

Chemical Isotope Labeling LC-MS for Tissue, Serum and Urine Metabolomics

by

Xiaohang Wang

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

Master of Science

Department of Chemistry  
University of Alberta

© Xiaohang Wang, 2018

## Abstract

Metabolomics is the ultimate reflection of organisms influenced by both genetic and environmental factors. Compared with other omics areas, it is the most appropriate and closest description of phenotype, which is the comprehensive characteristics of an organism. It is a powerful tool for global study of composition, dynamics and responses of metabolites in cells, biofluids, tissues and organs. Metabolomics has been widely used for studying the effects of system perturbations on organisms, such as environmental factors or diseases.

Metabolomics can be divided into two categories, targeted or untargeted. Untargeted metabolomics is global in scope and it simultaneously detects the entire set of metabolites, which is more significant and more promising. However, due to the complexity of metabolites, it is impossible to profile the metabolome by using one single platform. Since the whole metabolome can be divided into different submetabolomes based on different chemical functional groups, it is better to analyze each submetabolome separately to improve metabolite coverage. Of those methods, chemical isotope labeling (CIL) method has been developed due to its various advantages. CIL can add one isotope tag to target different submetabolomes to improve separation, sensitivity and capability of relative quantification. This “divide and conquer” technology enables the study of the whole metabolome using one platform, such as reverse phase liquid chromatography mass spectrometry (RPLC-MS) in positive mode only.

In this work, I applied CIL LC-MS technique to profile the amine/phenol submetabolome of tissue samples. Tissue metabolomics can reveal organ-specific metabolic fingerprints and play a crucial role in investigating specific diseases and sites

of toxicity. I first developed a workflow for carrying out tissue sample processing with a solvent system of methanol/dichloromethane/water. Then, dansylation labeling was applied to profile the amine/phenol submetabolome of mouse brain tissues with Alzheimer's disease (AD). The differences in submetabolome between AD transgenic and wild-type mice were investigated. Several metabolite biomarker candidates have been found with good discriminating power. Then, my developed workflow was also used for amine/phenol submetabolome profiling of rat brain, heart, liver, kidney and muscle tissues with Dexamethasone (Dex) treatment. The side effects of Dex treatment on metabolome were studied with these five kinds of tissues, and some common changes were observed using pathway analysis.

Biofluid samples, such as serum, plasma, urine, saliva, are considered as a pool of metabolites of the body that can reflect some systemic metabolic changes. These samples are relatively easy to obtain and widely used for metabolomics studies. In this work, I applied CIL LC-MS to analyze amine/phenol and carboxyl submetabolomes of serum samples from two cohorts of rheumatoid arthritis (RA) patients. We characterized the submetabolome changes between the early RA and healthy control groups and several potential biomarker candidates were discovered. Besides,  $^{13}\text{C}$ -/ $^{12}\text{C}$ -dansylation labeling LC-MS method were also applied for parallel profiling urine metabolome changes of a mouse model with AD. The metabolic differences between AD and control were observed for both male and female mouse urine samples. The findings of these works indicate that, CIL LC-MS is a comprehensive technique and it is applicable and promising for untargeted metabolomics.

## Acknowledgements

I would like to sincerely thank my supervisor, Dr. Liang Li, for his enthusiasm, motivation, support and patience on my research. I am very lucky to have him as my mentor, who led me to the world of analytical chemistry and metabolomics. I am grateful to work and learn in Dr. Li's lab for the past 3 years. Thanks to him for the joyful activities he organized, the great opportunities he provided and the professional guidance he gave. These memories made my master studies a wonderful experience.

I would like to thank my supervisory committee, Dr. Derrick Clive and Dr. Michael Serpe for their guidance and kind suggestions on my research projects. I would also like to thank Dr. Charles Lucy for participating in my oral defense and reviewing my thesis.

I would like to thank my group members for their kind help and friendship over the years. I would like to thank the all the current members with special thanks to Shuang Zhao and Wei Han, who gave me the lab training and were always willing to help me to overcome obstacles. I would like to thank Yunong Li and Hao Li for the kind assistance with data processing. I would like to thank Xinyun Gu, Dorothea Mung, Minglei Zhu, Xian Luo, Kevin Hooton, Adriana Zardini Buzatto, for the great time we were working together and for the happiness and excitements we experienced.

Finally, I would like to thank my family and friends, for their company and support during the course of my degree. I am gratefully indebted to my parents, Hongqing Wang and Shulan Wang. I would like to thank them for giving birth to me, for bringing me up with their unconditional love and for supporting me through my whole life. Lastly, I would also be appreciative to all of my friends for their encouragement and awesome friendship.

# Table of Contents

List of Figures .....	ix
List of Abbreviations .....	xiii
List of Symbols .....	xv
Chapter 1 Introduction .....	1
1.1 Multi “omics” .....	1
1.2 Metabolomics .....	1
1.2.1 Significances of Metabolomics .....	1
1.2.2 Analytical Platforms for Metabolomics .....	2
1.2.3 Categories of Metabolomics .....	4
1.2.4 Workflow for MS-based Untargeted Metabolomics .....	4
1.2.5 Chemical Isotope Labeling Metabolomics .....	6
1.3 Scope of the Thesis .....	7
Chapter 2 Development of Chemical Isotope Labeling LC-MS for Tissue Metabolomics and Its Application for Brain and Liver Metabolome Profiling in Alzheimer’s Disease Mouse Model .....	9
2.1 Introduction .....	9
2.2 Experimental .....	11
2.2.1 Chemicals and Reagents .....	11
2.2.2 Chicken Liver Sample Collection .....	11
2.2.3 Mouse Model and Sample Collection .....	11
2.2.4 Analytical Workflow .....	12
2.2.5 Tissue Extraction .....	13
2.2.6 Dansylation Labeling .....	13
2.2.7 LC-UV Sample Normalization .....	14
2.2.8 LC-MS .....	14
2.2.9 Data Processing, Metabolite Identification and Statistical Analysis .....	15
2.3 Results and Discussion .....	15
2.3.1 Tissue Metabolite Extraction .....	15

2.3.2. Mouse Liver and Brain Tissue Metabolome Analysis .....	18
2.3.3 Submetabolome Comparison of Tg and WT mice .....	20
2.4 Conclusions .....	25
Chapter 3 Chemical Isotope Labeling LC-MS for Metabolomics of Dexamethasone Side Effects in Rat Tissues .....	26
3.1 Introduction .....	26
3.2 Experimental Section.....	27
3.2.1 Sample Collection and Processing .....	27
3.2.2 Tissue Extraction .....	27
3.2.3 Dansyl Labeling.....	28
3.2.4 LC-UV Quantification.....	28
3.2.5 LC-FTICR-MS Analysis .....	28
3.2.6 Data Processing, Metabolite Identification and Statistical Analysis .....	29
3.3 Result and Discussion.....	29
3.3.1 Metabolite Detection .....	29
3.3.2 Comparative Metabolome Analysis between Dex-treated and Control Groups.....	30
3.3.3 Pathway Analysis .....	33
3.4 Conclusion.....	35
Chapter 4 Development of Chemical Isotope Labeling LC-MS Methods for Metabolomics Studies of Rheumatoid Arthritis Disease .....	37
4.1 Introduction .....	37
4.2 Experimental Section.....	38
4.2.1 Serum Sample Collection and Aliquoting .....	39
4.2.2 Dansyl Labeling.....	40
4.2.3 DmPA Bromide Labeling.....	40
4.2.4 LC-UV Quantification.....	40
4.2.5 LC-QTOF-MS Analysis .....	41
4.2.6 Data Processing, Metabolite Identification and Statistical Analysis .....	41
4.3 Results and Discussion.....	42

4.3.1. Analysis of Amine- and Phenol- Submetabolome.....	42
4.3.1.1 Metabolite Detection.....	42
4.3.1.2 Comparative Metabolome Analysis for RA Biomarker Discovery.....	42
4.3.1.3 Comparative Metabolome Analysis of RA with and without Methotrexate Treatment.....	49
4.3.1.4 Comparative Metabolome Analysis of RA with Anti-CCP Positive and Negative .....	50
4.3.1.5 Comparative Metabolome Analysis between Responders and Non- Responders .....	51
4.3.2 Analysis of Carboxyl Submetabolome .....	53
4.3.2.1 Metabolite Detection.....	53
4.3.2.2 Comparative Metabolome Analysis for RA Biomarker Discovery.....	54
4.3.2.3 Comparative Metabolome Analysis of RA with and without Methotrexate Treatment.....	58
4.3.2.4 Comparative Metabolome Analysis of RA with Anti-CCP Positive and Negative .....	60
4.3.2.5 Comparative Metabolome Analysis between Responders and Non- responders .....	61
4.4 Conclusion.....	61
Chapter 5 Development of Chemical Isotope Labeling LC-MS for Metabolite Biomarker Discovery of Alzheimer’s Disease in a Mouse Model .....	63
5.1 Introduction .....	63
5.2 Experimental Section.....	64
5.2.1 Sample Collection and Processing .....	64
5.2.2 Dansyl Labeling.....	65
5.2.3 LC-UV Quantification.....	65
5.2.4 LC-FTICR-MS Analysis .....	66
5.2.5 Data Processing, Metabolite Identification and Statistical Analysis .....	66
5.3 Results and Discussion.....	67
5.3.1 Metabolite Detection .....	67
5.3.2 Submetabolome Comparison of Tg and WT Mice.....	67

5.3.3 Receiver Operating Characteristic Curves .....	72
5.3.4 Time Series Analysis .....	73
5.3.5 Significance of Potential Metabolite Biomarkers .....	73
5.4 Conclusion .....	75
Chapter 6 Conclusions .....	76
Bibliography .....	78
Appendix .....	93

## List of Figures

- Figure 1.1 The scheme of four CIL reactions, including: (a) Dansyl chloride for amine/phenol submetabolome; (b) Dansyl chloride with base-activation for hydroxyl submetabolome; (c) p-Dimethylaminophenacyl (DmPA) bromide for carboxyl submetabolome; (d) Dansylhydrazine (DnsHz) for carbonyl submetabolome. ....6
- Figure 2.1. Workflow for tissue extraction and dansylation isotope labeling LC-MS. .... 12
- Figure 2.2. Effects of different tissue extraction conditions on peak pair detection of  $^{12}\text{C}$ -/ $^{13}\text{C}$ -dansyl labeled chicken liver samples: (a) two organic solvents, (b) the extraction solvent composition, (c) the ratio of solvent volume to tissue weight and (d) additional means of cell-breaking. Data are presented as the mean $\pm$ SD from triplicate experiments with duplicate injections (n=6). Here, M refers to methanol; W refers to water; D refers to dichloromethane. .... 16
- Figure 2.3. (a) Peak pair number detected against the injection amount of  $^{12}\text{C}$ -/ $^{13}\text{C}$ -labeled pooled tissue sample with triplicate injections (n=3). (b) Base-peak ion chromatogram (BPC) of a labeled QC sample. .... 19
- Figure 2.4. (a) PCA scores plot of liver samples. (b) PLS-DA scores plot of liver samples (R<sup>2</sup>=0.999, Q<sup>2</sup>=0.669). (c) PCA scores plot of brain samples. (d) PLS-DA scores plot of brain samples (R<sup>2</sup>=0.999, Q<sup>2</sup>=0.556). ....20
- Figure 2.5. Volcano plots of (a) liver samples and (b) brain samples. Metabolites with fold change (FC) > 1.2 are labeled in blue and metabolites with FC < 0.83 are labeled in red; both with q-value < 0.1. ....21
- Figure 2.6. (a) ROC curve built using 3 metabolite biomarker candidates—uridine, uridine-H<sub>2</sub>O and spermidine—for liver samples. (b) ROC curve built using 4 metabolite biomarker candidates—1,4-diaminobutane, histidine, 4-ethylphenol and 5-hydroxyindoleacetic acid—for brain samples. (c) ROC curve built using 7 metabolite biomarker candidates—1,4-diaminobutane, histidine, 4-ethylphenol, 5-hydroxyindoleacetic acid, 5-hydroxylysine, lipoamide and 4-guanidinobutanoic acid—for brain samples. ....22
- Figure 2.7. Box plots of the relative concentrations of (a) 1,4-diaminobutane, (b) histidine, (c) 4-ethylphenol and (d) 5-hydroxyindoleacetic acid in the WT group and the Tg group. ....24
- Figure 3.1. Workflow of tissue extraction, dansylation isotope labeling LC-MS and lipids analysis. ....27
- Figure 3.2. PCA scores plot of Dex vs. control of (a) brain, (b) heart, (c) kidney, (d) liver, (e) muscle tissue samples. ....31
- Figure 3.3. Volcano plot of (a) brain, (b) heart, (c) kidney, (d) liver, (e) muscle tissue samples. Metabolites with fold change (FC) > 1.2 were labeled in blue, and metabolites with FC < 0.83 were labeled in red, both with q-value < 0.05. ....32

Figure 3.4. Overview of metabolic pathway analysis of brain tissues.....	33
Figure 3.5. Pathway of arginine and proline metabolism of brain tissue. ....	33
Figure 3.6. (a) Box plots of the relative concentrations of (a) 1,4-diaminobutane, (b) spermidine, (c) ornithine, (d) proline.....	34
Figure 4.1. Workflow of the project. ....	38
Figure 4.2. (a) PCA and (b) PLS-DA scores plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 (R2=0.897, Q2=0.773 for PLS-DA plot) for dansylation labeling.....	42
Figure 4.3. Volcano plot of the changes of metabolites between (a) B1C and B1D, (b) B2C and B2D1, (c) B2C and B2D2. Metabolites with fold change (FC) > 1.5 or FC < 0.67 and q-value < 0.05 were highlighted. ....	44
Figure 4.4. ROC curve built using 4 common positively significant metabolites: cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine, for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2. ....	46
Figure 4.5. Box plots of the relative concentrations of (a) o-phosphoethanolamine, (b) cystine, (c) gamma glutamylglutamic acid, (d) glycyl-valine in the RA and control groups in 2 cohorts. ....	48
Figure 4.6. PLS-DA scores plot of (a) B1C vs. N (R2=0.988, Q2=0.913) and (c) B1C vs. Y (R2=0.981, Q2=0.915). The ROC curve built for (b) B1C vs. N using 19 identified significant metabolites and for (d) B1C vs. Y using 14 identified significant metabolites. Here, N refers to RA without MTX treatment and Y refers to RA with MTX treatment. ....	49
Figure 4.7. PLS-DA scores plot of (a) B1C vs. B1DP (positive) vs. B1DN (negative) (R2=0.964, Q2=0.852) and (b) B2D1P (1positive) vs. B2D1N (1negative) vs. B2D2P (2positive) vs. B2D2N (2negative) (R2=0.914, Q2=0.714). ....	50
Figure 4.8. (a) 2D and (b) 3D PLS-DA scores plot of B2D1 r vs. B2D1 nr vs. B2D2 r vs. B2D2 nr (R2=0.92, Q2=0.561). ....	52
Figure 4.9. Volcano plot of the changes of metabolites between (a) B2D1 r and B2D1 nr, (b) B2D2 r and B2D2 nr, (c) B2D1 nr and B2D2 nr, (d) B2D1 r and B2D2 r. Metabolites with fold change (FC) > 1.5 or FC < 0.67 and q-value < 0.05 were highlighted. ....	53
Figure 4.10. (a) PCA and (b) PLS-DA scores plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 (R2=0.773, Q2=0.592 for PLS-DA plot) for DmPA labeling.....	54
Figure 4.11. Volcano plot of the changes of metabolites between (a) B1C and B1D, (b) B2C and B2D1, (c) B2C and B2D2. Metabolites with fold change (FC) > 1.2 or FC < 0.83 and q-value < 0.1 were highlighted. ....	55
Figure 4.12. ROC curve built using 5 common significant metabolites, azelaic acid, acid-9545, acid-23681, acid-23895, acid-23900, for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2.....	57
Figure 4.13. Box plots of the relative concentrations of (a) azelaic acid, (b) acid-9545, (c) acid-23681, (d) acid-23895, (e) acid-23900 in the RA and control groups in 2	

cohorts. ....	58
Figure 4.14. PLS-DA scores plot of (a) B1C vs. N ( $R^2=0.943$ , $Q^2=0.746$ ) and (c) B1C vs. Y ( $R^2=0.951$ , $Q^2=0.801$ ). The ROC curve built for (b) B1C vs. N using 6 identified significant metabolites and for (d) B1C vs. Y using 14 identified significant ..... 59	59
Figure 4.15. PLS-DA scores plot of (a) B1C vs. B1DP (positive) vs. B1DN (negative) ( $R^2=0.881$ , $Q^2=0.633$ ) and (b) B2D1P (1positive) vs. B2D1N (1negative) vs. B2D2P (2positive) vs. B2D2N (2negative) ( $R^2=0.817$ , $Q^2=0.609$ ). .... 60	60
Figure 4.16. (a) 2D and (b) 3D PLS-DA scores plot of B2D1 r vs. B2D1 nr vs. B2D2 r vs. B2D2 nr ( $R^2=0.834$ , $Q^2=0.266$ ). .... 61	61
Figure 5.1. 2D and 3D PCA plots (a, b) and PLS-DA plots (c, d) of female mouse urine samples. .... 68	68
Figure 5.2. Volcano plot of metabolites with a fold change $> 1.5$ are marked in blue, and those with fold change $< 0.67$ are marked in red, both with a q-value $< 0.05$ (corresponding to a p-value $< 0.181$ ). .... 68	68
Figure 5.3. 2D and 3D PCA plots (a, b) and PLS-DA plots (c, d) of male mouse urine samples. .... 70	70
Figure 5.4. Volcano plot of metabolites with a fold change $> 1.5$ are marked in blue, and those with fold change $< 0.67$ are marked in red, both with a q-value $< 0.05$ (corresponding to a p-value $< 0.114$ ). .... 71	71
Figure 5.5. The receiver operating characteristic curve generated based on six common significant metabolites for (a) female and (b) male mouse urine samples. .... 72	72
Figure 5.6. 2D (a, c) and 3D (b,d) PLS-DA plots displaying the metabolic trajectory of eight collection time points among the female (a, b) and male (c, d) Tg mice. Box1 to Box8 represent 8 time points. .... 73	73
Figure 5.7. Box plots of the relative concentrations of methionine in the Tg and WT groups at eight different time points for female (a) and male (b) mice. B1 to B8 represented 8 time points. .... 74	74
Appendix Figure A2.1. (a) Mouse liver sample. (b) Mouse brain sample. .... 93	93
Appendix Figure A2.2. (a) PCA scores plot of liver samples. (b) PCA scores plot of brain samples. Red dots were labeled as the QC samples. Green dots were labeled as 5xFAD transgenic mice, and blue dots were labeled as wild type mice. .... 94	94
Appendix Figure A2.3. (a) 3D PCA scores plot of liver samples. (b) 3D PCA scores plot of brain samples. Red dots were labeled as 5xFAD transgenic mice, and green dots were labeled as wild type mice. .... 95	95
Appendix Figure A2.4. (a) Permutation test result of the ROC curve of Tg vs. WT of liver samples. (b) Permutation test result of the ROC curve of Tg vs. WT of brain samples with 4 metabolite candidates. (c) Permutation test result of the ROC curve of Tg vs. WT of brain samples with 7 metabolite candidates. .... 96	96
Appendix Figure A4.1. PCA plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 vs. QC.	

Two outliers of duplicates of sample 152 were redly circled, which were then excluded.....	97
Appendix Figure A4.2. A series of ROC curves built with different numbers of the top-ranked significant metabolites for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2.....	98
Appendix Figure A4.3. ROC curve built using 7 common significant metabolites, cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine, dansyl-702, dansyl-9056, dansyl-14033, for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2. ....	99
Appendix Figure A4.4. Volcano plot of the changes of metabolites between (a) B1C and MTX N, (b) B1C and MTX Y. Metabolites with fold change (FC) > 1.5 or FC < 0.67 and q-value < 0.05 were highlighted. ....	100
Appendix Figure A4.5. PCA scores plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 vs. QC.....	100
Appendix Figure A4.6. A series of ROC curves built with different numbers of the top-ranked significant metabolites for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2.....	101
Appendix Figure A4.7. Volcano plot of the changes of metabolites between (a) B1C and MTX N, (b) B1C and MTX Y. Metabolites with fold change (FC) > 1.2 or FC < 0.83 and q-value < 0.1 were highlighted. ....	102
Appendix Figure A4.8. (a) Volcano plot of the changes of metabolites between B2D1 nr and B2D2 nr. (b) A series of ROC curves built with different numbers of the top-ranked significant metabolites for B2D1 nr vs. B2D2 nr.....	102
Appendix Figure A5.1. Injection amount optimization curve. ....	103
Appendix Figure A5.2. (a)PCA scores plot of all the urine samples. Green dots represented the QC samples. Red dots represented Tg mouse urine samples, and blue dots represented WT mouse urine samples. (b) PCA scores plot of all the urine samples. Green dots were male mouse urine samples. Red dots were female mouse urine samples. ....	103

## List of Abbreviations

ACN	Acetonitrile
AD	Alzheimer's disease
APP	Amyloid precursor protein
AUC	Area under the curve
BPI	Base peak ion
CCP	Cyclic citrullinated peptide
CE	Capillary electrophoresis
CIL	Chemical isotope labeling
CRP	C-reactive protein
CSF	Cerebrospinal fluid
DAS	Disease activity score
DCM	Dichloromethane
Dex	Dexamethasone
DI	Direct infusion
DmPA	p-Dimethylaminophenacyl
DnsCl	Dansyl chloride
EML	Evidence-based Metabolome Library
ESI	Electrospray ionization
FC	Fold change
FTICR	Fourier transform ion cyclotron resonance
GC	Gas chromatography
HILIC	Hydrophilic interaction liquid chromatography
HMDB	Human metabolome database
HPLC	High performance liquid chromatography
LC	Liquid chromatography
LLE	Liquid-liquid extraction
MeOH	Methanol
MPA	Mobile phase A
MPB	Mobile phase B

MS	Mass spectrometry
MTX	Methotrexate
NMR	Nuclear magnetic resonance spectroscopy
PCA	Principal component analysis
PDA	Photodiode array
PLS-DA	Partial least squares discriminant analysis
ppm	Parts per million
QC	Quality control
QqQ	Triple quadrupole
QTOF	Quadrupole time of flight
RA	Rheumatoid arthritis
ROC	Receiver operating characteristic
RP	Reversed phase
RT	Retention time
SRM	Selected reaction monitoring
SPE	Solid-phase extraction
TEOA	Triethanolamine
Tg	Transgenic
UHPLC	Ultra high-performance liquid chromatography
UV	Ultraviolet
WT	Wild type

## List of Symbols

Da	Dalton
m/z	Mass-to-charge
v/v	Volume to volume ratio

# Chapter 1 Introduction

## 1.1 Multi “omics”

A living system is quite complicated with multiplex characteristics of life and scales of biological organizations. Biological functions can be explored through the molecules involved in physical and biochemical reactions in the body. The comprehensive assessment of a set of molecules can be achieved with various “omics” technologies, which include genomics, transcriptomics, proteomics and metabolomics.<sup>1</sup> Genomics was the first omics to appear and the most mature omics field. It studies the whole genomes and genetic variants caused by disease or medical treatment, through DNA sequencing and bioinformatics.<sup>2</sup> Nonetheless, many parts of genes have not been experimentally characterized and some of the relations between gene assignments and biochemical functions are still not clear, which may also be affected by other factors, such as environment.<sup>3</sup> Transcriptomics is the connection between genomics and proteomics. It focuses on identifying RNA transcripts and measuring and quantifying the expression of genes in different organisms and under different conditions.<sup>4</sup> The limitation of transcriptomics is that some knowledge of transcriptome is still based on gene predictions, and relatively small changes in RNA level may lead to significant protein changes in the organisms.<sup>5</sup> This indirect correlation needs to be further studied. Proteomics explores the entire proteome and analyzes protein’s abundance and interaction, and reflects the underlying genomics and transcriptomics.<sup>6</sup> Post-translational modifications can greatly increase the complexity of proteome studies. Metabolomics is the end point of the omics cascade.<sup>7</sup> It quantitatively studies the entire set of metabolites, including amino acids, carbohydrates, fatty acids and other products of metabolic functions. The four omics can provide comprehensive view of biological processes.

## 1.2 Metabolomics

### 1.2.1 Significances of Metabolomics

Metabolomics is an ultimate reflection of organisms influenced by both genetic and environmental factors. It is sensitive to the small changes in organisms’ processes (e.g. DNA duplication, RNA expression) or diet or environment, which may cause huge metabolite concentration variances.<sup>8</sup> Therefore, it is the most appropriate and closest

description of phenotype. Metabolomics is one powerful tool for global study of composition, dynamics and responses of metabolites, in cells, biofluids, tissues and organs.<sup>9</sup> It can be used for studying the effects of system perturbations on metabolic functions by environmental influences or toxin effects, as well as for diagnosing diseases or monitoring treatment.<sup>10</sup> Thus, metabolomics studies have evolved exponentially since the concept was introduced.

### **1.2.2 Analytical Platforms for Metabolomics**

Normally, the analytical techniques for metabolomics include detection and separation processes. For detection, there are mainly two platforms, nuclear magnetic resonance (NMR) and mass spectrometry (MS). For separation, there are three kinds of chromatography methods, capillary electrophoresis (CE), gas chromatography (GC), high-performance (HP) or ultra-high-performance (UP) liquid chromatography (LC). These high-throughput separation techniques are usually coupled to MS for metabolic profiling. No single technique can provide an ideal analysis for all metabolites. Each technique has its advantages and disadvantages. Next, I will introduce these platforms in more details.

NMR is a spectroscopic technique. NMR is based on energy absorption and re-emission of the atom nuclei affected by the changing of an external magnetic field.<sup>11</sup> With rich natural abundance of hydrogen in biological samples, high resolution <sup>1</sup>H NMR is the most widely used NMR technique. It can rapidly quantify and identify a wide range, from low-molecule weight to high-molecule weight, of metabolites in a single run with little requirement for sample preparation, while simultaneously detect lots of other compounds. NMR is an information-rich method especially for providing structural information, which can be used for characterizing and exploring biological processes.<sup>12</sup> Besides, it is a non-destructive technique, as a result, samples can be recovered for further analysis if necessary. NMR has been widely used for biomarker discovery and disease studies.<sup>10, 13</sup> However, the major limitation of NMR is the relatively low sensitivity. The detection can only be achieved above the micro-molar range, while in real world analysis, many metabolites will be below its detection limit.<sup>14</sup>

MS detects metabolites in the form of spectrum peaks with mass-to-charge ratios ( $m/z$ ). MS is the most widely used platform in metabolomics. It can provide high sensitivity and selectivity, as well as reproducible quantitative analysis and the

possibility for metabolite identification. High-resolution MS can provide accurate mass measurement, allowing for precise identification, and the most widely used MS are Fourier transform ion cyclotron mass spectrometers (FTICR-MS) and time of flight mass spectrometers (TOF-MS).<sup>15</sup> FTICR-MS offers extremely high resolution (100,000–1,000,000), but the slow data acquisition rate leads to low sensitivity. Besides, the high instrumental cost limits its application. In comparison, TOF-MS can provide high resolution as well as high sensitivity with a relatively low cost, thus TOF-MS is more popular for metabolomics. Hybrid instruments, such as quadrupole-TOF-MS, are widely used for rapid screening analysis<sup>16</sup>.

Different kinds of MS-based methods have been developed. Direct injection MS is a rapid technique that can analyze a large number of metabolites.<sup>17</sup> The obvious drawbacks of this method involve co-suppression and low ionization efficiencies.<sup>18</sup> Thus, MS is usually combined with separation techniques to reduce sample complexity, such as CE-MS, GC-MS, LC-MS. CE-MS is an emerging tool for metabolomics studies and has significant potential.<sup>19</sup> CE-MS can provide extremely high-resolution and analyze for almost all charged compounds.<sup>19</sup> However, the repeatability needs to be further improved.<sup>20</sup> GC-MS is a high-throughput technique and generally detects low-molecular-weight, volatile and thermally stable compounds, such as the analysis of breath.<sup>21</sup> The high-molecular-weight, non-volatile metabolites cannot be analyzed directly, and they need multiple chemical derivatization procedures to get volatility and chemical stability.<sup>22</sup> Because GC analysis is usually done under high temperatures, the sample stability is a major concern. LC-MS is the most widely used technology with great sensitivity. The recent UPLC system can further improve chromatographic resolution. Compared with GC-MS, LC-MS doesn't require sample volatility and has lower analysis temperature. Sample derivatization is usually not required, but when necessary, it can be useful to provide better separation and sensitivity.<sup>23</sup> Metabolites can be detected in both positive and negative ion modes, which can achieve high metabolome coverage. It usually uses electrospray ionization (ESI) as the ionization source, which is a soft ionization technique and leads to little fragmentation, but ESI is easily affected by ion suppression.<sup>24</sup> Various column chemistries have been developed. For example, hydrophilic interaction liquid chromatography (HILIC) can be used to separate polar metabolites, while reverse phase liquid chromatography (RPLC) can separate non-polar metabolites. However, one kind of column can be beneficial to

analyze only one kind of polarity of metabolites, which makes detection more complicated.

### **1.2.3 Categories of Metabolomics**

Based on experimental methods, metabolomics can be divided into two major categories: targeted and untargeted. Targeted metabolomics only focuses on a few selected compounds or particular metabolic pathways.<sup>25</sup> Commonly, it is a hypothesis driven approach and explores specific biological pathways. Typically, triple quadrupole (QQQ) mass spectrometry with selected reaction monitoring (SRM) is applied for routine targeted analysis.<sup>26</sup> In this case, only certain  $m/z$  values in mass spectra or certain regions of chromatogram are analyzed.<sup>27</sup> Targeted metabolomics is not a global approach, but can provide high sensitivity, high throughput and capability for absolute quantification. In comparison, untargeted metabolomics is global in scope and cares about the entire chromatogram and all  $m/z$  values.<sup>28</sup> It is usually hypothesis generating instead of hypothesis driven, aiming at detecting as many metabolites as possible.<sup>26</sup> Untargeted metabolomics determines the relative amount of all measurable unknown metabolites and carries out the identification. This technique is significantly attractive and suitable for diagnostic biomarker discovery as well as nonbiased metabolite fingerprinting in response to disease or genetic alterations.<sup>20</sup> Different kinds of biological samples have been studied, such as tissue,<sup>29</sup> serum,<sup>30</sup> urine,<sup>31</sup> cell,<sup>32</sup> with different advantages. Biofluids are usually easy to collect and widely used for different analysis, which are considered as a pool of metabolites of the organisms and can reflect systemic metabolic changes. In comparison, tissue samples can reveal organ-specific metabolic fingerprints,<sup>33</sup> and tissue metabolomics plays an important role in investigating specific diseases<sup>34</sup> and sites of toxicity<sup>35</sup>. Although untargeted metabolomics is promising and significantly developed in last several decades, no one single method can achieve the non-targeted analysis of all the metabolites, due to their diverse polarities, molecular weights and concentrations.

### **1.2.4 Workflow for MS-based Untargeted Metabolomics**

Experimental design and sample preparation play important roles in MS-based untargeted metabolomics study. Generally, the experiment steps include sampling, sample preparation, sample analysis, instrument analysis, metabolite identification and statistical analysis.<sup>15</sup> The first step is sampling, which primarily depends on the

experiment type and experimental design. Sampling is the basis of the project and many aspects should be considered. For example, how much sample would be sufficient, what kind of samples could be analyzed (e.g. serum, plasma, tissue, cell, saliva, etc.), what would be the effects of gender, age and diet, and what would need to be done to inhibit enzymatic activity and stop metabolism after sample collection (e.g. freezing or acidic treatments). Otherwise, biological variability and metabolite degradation would be major issues. The second step is sample preparation, which mainly extracts metabolites from complex matrix and removes interfering components (e.g. protein precipitation with organic solvent). When necessary, extraction methods, such as liquid-liquid extraction (LLE) and solid-phase extraction (SPE) can pre-concentrate certain metabolites. After sample preparation, the third step is instrumental analysis and data acquisition, in which samples are analyzed by direct injection MS, CE-MS, GC-MS or LC-MS, with LC-MS as the most popular one. In this process, metabolites are detected as peaks with  $m/z$  values in mass spectra and the intensities of chromatographic peaks are related to their concentrations. Then, metabolic data in mass spectra are exported to a standard and uniform format for further analysis. The data exportation includes peak picking, noise and background exclusion, alignment of chromatograms and mass spectra. The most widely used software is XCMS, a web-based software.<sup>36</sup> Here, each peak is referred to as a metabolite feature and extracted information is displayed in a peak table with information of each feature's retention time,  $m/z$ , intensity, as well as  $p$ -values and fold changes representing relative concentration differences. The next step is the identification of metabolites, which still remains a big challenge. In this step, accurate mass of metabolites is firstly searched in metabolite database, such as METLIN<sup>37</sup> and Human Metabolome Database (HMDB)<sup>38</sup> to get putative identification information. Then, further experimental data are used, including retention time and MS/MS, to get fragment patterns, before the data are compared with standards to definitely identify those metabolites of interest. Currently, large number of metabolites still cannot be matched and comprehensive metabolite identification is impractical. Finally, statistical tools are applied for sample classification and significant markers determination, such as principal component analysis (PCA), partial least square discriminant analysis (PLS-DA), analysis of variance (ANOVA) and volcano plot.

## 1.2.5 Chemical Isotope Labeling Metabolomics

As the widely used untargeted metabolomics technique, RPLC-MS has been applied for the studies of various kinds of samples, such as serum<sup>39</sup>, urine<sup>40</sup>, tissue<sup>41</sup> and cells<sup>42</sup> etc. However, due to the complexity of metabolites (various chemical and physical properties), conventional RPLC-MS can only handle medium polar and non-polar compounds, RPLC is not appropriate for extremely polar metabolites. Ionization also needs to be performed in both positive and negative mode to increase metabolome coverage, making the experiments very complicated.

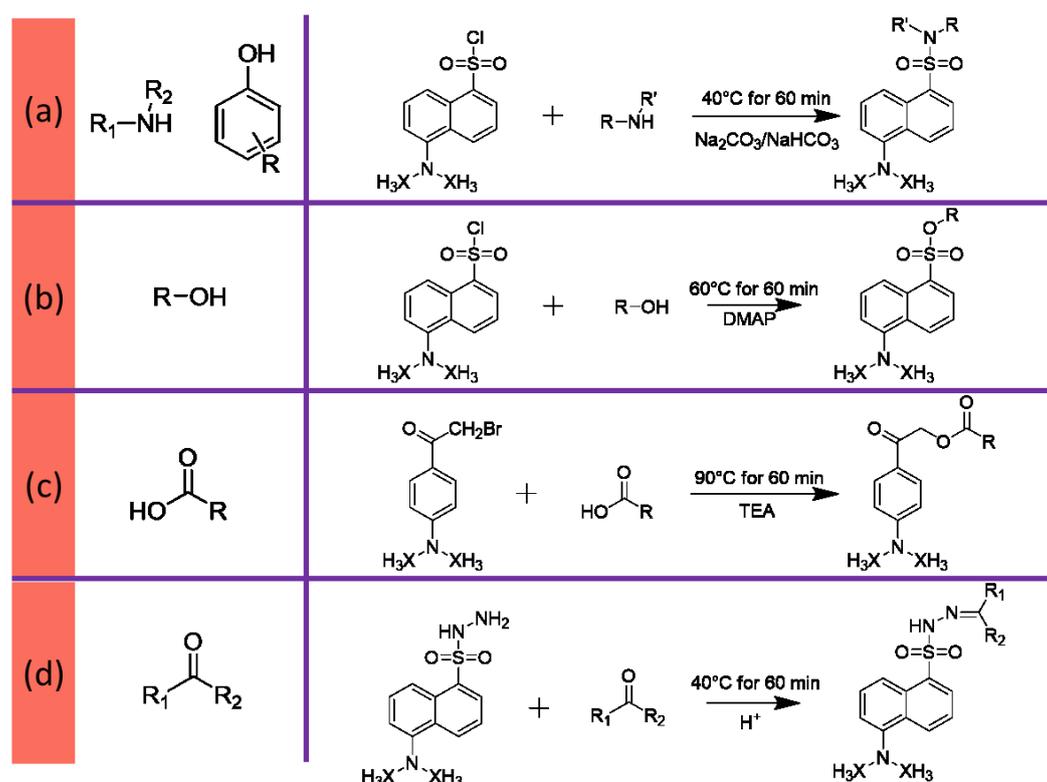


Figure 1.1 The scheme of four CIL reactions, including: (a) Dansyl chloride for amine/phenol submetabolome; (b) Dansyl chloride with base-activation for hydroxyl submetabolome; (c) p-Dimethylaminophenacyl (DmPA) bromide for carboxyl submetabolome; (d) Dansylhydrazine (DnsHz) for carbonyl submetabolome.

Previously, our lab developed a “divide and conquer” technology, in which the whole metabolome is divided into different submetabolomes based on chemical functional groups (Figure 1.1). Chemical isotope labeling (CIL) is used to analyze each submetabolome with high coverage and the combined results are used for the complete metabolome analysis. Dansyl chloride (DnsCl) for amine/phenol submetabolome<sup>23</sup>, is used here for the discussion of CIL LC-MS process and its benefits. In this approach, individual experimental samples are labeled with <sup>12</sup>C-DnsCl, while a pooled sample,

working as the reference and internal standard, is labeled with  $^{13}\text{C}$ -DnsCl. After mixing together and LC-MS analysis, metabolites are detected as peak pairs, which contain light peaks ( $^{12}\text{C}$  labeled) and heavy peaks ( $^{13}\text{C}$  labeled) with an  $m/z$  difference of 2.0067 Da. The intensity ratio of light and heavy peaks is used for relative quantification of metabolites. Dansyl labeling method has many advantages. Firstly, the aromatic group makes the metabolite more hydrophobic which increases retention in RPLC. Secondly, the tertiary amine enhances ESI chargeability, thus enhancing MS signal. Thirdly, the two carbons on the tag are either  $^{12}\text{C}$  or  $^{13}\text{C}$ , which are used for relative quantification. Dansyl labeling method simultaneously improves separation, detection and quantification, in which case, only positive ion mode RPLC is required. In the last decade, we have also developed  $^{12}\text{C}$ -/ $^{13}\text{C}$  DnsCl with base-activation for the hydroxyl submetabolome<sup>43</sup>,  $^{12}\text{C}$ -/ $^{13}\text{C}$ -dimethylaminophenacyl (DmPA) bromide labeling for the carboxyl submetabolome<sup>44</sup> and  $^{12}\text{C}$ -/ $^{13}\text{C}$  dansylhydrazine labeling for the carbonyl submetabolome<sup>45</sup>. These four submetabolomes cover more than 95% of the whole metabolome. Thus this comprehensive technique is applicable and promising for untargeted metabolomics.

### 1.3 Scope of the Thesis

The objective of this research work is to develop CIL LC-MS methods for untargeted metabolomics.

In Chapter 2, CIL LC-MS is used to profile the amine and phenol submetabolome of mouse brain and liver tissues of Alzheimer's disease (AD). A tissue extraction protocol was developed for tissue metabolomics and this protocol was applied to discover biomarkers for AD.

In Chapter 3, CIL LC-MS is used to profile the amine and phenol submetabolome of rat liver, heart, kidney, muscle and brain tissues. The metabolic differences between Dex-treatment and control were compared and the affected pathways were analyzed.

In Chapter 4, CIL LC-MS is used to profile the amine and phenol submetabolome of rheumatoid arthritis disease and discover diagnostic biomarkers. Several biomarkers were discovered with good discriminating power.

In Chapter 5, CIL LC-MS is used to profile the amine and phenol submetabolome of mouse urine samples of AD. Metabolic trajectory changes with the development of the disease were observed and several metabolite biomarkers were discovered.

## **Chapter 2 Development of Chemical Isotope Labeling LC-MS for Tissue Metabolomics and Its Application for Brain and Liver Metabolome Profiling in Alzheimer's Disease Mouse Model**

### **2.1 Introduction**

Metabolomics involves the comprehensive analysis of all the small molecule metabolites in a biological system. It is a rapidly emerging field for characterizing complex biochemical phenotypes influenced by environmental, genetic and other factors. Metabolomic profiling is carried out using different kinds of samples, including biofluids, cells and tissues, depending on the areas of applications.<sup>46-47</sup> Of those, tissue is at an organizational level between cell and organ that is composed of cells and extracellular matrix. Compared with biofluids containing a pool of metabolites of the body that reflect systemic metabolic changes, tissue samples can reveal organ-specific metabolic fingerprints<sup>48-49</sup>. Hence, tissue metabolomics plays a crucial role in many biological and clinical applications, such as investigating organ-specific diseases or functions<sup>50-51</sup> and sites of toxicity<sup>35</sup>.

Many tissue metabolomics studies have been reported using various analytical platforms with varying degrees of metabolome coverage, including nuclear magnetic resonance (NMR), gas chromatography-mass spectrometry (GC-MS), capillary electrophoresis-mass spectrometry (CE-MS) and liquid chromatography-mass spectrometry (LC-MS)<sup>48</sup>. Among them, NMR analyses of tissues detected < 100 metabolites (e.g., brain<sup>14</sup> and lymph node<sup>52</sup>). GC-MS has been used to analyze tissues with the detectability of a few hundreds of metabolites (e.g., the detection of 200 metabolites in colon tissues<sup>53</sup>). CE-MS measured more than 800 peaks in colon tissues and around 1000 peaks in stomach tissues<sup>54</sup>. Importantly, LC-MS has been more widely used in tissue metabolomics<sup>55-60</sup>. LC-MS can detect up to 6000 metabolite features from rat brain tissue extract<sup>61</sup>. A metabolite feature is defined as a unique m/z peak at a given chromatographic retention time.

Although traditional LC-MS analysis may detect a large number of metabolite features, it has long been a problem to differentiate the weak signal of very low-abundant metabolites from the background noise<sup>62-63</sup>. Also, a single metabolite may be detected in multiple forms, including adduct ions, in-source fragment ions, dimers,

trimers, giving a misleading number of detected metabolites and complicating the result interpretation. A recent study of yeast cell extracts showed that a vast majority of over 25,000 metabolite features detected in high-resolution LC-MS experiments could be filtered down to less than 1000 peaks that were likely to be from unique metabolites <sup>64</sup>. Another important issue is related to metabolite quantification using LC-MS. Without internal standards, conventional LC-MS suffers from ion suppression effects in MS ionization, as sample matrix or coeluting compounds can significantly interfere with the ionization of metabolites and thereby distort the quantification results.

It is clear that new technologies are needed to provide higher metabolome coverage and better quantitative capability for tissue metabolomics. Combining chemical isotope labeling (CIL) with LC-MS has recently shown to offer much enhanced analytical performance for targeted or non-targeted metabolite analysis <sup>65-76</sup>. Previously, we reported a high-performance CIL LC-MS technique for quantitative metabolomics <sup>23</sup>. This chemical derivatization method can significantly enhance the electrospray ionization signal and improve the RPLC separation. For example, using <sup>12</sup>C- and <sup>13</sup>C-dansyl labeling LC-MS, we can achieve high-coverage metabolic profiling of the amine/phenol-submetabolome with high quantification accuracy <sup>23</sup>. This technique has been successfully applied to metabolomics studies of various types of samples <sup>77-82</sup>. However, there is no report of a simple and robust analytical workflow that utilizes CIL LC-MS for studying tissue samples.

In this work, we report a workflow of a tissue extraction protocol tailored to CIL LC-MS for tissue metabolomics. We first developed and optimized the tissue extraction method using chicken liver. We then applied this workflow to profile the amine/phenol-submetabolome of mouse liver and brain tissue samples of an Alzheimer's disease (AD) model to validate our method. We chose AD mouse model to demonstrate the analytical performance of our method, as it represents a typical biological study of animal models where various types of tissue samples can be readily acquired for metabolomic profiling. AD is the most common cause of dementia. Early diagnosis remains a challenge. Transgenic mouse models are widely used to study AD pathophysiology. Since brain tissue can directly reflect brain-specific metabolic changes, we examined the metabolomic changes in AD mouse brain tissues that may reveal potential diagnostic biomarkers. We also examined the liver tissues to showcase

the applicability of our method for different types of tissue samples.

## **2.2 Experimental**

### **2.2.1 Chemicals and Reagents**

All the chemicals and reagents, unless otherwise stated, were from Sigma-Aldrich Canada (Markham, ON, Canada).  $^{13}\text{C}$ -dansyl chloride was synthesized in our lab with the procedures published previously <sup>23</sup> and is available from MCID.chem.ualberta.ca. LC-MS grade water, acetonitrile (ACN), and methanol (MeOH) were from Thermo Fisher Scientific (Edmonton, AB, Canada).

### **2.2.2 Chicken Liver Sample Collection**

Three batches of a single lobe of chicken liver were from a local grocery store in separate days for method development. The liver sample was cut into small pieces in the range of 100-200 mg with a clean surgery knife. Then, each tissue piece was separately transferred into a pre-weighed 2-mL Eppendorf vial and stored in a -80 °C freezer until further analysis.

### **2.2.3 Mouse Model and Sample Collection**

A widely-used transgenic mouse model, 5xFAD, from the Jackson Laboratory <sup>83</sup>, was studied. The transgenic mice co-express five familial Alzheimer's disease (FAD) mutations, K670N/M671L (Swedish) + I716V (Florida) + V717I (London) in the Amyloid Precursor Protein (APP) gene and M146L + L286V mutations in the Presenilin 1 (PS1) gene. Approximately at the age of 1.5 months, the 5xFAD mice started generating high levels of A $\beta$ 42, followed by amyloid deposition at about the age of 2 months. The liver and brain tissue samples were taken from 11 mice, five 5xFAD Tg-positive and six wild-type (WT) littermate controls, after the mice were euthanized at the age of 5 months. Liver samples (Appendix Figure A1.1a) were washed with 1x phosphate-buffered saline (PBS) for 3 times before each lobe was cut into 2 pieces, which were then snap-frozen in liquid nitrogen and stored in a -80 °C freezer. Finally, only one of the 2 pieces of Lobe 1 (Appendix Figure A1.1a) of each liver sample was used for the metabolomic analysis. For each brain sample (Appendix Figure A1.1b), half of the brain was snap-frozen in liquid nitrogen and stored in a -80 °C freezer till extraction, and the remaining half was fixed in formalin and archived as paraffin-embedded tissue blocks.

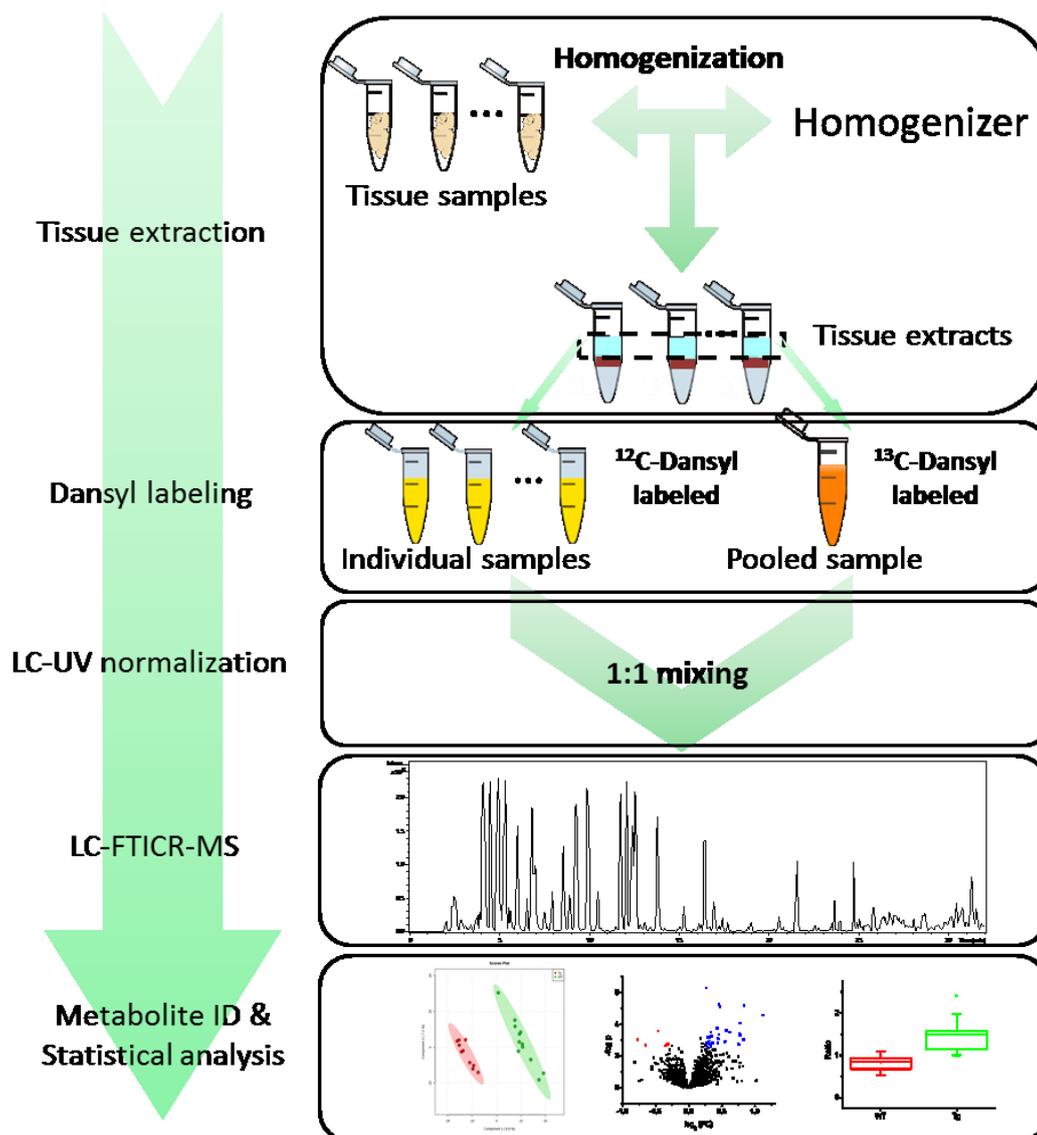


Figure 2.1. Workflow for tissue extraction and dansylation isotope labeling LC-MS.

## 2.2.4 Analytical Workflow

In our work, a differential CIL method was applied to achieve the relative metabolite quantification. The individual samples are labeled by  $^{12}\text{C}$ -dansyl chloride (light labeling), while a pooled sample, prepared by mixing small aliquots of individual samples, is labeled by  $^{13}\text{C}$ -dansyl chloride (heavy labeling). The same amount of  $^{13}\text{C}$ -labeled pooled sample is spiked to each individual sample to serve as a global internal standard. Figure 2.1 displays the overall workflow of this work. It contains the following steps: (1) tissue disruption and metabolite extraction, (2) dansylation labeling, (3) LC-UV quantification of total amount of labeled metabolites in each individual sample and pooled sample for sample amount normalization, (4) mixing of  $^{12}\text{C}$ -labeled individual samples and  $^{13}\text{C}$ -labeled pooled sample at equal amounts, (5) LC-FTICR-

MS analysis of  $^{12}\text{C}$ -/ $^{13}\text{C}$ - mixtures, (6) data processing using R programs, including peak pair picking, peak pair filtering, peak pair ratio calculation and peak pair grouping, (7) multi-variate statistical analysis using MetaboAnalyst <sup>84</sup> software, (8) metabolite identification based on the dansyl standard library <sup>85</sup>. The detailed experimental conditions in the workflow are described below.

### **2.2.5 Tissue Extraction**

Ice-cold solvents (methanol, dichloromethane (DCM) and water) were prepared in advance. After weighing out the tissue sample in an Eppendorf tube, we added 3 mL of methanol per gram of tissue and 0.64 mL of water per gram of tissue. The Eppendorf tube containing the sample and solvent was put into an ice-bath, and then the tissue sample was homogenized twice (10 s each time with a waiting period in between for temperature control), using a Bio-Gen PRO200 Homogenizer (PRO Scientific, USA). After the addition of DCM (3 mL/g tissue) and water (another 1.5 mL/g tissue), the mixture was vortexed twice (30 s each time), staying in the ice-bath in between. The mixture was then incubated in a -20 °C freezer for 15 min. Finally, the sample was centrifuged at 15000 g for 10 min at 4 °C to generate an upper aqueous layer (with polar metabolites) and a lower organic layer (with lipophilic compounds), with proteins and cellular debris precipitated on the bottom. The upper and lower layers were transferred into separate vials.

Three aliquots were taken from each of the upper layer samples: two 15  $\mu\text{L}$  aliquots for the amine/phenol-labeling LC-MS analysis, and one 30  $\mu\text{L}$  aliquot for making the pooled sample. All of the leftovers were stored at -80 °C for future analyses.

### **2.2.6 Dansylation Labeling**

The 15  $\mu\text{L}$  upper layer sample was thawed in a 4 °C freezer, vortexed to dissolve precipitates, and then the extraction solvent was removed using a Savant SC110A Speed Vac at room temperature. After that, the sample was re-dissolved to 25  $\mu\text{L}$  with water for dansylation labeling with the labeling protocol adapted from a previous report <sup>23</sup>. 25  $\mu\text{L}$  of individual tissue sample (experimental duplicates) or 25  $\mu\text{L}$  of the pooled sample was mixed with 12.5  $\mu\text{L}$  of ACN. Then 12.5  $\mu\text{L}$  of 250 mM sodium carbonate/sodium bicarbonate buffer was added to the samples. The solution was mixed with 25  $\mu\text{L}$  of freshly prepared  $^{12}\text{C}$ -dansyl chloride (DnsCl) solution (18 mg/mL) (for light labeling, individual samples) or  $^{13}\text{C}$ -DnsCl solution (18 mg/mL) (for heavy

labeling, pooled sample). After incubation for 45 min at 40 °C, 5 µL of 250 mM sodium hydroxide solution was added to the reaction mixture to quench the reaction. The solution was then incubated at 40 °C for another 10 min. Finally, 25 µL of formic acid (425 mM) in 1:1 ACN/H<sub>2</sub>O was added to make the solution acidic.

### 2.2.7 LC-UV Sample Normalization

The total amount of dansyl-labeled metabolites in each sample was measured using an LC-UV based protocol reported previously<sup>86</sup>. The instrument for detection was a Waters ACQUITY UPLC system with a photodiode array (PDA) detector. A Phenomenex Kinetex reversed-phase C18 column (50 mm × 2.1 mm, 1.7 µm particle size, 100 Å pore size) was used with a fast step-gradient. Mobile phase A was 0.1% (v/v) formic acid in 5% (v/v) ACN/water, and mobile phase B was 0.1% (v/v) formic acid in ACN. Started at 0% B for 1 min, the gradient was increased to 95% B within 0.01 min and held until 2.5 min to completely elute all labeled metabolites. Finally, the gradient was restored back to 0% B in 0.5 min and held for another 3 min. The flow rate was 0.45 mL/min and the total run time was 6 min. The peak area, which represents the total concentration of dansyl-labeled metabolites, was integrated using the Empower software (6.00.2154.003). According to the quantification results, <sup>12</sup>C-labeled individual sample and the <sup>13</sup>C-labeled pooled sample were mixed in equal mole amounts. The <sup>12</sup>C- and <sup>13</sup>C-labeled pooled samples were mixed in equal mole amounts to serve as a quality control (QC) sample.

### 2.2.8 LC-MS

Labeled tissue extracts were analyzed using a Bruker 9.4 T Apex-Qe FTICR mass spectrometer (Bruker, Billerica, MA), coupled with an Agilent capillary 1100 binary system (Agilent, Palo Alto, CA). An Agilent eclipse plus reversed-phase C18 column (100 × 2.1 mm, 1.8 µm particle size,) was used. LC mobile phase A was 0.1% (v/v) formic acid in 5% (v/v) ACN/water, and mobile phase B was 0.1% (v/v) formic acid in ACN. The gradient elution profile was: t = 0 min, 20% B; t = 3.5 min, 35% B; t = 18 min, 65% B; t = 24 min, 99% B, t = 32 min, 98% B. Flow rate was 180 µL/min. All mass spectra were collected in the positive ion mode. The MS conditions for FTICR-MS were: nitrogen nebulizer gas, 2.3 L/min; dry gas flow, 7.0 L/min; dry temperature, 195 °C; capillary voltage, 4200 V; spray shield, 3700 V; acquisition size, 256 k; mass scan range, m/z 200–1000; ion accumulation time, 1 s; TOF (AQS), 0.007

s; DC extract bias, 0.7 V. All the samples were injected in random order. QC samples and amino acid standards were injected every 10 sample runs to monitor the performance of the LC-MS.

### **2.2.9 Data Processing, Metabolite Identification and Statistical Analysis**

After LC-FTICR-MS analysis, the entire list of centroid peaks with collected information (e.g., retention times, m/z values, and peak intensities) was exported from Bruker Data Analysis software (Version 4.0). IsoMS<sup>87</sup> was used to pick peak pairs, to filter false-positive pairs including dimers and common adducts, and to calculate peak-pair intensity ratios. After the alignment of peak pairs from multiple samples using the Alignment program, the Zerofill program was applied to recover the high-confidence peak pair ratios lost during the previous data processing steps.

Based on the accurate mass and retention time, positive metabolite identification was performed using dansyl standard library search, which contains 273 unique dansylated amines/phenols. Based on the accurate mass search, putative identification was performed using the Human Metabolome Database (HMDB) ([www.hmdb.ca](http://www.hmdb.ca)) and the Evidence-based Metabolome Library (EML) (<http://www.mycompoundid.org/>).

Multivariate analyses, including principal components analysis (PCA), partial least squares discriminant analysis (PLS-DA) and receiver operating characteristic (ROC) curve analysis, were carried out using MetaboAnalyst software. The fold change and p-value between groups were calculated using Microsoft Excel. Volcano plots were constructed using OriginPro 8.0 (OriginLab). The q-value, multiple-testing-corrected p-value, was calculated using R and BioConductor ([www.bioconductor.org](http://www.bioconductor.org)).

## **2.3 Results and Discussion**

### **2.3.1 Tissue Metabolite Extraction**

Tissue samples derived from organs (e.g., liver, brain and muscle) contain an extracellular matrix as well as a variety of resident cell types and hence require disruption by homogenization before downstream processing. Various methods have been reported to break tissues and cells, including grinding with mortar and pestle and homogenization with bead beater or electric homogenizer. Of those, grinding tissues in liquid nitrogen by mortar and pestle is not applicable to high-throughput methods because of serial processing of samples and laborious manual intervention.<sup>88</sup> In

comparison, homogenization using a homogenizer is less labor-intensive and more efficient.<sup>33</sup> For the sample extraction, because we wanted to study polar metabolites and non-polar lipids (future work) simultaneously, we developed a protocol based on a widely-used method that can efficiently extract the two groups of compounds into different solvent phases.<sup>89</sup>

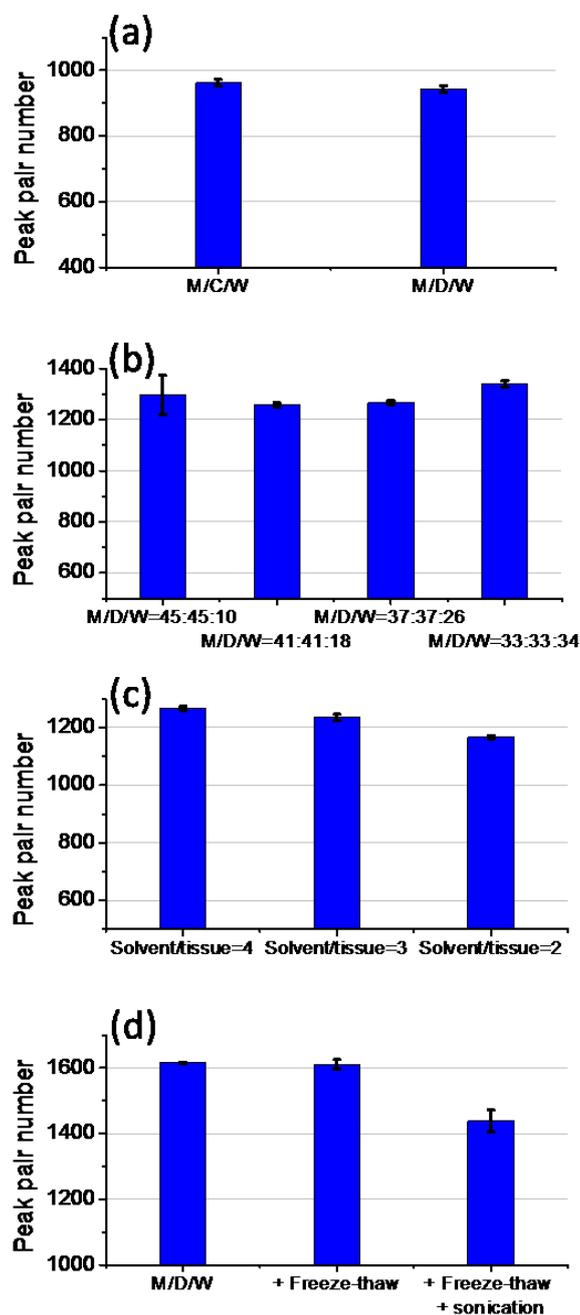


Figure 2.2. Effects of different tissue extraction conditions on peak pair detection of  $^{12}\text{C}/^{13}\text{C}$ -dansyl labeled chicken liver samples: (a) two organic solvents, (b) the extraction solvent composition, (c) the ratio of solvent volume to tissue weight and (d) additional means of cell-breaking. Data are presented as the mean $\pm$ SD from triplicate experiments with duplicate injections ( $n=6$ ). Here, M refers to methanol; W refers to water; D refers to dichloromethane.

For the development of this protocol tailored to CIL LC-MS, we evaluated the extraction efficiency of two organic solvents, the extraction solvent composition, the scale of the extraction solvent (i.e., the ratio of solvent volume to tissue weight) and mechanical tissue disruption methods. Because the amount of transgenic mouse tissues for AD studies was limited, three batches of chicken livers were used for the initial method development. To examine the effects of the extraction conditions on metabolite extraction efficiency for chicken liver tissues,  $^{12}\text{C}$ -/ $^{13}\text{C}$ -dansyl labeled upper aqueous layer samples were prepared and analyzed by LC-MS. The differences in total metabolite concentrations of the tissue extracts were normalized by injecting the same total mole amount according to the LC-UV measurement. We focused on finding a protocol that can extract the largest number of metabolites. The peak pair number obtained was thus used as the indicator of extraction efficiency of the corresponding extraction condition.

Firstly, we tried to replace chloroform used in the reported method<sup>89</sup> with DCM, as chloroform has been classified as a probable human carcinogen.<sup>90</sup> DCM is considered less hazardous yet reported with similar lipid extraction efficiency.<sup>90</sup> Figure 2.2a compares the extraction efficiency of two extraction solvents: methanol/chloroform/water (M/C/W) and methanol/DCM/water (M/D/W). The number of peak pairs detected using DCM ( $943\pm 10$ ;  $n=6$ ) was only 2.1% less than that using chloroform ( $963\pm 10$ ;  $n=6$ ), indicating comparable extraction efficiencies. Hence, DCM was used for subsequent studies.

Next, we tested the extraction efficiency of various extraction solvent compositions, mainly focusing on the effect of different water amounts, as water phase contains the metabolites of interest for CIL LC-MS. We used a different batch of chicken liver for this study. Figure 2.2b shows the results from four different solvent compositions where the methanol and DCM amounts were fixed, but the water amount was different. The four water amounts, 10%, 18%, 26% and 33%, gave similar peak pair numbers (i.e.,  $1297\pm 77$ ,  $1258\pm 8$ ,  $1267\pm 7$ , and  $1340\pm 11$ , respectively). The peak pair number from 33% of water was slightly higher than the others. However, when the water amount was too high, it might disturb the distribution of lipophilic compounds between the upper layer and lower layer supernatants. Therefore, 26% of water was deemed to be the most appropriate amount. Note that the number of peak pairs detected

in Figure 2.2b was higher than that in Figure 2.2a for the same solvent system. This difference could be attributed to the use of different livers (i.e., two different batches).

We also studied the effect of the ratio of extraction solvent vs. tissue weight (mL/g-tissue) on peak pair detection. With the extraction solvent composition fixed, we tried to decrease the volume of extraction solvent per gram of tissue so that the metabolite concentrations could be increase, which is beneficial for increasing chemical labeling efficiency. Figure 2.2c shows best condition should be the one with solvent/tissue=3, giving a peak pair number of  $1235\pm 11$ . This extraction solvent gave comparable metabolite extraction efficiency to that of using a mixture of M/D/W with solvent/tissue=4, and was chosen in the final protocol for tissue extraction.

We carried out a 3<sup>rd</sup> set of experiments to examine the effects of freeze-thaw and sonication for breaking cells and tissues on the number of peak pairs detected. In this case, another batch of chicken liver was used for the study. Figure 2.2d shows the peak pair numbers detected from three conditions: M/D/W alone, M/D/W with freeze-thaw cycles, M/D/W with freeze-thaw plus sonication (i.e.,  $1616\pm 2$ ,  $1611\pm 14$ , and  $1438\pm 34$ , respectively). Adding freeze-thaw did not yield more peak pairs. Sonication actually reduced the peak pair number, which might be due to the degradation or loss of metabolites during sonication.

Taken together, the optimal protocol for tissue sample preparation for CIL LC-MS is 3 mL/g-tissue of methanol and 0.64 mL/g-tissue of water before homogenization, and 3 mL/g-tissue of DCM and 1.5 mL/g-tissue of water after homogenization, followed by incubation and centrifugation.

### **2.3.2. Mouse Liver and Brain Tissue Metabolome Analysis**

After method optimization for sample preparation, the CIL LC-MS method was used to profile the amine/phenol-submetabolome of mouse brain and liver tissue samples. Eleven brain samples and eleven liver samples were collected from 5 Tg type mice and 6 non-Tg (WT) mice. In total, with duplicate  $^{12}\text{C}$ -labeling experiments, 22  $^{12}\text{C}$ -/ $^{13}\text{C}$ -mixtures per tissue type were analyzed. Since the injection amount of labeled metabolites into LC-MS can affect the number of peak pairs detected, we determined the optimal injection amount using the mixture of  $^{12}\text{C}$ - and  $^{13}\text{C}$ -labeled pooled sample. Figure 2.3a shows the number of peak pairs or metabolites measured against the injection amount. The number of peak pairs reached the maximum with an injection

amount of 6 nmol of labeled metabolites. As a result, all individual liver, brain and QC samples were analyzed using LC-MS with 6 nmol injection in each run. Figure 2.3b shows a typical base-peak ion chromatogram (BPC) of the dansyl-labeled QC sample. Many peaks were detected across the entire RPLC elution window, indicating the complexity of the amine/phenol submetabolome of the tissue samples.

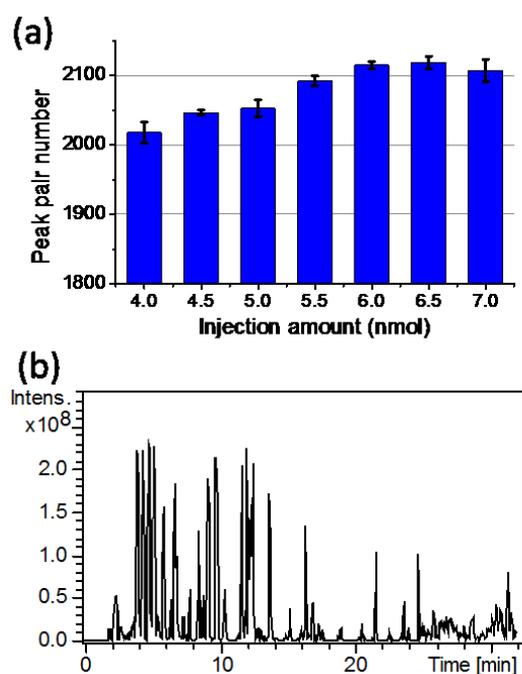


Figure 2.3. (a) Peak pair number detected against the injection amount of  $^{12}\text{C}$ -/ $^{13}\text{C}$ -labeled pooled tissue sample with triplicate injections ( $n=3$ ). (b) Base-peak ion chromatogram (BPC) of a labeled QC sample.

Our CIL LC-MS analysis achieved high submetabolome coverage for both mouse liver and brain samples. For liver samples, a total of 2319 peak pairs were commonly detected in more than 80% of the 22 runs. Among them, 89 were positively identified based on a dansyl standard library search (Supplemental Table S2.1A). For the remaining peak pairs, using accurate mass search with a mass accuracy tolerance of 10 ppm, 166 pairs were putatively identified by the HMDB library (Supplemental Table S2.1B), and 897 pairs were putatively identified by the EML library with one reaction (Supplemental Table S2.1C). Thus, among the 2319 peak pairs detected, we identified positively or putatively 1152 pairs (50%). Similarly, for brain samples, 1769 peak pairs or metabolites were commonly detected in more than 80% of the 22 samples. 78 metabolites were positively identified (Supplemental Table S2.2A). 113 peak pairs were putatively identified based on mass-match to the metabolite entries in the HMDB

library (Supplemental Table S2.2B), and 640 pairs were putatively identified based on the EML library search with one reaction (Supplemental Table S2.2C). Thus, out of the 1769 detected peak pairs, 831 (47%) could be positively or putatively identified.

The above results indicate that the tissue processing protocol combined with CIL LC-MS can be used to analyze a large number of amine-/phenol-containing metabolites. For future work, it is also possible to add other labeling methods<sup>43-45</sup> to increase the overall metabolome coverage.

### 2.3.3 Submetabolome Comparison of Tg and WT mice

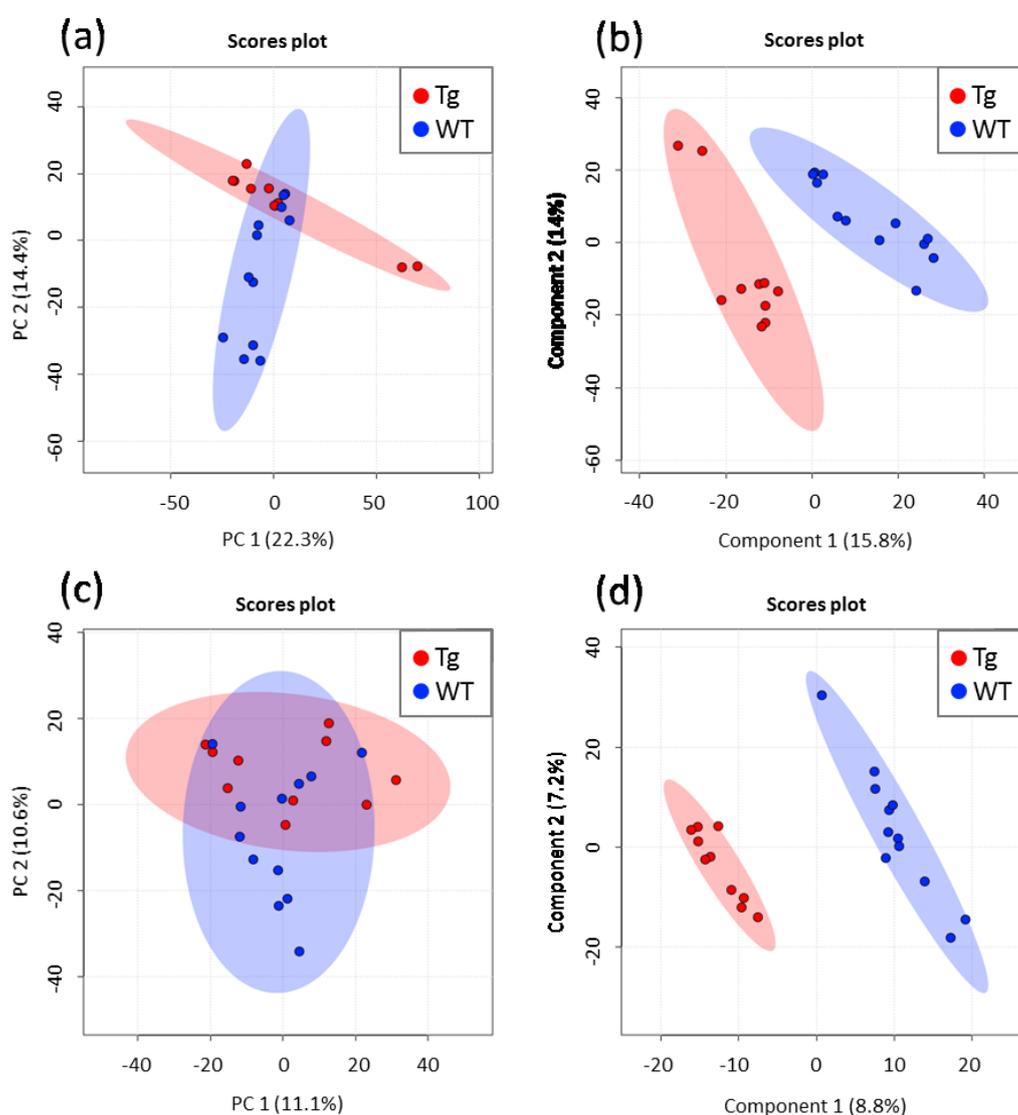


Figure 2.4. (a) PCA scores plot of liver samples. (b) PLS-DA scores plot of liver samples ( $R^2=0.999$ ,  $Q^2=0.669$ ). (c) PCA scores plot of brain samples. (d) PLS-DA scores plot of brain samples ( $R^2=0.999$ ,  $Q^2=0.556$ ).

Multivariate analyses, PCA and PLS-DA, were applied to study the

metabolomic differences between 5xFAD transgenic (Tg) mice and wild-type (WT) mice. From the PCA plots shown in Appendix Figure A2.2, it is clear that, for both liver and brain samples, the QC samples cluster together closely, indicating that the instrument performance was good and stable during the data collection period. Figure 2.4a and 2.4c show the PCA plots without QC. Some separation between Tg and WT groups for both liver and brain samples could be seen, indicating that there were metabolomic differences between the two groups. The separation could be more clearly seen in the 3D PCA plots (Appendix Figure A2.3). Using PLS-DA, the Tg and WT groups were well separated as shown in Figure 2.4c and 2.4d. The PLS-DA models were validated by cross-validation ( $R^2 = 0.999$  and  $Q^2 = 0.669$  for liver samples, and  $R^2 = 0.999$  and  $Q^2 = 0.556$  for brain samples).

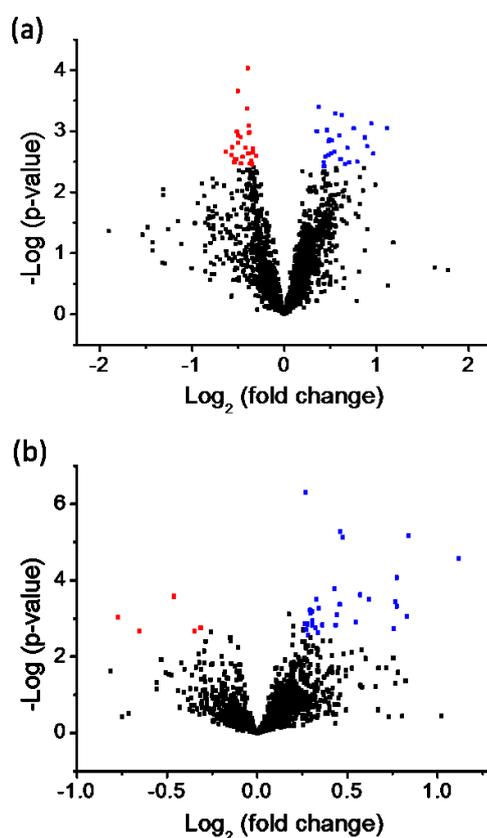


Figure 2.5. Volcano plots of (a) liver samples and (b) brain samples. Metabolites with fold change (FC) > 1.2 are labeled in blue and metabolites with FC < 0.83 are labeled in red; both with q-value < 0.1.

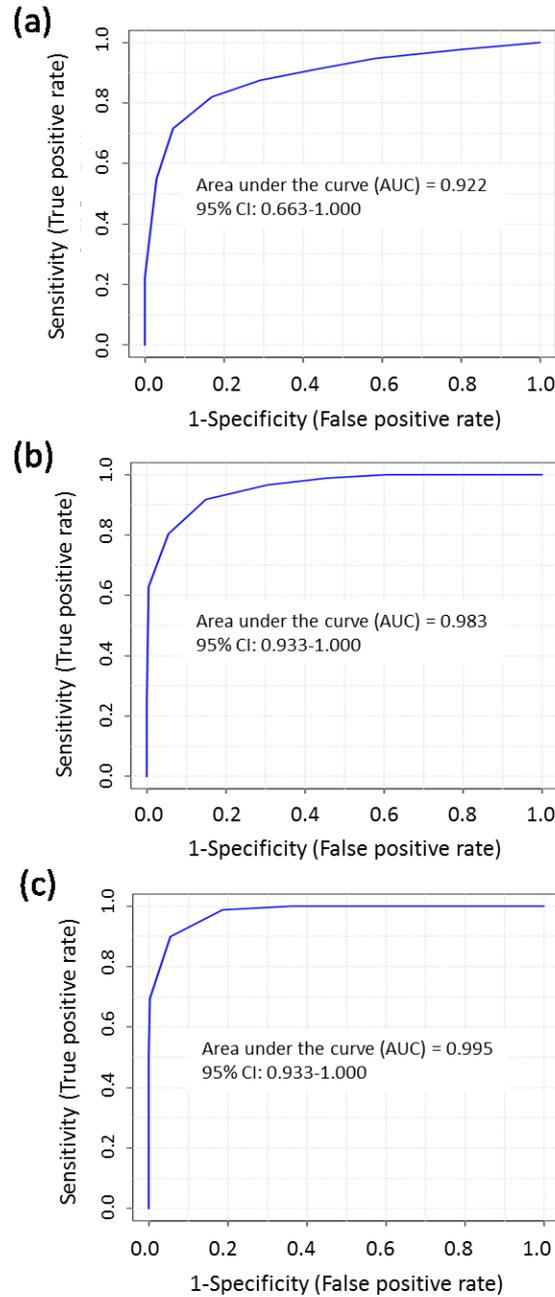


Figure 2.6. (a) ROC curve built using 3 metabolite biomarker candidates—uridine, uridine-H<sub>2</sub>O and spermidine—for liver samples. (b) ROC curve built using 4 metabolite biomarker candidates — 1,4-diaminobutane, histidine, 4-ethylphenol and 5-hydroxyindoleacetic acid—for brain samples. (c) ROC curve built using 7 metabolite biomarker candidates — 1,4-diaminobutane, histidine, 4-ethylphenol, 5-hydroxyindoleacetic acid, 5-hydroxylysine, lipoamide and 4-guanidinobutanoic acid—for brain samples.

We used the volcano plot to determine the significantly changed metabolites between the two groups. For liver samples, the volcano plot in Figure 2.5a shows 26

metabolites with fold change (FC) > 1.2,  $q < 0.1$  and  $p < 0.0038$  (up-regulated, in blue) and 26 metabolites with  $FC < 0.83$ ,  $q < 0.1$  and  $p < 0.003$  (down-regulated, in red). In comparison, for brain samples, there were 6 down-regulated and 33 up-regulated metabolites in the Tg group in Figure 2.5b. Among those significantly changed metabolites, three metabolites (uridine, uridine-H<sub>2</sub>O and spermidine) were positively identified in liver samples, while another four metabolites (1,4-diaminobutane, 5-hydroxyindoleacetic acid, 4-ethylphenol and L-histidine) were positively identified in brain samples.

The significantly changed metabolites could be potential biomarkers for discriminating Tg and wild type. The discrimination power was analyzed using receiver operating characteristic (ROC) curves with the random forest method. In liver samples, uridine, uridine-H<sub>2</sub>O and spermidine gave an area-under-the-curve (AUC) value of 0.867, 0.867 and 0.850, respectively. By combining the three metabolites into a biomarker panel, the discriminating power was increased and the AUC value was determined to be 0.922 within the range of 0.663-1 at the 95% confidence interval (Figure 2.6a). This panel could differentiate Tg from WT with sensitivity of 82.0% and specificity of 82.5%. The permutation test validated the diagnosis power of the biomarker panel (Appendix Figure A2.4a). For the brain samples, 1,4-diaminobutane, histidine, 4-ethylphenol and 5-hydroxyindoleacetic acid gave an AUC value of 0.983, 0.917, 0.917 and 0.913, respectively. Figure 2.6b shows the AUC value of the ROC curve built with the biomarker panel of the four metabolites was 0.983 within the range of 0.933-1 at the 95% confidence interval. The discrimination of Tg from WT had sensitivity of 88.0% and specificity of 88.5%. Thus, the biomarker panel of these four metabolites showed excellent performance.

To further improve the discriminating power, we added other three putatively identified significant metabolites, 5-hydroxylysine, lipoamide, and 4-guanidinobutanoic acid, which gave highly ranked individual AUC values (0.933, 0.933, and 0.932, respectively). The AUC value of the corresponding ROC curve of the seven metabolites was 0.992 (Figure 2.6c). Discrimination of Tg and WT was achieved at 91.0% sensitivity and 91.5% specificity. The ROC results were validated by permutation tests (Appendix Figure A2.4b).

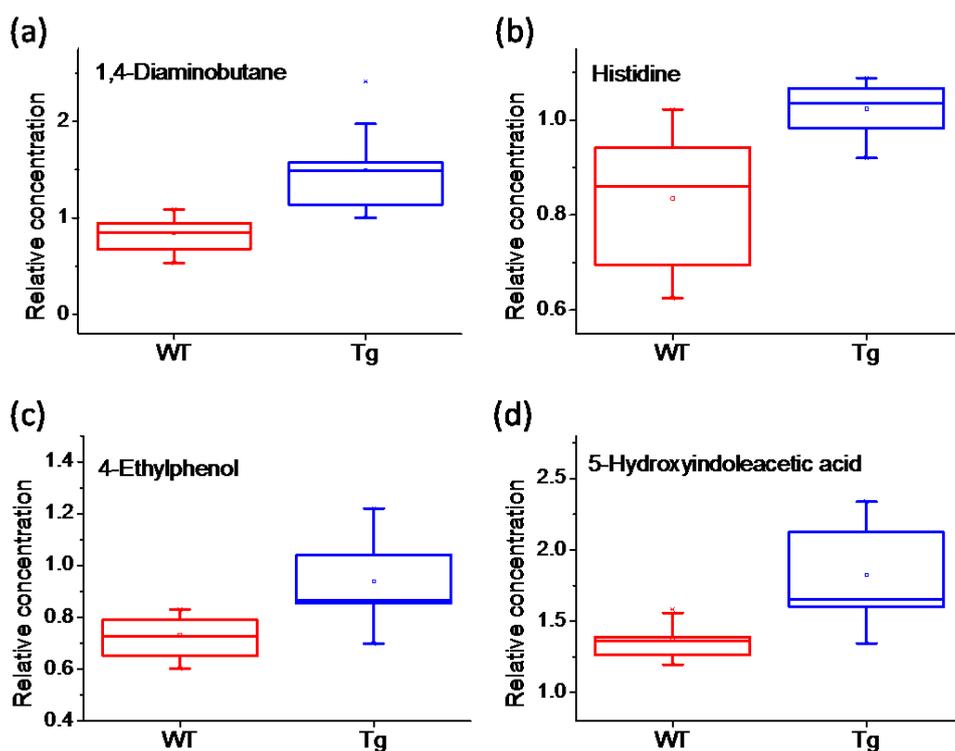


Figure 2.7. Box plots of the relative concentrations of (a) 1,4-diaminobutane, (b) histidine, (c) 4-ethylphenol and (d) 5-hydroxyindoleacetic acid in the WT group and the Tg group.

At the individual metabolite level, Figure 2.7 shows the box plots of relative concentrations of the four significantly charged metabolites in brain tissue samples of the WT group and Tg group. All of them were up-regulated in the Tg group. If some of these biomarker candidates can be validated in future studies with human samples, they could potentially serve as biomarkers for diagnosis of AD. The biological significance of these metabolites is briefly discussed as follows.

1,4-Diaminobutane, also known as putrescine, is a low molecular weight aliphatic amine and one of the simplest polyamines. It plays important roles in physiological functions, such as cell proliferation and apoptosis<sup>91</sup>, oxidative stress response and neuroprotection<sup>92</sup>. Putrescine in the brain may have regulatory activities to many receptors on the surface of neurons<sup>93</sup> and the increased concentration in the brain can cause ischemic brain damage. Also, it can disrupt the functions of gamma amino-butyric acid receptors<sup>93</sup> and be involved in the development of glutamate mediated neurotoxicity<sup>94</sup>. These results imply that the polyamines, especially putrescine, may have a relationship with AD pathogenesis.

Histidine is an essential amino acid for humans and other mammals, which has antioxidant functions. The decreased level of CSF histidine in AD patients has been reported.<sup>95</sup> 5-Hydroxyindoleacetic acid is a metabolic product of serotonin, which can serve as neurotransmitters, and both of them have been reported to be associated with AD. Several studies discovered the decreased CSF levels of 5-hydroxyindoleacetic acid in AD patients<sup>96-98</sup>. In our results, however, the concentration of histidine and 5-hydroxyindoleacetic acid was elevated in brain tissues of Tg mice.

4-Ethylphenol, belonging to polyphenols, has been reported to have antioxidative properties<sup>99</sup>. Although for now there are no studies showing relationships between ethylphenol and AD pathogenesis, its antioxidative functions might have positive effects in treatments.

## **2.4 Conclusions**

We have developed an analytical workflow for quantitative metabolomic profiling of tissue samples with high coverage. It includes a tissue extraction protocol coupled with a chemical isotope labeling LC-MS method. The workflow was applied for in-depth profiling of the amine/phenol-submetabolome of mouse brain and liver tissue samples. We detected a total of 2319 metabolites in more than 80% of the liver samples, and 1769 metabolites in brain samples. Significant metabolomic differences between Alzheimer's disease transgenic mice and wild-type mice have been observed. These results have demonstrated the excellent performance of our workflow for tissue metabolomics. To achieve a higher overall metabolome coverage, the method described herein should be applicable to profile other submetabolomes, such as carboxylic acids<sup>44</sup>, hydroxyls<sup>43</sup>, ketones and aldehydes<sup>45</sup>. Since our workflow can extract metabolites and lipids simultaneously, in the future, we will also work on the lipidomic analysis to provide comprehensive profiling of both the metabolome and the lipidome.

## Chapter 3 Chemical Isotope Labeling LC-MS for Metabolomics of Dexamethasone Side Effects in Rat Tissues

### 3.1 Introduction

Dexamethasone (Dex), first developed in 1957, is a non-selective synthetic glucocorticoid (GC) drug and corticosteroid medication.<sup>100</sup> Dex has both anti-inflammatory and immunosuppressant effects due to the ability to activate nuclear glucocorticoid receptors, thus regulating nutrient metabolism. Dex has been widely used for the treatment of allergic, immunological and inflammatory diseases, such as bronchospasm, rheumatoid arthritis and asthma.<sup>101</sup> Excess or long-term Dex administration can lead to side effects such as osteoporosis, insulin resistance, hyperglycaemia and diabetes.<sup>102</sup> However, a detailed influence on metabolism is still not clear. In this project, we explore the effect of Dex treatment on metabolomes.

Metabolomics is the study of a whole set of metabolites of an organism.<sup>103</sup> It can be used to measure the influence of diseases, environments and treatments on phenotypes. The study of the metabolism of pharmaceutical compounds and their effects on metabolite changes is known as pharmacometabolomics<sup>104</sup>, which is a powerful tool to monitor treatment. Dex can induce complicated metabolic pathway changes. For example, changes were revealed in a previous study using targeted metabolomics profiling to analyze serum samples of a rat model.<sup>105</sup> However, the metabolome coverage of this study was not high.

Our group reported a high-performance CIL LC-MS method for quantitative metabolomics profiling with high coverage. Dansylation labeling technique can be used for profiling amine/phenol submetabolome.<sup>23</sup> Dansylation technique can improve MS sensitivity, quantification precision as well as metabolome coverage. In this work, we applied dansylation LC-MS for evaluating the metabolic changes induced by Dex treatment. Five kinds of rat tissue samples, including brain, liver, heart, kidney and muscle, were analyzed and the metabolic differences between Dex-treated and control groups were observed and investigated.

## 3.2 Experimental Section

### 3.2.1 Sample Collection and Processing

Two groups (Dex-treated and control group) of male rats with age of 6-8 weeks and weight of 200-250 g were used. There were 6 rats in each group and they were kept under same standard environmental conditions. The Dex group and control group were intramuscularly injected with 2.5mg/kg of Dex or saline twice a week for 14 weeks, respectively. Then all mice were sacrificed and tissue samples were taken, snap-frozen in liquid nitrogen and stored in -80 °C until analysis. There were 5 kinds of samples, with 5 control and 4 Dex brain tissues, 6 control and 5 Dex kidney tissues, 5 control and 5 Dex heart tissues, 4 control and 4 Dex liver tissues, 6 control and 5 Dex muscle tissues. There were 49 tissue samples in total.

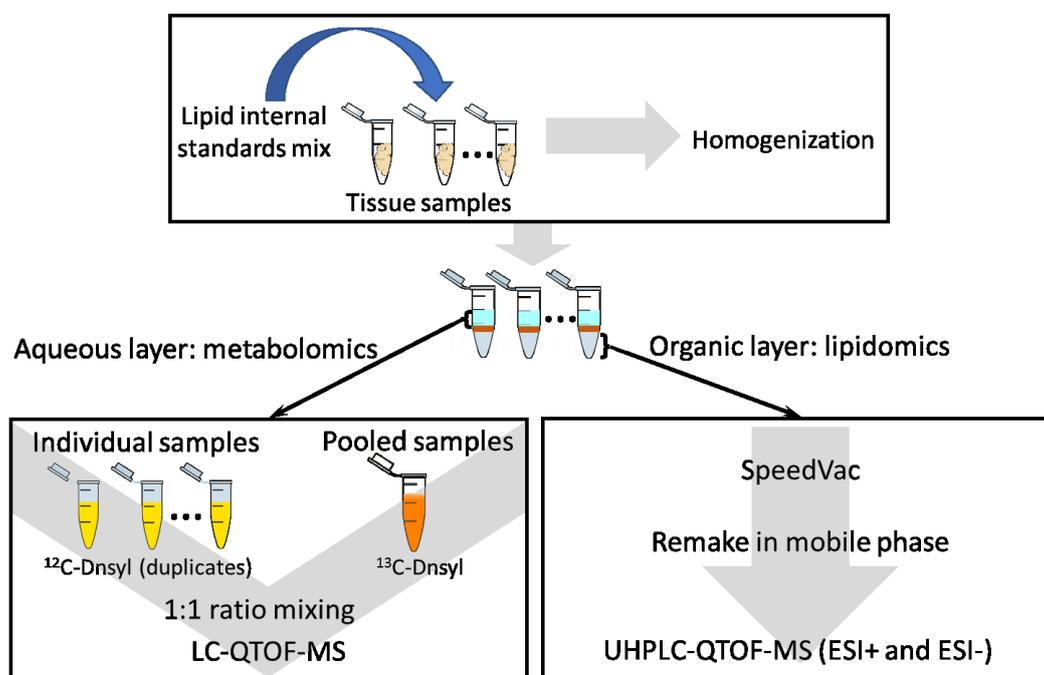


Figure 3.1. Workflow of tissue extraction, dansylation isotope labeling LC-MS and lipids analysis.

### 3.2.2 Tissue Extraction

Figure 3.1 displays the overall workflow. Ice-cold methanol, DCM and water were prepared first. After measuring the weight of a tissue sample, 2 mL of methanol and 0.42 mL of water per gram of tissue were added to the sample. The tissue sample was homogenized for 20 s using a Bio-Gen PRO200 Homogenizer (PRO Scientific, USA). After adding of 2 mL/g of DCM and 1 mL/g of water to the sample, the mixture

was vortexed 1 min. The sample was then incubated at -20 °C for 15 min. Finally, the solution was centrifuged at 16000 g for 10 min at 4 °C. The upper aqueous layer (with polar metabolites) and lower organic layer (with lipophilic compounds) were transferred into separate vials. Each upper layer was split into 4 aliquots: 2 x 15 µL for amines/phenols labeling and 30 µL for making the pooled samples for different kinds of tissues respectively. The leftover sample was stored at -80 °C for future analyses. The organic layer with lipids was analyzed by another person and is not discussed here.

### 3.2.3 Dansyl Labeling

The labeling method was adapted from the protocol in a previous report.<sup>23</sup> Fifteen µL of individual sample or pooled sample was thawed and vortexed, and then the solvent was removed at room temperature using a Savant SC110A Speed Vac. The sample was re-dissolved to 37.5 µL with 2:1 water/ACN. Then 12.5 µL of 250 mM sodium carbonate/sodium bicarbonate buffer was added to the samples. The solution was mixed with 25 µL of freshly prepared <sup>12</sup>C-DnsCl solution (18 mg/mL for light labeling, individual samples) or <sup>13</sup>C-DnsCl solution (18 mg/mL for heavy labeling, pooled sample). After incubation for 45 min at 40 °C, 5 µL of 250 mM sodium hydroxide solution was added to the reaction mixture, followed by incubation at 40 °C for another 10 min. Finally, 25 µL of formic acid (425 mM) in 1:1 ACN/H<sub>2</sub>O was added.

### 3.2.4 LC-UV Quantification

The total dansyl-labeled metabolite amount of each sample was normalized using a LC-UV protocol.<sup>86</sup> The instrument was a Waters ACQUITY UPLC system with a photodiode array (PDA) detector. A Phenomenex Kinetex reversed-phase C18 column (50 mm × 2.1 mm, 1.7 µm particle size, 100 Å pore size) was used with a fast step-gradient. The flow rate was 0.45 mL/min and the total run time was 6 min. The peak area integrated using the Empower software which was related to the total dansyl-labeled metabolite concentration. According to the quantification results, <sup>12</sup>C-labeled individual sample and the <sup>13</sup>C-labeled pooled sample were mixed in equal amounts. Besides, <sup>12</sup>C- and <sup>13</sup>C-labeled pooled samples were mixed in equal amounts serving as a QC sample.

### 3.2.5 LC-FTICR-MS Analysis

Samples were analyzed using a Bruker 9.4 T Apex-Qe FTICR mass

spectrometer (Bruker, Billerica, MA) with electrospray ionization (ESI), coupled with an Agilent capillary 1100 binary system (Agilent, Palo Alto, CA). An Agilent eclipse plus reversed-phase C18 column (100 × 2.1 mm, 1.8 μm particle size,) was used. All mass spectra were collected in the positive ion mode. All the samples were injected in random order. QC samples and amino acid standards were injected every 10 runs to monitor the performance of the LC-MS.

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

After LC-FTICR-MS analysis, the entire centroid peak lists were exported from Bruker Data Analysis software. IsoMS<sup>87</sup> was used to pick peak pairs, reduce redundant pairs, calculate peak-pair intensity ratios, and group the peak pairs found in adjacent spectra. After the alignment of same peak pairs from different runs using Alignment program, the Zerofill program was applied to fill in the missing values.

Based on accurate mass and retention time matches, positive metabolite identification was performed using dansyl standard library search, which contains 273 unique dansylated amines/phenols. Based on accurate mass search, putative identification was performed by using the Human Metabolome Database (HMDB) ([www.hmdb.ca](http://www.hmdb.ca)) and the Evidence-based Metabolome Library (EML) (<http://www.mycompoundid.org/>).

Principal components analysis (PCA) and pathway analysis were carried out using MetaboAnalyst 4.0 software (McGill University, Montreal, Canada). The fold change and p-value between groups were calculated using Microsoft Excel. Volcano plots were constructed using OriginPro 8.0 (OriginLab). The q-value, also known as multiple-testing-corrected p-value, was calculated using R and BioConductor ([www.bioconductor.org](http://www.bioconductor.org)).

## **3.3 Result and Discussion**

### **3.3.1 Metabolite Detection**

Since the injection amount of labeled metabolites onto the LC-MS can affect the detected number of peak pairs, we first determined the best injection amount of rat tissue samples. We injected increasing amounts of a mixture of <sup>12</sup>C- and <sup>13</sup>C-labeled pooled sample, formed by combining all individual aliquots from 5 kinds of tissue samples. The result showed that with the injection amount of 3 nmol of labeled

metabolites, the number of peak pairs reached a maximum. Thus, for each run, all individual brain, heart, kidney, liver, muscle and QC samples were analyzed using LC–MS with 3 nmol injection.

After LC-MS acquisition and data processing, different but similar number of peak pairs were detected for all 5 kinds of tissues. For brain tissues, a total of 1331 peak pairs were commonly detected in more than 80% of the 18 samples. Using dansyl standard library search, 62 metabolites were positively identified. For the remaining metabolites, using MyCompoundID and accurate mass search with a mass accuracy tolerance of 10 ppm, 659 metabolites were putatively identified in the HMDB library and 1264 metabolites were putatively identified in the EML database with 1 reaction. Thus, totally, 970 metabolites can be matched (73%). For heart tissues, 1444 metabolites were commonly measured and 1068 metabolites can be matched (74%). For kidney tissues, 1469 metabolites were commonly measured and 1103 metabolites can be matched (74%). For liver tissues, 1768 metabolites were commonly measured and 1310 metabolites can be matched (74%). For muscle tissues, 1625 metabolites were commonly measured and 1176 metabolites can be matched (72%). As a result, we achieved high submetabolome coverage and more than 70% of metabolites can be matched with different confidence levels for all 5 kinds of samples.

### **3.3.2 Comparative Metabolome Analysis between Dex-treated and Control Groups**

Multivariate analysis was used to visualize the metabolic differences between Dex-treated and control groups. Principal component analysis (PCA) plots in Figure 3.2 show a clear separation between Dex and control groups in brain, liver and muscle tissues, a separation with some overlapping data points in kidney tissues and little separation in heart tissues. The results suggest that, with Dex treatment, significant metabolic changes have occurred in brain, liver and muscle tissues, and some kind of changes have occurred in kidney, while small changes have occurred in heart. These observations can be inferred from clinical reports of GC therapy. For muscle tissues, it was reported that GC can induce weakness and atrophy.<sup>106</sup> For liver tissues, Dex can cause insulin resistance and lead to hyperglycaemia<sup>107</sup>, which may disturb liver functions. Further, detoxify of side effects of Dex can affect the function of liver and kidney. For brain tissues, GC treatment may cause neuropsychiatric diseases, such as

anxiety, depression and psychosis,<sup>108</sup> which may explain the huge metabolic differences between the two groups.

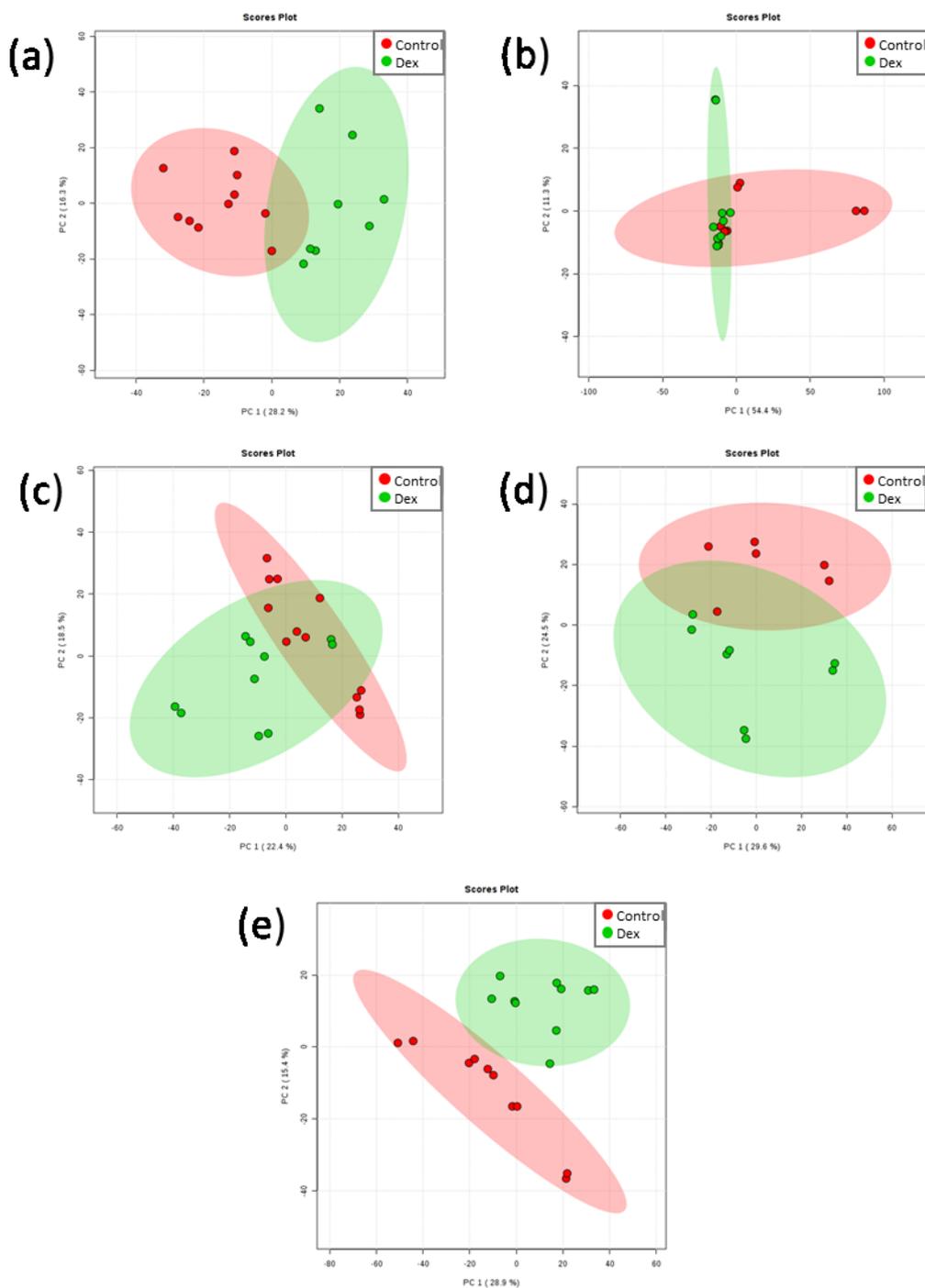


Figure 3.2. PCA scores plot of Dex vs. control of (a) brain, (b) heart, (c) kidney, (d) liver, (e) muscle tissue samples.

These observations can be further validated by univariate analysis using volcano plot. As we can see from Figure 3.3, metabolites with fold change (FC) > 1.2 are labeled in blue (up-regulated) or in red (down-regulated) with q-value < 0.05. A large number of significant metabolites were detected in brain, liver and muscle tissues, with 237

down-regulated and 290 up-regulated metabolites, 313 down-regulated and 68 up-regulated metabolites, 174 down-regulated and 245 up-regulated metabolites detected, respectively. For kidney tissues, there were 140 down-regulated and 22 up-regulated metabolites, while for heart tissues, only 1 down-regulated and 7 up-regulated metabolites, consistently with the results indicated in PCA results.

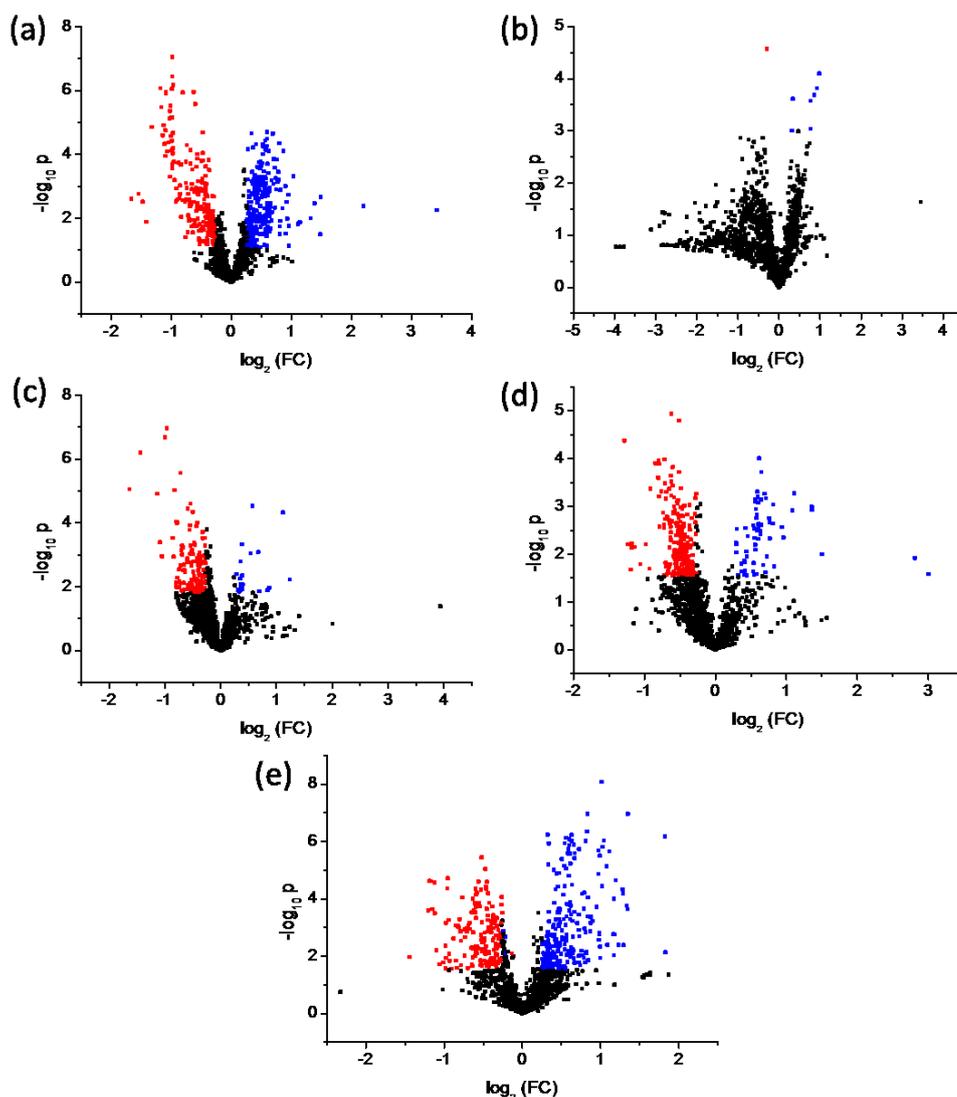


Figure 3.3. Volcano plot of (a) brain, (b) heart, (c) kidney, (d) liver, (e) muscle tissue samples. Metabolites with fold change (FC) > 1.2 were labeled in blue, and metabolites with FC < 0.83 were labeled in red, both with q-value < 0.05.

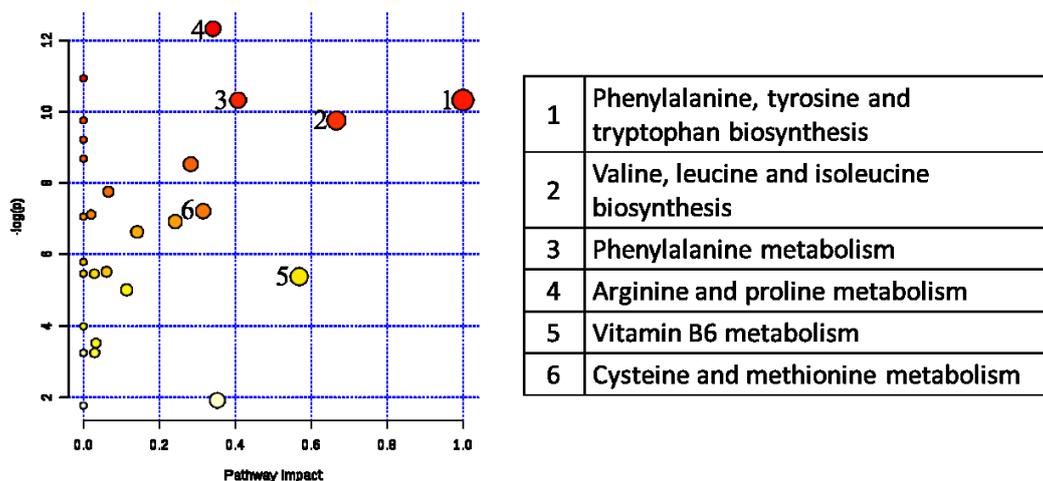


Figure 3.4. Overview of metabolic pathway analysis of brain tissues.

### 3.3.3 Pathway Analysis

Using all the positively identified metabolites, pathway analysis was carried out with Metaboanalyst 4.0 for all five kinds of tissues respectively. These compounds were matched to the rat pathway library, which includes 81 pathways. Depending on how many hits each pathway had and the importance factors of the hit compounds, a pathway impact and a p-value were calculated for each pathway. Taking brain's pathway result as an example in Figure 3.4,  $-\log(p\text{-value})$  is plotted against the pathway impact. We compared the affected pathways among all 5 kinds of tissues and found 6 commonly shared pathways with impact  $> 0.3$ , including “Arginine and proline metabolism”, “Vitamin B6 metabolism”, “Phenylalanine metabolism”, “Phenylalanine, tyrosine and tryptophan biosynthesis”, “Valine, leucine and isoleucine biosynthesis”, and “Cysteine and methionine metabolism”.

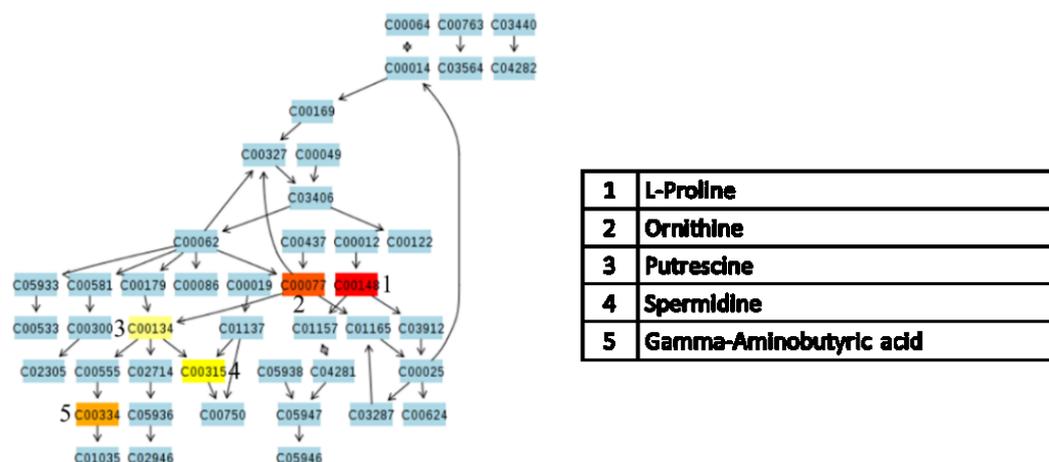


Figure 3.5. Pathway of arginine and proline metabolism of brain tissue.

Among these common pathways, we selected “Arginine and proline metabolism” as an example (Figure 3.5) for an in-depth discussion, which is one of the most affected pathways for all 5 kinds of tissues. On the pathway schematic, light blue means the metabolites cannot be positively identified, but is used as a background for enrichment analysis. The 5 metabolites with other colors (varying from yellow to red) are the positively identified metabolites with different levels of significance, including proline, ornithine, putrescine, spermidine and gamma-aminobutyric acid. A red metabolite has a more significant change between the Dex and control groups than a yellow metabolite. With the help of pathway analysis, we can study and compare the importance of some pathways and metabolites among the 5 kinds of tissues. This would facilitate our understanding of the effect of Dex treatment. Figure 3.6 shows the relative metabolite concentrations of putrescine, spermidine, ornithine, proline in the control and Dex groups for all 5 kinds of tissues. The significance of these metabolites is briefly discussed in the following section.

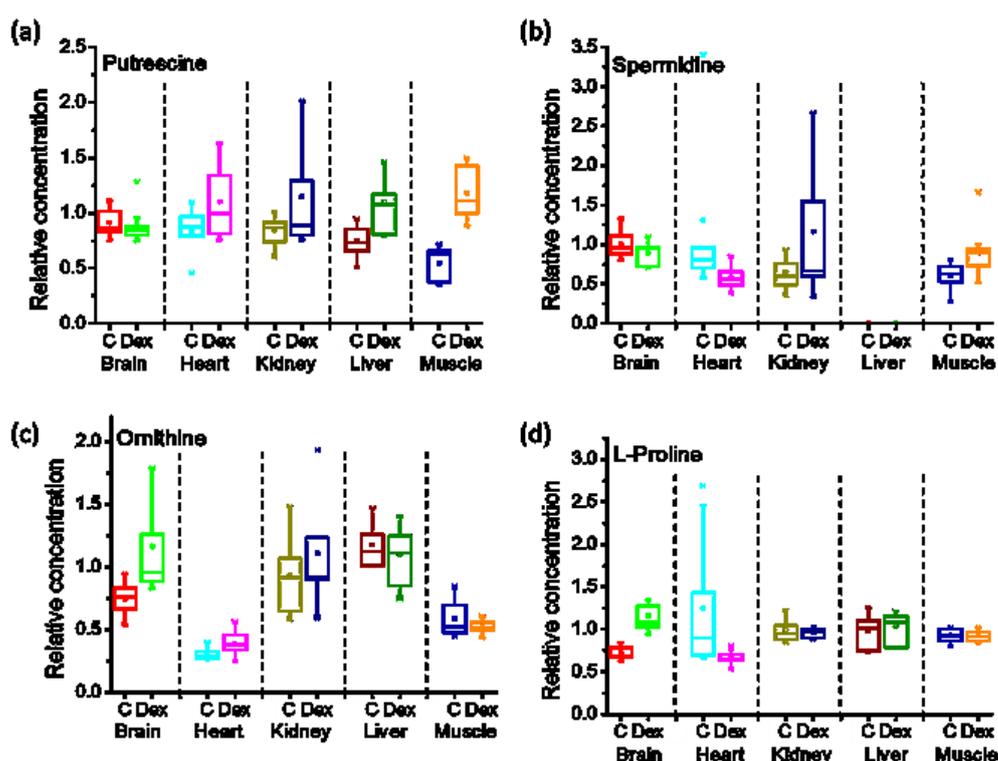


Figure 3.6. (a) Box plots of the relative concentrations of (a) 1,4-diaminobutane, (b) spermidine, (c) ornithine, (d) proline.

Putrescine is also known as 1,4-diaminobutane, and together with spermidine, they are all polyamines. Polyamines play crucial roles in cell proliferation and differentiation.<sup>109</sup> In this work, putrescine was up-regulated in Dex group in liver and

muscle tissues; while its concentration also increased in heart and kidney tissues, the increase was not significant. Thus, Dex treatment may affect proliferative capacity. Interestingly, the level of spermidine remained stable for all tissues (not detected in liver tissue). Previously, a study of Dex treatment on spinal cord injury reported similar results, in which increased putrescine was observed, while spermidine remained constant.<sup>110</sup> It may reflect the neurotrophic and regenerative effects of polyamines from Dex stimulation.

Ornithine is a non-proteinogenic amino acid. Derivatization of ornithine produces polyamines such as, putrescine, spermidine, essential for cell proliferation.<sup>111</sup> Dex treatment can reduce the expression and activity of ornithine decarboxylase in spleen, but not in the liver or kidney, which can catalyze decarboxylation of ornithine to produce polyamines.<sup>112</sup> From our results, the ornithine level was not affected, except in brain tissues in which it was slightly increased.

Proline is a non-essential proteinogenic amino acid. Hydroxylation of proline plays important roles in the stability of collagen. Injection of Dex and proline can suppress bone collagen synthesis.<sup>113</sup> Proline can also be related with neuropathophysiology of some disorders.<sup>114</sup> From our results, proline level in brain samples significantly increased after Dex treatment, which may cause brain damage.

### **3.4 Conclusion**

In this work, we applied CIL LC-MS method to study the effect of Dex treatment on amine/phenol submetabolome of rats with high coverage. Five kinds of rat tissue samples were analyzed, including brain, heart, kidney, liver and muscle. Significant metabolic differences were observed between Dex and control groups in brain, liver and muscle tissues, while some metabolic differences were detected in kidney tissues and minor differences for heart tissues. We also performed pathway analysis using positively identified metabolites for all 5 kinds of tissues and many common pathways were found, which means the Dex treatment had some common effects on different tissues. Of those pathways, one of the most affected pathways, Arginine and proline metabolism, was taken for in-depth discussion. The biological significance of involved identified metabolites, including putrescine, spermidine, ornithine, proline, were investigated. In the future, we will carry out other labeling techniques, such as carboxylic acids<sup>44</sup>, hydroxyls<sup>43</sup>, ketones and aldehydes<sup>45</sup>, to achieve

higher metabolome coverage. We will also combine the results from lipidomic studies, which will help us understand more about the effect of Dex treatment.

## Chapter 4 Development of Chemical Isotope Labeling LC-MS Methods for Metabolomics Studies of Rheumatoid Arthritis Disease

### 4.1 Introduction

Rheumatoid arthritis (RA) is a common and chronic symmetric polyarticular arthritis. In this autoimmune disease, tissues are mistakenly attacked by the immune system.<sup>115</sup> An estimated 24.5 million of people are affected by RA worldwide.<sup>116</sup> RA mainly affects the small joints of hands, feet and cervical spine, causing inflammation in the synovium. The cause of RA is unknown. Hypothesis include genetic and environmental factors, e.g. smoking. Currently, there is no cure for RA. Treatment just helps reduce joint pain and inflammation. What is worse, the diagnosis of RA is difficult. Diagnosis is based on comprehensive clinical features, including joint swelling, X-ray testing, rheumatoid factor testing, etc., while early diagnosis is hampered by accuracy and sensitivity of currently available biomarkers. In this work, we hunt for metabolite biomarkers for early stage diagnosis of RA with good sensitivity and precision.

Current conventional biomarkers includes the measurement of RA factor, anti-cyclic-citrullinated-peptide (anti-CCP) antibody testing,<sup>117</sup> and the antibody against the Fc region of IgG.<sup>118</sup> Besides, inflammation markers, erythrocyte sedimentation rates, and C-reactive protein (CRP) can also be used to monitor disease development of RA.<sup>119</sup> However, the relatively high false positive rate remains a significant problem. Metabolomics is a powerful tool to characterize the complex biochemical phenotypes, and is an emerging field for biomarker discovery. Many metabolomics studies of RA have been reported.<sup>120-121</sup> But, their coverage of the metabolome is not high. As a result, new techniques are required to improve metabolome coverage as well as quantification accuracy.

Previously, we reported the use of high-performance chemical isotope labeling (CIL) LC-MS method for the profiling of amine/phenol submetabolome.<sup>23</sup> This derivatization method can significantly improve LC separation, MS sensitivity, and also provide precise relative quantification result. In this work, we applied CIL LC-MS to analyze amine/phenol and carboxyl submetabolomes of serum samples from two cohorts of RA patients. We first characterized the submetabolome changes between the

early RA and healthy control groups, and identified several significantly changed metabolites as biomarker candidates. Then, we compared the metabolic differences between RA patients on methotrexate (MTX) treatment and without MTX treatment and healthy controls, and between RA patients with anti-CCP antibody positive and with anti-CCP antibody negative and controls. The discovered metabolites were consistently discriminatory, validating their good consistency. We also compared the metabolite changes between responders and non-responders and discovered potential biomarkers for disease activity, which may help the administration of medication in the clinic. These findings indicate that RA patients can be well differentiated from healthy controls and the discovered biomarkers can be helpful for both diagnosis and medication administration if further validated.

## 4.2 Experimental Section

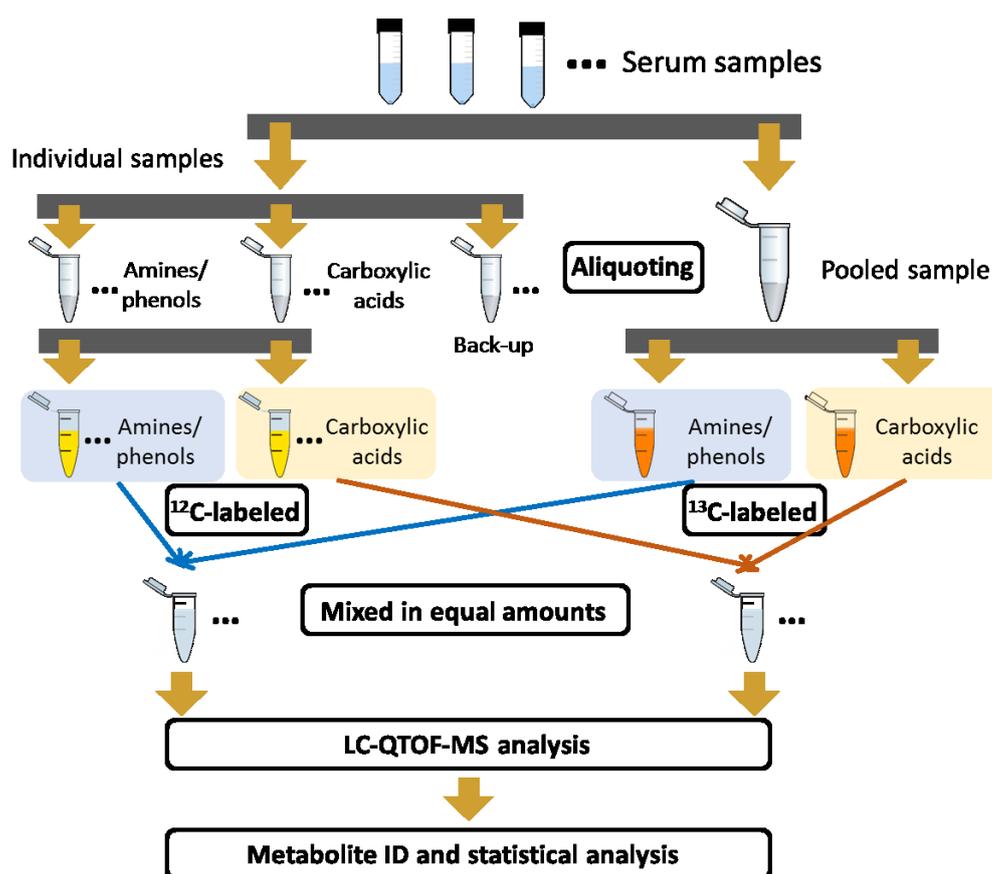


Figure 4.1. Workflow of the project.

Figure 4.1 shows the whole workflow of this work. It mainly includes the following steps: (1) serum sample collection and aliquoting; (2) dansylation and DmPA bromide labeling; (3) LC-UV quantification of total amount of dansyl labeled

metabolites in each individual sample and pooled sample for sample amount normalization; (4) mixing of  $^{12}\text{C}$ -labeled individual samples and  $^{13}\text{C}$ -labeled pooled sample in equal amounts; (5) LC-QTOF-MS analysis of  $^{12}\text{C}$ -/ $^{13}\text{C}$ - mixtures; (6) data processing using R programs, including peak pair picking, peak pair filtering, peak pair ratio calculation and peak pair grouping; (7) metabolite identification; and (8) uni- and multi-variate statistical analysis. Detailed experimental conditions are described below.

#### **4.2.1 Serum Sample Collection and Aliquoting**

A total of 250 serum samples were taken from two RA cohorts. Cohort A samples were collected from 50 RA patients, with 39 females (mean age 49.9) and 11 males (mean age 47.8), symptom duration <3 years, disease activity score (DAS) >3.7, naïve to b-DMARD, and 50 age and sex-matched healthy controls. They were labeled as B1D and B1C, respectively. Cohort B samples were collected from 50 RA patients, with 40 females (mean age 53.4) and 10 males (mean age 57.2), symptom duration <5 years, both pre- and post- (3 months) treatment with TNFi, as well as another 50 age and sex-matched healthy controls. Cohort B samples were labeled as B2D1, B2D2, B2C, respectively. Also, the cohort A samples were divided into with-methotrexate (MTX) treatment (Y) group and without-MTX treatment (N) group. Moreover, from the perspective of anti-CCP antibody testing, the two cohorts of RA patients can be divided into anti-CCP positive (P) and anti-CCP negative (N) groups. Besides, cohort B RA samples with disease activity score (DAS) change of >1.2 and Post DAS < 3.2 are defined as responders (r), others are defined as non-responders (nr). So, cohort B RA samples include B2D1 r, B2D1 nr, B2D2 r, B2D2 nr, four groups in total. The analysis about RA, MTX treatment, anti-CCP responders and non-responders was all performed in this work.

Samples were non-fasting blood samples, taken at least 2 hours after rising from the bed at different facilities. After centrifuge, serum samples were taken and aliquoted and stored at -80 °C and shipped. In our lab, each serum sample was split into 8 aliquots, with 2 x 15  $\mu\text{L}$  for amines/phenols analysis, 2 x 15  $\mu\text{L}$  for carboxylic acids analysis, 2 x 15  $\mu\text{L}$  for back-up, 35  $\mu\text{L}$  for making the pooled sample. All the remaining samples were stored leave in. These aliquots were all stored in -80 °C freezer until further analysis.

#### 4.2.2 Dansyl Labeling

The dansyl labeling method was adapted from a previously reported protocol.<sup>23</sup> Briefly, 15  $\mu\text{L}$  of individual serum sample (experimental duplicates) or 15  $\mu\text{L}$  of the pooled sample was thawed in 4  $^{\circ}\text{C}$  freezer and then spun down. The sample was mixed with 45  $\mu\text{L}$  of methanol, and then the mixture was stored at -20  $^{\circ}\text{C}$  for 2 hours to precipitate the proteins. After this, the mixture was centrifuged at 16,000 g for 15 min and 45  $\mu\text{L}$  of supernatants was taken and dried using a Speed-Vac centrifugal evaporator. The sample was re-dissolved to 37.5  $\mu\text{L}$  with 2:1 water/ACN. Then, 12.5  $\mu\text{L}$  of 250 mM sodium carbonate/sodium bicarbonate buffer was added to the samples to adjust the pH to 10. The solution was mixed with 25  $\mu\text{L}$  of freshly prepared  $^{12}\text{C}$ -DnsCl solution (18 mg/mL) (for light labeling, individual samples) or  $^{13}\text{C}$ -DnsCl solution (18 mg/mL) (for heavy labeling, pooled sample). After incubation for 45 min at 40  $^{\circ}\text{C}$ , 5  $\mu\text{L}$  of 250 mM sodium hydroxide solution was added to the reaction mixture to quench the reaction. The solution was then incubated at 40  $^{\circ}\text{C}$  for another 10 min. Finally, 25  $\mu\text{L}$  of formic acid (425 mM) in 1:1 water/ACN was added to make the solution acidic.

#### 4.2.3 DmPA Bromide Labeling

The DmPA bromide labeling method was adapted from a previous protocol.<sup>122</sup> First, 15  $\mu\text{L}$  of the individual or pool sample was thawed at 4  $^{\circ}\text{C}$  and then spun down. The sample was mixed with 45  $\mu\text{L}$  of ACN. The mixture was then stored at -20  $^{\circ}\text{C}$  for 2 h to precipitate the proteins, and centrifuged at 16,000 g for 15 min. 45  $\mu\text{L}$  of supernatant was taken and mixed with 10  $\mu\text{L}$  of 0.5 M triethanolamine and 25  $\mu\text{L}$  of freshly prepared  $^{12}\text{C}$ -DmPA bromide solution (10 mg/mL) (for light labeling) or  $^{13}\text{C}$ -DmPA bromide solution (10 mg/mL) (for heavy labeling). The mixture was then incubated at 80  $^{\circ}\text{C}$  for one h. After cooling down at -20 $^{\circ}\text{C}$ , 20  $\mu\text{L}$  of 0.2 M Tri-Gly was added followed by 30-min incubation at 80  $^{\circ}\text{C}$  to quench the reaction.

#### 4.2.4 LC-UV Quantification

The total dansyl-labeled metabolite amount of each sample was quantified and normalized using a protocol previously reported based on LC–UV measurement.<sup>86</sup> The instrument for detection was a Waters ACQUITY UPLC system with photodiode array (PDA) detector. A Phenomenex Kinetex reversed-phase C18 column (50 mm  $\times$  2.1 mm,

1.7  $\mu\text{m}$  particle size, 100  $\text{\AA}$  pore size) was used to achieve a fast step-gradient. According to the quantification results,  $^{12}\text{C}$ -labeled individual samples and the  $^{13}\text{C}$ -labeled pooled sample were mixed in equal amounts for both dansyl and DmPA labeling, respectively. Besides,  $^{12}\text{C}$ - and  $^{13}\text{C}$ -labeled pooled samples were mixed in equal amounts serving as a QC sample.

#### **4.2.5 LC-QTOF-MS Analysis**

Samples were analyzed using a Maxis II Quadrupole Time-of-flight (QTOF) mass spectrometer (Bruker, Billerica, MA) with electrospray ionization (ESI), coupled with an UltiMate 3000 UHPLC (Thermo Scientific, MA). An Agilent Eclipse Plus RP C18 column (100  $\times$  2.1 mm, 1.8  $\mu\text{m}$  particle size,) was used. All mass spectra were collected in the positive ion mode. All the samples were injected in random order. QC samples and amino acid standards were injected every 10 runs to monitor the performance of the LC-MS. For DmPA labeling,  $^{12}\text{C}$ - and  $^{13}\text{C}$ -labeled blank samples were also injected every 10 runs to be used for blank peak pair subtraction.

#### **4.2.6 Data Processing, Metabolite Identification and Statistical Analysis**

After LC-QTOF-MS analysis, entire centroid peak lists were exported from Bruker Data Analysis software. IsoMS<sup>87</sup> was used to pick peak pairs, calculate peak-pair intensity ratios, group the peak pairs found in adjacent spectra. After the alignment of same peak pairs from different runs using Alignment program, the Zerofill program was applied to fill in missing values.

Based on accurate mass and retention time matches, positive metabolite identification was performed using dansyl standard library search, which contains 273 unique dansylated amines/phenols, and acid standard library search, which contains 188 acid standards. Based on accurate mass search, putative identification was performed by searching the Human Metabolome Database (HMDB) ([www.hmdb.ca](http://www.hmdb.ca)) and the Evidence-based Metabolome Library (EML) using MycompoundID (<http://www.mycompoundid.org/>).

Principal components analysis (PCA), partial least squares discriminant analysis (PLS-DA) and receiver operating characteristic (ROC) curve analysis was carried out using MetaboAnalyst<sup>84</sup> software (McGill University, Montreal, Canada). The fold change and p-value between groups was calculated using Microsoft Excel.

Volcano plots were constructed using OriginPro 8.0 (OriginLab). The q-value, multiple-testing-corrected p-value, was calculated using R and BioConductor (www.bioconductor.org).

### 4.3 Results and Discussion

#### 4.3.1. Analysis of Amine- and Phenol- Submetabolome

##### 4.3.1.1 Metabolite Detection

An average of  $5804 \pm 302$  peak pairs or metabolites were detected in duplicate experiments. In total, 3415 metabolites were commonly detected in more than 80% of the 500 samples. Using the dansyl standard library search, we positively identified 116 metabolites. For the remaining metabolites, using MyCompoundID in the HMDB library with accurate mass search and a mass accuracy tolerance of 10 ppm, we putatively identified 659 metabolites. In the EML database with 1 reaction, with accurate mass search and a mass accuracy tolerance of 10 ppm, we putatively identified 1264 metabolites. Thus, for the measured 3415 metabolites, a total of 2039 metabolites can be matched (60%).

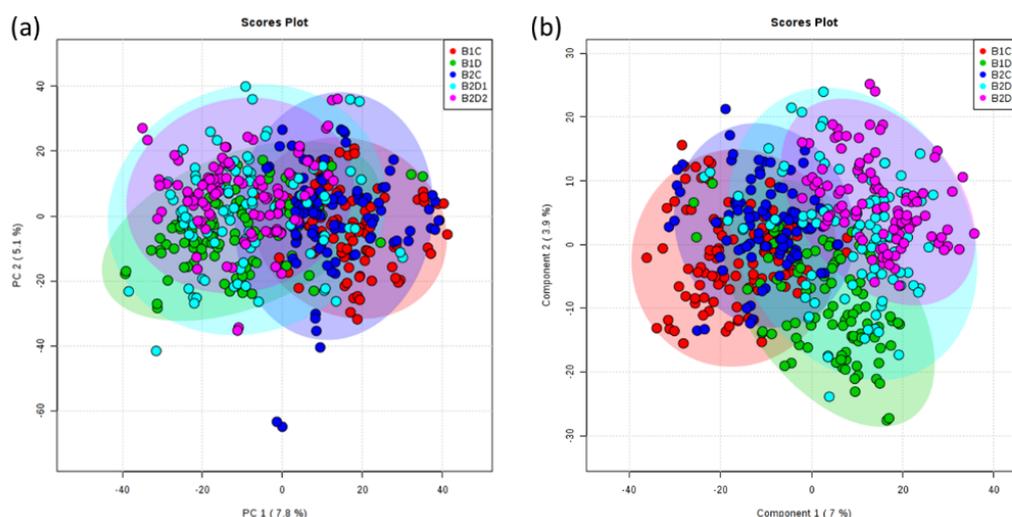


Figure 4.2. (a) PCA and (b) PLS-DA scores plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 ( $R^2=0.897$ ,  $Q^2=0.773$  for PLS-DA plot) for dansylation labeling.

##### 4.3.1.2 Comparative Metabolome Analysis for RA Biomarker Discovery

We first compared metabolome differences between healthy control and RA groups, and Appendix Figure A4.1 shows PCA scores plot of the whole data sets with QC samples. The QC samples cluster together closely, suggesting the good instrument performance and data quality. However, two outliers of duplicates of sample #152, in

B2D1 group, were observed. This sample also had abnormal results in terms of concentration. It turned out to be a urine sample, wrongly collected during sample preparation. Thus, it was excluded.

After data exclusion, Figure 4.2 displays PCA and PLS-DA scores plots of all the five groups of samples. In the PCA plot, we can see two control groups from two cohorts cluster together on the right, while three RA groups cluster together on the left. There was a separation between control and RA groups, meaning there were significant metabolic differences between the two groups. The separation can be seen more clearly in the PLS-DA plot, although there are still many overlapping data points. Besides, the separation between the control and RA groups was observed for both two cohorts, which further validated the metabolomic differences.

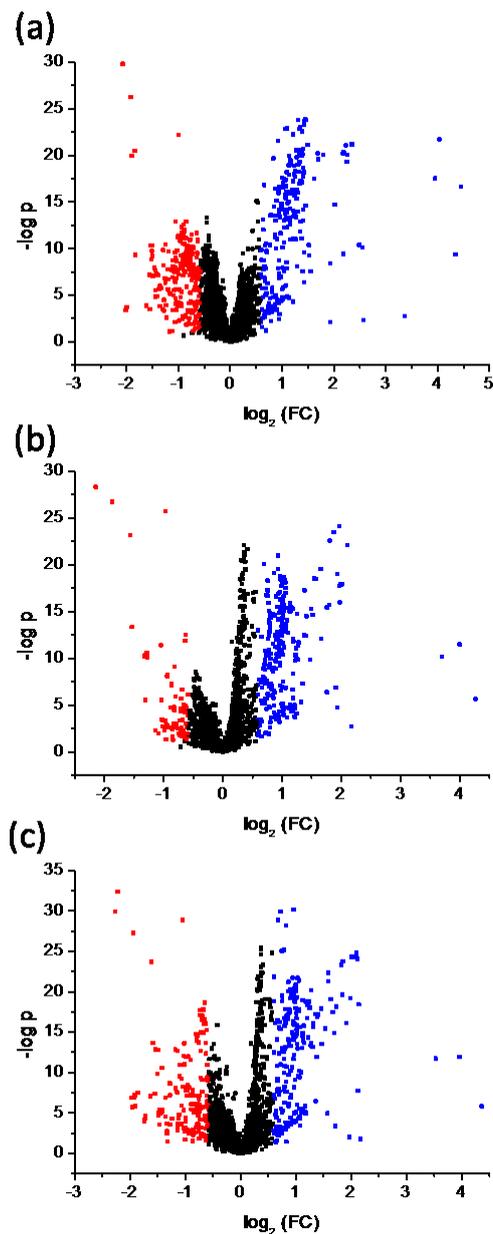
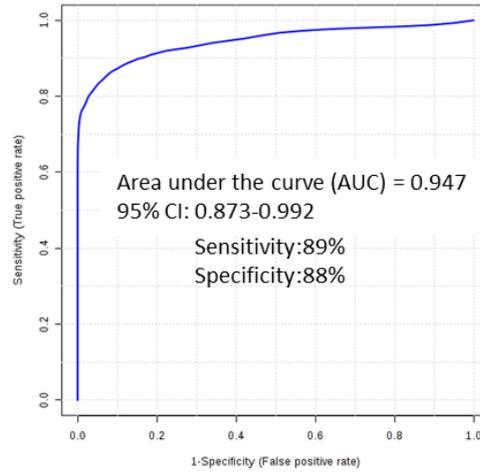


Figure 4.3. Volcano plot of the changes of metabolites between (a) B1C and B1D, (b) B2C and B2D1, (c) B2C and B2D2. Metabolites with fold change (FC) > 1.5 or FC < 0.67 and q-value < 0.05 were highlighted.

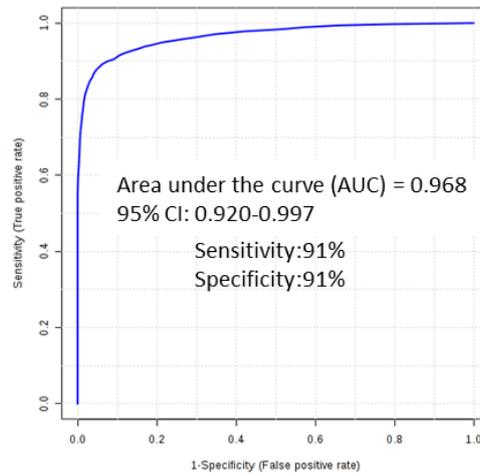
Then univariate analysis, volcano plot, was used to determine significantly changed metabolites that separated RA from control groups. Figure 4.3 shows three volcano plots between control and RA for two cohorts of samples. Of those, metabolites with an average fold change greater than 1.5 (up-regulated) are marked in blue, and those with fold change smaller than 0.67 (down-regulated) are marked in red, both with q-values smaller than 0.05. When comparing B1C versus B1D (Figure 4.3a), 256 metabolites were down-regulated, and 247 were up-regulated in B1D. Among these 503 significant metabolites, 18 can be positively identified (see Supplemental Table

S4.1 for the list). When comparing B2C versus B2D1 (Figure 4.3b), 92 metabolites were down-regulated, and 278 were up-regulated in B2D1. Among these 370 significant metabolites, eight can be positively identified (see Supplemental Table S4.2 for the list). When comparing B2C versus B2D2 (Figure 4.3c), 184 metabolites were down-regulated, and 259 were up-regulated in B2D2. Among these 443 significant metabolites, 18 can be positively identified (see Supplemental Table S4.3 for the list). We compared the metabolites listed in Supplemental Table S4.1, S4.2 and S4.3 and found four common metabolites, cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine, with similar fold changes.

(a)



(b)



(c)

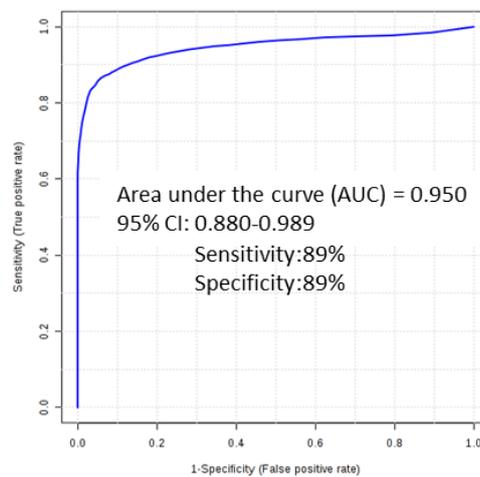


Figure 4.4. ROC curve built using 4 common positively significant metabolites: cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine, for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2.

Receiver operating characteristic (ROC) curves were generated by random forest method to test differentiating power of biomarker candidates. Firstly, we input all the significant metabolites into Metaboanalyst, and then different numbers of the top-ranked metabolites were used to build ROC curves. Appendix Figure A4.2 shows a series of ROC curves built with different numbers of variables. However, even with only top five metabolites, the discrimination reached highest with AUC around 1 for all the three kinds of comparisons, meaning the metabolic differences between RA and control were very significant. For clinical diagnostic purposes, we want to achieve high sensitivity and specificity while using a minimal number of biomarkers. Since we had four common identified significant metabolites, we tested the discriminating power of the biomarker panel formed with these four metabolites. From Figure 4.4a, for B1C versus B1D, the AUC value of the ROC curve was 0.947, within the range of 0.874-0.993 at the 95% confidence interval, and the discrimination of RA from control can be achieved at 89% sensitivity and 88% specificity. For B2C versus B2D1 (Figure 4.4b), the AUC value of the ROC curve was 0.968, within the range of 0.92-0.997 at the 95% confidence interval, and the discrimination can be achieved with both sensitivity and specificity at 91%. For B2C versus B2D2 (Figure 4.4c), the AUC value of the ROC curve was 0.951, within the range of 0.879-0.991 at the 95% confidence interval, and the discrimination can be achieved with both sensitivity and specificity at 89%. The results suggest the biomarker panel with these four metabolites had good performance for both cohort A and cohort B samples. Besides, as we can see from Figure 4.5, the box plots of these four metabolites show that they constantly up- or down-regulated in both cohort A and B, and the differences between RA and control were very significant. We also added more metabolites into the biomarker panel to further improve discriminating power. Three common unidentified significant metabolites, dansyl-702, dansyl-9056, dansyl-14033, were found with highly ranked univariate AUC values among all the three kinds of comparisons. As we can see from Appendix Figure A4.3, the ROC curves of newly formed biomarker panel with these seven common metabolites had much improved performance for all the three kinds of comparisons with AUC value of almost 1. These seven metabolites don't correlate with age, gender and symptom durations. If those three unidentified metabolites could be finally identified, and all the seven metabolites could be further validated using large cohorts of samples, they would serve as good biomarkers for the early diagnosis of RA.

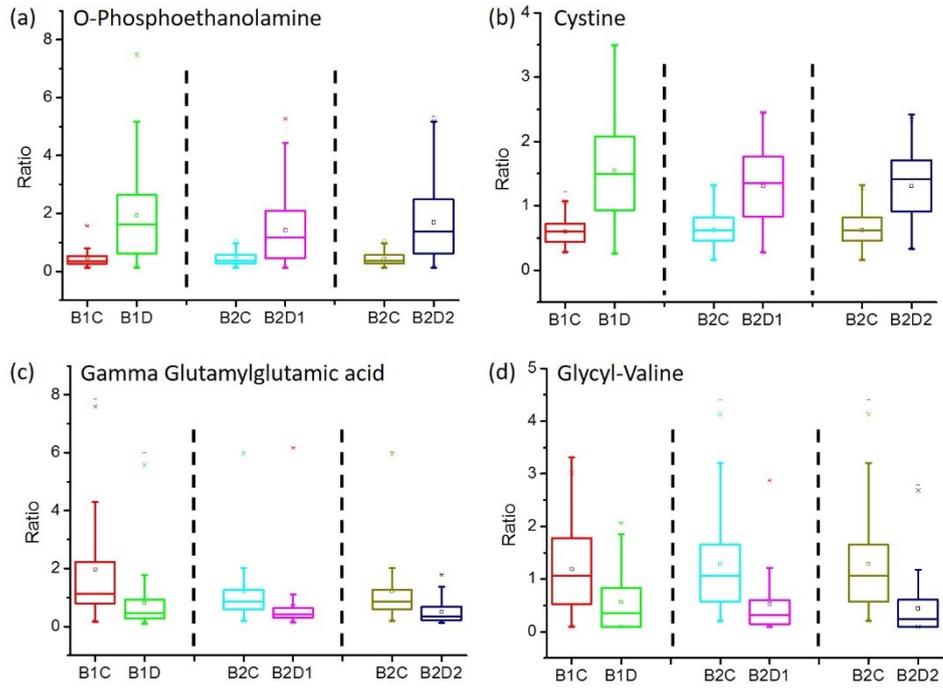


Figure 4.5. Box plots of the relative concentrations of (a) o-phosphoethanolamine, (b) cystine, (c) gamma glutamylglutamic acid, (d) glycyl-valine in the RA and control groups in 2 cohorts.

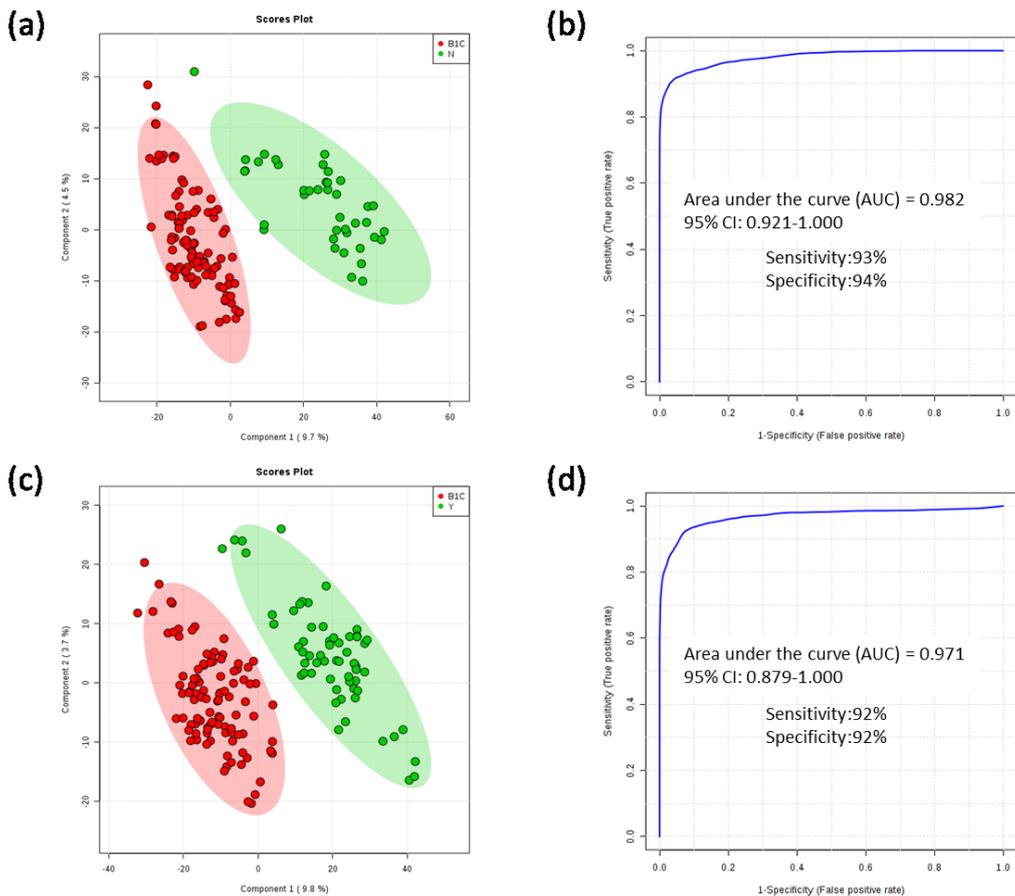


Figure 4.6. PLS-DA scores plot of (a) B1C vs. N ( $R^2=0.988$ ,  $Q^2=0.913$ ) and (c) B1C vs. Y ( $R^2=0.981$ ,  $Q^2=0.915$ ). The ROC curve built for (b) B1C vs. N using 19 identified significant metabolites and for (d) B1C vs. Y using 14 identified significant metabolites. Here, N refers to RA without MTX treatment and Y refers to RA with MTX treatment.

#### 4.3.1.3 Comparative Metabolome Analysis of RA with and without Methotrexate Treatment

Methotrexate (MTX) is an immune system suppressant and it can be used to treat RA,<sup>123</sup> with relatively low costs, especially compared with much more expensive biological treatments. In this work, for cohort A RA patient, 30 of them had MTX treatment (Y), while the remaining 20 patients were not treated with MTX (N). We compared N versus control and Y versus control, and then studied the differences between these two comparisons.

As we can see from Figure 4.6, from PLS-DA plot, we observed a clear separation between B1C and N, and between B1C and Y. And from volcano plot (Appendix Figure A4.4), when comparing B1C versus N 272 down-regulated metabolites and 286 up-regulated metabolites were detected in without MTX treatment RA patients. While for B1C versus Y, 260 down-regulated metabolites and 218 up-regulated metabolites were detected in with MTX treatment RA patients. 19 significant metabolites were positively identified for B1C versus N and 14 metabolites were identified for B1C versus Y. Using these metabolites, respectively, the AUC value of the ROC curve was 0.982 with sensitivity at 93% and specificity at 94% for B1C versus N, and the AUC value of the ROC curve was 0.971 with both sensitivity and specificity at 92% for B1C versus Y. These results suggest that, for both with and without MTX treatment RA samples, the metabolic differences between RA and control were very significant. We also compared the identified significant metabolites and found 12 common metabolites between these two comparisons. These metabolites were listed in Supplemental Table S4.4, with the information of fold changes and p value. Most of these metabolites had a little smaller although similar fold changes in the comparison of with MTX treatment versus control, which means the MTX treatment had little effect on patients' metabolome. The four previously detected diagnostic biomarkers, cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine, for B1C versus B1D were also commonly detected here.

#### 4.3.1.4 Comparative Metabolome Analysis of RA with Anti-CCP Positive and Negative

Anti-CCP antibody assays, with high sensitivity, specificity and reproducibility, are important for the diagnosis of RA.<sup>117</sup> In this work, there were 26 anti-CCP positive patients and 23 anti-CCP negative patients in cohort A. In cohort B, there were 36 anti-CCP positive patients and 14 anti-CCP negative patients. Anti-CCP positive patients usually had more severe symptoms. Here, we compared the metabolic differences between anti-CCP positive patients and healthy controls, and between anti-CCP negative and healthy controls. And the discovered metabolite biomarkers could work together with anti-CCP assays for the diagnosis of RA.

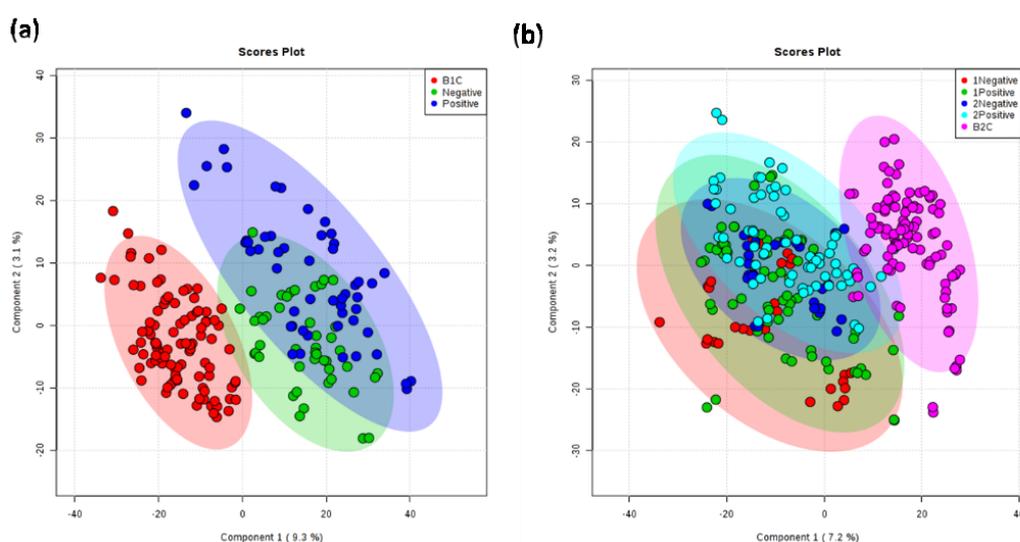


Figure 4.7. PLS-DA scores plot of (a) B1C vs. B1DP (positive) vs. B1DN (negative) ( $R^2=0.964$ ,  $Q^2=0.852$ ) and (b) B2D1P (1positive) vs. B2D1N (1negative) vs. B2D2P (2positive) vs. B2D2N (2negative) ( $R^2=0.914$ ,  $Q^2=0.714$ ).

From Figure 4.7, we can see that, for both 2 cohorts, there was a clear separation on PLS-DA plot between RA and healthy controls, which has also been indicated from previous comparisons. For cohort A samples (Figure 4.7a), we can also see some separations between anti-CCP negative and anti-CCP positive groups. We also identified significant metabolites with fold change  $> 1.5$  or  $< 0.67$  and  $q$ -value  $< 0.05$ . The common identified significant metabolites were listed in Supplemental Table S4.5. We can see that the fold changes of these metabolites were very similar, meaning the good consistency of these metabolites and the metabolic differences between anti-CCP positive and anti-CCP negative were not huge. Similarly, for cohort B samples, we can see some not obvious separations between anti-CCP negative and anti-CCP positive for

both pre- and post-treatment RA samples (Figure 4.7b). In the list of common significant metabolites (Supplemental Table S4.6), we saw similar fold changes. What is worth to mention is that, the four previously detected diagnostic biomarkers, cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine, for RA versus healthy controls were also commonly detected here. These results suggest that the metabolic differences between either positive or negative anti-CCP RA patients and healthy controls can be constantly detected, and the metabolite biomarkers could be used for the diagnosis of RA.

#### **4.3.1.5 Comparative Metabolome Analysis between Responders and Non-Responders**

The Disease Activity Score (DAS) is a clinical assessment of RA disease activity, which combines information from swollen joints, tender joints, the acute phase response and general health.<sup>124</sup> In clinic, to evaluate RA clinical trials, the change in DAS and the level of DAS is used. According to European League Against Rheumatism (EULAR) criteria<sup>125</sup>, after treatment, the RA patients with DAS change  $> 1.2$  and Post DAS  $< 3.2$  are defined as responders (r). Others are defined as non-responders (nr). In this work, for cohort B RA patients, there were 25 responders and 25 non-responders. We compared the metabolic differences between r and nr for both pre- and post-treatment samples, and also the metabolic differences between pre- and post-treatment for both r and nr samples. These comparisons should be useful for drug administration and treatment monitoring, if we could find some metabolic changes.

Figure 4.8 shows the PLS-DA scores plots for the comparison between B2D1 r and B2D1 nr and B2D2 r and B2D2 nr. We can see that B2D1 nr and B2D1 r overlap together badly, while B2D2 nr and B2D2 r displays some separation. Besides, we can see a better separation between B2D1 r and B2D2 r, than between B2D1 nr and B2D2 r. These results indicate that with treatment, the responders had greater metabolic changes.

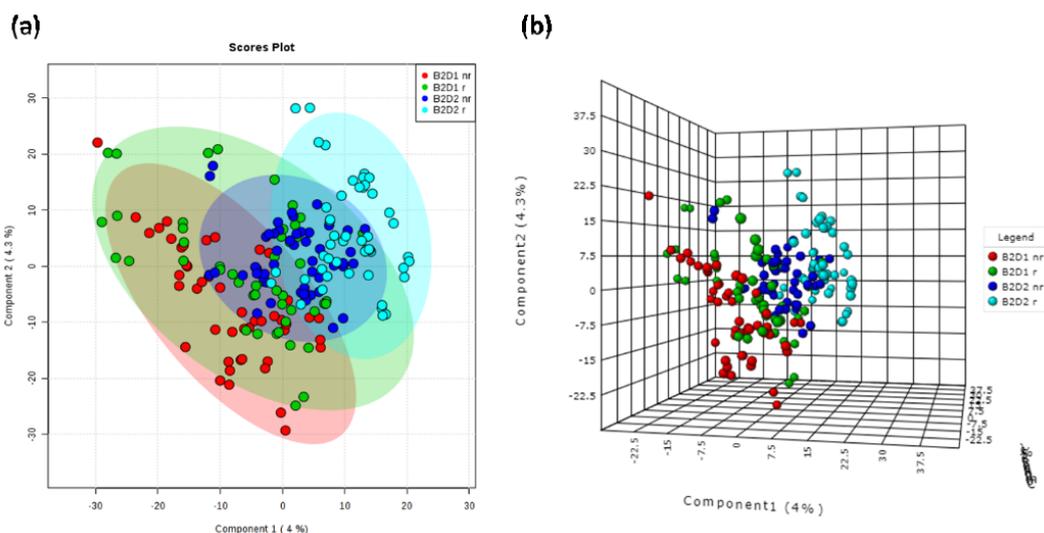


Figure 4.8. (a) 2D and (b) 3D PLS-DA scores plot of B2D1 r vs. B2D1 nr vs. B2D2 r vs. B2D2 nr ( $R^2=0.92$ ,  $Q^2=0.561$ ).

Since the not large number of samples for these comparisons, univariate analysis should be more reliable. Figure 4.9 shows the volcano plots of the four kinds of comparisons. For B2D1, when comparing r versus nr (Figure 4.9a), 3 down-regulated metabolites and 6 up-regulated metabolites were detected in B2D1 r, while comparing B2D2 r versus B2D2 nr (Figure 4.9b), 13 down-regulated metabolites and 3 up-regulated metabolites were detected in B2D2 r. Among these significant metabolites, none can be positively identified for B2D1, while 1 metabolite, phenyl-Leucine, can be identified for B2D2. Since the metabolic differences between r and nr in pre-treatment samples can be especially helpful for drug administration in the beginning of treatment, we studied more about the 9 unidentified metabolites. We found one metabolite, dansyl-11441, had a good univariate AUC value of 0.78. It was also detected to have good power to differentiate RA from controls (AUC of 0.84 for B2C versus B2D1 and AUC of 0.83 for B2C versus B2D2). These results indicate that, the metabolomic differences between responders and non-responders for both before and after treatment RA patients were not so significant. But we can still find some encouraging results. One reason for that was the relatively small sets of samples. In the next step, we will add cohort C, which contains more samples of responders and non-responders. We expect to obtain more results from these comparisons with more samples.

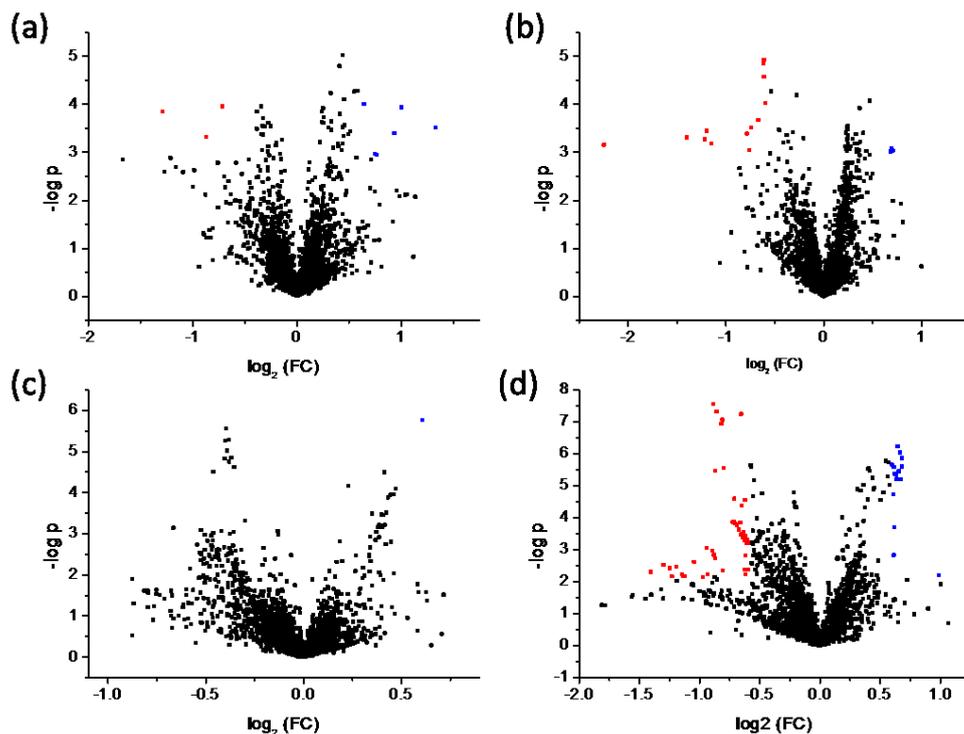


Figure 4.9. Volcano plot of the changes of metabolites between (a) B2D1 r and B2D1 nr, (b) B2D2 r and B2D2 nr, (c) B2D1 nr and B2D2 nr, (d) B2D1 r and B2D2 r. Metabolites with fold change (FC) > 1.5 or FC < 0.67 and q-value < 0.05 were highlighted.

We also compared the metabolomic changes due to treatment for both non-responders and responders. In Figure 4.9c, for non-responders, there was only one up-regulated metabolite detected in post-treatment group and it cannot be identified. In comparison, for responders, 51 down-regulated and 16 up-regulated metabolites were detected in post-treatment group, as we can see from Figure 4.9d. Four of them, theophylline, phenyl-Leucine, phenylalanylphenylalanine, histidinyl-alanine, were positively identified. These results suggest that responders showed significant metabolic differences after treatment, while the metabolome of non-responders remained relatively constant for both pre- and post-treatment. These results are important for the correlation analysis between DAS changes and metabolite concentration changes, from which we can study the disease activity. Still, as mentioned before, we expect to gain more results after the analysis of cohort C samples.

### 4.3.2 Analysis of Carboxyl Submetabolome

#### 4.3.2.1 Metabolite Detection

Compared with dansyl labeling, DmPA labeling has much higher background,

which may be due to the contaminants from experimental plastic vials. As a result, blank subtraction was performed after zero-filling. After that, totally, 2114 metabolites were commonly detected in more than 80% of the 500 samples. Using acid standard library search, we positively identified 21 metabolites. For the remaining metabolites, using MyCompoundID in the HMDB library with accurate mass search and a mass accuracy tolerance of 10 ppm, we putatively identified 333 metabolites. In the EML database with 1 reaction, with accurate mass search and a mass accuracy tolerance of 10 ppm, we putatively identified 1206 metabolites. Thus, for the measured 2114 metabolites, a total of 1560 metabolites can be matched (74%).

#### 4.3.2.2 Comparative Metabolome Analysis for RA Biomarker Discovery

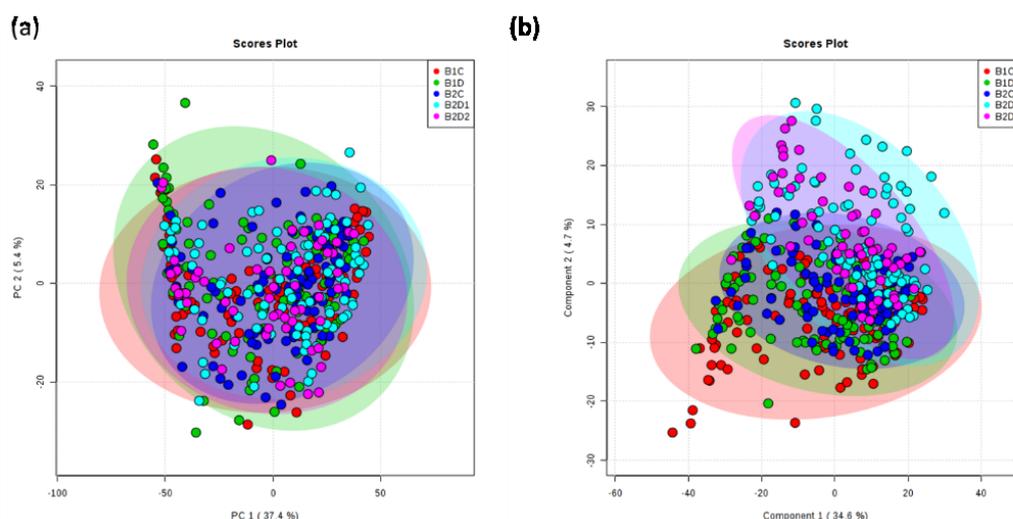


Figure 4.10. (a) PCA and (b) PLS-DA scores plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 ( $R^2=0.773$ ,  $Q^2=0.592$  for PLS-DA plot) for DmPA labeling.

Multivariate analysis was first performed to compare metabolic differences between RA and healthy controls. Appendix Figure A4.5 displays PCA scores plot of the whole data sets with QC samples. The good cluster of QCs indicated the good instrument performance. Figure 4.10 shows PCA and PLS-DA plots of all the 5 groups of samples without QC. From the PCA plot, the 5 groups of samples almost overlap together and the separation between RA and control is not clear. From the PLS-DA plot, we can see a separation between RA and control, and it was observed for both two cohorts. Besides, the two control groups almost overlap together, which was expected and validated the observed metabolomic differences.

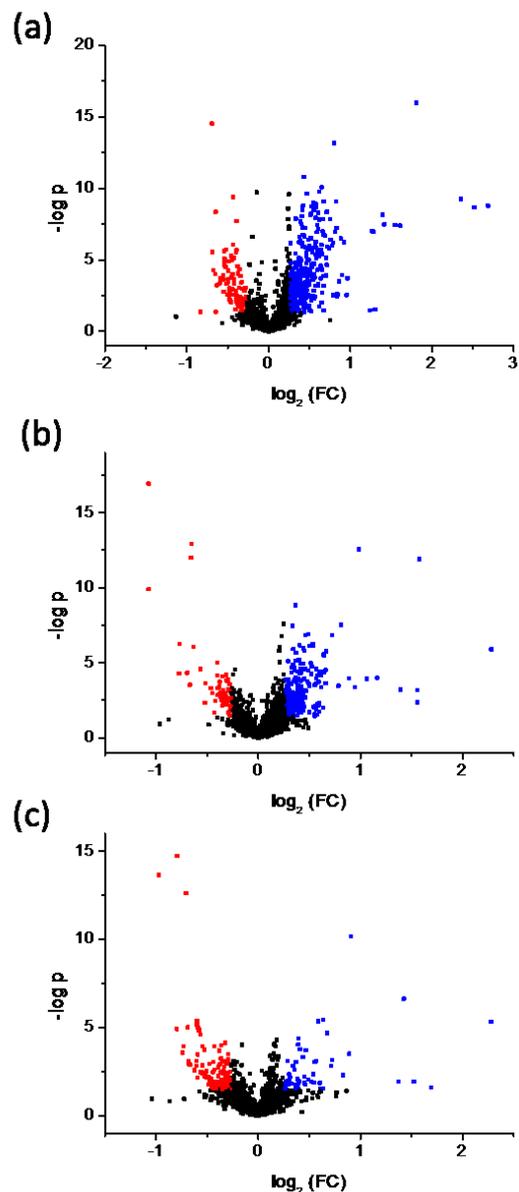
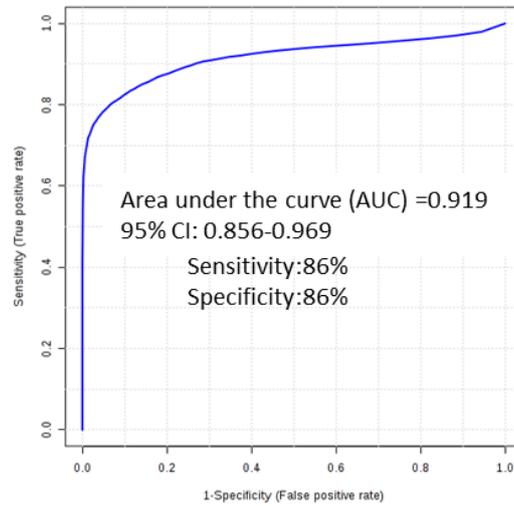


Figure 4.11. Volcano plot of the changes of metabolites between (a) B1C and B1D, (b) B2C and B2D1, (c) B2C and B2D2. Metabolites with fold change (FC) > 1.2 or FC < 0.83 and q-value < 0.1 were highlighted.

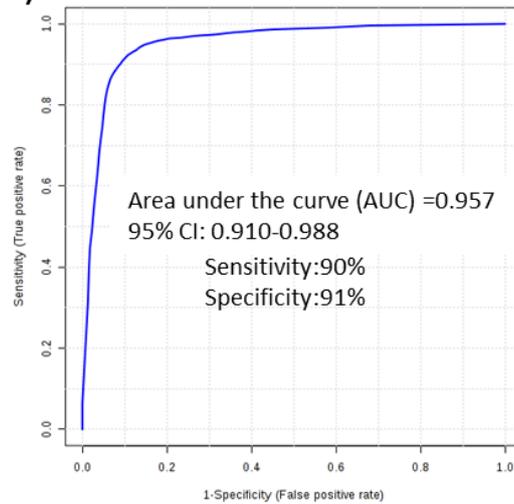
Then volcano plots were used to determine significantly changed metabolites between RA and control groups. Metabolites with an average fold change greater than 1.2 or smaller than 0.83 and q-values smaller than 0.1 were considered as significant. When comparing B1C versus B1D (Figure 4.11a), 96 metabolites were down-regulated, and 388 were up-regulated in B1D. Among these 484 significant metabolites, 9 can be positively identified. When comparing B2C versus B2D1 (Figure 4.11b), 57 down-regulated and 244 up-regulated metabolites were found in B2D1, and 6 can be positively identified. When comparing B2C versus B2D1 (Figure 4.11c), 159 down-

regulated and 53 up-regulated metabolites were found, and 4 of them can be positively identified. We compared these identified significant metabolites for all the 3 kinds of comparisons, 1 metabolite, azelaic acid, were found with similar fold changes.

(a)



(b)



(c)

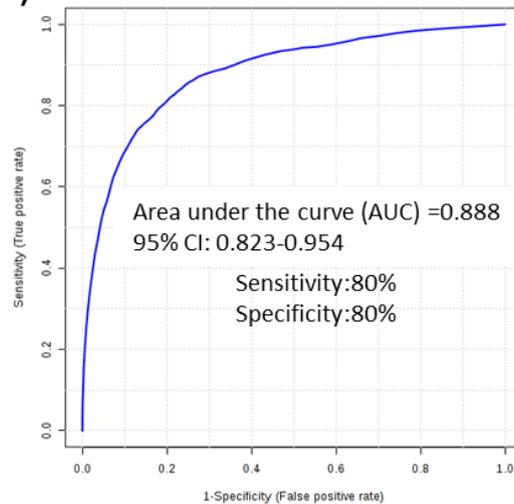


Figure 4.12. ROC curve built using 5 common significant metabolites, azelaic acid, acid-9545, acid-23681, acid-23895, acid-23900, for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2.

ROC curves were built to test the differentiating power of potential biomarker candidates. Using all the significant metabolites, Appendix Figure A4.6 shows a series of ROC curves built with different numbers of variables. When more variables were included for the curve building, an increased discrimination power is observed. With top 15 metabolites, the performance almost reaches maximum with AUC value greater than 0.96 for all the 3 kinds of comparisons. Then we tried to reduce the number of metabolites in the biomarker panel. Previously, we found one common identified significant metabolite, azelaic acid. However, with only this metabolite, the discrimination power of ROC was not high (AUC $\approx$ 0.5). Then we added 4 common unidentified significant metabolites, acid-9545, acid-23681, acid-23895, acid-23900, which were selected from top 30 metabolites ranked by univariate AUC values among all the three kinds comparisons, to form a stronger biomarker panel. From Figure 4.12a, for B1C versus B1D, the AUC value of the ROC curve was 0.919, within the range of 0.856-0.969 at the 95% confidence interval, and the discrimination of RA from control can be achieved at 86% sensitivity and 86% specificity. For B2C versus B2D1 (Figure 4.12b), the AUC value of the ROC curve was 0.957 with sensitivity at 90% and specificity at 91%. For B2C versus B2D2 (Figure 4.12c), the AUC value of the ROC curve was 0.888 with both sensitivity and specificity at 80%. Besides, as we can see from Figure 4.13, the box plots of these 5 metabolites show that, they constantly up- or down-regulated in both cohort A and B, and the differences between RA and control were very significant. These results suggested the biomarker panel with these 5 metabolites had good performance for both cohort 1 and cohort 2 samples.

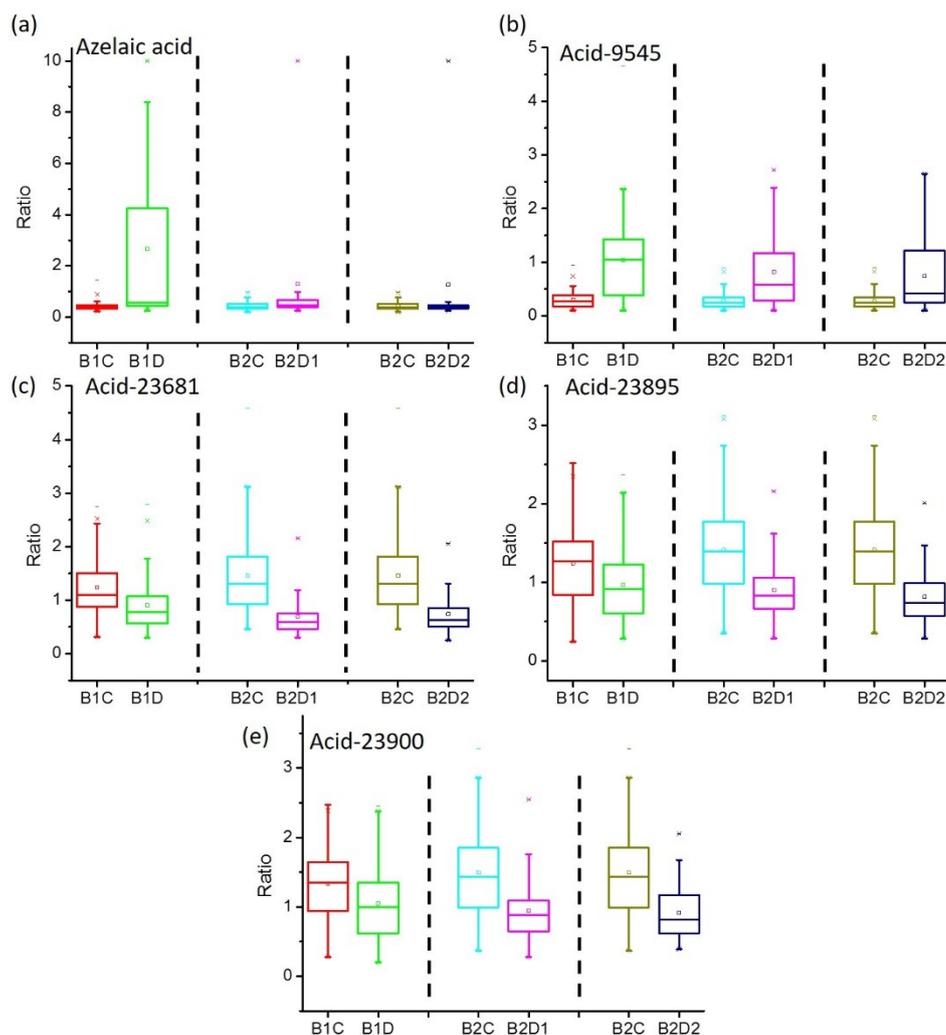


Figure 4.13. Box plots of the relative concentrations of (a) azelaic acid, (b) acid-9545, (c) acid-23681, (d) acid-23895, (e) acid-23900 in the RA and control groups in 2 cohorts.

#### 4.3.2.3 Comparative Metabolome Analysis of RA with and without Methotrexate Treatment

The comparison of RA with MTX treatment versus control and RA without MTX treatment versus control for DmPA labeling data was performed. Any differences were studied. From PLS-DA plot of Figure 4.14a, c, we observed a separation between B1C and MTX N, and between B1C and MTX Y. And from volcano plot (Appendix Figure A4.7), when comparing B1C versus MTX N, 116 down-regulated metabolites and 250 up-regulated metabolites were detected in without MTX treatment RA patients, while for B1C versus Y, 140 down-regulated metabolites and 429 up-regulated metabolites were detected in with MTX treatment RA patients. Six significant metabolites were positively identified for B1C versus N and 13 metabolites were

identified for B1C versus Y. Using these metabolites, respectively, the AUC value of the ROC curve was 0.877 with both sensitivity and specificity at 79% for B1C versus N (Figure 4.14b), and the AUC value of the ROC curve was 0.861 with both sensitivity and specificity at 79% for B1C versus Y (Figure 4.14d). These results suggest that, for both with and without MTX treatment RA samples, the metabolic differences between RA and control were significant. Besides, we also compared the identified significant metabolites and found 4 common metabolites between these two comparisons and these metabolites were listed in Supplemental Table S4.7, with the information of fold changes and p value. We can see that, most of these metabolites had similar fold changes, meaning the effect of MTX treatment was not strong.

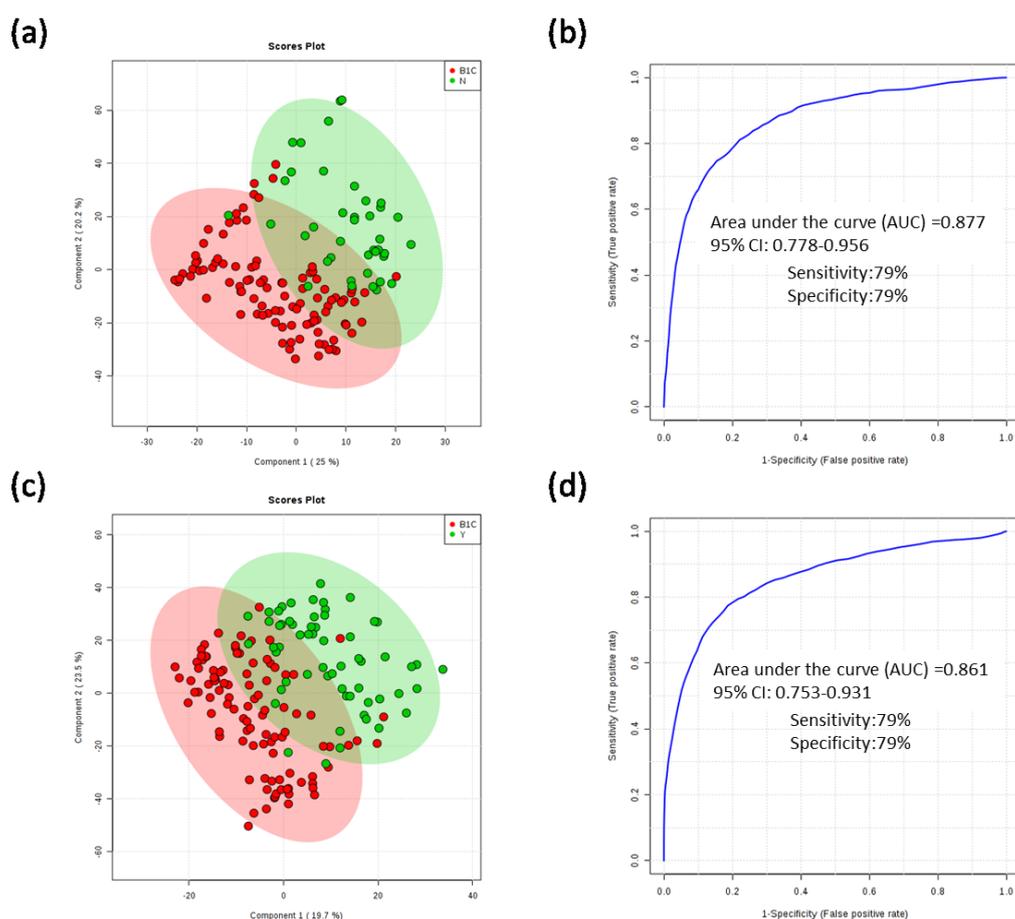


Figure 4.14. PLS-DA scores plot of (a) B1C vs. N ( $R^2=0.943$ ,  $Q^2=0.746$ ) and (c) B1C vs. Y ( $R^2=0.951$ ,  $Q^2=0.801$ ). The ROC curve built for (b) B1C vs. N using 6 identified significant metabolites and for (d) B1C vs. Y using 14 identified significant

#### 4.3.2.4 Comparative Metabolome Analysis of RA with Anti-CCP Positive and Negative

We also compared the metabolic differences between anti-CCP positive RA patients and healthy controls, and between anti-CCP RA negative and healthy controls for DmPA labeling results. From the PLS-DA plot of Figure 4.15, we can see that, for both 2 cohorts, there was a separation between RA and healthy controls. For cohort A samples (Figure 4.15a), we can also see some separations between anti-CCP negative and anti-CCP positive groups. We also identified significant metabolites with fold change  $> 1.2$  and  $q$ -value  $< 0.1$ , and the common identified significant metabolites were listed in Supplemental Table S4.8a. The fold changes of these metabolites were very similar, meaning the good consistency of these metabolites and the metabolic differences between anti-CCP positive and anti-CCP negative were not that significant. Similarly, for cohort B samples, we can see some not obvious separations between anti-CCP negative and anti-CCP positive for both pre- and post-treatment RA samples (Figure 4.15b), and from the list of common significant metabolites (Supplemental Table S4.8b, c), we saw similar fold changes. These results indicate that the metabolic differences between RA patients, either anti-CCP positive or negative, and healthy controls can be constantly detected, which is of great importance for the diagnosis of RA in clinic.

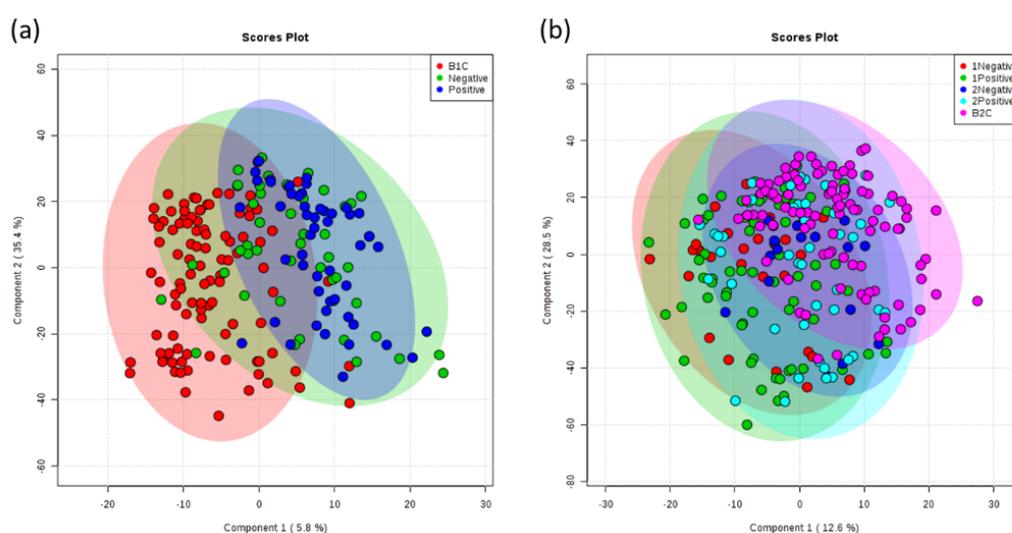


Figure 4.15. PLS-DA scores plot of (a) B1C vs. B1DP (positive) vs. B1DN (negative) ( $R^2=0.881$ ,  $Q^2=0.633$ ) and (b) B2D1P (1positive) vs. B2D1N (1negative) vs. B2D2P (2positive) vs. B2D2N (2negative) ( $R^2=0.817$ ,  $Q^2=0.609$ ).

### 4.3.2.5 Comparative Metabolome Analysis between Responders and Non-responders

Similar as we did for dansyl labeling samples, we compared the metabolic differences between responders and non-responders for both pre- and post- treatment samples, and also the metabolic differences between pre- and post- treatment for both r and nr samples for DmPA labeling. Figure 4.16 shows the PLS-DA score plots for the comparison between B2D1 r and B2D1 nr and B2D2 r and B2D2 nr. We can see that both B2D1 nr and B2D1 r, and B2D2 nr and B2D2 r overlap together badly, meaning small metabolic differences between responders and non-responders in either pre- and post-treatment. Besides, we observed a better separation between B2D1 r and B2D2 r, and B2D1 nr and B2D2 nr. The results indicate that treatment caused some metabolic changes. We then used univariate analysis to discover significant metabolites for all the 4 kinds of comparisons. However, we can only determine significant metabolites for B2D1 nr vs. B2D2 nr, in which there were 173 down-regulated and 1 up-regulated metabolites in Appendix Figure 4.8a. Using these metabolites, the best performance of ROC curve only reached an AUC value of 0.66 in Appendix Figure 4.8b, which means the discriminating power of these metabolites is not high. These results indicate that there were no significant metabolomic differences between responders and non-responders for both before and after treatment RA patients for carboxy submetabolome.

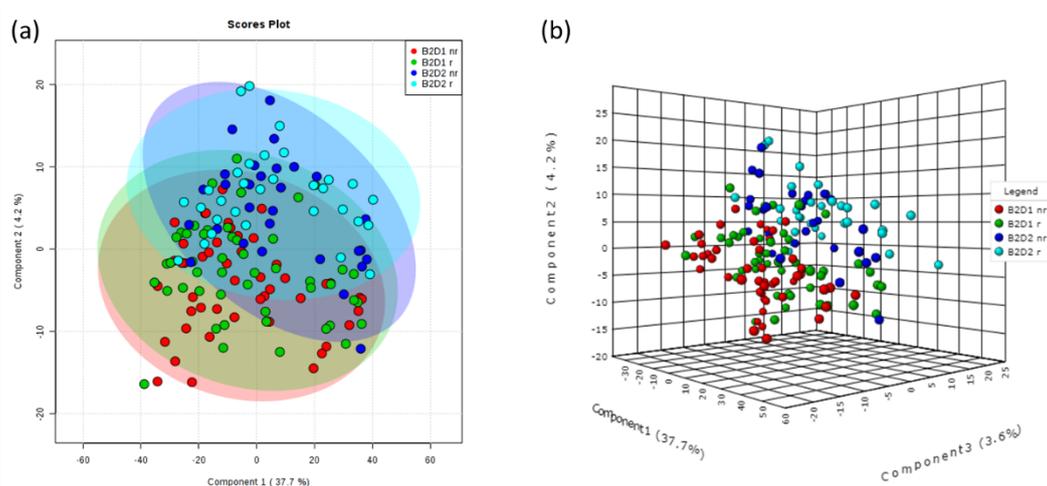


Figure 4.16. (a) 2D and (b) 3D PLS-DA scores plot of B2D1 r vs. B2D1 nr vs. B2D2 r vs. B2D2 nr ( $R^2=0.834$ ,  $Q^2=0.266$ ).

## 4.4 Conclusion

In this work, CIL LC-MS method was applied to profile amine/phenol and

carboxyl submetabolomes in serums of rheumatoid arthritis (RA) disease. We observed significant metabolic differences between RA patients and healthy controls for both two cohorts of samples and both two submetabolomes. For dansyl labeling, four common identified significant metabolites were discovered, including cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine. And for DmPA labeling, five common significant metabolites, azelaic acid, acid-9545, acid-23681, acid-23895, acid-23900 were discovered. These metabolites can be potential biomarkers if further validated. We also found several metabolites can consistently differentiate control from both with and without MTX RA samples. Besides, we found metabolic differences between RA patients, either anti-CCP positive or negative, and healthy controls can be constantly detected, which was good for the diagnosis. The four discovered diagnostic biomarkers also showed up constantly in both MTX and anti-CCP analysis, indicating the good consistency of these biomarkers. Due to the small sample set, the comparison results were not so significant. In the future, we will add cohort C, which have more samples and can help validate the discovered potential biomarkers as well as study more about the differences between responder and non-responders. We will also do the correlation analysis between DAS changes, CRP changes and metabolite concentration changes to understand more about the disease activity. We will also try to profile other submetabolomes, such as hydroxyls<sup>43</sup>, ketones and aldehydes<sup>45</sup>, to reach higher metabolome coverage.

## Chapter 5 Development of Chemical Isotope Labeling LC-MS for Metabolite Biomarker Discovery of Alzheimer's Disease in a Mouse Model

### 5.1 Introduction

Alzheimer's disease (AD) is a chronic neurodegenerative disease and the leading cause of dementia. Currently, 5.7 million Americans are living with AD dementia, and the number of patients is expected to grow to more than 100 million worldwide by 2050.<sup>126</sup> The clinical features of AD include progressive memory and cognitive impairment, problems with language and visuospatial function, and increasing frailty.<sup>127</sup> Although the detailed mechanisms of AD remain to be solved, two histological features, neurofibrillary tangle and neuritic plaques, are hallmarks. Neurofibrillary tangle is a kind of intracellular filamentous lesion, composed of over-phosphorylated microtubule-associated tau protein.<sup>128</sup> Neuritic plaques contain deposits of amyloid  $\beta$ -peptides ( $A\beta$ ), which are the cleavage products of amyloid precursor protein (APP).<sup>129</sup> The diagnosis of AD in early-stage, mild cognitive impairment or even pre-clinical, is difficult. Currently, there is no pharmacologic treatment available for AD. Many failures in drug development have happened, with the majority focusing on  $A\beta$ . Many researchers believe that the failed clinical trials may be due to the too late administration of treatment.<sup>130</sup> Medication at the early stage might offer better chance of slowing or even stopping the progression of AD.<sup>131</sup> Thus, the discovery of highly sensitive biomarkers is important, and effective biomarkers can be useful for both the early diagnosis and monitoring treatment of AD.

Traditional biomarker discovery studies are focused on transcriptomic or proteomic assays.<sup>132-133</sup> Metabolomics is a powerful tool to characterize the complex biochemical phenotypes influenced by both environment and genetic factors.<sup>8</sup> It is an emerging field for the biomarker discovery of AD. Many studies of AD have been reported using different kinds of analytical techniques, such as NMR<sup>10</sup>, GC/MS<sup>134</sup>, direct infusion electrospray mass spectrometry (DI-ESI-MS)<sup>135</sup> and LC-MS<sup>136</sup>. These studies implied that the metabolomic changes can be associated with AD and metabolomics is able to provide promising biomarkers. However, current techniques can only achieve very limited metabolite coverage, and new technologies are required to provide higher coverage and better quantitative capability.

Previously, our group reported a chemical isotope labeling (CIL) LC-MS method to achieve accurate quantification and high metabolome coverage based on  $^{13}\text{C}$ -/ $^{12}\text{C}$ -isotope dansylation labeling.<sup>23</sup> This derivatization method can greatly enhance the ESI signal response, reversed-phase (RP) LC separation performance, as well as measurement precision. It has been well applied to different kinds of biological samples, including serum<sup>77</sup>, CSF<sup>78</sup>, saliva<sup>80</sup>, sweat<sup>81</sup> and cells<sup>82</sup>.

The transgenic mice model is the most widely used animal model for AD studies. Previously we reported one urine metabolomics study using TgCRND8 transgenic mice.<sup>79</sup> The urine samples were collected at three different ages, and a clear metabolomic trajectory change was observed. Based on the promising findings of this work, we wanted to do further in-depth investigations of metabolic changes due to AD using this mouse model, not only just urine metabolomics, but serum metabolomics, with larger numbers of samples. Here, we report the use of  $^{13}\text{C}$ -/ $^{12}\text{C}$ -dansylation labeling coupled with LC-MS method for parallel profiling of serum and urine metabolome changes. The metabolic differences between AD and control were observed for both male and female mouse urine samples. The trajectory changes with the development of disease were also studied. The discovered metabolite biomarkers have good discriminating power and they could be used to guide the selection of human metabolite biomarkers for early diagnosis of AD.

## **5.2 Experimental Section**

### **5.2.1 Sample Collection and Processing**

In this experiment, the TgCRND8<sup>137</sup> transgenic (Tg) mouse model of early-onset AD with A $\beta$  amyloid deposition was used, as well as non-Tg wild type (WT) littermate controls. Tg mouse model expresses a double mutant form of amyloid precursor protein (APP) 695 (KM670/671NLV717F) isoform, and can develop amyloid deposits as early as 2–3 months, which can be used for the discovery of potential biomarkers at the early stage of AD.

There were 46 mice, with 24 Tg and 22 WT, as well as 23 females and 23 males, as Table 5.1 shows. The urine and serum samples were collected from the age of 8-9 weeks to the age of 25-26 weeks biweekly to represent different stages of disease development. In total, there were 8 collection time points. Here, I only focused on the

analysis of urine samples, while another group member will analyze serum samples. For the collection of urine samples, the mouse was lifted one-by-one from the home cage and placed into a clean and new disposable plastic cage. The mouse was then left inside the cage to urinate spontaneously. After that, urine was immediately pipetted from the floor of the cage into a 1.5 mL Eppendorf tube and snap-frozen on dry ice.

Table 5.1. Sample information about mouse urine and serum samples.

<b>Ages</b>	<b>Total No. of urine (or serum) samples</b>	<b>No. of Tg samples</b>	<b>No. of WT samples</b>	<b>No. of M samples</b>	<b>No. of F samples</b>
Week 8-9	46	24	22	23	23
Week 10-11	46	24	22	23	23
Week 12-13	43	21	22	21	22
Week 14-15	41	19	22	19	22
Week 16-17	38	18	20	16	22
Week 18-19	38	18	20	16	22
Week 20-21	38	18	20	16	22
Week 25-26	36	16	20	15	21

### 5.2.2 Dansyl Labeling

Samples were diluted first. Briefly, 20  $\mu\text{L}$  of supernatant of individual mouse urine sample was diluted to 80  $\mu\text{L}$  by adding 60  $\mu\text{L}$  of water. A pooled sample was prepared by aliquoting 15  $\mu\text{L}$  of each of the diluted individual mouse urine samples and mixing them. The labeling method was adapted from a protocol in a previous report.<sup>23</sup> 25  $\mu\text{L}$  of individual mouse urine sample (experimental duplicates) or 25  $\mu\text{L}$  of the pooled sample was mixed with 12.5  $\mu\text{L}$  of ACN. Then, 12.5  $\mu\text{L}$  of 250 mM sodium carbonate/sodium bicarbonate buffer was added to the samples. The solution was mixed with 25  $\mu\text{L}$  of freshly prepared  $^{12}\text{C}$ -dansyl chloride (DnsCl) solution (18 mg/mL, for light labeling, individual samples) or  $^{13}\text{C}$ -DnsCl solution (18 mg/mL, for heavy labeling, pooled sample). After incubation for 45 min at 40  $^{\circ}\text{C}$ , 5  $\mu\text{L}$  of 250 mM sodium hydroxide solution was added to the reaction mixture. The solution was then incubated at 40  $^{\circ}\text{C}$  for another 10 min. Finally, 25  $\mu\text{L}$  of formic acid (425 mM) in 1:1 ACN/ $\text{H}_2\text{O}$  was added.

### 5.2.3 LC-UV Quantification

The total dansyl-labeled metabolite amount of each sample was normalized

using a protocol previously reported,<sup>86</sup> based on LC–UV. The instrument for detection was a Waters ACQUITY UPLC system with photodiode array (PDA) detector. A Phenomenex Kinetex reversed-phase C18 column (50 mm × 2.1 mm, 1.7 μm particle size, 100 Å pore size) was used to achieve a fast step-gradient. The flow rate was 0.45 mL/min and the total run time was 6 min. According to the quantification results, <sup>12</sup>C-labeled individual sample and the <sup>13</sup>C-labeled pooled sample were mixed in equal amounts. Besides, <sup>12</sup>C- and <sup>13</sup>C-labeled pooled samples were mixed in equal amounts serving as a quality control (QC) sample.

#### 5.2.4 LC-FTICR-MS Analysis

Samples were analyzed using a Bruker 9.4 T Apex-Qe FTICR mass spectrometer (Bruker, Billerica, MA) with electrospray ionization (ESI), coupled with an Agilent capillary 1100 binary system (Agilent, Palo Alto, CA). An Agilent eclipse plus reversed-phase C18 column (100 × 2.1 mm, 1.8 μm particle size,) was used. All the samples were injected in random order. QC samples and amino acid standards were injected every 10 runs to monitor the performance of the LC-MS.

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

After LC-FTICR-MS analysis, the entire centroid peak lists, with information (e.g. retention time, m/z, and peak intensity), were exported from Bruker Data Analysis software. IsoMS<sup>87</sup> was used to pick peak pairs, reduce false-positive pairs such as dimers and common adducts, calculate peak-pair intensity ratios, and group the peak pairs found in adjacent spectra. After the alignment of same peak pairs from different runs using Alignment program, the Zerofill program was applied to fill in missing values. Peak pairs were finally reconstructed and the chromatographic peak ratios were determined using IsoMS-Quant program.

Based on accurate mass and retention time matches, positive metabolite identification was performed using dansyl standard library search, which contains 273 unique dansylated amines/phenols. Based on accurate mass search, putative identification was performed by using the Human Metabolome Database (HMDB) ([www.hmdb.ca](http://www.hmdb.ca)) and the Evidence-based Metabolome Library (EML) (<http://www.mycompoundid.org/>).

Principal components analysis (PCA), partial least squares discriminant

analysis (PLS-DA) and receiver operating characteristic curves were carried out using MetaboAnalyst software 3.0 (McGill University, Montreal, Canada). The fold change and p-value between groups were calculated using Microsoft Excel. Volcano plots were constructed using OriginPro 8.0 (OriginLab). The q-value, multiple-testing-corrected p-value, was calculated using R and BioConductor ([www.bioconductor.org](http://www.bioconductor.org)).

## **5.3 Results and Discussion**

### **5.3.1 Metabolite Detection**

The injection amount of labeled metabolites onto the LC–MS can affect the number of peak pairs detected in a sample. To determine the best injection amount of urine samples, we injected increasing amount of mixture of <sup>12</sup>C- and <sup>13</sup>C-labeled pooled sample for the optimization. In Appendix Figure A5.1, it was found that with the injection amount of 11.8 nmol, the number of peak pairs reached maximum. Thus, for each run, all individual urine and QC samples were analyzed using LC–MS with 20 nmol injection.

Experimental duplicates were applied for all the urine samples. In total, there were 600 sample runs analyzed by LC-MS. From the 600 sample injections, a total of 25188 peak pairs were detected with an average of  $10086 \pm 2278$  peak pairs for each sample. It means that our method could achieve high submetabolome coverage. Among those, 3530 peak pairs were detected in more than 80% of the samples. These were subjected to multivariate and univariate statistical analysis.

### **5.3.2 Submetabolome Comparison of Tg and WT Mice**

Principal component analysis (PCA) was performed to gain a general idea of the whole data sets. Appendix Figure A5.2a shows the PCA plot of all the urine samples. The QC samples cluster together closely, indicating that the instrument performance was good and the quality of the data should be satisfactory. We observed a separation between Tg and WT mice samples. However, some kind of in-group separations were more obvious, and it was inferred due to the gender differences. It was confirmed by labeling the samples as “female” and “male” (Appendix Figure A5.2b), which has also been reported previously.<sup>79</sup> As a result, we studied female and male mouse samples

separately.

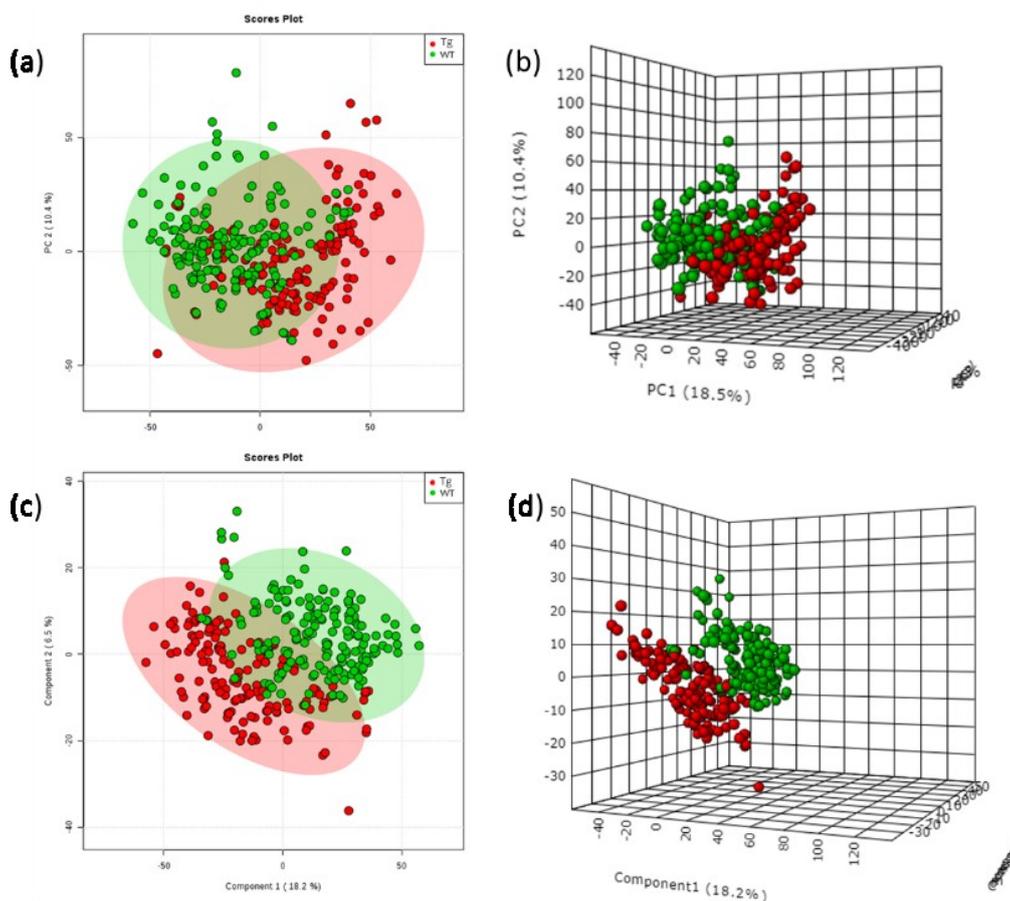


Figure 5.1. 2D and 3D PCA plots (a, b) and PLS-DA plots (c, d) of female mouse urine samples.

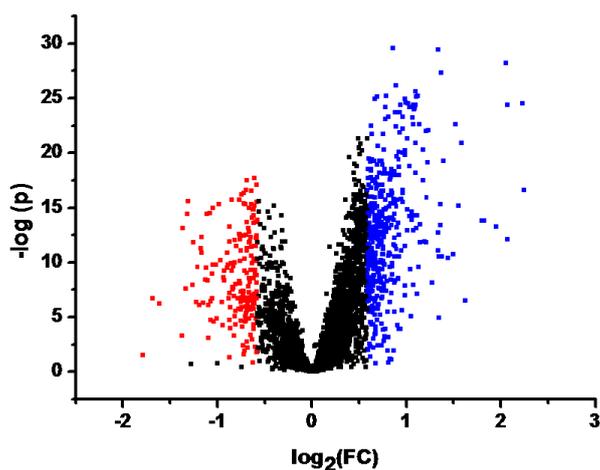


Figure 5.2. Volcano plot of metabolites with a fold change  $> 1.5$  are marked in blue, and those with fold change  $< 0.67$  are marked in red, both with a q-value  $< 0.05$  (corresponding to a p-value  $< 0.181$ ).

For the female mouse urine samples, Figure 5.1 shows 2D and 3D PCA plots

and PLS-DA plots, generated from metabolome data set with the metabolites showing in more than 80% of female mouse urine samples. From PCA plots, there is a separation between Tg and WT groups, and the separation is more obvious in 3D plot. It means that there were significant submetabolome differences between two groups due to the disease. In PLS-DA plots, a more significant separation between the two was observed, and they were completely separated in 3D plot. Besides, the PLS-DA plot had an excellent R<sup>2</sup> value (goodness of fit) and Q<sup>2</sup> value (goodness of predictability) of 0.903 and 0.781, respectively, which validated the visual separation between the two. And the model was further validated by passing the permutation test. The univariate analysis was then applied to determine significantly changed metabolites, which contributed to the separation between Tg and WT groups. In Figure 5.2, a volcano plot displays significant metabolites with a fold change (FC) of > 1.5 or < 0.67 and a q-value < 0.05 (meaning false discovery rate was < 5% and the corresponding p-value was < 0.181). In total, there were 434 metabolites with an increased concentration (up-regulated) and 188 metabolites with a decreased concentration (down-regulated) in Tg compared to WT groups.

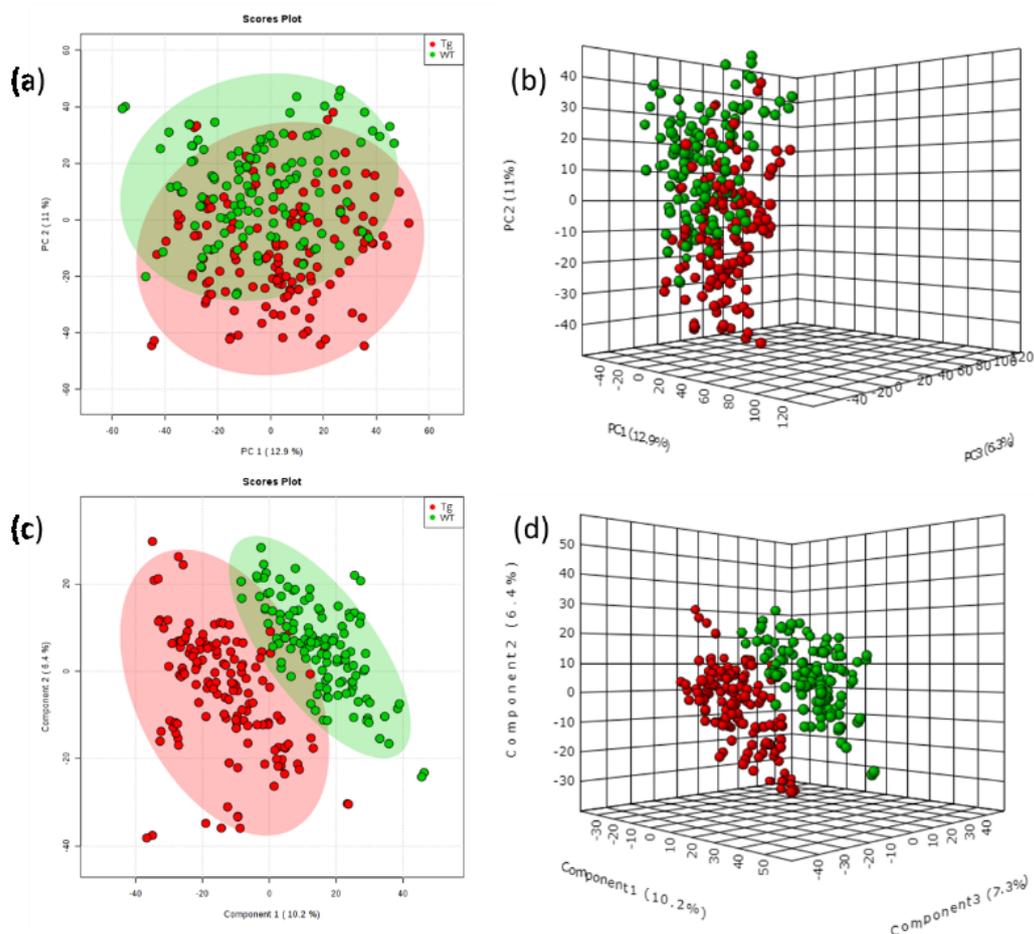


Figure 5.3. 2D and 3D PCA plots (a, b) and PLS-DA plots (c, d) of male mouse urine samples.

For the male mouse urine samples, a similar result was observed from the 2D and 3D PCA plots and PLS-DA plots in Figure 5.3. A separation, although with lots of overlapping data points, was seen in the PCA plot, and a better separation was seen from the PLS-DA plot between Tg and WT groups. The model was validated by high  $R^2$  and  $Q^2$  values of 0.915 and 0.822, respectively, and permutation test. Besides, in Figure 5.4, a volcano plot shows that metabolites with a fold change (FC) of  $> 1.5$  or  $< 0.67$  and a  $q$ -value  $< 0.05$  (the corresponding  $p$ -value of  $< 0.114$ ) were highlighted. In total, there were 168 up-regulated metabolites and 134 down-regulated metabolites.

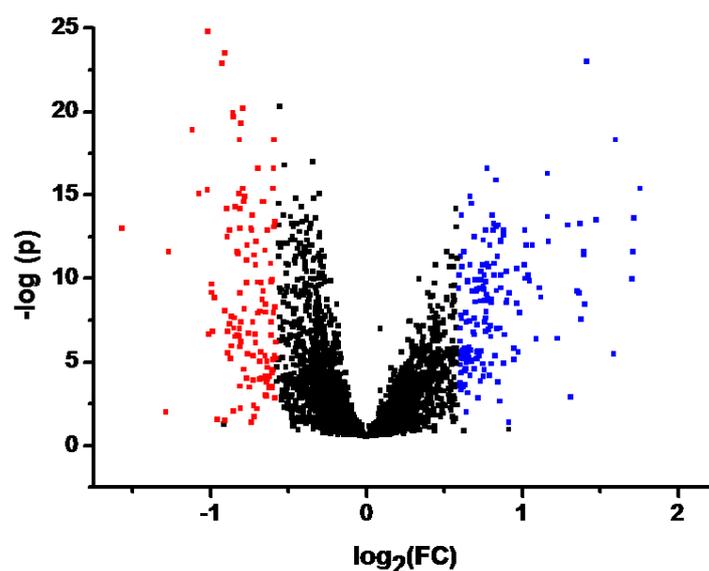


Figure 5.4. Volcano plot of metabolites with a fold change > 1.5 are marked in blue, and those with fold change < 0.67 are marked in red, both with a q-value < 0.05 (corresponding to a p-value < 0.114).

Table 5.2. List of common identified significant metabolites.

HMDB.No.	Name	Female		Male	
		P value	Fold change	P value	Fold change
HMDB00099	Cystathionine	1.77E-12	2.32	1.34E-08	1.70
HMDB00118	Homovanillic acid	6.19E-15	2.05	6.48E-13	1.79
HMDB00696	Methionine	1.60E-14	3.54	1.88E-07	2.13
HMDB01257	Spermidine	2.86E-08	1.67	1.17E-06	1.56
HMDB28689	Alanyl-Histidine	5.73E-29	4.17	1.38E-06	1.64
HMDB01370	Diaminopimelic acid	5.94E-07	0.33	3.40E-10	0.51

For the female mouse urine samples, among 622 significant metabolites, 19 metabolites were positively identified based on dansyl standard library search. And for the male mouse urine samples, among 312 significant metabolites, eight of them can be positively identified. Among those positively identified metabolites, six of them were commonly shared by both female and male mouse urine samples. Table 5.2 shows the common significant metabolites. What is interesting to mention is that, these metabolites displayed greater changes for female samples than those for male samples, which may infer that female mice were affected more intensely by AD than male mice.

It can also be implied from the greater number of significant metabolites for female samples (622) than for male samples (312).

### 5.3.3 Receiver Operating Characteristic Curves

Using these six common metabolites, receiver operating characteristic (ROC) curve was generated to test the diagnosis power of these metabolite biomarker candidates. The classification model was built using the random forest method. As shown in Figure 5.5a, for female mouse urine samples, the area-under-the-curve (AUC) value of ROC curve was 0.892, within the range of 0.836-0.938 at the 95% confidence interval. Discrimination of AD from WT can be achieved at 80% sensitivity and 80% specificity. For male mouse urine samples, displayed in Figure 5.5b, AUC value of ROC curve was 0.881, within the range of 0.826 - 0.922 at the 95% confidence interval, and discrimination of AD from WT can be achieved with both sensitivity and specificity at 81%. The six common significant metabolites with high differentiation power can be biomarker candidates.

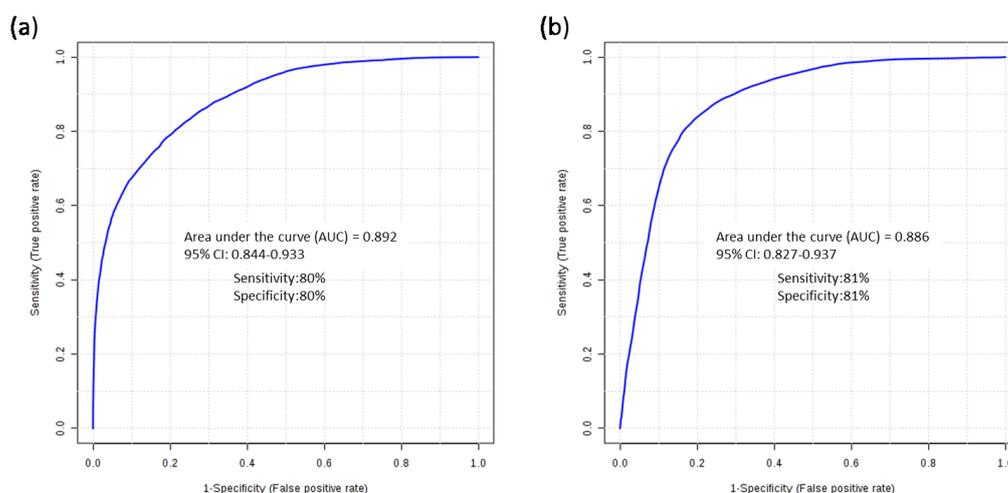


Figure 5.5. The receiver operating characteristic curve generated based on six common significant metabolites for (a) female and (b) male mouse urine samples.

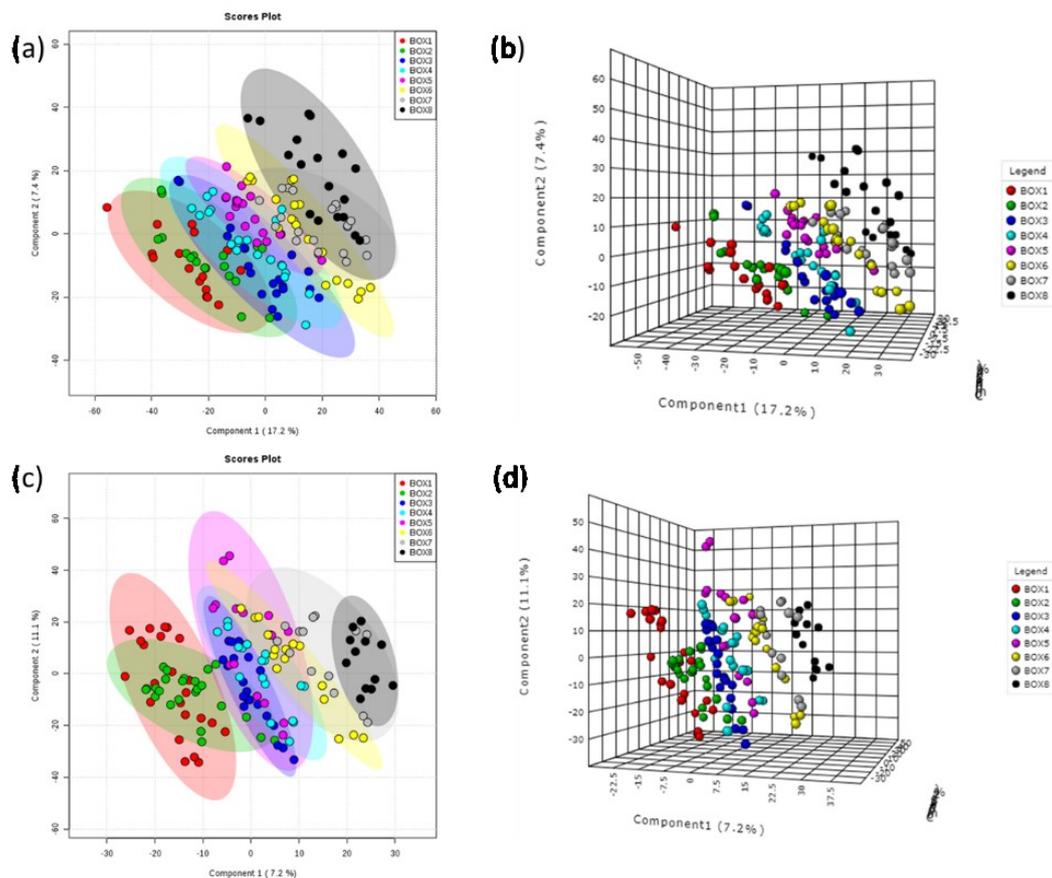


Figure 5.6. 2D (a, c) and 3D (b,d) PLS-DA plots displaying the metabolic trajectory of eight collection time points among the female (a, b) and male (c, d) Tg mice. Box1 to Box8 represent 8 time points.

### 5.3.4 Time Series Analysis

We were also interested in the metabolic trajectory changes with the development of the disease. Using PLS-DA plots, Figure 5.6 shows the metabolic trajectory of eight collection time points from female (Figure 5.6a, b) and male (Figure 5.6c, d) Tg mice. For both female and male mice, from 2D PLS-DA plots (Figure 5.6a, c), we can see a clear progression of eight groups from left-bottom side to the right-top side, which represents a metabolomic trajectory changing from the age of week 8-9, 10-11 ... to week 25-26. The changes were seen more obviously from three-dimensional (3D) PLS-DA plots (Figure 5.6b, d).

### 5.3.5 Significance of Potential Metabolite Biomarkers

Six positively identified significant metabolites, methionine, cystathionine, homovanillic acid, spermidine, alanyl-histidine, diaminopimelic acid, were chosen as the biomarker candidates. Among them, diaminopimelic acid is down-regulated and the

other five metabolites are up-regulated in the Tg group. Taking methionine as an example, Figure 5.7 shows its relative concentrations in the Tg and WT groups at eight different time points. We can see that in the WT group, the concentrations are lower but more constant than those in the Tg group. Besides, the concentrations fluctuated more intensely in the Tg group, which may be due to the effect of the disease development. These metabolites provide insights into disease pathogenesis. The biological significance of these 6 common significant metabolites is briefly discussed below.

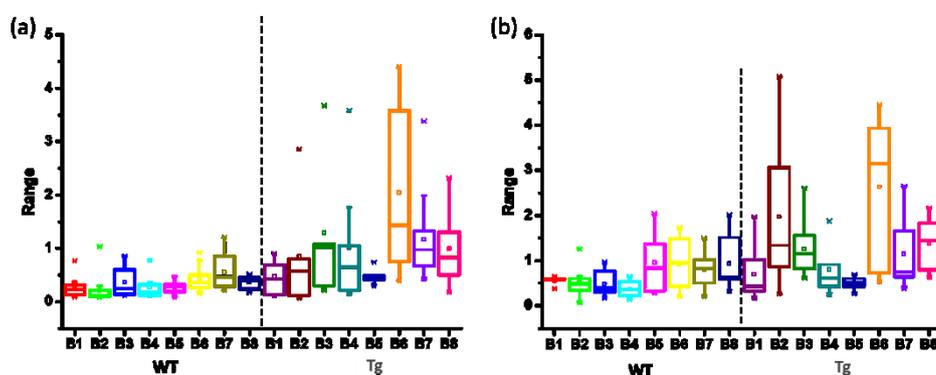


Figure 5.7. Box plots of the relative concentrations of methionine in the Tg and WT groups at eight different time points for female (a) and male (b) mice. B1 to B8 represented 8 time points.

Methionine is an important sulfur-containing amino acid that plays essential roles in cell physiology and can serve as an antioxidant.<sup>138</sup> High concentration of methionine in blood can lead to brain alterations and memory impairment, and diet high in methionine may contribute to neurodegeneration, causing greater risk of AD in a mouse model.<sup>139</sup> Methionine was also detected as an up-regulated metabolites in our previous experiment.<sup>79</sup> Besides, the increased concentration of methionine was also confirmed with human CSF samples.<sup>134</sup> These results reveal that the up-regulated concentration of methionine may have close relations to AD.

Cystathionine is a dipeptide formed by serine and homocysteine. The transsulfuration of methionine produces homocysteine, and it can couple with serine to form cystathionine. It means the pathways of cystathionine and methionine have close relations, which may also have close relations with AD. Besides, cystathionine beta-synthase was reported to be associated with AD, although the relationship between the two remains to be resolved.<sup>140</sup>

The other four metabolites have also been reported to be associated with AD. Homovanillic acid is a dopamine metabolite and the depletion of dopamine is related with neurodegenerative condition.<sup>98</sup> Spermidine is a polyamine, playing critical roles in nerve growth and regenerations. Increased levels of spermidine in AD patients has been reported.<sup>141</sup> Alanyl-histidine is an endogenous antioxidant and antiglycating agent, and it is found to have functions of rescuing cognitive deficits.<sup>142</sup> Diaminopimelic acid is an amino acid derivative related to the cognitive decline.<sup>143</sup>

#### **5.4 Conclusion**

In this work, a high-performance chemical isotope labeling method coupled with LC-MS was developed and applied to mouse urine samples to profile amine/phenol submetabolome. We observed significant metabolomic differences between control and AD groups for both female and male mice. Besides, a clear metabolomic trajectory change from the age of week 8-9, 10-11 ... to week 25-26 was also observed, which indicated that the disease development can be monitored. Many significant metabolites were found and identified. Of those, 6 metabolites were commonly detected in both female and male mice with similar fold changes. The biological significances of these 6 metabolite biomarker candidates were briefly investigated and they can serve as a guidance for human biomarker discovery if further validated. In the future, we will also compare the results from serum samples to get more comprehensive results.

## Chapter 6 Conclusions

With high-performance chemical isotope labeling (CIL) LC-MS method, metabolomics has been investigated for different kinds of samples and different diseases. In Chapter 2, a workflow for carrying out tissue sample processing and high-coverage metabolome profiling was developed. This workflow involves a solvent system of methanol/DCM/water to process a tissue sample, high-performance differential CIL of tissue extracts, and high-resolution LC-MS for labeled metabolite detection. To demonstrate the potential applications of this workflow to tissue metabolomics, we examined amine/phenol submetabolome differences between the liver and brain tissues of mouse model of Alzheimer's disease. A total of 2319 metabolites were commonly detected in more than 80% of the liver samples, and 1769 metabolites were commonly detected in brain samples. Significant metabolomic differences between transgenic mice and wild type mice were observed. Several metabolite biomarker candidates have been found with good discriminating power. These results suggest the good performance of that CIL LC-MS workflow for tissue metabolomics with high coverage and good quantification capability.

In Chapter 3, the CIL LC-MS method was applied for evaluating the amine/phenol submetabolome changes induced by Dex treatment. Five kinds of rat tissue samples, including brain, liver, heart, kidney and muscle, were analyzed. The metabolic differences between Dex-treated and control groups were observed and significantly changed metabolites were determined and some of them were identified. Using all the positively identified metabolites, pathway analysis was carried out for all five kinds of tissues, respectively, and commonly affected pathways were investigated. Of those, arginine and proline metabolism, was taken for in-depth discussion. The biological significance of involved identified metabolites, including putrescine, spermidine, ornithine, proline, were discussed. These results suggest Dex treatment can profoundly induce metabolic changes in many pathways.

In Chapter 4, we applied the CIL LC-MS method to analyze amine/phenol and carboxyl submetabolomes of 250 serum samples from two cohorts of rheumatoid arthritis (RA) patients. In this work, we firstly characterized the submetabolome changes between the early RA and healthy control groups and discovered 4 biomarker candidates. Then, we compared the metabolic differences between RA patients with

methotrexate (MTX) treatment and without MTX treatment and healthy controls, and between RA patients with anti-CCP antibody positive and with anti-CCP antibody negative and controls. We also compared the metabolite changes between responders and non-responders, which may help the administration of medication in the clinic. These findings indicate that RA patients can be well differentiated from healthy controls and the discovered biomarkers can be helpful for both diagnosis and medication administration if further validated.

In Chapter 5,  $^{13}\text{C}$ -/ $^{12}\text{C}$ -dansylation labeling coupled with LC-MS method was used for parallel profiling of serum and urine metabolome changes. 3530 peak pairs were detected in more than 80% of the samples. The metabolic differences between AD and control were observed for both male and female mouse urine samples. The metabolic trajectory changes with the development of disease were also observed. The discovered 6 metabolite biomarkers, methionine, cystathionine, homovanillic acid, spermidine, alanyl-histidine, diaminopimelic acid, have good discriminating power and they could be used to guide the selection of human metabolite biomarkers for early diagnosis of AD.

## Bibliography

1. Hasin, Y.; Seldin, M.; Lusis, A., Multi-omics approaches to disease. *Genome Biology* **2017**, *18* (1), 83.
2. Lockhart, D. J.; Winzler, E. A., Genomics, gene expression and DNA arrays. *Nature* **2000**, *405* (6788), 827.
3. Fiehn, O., Metabolomics—the link between genotypes and phenotypes. In *Functional Genomics*, Springer: 2002; pp 155-171.
4. Martin, J. A.; Wang, Z., Next-generation transcriptome assembly. *Nature Reviews Genetics* **2011**, *12* (10), 671.
5. Schwanhäusser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M., Global quantification of mammalian gene expression control. *Nature* **2011**, *473* (7347), 337.
6. Vitzthum, F.; Behrens, F.; Anderson, N. L.; Shaw, J. H., Proteomics: from basic research to diagnostic application. A review of requirements & needs. *Journal of Proteome Research* **2005**, *4* (4), 1086-1097.
7. Dettmer, K.; Hammock, B. D., Metabolomics--a new exciting field within the "omics" sciences. *Environ Health Persp* **2004**, *112* (7), A396.
8. Orešič, M., Metabolomics, a novel tool for studies of nutrition, metabolism and lipid dysfunction. *Nutrition, Metabolism and Cardiovascular Diseases* **2009**, *19* (11), 816-824.
9. Turi, K. N.; Romick-Rosendale, L.; Ryckman, K. K.; Hartert, T. V., A review of metabolomics approaches and their application in identifying causal pathways of childhood asthma. *Journal of Allergy and Clinical Immunology* **2018**, *141* (4), 1191-1201.
10. Barba, I.; Fernandez-Montesinos, R.; Garcia-Dorado, D.; Pozo, D., Alzheimer's disease beyond the genomic era: nuclear magnetic resonance (NMR) spectroscopy-based metabolomics. *Journal of Cellular and Molecular Medicine* **2008**, *12* (5a), 1477-1485.
11. Bothwell, J. H.; Griffin, J. L., An introduction to biological nuclear magnetic resonance spectroscopy. *Biological Reviews* **2011**, *86* (2), 493-510.
12. Wining, H.; Roldán-Marín, E.; Dragsted, L. O.; Viereck, N.; Poulsen, M.; Sánchez-Moreno, C.; Cano, M. P.; Engelsen, S. B., An exploratory NMR nutri-metabonomic investigation reveals dimethyl sulfone as a dietary biomarker for onion

- intake. *Analyst* **2009**, *134* (11), 2344-2351.
13. Jung, J. Y.; Lee, H.-S.; Kang, D.-G.; Kim, N. S.; Cha, M. H.; Bang, O.-S.; Hwang, G.-S., <sup>1</sup>H-NMR-based metabolomics study of cerebral infarction. *Stroke* **2011**, STROKEAHA. 110.598789.
  14. Graham, S. F.; Holscher, C.; Green, B. D., Metabolic signatures of human Alzheimer's disease (AD): <sup>1</sup>H NMR analysis of the polar metabolome of post-mortem brain tissue. *Metabolomics* **2014**, *10* (4), 744-753.
  15. Dettmer, K.; Aronov, P. A.; Hammock, B. D., Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews* **2007**, *26* (1), 51-78.
  16. Scholz, M.; Gatzek, S.; Sterling, A.; Fiehn, O.; Selbig, J., Metabolite fingerprinting: detecting biological features by independent component analysis. *Bioinformatics* **2004**, *20* (15), 2447-2454.
  17. Kirwan, J. A.; Weber, R. J.; Broadhurst, D. I.; Viant, M. R., Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. *Scientific Data* **2014**, *1*, 140012.
  18. Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X., Modern analytical techniques in metabolomics analysis. *Analyst* **2012**, *137* (2), 293-300.
  19. Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T., Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *Journal of Proteome Research* **2003**, *2* (5), 488-494.
  20. Ohashi, H.; Hasegawa, M.; Wakimoto, K.; Miyamoto-Sato, E., Next-generation technologies for multiomics approaches including interactome sequencing. *Biomed Res Int* **2015**.
  21. Deng, C.; Zhang, X.; Li, N., Investigation of volatile biomarkers in lung cancer blood using solid-phase microextraction and capillary gas chromatography–mass spectrometry. *Journal of Chromatography B* **2004**, *808* (2), 269-277.
  22. Dunn, W. B.; Ellis, D. I., Metabolomics: current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry* **2005**, *24* (4), 285-294.
  23. Guo, K.; Li, L., Differential <sup>12</sup>C-/<sup>13</sup>C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal Chem* **2009**, *81* (10), 3919-3932.
  24. El-Aneed, A.; Cohen, A.; Banoub, J., Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy*

*Reviews* **2009**, *44* (3), 210-230.

25. Nicholson, J. K.; Connelly, J.; Lindon, J. C.; Holmes, E., Metabonomics: a platform for studying drug toxicity and gene function. *Nature Reviews Drug Discovery* **2002**, *1* (2), 153.

26. Patti, G. J.; Yanes, O.; Siuzdak, G., Innovation: Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology* **2012**, *13* (4), 263.

27. Styczynski, M. P.; Moxley, J. F.; Tong, L. V.; Walther, J. L.; Jensen, K. L.; Stephanopoulos, G. N., Systematic identification of conserved metabolites in GC/MS data for metabolomics and biomarker discovery. *Anal Chem* **2007**, *79* (3), 966-973.

28. Alonso, A.; Marsal, S.; Julià, A., Analytical methods in untargeted metabolomics: state of the art in 2015. *Frontiers in Bioengineering and Biotechnology* **2015**, *3*, 23.

29. González-Domínguez, R.; García-Barrera, T.; Vitorica, J.; Gómez-Ariza, J. L., Region-specific metabolic alterations in the brain of the APP/PS1 transgenic mice of Alzheimer's disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **2014**, *1842* (12), 2395-2402.

30. Nishiumi, S.; Kobayashi, T.; Ikeda, A.; Yoshie, T.; Kibi, M.; Izumi, Y.; Okuno, T.; Hayashi, N.; Kawano, S.; Takenawa, T., A novel serum metabolomics-based diagnostic approach for colorectal cancer. *Plos One* **2012**, *7* (7), e40459.

31. Zhang, L.; Li, L.; Kong, H.; Zeng, F., Urinary metabolomics study of renal cell carcinoma based on gas chromatography-mass spectrometry. *Nan fang yi ke da xue xue bao= Journal of Southern Medical University* **2015**, *35* (5), 763-766.

32. Hounoum, B. M.; Blasco, H.; Emond, P.; Mavel, S., Liquid chromatography–high-resolution mass spectrometry-based cell metabolomics: Experimental design, recommendations, and applications. *TrAC Trends in Analytical Chemistry* **2016**, *75*, 118-128.

33. Lin, C. Y.; Wu, H.; Tjeerdema, R. S.; Viant, M. R., Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics. *Metabolomics* **2007**, *3* (1), 55-67.

34. Viant, M. R.; Rosenblum, E. S.; Tjeerdema, R. S., NMR-based metabolomics: a powerful approach for characterizing the effects of environmental stressors on organism health. *Environmental Science & Technology* **2003**, *37* (21), 4982-4989.

35. Waters, N. J.; Holmes, E.; Waterfield, C. J.; Farrant, R. D.; Nicholson, J. K.,

NMR and pattern recognition studies on liver extracts and intact livers from rats treated with  $\alpha$ -naphthylisothiocyanate. *Biochemical Pharmacology* **2002**, *64* (1), 67-77.

36. Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G., XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **2006**, *78* (3), 779-787.

37. Smith, C. A.; O'Maille, G.; Want, E. J.; Qin, C.; Trauger, S. A.; Brandon, T. R.; Custodio, D. E.; Abagyan, R.; Siuzdak, G., METLIN: a metabolite mass spectral database. *Ther. Drug. Monit.* **2005**, *27* (6), 747-751.

38. Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S., HMDB: the human metabolome database. *Nucleic Acids Research* **2007**, *35* (suppl\_1), D521-D526.

39. Siskos, A. P.; Jain, P.; Römisch-Margl, W.; Bennett, M.; Achaintre, D.; Asad, Y.; Marney, L.; Richardson, L.; Koulman, A.; Griffin, J. L., Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. *Anal. Chem.* **2016**, *89* (1), 656-665.

40. Soumeh, E. A.; Hedemann, M. S.; Poulsen, H. D.; Corrent, E.; van Milgen, J.; Nørgaard, J. V., Nontargeted LC-MS Metabolomics Approach for Metabolic Profiling of Plasma and Urine from Pigs Fed Branched Chain Amino Acids for Maximum Growth Performance. *Journal of Proteome Research* **2016**, *15* (12), 4195-4207.

41. Ebshiana, A. A.; Snowden, S. G.; Legido-Quigley, C., Multimodal Metabolomics Combining UPLC-qToF-MS and GC-MS Data in Plasma and Brain Tissue. **2017**.

42. Mir, S. A.; Rajagopalan, P.; Jain, A. P.; Khan, A. A.; Datta, K. K.; Mohan, S. V.; Lateef, S. S.; Sahasrabudhe, N.; Somani, B.; Prasad, T. K., LC-MS-based serum metabolomic analysis reveals dysregulation of phosphatidylcholines in esophageal squamous cell carcinoma. *Journal of Proteomics* **2015**, *127*, 96-102.

43. Zhao, S.; Luo, X.; Li, L., Chemical isotope labeling LC-MS for high coverage and quantitative profiling of the hydroxyl submetabolome in metabolomics. *Anal. Chem.* **2016**, *88* (21), 10617-10623.

44. Guo, K.; Li, L., High-performance isotope labeling for profiling carboxylic acid-containing metabolites in biofluids by mass spectrometry. *Anal. Chem.* **2010**, *82* (21), 8789-8793.

45. Zhao, S.; Dawe, M.; Guo, K.; Li, L., Development of High-Performance

Chemical Isotope Labeling LC–MS for Profiling the Carbonyl Submetabolome. *Anal. Chim.* **2017**, *89* (12), 6758-6765.

46. Causon, T. J.; Hann, S., Review of sample preparation strategies for MS-based metabolomic studies in industrial biotechnology. *Anal. Chim. Acta* **2016**, *938*, 18-32.

47. Yan, M.; Xu, G., Current and future perspectives of functional metabolomics in disease studies review. *Anal. Chim. Acta* **2018**, <https://doi.org/10.1016/j.aca.2018.04.006>, in press.

48. Naz, S.; dos Santos, D. C. M.; Garcia, A.; Barbas, C., Analytical protocols based on LC-MS, GC-MS and CE-MS for nontargeted metabolomics of biological tissues. *Bioanalysis* **2014**, *6* (12), 1657-1677.

49. Cudjoe, E.; Bojko, B.; Togunde, P.; Pawliszyn, J., In vivo solid-phase microextraction for tissue bioanalysis. *Bioanalysis* **2012**, *4* (21), 2605-2619.

50. Xiang, L.; Zhang, H.; Wei, J.; Tian, X. Y.; Luan, H.; Li, S.; Zhao, H.; Cao, G.; Chung, A. C. K.; Yang, C.; Huang, Y.; Cai, Z., Metabolomics studies on db/db diabetic mice in skeletal muscle reveal effective clearance of overloaded intermediates by exercise. *Anal. Chim. Acta* **2018**, <https://doi.org/10.1016/j.aca.2017.11.082>, in press.

51. Huang, Q.; Tan, Y.; Yin, P.; Ye, G.; Gao, P.; Lu, X.; Wang, H.; Xu, G., Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics. *Cancer Research* **2013**, *73* (16), 4992-5002.

52. Zhang, H.; Cui, L.; Liu, W.; Wang, Z.; Ye, Y.; Li, X.; Wang, H., <sup>1</sup>H NMR metabolic profiling of gastric cancer patients with lymph node metastasis. *Metabolomics* **2018**, *14* (4), 47.

53. Lubes, G.; Goodarzi, M., GC–MS based metabolomics used for the identification of cancer volatile organic compounds as biomarkers. *J Pharmaceut Biomed* **2017**.

54. Hirayama, A.; Kami, K.; Sugimoto, M.; Sugawara, M.; Toki, N.; Onozuka, H.; Kinoshita, T.; Saito, N.; Ochiai, A.; Tomita, M., Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Research* **2009**, *69* (11), 4918-4925.

55. Cao, Q. J. W.; Ouyang, C. Z.; Zhong, X. F.; Li, L. J., Profiling of small molecule metabolites and neurotransmitters in crustacean hemolymph and neuronal tissues using reversed-phase LC-MS/MS. *Electrophoresis* **2018**, *39* (9-10), 1241-1248.

56. Wells, A.; Barrington, W. T.; Dearth, S.; May, A.; Threadgill, D. W.; Campagna,

S. R.; Voy, B. H., Tissue Level Diet and Sex-by-Diet Interactions Reveal Unique Metabolite and Clustering Profiles Using Untargeted Liquid Chromatography Mass Spectrometry on Adipose, Skeletal Muscle, and Liver Tissue in C57BL6/J Mice. *J. Proteome Res.* **2018**, *17* (3), 1077-1090.

57. Sanchez-Lopez, E.; Happe, H.; Steenvoorden, E.; Crego, A. L.; Marina, M. L.; Peters, D. J. M.; Mayboroda, O. A.; Dutch, D. C., A cross-platform metabolomics workflow for volume-restricted tissue samples: application to an animal model for polycystic kidney disease. *Mol. BioSyst.* **2017**, *13* (10), 1940-1945.

58. Ciborowski, M.; Kisluk, J.; Pietrowska, K.; Samczuk, P.; Parfieniuk, E.; Kowalczyk, T.; Kozlowski, M.; Kretowski, A.; Niklinski, J., Development of LC-QTOF-MS method for human lung tissue fingerprinting. A preliminary application to nonsmall cell lung cancer. *Electrophoresis* **2017**, *38* (18), 2304-2312.

59. Leuthold, P.; Schaeffeler, E.; Winter, S.; Buttner, F.; Hofmann, U.; Murdter, T. E.; Rausch, S.; Sonntag, D.; Wahrheit, J.; Fend, F.; Hennenlotter, J.; Bedke, J.; Schwab, M.; Haag, M., Comprehensive Metabolomic and Lipidomic Profiling of Human Kidney Tissue: A Platform Comparison. *J. Proteome Res.* **2017**, *16* (2), 933-944.

60. Huang, T. J.; Toro, M.; Lee, R.; Hui, D. S.; Edwards, J. L., Multi-functional derivatization of amine, hydroxyl, and carboxylate groups for metabolomic investigations of human tissue by electrospray ionization mass spectrometry. *Analyst* **2018**, *143* (14), 3408-3414.

61. Xie, B. E.; Wang, Y. Y.; Jones, D. R.; Dey, K. K.; Wang, X. S.; Li, Y. X.; Cho, J. H.; Shaw, T. I.; Tan, H. Y.; Peng, J. M., Isotope Labeling-Assisted Evaluation of Hydrophilic and Hydrophobic Liquid Chromatograph-Mass Spectrometry for Metabolomics Profiling. *Anal. Chem.* **2018**, *90* (14), 8538-8545.

62. Li, Z. C.; Lu, Y.; Guo, Y. F.; Cao, H. J.; Wang, Q. H.; Shui, W. Q., Comprehensive evaluation of untargeted metabolomics data processing software in feature detection, quantification and discriminating marker selection. *Anal. Chim. Acta* **2018**, *1029*, 50-57.

63. DeFelice, B. C.; Mehta, S. S.; Samra, S.; Cajka, T.; Wancewicz, B.; Fahrman, J. F.; Fiehn, O., Mass Spectral Feature List Optimizer (MS-FLO): A Tool To Minimize False Positive Peak Reports in Untargeted Liquid Chromatography-Mass Spectroscopy (LC-MS) Data Processing. *Anal. Chem.* **2017**, *89* (6), 3250-3255.

64. Mahieu, N. G.; Patti, G. J., Systems-Level Annotation of a 25 000 Features to

Fewer than G Metabolomics Data Set Reduces 1000 Unique Metabolites. *Anal. Chem.* **2017**, *89* (19), 10397-10406.

65. Arrivault, S.; Guenther, M.; Fry, S. C.; Fuenfgeld, M.; Veyel, D.; Mettler-Altmann, T.; Stitt, M.; Lunn, J. E., Synthesis and Use of Stable-Isotope-Labeled Internal Standards for Quantification of Phosphorylated Metabolites by LC-MS/MS. *Anal. Chem.* **2015**, *87* (13), 6896-6904.

66. Chu, J. M.; Qi, C. B.; Huang, Y. Q.; Jiang, H. P.; Hao, Y. H.; Yuan, B. F.; Feng, Y. Q., Metal Oxide-Based Selective Enrichment Combined with Stable Isotope Labeling-Mass Spectrometry Analysis for Profiling of Ribose Conjugates. *Analytical Chemistry* **2015**, *87* (14), 7364-7372.

67. Dai, W. D.; Huang, Q.; Yin, P. Y.; Li, J.; Zhou, J.; Kong, H. W.; Zhao, C. X.; Lu, X.; Xu, G. W., Comprehensive and Highly Sensitive Urinary Steroid Hormone Profiling Method Based on Stable Isotope-Labeling Liquid Chromatography Mass Spectrometry. *Analytical Chemistry* **2012**, *84* (23), 10245-10251.

68. Hao, L.; Johnson, J.; Lietz, C. B.; Buchberger, A.; Frost, D.; Kao, W. J.; Li, L. J., Mass Defect-Based N,N-Dimethyl Leucine Labels for Quantitative Proteomics and Amine Metabolomics of Pancreatic Cancer Cells. *Analytical Chemistry* **2017**, *89* (2), 1138-1146.

69. Leng, J. P.; Wang, H. Y.; Zhang, L.; Zhang, J.; Wang, H.; Guo, Y. L., A highly sensitive isotope-coded derivatization method and its application for the mass spectrometric analysis of analytes containing the carboxyl group. *Anal. Chim. Acta* **2013**, *758*, 114-121.

70. Mochizuki, T.; Todoroki, K.; Inoue, K.; Min, J. Z.; Toyo'oka, T., Isotopic variants of light and heavy L-pyroglutamic acid succinimidyl esters as the derivatization reagents for DL-amino acid chiral metabolomics identification by liquid chromatography and electrospray ionization mass spectrometry. *Anal. Chim. Acta* **2014**, *811*, 51-59.

71. Tayyari, F.; Gowda, G. A. N.; Gu, H. W.; Raftery, D., N-15-Cholamine-A Smart Isotope Tag for Combining NMR- and MS-Based Metabolite Profiling. *Analytical Chemistry* **2013**, *85* (18), 8715-8721.

72. Wagner, M.; Ohlund, L. B.; Shiao, T. C.; Vezina, A.; Annabi, B.; Roy, R.; Sleno, L., Isotope-labeled differential profiling of metabolites using N-benzoyloxysuccinimide derivatization coupled to liquid chromatography/high-

resolution tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2015**, *29* (18), 1632-1640.

73. Wong, J. M. T.; Malec, P. A.; Mabrouk, O. S.; Ro, J.; Dus, M.; Kennedy, R. T., Benzoyl chloride derivatization with liquid chromatography-mass spectrometry for targeted metabolomics of neurochemicals in biological samples. *J. Chromatogr. A* **2016**, *1446*, 78-90.

74. Yang, W. C.; Regnier, F. E.; Jiang, Q.; Adamec, J., In vitro stable isotope labeling for discovery of novel metabolites by liquid chromatography-mass spectrometry: Confirmation of gamma-tocopherol metabolism in human A549 cell. *J. Chromatogr. A* **2010**, *1217* (5), 667-675.

75. Yuan, W.; Edwards, J. L.; Li, S. W., Global profiling of carbonyl metabolites with a photo-cleavable isobaric labeling affinity tag. *Chem. Commun.* **2013**, *49* (94), 11080-11082.

76. Yuan, W.; Zhang, J. X.; Li, S. W.; Edwards, J. L., Amine Metabolomics of Hyperglycemic Endothelial Cells using Capillary LC-MS with Isobaric Tagging. *Journal of Proteome Research* **2011**, *10* (11), 5242-5250.

77. Han, W.; Sapkota, S.; Camicioli, R.; Dixon, R. A.; Li, L., Profiling novel metabolic biomarkers for Parkinson's disease using in-depth metabolomic analysis. *Movement Disord* **2017**, *32* (12), 1720-1728.

78. Wu, Y.; Streijger, F.; Wang, Y.; Lin, G.; Christie, S.; Mac-Thiong, J.-M.; Parent, S.; Bailey, C. S.; Paquette, S.; Boyd, M. C., Parallel metabolomic profiling of cerebrospinal fluid and serum for identifying biomarkers of injury severity after acute human spinal cord injury. *Scientific Reports* **2016**, *6*, 38718.

79. Peng, J.; Guo, K.; Xia, J.; Zhou, J.; Yang, J.; Westaway, D.; Wishart, D. S.; Li, L., Development of isotope labeling liquid chromatography mass spectrometry for mouse urine metabolomics: quantitative metabolomic study of transgenic mice related to Alzheimer's disease. *Journal of Proteome Research* **2014**, *13* (10), 4457-4469.

80. Zheng, J.; Dixon, R. A.; Li, L., Development of isotope labeling LC-MS for human salivary metabolomics and application to profiling metabolome changes associated with mild cognitive impairment. *Anal Chem* **2012**, *84* (24), 10802-10811.

81. Hooton, K.; Li, L., Nonocclusive Sweat Collection Combined with Chemical Isotope Labeling LC-MS for Human Sweat Metabolomics and Mapping the Sweat Metabolomes at Different Skin Locations. *Anal. Chem.* **2017**, *89* (15), 7847-7851.

82. Luo, X.; Gu, X.; Li, L., Development of a simple and efficient method of harvesting and lysing adherent mammalian cells for chemical isotope labeling LC-MS-based cellular metabolomics. *Anal. Chim. Acta.* **2017**.
83. Oakley, H.; Cole, S. L.; Logan, S.; Maus, E.; Shao, P.; Craft, J.; Guillozet-Bongaarts, A.; Ohno, M.; Disterhoft, J.; Van Eldik, L., Intraneuronal  $\beta$ -amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: potential factors in amyloid plaque formation. *Journal of Neuroscience* **2006**, *26* (40), 10129-10140.
84. Xia, J.; Psychogios, N.; Young, N.; Wishart, D. S., MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Research* **2009**, *37* (suppl\_2), W652-W660.
85. Huan, T.; Wu, Y.; Tang, C.; Lin, G.; Li, L., DnsID in MyCompoundID for rapid identification of dansylated amine- and phenol-containing metabolites in LC-MS-based metabolomics. *Anal. Chem.* **2015**, *87* (19), 9838-9845.
86. Wu, Y.; Li, L., Determination of Total Concentration of Chemically Labeled Metabolites as a Means of Metabolome Sample Normalization and Sample Loading Optimization in Mass Spectrometry-Based Metabolomics. *Anal. Chem.* **2012**, *84* (24), 10723-10731.
87. Zhou, R.; Tseng, C.-L.; Huan, T.; Li, L., IsoMS: automated processing of LC-MS data generated by a chemical isotope labeling metabolomics platform. *Anal. Chem.* **2014**, *86* (10), 4675-4679.
88. Liebeke, M.; Bundy, J. G., Tissue disruption and extraction methods for metabolic profiling of an invertebrate sentinel species. *Metabolomics* **2012**, *8* (5), 819-830.
89. Beckonert, O.; Keun, H. C.; Ebbels, T. M.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K., Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols* **2007**, *2* (11), 2692-2703.
90. Cequier-Sánchez, E.; RODRíguez, C.; Ravelo, A. n. G.; Zarate, R., Dichloromethane as a solvent for lipid extraction and assessment of lipid classes and fatty acids from samples of different natures. *J Agr Food Chem* **2008**, *56* (12), 4297-4303.
91. Seiler, N.; Raul, F., Polyamines and apoptosis. *Journal of Cellular and*

*Molecular Medicine* **2005**, 9 (3), 623-642.

92. Jänne, J.; Alhonen, L.; Pietilä, M.; Keinänen, T. A., Genetic approaches to the cellular functions of polyamines in mammals. *The FEBS Journal* **2004**, 271 (5), 877-894.
93. Takano, K.; Ogura, M.; Nakamura, Y.; Yoneda, Y., Neuronal and glial responses to polyamines in the ischemic brain. *Current Neurovascular Research* **2005**, 2 (3), 213-223.
94. Chen, G. G.; Fiori, L. M.; Moquin, L.; Gratton, A.; Mamer, O.; Mechawar, N.; Turecki, G., Evidence of altered polyamine concentrations in cerebral cortex of suicide completers. *Neuropsychopharmacology* **2010**, 35 (7), 1477.
95. Fonteh, A.; Harrington, R.; Tsai, A.; Liao, P.; Harrington, M., Free amino acid and dipeptide changes in the body fluids from Alzheimer's disease subjects. *Amino Acids* **2007**, 32 (2), 213-224.
96. Volicer, L.; Drenfeld, L. K.; Freedman, M.; Albert, M. L.; Langlais, P. J.; Bird, E. D., Serotonin and 5-hydroxyindoleacetic acid in csf: Difference in parkinson's disease and dementia of the alzheimer's type. *Archives of Neurology* **1985**, 42 (2), 127-129.
97. Joerg, S. H.; Stefanie, G.; Tomas, M.-T., 5-Hydroxyindoleacetic acid and homovanillic acid concentrations in cerebrospinal fluid in patients with Alzheimer's disease, depression and mild cognitive impairment. *Neuroendocrinology Letters* **2004**, 25 (6).
98. Morimoto, S.; Takao, M.; Hatsuta, H.; Nishina, Y.; Komiya, T.; Sengoku, R.; Nakano, Y.; Uchino, A.; Sumikura, H.; Saito, Y.; Kanemaru, K.; Murayama, S., Homovanillic acid and 5-hydroxyindole acetic acid as biomarkers for dementia with Lewy bodies and coincident Alzheimer's disease: An autopsy-confirmed study. *Plos One* **2017**, 12 (2), e0171524.
99. Fukumori, R.; Nakamichi, N.; Takarada, T.; Kambe, Y.; Matsushima, N.; Moriguchi, N.; Yoneda, Y., Inhibition by 2-Methoxy-4-ethylphenol of Ca<sup>2+</sup> Influx Through Acquired and Native N-Methyl-D-aspartate Receptor Channels. *Journal of Pharmacological Sciences* **2010**, 112 (3), 273-281.
100. Rafacho, A.; Ortsäter, H.; Nadal, A.; Quesada, I., Glucocorticoid treatment and endocrine pancreas function. *Journal of Endocrinology* **2014**, JOE-14-0373.
101. Wu, D.-Y.; Ou, C.-Y.; Chodankar, R.; Siegmund, K. D.; Stallcup, M. R.,

Distinct, genome-wide, gene-specific selectivity patterns of four glucocorticoid receptor coregulators. *Nuclear Receptor Signaling* **2014**, *12*.

102. Ciriaco, M.; Ventrice, P.; Russo, G.; Scicchitano, M.; Mazzitello, G.; Scicchitano, F.; Russo, E., Corticosteroid-related central nervous system side effects. *Journal of Pharmacology & Pharmacotherapeutics* **2013**, *4* (Suppl1), S94.

103. Brown, D. G.; Rao, S.; Weir, T. L.; O'Malia, J.; Bazan, M.; Brown, R. J.; Ryan, E. P., Metabolomics and metabolic pathway networks from human colorectal cancers, adjacent mucosa, and stool. *Cancer & Metabolism* **2016**, *4* (1), 11.

104. Jacob, M.; Lopata, A. L.; Dasouki, M.; Abdel Rahman, A. M., Metabolomics toward personalized medicine. *Mass Spectrometry Reviews* **2017**.

105. Malkawi, A. K.; Alzoubi, K. H.; Jacob, M.; Matic, G.; Ali, A.; Al Faraj, A.; Almuhanha, F.; Dasouki, M.; Abdel Rahman, A. M., Metabolomics Based Profiling of Dexamethasone Side Effects in Rats. *Frontiers in Pharmacology* **2018**, *9*, 46.

106. Lane, R.; Mastaglia, F., Drug-induced myopathies in man. *Lancet (London, England)* **1978**, *2* (8089), 562-566.

107. Ferris, H. A.; Kahn, C. R., New mechanisms of glucocorticoid-induced insulin resistance: make no bones about it. *The Journal of Clinical Investigation* **2012**, *122* (11), 3854-3857.

108. Bakker, M. F.; Jacobs, J. W.; Welsing, P. M.; Verstappen, S. M.; Tekstra, J.; Ton, E.; Geurts, M. A.; van der Werf, J. H.; van Albada-Kuipers, G. A.; Jahangier-de Veen, Z. N., Low-dose prednisone inclusion in a methotrexate-based, tight control strategy for early rheumatoid arthritis: a randomized trial. *Annals of Internal Medicine* **2012**, *156* (5), 329-339.

109. Pegg, A. E., Recent advances in the biochemistry of polyamines in eukaryotes. *Biochemical Journal* **1986**, *234* (2), 249.

110. Gonzalez, S.; Coirini, H.; Deniselle, M. C. G.; Gonzalez, S.; Calandra, R.; De Nicola, A. F., Time-dependent effects of dexamethasone on glutamate binding, ornithine decarboxylase activity and polyamine levels in the transected spinal cord. *The Journal of Steroid Biochemistry and Molecular Biology* **1995**, *55* (1), 85-92.

111. Fajardo, I.; Urdiales, J. L.; Medina, M. A.; Sanchez-Jimenez, F., Effects of phorbol ester and dexamethasone treatment on histidine decarboxylase and ornithine decarboxylase in basophilic cells. *Biochemical Pharmacology* **2001**, *61* (9), 1101-1106.

112. Stefanelli, C.; Flamigni, F.; Carati, D.; Rossoni, C.; Caldarera, C. M., Effects of

dexamethasone on spermidine N1-acetyltransferase and ornithine decarboxylase activities in rat spleen. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1987**, *930* (1), 79-86.

113. Advani, S.; LaFrancis, D.; Bogdanovic, E.; Taxel, P.; Raisz, L.; Kream, B., Dexamethasone suppresses in vivo levels of bone collagen synthesis in neonatal mice. *Bone* **1997**, *20* (1), 41-46.

114. Wyse, A. T.; Netto, C. A., Behavioral and neurochemical effects of proline. *Metabolic Brain Disease* **2011**, *26* (3), 159.

115. Firestein, G. S., Evolving concepts of rheumatoid arthritis. *Nature* **2003**, *423*, 356.

116. Vos, T.; Allen, C.; Arora, M.; Barber, R. M.; Bhutta, Z. A.; Brown, A.; Carter, A.; Casey, D. C.; Charlson, F. J.; Chen, A. Z., Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* **2016**, *388* (10053), 1545-1602.

117. Niewold, T. B.; Harrison, M. J.; Paget, S. A., Anti-CCP antibody testing as a diagnostic and prognostic tool in rheumatoid arthritis. *QJM: An International Journal of Medicine* **2007**, *100* (4), 193-201.

118. Scott, D., Prognostic factors in early rheumatoid arthritis. *Rheumatology (Oxford, England)* **2000**, *39*, 24-29.

119. Uchida, T.; Fukawa, A.; Uchida, M.; Fujita, K.; Saito, K., Application of a novel protein biochip technology for detection and identification of rheumatoid arthritis biomarkers in synovial fluid. *Journal of Proteome Research* **2002**, *1* (6), 495-499.

120. Wang, Z.; Chen, Z.; Yang, S.; Wang, Y.; Yu, L.; Zhang, B.; Rao, Z.; Gao, J.; Tu, S., 1H NMR-based metabolomic analysis for identifying serum biomarkers to evaluate methotrexate treatment in patients with early rheumatoid arthritis. *Experimental and Therapeutic Medicine* **2012**, *4* (1), 165-171.

121. Dervieux, T.; Greenstein, N.; Kremer, J., Pharmacogenomic and metabolic biomarkers in the folate pathway and their association with methotrexate effects during dosage escalation in rheumatoid arthritis. *Arthritis & Rheumatology* **2006**, *54* (10), 3095-3103.

122. Peng, J.; Li, L., Liquid–liquid extraction combined with differential isotope dimethylaminophenacyl labeling for improved metabolomic profiling of organic acids.

*Anal. Chim. Acta.* **2013**, *803*, 97-105.

123. Van Ede, A. E.; Laan, R. F.; Blom, H. J.; De Abreu, R. A.; van de Putte, L. B. Methotrexate in rheumatoid arthritis: An update with focus on mechanisms involved in toxicity, *Seminars in Arthritis and Rheumatism*, Elsevier: 1998; pp 277-292.

124. Fransen, J.; Van Riel, P., The Disease Activity Score and the EULAR response criteria. *Clinical and Experimental Rheumatology* **2005**, *23* (5), S93.

125. Van Gestel, A.; Prevoo, M.; Van't Hof, M.; Van Rijswijk, M.; Van de Putte, L.; Van Riel, P., Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis: comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism criteria. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology* **1996**, *39* (1), 34-40.

126. Association, A. s., 2018 Alzheimer's disease facts and figures. *Alzheimer's & Dementia* **2018**, *14* (3), 367-429.

127. Selkoe, D. J., Alzheimer's disease: genes, proteins, and therapy. *Physiological reviews* **2001**, *81* (2), 741-766.

128. Ballatore, C.; Lee, V. M.-Y.; Trojanowski, J. Q., Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nature Reviews Neuroscience* **2007**, *8* (9), 663.

129. Love, S.; Louis, D.; Ellison, D. W., *Greenfield's Neuropathology, 2-Volume Set*. CRC Press: 2008.

130. Trushina, E.; Mielke, M. M., Recent advances in the application of metabolomics to Alzheimer's Disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **2014**, *1842* (8), 1232-1239.

131. Karlawish, J.; Jack, C. R.; Rocca, W. A.; Snyder, H. M.; Carrillo, M. C., Alzheimer's disease: The next frontier—Special Report 2017. *Alzheimer's & Dementia* **2017**, *13* (4), 374-380.

132. Blennow, K., Cerebrospinal fluid protein biomarkers for Alzheimer's disease. *NeuroRX* **2004**, *1* (2), 213-225.

133. Tapiola, T.; Alafuzoff, I.; Herukka, S. K.; Parkkinen, L.; Hartikainen, P.; Soininen, H.; Pirttilä, T., Cerebrospinal fluid  $\beta$ -amyloid 42 and tau proteins as biomarkers of Alzheimer-type pathologic changes in the brain. *Archives of Neurology* **2009**, *66* (3), 382-389.

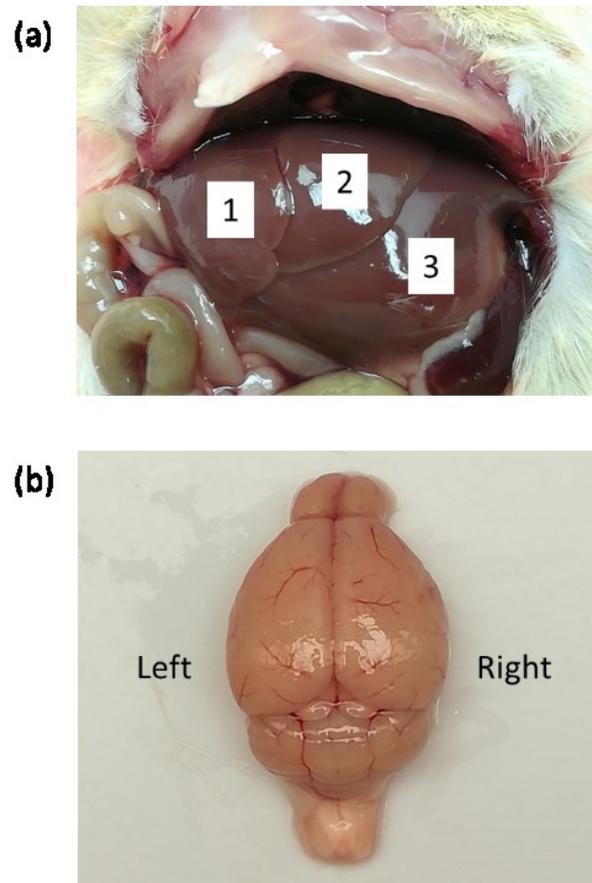
134. Orešič, M.; Hyötyläinen, T.; Herukka, S.; Sysi-Aho, M.; Mattila, I.; Seppänen-Laakso, T.; Julkunen, V.; Gopalacharyulu, P.; Hallikainen, M.; Koikkalainen, J., Metabolome in progression to Alzheimer's disease. *Translational Psychiatry* **2011**, *1* (12), e57.
135. González-Domínguez, R.; García-Barrera, T.; Gómez-Ariza, J., Using direct infusion mass spectrometry for serum metabolomics in Alzheimer's disease. *Anal. Bioanal. Chem.* **2014**, *406* (28), 7137-7148.
136. Chu, H.; Zhang, A.; Han, Y.; Lu, S.; Kong, L.; Han, J.; Liu, Z.; Sun, H.; Wang, X., Metabolomics approach to explore the effects of Kai-Xin-San on Alzheimer's disease using UPLC/ESI-Q-TOF mass spectrometry. *Journal of Chromatography B* **2016**, *1015*, 50-61.
137. Chishti, M. A.; Yang, D.-S.; Janus, C.; Phinney, A. L.; Horne, P.; Pearson, J.; Strome, R.; Zucker, N.; Loukides, J.; French, J., Early-onset amyloid deposition and cognitive deficits in transgenic mice expressing a double mutant form of APP695. *Journal of Biological Chemistry* **2001**.
138. Tapia-Rojas, C.; Lindsay, C. B.; Montecinos-Oliva, C.; Arrazola, M. S.; Retamales, R. M.; Bunout, D.; Hirsch, S.; Inestrosa, N. C., Is L-methionine a trigger factor for Alzheimer's-like neurodegeneration?: Changes in A $\beta$  oligomers, tau phosphorylation, synaptic proteins, Wnt signaling and behavioral impairment in wild-type mice. *Molecular Neurodegeneration* **2015**, *10*, 62.
139. Tapia-Rojas, C.; Lindsay, C. B.; Montecinos-Oliva, C.; Arrazola, M. S.; Retamales, R. M.; Bunout, D.; Hirsch, S.; Inestrosa, N. C., Is L-methionine a trigger factor for Alzheimer's-like neurodegeneration?: Changes in A $\beta$  oligomers, tau phosphorylation, synaptic proteins, Wnt signaling and behavioral impairment in wild-type mice. *Molecular Neurodegeneration* **2015**, *10* (1), 62.
140. Beyer, K.; Lao, J. I.; Carrato, C.; Rodriguez-Vila, A.; Latorre, P.; Mataró, M.; Llopis, M. A.; Mate, J. L.; Ariza, A., Cystathionine beta synthase as a risk factor for Alzheimer disease. *Current Alzheimer Research* **2004**, *1* (2), 127-133.
141. Morrison, L. D.; Kish, S. J., Brain polyamine levels are altered in Alzheimer's disease. *Neuroscience letters* **1995**, *197* (1), 5-8.
142. Herculano, B.; Tamura, M.; Ohba, A.; Shimatani, M.; Kutsuna, N.; Hisatsune, T.,  $\beta$ -alanyl-L-histidine rescues cognitive deficits caused by feeding a high fat diet in a transgenic mouse model of Alzheimer's disease. *Journal of Alzheimer's Disease* **2013**,

33 (4), 983-997.

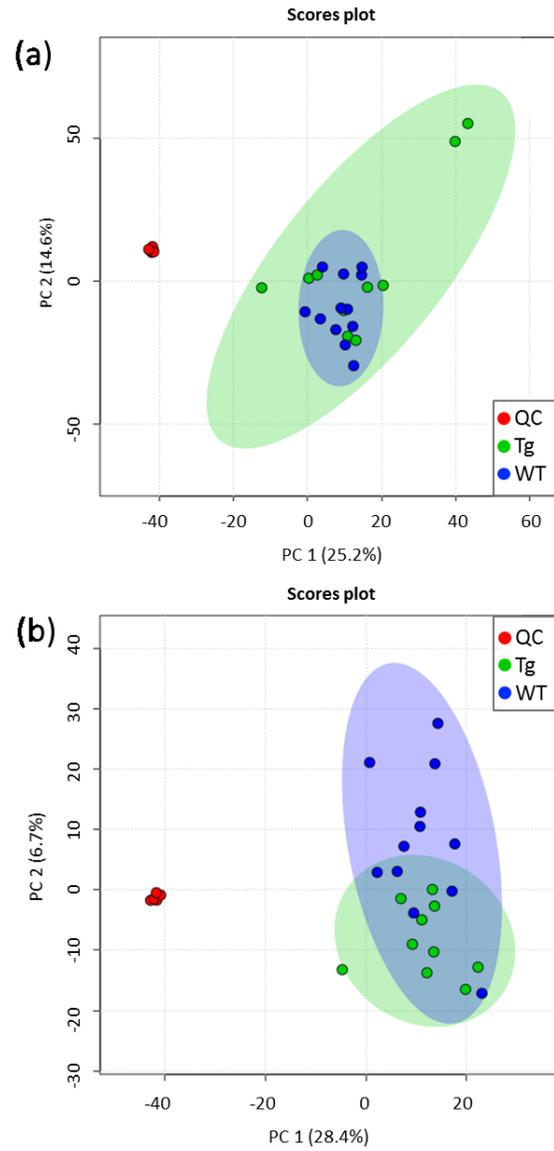
143. Xu, R.; Wang, Q., Towards understanding brain-gut-microbiome connections in Alzheimer's disease. *BMC Systems Biology* **2016**, *10* (3), 63.

## Appendix

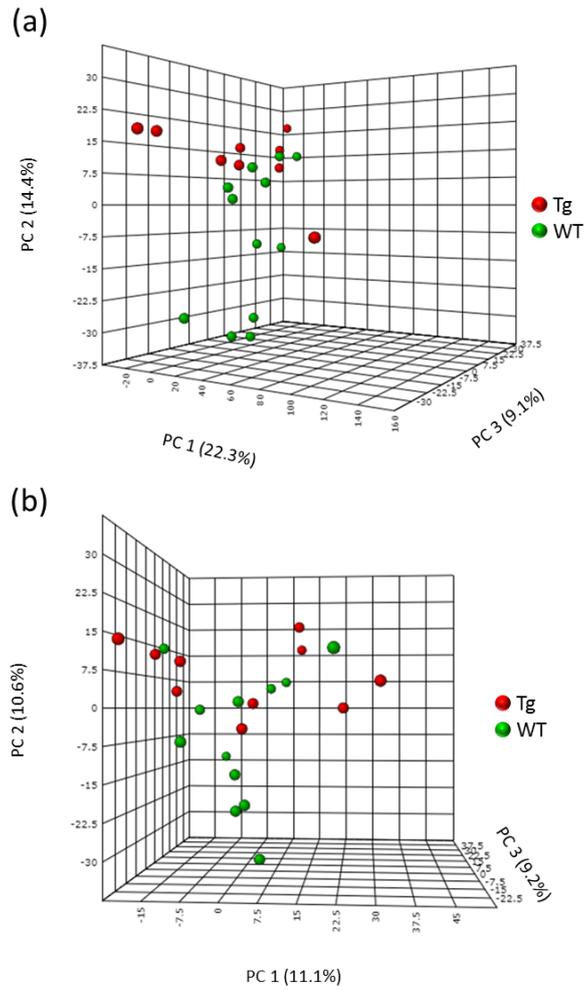
Chapter 2 Development of Chemical Isotope Labeling LC-MS for Tissue Metabolomics and Its Application on Metabolite Biomarker Discovery of Alzheimer's Disease on a Mouse Model



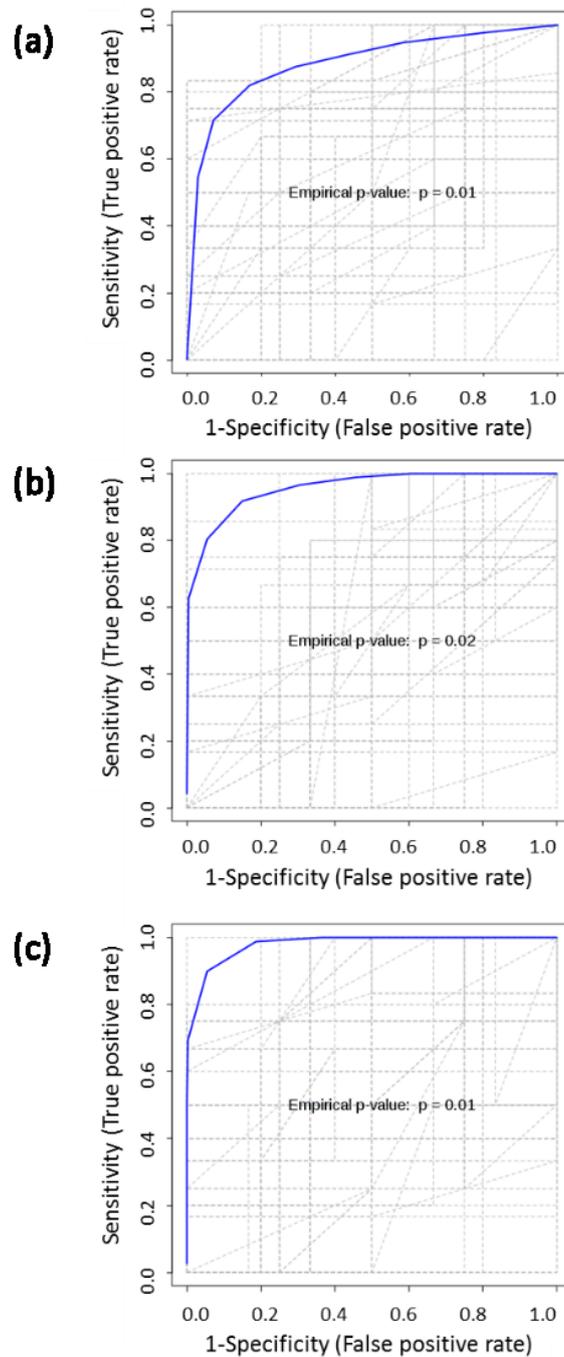
Appendix Figure A2.1. (a) Mouse liver sample. (b) Mouse brain sample.



Appendix Figure A2.2. (a) PCA scores plot of liver samples. (b) PCA scores plot of brain samples. Red dots were labeled as the QC samples. Green dots were labeled as 5xFAD transgenic mice, and blue dots were labeled as wild type mice.

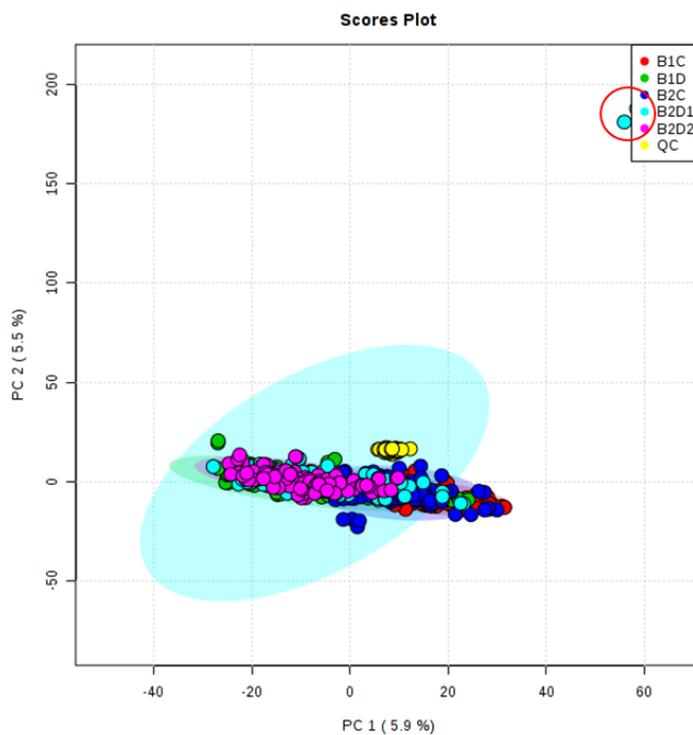


Appendix Figure A2.3. (a) 3D PCA scores plot of liver samples. (b) 3D PCA scores plot of brain samples. Red dots were labeled as 5xFAD transgenic mice, and green dots were labeled as wild type mice.

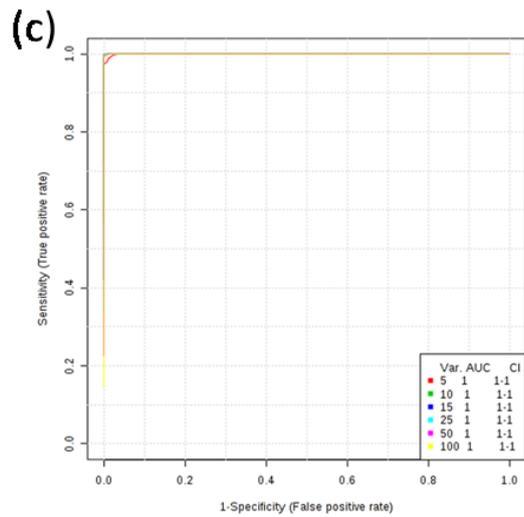
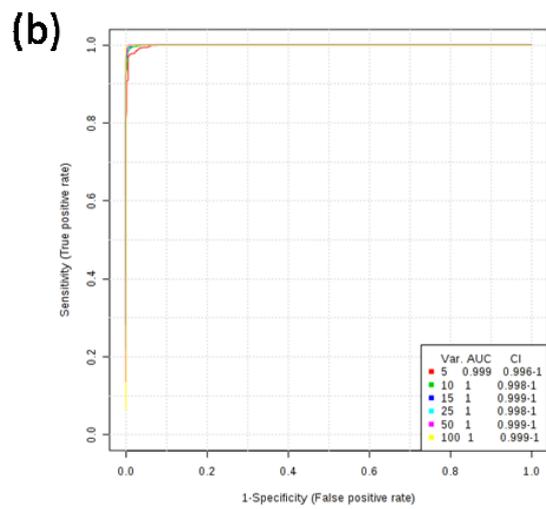
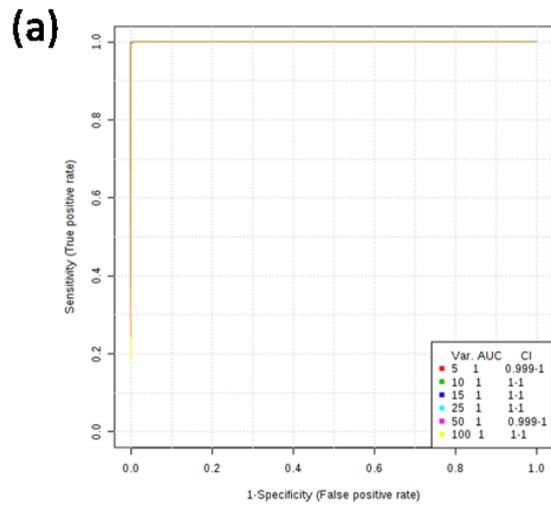


Appendix Figure A2.4. (a) Permutation test result of the ROC curve of Tg vs. WT of liver samples. (b) Permutation test result of the ROC curve of Tg vs. WT of brain samples with 4 metabolite candidates. (c) Permutation test result of the ROC curve of Tg vs. WT of brain samples with 7 metabolite candidates.

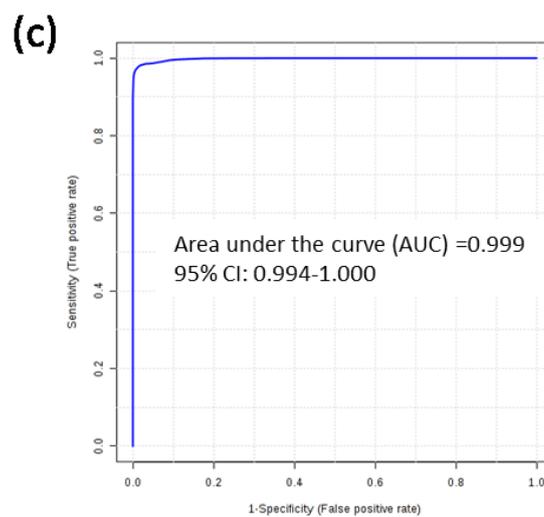
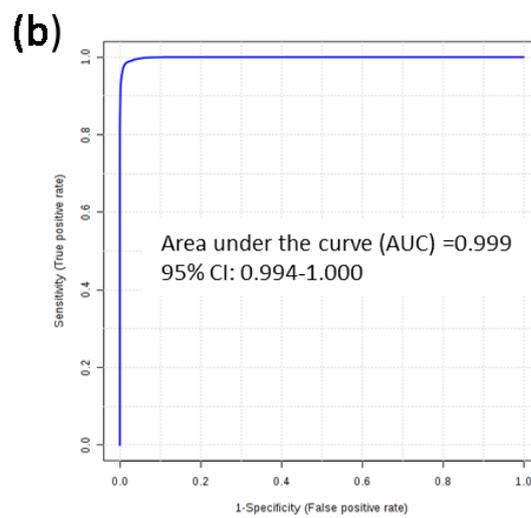
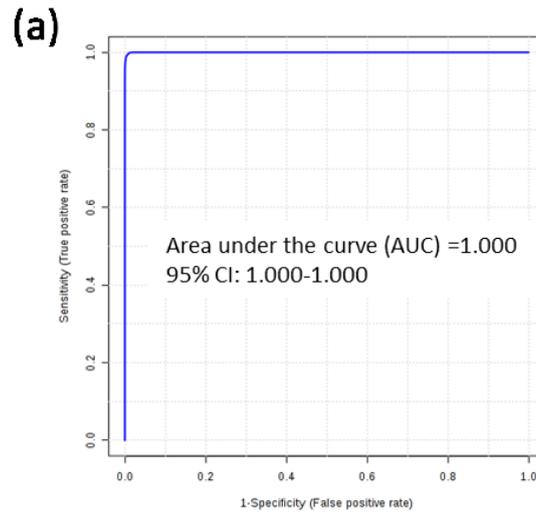
Chapter 4 Development of Chemical Isotope Labeling LC-MS Methods for Metabolomics Studies of Rheumatoid Arthritis Disease



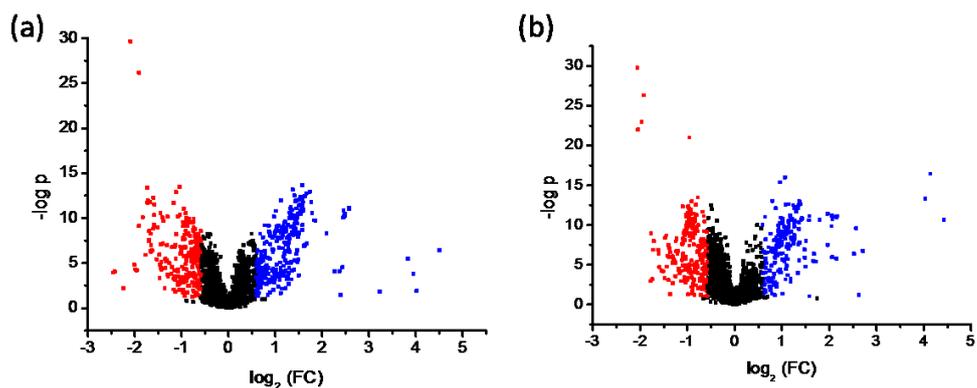
Appendix Figure A4.1. PCA plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 vs. QC. Two outliers of duplicates of sample 152 were redly circled, which were then excluded.



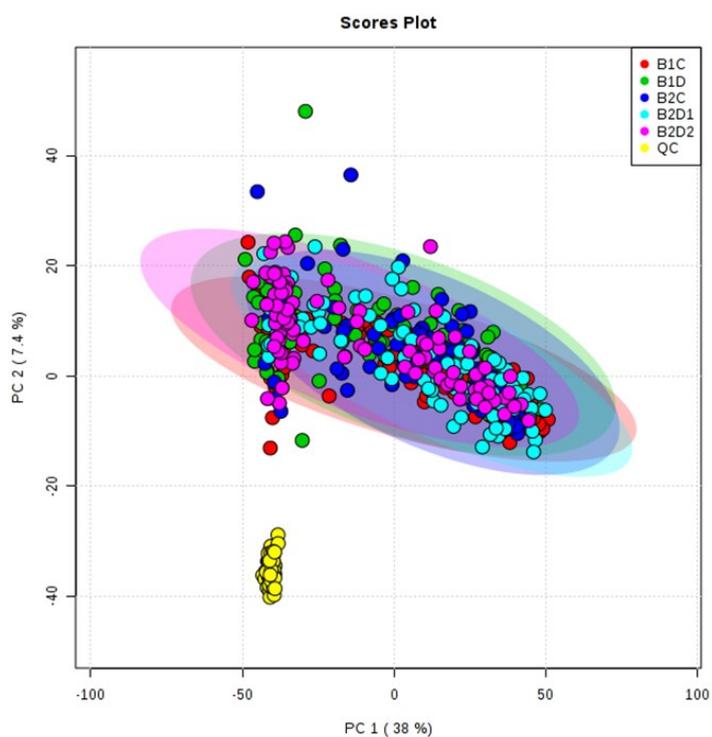
Appendix Figure A4.2. A series of ROC curves built with different numbers of the top-ranked significant metabolites for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2.



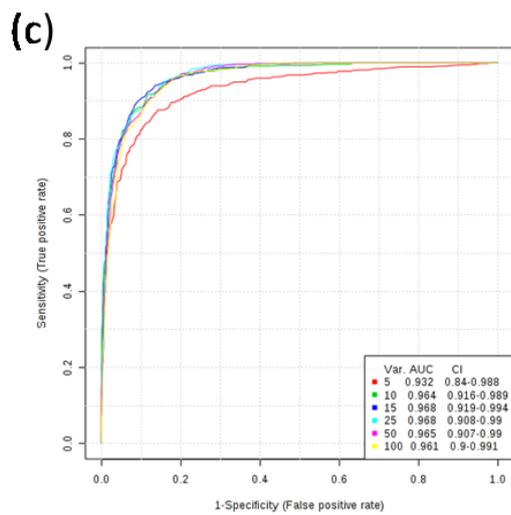
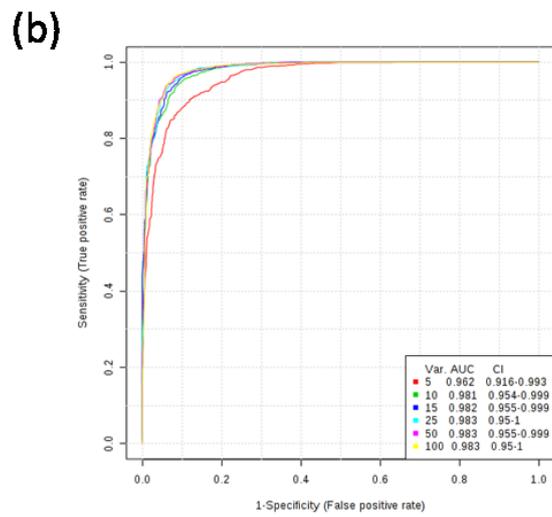
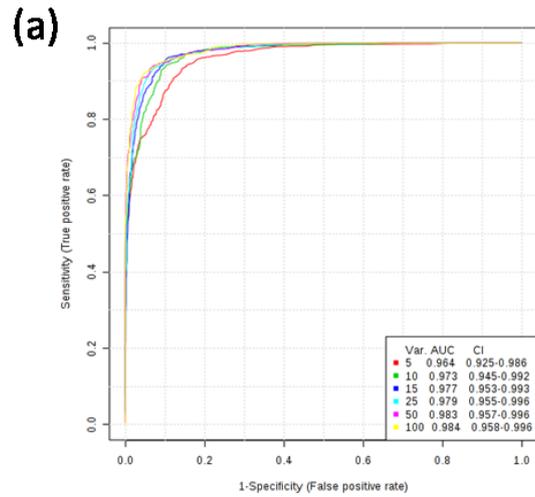
Appendix Figure A4.3. ROC curve built using 7 common significant metabolites, cystine, o-phosphoethanolamine, gamma glutamylglutamic acid, glycyl-valine, dansyl-702, dansyl-9056, dansyl-14033, for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2.



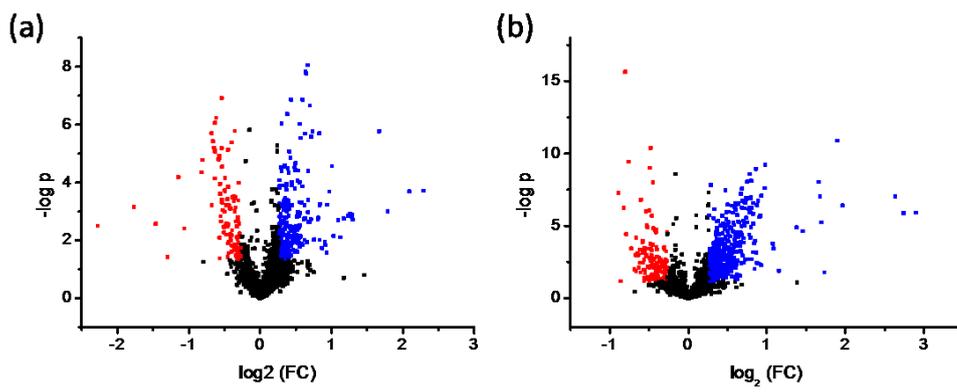
Appendix Figure A4.4. Volcano plot of the changes of metabolites between (a) B1C and MTX N, (b) B1C and MTX Y. Metabolites with fold change (FC) > 1.5 or FC < 0.67 and q-value < 0.05 were highlighted.



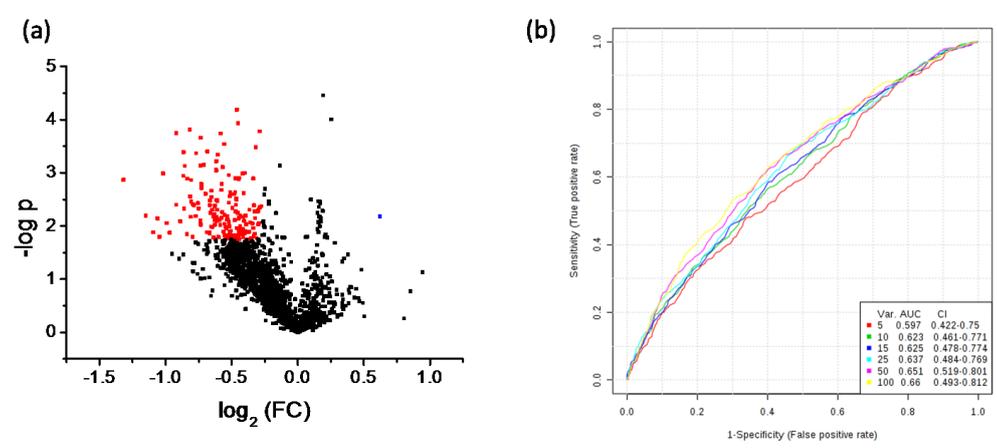
Appendix Figure A4.5. PCA scores plot of B1C vs. B1D vs. B2C vs. B2D1 vs. B2D2 vs. QC.



Appendix Figure A4.6. A series of ROC curves built with different numbers of the top-ranked significant metabolites for (a) B1C vs. B1D, (b) B2C vs. B2D1, (c) B2C vs. B2D2

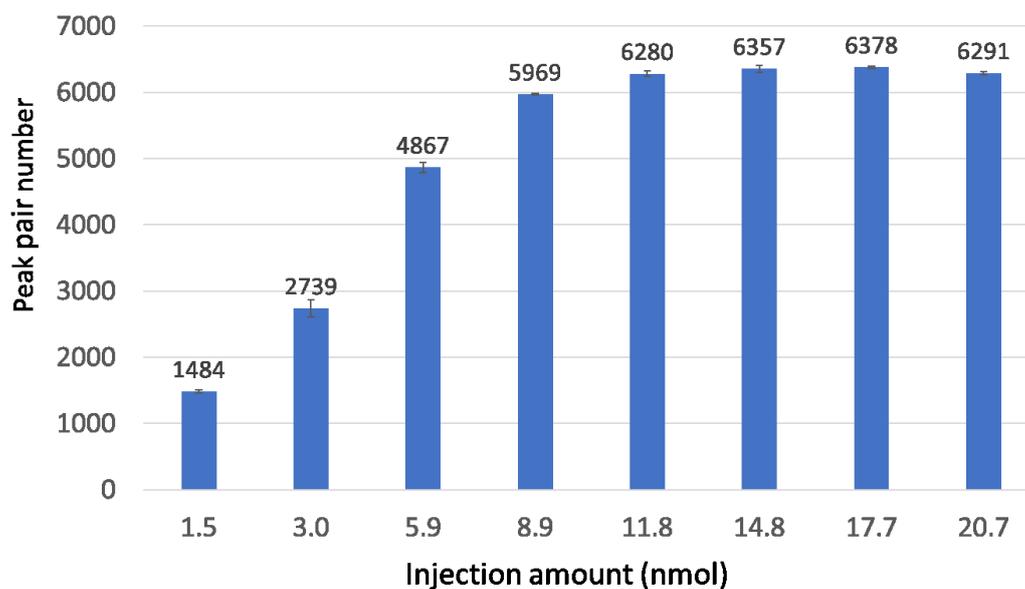


Appendix Figure A4.7. Volcano plot of the changes of metabolites between (a) B1C and MTX N, (b) B1C and MTX Y. Metabolites with fold change (FC) > 1.2 or FC < 0.83 and q-value < 0.1 were highlighted.

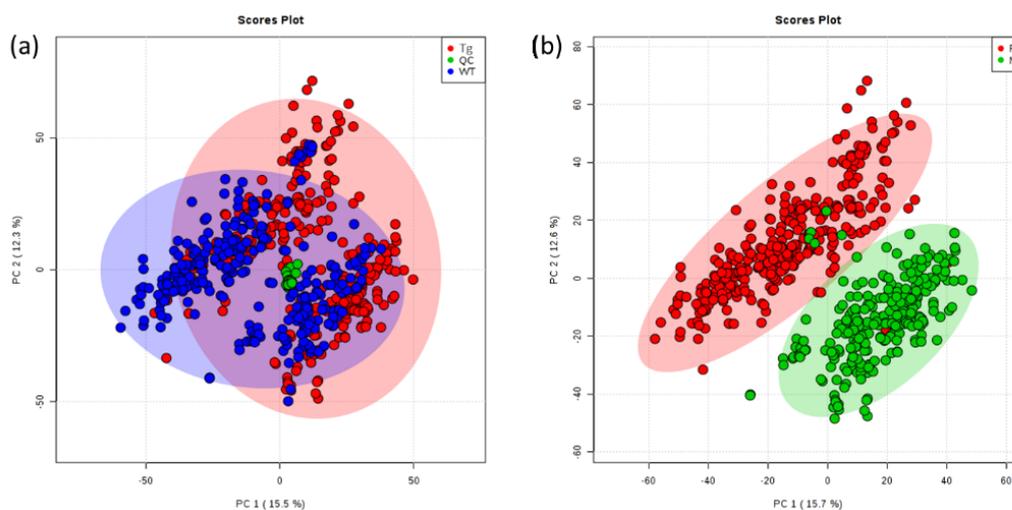


Appendix Figure A4.8. (a) Volcano plot of the changes of metabolites between B2D1 nr and B2D2 nr. (b) A series of ROC curves built with different numbers of the top-ranked significant metabolites for B2D1 nr vs. B2D2 nr.

Chapter 5 Development of Chemical Isotope Labeling LC-MS for Metabolite Biomarker Discovery of Alzheimer's Disease in a Mouse Model



Appendix Figure A5.1. Injection amount optimization curve.



Appendix Figure A5.2. (a) PCA scores plot of all the urine samples. Green dots represented the QC samples. Red dots represented Tg mouse urine samples, and blue dots represented WT mouse urine samples. (b) PCA scores plot of all the urine samples. Green dots were male mouse urine samples. Red dots were female mouse urine samples.