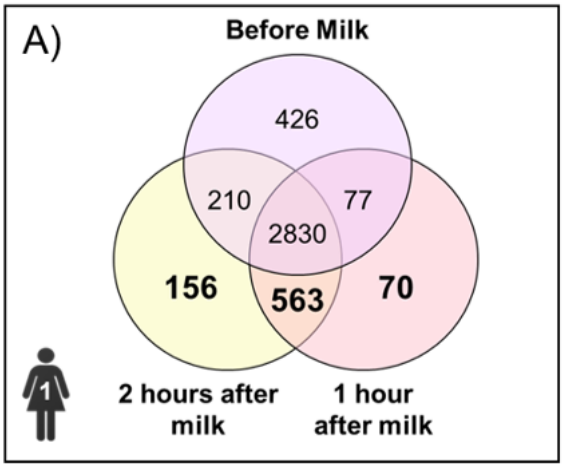
### 3.3.2 Determination of Unique and Common Metabolites to Urine and Milk

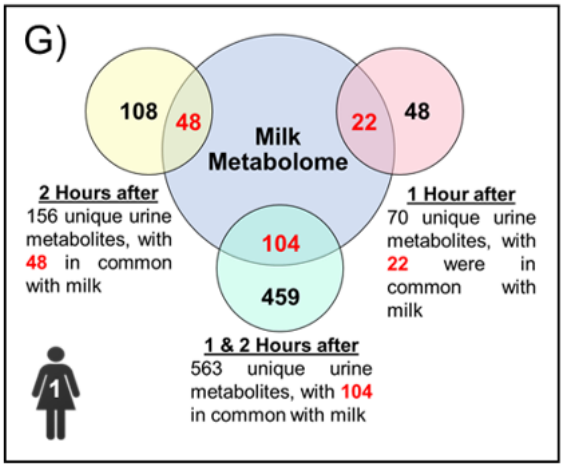
All individual urine samples were labeled with $^{12}$C-dansyl chloride while the pooled sample (generated by taking equal aliquots of all samples) was labeled with $^{13}$C-dansyl chloride. After labeling, each individual and pooled sample was quantified by LC-UV and the concentration of labeled metabolites was determined. The samples were then mixed in equimolar amounts of $^{12}$C-individual sample and $^{13}$C-pooled sample and analyzed by LC-MS as single injections.

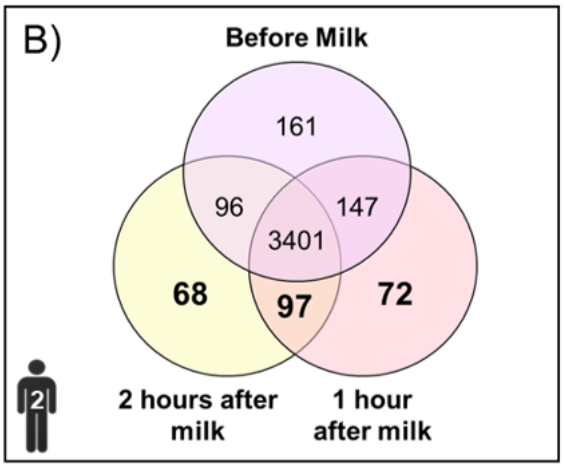
There are many other external factors that can affect the results of this study such as the individual's previous meal, their metabolism, intestinal uptake, gut microbial composition and activity, and tolerance to milk and milk products etc. Along with inter-day variability within the same individual, there is also inter-individual variability which is much more pronounced. Therefore, the resulting dataset was grouped by each individual, and analysed independently. In

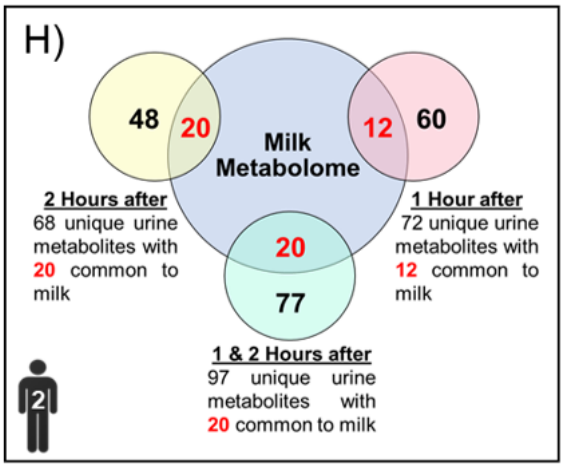
order to determine any milk metabolites excreted from the urine, the dataset from the urine study was aligned with the cow milk profiling dataset obtained from the study in chapter 2. Using the same criteria mentioned in chapter 2, the milk metabolites found consistently across 80% of all milk samples was compared with the urine metabolites before, 1 hour and 2 hours after milk intake. Firstly, I determined the number of unique metabolites to each collection time point (Figure 3.2 A-F). These unique metabolites include both urine and milk metabolites.

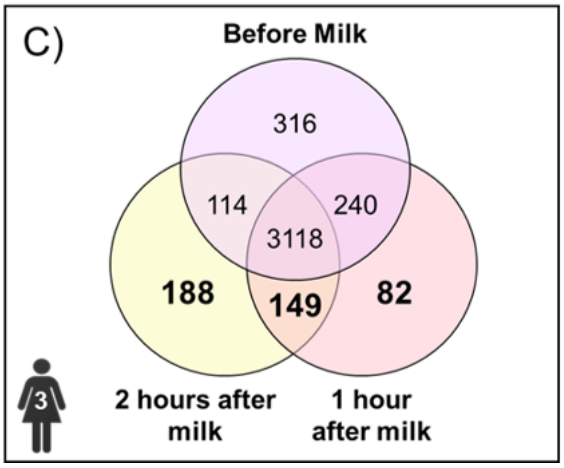
Then I compared the unique metabolites for each time point with the milk metabolites and determined the number of metabolites common to milk (Figure 3.2 G-L). As an example, Figure 3.2A shows the Venn diagram for volunteer #1. As the purpose of this study is to see the effects of milk on the urine metabolome, I only focused on the unique metabolites observed in the 1 hour, 1 and 2 hour and 2 hours after milk consumption urine samples. The numbers of those metabolites observed is bolded in Figure 3.2A. There are 70 metabolites unique to the 1-hour-after-milk urine samples, 156 metabolites unique to the 2-hours-after-milk urine samples and 563 metabolites found in both the 1-and-2-hours-after-milk urine samples. The above-mentioned metabolites were then compared with the milk metabolites obtained from the milk profiling set. Figure 3.2G shows the distribution of metabolites (1 hour, 2 hours and both 1 and 2 hours after milk intake) that are in common with milk (shown in red) or unique to urine (shown in black). Of the 70 metabolites unique to the 1-hour-after samples, 22 of them were in common with milk. 48 of the 156 unique metabolites in the 2-hours-after samples were in common with milk. Lastly, 104 metabolites out of the 563 unique metabolites to the 1-and-2-hours-after samples were matched with milk metabolites.

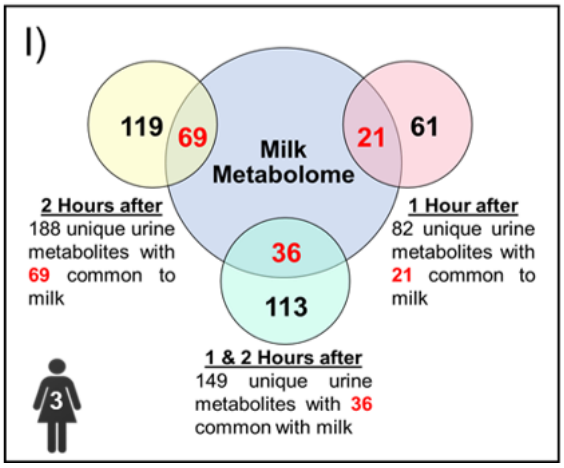

A)
Before Milk
426
210
77
2830
156
563
70
1
2 hours after milk
1 hour after milk


G)
108
48
Milk Metabolome
22
48
2 Hours after
156 unique urine metabolites, with 48 in common with milk
1 Hour after
70 unique urine metabolites, with 22 were in common with milk
104
459
1 & 2 Hours after
563 unique urine metabolites, with 104 in common with milk
1


B)
Before Milk
161
96
147
3401
68
97
72
2
2 hours after milk
1 hour after milk


H)
48
20
Milk Metabolome
12
60
2 Hours after
68 unique urine metabolites with 20 common to milk
1 Hour after
72 unique urine metabolites with 12 common to milk
20
77
1 & 2 Hours after
97 unique urine metabolites with 20 common to milk
2


C)
Before Milk
316
114
240
3118
188
149
82
3
2 hours after milk
1 hour after milk


I)
119
69
Milk Metabolome
21
61
2 Hours after
188 unique urine metabolites with 69 common to milk
1 Hour after
82 unique urine metabolites with 21 common to milk
36
113
1 & 2 Hours after
149 unique urine metabolites with 36 common with milk
3

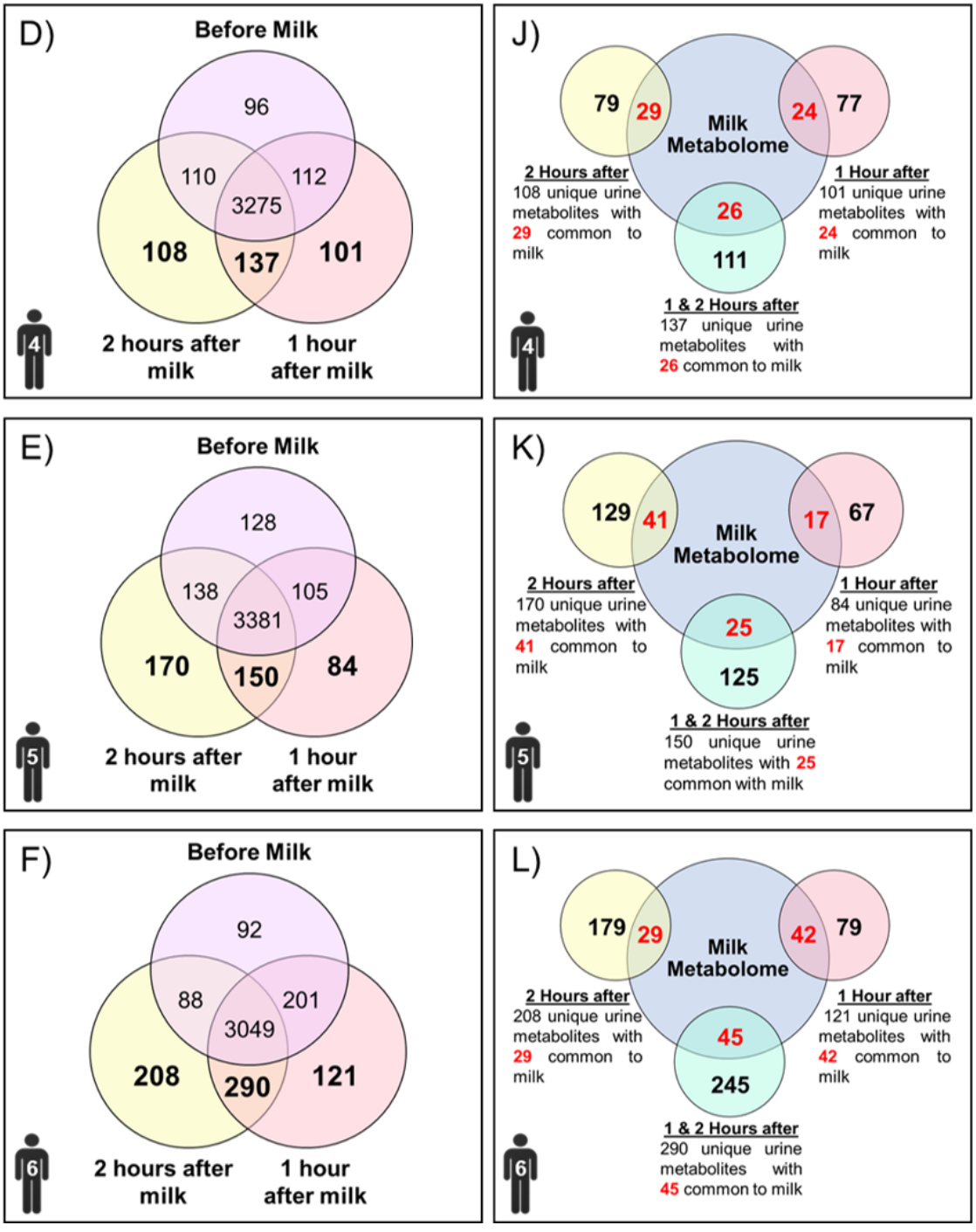
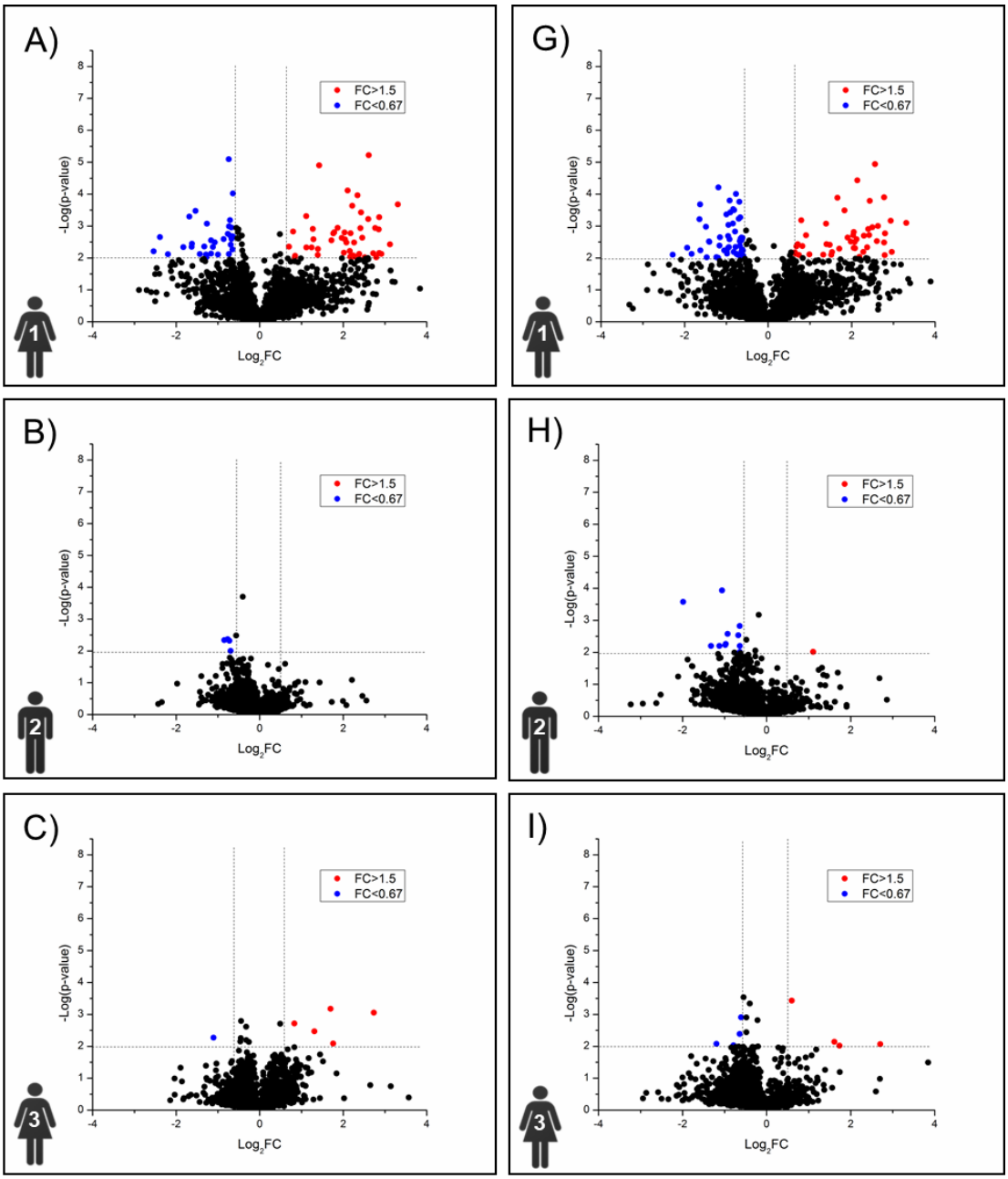



Figure 3.2 The distribution of metabolites 1 and 2 hours after milk intake for the 6 volunteers. Plots A-F show the distribution of metabolites in each collection time point. The metabolites unique to the after-milk samples are bolded in black. Plots G-L show the number of the unique metabolites generated from A-F that were found to be in common with milk metabolites determined from the previous study in chapter 2. These metabolites are bolded in red. In the corner of each plot is the volunteer number and gender.

### 3.3.3 Univariate analysis

Univariate analysis was used to determine metabolites with significant concentration changes after drinking milk. Firstly, the common metabolites between the 3 different collection time points was determined. Then the significant changed metabolites between "before" and "after" milk intake was identified and plotted on Volcano Plots for each individual following the criteria: p-value < 0.01, fold change (FC) > 1.5 and FC <0.67 refers to significantly increased and decreased, respectively. Figure 3.3 A-F show the plots of the significantly changed metabolites 1-hour-after-milk intake and Figure 3.3 G-L shows the plots for the metabolites changing 2-hours-after-milk intake. For volunteer #1, the individual had 47 and 31 significantly increased and decreased metabolites, 1-hour after milk intake. 2-hours after intake the numbers were 46 and 52, increased and decreased, respectively (Figure 3.3 A, G). For volunteer #2, they had 0 increased and 4 decreased metabolites at 1-hour after milk and 1 increased and 11 decreased metabolites at 2-hours after milk intake (Figures 3.3 B, H). Volunteer #3 had 7 increased and 1 decreased metabolites after 1-hour of drinking milk (Figure 3.3 C). Those numbers changed to 4 increased and 4 decreased 2-hours after milk (Figure 3.3 I). Volunteer #4 had 17 increased and 0 decreased after 1-hour post milk intake and had 7 increased and 1 decreased 2-hours post-milk (Figures 3.3 D, J). Volunteer #5 had 14 increased metabolites 1-hour post-milk intake and 0 decreased (Figure 3.3 E). 2-hours post milk intake those numbers changed to 67 increased and 6 decreased (Figure 3.3 K). Lastly, volunteer #6 had 13 increased and 3 decreased metabolites after 1-hour of consuming milk (Figure 3.3 F). 2-hours after milk intake those numbers were 42 increased and 12 decreased (Figure 3.3 L).


A)
G)
B)
H)
C)
I)
FC>1.5
FC<0.67
-Log(p-value)
Log2FC
1
2
3

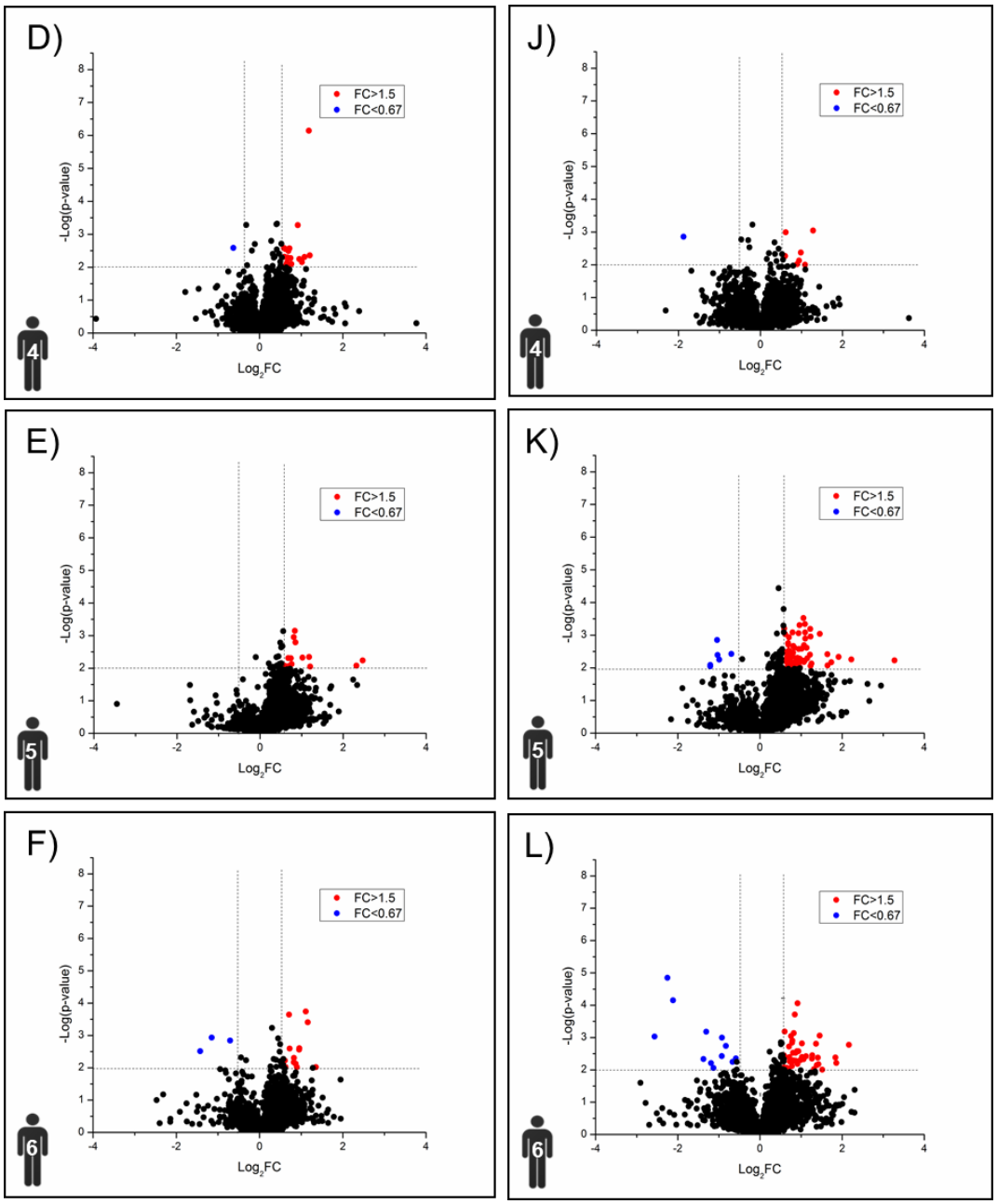



Figure 3.3 Univariance analysis of the urine samples 1 and 2 hours after milk intake. Plots A-F show the volcano plots for the 6 individual's urine samples collected 1 hour after milk intake. Plots G-L show the volcano plots for the urine samples collected 2 hours after milk intake. The criteria for determination of significant metabolites were as follows: a p-value <0.01, a fold-change (FC) of >1.5 was determined as increased (red data points) and FC<0.67 as decreased (blue data points). The volunteer's number and gender is shown in the bottom left corner of each plot.

Compared to the total number of metabolites observed in the urine or milk metabolome, the number of overlapped features determined were much lower (~3000 metabolites vs <110 metabolites). As well, the number of metabolites with significantly changed concentrations was also small. This was, in a way, expected. Past studies have found that dietary effects on the metabolome are less pronounced than say, a drug effect and can be easily masked by other variables such as inter-individual and inter-day variation within the same individual.[1] As well, there is no observed pattern to the changes in the urine metabolome across individuals thus demonstrating the unique individual responses to milk consumption. It should, however, be noted that the effects are more complex since food is not just comprised of one component and those components may be involved in various metabolic pathways. Milk, for example, is comprised of a variety of proteins, lipids, sugars, vitamins and minerals.[17] Also, the number of milk metabolites excreted out refers to any metabolites that have not undergone bio-transformations during the ingestion, digestion, absorption and elimination process. Our method, unfortunately, cannot account for bio-transformed metabolites.

### 3.4 Conclusion and Future Work

Overall the results of this study show no significant effect of milk consumption on the human urine metabolome 1 or 2 hours after intake to warrant large changes in the metabolomic profiles of the individuals. However, the method is still able to detect the subtle and individually unique changes that do occur in the urine metabolome after drinking milk. For the future work, a more carefully designed double-blind (if possible) study needs to be designed with a larger group of volunteers. As well, different labeling chemistries can be used to investigate the effects of milk consumption on the other submetabolomes.

### 3.5 Literature Cited


(1) Claus, S. P.; Swann, J. R. *Annu. Rev. Food Sci. Technol.* **2013**, *4*, 381–399.

(2) Vrijheid, M. *Thorax* **2014**, *69*, 876–878.

(3) McNiven, E. M. S.; German, J. B.; Slupsky, C. M. *J. Nutr. Biochem.* **2011**, *22*, 995–1002.

(4) Jones, D. P.; Park, Y.; Ziegler, T. R. *Annu. Rev. Nutr.* **2012**, *32*, 183–202.

(5) Astarita, G.; Langridge, J. *J. Nutrigenet. Nutrigenomics* **2013**, *6*, 181–200.

(6) Jessri, M.; Ng, A. P.; L'Abbé, M. R. *Nutrients* **2017**, *9*, 910.

(7) Lock, K.; Pomerleau, J.; Causer, L.; Altmann, D. R.; McKee, M. *Bull. World Health Organ.* **2005**, *83*, 100–108.

(8) Michaëlsson, K.; Wolk, A.; Langenskiöld, S.; Basu, S.; Warensjö Lemming, E.; Melhus, H.; Byberg, L. *BMJ.* **2014**, *349*, g6015.

(9) Okada, T. *Am. J. Clin. Nutr.* **2004**, *80*, 1088-9-90.

(10) Rozenberg, S.; Body, J.-J.; Bruyère, O.; Bergmann, P.; Brandi, M. L.; Cooper, C.; Devogelaer, J.-P.; Gielen, E.; Goemaere, S.; Kaufman, J.-M.; Rizzoli, R.; Reginster, J.-Y. *Calcif. Tissue Int.* **2016**, *98*, 1–17.

(11) Guo, K.; Li, L. *Anal. Chem.* **2009**, *81*, 3919–3932.

(12) Wu, Y.; Li, L. *Anal. Chem.* **2012**, *84*, 10723–10731.

(13) Zhou, R.; Tseng, C.-L.; Huan, T.; Li, L. *Anal. Chem.* **2014**, *86*, 4675–4679.

(14) Huan, T.; Li, L. *Anal. Chem.* **2015**, *87*, 1306–1313.

(15) Huan, T.; Li, L. *Anal. Chem.* **2015**, *87*, 7011–7016.

(16) Huan, T.; Wu, Y.; Tang, C.; Lin, G.; Li, L. *Anal. Chem.* **2015**, *87*, 9838–9845.

(17) Pereira, P. C. *Nutrition* **2014**, *30*, 619–627.

# Chapter 4 Development of a Quantitative Metabolomics Method for Detecting Potential Milk Adulterant in Human Milk

## 4.1 Introduction

A mother's milk is considered the "gold standard" of a perfect source of nutrition for a newborn. However, in some cases a mother may be unable to provide sufficient amounts of milk for their child or the milk is considered unsafe for the baby.[1,2] Commercial infant-formula is an alternative when maternal breast milk is not available. Because of the many benefits correlated to breast milk consumption for infants, human milk is still preferred by many mothers.[3] Donor human milk (DHM) is recommended as the first choice of milk in place of the unavailable maternal milk.[4] In North America, the Human Milk Banking Association of North America (HMBANA), a professionally governed non-profit organization, has helped establish a total of 18 human milk banks in the US and Canada. The milk obtained by these milk banks go through vigorous screening before distribution. Due to the vigorous screening and multi-step strict processing, along with the reliance of donations, donor human milk is very limited in supply. Because of this, the milk is mainly available by a physician's prescription and for high-risk infants (typically premature or ill) residing in the Neonatal Intensive Care Unit (NICU).

With the limited access to human milk from official breast milk banks, individuals are looking to the community for DHM.[4,5] Human milk sharing allows for easier and more affordable access to DHM. There is a market for the buying and selling of this "superfood", populated by the Internet. The cost can be as high as $5 USD per ounce. However, there are many health-related risks associated with milk sharing and consumption.[4,6]

*A version of this has been submitted for publishing as "Development of a Quantitative Metabolomics Method for Detecting Potential Milk Adulterant in Human Milk" Mung, D.; Li, L. on August 26, 2017. I prepared all the samples, ran the samples, performed data processing and statistical analysis, generated the figures and wrote and edited the manuscript.

Purchased human milk may be subjected to milk adulteration (i.e., diluted with another form of milk). Recent studies have detected cow DNA in 10% of the breast milk samples purchased online in amounts (~10% contamination) that would suggest intentional adulteration.[7] Sellers may be tempted to adulterate their human milk with other milks to increase the amount of "human milk" available to be sold to inflate their profit. There are many risks associated with infant consumption of plant or animal adulterated human milk including iron deficiency, dehydration, increased potential renal solute load (PRSL), and allergic reactions.[8-13]

There is a clear need for rapid detection of milk adulterants in human milk. DNA-based analytical techniques are powerful in identifying milk of different species as well as offering semi-quantitative information on the level of contaminant or adulterant in human milk.[7] However, they require a targeted DNA probe for each adulterant and do not detect adulterants with no available probe. An alternative approach for differentiating milk of different species is to examine the proteome[14-17] or metabolome[18-23] of milk samples. Because of the possibility of acquiring metabolomic data quickly, metabolomics is a promising method for rapid detection of adulterants in human milk. The hypothesis is that the human milk metabolome can serve as a fingerprint and any significant deviation of the metabolome of an unknown milk sample from the human milk metabolome may indicate contamination or adulteration.

There are reports of using NMR[20,21], GC-MS[18,22,23] and LC-MS[19,20] to study the metabolome differences among different types of milk. However, due to limited metabolic coverage of these techniques, comparison of the metabolome profiles of different species may not be sufficiently sensitive to detect a small amount of adulterant. To our knowledge, there is no report of examining the possibility of using metabolomics for detecting contaminants or adulterants in human breast milk. In this work, we report a high-performance chemical isotope

labeling liquid chromatography mass spectrometry (CIL LC-MS) method to profile the amine/phenol submetabolomes of human milk and five potential human milk adulterants (soy, almond, cow, goat and infant formula). Compared to conventional LC-MS, CIL LC-MS generates the submetabolomic profiles with high coverage and high quantification accuracy and precision.[24] Thus, we wish to examine the possibility of using the submetabolome profiles to detect contaminant in human milk mixed with different amounts of soy, almond, cow, goat or infant formula. We demonstrate that this CIL LC-MS submetabolomic profiling method can detect adulterant at a level of as low as 5% in human milk. This level of detection should be adequate for milk adulterant analysis, as below this level does not make economic sense for volume-based adulteration.[25]

## 4.2 Experimental

### 4.2.1 Chemicals and Reagents

All the chemicals and reagents, unless otherwise stated, were purchased from Sigma-Aldrich Canada. For dansylation labeling, the $^{12}$C-labeling reagent (dansyl chloride) was purchased from Sigma-Aldrich and the $^{13}$C-labeling reagent was synthesized according to the method published previously.[24]

### 4.2.2 Milk Sample Collection and Preparation

All cow, goat, soy, almond and infant formula milk was purchased from local grocery stores and the human breast milk was donated from ten different healthy volunteers. Ethics approval for working with human samples was given by the University of Alberta Ethics Committee. For the cow, goat, soy, almond and infant formula milk, three different brands were purchased. For the cow milk, two different fat percentages (skim and 3.25%) of milk were also

purchased for each of the three brands. All milk samples were aliquoted into 50 mL collection tubes and stored in a -80 °C freezer until use.

A previously determined method for cow milk metabolite extraction was utilized for this study.[26] The metabolite extraction method, involving protein precipitation by methanol followed by lipid removal with dichloromethane, was found to be applicable to these different types of milk samples. Briefly, 100 µL of milk was mixed with 300 µL of methanol, vortexed and incubated at -20 °C for 15 min. The sample was centrifuged at 20 800 *g* for 15 min and 350 µL of the supernatant was taken into a new vial. 500 µL of dichloromethane and 50 µL of water were added to the supernatant, vortexed and left at room temperature for 10 min. 350 µL of the aqueous layer was then taken into a new vial and dried down using a SpeedVac. The sample was re-dissolved in 50 µL of water for dansylation labeling.

### 4.2.3 Dansylation Labeling

Dansyl chloride was used as the labeling reagent to react mainly amine- and phenol-containing metabolites to form dansyl-amine or dansyl-phenol derivative.[24] 50 µL of a processed milk sample was mixed with 25 µL of 250 mM sodium bicarbonate buffer and 25 µL of acetonitrile and the solution was vortexed. 50 µL of 18 mg/mL $^{12}$C- or $^{13}$C-dansyl chloride in acetonitrile was added, vortexed and incubated at 40 °C for 45 min. To quench the excess dansyl chloride, 10 µL of 250 mM sodium hydroxide was added and the solution was incubated at 40 °C for 10 min. Finally, 50 µL of 425 mM formic acid was added to acidify the sample. Individual samples were labeled with $^{12}$C-dansyl chloride and a pooled sample was labeled with $^{13}$C-danysl chloride.

### 4.2.4 LC-UV Quantification

For LC-UV, a Waters ACQUITY UPLC system with a photodiode array (PDA) detector was used for the quantification of dansyl labeled metabolites for sample amount normalization as described earlier.[27] Briefly, 4 µL of each labeled sample was injected onto a Phenomenex Kinetex C18 column (2.1 mm × 5 cm, 1.7 µm particle size) for a fast step-gradient run. Solvent A was 0.1% (v/v) formic acid in 5% (v/v) ACN, and solvent B was 0.1% (v/v) formic acid in ACN. The gradient started with 0% B for 1 min and was increased to 95% within 0.01 min and held at 95% B for 1 min to ensure complete elution of all labeled metabolites. The flow rate used was 0.45 mL/min. The peak area related to the total labeled metabolite concentration in the sample was integrated using the Empower software (6.00.2154.003). Based on the quantification results, the $^{12}$C-labeled sample and the $^{13}$C-labeled pool were mixed in equal amounts.

### 4.2.5 LC-MS

All LC-MS experiments were performed on an Dionex UltiMate 3000 UHPLC (Thermo Scientific, MA) connected to a Bruker Impact HD Quadrupole Time-Of-Flight (QTOF) mass spectrometer (Billercia, MA) with an ESI source. A reverse phase column (Agilent Eclipse Plus C18 column, 2.1 mm x 10 cm, 1.8 µm particle size, 95 Å pore size) was used for liquid chromatography separation of labeled metabolites. Mobile phase A was made up of 5% (v/v) acetonitrile and 0.1% (v/v) formic acid in water. Mobile phase B consists of 0.1% (v/v) formic acid in acetonitrile. The 32-min gradient conditions were: 0 min (20% B), 0-3.5 min (20-35% B), 3.5-18 min (35-65% B), 18-24 min (65-99% B) and 24-32 min (99% B). The column was then re-equilibrated at 20% B for 15 min. The flow rate was 180 µL/min.

To determine the optimal amount injection, a 1:1 mixed $^{12}C$-/$^{13}C$-labeled milk sample with a known concentration of 1.23 mM (measured by LC-UV) was injected in 5 µL increments from 5 to 30 µL. Injection triplicate runs were performed for gauging the technical reproducibility. From the plot of the peak pair number detected as a function of sample amount injected (Appendix Figure A4.1), we found that peak pair number saturation occurred when 24.6 nmol in 20 µL of sample was injected. However, many samples were less concentrated than the 1.23 mM pooled sample used for the injection optimization experiment and therefore will require a much larger injection volume. This will lead to peak broadening. Thus, in subsequent experiments, we injected 20 nmol of labeled sample into the LC-MS for analysis as the difference in the detected number of metabolites is minimal.

### 4.2.6 Data Analysis

Bruker DataAnalysis software 4.3 was used to extract MS spectral peaks. An in-house software tool, IsoMS, was used to process the raw data generated from multiple LC-MS runs by peak picking, peak pairing, peak-pair filtering to remove redundant peaks of the same metabolite, such as adduct ions, dimer, multimers, to retain only $[M+H]^+$ for a labeled metabolite (i.e., one peak pair corresponds to one metabolite), and peak-pair intensity ratio calculation.[28] The same peak pairs detected from multiple samples were then aligned to produce a CSV file containing the metabolite information and peak ratios relative to a control (i.e., a pooled sample). A zero-fill program was used to find missing peak pairs from the raw mass spectral data, filling in the missing values.[29] Finally, peak ratios were re-calculated by using chromatographic peaks, instead of mass spectral peaks, with IsoMS-Quant.[30] The final metabolite-intensity data file was then exported to MetaboAnalyst version 3.0 (McGill University, Montreal, Canada) for multivariate statistical

analysis. Principal component analysis (PCA) were used to analyze the data after interquartile range data filtering, sample normalization by sum and auto-scaling data scaling.

Positive metabolite identification was performed based on mass and retention time matching to the dansyl standards library consisting of 273 unique standards with 315 entries.[31] This library with information on MS, MS/MS and ion chromatogram for each dansyl labeled standard is freely accessible at www.MyCompoundID.org. Putative identification was done based on accurate mass match to the metabolites in the human metabolome database (HMDB) (8,021 known human endogenous metabolites) and the Evidence-based Metabolome Library (EML) (375,809 predicted human metabolites with one reaction) using MyCompoundID.[32] The mass accuracy tolerance window was set at 8 ppm and the retention time tolerance window set to 20 s for the definitive identification searches and for the putative searches a mass tolerance of 0.005 Da was used.

### 4.3 Results and Discussion

#### 4.3.1 Metabolome Profiles of Different Types of Milk

Using the human milk metabolome or submetabolome as a fingerprint for detecting the presence of an adulterant, analyzing a large number of metabolites with high quantification accuracy and precision is important in order to more readily reveal small changes in the metabolome profile when a small amount of adulterant is added to human milk. Dansylation LC-MS offers a means of analyzing a larger number of the amine/phenol-containing metabolites in milk. We examined the differences of the amine/phenol submetabolomes among different types of milk that may be used as a human milk adulterant/diluent. This profiling work includes 6 different types of milk: cow, goat, almond, soy, infant formula and human breast milk; some have different brands and different fat contents. The 10 samples of human milk were donated from different

individuals collected at various time points postpartum. Experimental triplicates were performed on all samples for a total of 60 labeled samples analyzed by LC-MS. Each individual sample was labeled with $^{12}$C-dansyl chloride. The pooled sample was generated by combining equal volume aliquots of all the different milk samples, metabolite extraction, and $^{13}$C-dansylation. All labeled samples were then quantified by LC-UV and the concentrations of labeled metabolites were determined. The individual samples were then mixed with equal mole amounts of the pooled sample.

The average number of metabolites detected per milk type is as follows: 4318±198 (n=9) goat metabolites, 4256±136 (n=18) cow metabolites, 3080±470 (n=9) almond metabolites, 4129±297 (n=9) soy metabolites, 4444±563 (n=9) infant formula metabolites and 4020±375 (n=30) human metabolites. These results indicate that a high coverage analysis of the amine/phenol submetabolome was achieved for all the milk samples.

We next examined whether the submetabolomes are different from each other for different types of milk samples. Figure 4.1A shows the PCA score plots of the six different types of milk with the quality control (QC) injections (Figure 4.1A) and without the QC (Figure 4.1B). From the plots, we can see that the QC data clusters well for samples run over the span of three months, indicating good technical reproducibility as well as the stability of the labeled samples during storage. Each type of milk cluster together tightly, but different types of milk separate from each other. The animal milk (cow and goat) cluster closely and are apart from the plant-based milk (almond and soy) which form their own clusters. Although the milk samples seemingly group by source (animal vs. plant), they still cluster individually by milk type. Human milk groups together on its own and shows more variability within the milk type. This is most likely due to genomic

diversity as well as external factors, such as the environment and diet, which was not controlled for this study.

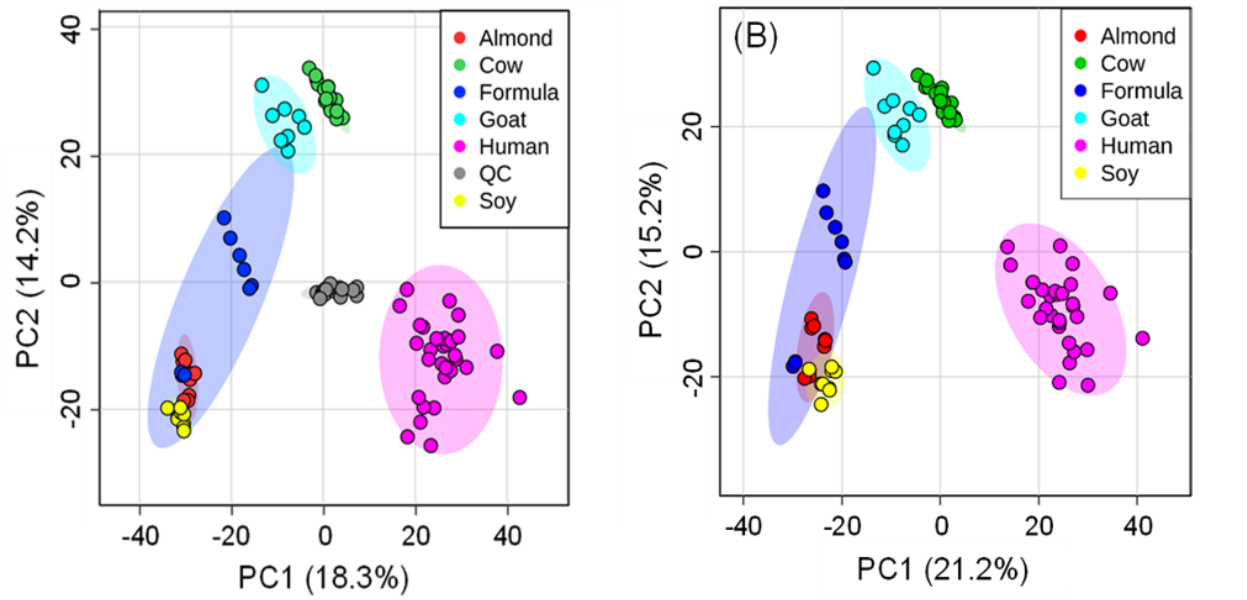



Figure 4.1 PCA plots of different types of milk studied in this work with QC data (A) and without QC data (B).

All the samples cluster together well by milk type with the exception of formula milk. Two of the three brands tested in this study cluster between the soy and cow milk data-points. This is expected as the ingredients listed on the labels for infant formula typically includes both cow and soy-based components. However, one of the brands separates away from the other formula milk, and is found closer to soy milk. The finding was confirmed by analyzing different lots of this brand of formula purchased over 12 months apart. This may suggest a more dominant soy-based or other plant-based milk recipe for this brand of infant formula, which does boast a more digestible formula; the dominance of plant-based components may be responsible for that. This example illustrates that the submetabolome profiles can be useful to differentiate these brands.

### 4.3.2 Metabolite Differences of Different Types of Milk

The data from the metabolome profiling of the different types of milk was sorted and the metabolites unique to each milk type were determined. The following criteria were used: the metabolite must be detected in two of the three experimental triplicates for every sample (i.e., brand) of that milk type. For example, three brands of soy milk were purchased. Each brand was extracted three times for experimental triplicates. After analysis, if metabolite A is found in two of the three experimental replicates for one brand it is counted as a "real" metabolite of that brand. If that same metabolite A is also detected as a "real" metabolite in the other two brands, using the same criteria, it is determined as a soy milk metabolite (versus a brand-specific metabolite if it was only found in the first brand).

Using the criteria mentioned above, the following number of metabolites was determined for each type of milk: 2925 metabolites in goat milk, 2670 in cow milk, 1348 in almond milk, 2562 in soy milk, 1553 in formula milk and 1043 in human milk. The potential human adulterant metabolites were then compared to human milk. Figure 4.2 shows the number of common metabolites between the milk type and human milk in black. In red are the unique metabolites to that specific milk (i.e., only found in that milk type and not in any of the other types of milk). Of the 2670 metabolites detected in cow milk, 851 were in common with human milk and 352 were unique to cow milk. For the 1348 almond metabolites, 129 were unique to almond milk while 488 were in common with human milk. Goat milk had a total of 2925 metabolites, 834 of which were in common to human milk and 510 unique to goat. Infant formula had 518 and 252 metabolites in common with human milk and unique to formula milk, respectively, out of the 1553 detected. Lastly, soy milk had 583 common metabolites with human milk and 971 unique metabolites to soy milk, out of the 2562 detected.

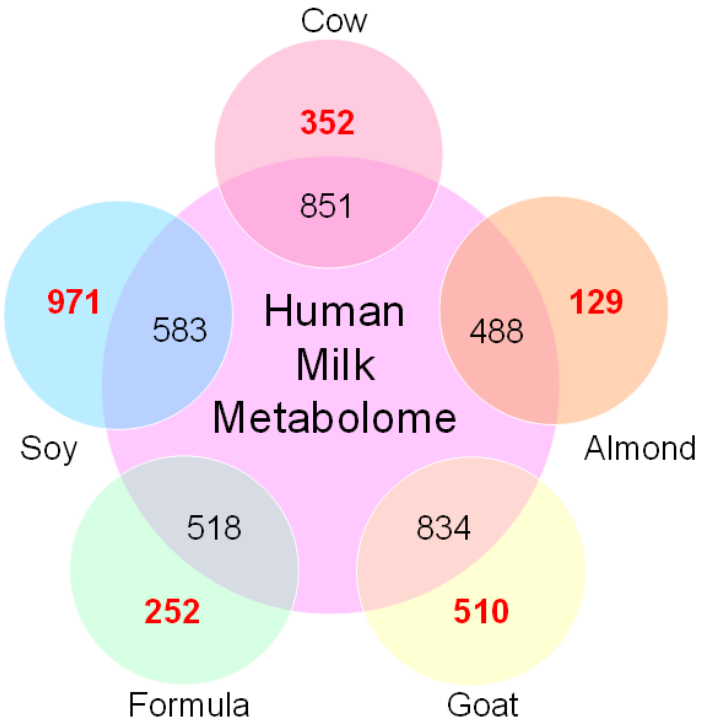



Figure 4.2 Comparison of the numbers of unique and common metabolites for the five different milk adulterants (cow, almond, goat, formula and soy) to those of human milk.

We used a standard library (dansyl standards) and two databases (HMDB and EML) to identify the metabolites detected in each type of milk. Of the 1043 human metabolites, 47 and 863 were matched to the dansyl library and HMDB/EML, respectively (see Supplemental Table S4.1 for the lists). For the 1348 almond metabolites, 43 were matched to the dansyl library and 1076 matched to HMDB/EML (Supplemental Table S4.2). Of the 2562 soy metabolites, 2026 were putatively matched to HMDB/EML while 73 were matched to the dansyl library (Supplemental Table S4.3). The 2670 cow milk metabolites were searched and 78 and 2170 were matched to the dansyl library and HMDB/EML, respectively (Supplemental Table S4.4). 49 infant formula milk metabolites were matched to the dansyl library and 1255 matched to HMDB/EML out of the 1553 formula milk metabolites searched (Supplemental Table S4.5). Lastly of the 2925 goat metabolites

searched, 78 were matched to the dansyl library and 2387 matched to HMDB/EML (Supplemental Table S4.6).

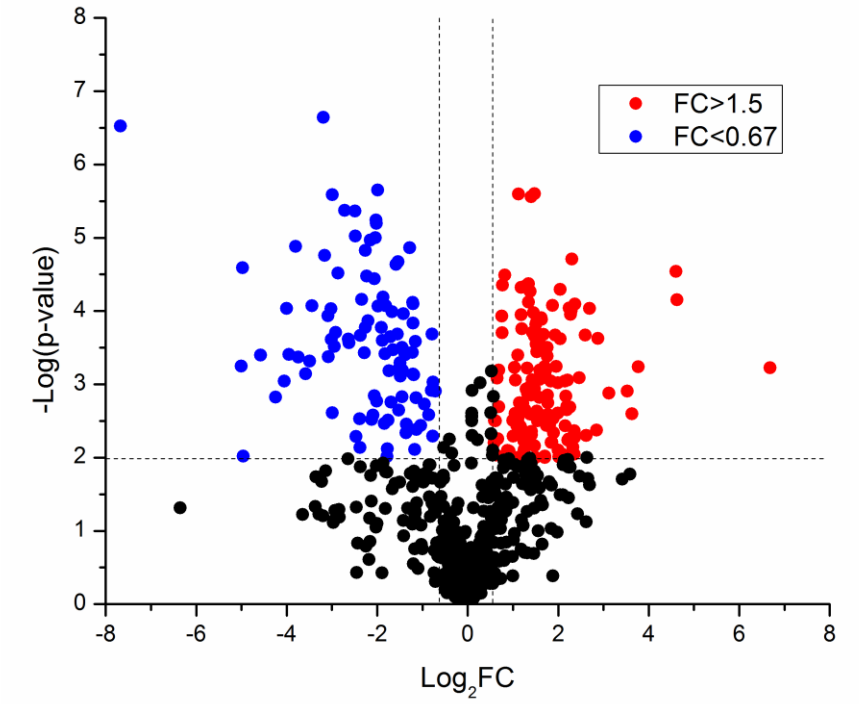



Figure 4.3 Volcano plot for comparison of pooled cow milk and pooled human milk.

The above results show that there are many common metabolites found in different milk, but there are also a large number of unique metabolites found in each type of milk. Moreover, the concentrations of metabolites can be different among different types of milk. One example is shown in Figure 4.3 where the common metabolites found in pooled cow milk and pooled human milk are compared using a volcano plot. There are 146 metabolites (blue dots) having higher concentrations in cow milk than human milk and 103 metabolites (red dots) with lower concentrations, using a criterion of fold-change of greater than 1.5 or less than 0.67 with p-value of less than 0.01. We next explore this metabolome compositional difference to see if we can detect the presence of potential adulterant in human milk.

### 4.3.3 Human Milk mixed with Different Amounts of Potential Adulterants

To investigate whether the CIL-LCMS method is able to distinguish between a non-adulterated and adulterated human milk metabolome at various contamination levels, we prepared a series of mixtures of human milk and potential adulterant (see Figure 4.4). Firstly, all the different samples were pooled by milk type. Then the pooled human milk was diluted with a milk adulterant in the following percentages: 5%, 10%, 25%, 50%, and 75%. Experimental triplicates were performed for each level of milk adulteration (5 in total) for the 4 potential milk adulterants and the 6 pooled milk samples (analyzed for reference) for a total of 93 samples for LC-MS analysis.

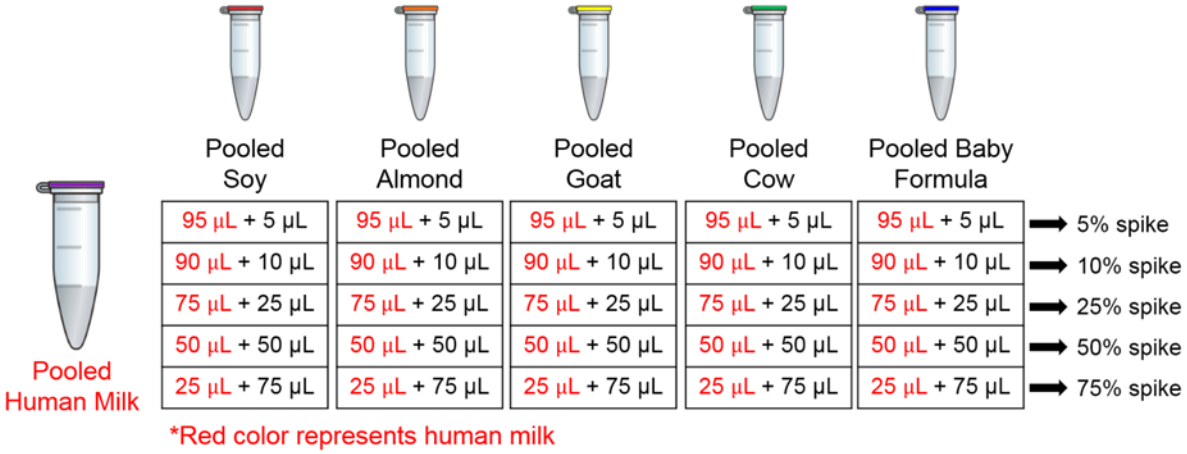



Figure 4.4 Binary mixtures prepared to study the performance of the CIL LC-MS method for differentiating human milk containing varying amounts of adulterant and non-adulterated human milk.

The four potential milk adulterants were chosen based on physical similarities to human milk. From experience, the color and consistency along with transparency of human milk varies depending on the length of time postpartum, maternal weight, diet and time of collection (day vs night).[33] Figure 4.5 shows the PCA analysis of the different levels of adulterated human milk for

different types of milk adulterants. The most common human milk diluent is cow milk for a variety of reasons. There are obvious similarities in the color and consistency (4 different fat percentages to choose from) of the milk. It is also easily accessible and it is the cheapest in price. A recent publication found that approximately 11% of the human donor milk samples purchased online tested positive for cow milk in levels that would suggest intentional contamination (>10% dilution).[7]

Figure 4.5A shows the PCA scores plot of the human milk spiked with varying levels of pooled cow milk. On 2-dimensional plots, the first principal component (PC1) describes the direction that contains the most amount of variation between the metabolites. The second principal component (PC2) captures the direction that contains the next most variation. From the scores plot we see a slight separation between the 5% contaminated human milk and the pure human milk data points on PC2, but no separation on PC1. With 10% contamination, we see a large separation on PC2, but not PC1. Adding 25% or more cow milk, we see an obvious separation on PC1. Based on the PCA plot shown in Figure 4.5A, we can clearly differentiate the adulterated human milk containing 10% or more cow milk from non-adulterated human milk. To differentiate human milk and human milk with 5% cow milk, we used univariate analysis (i.e., volcano plots) to find 5 metabolites in common (see Supplemental Table S4.7 including putative identities) between cow and human milk with concentrations much higher in cow milk than in human milk.

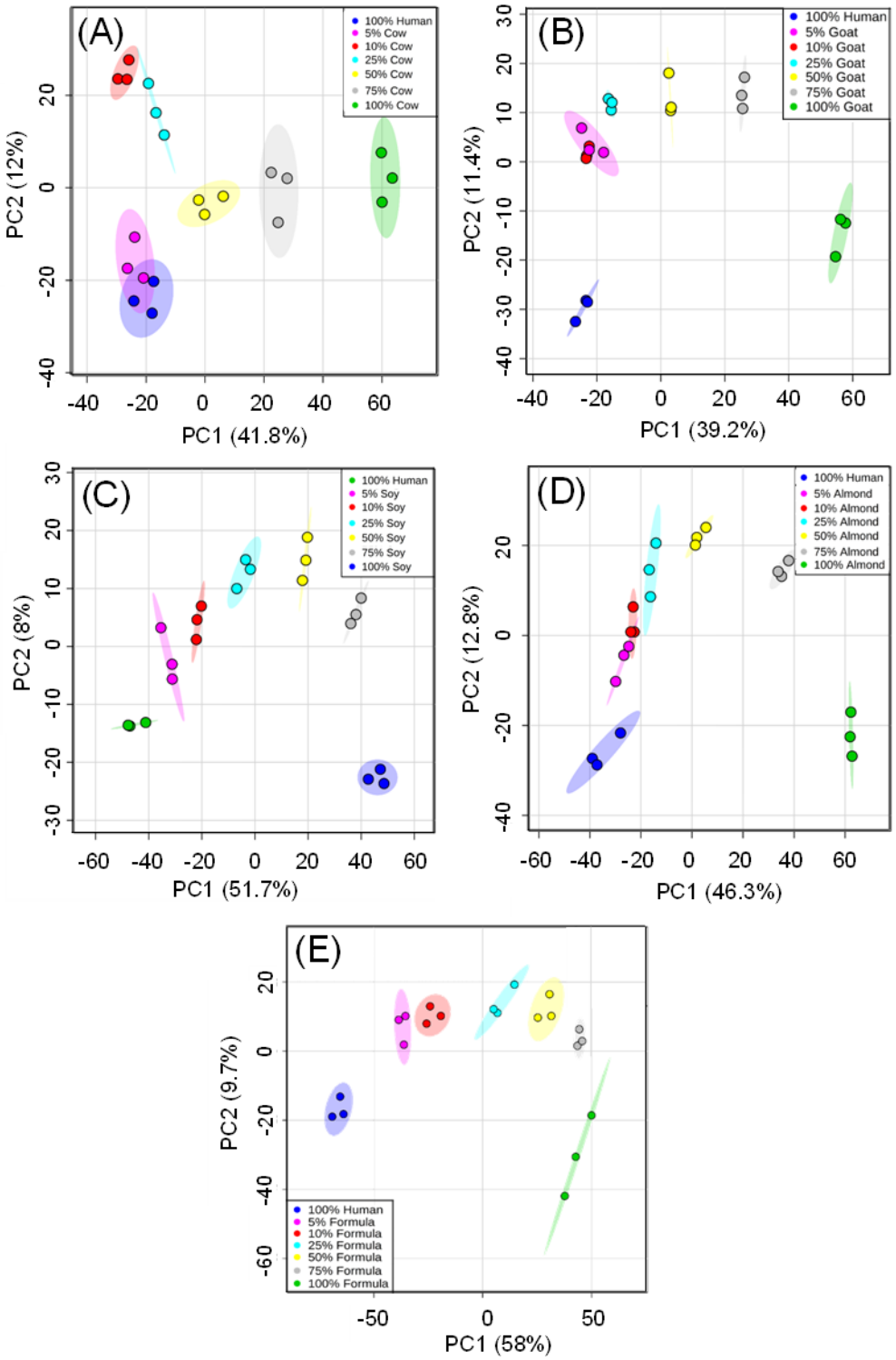



Figure 4.5 PCA plots for the pooled human milk (100%) and adulterated human milk with the following adulterant: (A) cow, (B) goat, (C) soy, (D) almond and (E) infant formula. The level of adulterant varied from 5%, 10%, 25%, 50%, to 75%.

Figure 4.6 shows 5 common metabolites detected in the human milk samples with different levels of cow milk contamination. Our technique is able to detect differences in the fold changes of these 5 metabolites as the human milk is spiked with increasing amounts of cow milk. Thus, even at the 5% level, the submetabolome profiles can be used to detect the adulterated human milk.

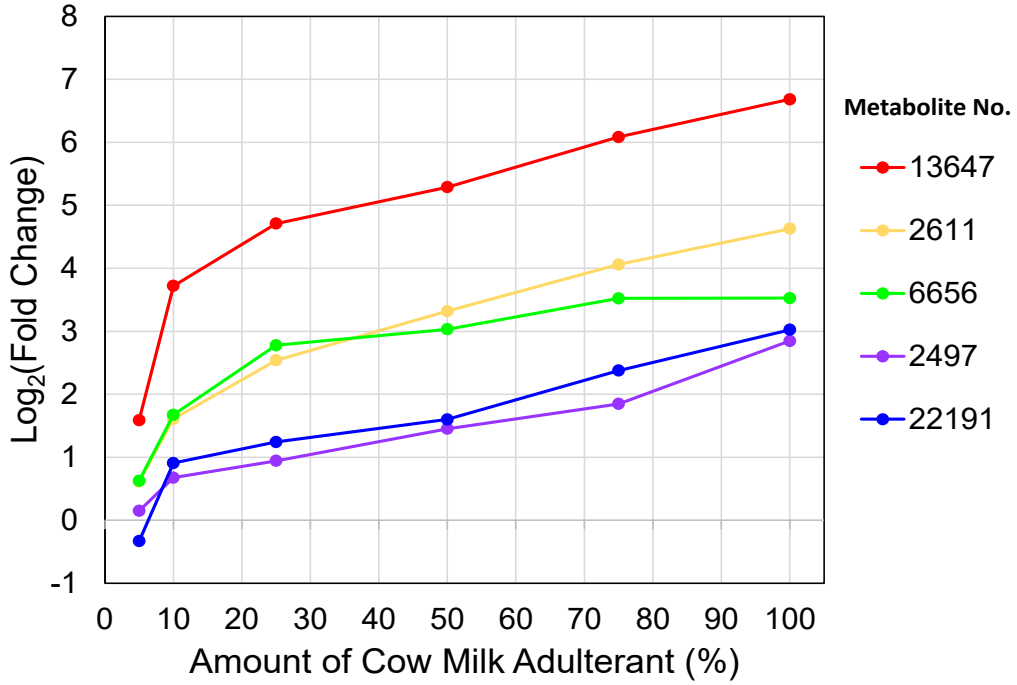



Figure 4.6 Fold changes of 5 metabolites with increased levels of cow milk added to the human milk. The putative IDs of metabolites with corresponding Metabolite No. are provided in Supplemental Table S4.7.

The next adulterant we studied was goat milk. Again, human milk was spiked with varying amounts of goat milk. Figure 4.5B shows the PCA scores plot for the data from this experiment. The human milk diluted with 5% or more goat milk can be readily separated from human milk.

Moving onto the plant based milk types, Figure 4.5C and D show the PCA scores plots for human milk adulterated with soy and almond milk, respectively. For human milk spiked with soy milk we can see separation between all milk samples (non-adulterated human and 5-100% soy

milk). For almond milk, we also see clear separation between non-adulterated human milk and human milk with 5% or more almond milk. Lastly, we investigated commercially available infant formula which consists of cow and soy components. Here we see individual clustering of each milk sample. We see separation between non-adulterated human milk and human milk with ≥5% formula milk.

Overall, we are able to detect significant differences between human milk and milk that has been adulterated with over 5% of foreign milk chosen in this study using multivariate or univariate analysis. It should be noted that it is less likely for purchased donor milk to contain levels less that 5-10% as the costs of milk adulteration would not outweigh the profits (i.e. transportation fees, cost of the milk purchased, recycling fee associated with the container, time and effort spent overall, etc.).

### 4.4 Conclusions

We have developed a CIL LC-MS metabolomics method to determine differences between the amine/phenol submetabolomes of human, soy, almond, cow, goat and infant formula milk. We have shown that this method is able to detect metabolomic differences between non-adulterated human milk and human milk containing 5% or more potential milk adulterant. It is a relatively simple and straightforward method. The overall strategy described in this work should also be applicable to analyze other types of food or drinks in the general application area of food authentication.[34,35] Future technical development work will focus on optimizing the sample processing procedure to shorten the overall analysis time to less than 1 hour per sample.

## 4.5 Literature Cited


(1) Briere, C.-E.; McGrath, J.; Cong, X.; Cusson, R. *J. Obstet. Gynecol. neonatal Nurs. JOGNN* **2014**, *43*, 272–281.

(2) Parker, M. G.; Patel, A. L. *Semin. Perinatol.* **2017**, *41*, 175–186.

(3) Kair, L. R.; Flaherman, V. J. *J. Hum. Lact.* **2017**, 89033441771641.

(4) St-Onge, M.; Chaudhry, S.; Koren, G. *Can. Fam. Physician* **2015**, *61*, 143–146.

(5) Geraghty, S. R.; McNamara, K. A.; Dillon, C. E.; Hogan, J. S.; Kwiek, J. J.; Keim, S. A. *Breastfeed. Med.* **2013**, *8*, 474–478.

(6) Kim, J.; Unger, S. *Paediatr. Child Health* **2010**, *15*, 595–602.

(7) Keim, S. A.; Kulkarni, M. M.; McNamara, K.; Geraghty, S. R.; Billock, R. M.; Ronau, R.; Hogan, J. S.; Kwiek, J. J. *Pediatrics* **2015**, *135*, e1157-62.

(8) Ziegler, E. E. *Nestle Nutr. Workshop Ser. Pediatr. Program.* **2007**, *60*, 185-96-9.

(9) Basnet, S.; Schneider, M.; Gazit, A.; Mander, G.; Doctor, A. *Pediatrics* **2010**, *125*, e973-7.

(10) Doron, D.; Hershkop, K.; Granot, E. *Clin. Nutr.* **2001**, *20*, 259–261.

(11) Ziegler, E. E. *Nutr. Rev.* **2011**, *69*, S37–S42.

(12) Christie, L.; Hine, R. J.; Parker, J. G.; Burks, W. *J. Am. Diet. Assoc.* **2002**, *102*, 1648–1651.

(13) Ho, M. H.-K.; Wong, W. H.-S.; Chang, C. *Clin. Rev. Allergy Immunol.* **2014**, *46*, 225–240.

(14) Anagnostopoulos, A. K.; Katsafadou, A. I.; Pierros, V.; Kontopodis, E.; Fthenakis, G. C.;

Arsenos, G.; Karkabounas, S. C.; Tzora, A.; Skoufos, I.; Tsangaris, G. T. *J. Proteomics* **2016**, *147*, 76–84.

(15) Gaso-Sokac, D.; Kovac, S.; Josic, D. *Food Technol. Biotechnol.* **2011**, *49*, 397–412.

(16) Caira, S.; Nicolai, M. A.; Lilla, S.; Calabrese, M. G.; Pinto, G.; Scaloni, A.; Chianese, L.; Addeo, F. *Food Chem.* **2017**, *230*, 482–490.

(17) Yang, Y.; Bu, D.; Zhao, X.; Sun, P.; Wang, J.; Zhou, L. *J. Proteome Res.* **2013**, *12*, 1660–1667.

(18) Scano, P.; Murgia, A.; Pirisi, F. M.; Caboni, P. *J. Dairy Sci.* **2014**, *97*, 6057–6066.

(19) Inoue, K.; Tanada, C.; Sakamoto, T.; Tsutsui, H.; Akiba, T.; Min, J. Z.; Todoroki, K.; Yamano, Y.; Toyo'oka, T. *Food Chem.* **2015**, *181*, 318–324.

(20) Yang, Y.; Zheng, N.; Zhao, X.; Zhang, Y.; Han, R.; Yang, J.; Zhao, S.; Li, S.; Guo, T.; Zang, C.; Wang, J. *J. Proteomics* **2016**, *136*, 174–182.

(21) Li, Q.; Yu, Z.; Zhu, D.; Meng, X.; Pang, X.; Liu, Y.; Frew, R.; Chen, H.; Chen, G. *J. Sci. Food Agric.* **2017**, *97*, 2875–2882.

(22) Murgia, A.; Scano, P.; Contu, M.; Ibba, I.; Altea, M.; Bussu, M.; Demuru, M.; Porcu, A.; Caboni, P. *LWT - Food Sci. Technol.* **2016**, *74*, 427–433.

(23) Scano, P.; Murgia, A.; Demuru, M.; Consonni, R.; Caboni, P. *Food Res. Int.* **2016**, *87*, 76–82.

(24) Guo, K.; Li, L. *Anal. Chem.* **2009**, *81*, 3919–3932.

(25) Everstine, K.; Spink, J.; Kennedy, S. *J. Food Prot.* **2013**, *76*, 723–735.

(26) Mung, D.; Li, L. *Anal. Chem.* **2017**, *89*, 4435–4443.

(27) Wu, Y.; Li, L. *Anal. Chem.* **2012**, *84*, 10723–10731.

(28) Zhou, R.; Tseng, C.-L.; Huan, T.; Li, L. *Anal. Chem.* **2014**, *86*, 4675–4679.

(29) Huan, T.; Li, L. *Anal. Chem.* **2015**, *87*, 1306–1313.

(30) Huan, T.; Li, L. *Anal. Chem.* **2015**, *87*, 7011–7016.

(31) Huan, T.; Wu, Y.; Tang, C.; Lin, G.; Li, L. *Anal. Chem.* **2015**, *87*, 9838–9845.

(32) Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; Lin, G. *Anal. Chem.* **2013**, *85*, 3401–3408.

(33) Ballard, O.; Morrow, A. L. *Pediatr. Clin. North Am.* **2013**, *60*, 49–74.

(34) Riedl, J.; Esslinger, S.; Fauhl-Hassek, C. *Anal. Chim. Acta* **2015**, *885*, 17–32.

(35) Cuadros-Rodríguez, L.; Ruiz-Samblás, C.; Valverde-Som, L.; Pérez-Castaño, E.; González-Casado, A. *Anal. Chim. Acta* **2016**, *909*, 9–23.

# Chapter 5 CIL LC-MS for the Absolute Quantification of Propionic Acid in Human and Rat Plasma Samples

## 5.1 Introduction

Systems biology is the holistic study of the biological system of an organism at the molecular level.[1] It involves the integration from a variety of "omics" technologies including the following (from upstream to downstream): the study of genomics (DNA), transcriptomics (gene expression) proteomics (protein expression), and lastly metabolomics (small molecule products or intermediates of cellular pathways).[1] However the state of an organism is inherently complex. There is a bidirectional relationship between intrinsic and extrinsic factors that needs to be regulated in order to maintain physiological homeostasis.[1,2] The metabolome offers a glimpse of the downstream end products of the multi-interaction relationship between genes, proteins and the environment.[1,3] It is complex and dynamic making it more representative of the organism's phenotype.[1] The metabolome is heavily influenced by the relationship between the host's intrinsic factors (genes) and the environment (exposome and lifestyle).[1,4] The exposome encompasses any eternal exposures the organism may be subjected to (which include diets, medications (xenobiotics), gut microbiota composition, toxins, allergens, pollutants, exercise and stress etc.)[1,4] Microbiomes (external and internal) play a large role in the health of an organism.[5] Microbiome communities occupy every aspect of our daily lives from the items we touch, the food we ingest, air we breathe; inhabiting our skin and colonising in our gut.[5] There is a codependency between the host and the vast colonies of microorganisms that have co-evolved alongside the host.[6] The microbiota plays a role in helping produce essential nutrients, assist with metabolism, boost the immune system and develop and maintain the gastrointestinal tract.[2,6]

*I optimized the sample labeling method, generated the LC-MS method, prepared the samples, collected the data, processed the data, wrote and edited this chapter

Research has shown the importance of the gut microbiome in prevention of a variety of disease and conditions including the some of the following: obesity, asthma, cancer, and digestive disorders.[5,7,8] Another role of the gut microbiota is to break down undigested carbohydrates (prebiotics known as fibre) thereby recovering what would have been wasted energy.[6,8,9] Fibre metabolism helps increase the blood levels of short chain fatty acids (SCFAs) such as propionate, butyrate and acetate.[8,9] These SCFAs have been found to play a variety of roles in maintaining the health and prevention of disease in the organism. SCFAs that are metabolized in the liver, tissues, and muscles provide about 8% of the daily energy requirements for the host.[9] They also play a role in maintaining the intestines and have an anti-microbial, anti-tumorigenic and anti-inflammatory effects thereby modulating metabolism, intestinal health and the immune system.[10]

In this chapter, I report a chemical isotope labeling LC-MS workflow for the absolute quantification of propionic acid that is applicable to both rat and human plasma samples. P-dimethylaminophenacyl (DmPA) bromide was the reagent used for the labeling of propionic acid as the non-labeling approach proved ineffective due to low sensitivity and inability to retain well on the reverse phase LC column.

## 5.2 Experimental

### 5.2.1 Chemical and Reagents

The deuterated propionic acid standard was purchased from Cambridge Laboratories Canada. All other chemicals and reagents, unless otherwise stated, were purchased from Sigma-Aldrich Canada. For dansylation labeling, the $^{12}C$-labeling reagent (p-dimethylaminophenacyl (DmPA) bromide) was synthesized according to the method published previously.[11]

### 5.2.2 Rat and Human Samples

All rat and human plasma samples were received by Dr. Juan Gonzalez Abraldes. Informed consent was obtained from individual volunteers, and ethics approval was obtained from the University of Alberta in compliance with the University of Alberta Heath Information policy.

### 5.2.3 DmPA Labeling

15 µL of plasma spiked with 1 µL of 0.2 mM deuterated PA (internal control) was mixed with 50 µL of acetonitrile for protein precipitation. The solution was vortexed for 30 seconds, spun down and incubated in the -20 °C freezer for 20 minutes. It was then centrifuged at 14 000 rpm for 15 minutes. 40 µL of the supernatant was then removed into a new vial to which 10 µL of 1.0 M triethanolamine (TEOA) and 25 µL of 15 mg/mL DmPA was added. The solution was then vortexed for another 30 seconds and incubated at 80 °C for 1 hour. After the labeling reaction, the solution was centrifuged at 14 000 rpm for 5 minutes. 2 µL of the solution was then injected into the RPLC-MS for analysis.

### 5.2.4 LC-MS

All LC-MS experiments were performed on an Dionex UltiMate 3000 UHPLC (Thermo Scientific, MA) connected to a Bruker Impact HD Quadrupole Time-Of-Flight (QTOF) mass spectrometer (Billercia, MA) with an ESI source. A reverse phase column (Agilent Eclipse Plus C18 column, 2.1 mm x 10 cm, 1.8 µm particle size, 95 Å pore size) was used for liquid chromatography separation of labeled metabolites. Mobile phase A was made up of 5% (v/v) acetonitrile and 0.1% (v/v) formic acid in water. Mobile phase B consists of 0.1% (v/v) formic acid in acetonitrile. The 35-min gradient conditions were as follows: 0 min (20% B), 0-3.5 min (20-35% B), 3.5-10 min (35-50% B), 10-14 min (50-99% B), 14-26.5 min (99% B), 26.5-30 min

(99-20% B) and 30-35 min (20% B). The flow rate ranged from 250 for the analysis and equilibration segments and 300 µL/min for the clean segments.

### 5.2.5 Data Processing

Bruker DataAnalysis software 4.3 was used to obtain the extracted ion chromatogram for DmPA labeling propionic acid and DmPA-deuterated propionic acid. A mass tolerance of ±0.005 Da was used.

## 5.3 Results and Discussion

### 5.3.1 Labeled Approach for Organic Acids

SCFAs, such as propionate, are small, volatile, polar compounds which do not retain well on a reverse phase (RP) column. Previously, p-dimethylaminophenacyl (DmPA) bromide has been shown to label carboxylic acid containing metabolites thereby increase their retention ability on a RP column and increasing ESI efficiency in positive ion mode.[11] A modified labeling procedure was determined by a fellow lab-mate, Wei Han. The modified procedure was as follows: 15 µL of plasma was mixed with 50 µL of acetonitrile for protein precipitation. The sample was incubated at -20 °C for 20 minutes. Then it was centrifuged at 14 000 rpm for 10 minutes. After which, the 45 µL of supernatant was removed into a new tube and the protein pellet was discarded. 10 µL of 1.0 M triethanolamine (TEOA) and 25 µL of 25 mg/mL DmPA reagent was mixed with the supernatant, vortexted and incubated at 80 °C for 1 hour. After which the labeled solution is ready for LC-MS analysis. This method was further optimized by myself.

### 5.3.2 Determination of the LC-MS Method

As this study focuses on the targeted analysis of propionic acid (PA), the LC-MS method was tailored to that. The LC-MS run is shown below in figure 5.1. In blue is the change in the %

of MPB and in red is the change in % MPA. The LC-MS run was split into four segments (as colored below) including: a mass calibration/waste segment (in red), a MS collection segment (in yellow), a clean segment (in green) followed by a re-equilibration segment (in blue). The LC run starts at 80:20 ratio of MPA to MPB. The first waste segment begins with a 30 second sodium formate calibration section to allow for mass calibration of the individual run. During this segment, the solution eluting off the column is directed to the waste line. At a flow rate of 180 µL/min, PA elutes out from 10.6-11.0 minutes. In order to shorten the run time, the flow rate was increased to 250 µL/min. At an 80:20 mixture of MPA and MPB the pressure is ~ 250 bar. Ideally the column should be able to handle pressures higher than 250 bar however there is a risk of leaking at the fittings which is what I encountered when running at a flow rate of >250 µL/min and therefore did not attempt to switch the flow rate any higher. At this flow rate, PA elutes out at 8.7-9.1 minutes. Therefore at 8.5 minutes, the divert valve is switched from waste to the MS source and MS information is collected from 8.5-9.3 minutes during the "MS collection segment". After the elution of PA, the valve is switched back to waste for the "clean segment". At this point MPB is ramped up to 99% and the flow rate increased to 300 µL/min and will remain so from 14-26.5 minutes labeled sample from the RP column. The last segment is the re-equilibration segment which includes decreasing the flow rate back to 250 µL/min and returning the MPB and MPA to 20% and 80%, respectively.

For determination of the injection volume, 5 µL was chosen during the initial testing of the labeling and LC-MS method. Many peaks were saturated and the column carryover was a potential issue with repeat injections, therefore the injection volume was decreased to 2 uL to avoid dirtying the LC (therefore avoiding additional clean runs) and MS.

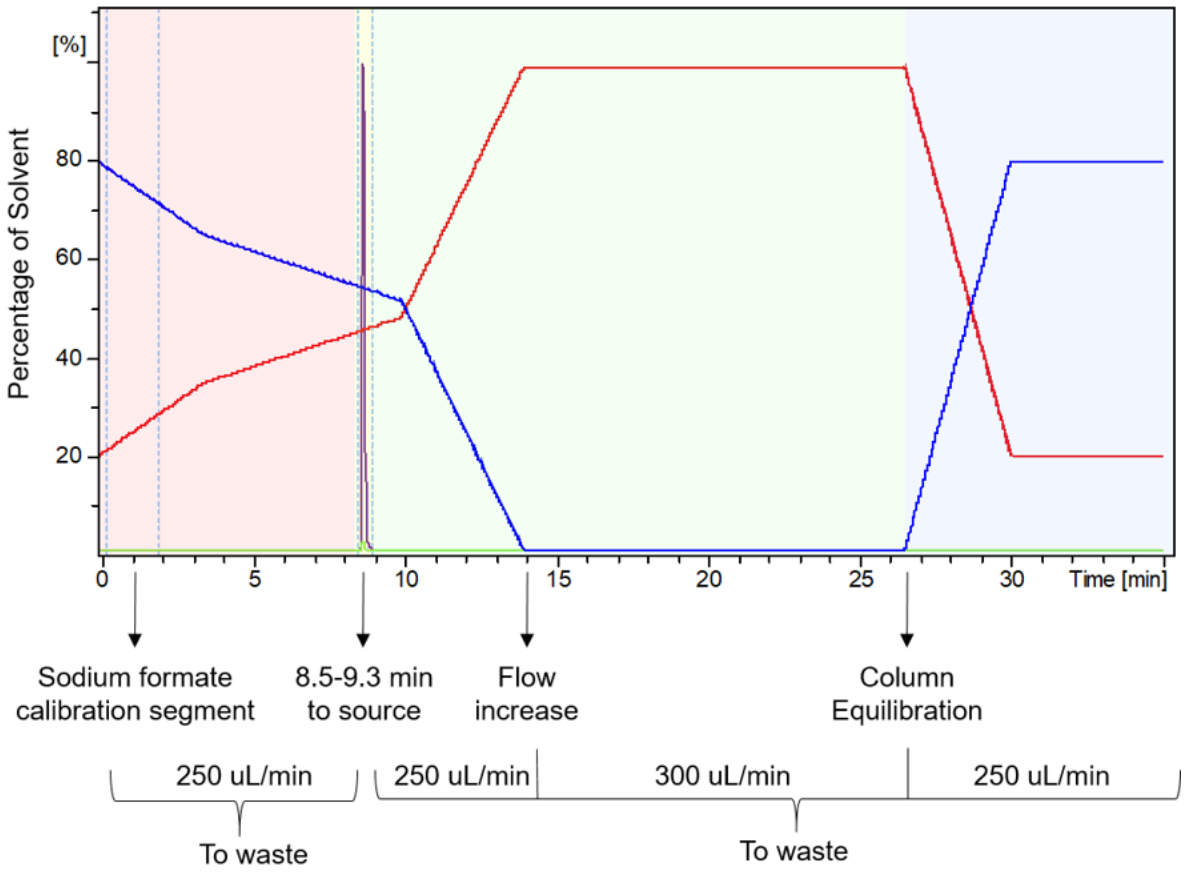



Figure 5.1 The LC-MS method for the targeted quantification of DmPA-labeled propionic acid. In blue and red is the percentage of mobile phase A and B, respectively.

### 5.3.3 Optimization of the Labeling Conditions

Once the LC-MS method was determined, the following labeling conditions were optimized: temperature, TEOA and DmPA concentration. For the optimization of each condition, only the condition of interest was changed while the remaining conditions for the labeling procedure were held constant. Experimental triplicates (n=3) were also performed for each condition. Different protein precipitation solvent was not compared due to the high temperature of the reaction. Therefore, temperature was first tested to investigate if the reaction could proceed sufficiently at lower temperatures. From Figure 5.2A we see that lowering the reaction temperature resulted in a lower peak area detected for propionic acid (PA). The reaction temperature of 80 °C corresponded to the highest peak area. Therefore, in regards to the protein precipitation solvent,

acetonitrile remained the best choice due to its high boiling point of ~82 °C. Other solvents such as methanol or acetone who have much lower boiling points (65 °C and 56 °C, respectively) would result in increased sample loss during the high temperature incubation period. With acetonitrile protein precipitation of heparin-containing plasma, the protein does not form the most compact pellet at the bottom of the vial. Therefore, after protein precipitation, 40 µL of supernatant was removed (instead of the previous 45 µL) to avoid accidental disturbance of the pellet.

Next, the TEOA concentration was varied. Figure 5.2B shows the peak area detected for PA remained relatively the same when the concentration of TEOA was varied from 0.1-1.0 M. Lastly, the concentration of labeling reagent DmPA was investigated. Figure 5.2C shows the peak area for PA increasing as the concentration of DmPA decreased from 25 mg/mL to 15 mg/mL.

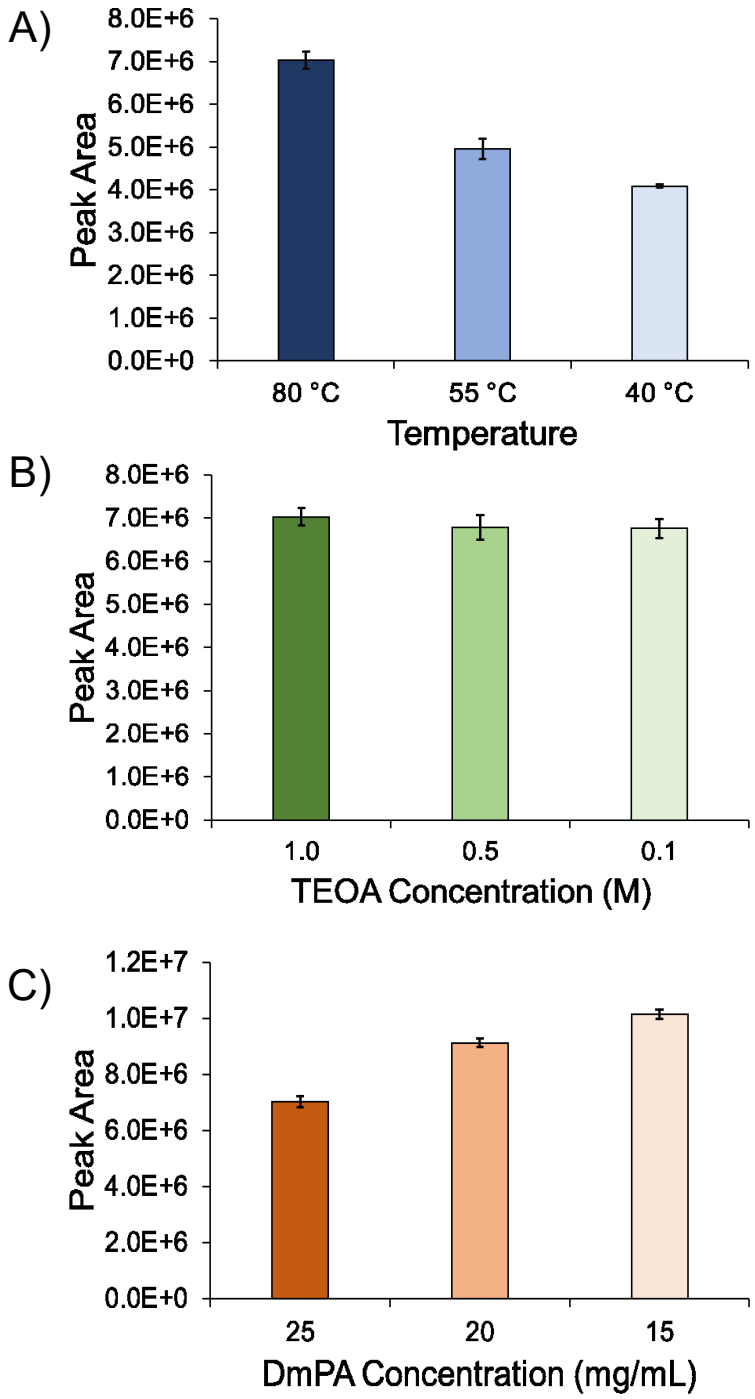



Figure 5.2 The peak area of propionic acid at varying experimental conditions including the following: (A) temperature, (B) TEOA concentration and (C) DmPA concentration. Experimental triplicates (n=3) were performed.

With the following information, a new optimized method was determined which involved the use of 0.1 M TEOA, 15 mg/mL DmPA and a reaction temperature of 80 °C. Comparing all the data (Figure 5.2 A-C) from the varying conditions together (Figure 5.3), we observed that although the newly determined method did correspond to a higher peak area for PA observed when compared to the old method, it however did not correspond to the highest peak area for PA overall. Rather the method used during the DmPA optimization experiment (using 15 mg/mL shown in pale orange) proved best for the labeling of PA and was subsequently used for the analysis of the real samples.

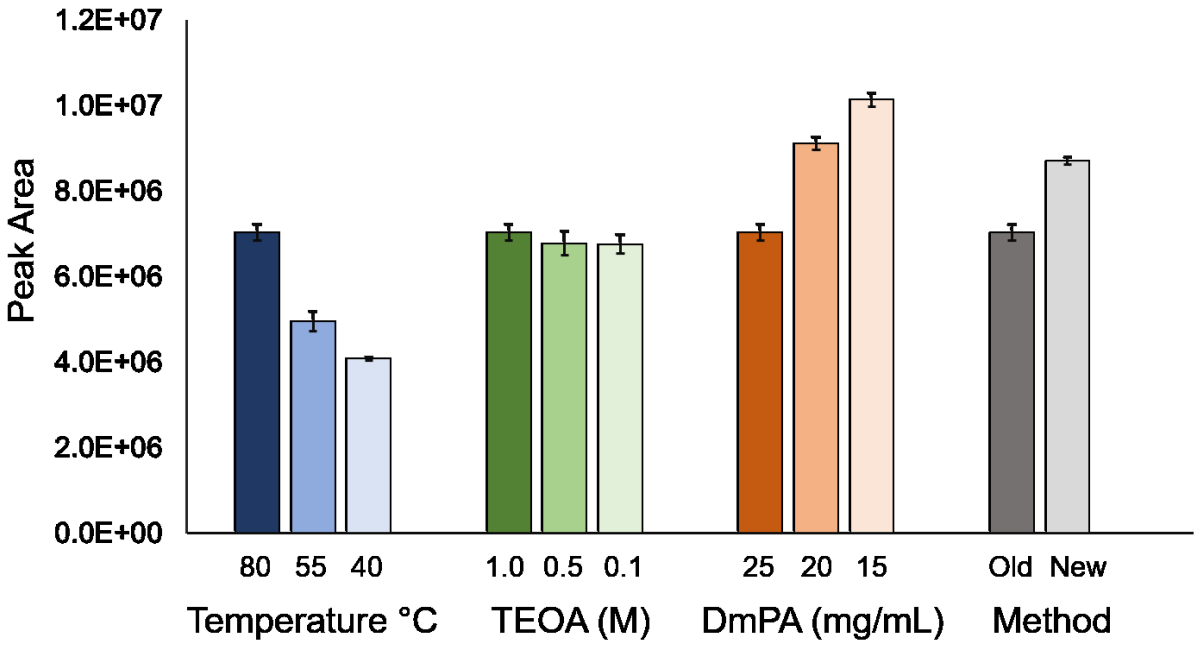



Figure 5.3 Comparison of the peak area for propionic acid obtained from all the optimization experiments including: temperature, TEOA and DmPA concentration and from the newly determined method versus the original method. Experimental triplicates (n=3) was performed.

### 5.3.4 Final Labeling Procedure for Quantitative Analysis of Propionic Acid

The final labeling procedure is as follows. 15 µL of plasma spiked with 1 µL of 0.2 mM deuterated PA (internal control) was mixed with 50 µL of acetonitrile for protein precipitation. The solution was vortexed for 30 seconds, spun down and incubated in the -20 °C freezer for 20 minutes. It was then centrifuged at 14 000 rpm for 15 minutes. 40 µL of the supernatant was then removed into a new while to which 10 µL of 1.0 M TEOA and 25 µL of 15 mg/mL was added. The solution was then vortexed for another 30 seconds and incubated at 80 °C for 1 hour. After the labeling reaction, the solution was centrifuged at 14 000 rpm for 5 minutes to spin down any remaining particles (labeled proteins that precipitate out of solution). 2 µL of the solution was then injected into the RPLC-MS for analysis. Using Bruker's DataAnalysis 4.3, the EICs obtained for DmPA-labeled PA and DmPA-labeled deuterated PA was observed to co-elute at a retention time of ~8.8 minutes (Figure 5.4A). The mass spectrum shows the two labeled metabolites as a peak pair with a mass difference of 2.01 Da between the light and heavy peaks (Figure 5.4B).

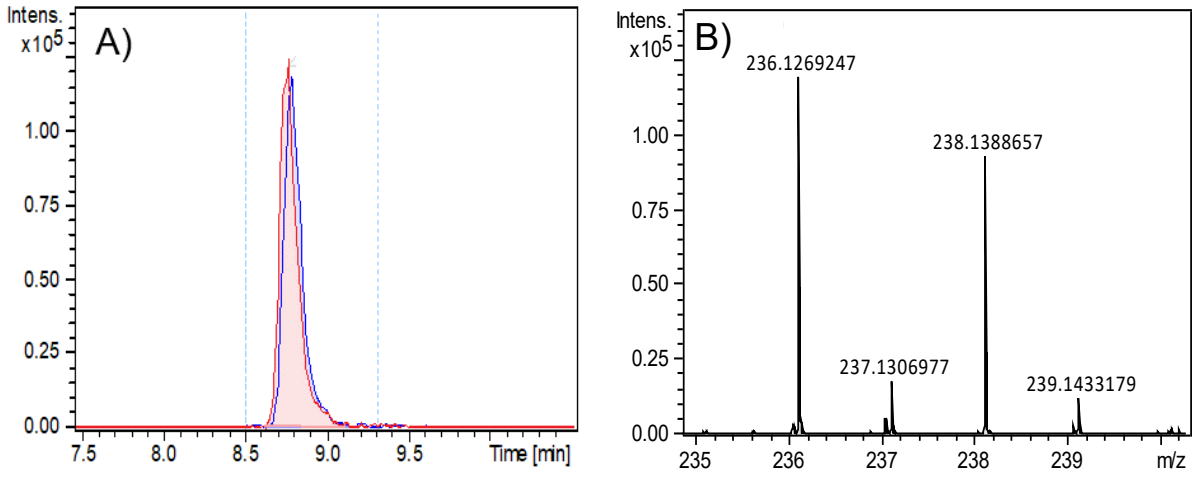



Figure 5.4 (A) The extracted ion chromatograms of the co-eluting DmPA-labeled propionic acid (blue) and DmPA-labeled deuterated propionic acid (red) and (B) the respective mass spectrum for the labeled metabolites observed as a peak pair.

## 5.4 Conclusion and Future Work

In this work, a chemical isotope labeling method was developed for the targeted quantification of propionic acid in plasma. The labeling conditions, temperature and reagent concentrations, were optimized. A 35-minute LC-MS method was also generated for the analysis of the labeled sample. For my future work, I want to extend and optimize this method for the targeted quantification analysis of the other SCFAs found in plasma.

## 5.5 Literature Cited


(1) Claus, S. P.; Swann, J. R. *Annu. Rev. Food Sci. Technol* **2013**, *4*, 381–399.

(2) Renz, H.; Holt, P. G.; Inouye, M.; Logan, A. C.; Prescott, S. L.; Sly, P. D. *J. Allergy Clin. Immunol.* **2017**, *140*, 24–40.

(3) Collino, S.; Martin, F.-P. J.; Kochhar, S.; Rezzi, S. *Nutrients* **2009**, *1*, 101–110.

(4) Swann, J. R.; Claus, S. P. *Sci. Prog.* **2014**, *97*, 41–47.

(5) Dai, D.; Prussin, A. J.; Marr, L. C.; Vikesland, P. J.; Edwards, M. A.; Pruden, A. *Environ. Sci. Technol.* **2017**, acs.est.7b01097.

(6) Fernstrand, A. M.; Bury, D.; Garssen, J.; Verster, J. C. *Food Nutr. Res.* **2017**, *61*, 1297053.

(7) Zitvogel, L.; Daillère, R.; Roberti, M. P.; Routy, B.; Kroemer, G. *Nat. Rev. Microbiol.* **2017**, *15*, 465–478.

(8) Gomez-Llorente, M. A.; Romero, R.; Chueca, N.; Martinez-Cañavate, A.; Gomez-Llorente, C. *Int. J. Mol. Sci.* **2017**, *18*, 1490.

(9) Slavin, J. *Nutrients* **2013**, *5*, 1417–1435.

(10) Tan, J.; McKenzie, C.; Potamitis, M.; Thorburn, A. N.; Mackay, C. R.; Macia, L. *Adv. Immunol.* **2014**, *121*, 91–119.

(11) Guo, K.; Li, L. *Anal. Chem.* **2010**, *82*, 8789–8793.

(12) Guo, K.; Li, L. *Anal. Chem.* **2010**, *82*, 8789–8793.

# Chapter 6 Conclusions and Future Work

## 6.1 Thesis Summary

Metabolomics using reverse phase liquid chromatography mass spectrometry (RPLC-MS) is being merged with other disciplines such as food and nutritional sciences. In chapter 2, I report an analytical workflow for in-depth profiling of the milk metabolome based on chemical isotope labeling (CIL) and liquid chromatography mass spectrometry (LC-MS) with a focus of using dansylation labeling to target the amine/phenol submetabolome. An optimal sample preparation method including the use of methanol at a 3:1 ratio of solvent to milk for protein precipitation and dichloromethane for lipid removal was developed. Experimental replicates found this workflow to be accurate and precise and applicable to profile milk metabolomes of different species (cow, goat and human) and types. In the metabolomic analysis of 36 samples from different categories of cow milk (brands, batches and fat percentages) with experimental triplicates, a total of 7104 peak pairs or metabolites could be detected with an average of 4573±505 (n=108) pairs detected per LC-MS run. Among them, 3820 peak pairs were consistently detected in over 80% of the samples with 70 metabolites positively identified by mass and retention time matches to the dansyl standard library and 2988 pairs with their masses matched to the human metabolome libraries.

In chapter 3, CIL LC-MS was used to determine the dietary effects of cow milk consumption on the human urine metabolome, 1 and 2 hours after intake. Six healthy volunteers were analyzed individually. Dansylation labeling was used to profile the urine amine and phenol submetabolomes. The data from the urine study was then merged with the cow milk metabolome data obtained in chapter 2 in order to detect cow milk metabolites excreted in the urine. Overall there is no significant differences between the before and after-drinking milk urine amine and

phenol submetabolomes when using multivariate analysis. However, subtle differences can be detected in the urine metabolome with the use of Venn Diagrams and Volcano Plots.

Chapter 4 involved the use of CIL LC-MS to profile the metabolome of a foreign milk sample, followed by multivariate or univariate comparison of the resultant metabolomic profile with that of human milk to determine the differences. Using dansylation LC-MS to profile the amine/phenol submetabolome, we could detect an average of 4129±297 (n=9) soy metabolites, 3080±470 (n=9) almond metabolites, 4256±136 (n=18) cow metabolites, 4318±198 (n=9) goat metabolites, 4444±563 (n=9) infant formula metabolites, and 4020±375 (n=30) human metabolites. This high level of coverage allowed us to readily differentiate the six different types of samples. From the analysis of binary mixtures of human milk containing 5, 10, 25, 50 and 75% other type of milk, we demonstrated that this method could be used to detect the presence of as low as 5% adulterant in human milk. We envisage that this method could be applied to detect contaminant or adulterant in other types of food or drinks.

Lastly, chapter 5 involved a targeted approach for metabolomics involving the absolute quantification of propionic acid in plasma samples using RPLC-MS. P-dimethylaminophenacyl (DmPA) bromide was used to label propionic acid and deuterated propionic acid standard spiked into the sample. The following conditions were found to be optimal for labeling: 80 °C reaction temperature, 1M concentration of TEOA and 15 mg/mL concentration of DmPA reagent. An RPLC-MS method was also determined for the targeted analysis of propionic acid.

### 6.2 Future Work

I hope to expand the coverage of the different milk metabolomes (chapter 2 and 4) by utilizing other chemical isotope labeling methods to target the other submetabolomes. For chapter

4 the milk adulteration project, I hope to expand the study and investigate other foreign milk (coconut, hazelnut, cashew, hemp and rice milk etc.).

For the nutrimetabonomics project, I would like to carefully design a more thorough experiment involving more participants and holding constant as many factors as possible without being too invasive to the volunteers (i.e. ask for absolutely no milk or milk-product intake at least 24 hours prior to the urine experiment). As well, I hope to also utilize different labeling chemistries to investigate the effects of milk consumption on other urine submetabolomes.

Lastly for the absolute quantification project, I hope to generate a method to absolutely quantify other short chain fatty acids such as formic acid, acetic acid, butyric acid, valeric acid.

# Bibliography


(1) Ahmed, M. U.; Saaem, I.; Wu, P. C.; Brown, A. S. Crit. Rev. *Biotechnol*. **2014**, *34*, 180–196.

(2) Anagnostopoulos, A. K.; Katsafadou, A. I.; Pierros, V.; Kontopodis, E.; Fthenakis, G. C.; Arsenos, G.; Karkabounas, S. C.; Tzora, A.; Skoufos, I.; Tsangaris, G. T. J. *Proteomics* **2016**, *147*, 76–84.

(3) Andreas, N. J.; Hyde, M. J.; Gomez-Romero, M.; Lopez-Gonzalvez, M. A.; Villasenor, A.; Wijeyesekera, A.; Barbas, C.; Modi, N.; Holmes, E.; Garcia-Perez, I. *Electrophoresis* **2015**, *36*, 2269-2285.

(4) Astarita, G.; Langridge, J. J. Nutrigenet. *Nutrigenomics* **2013**, *6*, 181–200.

(5) Ballard, O.; Morrow, A. L. *Pediatr. Clin. North Am*. **2013**, *60*, 49–74.

(6) Basnet, S.; Schneider, M.; Gazit, A.; Mander, G.; Doctor, A. *Pediatrics* **2010**, *125*, e973-7.

(7) Bernabucci, U.; Basirico, L.; Morera, P.; Dipasquale, D.; Vitali, A.; Piccioli Cappelli, F.; Calamari, L. *J Dairy Sci* **2014**, *98*, 1815-1827.

(8) Binkley, J. *Chromatogr.* Online **2010**, 1–7.

(9) Briere, C.-E.; McGrath, J.; Cong, X.; Cusson, R. J. *Obstet. Gynecol. neonatal Nurs. JOGNN* **2014,** *43*, 272–281.

(10) Bruheim, P.; Kvitvang, H. F. N.; Villas-Boas, S. G. *J Chromatogr A* **2013**, *1296*, 196-203.

(11) Caboni, P.; Murgia, A.; Porcu, A.; Demuru, M.; Pulina, G.; Nudda, A. *J Dairy Sci* **2016**, *99*, 6046-6051.

(12) Caira, S.; Nicolai, M. A.; Lilla, S.; Calabrese, M. G.; Pinto, G.; Scaloni, A.; Chianese, L.; Addeo, F. *Food Chem*. **2017**, *230*, 482–490.

(13) Cappozzo, J. C.; Koutchma, T.; Barnes, G. *J Dairy Sci* **2015**, *98*, 5068-5079.

(14) Cesare Marincola, F.; Dessì, A.; Corbu, S.; Reali, A.; Fanos, V. *Clin. Chim. Acta* **2015**, *451*, 103–106.

(15) Cevallos-Cevallos, J. M.; Reyes-De-Corcuera, J. I. *Advances in food and nutrition research* **2012**, *67*, 1–24.

(16) Christie, L.; Hine, R. J.; Parker, J. G.; Burks, W. J. Am. *Diet. Assoc*. **2002**, *102*, 1648–1651.

(17) Cilla, A.; Quintaes, K. D.; Barbera, R.; Alegria, A. *Crit Rev Food Sci Nutr* **2016**, *56*, 1880-1892.

(18) Claus, S. P.; Swann, J. R. *Annu. Rev. Food Sci. Technol* **2013**, *4*, 381–399.

(19) Claus, S. P.; Swann, J. R. *Annu. Rev. Food Sci. Technol*. **2013**, *4*, 381–399.

(20) Collino, S.; Martin, F.-P. J.; Kochhar, S.; Rezzi, S. *Nutrients* **2009**, *1*, 101–110.

(21) Croissant, A. E.; Washburn, S. P.; Dean, L. L.; Drake, M. A. *J Dairy Sci* **2007**, *90*, 4942-4953.

(22) Crutchfield, C. A.; Thomas, S. N.; Sokoll, L. J.; Chan, D. W. *Clin. Proteomics* **2016**, *13*, 1.

(23) Cuadros-Rodríguez, L.; Ruiz-Samblás, C.; Valverde-Som, L.; Pérez-Castaño, E.; González-Casado, A. *Anal. Chim. Acta* **2016**, *909*, 9–23.

(24) Dai, D.; Prussin, A. J.; Marr, L. C.; Vikesland, P. J.; Edwards, M. A.; Pruden, A. *Environ. Sci. Technol*. **2017**, *151*, 7759-7774.

(25) Dai, W. D.; Huang, Q.; Yin, P. Y.; Li, J.; Zhou, J.; Kong, H. W.; Zhao, C. X.; Lu, X.; Xu, G. W. *Anal Chem* **2012**, *84*, 10245-10251.

(26) Dinis-Oliveira, R. J. *Drug Metab. Rev*. **2012**, *48*, 568–576.

(27) Doron, D.; Hershkop, K.; Granot, E. *Clin. Nutr.* **2001**, *20*, 259–261.

(28) Everstine, K.; Spink, J.; Kennedy, S. J. *Food Prot.* **2013,** *76*, 723–735.

(29) Fan, R. J.; Guan, Q.; Zhang, F.; Leng, J. P.; Sun, T. Q.; Guo, Y. L. *Anal Chim Acta* **2016,** *908*, 132-140.

(30) Fang, Z.-Z.; Gonzalez, F. J. Arch. *Toxicol*. **2014,** *88*, 1491–1502.

(31) Fernstrand, A. M.; Bury, D.; Garssen, J.; Verster, J. C. *Food Nutr. Res.* **2017**, *61*, 1297053.

(32) Forcisi, S.; Moritz, F.; Kanawati, B.; Tziotis, D.; Lehmann, R.; Schmitt-Kopplin, P. *J. Chromatogr*. A **2013**, *1292*, 51–65.

(33) Franklin, S.; Vondriska, T. M. *Circ. Cardiovasc. Genet.* **2011**, *4*, 576.

(34) Gagneur, J.; Stegle, O.; Zhu, C.; Jakob, P.; Tekkedil, M. M.; Aiyar, R. S.; Schuon, A.-K.; Pe'er, D.; Steinmetz, L. M. *PLoS Genet.* **2013**, *9*, e1003803.

(35) Gaso-Sokac, D.; Kovac, S.; Josic, D. *Food Technol. Biotechnol.* **2011**, *49*, 397–412.

(36) Geraghty, S. R.; McNamara, K. A.; Dillon, C. E.; Hogan, J. S.; Kwiek, J. J.; Keim, S. A. *Breastfeed. Med*. **2013**, *8*, 474–478.

(37) Gibbons, H.; O'Gorman, A.; Brennan, L. *Curr Opin Lipidol* **2015***, 26,* 30-34.

(38) Gill, B. D.; Indyk, H. E. *International Dairy Journal* **2007,** *17*, 596-605.

(39) Gill, B. D.; Indyk, H. E.; Manley-Harris, M. *International Dairy Journal* **2012**, *24,* 40-43.

(40) Gomez-Llorente, M. A.; Romero, R.; Chueca, N.; Martinez-Cañavate, A.; Gomez-Llorente, C. *Int. J. Mol. Sci*. **2017**, *18,* 1490.

(41) Gowda, G. A. N.; Djukovic, D. *Methods Mol. Biol.* **2014**, *1198*, 3–12.

(42) Guo, K.; Li, L. *Anal. Chem*. **2009**, *81*, 3919–3932.

(43) Guo, K.; Li, L. *Anal. Chem*. **2010**, *82*, 8789–8793.

(44) Hampel, D.; Allen, L. H. *Crit Rev Food Sci Nutr* **2016**, *56*, 494-511.

(45) Hao, L.; Zhong, X. F.; Greer, T.; Ye, H.; Li, L. J. *Analyst* **2015**, *140*, 467-475.

(46) Hess, J. M.; Jonnalagadda, S. S.; Slavin, J. L. *Compr. Rev. Food. Sci. Food Saf*. **2016**, *15*, 251-268.

(47) Ho, M. H.-K.; Wong, W. H.-S.; Chang, C. *Clin. Rev. Allergy Immunol*. **2014**, *46*, 225–240.

(48) Horgan, R. P.; Kenny, L. C. *Obstet. Gynaecol*. **2011**, *13*, 189–195.

(49) Huan, T.; Li, L. *Anal. Chem*. **2015**, *87*, 7011–7016.

(50) Huan, T.; Li, L. *Anal. Chem*. **2015**, *87*, 1306–1313.

(51) Huan, T.; Wu, Y.; Tang, C.; Lin, G.; Li, L. *Anal. Chem.* **2015,** *87*, 9838–9845.

(52) Inoue, K.; Tanada, C.; Sakamoto, T.; Tsutsui, H.; Akiba, T.; Min, J. Z.; Todoroki, K.; Yamano, Y.; Toyo'oka, T. *Food Chem*. **2015**, *181*, 318–324.

(53) Jessri, M.; Ng, A. P.; L'Abbé, M. R. *Nutrients* **2017,** *9*, 910.

(54) Jones, D. P.; Park, Y.; Ziegler, T. R. *Annu. Rev. Nutr*. **2012**, *32*, 183–202.

(55) Kair, L. R.; Flaherman, V. J. J. *Hum. Lact*. **2017**, 89033441771641.

(56) Keim, S. A.; Kulkarni, M. M.; McNamara, K.; Geraghty, S. R.; Billock, R. M.; Ronau, R.; Hogan, J. S.; Kwiek, J. J. *Pediatrics* **2015**, *135*, e1157-62.

(57) Kim, J.; Unger, S. *Paediatr. Child Health* **2010**, *15*, 595–602.

(58) Kim, S.; Kim, J.; Yun, E. J.; Kim, K. H. *Curr Opin Biotechnol* **2016**, *37*, 16-23.

(59) Klein, M. S.; Almstetter, M. F.; Schlamberger, G.; Nurnberger, N.; Dettmer, K.; Oefner, P. J.; Meyer, H. H. D.; Wiedemann, S.; Gronwald, W. *J Dairy Sci* **2010**, *93*, 1539-1550.

(60) Lenz, E. M.; Wilson, I. D. *Journal of Proteome Research.* **2007**, *6,* 443–458.

(61) Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; Lin, G. *Anal. Chem*. **2013**, *85*, 3401–3408.

(62) Li, Q.; Yu, Z.; Zhu, D.; Meng, X.; Pang, X.; Liu, Y.; Frew, R.; Chen, H.; Chen, G. *J. Sci. Food Agric*. **2017**, *97*, 2875–2882.

(63) Liu, P.; Huang, Y. Q.; Cai, W. J.; Yuan, B. F.; Feng, Y. Q. *Anal Chem* **2014**, *86*, 9765-9773.

(64) Lock, K.; Pomerleau, J.; Causer, L.; Altmann, D. R.; McKee, M. *Bull. World Health Organ*. **2005**, *83*, 100–108.

(65) Mayne, S. T.; Playdon, M. C.; Rock, C. L. *Nat Rev Clin Oncol* **2016**, *13*, 504-515.

(66) McNiven, E. M. S.; German, J. B.; Slupsky, C. M. *J. Nutr. Biochem*. **2011**, *22*, 995–1002.

(67) Michaëlsson, K.; Wolk, A.; Langenskiöld, S.; Basu, S.; Warensjö Lemming, E.; Melhus, H.; Byberg, L. *BMJ* **2014**, *349,* g6015.

(68) Mung, D.; Li, L. *Anal. Chem*. **2017**, *89*, 4435–4443.

(69) Munir, M. T.; Yu, W.; Young, B. R.; Wilson, D. I.; I2C *Trends Food Sci Technol* ***2015***, 43, 205-218.

(70) Murgia, A.; Scano, P.; Contu, M.; Ibba, I.; Altea, M.; Bussu, M.; Demuru, M.; Porcu, A.; Caboni, P. *LWT - Food Sci. Technol*. **2016**, *74*, 427–433.

(80) Nasrollahi, S. M.; Imani, M.; Zebeli, Q. *J Dairy Sci* **2015**, *98*, 8926-8939.

(81) Okada, T. Am. *J. Clin. Nutr*. **2004**, *80*, 1088-9-90.

(82) Oliver, S. P.; Boor, K. J.; Murphy, S. C.; Murinda, S. E. *Foodborne Pathog Dis* **2009**, *6,* 793-806.

(83) Palladino, R. A.; Buckley, F.; Prendiville, R.; Murphy, J. J.; Callan, J.; Kenny, D. A. *J Dairy Sci* **2010**, *93*, 2176-2184.

(84) Parker, M. G.; Patel, A. L. *Semin. Perinatol.* **2017**, *41*, 175–186.

(85) Pereira, P. C. *Nutrition* **2014**, *30*, 619–627.

(86) Putri, S. P.; Nakayama, Y.; Matsuda, F.; Uchikata, T.; Kobayashi, S.; Matsubara, A.; F ukusaki, E. *J. Biosci. Bioeng.* **2013**, *115*, 579–589.

(87) Renz, H.; Holt, P. G.; Inouye, M.; Logan, A. C.; Prescott, S. L.; Sly, P. D. *J. Allergy Clin. Immunol.* **2017**, *140,* 24–40.

(88) Riedl, J.; Esslinger, S.; Fauhl-Hassek, C. *Anal. Chim. Acta* **2015**, *885*, 17–32.

(89) Rocchi, S.; Caretti, F.; Gentili, A.; Curini, R.; Perret, D.; Perez-Fernandez, V. *Food Chem* **2016**, *211*, 455-464.

(90) Roche, J. R.; Turner, L. R.; Lee, J. M.; Edmeades, D. C.; Donaghy, D. J.; Macdonald, K. A.; Penno, J. W.; Berry, D. P. *Anim. Prod. Sci* **2009**, *49*, 222-232.

(91) Rochfort, S. *J. Nat. Prod*. **2005**, *68*, 1813–1820.

(92) Rotroff, D. M.; Motsinger-Reif, A. A. *Int. J. Genomics* **2016**, *2016*, 1715985.

(93) Rozenberg, S.; Body, J.-J.; Bruyère, O.; Bergmann, P.; Brandi, M. L.; Cooper, C.; Devogelaer, J.-P.; Gielen, E.; Goemaere, S.; Kaufman, J.-M.; Rizzoli, R.; Reginster, J.-Y. *Calcif. Tissue Int*. **2016**, *98*, 1–17.

(94) Scano, P.; Murgia, A.; Demuru, M.; Consonni, R.; Caboni, P. *Food Res. Int.* **2016**, *87*, 76–82.

(95) Scano, P.; Murgia, A.; Pirisi, F. M.; Caboni, P. *J. Dairy Sci.* **2014**, *97*, 6057–6066.

(96) Slavin, J. *Nutrients* **2013**, *5*, 1417–1435.

(97) Smilowitz, J. T.; O'Sullivan, A.; Barile, D.; German, J. B.; Lonnerdal, B.; Supsky, C. M. *J Nutr* **2013**, *143*, 1709-1718.

(98) Sousa, S. G.; Delgadillo, I.; Saraiva, J. A. *Crit Rev Food Sci Nutr* **2016**, *56*, 1043-1060.

(99) Spevacek, A. R.; Smilowitz, J. T.; Chin, E. L.; Underwood, M. A.; German, J. B.; Slupsky, C. M. *J Nutr* **2015**, *145*, 1698-1708.

(100) St-Onge, M.; Chaudhry, S.; Koren, G. *Can. Fam. Physician* **2015**, *61*, 143–146.

(101) Sundekilde, U. K.; Downey, E.; O'Mahony, J. A.; O'Shea, C. A.; Ryan, C. A.; Kelly, A. L.; Bertram, H. C. *Nutrients* **2016**, *8*, 15.

(102) Swann, J. R.; Claus, S. P. *Sci. Prog*. **2014**, *97*, 41–47.

(103) Tan, J.; McKenzie, C.; Potamitis, M.; Thorburn, A. N.; Mackay, C. R.; Macia, L. *Adv. Immunol*. **2014**, *121*, 91–119.

(104) Tang, Y. A.; Li, L. *Anal Chim Acta* **2013**, *792*, 79-85.

(105) Tayyari, F.; Gowda, G. A. N.; Gu, H. W.; Raftery, D. *Anal Chem* **2013**, *85,* 8715-8721.

(106) Tian, H.; Zheng, N.; Wang, W. Y.; Cheng, J. B.; Li, S. L.; Zhang, Y. D.; Wang, J. Q. *Sci Rep* **2016**, *61*, 1149-1158.

(107) Villasenor, A.; Garcia-Perez, I.; Garcia, A.; Posma, J. M.; Fernandez-Lopez, M.; Nicholas, A. J.; Modi, N.; Holmes, E.; Barbas, C. *Anal Chem* **2014**, *86*, 8245-8252.

(108) Vrijheid, M. *Thorax* **2014**, *69*, 876–878.

(109) Vuckovic, D. *Anal Bioanal Chem* **2012**, *403*, 1523-1548.

(110) Wang, K.; Huang, C.; Nice, E. *Biomed. Chromatogr.* **2014**, *28*, 848–857.

(111) Warner, D.; Podesta, S. C.; Hatew, B.; Klop, G.; van Laar, H.; Bannink, A.; Dijkstra, J.J *Dairy Sci* **2015,** *98*, 3383-3393.

(112) Wu, J. F.; Domellof, M.; Zivkovic, A. M.; Larsson, G.; Ohman, A.; Nording, M. L. *Biochem Biophys Res Commun* **2016**, *469*, 626-632.

(113) Wu, Y.; Li, L. *Anal. Chem*. **2012**, *84*, 10723–10731.

(114) Wu, Y.; Li, L. *Anal. Chem.* **2014,** *86,* 9428–9433.

(115) Wu, Y.; Li, L. *J. Chromatogr. A* **2016**, *1430*, 80–95.

(116) Yang, Y. X.; Zheng, N.; Zhao, X. W.; Zhang, Y. D.; Han, R. W.; Yang, J. H.; Zhao, S. G.; Li, S. L.; Guo, T. J.; Zang, C. J., et al. *J. Proteomics* **2016,** *136,* 174-182.

(117) Yang, Y.; Bu, D.; Zhao, X.; Sun, P.; Wang, J.; Zhou, L. *J. Proteome Res*. **2013**, *12*, 1660–1667.

(118) Zhang, H. M.; Nan, Z. R.; Hui, G. Q.; Liu, X. H.; Sun, Y. *Genet. Mol. Res.* **2014**, *13*, 198–204.

(119) Zhang, T.; Watson, D. G. *Analyst* **2015**, *140*, 2907–2915.

(120) Zhao, S.; Luo, X.; Li, L. *Anal Chem* **2016**, *88,* 10617-10623.

(121) Zheng, H.; Clausen, M. R.; Dalsgaard, T. K.; Bertram, H. C. *Nutrients* **2015**, *7*, 4875-4896.

(122) Zhou, R.; Tseng, C.-L.; Huan, T.; Li, L. *Anal. Chem.* **2014**, *86*, 4675–4679.

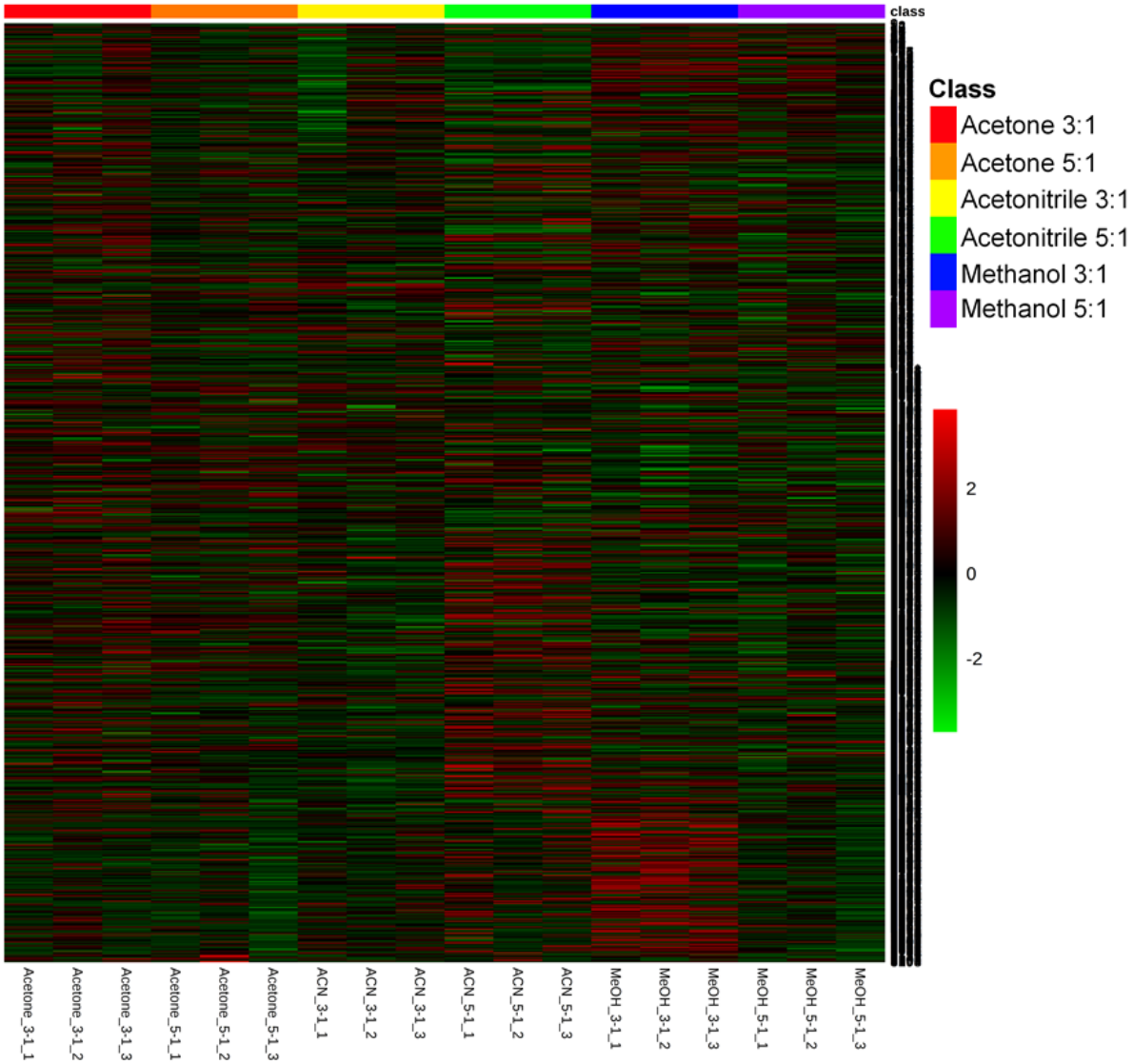
(123) Ziegler, E. E. *Nestle Nutr. Workshop Ser. Pediatr. Program*. **2007**, *60,* 185-96-9.

(124) Ziegler, E. E. *Nutr. Rev.* **2011**, *69*, S37–S42.

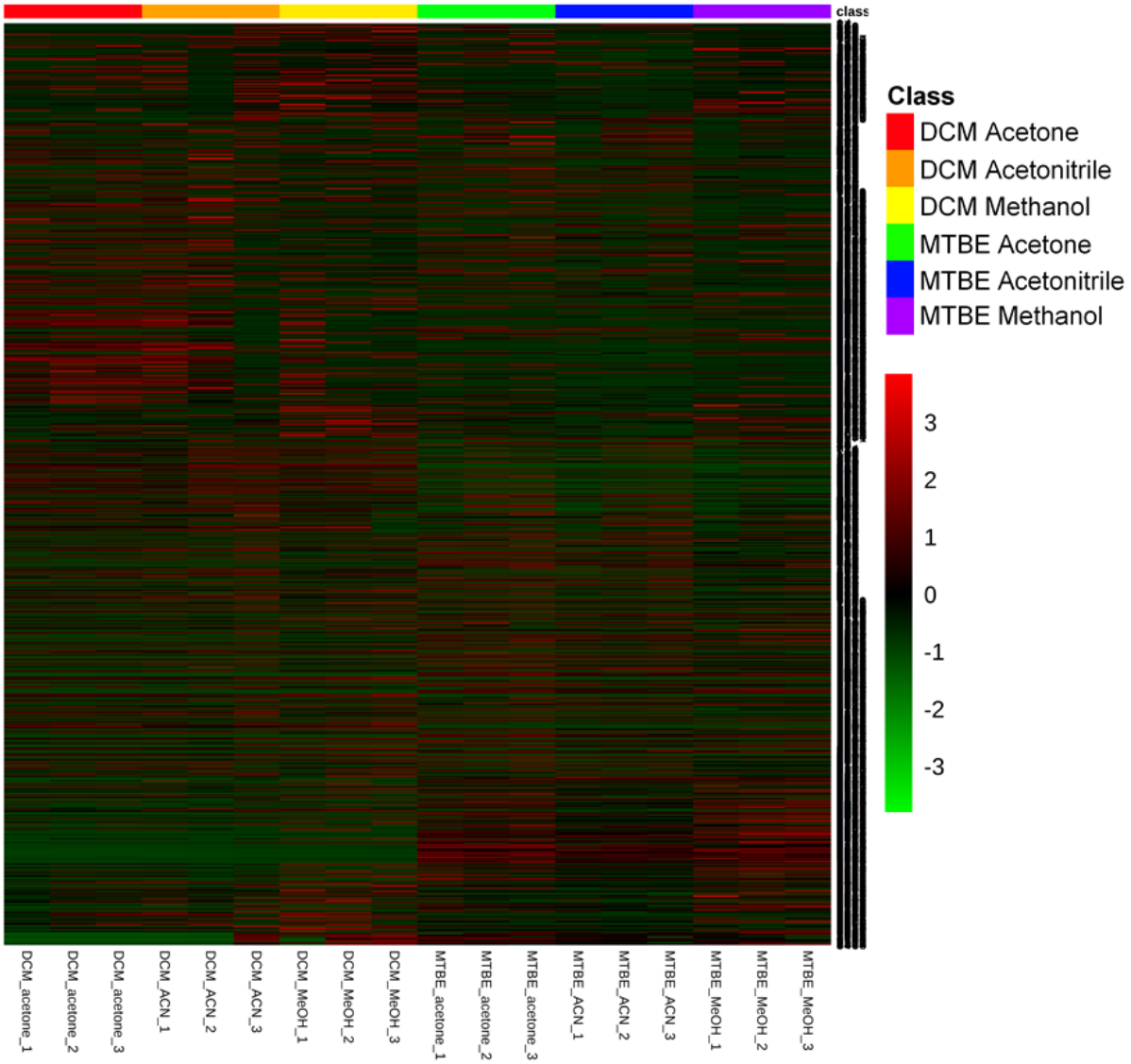
(125) Zitvogel, L.; Daillère, R.; Roberti, M. P.; Routy, B.; Kroemer, G. *Nat. Rev. Microbiol.* **2017**, *15*, 465–478.

# Appendix

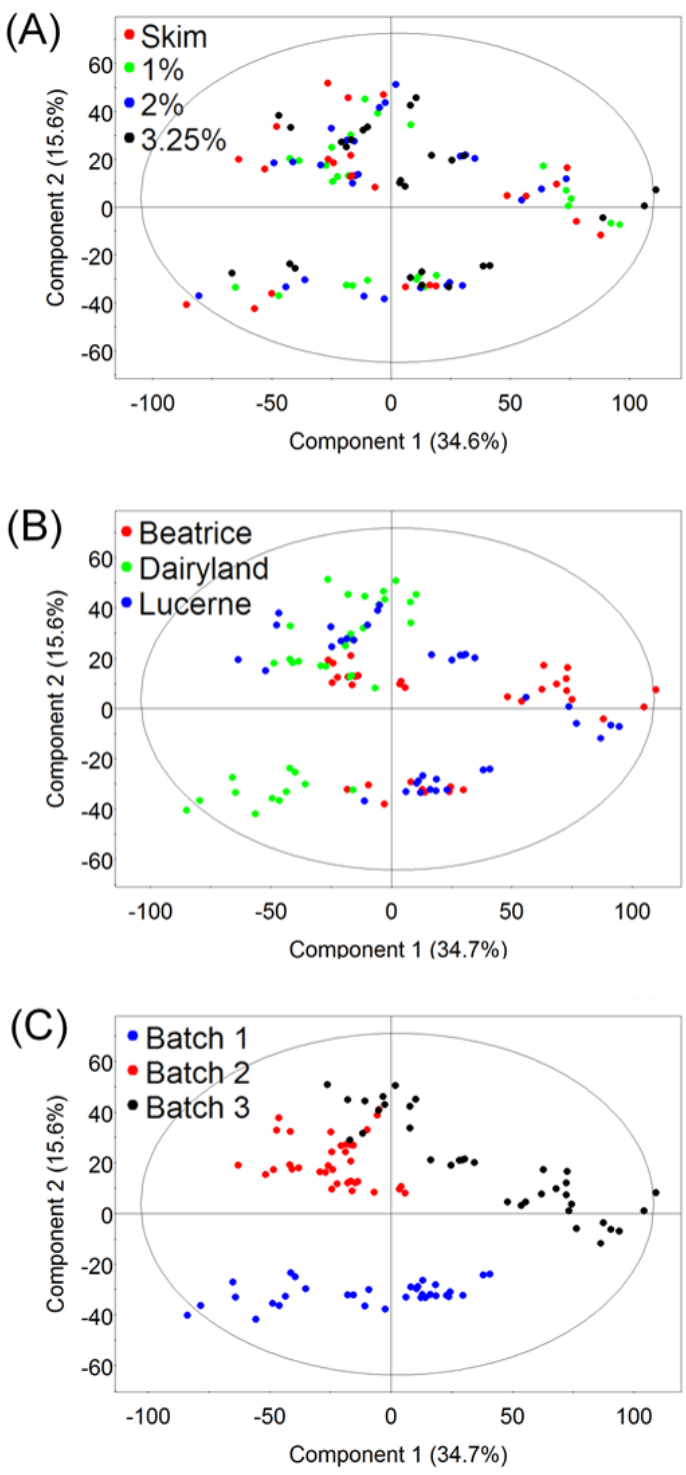
Chapter 2 CIL LC-MS for Milk Metabolomics: Comprehensive and Quantitative Profiling of the Amine/Phenol Submetabolome



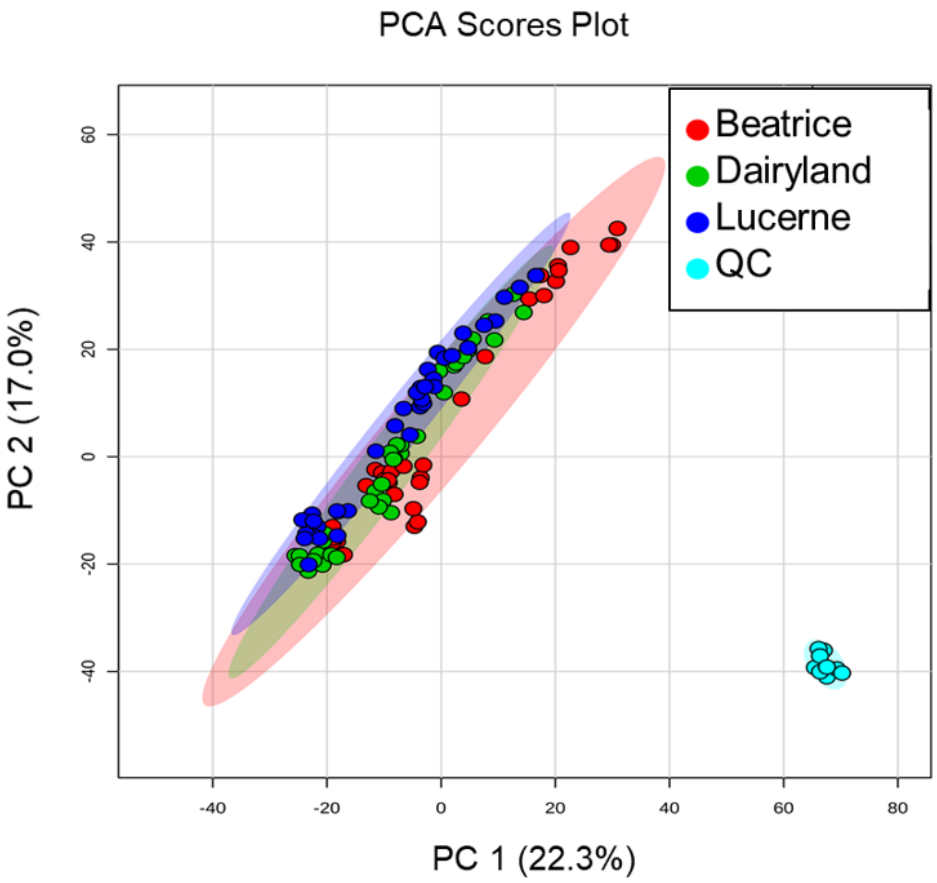

Appendix Figure A2.1. Heatmap of the absolute intensities (light-labeled metabolites) of the milk metabolites extracted by different protein precipitation solvent (acetone, acetonitrile and methanol) and different solvent to sample ratio (3:1 and 5:1).



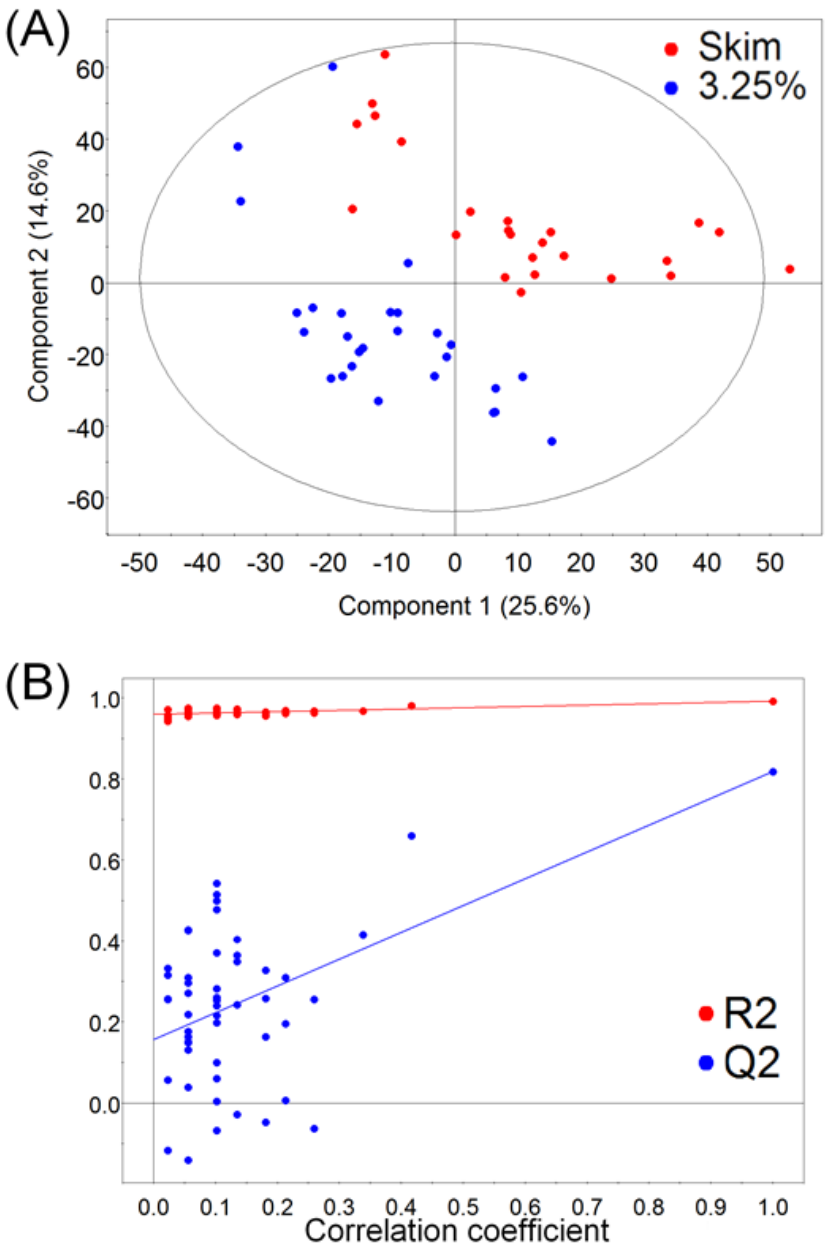

Appendix Figure A2.2. Heatmap of the absolute intensities (light) of the milk metabolites after lipid removal with either dichloromethane (DCM) and methyl-tert-butyl ether (MTBE) in combination with the various protein precipitation solvents.



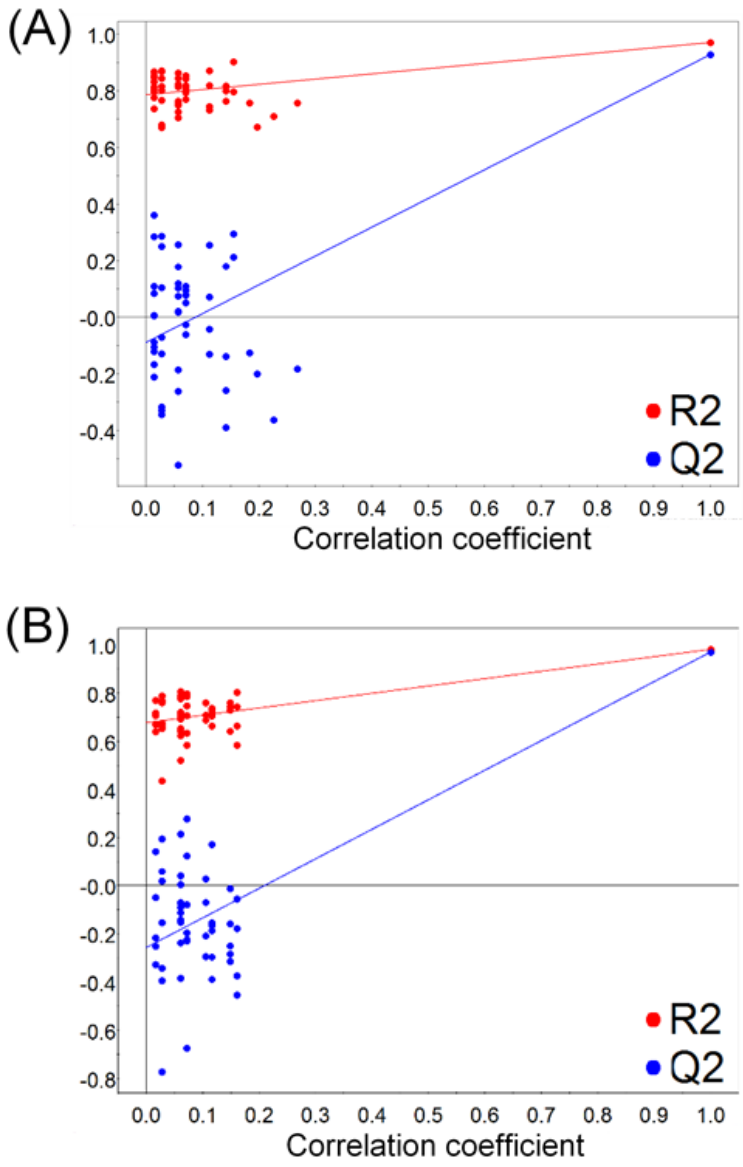

Appendix Figure A2.3. PCA plots of the milk metabolome data set with colored data points according to (A) fat contents, (B) brands and (C) batches.




Appendix Figure A2.4. PCA plot of all the milk samples grouped by brand along with the QC samples. The QC was prepared from one cow milk sample split into two aliquots: one labelled with $^{12}$C-dansylation and another labelled with $^{13}$C-dansylation. The two labelled aliquots were mixed 1:1 and analyzed after every 10 injections. Please note that, for the profiling milk samples, a pooled sample labeled by $^{13}$C-dansylation was used as a reference which was different from the $^{13}$C-labeled QC sample

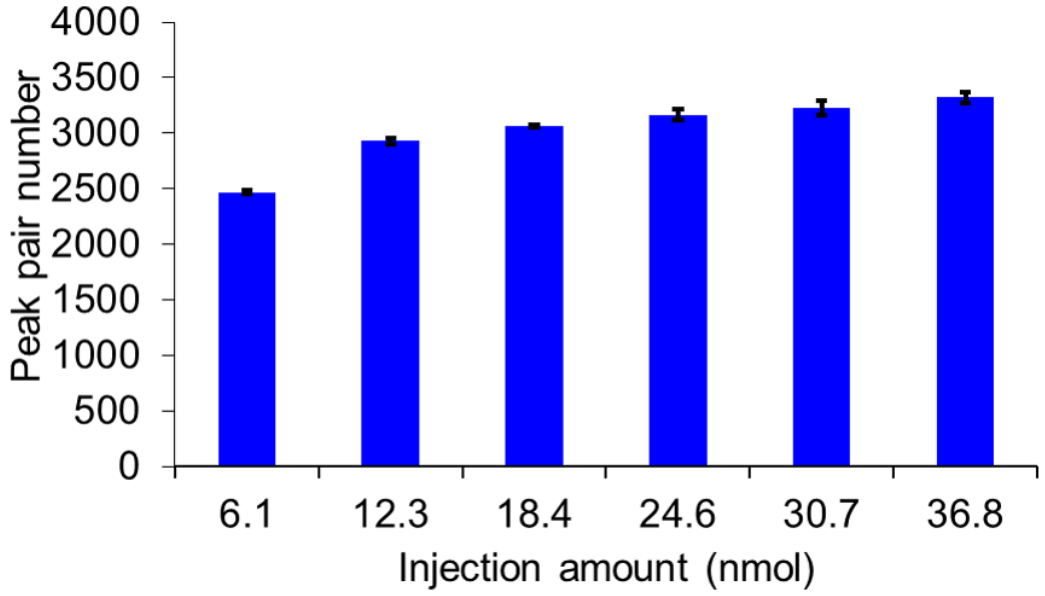



Appendix Figure A2.5. Statistical analysis of the skim and 3.25% fat content milk with the corresponding (A) PLS-DA plot and (B) PLS-DA model validation results.




Appendix Figure A2.6. Validation plots of the milk metabolome data set with colored data points grouped according to (A) brands and (B) batches.

Chapter 4 Development of a Quantitative Metabolomics Method for Detecting Potential Milk Adulterant in Human Milk




Appendix Figure A4.1. Injection optimization curve for determining the optimal sample loading to LC-MS.